KNOWLEDGE EXTRACTION FROM EXPERIMENTAL AND COMPUTATIONAL DATA FOR SELECTIVE CO OXIDATION AND WATER GAS SHIFT REACTION USING DATA MINING TECHNIQUES

by

Mehmet Erdem Günay

B.S., Chemical Engineering, Boğaziçi University, 2002

M.S., Chemical Engineering, Boğaziçi University, 2005

Submitted to the Institute for Graduate Studies in

Science and Engineering in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

Graduate Program in Chemical Engineering

Boğaziçi University

2012

ACKNOWLEDGEMENTS

First and foremost I would like to offer my sincerest gratitude to my thesis supervisor Prof. Ramazan Yıldırım who has supported me throughout my thesis with his invaluable guidance and experience. This dissertation would not have been possible without his inspiration and effort.

I would like to express my great appreciations for Prof. Ahmet Erhan Aksoylu, Asst. Prof. Burak Alakent and Prof. Mahir Arıkol for their sincere support and guidance during the progression of this work. It was a great opportunity for me to learn from their experience and knowledge. I am thankful to Prof. Hüsnü Atakül and Prof İsmail Boz for their valuable comments and suggestions; their support is highly appreciated. I am also grateful to Prof. Zeynep İlsen Önsan, Assoc. Prof. Ahmet Kerim Avcı and Asst. Prof. Kerem Uğuz for their constructive comments at different stages of my research.

I would like to thank B. Kerem Aksakal, Merve Ayvaz, Şeyma Özkara-Aydınoğlu, M. Selcen Başar, Tuğba Davran-Candan, Burcu Selen-Çağlayan, Feyza Gökaliler, Mustafa Karakaya, Çağla Odabaşı, Göktuğ N. Özyönüm, A. İpek Paksoy, Aslıhan Sümer, Eyüp Şimşek, Sadi T. Tezcanlı, K. Erdem Uğuz, Ali Uzun and all the other past and present members of the CATREL team. I was very lucky to be a part of such a great team.

Special thanks to Yakup Bal, Belgin Balkan, Nurettin Bektaş, Bilgi Dedeoğlu, Melike Gürbüz and Esma Toprak for their technical aid as well as their heartfelt friendship.

Finally, I wish to thank to my dearest family members starting from my wife Gaye Günay; my parents İsmail H. Günay and Leyla Günay; my siblings Bilge, Emre and Ülkü for their motivation and encouragement all through the way.

ABSTRACT

KNOWLEDGE EXTRACTION FROM EXPERIMENTAL AND COMPUTATIONAL DATA FOR SELECTIVE CO OXIDATION AND WATER GAS SHIFT REACTION USING DATA MINING TECHNIQUES
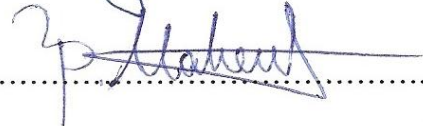
The objective of this dissertation is to apply various data mining techniques for knowledge extraction from experimental and computational data to improve catalyst design and testing conditions for selective CO oxidation and water gas shift reactions. First, the experimental data produced in our laboratory were analyzed in three parts: selective CO oxidation over Pt catalysts by neural networks, selective CO oxidation over Au catalysts by decision trees and neural networks, and water gas shift reaction over Pt catalysts by neural networks. In all these works, the models successfully established the effects and relative significances of the catalyst variables in a reasonable agreement with the literature. Then, the data generated by density functional theory for selective CO oxidation over Au nanoparticles were modeled using neural networks and logistic regression, both of which demonstrated that the relations between structural properties of the Au nanoparticles and adsorption energies of CO and O_2 on these particles can be successfully established. Finally, the published data by the other investigators were modeled to assess the possibility of predicting the outcome of an experiment, which was not done yet, from the knowledge accumulated in the literature. First, the publications on selective CO oxidation over Cu based catalysts were modeled by neural networks. Then, the publications on the same reaction for noble metal-gold based catalysts were modeled by genetic algorithm supported clustering together with neural networks. In both works, the experimental CO conversions reported in each publication were successfully predicted by the neural networks trained using the data from the remaining publications. The relative significances of the input variables and the major trends associated with these variables were also successfully estimated. It was then concluded that the knowledge accumulated in the published data through years can be extracted using some suitable data mining tools to help the researchers in planning their future experimental works more effectively.

ÖZET

SEÇİMLİ CO OKSİDASYONU VE SU GAZ DEĞİŞİM REAKSİYONLARINA AİT DENEYSEL VE HESAPLAMALI VERİLERDEN VERİ MADENCİLİĞİ TEKNİKLERİ KULLANARAK BİLGİ ÇIKARIMI

Bu çalışmada seçimli CO oksidasyonu ve su-gaz değişim reaksiyonlarına ait deneysel ve hesaplamalı verilere çeşitli veri madenciliği teknikleri uygulayarak bilgi çıkarımı yapılması ve dolayısıyla katalizör tasarımı ve proseslerinin geliştirilmesi amaçlanmıştır. Öncelikle, laboratuarımızda üretilmiş deneysel veriler üç ayrı kısımda incelenmiştir: Pt bazlı katalizörlerde seçimli CO oksidasyonunun sinir ağlarıyla modellenmesi, Au bazlı katalizörlerde aynı reaksiyonun karar ağaçları ve sinir ağlarıyla modellenmesi, Pt bazlı katalizörlerde su-gaz değişim reaksiyonun sinir ağlarıyla modellenmesi. Bu çalışmaların üçünde de, katalizör değişkenlerinin etkisi ve önem dereceleri literatürdeki bilgilerle büyük bir uyum içinde bulunmuştur. Sonraki aşamada, Au kümecikleri üzerinde seçimli CO oksidasyonu için kuantum mekaniksel yöntemlerle hesaplanmış veriler, sinir ağları ve lojistik regresyon yöntemleriyle modellenmiştir. İki metot da Au kümeciklerinin yapısal özellikleriyle kümecikler üzerindeki CO ve O_2 adsorpsiyon enerjilerinin ilişkisini başarılı bir şekilde bulmuştur. Son olarak, başka araştırmacıların yayınlanmış çalışmalarına ait deneysel sonuçları tahmin etmeye yönelik modeller kurulmuş ve literatürde birikmiş bilginin çıkarımı amaçlanmıştır. İlk aşamada Cu bazlı katalizörler üzerinde seçimli CO oksidasyonu sinir ağlarıyla modellenmiştir. Ardından, soy metal ve altın bazlı katalizörlerde aynı reaksiyon genetik algoritma destekli gruplandırma ve sinir ağlarıyla modellenmiştir. İki çalışmada da bilimsel yayınlarda rapor edilmiş deneysel veriler, diğer yayınların sonuçlarıyla eğitilmiş sinir ağları tarafından başarılı şekilde tahmin edilmiştir. Değişkenlerin önem analizi ve bu verilere ait ana trendler de başarılı şekilde çıkartılmıştır. Sonuç olarak, yıllar boyunca literatürde birikmiş olan bilgi, uygun bir veri madenciliği yöntemi kullanarak başarılı şekilde çıkartılabilir ve böylelikle bu konuda çalışma yapacak olan araştırmacıların geleceğe yönelik deneylerini daha efektif bir şekilde planlanması ve tasarlaması sağlanabilir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

r	Pearson correlation coefficient
R^2	Coefficient of determination
α	Momentum term
η	Learning rate
δ	Error

LIST OF ACRONYMS/ABBREVIATIONS

CI	Co-impregnation
CP	Co-precipitation
DFT	Density functional theory
DP	Deposition precipitation
HM	Hydrothermal method
HP	Homogenous precipitation
IWI	Incipient to wetness impregnation
PEM	Polymer Electrolyte Membrane
RMSE	Root mean square error
SGP	Sol-gel precipitation
SI	Sequential impregnation
SP	Sequential precipitation
UG	Urea gelation
UNC	Urea nitrate combustion
SE	Standard error
SSE	Sum of square errors
WGS	Water gas shift reaction
WI	Wet impregnation

1. INTRODUCTION

The world energy demand has rapidly increased together with the increase of the world population in the recent years. Today, most of the energy is supplied by fossil fuels even though they are quite harmful to the environment due to their emissions that cause global warming, air pollution and acid rains. In addition, the petroleum sources are limited and they are estimated to be depleted in the near future. Therefore, it is very important to search new and more sustainable fuel alternatives and energy conversion technologies, such as the use of hydrogen in a Polymer Electrolyte Membrane (PEM) fuel cell.

Hydrogen is the most abundant element in the universe, contrary to that free hydrogen is quite scarce in the earth atmosphere. The efficient production of hydrogen gas has been researched for years, and there are various methods developed such as electrolysis of water, the processing of biomass and biological wastes and thermal decomposition of water. However, the safe storage of hydrogen is technically quite difficult especially for transportation applications. Therefore, on-site generation of hydrogen from a conventional fuel using a fuel processor is considered to be the best option in the near future for mobile and small-medium scale stationary applications of fuel cells [1]. The hydrogen produced this way contains CO, CO_2 and H_2O, all of which can affect the performance of the fuel cell. Among these side products, the presence of CO, even in trace amounts, damages the anode catalyst of the PEM fuel cell; hence, it should be totally eliminated from the hydrogen stream. High and low water gas shift (WGS) reactions, which are generally applied consecutively, can reduce the amount of CO to about 1% (still too high) [2]. Pt, Rh, Pd, Ru and Au based catalysts supported over CeO_2 and Al_2O_3 [3]; Pt over ceria-zirconia [4]; Cu/Al_2O_3 promoted by Mn [5]; Cu and Ni over CeO_2 [6, 7]; Au over ZnO and Fe_2O_3 [8]; Cu-ZnO over Al_2O_3, MgO, SiO_2-Al_2O_3, SiO_2-MgO, β-zeolite, and CeO_2 [9] are some examples that have been studied for WGS reactions within the last decade.

The last step of hydrogen purification is the catalytic selective CO oxidation in the presence of excess amount of hydrogen, which is a crucial process still being improved by the use of more efficient catalysts. There are three main types of catalysts for selective CO

oxidation, which are noble metal catalysts (Pt, Pd, Rh, Ir, etc.), gold based catalysts and copper based catalysts.

Pt catalysts promoted by the oxides of various metals such as Co [10, 11, 12], Ce [12, 13], Mg [14], Mn [15], Fe [16, 17], Ni [10, 18], K [19] and Zr [20] over a suitable support like Al_2O_3 are some part of the extensive studies on noble metal catalysts for the selective CO oxidation. There are also plenty of successful studies in the literature on gold based catalysts: Au over FeO_x [21, 22, 23]; Au over CeO_x [23, 24]; Au over MnO_x [22, 23]; Au over NiO_x and CoO_x, $Mg(OH)_2$, MgO, TiO_x, Al_2O_3 and SnO_x [23]; Au over MgF_2-MgO [25]; Au/Al_2O_3 promoted by MgO, MnO_x and FeO_x [26]; Au/Al_2O_3 promoted by Li_2O, Rb_2O, BaO and CeO_x [27]. Copper based catalysts have been also studied extensively especially in the last decade, due to copper being less precious in terms of economical point of view and being highly active when combined with ceria. There are plenty of works on Cu based catalysts reported to be quite effective for the preferential CO oxidation [28, 29, 30].

Catalyst design is a complex process involving various interacting factors such as active metal type and loading, promoter type and loading, support type, preparation method and pretreatment conditions. Any variation in these factors may cause significant changes in the activity and selectivity of the catalyst. The same is true for the operating conditions; different combinations of reaction variables like temperature, feed flow rate, feed composition and the amount of the catalyst can produce completely different catalytic performance.

Due to these complex factors, coming up with a high activity catalyst for a process is a challenging accomplishment that needs quite many experiments requiring significant time, labor and resources. Although such effort is needed, it is quite difficult to interpret the raw catalytic reaction containing the information about the complex interactions among the catalyst preparation and operating variables. Data mining is the branch of computer science to extract such knowledge, which is too hard to observe by naked eyes, providing the discovery of useful information, such as correlations, trends and patterns from the data.

The applications of the data mining methods have become quite widespread due to the great development in computer hardware and software in the last two decades. Even the methods demanding a massive number of computational calculations are now performed quite easily by the high capacity computer processors. Thus, data mining is now easily applicable in almost every field of science including the field of catalysis. Multiple regression, artificial neural networks and logistic regression methods together with data clustering and classification methods such as decision trees and k-means clustering techniques can be implemented for catalytic data.

A multiple regression model uses linear surfaces such as a plane or hyper plane to approximate the relationship between a continuous target variable and a set of predictor variables [31]. Several successful applications of multiple regression modeling for catalytic data have been reported in the literature [32-36].

Artificial neural networks, which have superior ability on non-linear regression and data approximation, are among the most effective knowledge extraction tools [31, 37]. Some of the successful implementations of neural networks in the area of chemistry were reviewed by Burns and Whitesides [38] in as early as 1993. Similarly, Himmelblau [39] reported numerous works related to the application of neural network modeling in the field of chemical engineering in 2000. A recent publication of Baerns and Holena [40] reviewed and described the use of artificial neural networks for the study of catalytic materials. After the pioneering work of Hattori and Kito [41], numerous successful applications of neural networks have been also reported in the field of catalysis [42-49], including some of our recent works [34, 50-53].

Modular neural networks, which divide the problem into subtasks such as separating the independent input variables into different groups, seems to be also very suitable to model the catalytic works [54]. Similar variables (for example, catalyst preparation or operating variables) can be collected in the same groups, and each group can be processed differently (i.e. connecting to the different hidden layers, using different activation function and so on) and combined later so that the catalytic performance can be modeled more accurately. Another advantage of the modular networks is that they require the use of less number of weights (connections) than the monolithic networks for a given data set. This

can reduce the risk of over-fitting (i.e. memorizing), improve the generalization and prediction ability of the network, and it may also decrease the computational time required for the training [54, 55].

Multiple logistic regression can also be used in the field of catalysis to model the relation between the predictor variables and a categorical response variable [56, 57, 58]. In the last decade, logistic regression has been applied in fields such as microbiology [59, 60], epidemiological studies [61, 62], environmental issues [63, 64], and catalysis [65].

Classification and clustering techniques are invaluable tools for data mining, considering that they can not only derive information from the experimental data but also pre-process the data to improve the effectiveness of regression techniques explained above. One of the data mining tools that can be used for classification of catalytic experimental data is the decision tree, which is a collection of decision nodes connected by branches and extending downward from the root node until terminating in leaf nodes. Attributes are tested at decision nodes and each possible outcome results in a branch. Every branch leads to another decision node or to a terminating leaf [56, 66]. Decision trees can help to derive simple but valuable rules, such as finding the variables, which lead to low, medium or high catalytic performance levels. Although there are various applications of decision trees in diverse areas, the use of them is quite new in the field of catalysis [65, 67, 68].

The use of k-nearest neighbor classification is another example of instance based learning, in which the training data is stored, so that a classification for a new unclassified record may be found simply by comparing it to the most similar records in the training set [31]. Similarly, k-means clustering algorithm is a method to group records into classes, in such way that the members in the same class are similar to each other and dissimilar to the members of the other classes [31]. It is a quite important technique to analyze large and complex data sets in more simple clusters and mostly used for a preliminary step of data mining.

Based on the principle of survival of the fittest, genetic algorithms are one of the most robust algorithms to search for the global optimum solution. First, a random initial population of a definite size, the members of which are the candidates of the solution of the

particular problem, is created. Then, the fittest individuals are selected to reproduce, some of the individuals mutate, and new members immigrate to the population replacing the ones with the worst fitness [56, 69]. At the end, a new generation is formed to replace the old one. During the evolution of the generations, survival strength of the population tends to increase. After the termination criteria are met, the solution with the highest fitness value is stated as the optimum solution [56]. Genetic algorithm can be successfully combined with k-means clustering algorithm to improve the optimization of the clustering process [69, 70].

This dissertation presents the applications of several data mining techniques, individually or together, in the field of catalysis to analyze the experimental data in a more systematic manner and to extract the essential knowledge to improve the catalyst design and testing conditions. In the Literature Survey (Section 2), the details on hydrogen production, low temperature CO oxidation catalysts and various data mining techniques with their algorithms are explained. The details of all the experimental and computational works are presented in the Experimental and Computational Data (Section 3) and Computational Details (Section 4), respectively. The results are presented in three main sections. In the first part, two sets of experimental data on selective CO oxidation (promoted Pt based and promoted Au based catalysts) and one set of experimental data for water gas shift reaction, all of which had been generated in our laboratories, were modeled and discussed.

For the Pt based catalysts, the experimental data for selective CO oxidation over Pt/Al_2O_3 catalysts promoted by the oxides Co, Ce, Mg, Fe, Mn, Zr, K and Ni were studied. Modular neural network modeling was applied to model the experimental data (Section 5.1.1). For the Au based catalysts; the experimental data for the selective CO oxidation over $Au/MO_x/Al_2O_3$ (M=Mg, Mn, Co, Ce, Fe, Ni) were analyzed by decision trees and modular neural networks (Section 5.1.2). Similarly, WGS activity of $Pt-Ce/Al_2O_3$ catalysts in the absence and presence of a second promoter (K, Ni, Co) was also analyzed by modular neural networks to demonstrate that this technique can be used to predict the results of unstudied conditions as well as to understand the effects of preparation and operating conditions on the performance of WGS catalysts (Section 5.1.3).

In the second part of Results and Discussion, the data generated for Au clusters by Density Functional Theory (DFT) are investigated over 2-10 atom Au clusters to determine the structure and activity relationships and investigate the CO and O_2 adsorptions on the catalyst surface. The structures and energies of the anionic, neutral and cationic clusters were determined, and then CO and O_2 adsorption energies were calculated at various sites of the clusters. As a result, a set of user defined descriptors, computationally determined structural properties and computationally determined adsorption energies were formed. Then, the data were modeled using a series of artificial neural networks explaining the structure and activity correlations (Section 5.2.1). Multiple logistic regression methodology was also applied to model the O_2 adsorption on Au_2-Au_{10} clusters, which was the first time such a technique is employed for DFT studies. The modeling was performed in two main sections analyzing the stable-unstable and weak-strong adsorptions. Each modeling was performed in two sub-sections; with the predictor variables of descriptors and with the predictor variables of structural properties (Section 5.2.2).

In the third part of Results and Discussion, the data mined out from the publications in the literature on selective CO oxidation are modeled to extract the valuable experience and knowledge accumulated in the publications and for predicting the outcome of the unstudied experimental conditions. The analysis is presented in two sub-sections: knowledge extraction from Cu based catalysts (Section 5.3.1) and noble metal-gold based catalysts (Section 5.3.2).

2. LITERATURE SURVEY

2.1. Hydrogen Production and Purification

The energy source of a fuel cell is simply the hydrogen, which is generally produced from hydrocarbons (most commonly from methane). There are three processes that can convert hydrocarbons to hydrogen: steam reforming (SMR), partial oxidation and auto-thermal reforming (ATR) [1, 71, 72].

$$C_nH_m + nH_2O \rightarrow nCO + (n + \frac{m}{2})H_2 \quad \text{(Steam reforming)} \tag{2.1}$$

$$C_nH_m + \frac{n}{2}O_2 \rightarrow nCO + \frac{m}{2}H_2 \quad \text{(Partial oxidation)} \tag{2.2}$$
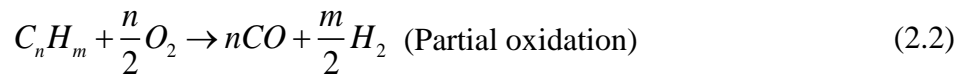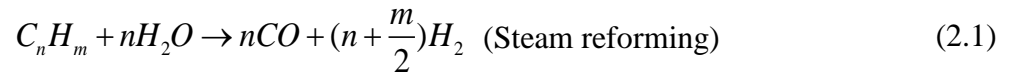
SMR is an endothermic and catalytic conversion of a hydrocarbon and steam to H_2 and CO (Equation 2.1) at 750-800°C. Partial oxidation, on the other hand, is an exothermic and non-catalytic reaction of a hydrocarbon and O_2 to produce a syngas mixture (Equation 2.2). ATR combines steam reforming and partial oxidation in one reactor; the endothermic reforming reactions proceed with the assistance of the internal combustion or oxidation of a portion of the feed hydrocarbons. SMR and partial oxidation processes produce syngas mixtures having different compositions. SMR has a much higher H_2/CO product ratio. Due to this advantage, SMR dominates other syngas producing methods in the market [73, 74]. Although the specific composition of the reformer effluent depends on the technology and operating conditions chosen, steam reforming outlet composition is about 55% H_2, 10% CO, 5% CO_2, and 30% H_2O [73].

The presence of CO in the hydrogen stream is a very serious problem for the fuel cell. Even at trace levels such as 5-10 ppm, CO poisons the noble metal anode of the fuel cell by forming chemisorption bonds with the metal surface [75]; thus, CO should be totally eliminated [2, 76]. It was suggested that most of the CO present in the reformer effluent can be removed by water-gas shift (WGS) reaction (Equation 2.3) followed by

preferential CO oxidation (PROX) for complete elimination of the remaining CO (Figure 2.1) [77].
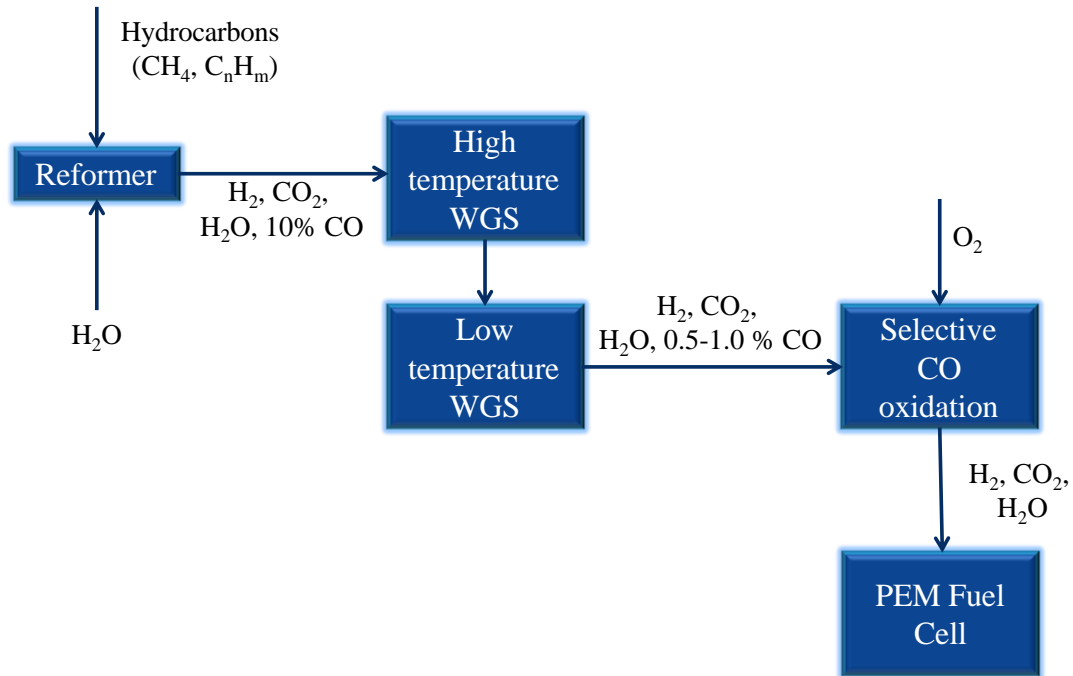
Figure 2.1. Scheme for H_2 production and purification.

Commercially, WGS is conducted in a two-step process: a high temperature shift (operating at 350-550 ^{o}C) using Fe-Cr based catalyst, and if necessary, a low temperature shift to reduce the carbon monoxide content further (operating at 200-300 ^{o}C) using a Cu-Zn catalyst. However, these catalysts are not suitable for fuel cell applications due to their low activity and stability as well as their special pretreatment and regeneration requirements [77]. Hence, more active and stable WGS catalysts need to be developed for PEM fuel cell applications. Pt, Rh, Pd, Ru and Au supported over CeO_2 and Al_2O_3 [3]; Pt over ceria-zirconia [4]; Cu/Al_2O_3 promoted by Mn [5]; Cu and Ni over CeO_2 [6, 7]; Au over ZnO and Fe_2O_3 [8]; Cu-ZnO over Al_2O_3, MgO, SiO_2-Al_2O_3, SiO_2-MgO, β-zeolite, and CeO_2 [9] are some of the examples that have been studied for the low temperature WGS studies within the last decade.

While each water gas shift reaction step reduces the CO content, they also provide additional hydrogen to the fuel cell. High temperature shift reactor reduces the level of CO to 3-5%; in addition, low temperature shift reactor reduces the CO level to 0.5-1% [2, 75].

Although, at the end of these two steps, CO level becomes approximately one tenth of the CO content in the syngas, this amount is still too high to be processed in a fuel cell. Hence, the water gas shift effluent is further purified by the preferential CO oxidation (Equation 2.4). Since there is also significant amount of H_2 present in the same stream, oxidation of H_2 (Equation 2.5) must be as low as possible.

$$CO + H_2O \rightarrow CO_2 + H_2 \text{ (Water gas shift reaction)} \tag{2.3}$$

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 \text{ (Preferential oxidation of CO)} \tag{2.4}$$

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \text{ (Undesired } H_2 \text{ oxidation)} \tag{2.5}$$
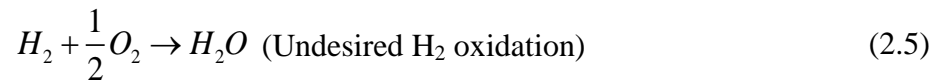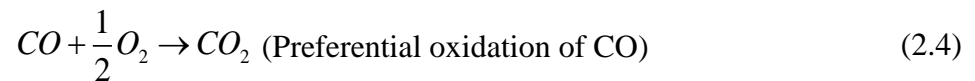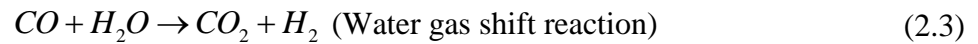
2.2. Low Temperature Preferential CO Oxidation in H_2 Rich Streams

The exit stream of the water gas shift reaction contains a high amount of hydrogen gas together with some amount of CO_2 and H_2O [1]. The possible catalyst that is to be used for preferential oxidation of CO must be able to provide a very high CO conversion in these feed conditions.

There have been hundreds of published papers covering the effects of various factors on CO oxidation activity and selectivity. For example, almost all of the noble metals over a support like Al_2O_3, TiO_2, and MgO have been studied for this purpose [6, 76]. Especially, platinum is the most studied noble metal while gold has drawn more attention in recent years as a more abundant and cheaper alternative [76]. In addition, some Cu based catalyst have also been studied and reported to be quite effective [28, 29, 30].

2.2.1. Noble Metal Catalysts

Examples to noble metal catalysts are Pt, Pd, Rh, Ir or Ru based catalysts, among which Pt is the most commonly studied one. The widely accepted Langmuir–Hinshelwood type reaction mechanism of the CO oxidation on Pt can be described by the following kinetics [78].

- adsorption of CO: $CO + * \rightarrow CO_{ads}$,

- dissociative adsorption of oxygen: $O_2 + 2* \rightarrow 2O_{ads}$,

- reaction (Langmuir–Hinshelwood): $CO_{ads} + O_{ads} \rightarrow CO_{2ads} + *$,

- desorption of CO$_2$: $CO_{2ads} \rightarrow CO_2 + *$

where ''$*$'' denotes a free adsorption site.

H_2 oxidation is strongly inhibited by the presence of CO, because CO chemisorption on noble metal surfaces is much stronger than H_2 or O_2 chemisorption, thus the reaction rate is mostly determined by CO oxidation kinetics [1]. CO covers the metal surface and prevents the adsorption of H_2 or O_2; hence, the reaction does not occur unless the temperature is high enough to desorb some of the CO on the surface allowing the adsorption of O_2. However, increasing the temperature to high levels can also cause the adsorption of H_2 on the metal surface, decreasing the selectivity towards CO oxidation.

There are countless numbers of published studies in the literature searching a high activity noble metal catalyst for preferential CO oxidation. Pt catalysts promoted by the oxides of various metals such as Co [10, 11, 12], Ce [12, 13], Mg [14], Mn [15], Fe [16, 17], Ni [10, 18], K [19] and Zr [20] over a suitable support like Al_2O_3 are some part of the extensive studies on noble metal catalysts for the selective CO oxidation. In addition to these, zeolite supported Pt catalyst [79]; magnesium supported Rh catalyst [80]; silica aerogel supported Pt catalyst [81]; zeolite supported Pd and Ru catalysts [82]; zinc, gallium, silica, ceria supported Pd catalysts [83] are some other published works on noble metal catalysts.

Although Pt based catalysts are quite resistant to H_2 oxidation, increasing the temperature to high levels can cause the adsorption of H_2 on the metal surface (decreasing the selectivity towards CO oxidation). Addition of H_2O usually has a slightly positive effect on CO conversion, which is possibly due to the fact that hydrated support can generate additional free Pt-sites for the CO oxidation [84]. CO_2 present in the feed stream on the other hand, generally has a negative effect on the performance of the catalyst. The magnitude of the negative effect on the activity is strongly dependent on the nature of the

catalyst. The highest conversion values for the Pt based catalysts are usually attained at mid-level temperatures.

2.2.2. Gold Group Catalysts

Au based catalysts exhibit a very high activity at the PEM fuel cell operating temperature; however, they are not very resistant to the presence of CO_2 and H_2O, which cause the deactivation of the catalyst [85]. The activity for H_2 oxidation was only dependent on the gold metal surface exposed, whereas the choice of support oxide also plays a role in CO oxidation. Hydrogen is taken not to chemisorb on Au, however CO chemisorbs weakly on the metal. Thus, the catalytic activity for CO may be a result of the interaction of species adsorbed on the metal and the support. The interaction of Au with the reducible oxide support may change the surface properties of Au particles so that they favor CO adsorption [1].

As in the case of platinum based catalysts, there are plenty of published experimental studies in the literature on gold based catalysts such as Pt, Pd, Rh, Ir or Ru [21, 22, 23]; Au over CeO_x [23, 24]; Au over MnO_x [22, 23]; Au over NiO_x and CoO_x, $Mg(OH)_2$, MgO, TiO_x, Al_2O_3 and SnO_x [23]; Au over MgF_2-MgO [25], Au/Al_2O_3 promoted by MgO, MnO_x and FeO_x [26]; Au/Al_2O_3 promoted by Li_2O, Rb_2O, BaO and CeO_x [27].

Gold based catalysts are more resistant to H_2 oxidation and achieve high activity at temperatures lower than those required by the Pt-group metals. These catalysts are also unaffected by the presence of H_2O; however, they lose their activity in the presence of CO_2, which is explained by co-adsorption of CO_2 on the gold particles [1].

2.2.3. Copper Based Catalysts

Copper based catalysts are composed of cerium-containing mixed oxides (Cu/CeO_2), which were reported to show promising properties in terms of activity, selectivity and resistance to CO_2 and H_2O [86]. Cu/CeO_2 mixed oxide catalysts exhibit precious metal-like behavior in selective CO oxidation. The significant activity enhancement of complete CO

oxidation by these catalysts is a result from the interaction of copper and CeO_2, such that a synergistic effect occurs between these materials [87].

Although, the number of studies on copper based catalysts before the year 2000 is a little bit limited, there is a growing number of papers in this group of catalysts in the recent years [28, 29, 30], probably due to copper being less precious in terms of economical point of view: the cost and limited availability of the platinum or the gold group metals discourage their widespread applications [87].

Copper based catalysts exhibit higher performance at the temperatures near 200°C, even though the selectivity is slightly lower at these temperatures [30]. The addition of CO_2 or H_2O to the reaction stream has generally negative effect on the performance of the catalyst [88].

2.3. Density Functional Theory for Catalyst Development

Density functional theory is a tool that has been widely applied to study the catalytic processes in the recent years; it helps to determine the physical and chemical properties of the catalysts as well as the thermodynamic/kinetic parameters of catalytic reactions from the first principles. The cluster model, which uses very small clusters (with the atom size of 10s) to represent the catalyst, is often employed in DFT studies because it physically suits well to CO oxidation over Au based catalysts: it is well established that Au is active for CO oxidation when it is in the form of nano sized clusters or finely dispersed on metal-oxides [89-94]. Experimental studies indicate that the catalytic systems consisting of Au particles with diameters less than 5 nm dispersed on a support like Al_2O_3, ZrO_2 and TiO_2 in the absence or presence of a promoter such as MgO, MnO_x, and FeO_x are highly active for CO oxidation or selective CO oxidation in hydrogen rich streams [95, 96]. There are numerous computational studies for the CO oxidation over Au catalysts using the cluster model and DFT methods. Some of these studies were focused on the adsorption of CO [97, 98] or O_2 [99, 100] while the others are about the co-adsorption of these species [98] or the reaction between them [101-104].

Although the tools and the parameters involved in DFT computations are quite different from the experimental work, the basic approach quite similar: The effects of some user defined descriptors (equivalent to the preparation variables in experimental work) such as cluster size, charge of the cluster and coordination number of the atom on which the adsorption or reaction take place, are investigated to understand the catalytic system. There are even some established trends and generalized observations that are also supported by the studies in the literature. For instance, it is known that the anionic clusters with even number of Au atoms adsorb O_2 more strongly [99, 105] while the adsorption of CO is stronger for cationic clusters [97]. Similarly, both CO and O_2 prefer the low coordinated Au atoms for adsorption [106, 107]. Again, similar to the experimental studies, it is known that the descriptors affect some structural properties such as binding energy, HOMO-LUMO gap, bond length, and so on [108, 109, 110]. For example, it is commonly believed that the binding energy generally increases with increasing cluster size [108] while HOMO-LUMO gap exhibits a zigzag behavior depending on the presence of unpaired electron [111]. As a result, not only the experimental works but also the quantum mechanically generated data and information can be used for the knowledge extraction to understand the catalytic systems better.

2.4. Data Mining Methods for Knowledge Extraction

Data mining is the discovery of useful information, such as correlations, trends, patterns from the data obtained by experimental and observational methods using statistical and mathematical methods. Some of the common tasks for data mining are description, clustering, classification, estimation (prediction) and association.

Description is the search for patterns and trends lying within the data to gain an insight. Exploratory data analysis (EDA) is a graphical method to find such information. The main role of EDA is to explore the data to reveal its structural secrets by using graphical methods. The particular graphical techniques employed in EDA are often quite simple, consisting of various techniques of: plotting the raw data (such as data traces, histograms and probability plots), or plotting some simple statistics (such as mean plots, standard deviation plots and box plots) [31].

Classification is quite similar to estimation however the target variable is categorical like "good or bad", "hot, mild or cold". Classification usually requires multi-dimensional plots with more than three dimensions and hence the use of more sophisticated methods such as k-nearest neighbor algorithm, decision trees or neural networks [31, 57].

Clustering is the concept used for grouping of records, observations or cases into similar objects. A cluster is a data group in which every data point is similar to each other but dissimilar to the data in other clusters. The purpose of clustering is not to make a prediction or an estimation of the target variable; instead, the entire data is segmented into homogenous subgroups. Clustering is the most common unsupervised data mining method. Some of the clustering methods are k-means clustering and Kohonen Neural Networks [31].

Estimation is the attempt to find a numerical target variable by constructing a model with complete records of the predictors and the target. Simple linear regression, multiple regression, regression trees, k-nearest neighbor methods and artificial neural networks are some of the tools that can be used for estimation. Prediction is completely identical to estimation; sometimes the term prediction is used for the attempt to find the results lying in the future [31, 57].

Association or affinity analysis search for the rules quantifying the relationship between two or more attributes. Association rules take the form "if a thesis is true then this consequence is true". Priori algorithm and the generalized rule induction algorithm are the two possible algorithms that can be applied for association [31].

Next, a brief introduction to some of the data mining techniques will be given in the order of the data mining tasks: classification, clustering and prediction (or estimation). Although artificial neural networks were the most applied method in this dissertation, it will be introduced after the classification and clustering methods (under the prediction methods) in order not to break the order of the data mining tasks.

2.4.1. Classification by Decision Trees

One of the most attractive classification methods is the construction of a decision tree, which is a collection of decision nodes, connected by branches, extending downward from the root node until terminating in leaf nodes. A decision tree starts from the root node, which is at the top of the diagram. Attributes are tested at decision nodes and each possible outcome results in a branch. Every branch leads to another decision node or to a terminating leaf node [31]. In most cases, the interpretation of the results summarized in a tree is very simple. This simplicity is useful not only for purposes of rapid classification of new data, but can also often yield a much simpler "model" for explaining why observations are classified or predicted in a particular manner [112].

Decision trees are constructed by using supervised learning; hence, a training data set is needed with the values of the target variable. The training data set should be rich and varied while the target variable must be discrete or the boundaries must be clear. There are many decision tree algorithms and one of the common one is the C4.5 [31].
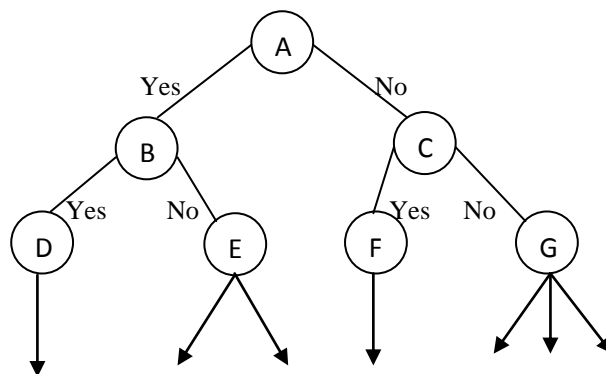
Figure 2.2. A simple decision tree.

The main problem for a decision tree is to decide how to split a root to branches, or a branch to other sub-branches. The C4.5 algorithm uses the concept of information gain or entropy reduction to select the optimal split. First several candidate splits are created. A variable X having k possible values have probabilities of $p_1, p_2, ...,p_k$ and the entropy of X is defined as:

$$H(X) = -\sum_j p_j \log_2(p_j) \tag{2.6}$$

The candidate split *S*, which divides the training set into *T* subsets *(T₁, T₂,...,Tₖ)* has a total entropy calculated as:

$$H_S(T) = \sum_{i=1}^{k} P_i H_S(T_i) \tag{2.7}$$

Where, P_i represents the proportion of the records in subset i and the main aim is find an optimum split that minimizes H_s.

The error rate for a leaf node, which is the fraction of wrong classifications to the total records in that node, is minimized by choosing the best split as described above. The total error rate of entire tree is the weighted average of the individual leaf error rates [56]. Although, a larger tree can have a very small error rate for the training data, the rule it proposes can be too complex and have the lack of generalization ability (the ability to estimate the data that was seen during the training) [56, 66]. For this reason, the tree size is optimized to have a sufficiently low training error and the highest generalization ability [67, 68]. The data can be randomly divided into two equal sets: one set to be used for training and constructing the decision tree model while the other to be used for testing the generalization ability of the model.

2.4.2. Classification by k-Nearest Neighbor Algorithm

The method of k-nearest neighbor is an example of instance based learning, in which the training data is stored, so that a classification for a new unclassified record may be found simply by comparing it to the most similar records in the training set [31].

In k-nearest neighbor modeling, the distance between an unknown sample and each sample in the training set is calculated, and the unknown sample is classified into the class containing the majority of the k-nearest neighbors in the training set as it is shown in Figure 2.3 [113]. The most commonly used distance function is the Euclidian distance:

$$d_{Euclidean}(x, y) = \sqrt{\sum_i (x_i - y_i)^2} \qquad (2.8)$$

Where, $x = x_1, x_2, ..., x_m$, and $y = y_1, y_2, ..., y_m$ represent the m attribute values of two records, which are compared with each other. By calculating the Euclidian distance of an unclassified new record to each training data in the data set, one can determine which records are the most similar to the unclassified data. According to the simple unweighted voting method, the closest candidate k number of data is chosen as the candidate categories for the classification of the new data [31].

For the example in Figure 2.3, k value of 3 is chosen and the closest three points are one A and two B's. Thus, the classification for the new record in "category B" can be concluded with a 66.67% confidence. In contrast to the unweighted voting, weighted voting considers the closeness of the new record to the possible categorical candidates by calculating the inverse square of the distances as shown in Equation 2.9; where, j is the number of possible candidates of the category Z closest to the new record. The weighted voting thereby gives only one categorical candidate with the highest inverse square distance.

$$d_{inverse}(new, Z) = \sum_j \frac{1}{\sqrt{\sum_i (x_i - y_i)^2}} \qquad (2.9)$$
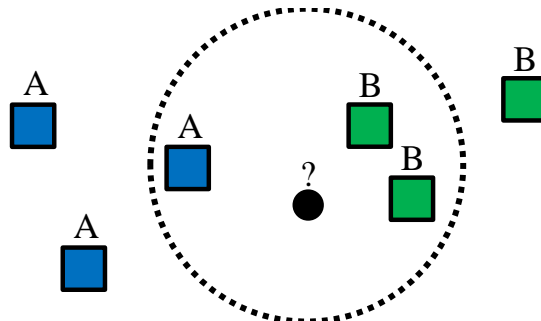
Figure 2.3. An unclassified data and three close neighbors.

2.4.3. Classification by Multiple Logistic Regression

While multiple linear regression modeling approximates the relationship between predictor variables to a continuous response (output) variable, multiple logistic regression makes an approximation between predictor variables to a categorical response [57, 58]. The general form of the multiple logistic regression for "m" independent variables is shown in Equation 2.10.

$$\Pi(X) = \frac{\exp(\beta \cdot X)}{1 + \exp(\beta \cdot X)} \tag{2.10}$$

Where; $\beta \cdot X$ is $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... \beta_m x_m$, $\Pi(X)$ represents the predicted response variable, $\beta_0, \beta_1, \beta_2, ..., \beta_m$ are the coefficients of multiple logistic regression and $x_1, x_2, ..., x_m$ are the predictor variables. Since, $\Pi(X)$ approaches to 1 and 0 when the exponential term approaches to infinity and zero respectively, the numerical value of $\Pi(X)$ indicates the probability of the response variable being positive or negative. The multiple logistic regression coefficients (β) are calculated by the maximum likelihood estimation of the log likelihood function as shown in Equation 2.11; where, n is the number of data points, Y_j and X_j represent the j$^{\text{th}}$ experimental response variable and the predictor variables, respectively [57, 58].

$$\ln\left[L(\beta)\right] = \sum_{j=1}^{n} \left\{ Y_j \ln\left[\Pi(X_j)\right] + (1 - Y_j)\ln\left[1 - \Pi(X_j)\right] \right\} \tag{2.11}$$

The significance of the whole model is examined by calculating the difference between the deviance of the model with the predictor variables ($L(\beta)_F$) and without the predictor variables ($L(\beta)_R$), which is computed as shown in Equation 2.12.

$$G = -2\ln\left[\frac{L(\beta)_R}{L(\beta)_F}\right] \tag{2.12}$$

G follows a chi-square (χ^2) distribution with "m" degrees of freedom. For a 95% confidence interval (α=0.05), if $G \le \chi^2(1-\alpha, m)$ the null hypothesis is true (all the coefficients of multiple logistic regression are zero). On the other hand; if $G > \chi^2(1-\alpha, m)$, at least one perhaps all the coefficients are different from zero.

The standard error (*SE*) of each coefficient is calculated by taking the square root of the diagonal elements of the inverse matrix of *–H,* as shown in Equation 2.13.

$$SE(\beta) = \sqrt{diag\left(inv(-H)\right)} \tag{2.13}$$

Where, *H* is the second order partial derivative matrix of the log likelihood function with regard to $\beta_0, \beta_1, \beta_2, ..., \beta_m$ (Equation 2.14) [57].

$$H = \left[\frac{\partial\left(\ln\left[L(\beta)\right]\right)^2}{\partial\beta_i \partial\beta_j} \right]_{(m+1)\times(m+1)} \tag{2.14}$$

The significance of each predictor variable is calculated by Wald test [57, 58]. The procedure of the test is to first calculate the ratio of the coefficients with their standard errors ($Z_{wald}{}^i$), then checking the two tailed p value of this ratio being whether below 0.05 (for 95% confidence interval) or not. The upper and lower limits of the coefficients are then calculated with a 95% confidence interval (Equation 2.15). The interval shows whether there is a probability for the coefficients ever being zero.

$$\beta_i = \beta_i \mp Z(1-\alpha/2)\cdot SE(\beta_i) \tag{2.15}$$

Odds ratio, which approximates the change of the probability of the response variable with respect to a unit change of the predictor variables, is used to analyze the variable effects.

2.4.4. Optimization by Genetic Algorithms

Genetic algorithms were inspired by Darwin's theory of evolution. They can be applied to solve a variety of optimization problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, non-differentiable, stochastic or highly nonlinear. Genetic algorithms are one of the multi-dimensional global optimization methods, and they have less tendency to become stuck in local minima [56].

Genetic algorithms work with a population $P=(x_1, x_2,...,x_p)$, in which each population member (x_i) is called an individual member containing the values of the variables of the function to be optimized [114]. Solutions from one population are taken and used to form a new population. This is motivated by an expectation that the new population will be better than the old one (i.e. more close to the optimum solution).

The genetic algorithm uses three main types of rules to create the next generation from the current population: selection, crossover and mutation. Selection is the process to choose the individuals, called parents that contribute to the population at the next generation. These are the candidate solutions to the problem and the fitter the chromosome the more likely it will be selected for reproduction. Crossover combines two parents to form children for the next generation. Mutation applies random changes to individual parents to form children [56].

Genetic algorithm starts by specifying a crossover probability of p_c (usually higher than 0.7), and a mutation probability of p_m (approximately 0.001). Then, the initial population is generated containing n chromosomes (candidate solutions) each having the length L (number of variables in the optimization problem). The function to be optimized (fitness function), f(x), is calculated for each chromosome in other words for each candidate solutions. The parents that will create the next generation can be selected by various methods such as the roulette wheel method. The procedure starts with the calculation of each chromosome's probability of selection (Equation 2.16). Then, parent chromosomes are chosen randomly according to these probabilities [56].

$$p_{si} = \frac{f(x_i)}{\sum_i f(x_i)} \qquad (2.16)$$

The next step is the crossover that occurs on the randomly chosen locus (crossover point), with the probability of p_c. There are several crossover methods and some of them are described below. For the single point crossover, one crossover point is selected, binary string from beginning of chromosome to the crossover point is copied from one parent and the rest is copied from the second parent (Figure 2.4). For the two point crossover, two crossover points are selected, binary string from beginning of chromosome to the first crossover point is copied from one parent, the part from the first to the second crossover point is copied from the second parent and the rest is copied from the first parent (Figure 2.5). For the uniform crossover, bits are randomly copied from the first or from the second parent (Figure 2.6).
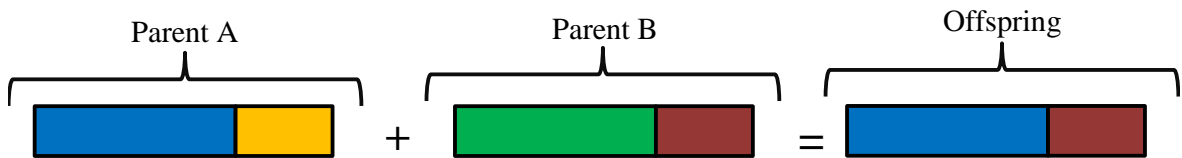
Figure 2.4. Single point crossover.
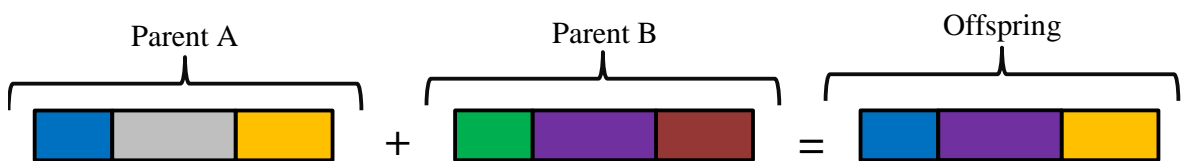
Figure 2.5. Two point crossover.
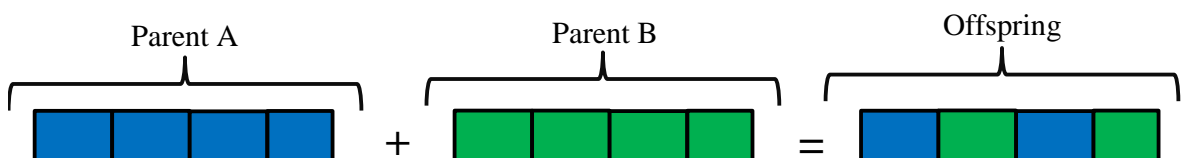
Figure 2.6. Uniform crossover.

The next step is the mutation, in which each of the offspring on each locus is randomly mutated with the probability of p_m. This increases the genetic variety, and helps

to prevent to be stuck in local optimal point while the solution is evolving towards the optimum point.

Selection, crossover, mutation steps are repeated until the number of individuals equal to the initial population size is achieved. This is called a new generation, which is entirely different than the prior generation. Each cycle through this algorithm is another generation and after testing several generations, a chromosome giving the best fitness function is found. In other words, the optimum solution of the optimization problem is achieved [56].

2.4.5. Clustering by k-means Clustering

K-means clustering is an unsupervised data mining method that groups the data into clusters, where the records are similar to those in the same cluster and dissimilar to those in the other clusters. Clustering is often used as a preliminary step in data mining process, with the resulting clusters being used as further inputs into a different technique such as neural networks [31].

In order to prevent any attribute dominate the others, the data should be normalized before applying k-means clustering as shown in Equation 2.17 (min-max normalization). Then, the degree of similarity between two data points is determined by calculating the Euclidian distance between them (Equation 2.18).

$$X^* = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{2.17}$$

$$d_{Euclidian} = \sqrt{\sum_i (x_i - y_i)^2} \tag{2.18}$$

Where; $x = x_1, x_2, ..., x_m$ and $y = y_1, y_2, ..., y_m$ represent the m attribute values of two records. k-means clustering algorithm is shown in Figure 2.7 [31].

Determining the number of clusters that the data is to be separated is a problem in k-means clustering algorithms. With a small number of k, no distinctive information can be

obtained from the data and the total SSE will be high. On the other hand, with a k value that is too high, local and common behavior of the records will be missed and there is a risk for even the noisy data to be clustered separately [31]. Hence, an optimal number of clusters is needed for a particular clustering problem.
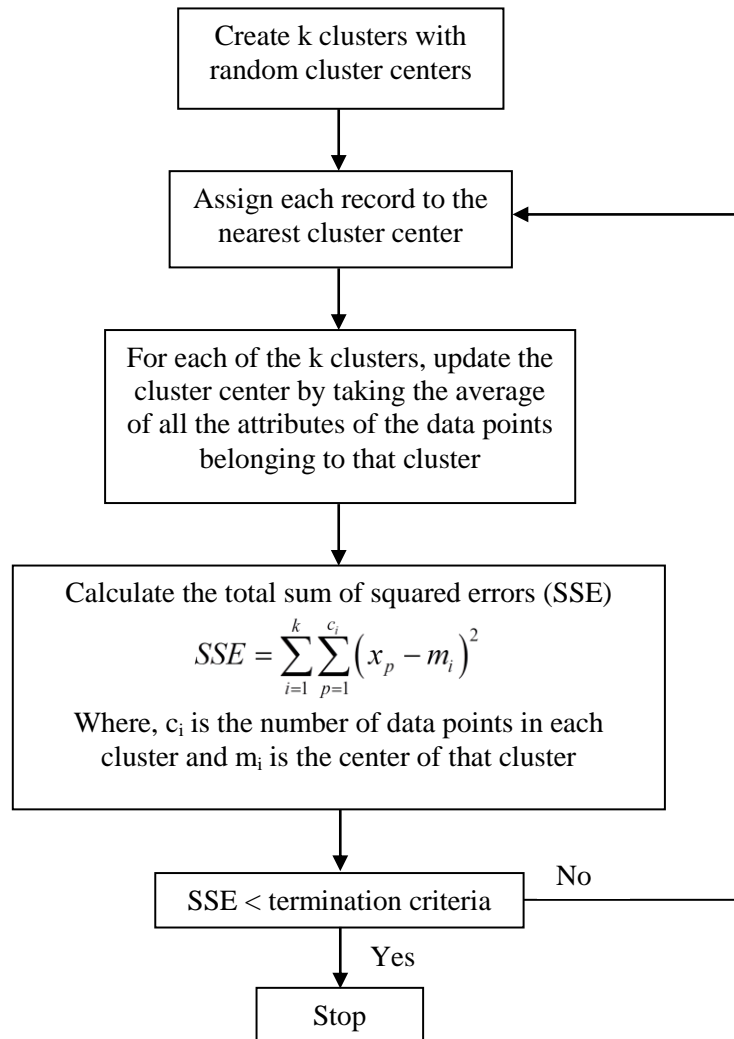
Figure 2.7. k-means clustering algorithm.

2.4.6. Genetic Algorithm Supported Clustering

The traditional k-means clustering algorithm techniques, which group the clusters according to the Euclidian distances as shown above, can easily be stuck to a local optimal solution due to the random influence of the initial clusters. Hence, in order to find the global optimum solution the k-means algorithms can be combined with genetic algorithms,

the combination of which is reported to be superior to the conventional k-means algorithm [69, 70].

The genetic algorithm supported clustering algorithm that was developed for general purpose is shown in Figure 2.8. First, "k" numbers of cluster centers are created, and then these cluster centers are combined to form a string of the population (candidate solution of the clustering problem). The procedure is repeated until "p" numbers of strings are formed. The next step is to calculate the Euclidian distance (Equation 2.18) of each record (data to be clustered) to the closest cluster center in each string. The sum of the distances indicates the sum of squared errors (SSE) of that string. The lower the SSE value of a string, the more successful the placement of the entire data to the clusters of that string is. Thus, the fitness of each string is SSE^{-1}, which is the objective function to be maximized.

Selection, crossover and mutation processes are applied in each step to increase the fitness of the population like the traditional genetic algorithms. Based on a crossover probability, the strings that are going to crossover with each other are selected by the Roulette Wheel method according to their fitness values [56]. Then, each string is mutated based on a mutation probability, which increases the genetic variety by creating different candidate solutions. In addition to these, elimination and immigration are also processed through the algorithm. In order to ensure that none of the strings contain empty clusters (without any member), such strings are eliminated from the population and random new strings immigrate to the population as substitutes. Finally, a new generation is formed but whether this is going to be accepted or not depends on the comparison of the average fitness of the new generation with the old one. If the average fitness improves, the generation is accepted; otherwise the old one is called back; thus, elitism is implemented on each generation [69, 70]. The termination criterion is the total number of generations created. The string with the highest fitness value ever encountered in all the generations is the optimum solution.

```
┌─────────────────────────────────┐
│  Create an initial population    │
│  of size p, (p strings each      │
│  composed of randomly            │
│  generated k cluster centers).   │
└─────────────────────────────────┘
              │ i = 0
              ▼
┌─────────────────────────────────┐
│  For each string, calculate the  │
│  total Euclidian distances       │
│  between each record and the     │
│  closest cluster center to that  │
│  record (the total SSE of the    │
│  string).                        │
└─────────────────────────────────┘
              │
              ▼
┌─────────────────────────────────┐
│  Calculate the fitness of each   │
│  string                          │
└─────────────────────────────────┘
```

$$f_{1,2,\dots p} = \frac{1}{SSE_{1,2,\dots p}}$$

```
┌─────────────────────────────────┐
│  Apply crossover, based on the   │
│  crossover probability μc.       │
│  Select parents by the roulette  │
│  wheel method (according to      │
│  their fitness values) and       │
│  reproduce them (one point       │
│  crossover).                     │
└─────────────────────────────────┘
              │
              ▼
┌─────────────────────────────────┐
│  Apply mutation on each string,  │
│  based on the mutation           │
│  probability, μm.                │
└─────────────────────────────────┘
              │
              ▼
┌─────────────────────────────────┐
│  Substitute the strings that     │
│  have empty clusters by          │
│  immigrating randomly created    │
│  new strings.                    │
└─────────────────────────────────┘
              │
              ▼
┌─────────────────────────────────┐
│  Is the average fitness of all   │
│  the strings maximum up to now?  │
└─────────────────────────────────┘
       │ Yes              │ No
       ▼                  ▼
┌──────────────────┐  ┌──────────────────┐
│ Accept the new   │  │ Reject the new   │
│ generation       │  │ generation, call │
│                  │  │ back the previous│
│                  │  │ one              │
└──────────────────┘  └──────────────────┘
       │                  │
       ▼                  ▼
┌─────────────────────────────────┐      ┌──────────────────────┐
│  Termination criteria reached?   │─Yes→│ Select the string    │
│                                  │      │ with the highest     │
│                                  │      │ fitness as the       │
│                                  │      │ optimum solution     │
└─────────────────────────────────┘      └──────────────────────┘
              │ No
              ▼
         ┌──────────┐
         │  i=i+1   │
         └──────────┘
```

Figure 2.8. Genetic algorithm supported clustering algorithm.

2.4.7. Estimation and Prediction by Multiple Linear Regression

Multiple regression is a method that linearly approximates the relationship of multiple independent predictor variables with a continuous target variable. The general form of multiple linear regression is shown in Equation 2.19 [56].

$$y = \beta \cdot X \tag{2.19}$$

Where; $\beta \cdot X$ is $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... \beta_m x_m$, y represents the predicted response variable; $\beta_0, \beta_1, \beta_2, ..., \beta_m$ are the coefficients of multiple regression and $x_1, x_2, ..., x_m$ are the predictor variables. Finding the coefficients of multiple regression is quite straightforward but the general form should be represented in matrix form (Equation 2.20).

$$\left[y \right]_{n,1} = \left[X' \right]_{n,(m+1)} \times \left[\beta' \right]_{(m+1),1} \tag{2.20}$$

Where; n is the number of data points, m is the number of predictor variables, $\left[\beta' \right]_{(m+1),1}$ is the coefficient matrix, composed of $\beta_0, \beta_1, \beta_2, ..., \beta_m$, $\left[X' \right]_{n,(m+1)}$ is the matrix of predictor variables, in which the first column is composed only of ones in order to achieve a constant term when multiplied by the coefficient matrix. Finally, the calculation of the multiple regression coefficients is shown in Equation 2.21 and Equation 2.22, where X'^T is the transpose of the matrix, X'.

$$X'^T \times X' \times \beta' = X'^T \times y \tag{2.21}$$

$$\beta' = \left(X'^T \times X' \right)^{-1} \times X'^T \times y \tag{2.22}$$

2.4.8. Estimation and Prediction by Artificial Neural Networks

Artificial neural networks, which were inspired from the learning that occurs in biological neurons, have superior ability on non-linear regression and data approximation; hence it is one of the most effective knowledge extraction tools [115]. Before explaining

what is done in artificial neural networks, the mechanism of biological neurons should be well understood. A typical biological neuron collects signals from others through a host of fine structures called dendrites (Figure 2.9). The neuron sends out spikes of electrical activity through a long, thin stand known as an axon, which splits into thousands of branches. A particular neuron is simple in structure; however, a very large number of neurons connected with each other can perform any complex task. The total number of neurons in the central nervous system ranges from under 300 for small free-living metazoans such as round worms, 30–100 million for the common octopus and small mammals such as mice to well over 200 billion for whales and elephants. Estimates for the human brain averages 95–100 billion [116]. An artificial neuron is very similar to its biological counterpart; it simply mimics the type of nonlinear learning that occurs in the networks of neurons found in nature [37].

Probabilistic neural networks, radial basis function networks, kohonen networks, hetero-associative networks, recurrent networks and modular neural networks are some types of neural networks [115]. The most commonly applied one is the multilayer feed forward perceptron network.

2.4.8.1 Multilayer Feed Forward Perceptron Network. A typical feed-forward neural network consists of three layers named as input, hidden and output layers (Figure 2.9). Input layer is the place where the vector of predictor variable values $(x_1...x_p)$ is presented to the input layer. In addition to the predictor variables, there is a constant input having a value of 1, called the bias that is fed to each of the hidden layers. Hidden layers are where the actual processing is done via a system of weighted connections, and output layer is where the final solution is given out.

What is done in a single neuron is quite simple like it is shown in Figure 2.10. First, input variables are multiplied by corresponding weights and the sum is calculated (Equation 2.23). Then, by using an activation function, such as logistic sigmoid function, (Equation 2.24) the output is calculated. The output of one neuron becomes input to another neuron [37].
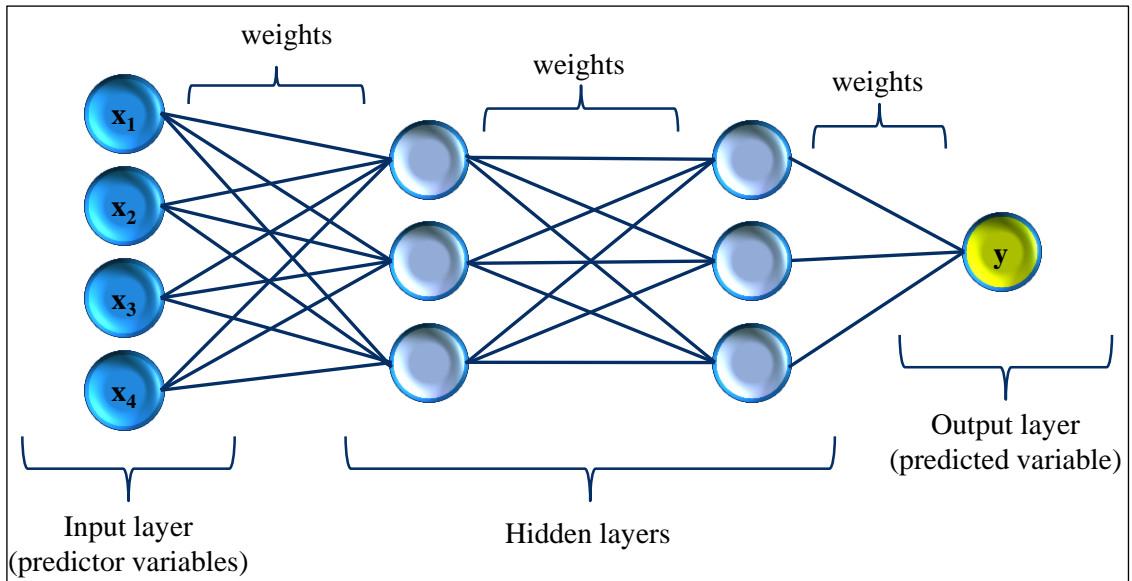
Figure 2.9. A typical feed forward neural network.
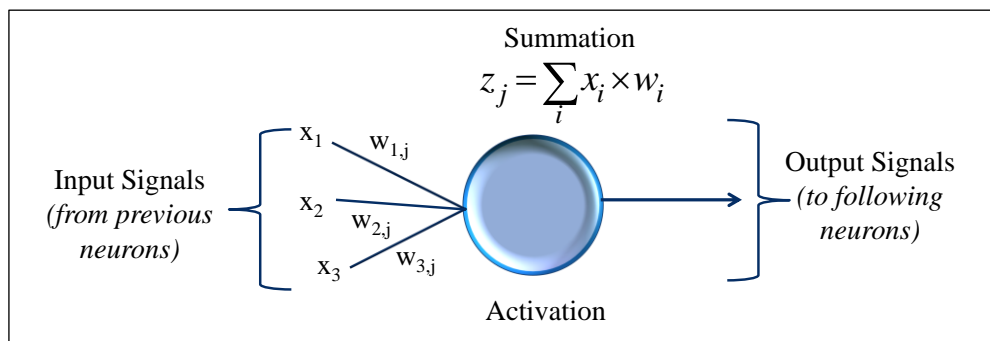
Figure 2.10. A single neuron.

$$net_j = \sum_i w_i \times x_i \tag{2.23}$$

$$y = \frac{1}{1 + \exp(-net_j)} \tag{2.24}$$

Supervised neural networks learn by examples thus require a large data set including the target values. As each observation from the training set is processed through the network an output value is produced. Comparing the predicted output with the real target value gives the error. The comparison can be done mathematically by calculating the sum of square error.

$$E = \frac{1}{2}\sum_j \left(t_j - y_j\right)^2 \qquad\qquad (2.25)$$

The main aim of the learning process is to adapt the weights of the neural network so that any given input can produce an output with small error. By processing the sum of square error through the network the network can adapt its weights. There are many learning algorithms and the most common one is the backpropagation learning algorithm, which is a simple gradient descend technique that minimizes the mean squared error defined in Equation 2.25. However, there are also many various second order approaches such as Levenberg-Marquardt, which is an extremely faster method [117, 118].

The backpropagation algorithm defines two sweeps of the network (Figure 2.11): first a forward sweep from the input layer to the output layer, and then a backward sweep from the output layer to the input layer. The backward sweep is similar to the forward sweep, except that the error values are propagated back through the network to determine how the weights are to be changed during training. During the backward sweep, values pass along the weighted connections in the reverse direction to that, which was taken during the forward sweep: for example, one particular neuron in a layer will send activation to every unit in the next layer during the forward sweep, likewise during the backward sweep that neuron will receive error signals from every unit in the next layer [37].

The first stage of the backpropagation algorithm is to initialize the weights to small random values, (i.e. between -0.3 and +0.3). During the backpropagation process, weights of the neural network are updated with each training data point in the direction that leads to the most rapid decrease in the error (Equation 2.25). The weight update process is given in Equation 2.26-31; where η is the learning rate (a real number), x is the signal coming to the neuron and $\sum\left(W_{jk} \times \delta_j\right)$ is the error propagated from the output neuron to the hidden layer neurons [37].

Step 1
Read first input pattern and
associated output pattern

Step 2
For input layer assign as net input
to each unit its corresponding
element in the input vector. The
output for each unit is its net input.

$$net_j = w_0 + \sum_{i=1}^{n} x_i w_{ij}$$

Step 3
For the first hidden layer
units calculate the net input

$$o_j = \frac{1}{1 + \exp(-net_j)}$$

These steps are
repeated for all
subsequent hidden
layers

Step 4
For the first hidden layer
units calculate the net
output

$$\delta_j = (t_j - o_j) \times o_j \times (1 - o_j)$$

Step 5
For the output layer units calculate
the net input and output

This step is
repeated for all
subsequent hidden
layers

$$\delta_j = o_j \times (1 - o_j) \sum_k \delta_k w_{kj}$$

Step 6
For last hidden layer calculate the
error for each unit

No

$$\Delta w_{ij}(n+1) = \eta \times \delta_j \times o_i + \alpha \times \Delta w_{ij}(n)$$

Last Pattern?

Step 7
For all layers update weights for each unit

No

Yes

Maximum
Iteration
Number
Reached?

Yes

Stop

Figure 2.11. The backpropagation algorithm.

$$w_{new} = w_{old} + \eta \frac{\partial E}{\partial w} \tag{2.26}$$

$$\frac{\partial E}{\partial w} = \delta \times x \tag{2.27}$$

$$w_{new} = w_{old} + \Delta w \tag{2.28}$$

$$\Delta w = \eta \times \delta \times x \tag{2.29}$$

$$\delta = output_j \left(1 - output_j\right)\left(actual_j - output_j\right) \quad \text{for output layer nodes} \tag{2.30}$$

$$\delta = output_j \left(1 - output_j\right) \times \sum \left(W_{jk} \times \delta_j\right) \quad \text{for hidden layer nodes} \tag{2.31}$$

$$\Delta w(n+1) = \eta \times \delta \times x + \alpha \times \Delta w(n) \tag{2.32}$$

After repeating this algorithm with all the samples (training data points) for many times, the sum of E gradually converges to a minimum value, so that the network becomes a fully trained one. An additional parameter that is commonly used for neural network training is the momentum term (α). This term adds a proportion of the previous weight change to the weight update rule, which prevents the oscillation of the weight changes (Equation 2.32). Hence, the weight change for pattern (n+1) is dependent on the weight change for pattern (n).

2.4.8.2 Levenberg-Marquardt Method of Training. The method of Levenberg Marquardt training is a blend of gradient descend and Gauss-Newton iteration. The error achieved by comparing the predicted output with the real target value is given in Equation 2.33. The Jacobian matrix of the error vector for all the patterns is shown in Equation 2.34, where w consists of all the weights of the neural network (Equation 2.35). The weight update rule is presented in Equation 2.36, where, $J^T.J$ is the Hessian matrix for the error vectors and η is the Levenberg-Marquardt learning parameter.

When η is zero, the method becomes exactly the same as Newton's method using only the Hessian matrix. When η is large, the method becomes the same as the gradient descent with a small step size [117, 118]. Gradient descent method has higher performance far from the minimum and Newton's method is faster and more accurate near an error minimum. Thus, η is decreased after each successful step and is increased only when a

tentative step increases the error. In this way, error always decreases after each iteration [117, 118].

$$E = \frac{1}{2}\sum_j \left(t_j - y_j\right)^2 = \frac{1}{2}\sum_j e_j^2 \tag{2.33}$$

$$J = \begin{bmatrix} \dfrac{\partial e_1}{\partial w_1}, \dfrac{\partial e_1}{\partial w_2}, \dots \dfrac{\partial e_1}{\partial w_N} \\[2mm] \dfrac{\partial e_2}{\partial w_1}, \dfrac{\partial e_2}{\partial w_2}, \dots \dfrac{\partial e_2}{\partial w_N} \\[2mm] \dots\dots\dots\dots\dots\dots \\[2mm] \dfrac{\partial e_n}{\partial w_1}, \dfrac{\partial e_n}{\partial w_2}, \dots \dfrac{\partial e_n}{\partial w_N} \end{bmatrix} \tag{2.34}$$

$$w = \left[w_1, w_2, w_3 \dots w_N\right]^T \tag{2.35}$$

$$w_{new} = w_{old} + \left(J^T \cdot J + \eta \cdot I\right)^{-1} \cdot J^T \cdot e \tag{2.36}$$

<u>2.4.8.3. Partially Connected Neural Networks.</u> Modularity is a very important concept in nature. Modularity can be defined as subdivision of a complex object into simpler objects. The subdivision is determined either by the structure or function of the object and its subparts. Modularity in neural networks is linked to the notion of local computation. That means that each module is an independent system interacting with other in a whole architecture in order to perform a more complex function. The structure of the modular system is similar to architectures known from logical neural networks. The new network is not fully connected and therefore the number of weight connections is much less than in a monolithic multilayer perceptron [54, 55].

The term ''Modular neural networks'' is very fuzzy. It is used in a lot of ways and with different structures. Everything that is not monolithic is said to be modular. One of the main ideas of this approach is to construct a network of modules where all the modules are also neural networks. The architecture of a single module is simpler and smaller than the one of a monolithic network. The tasks are modified in such a way that training a subtask is easier than training the complete task (Figure 2.12). The modules can be processed separately having different number of neurons and different activation functions,

and their outputs can be combined in another layer to be processed together (working in parallel) [54, 55].
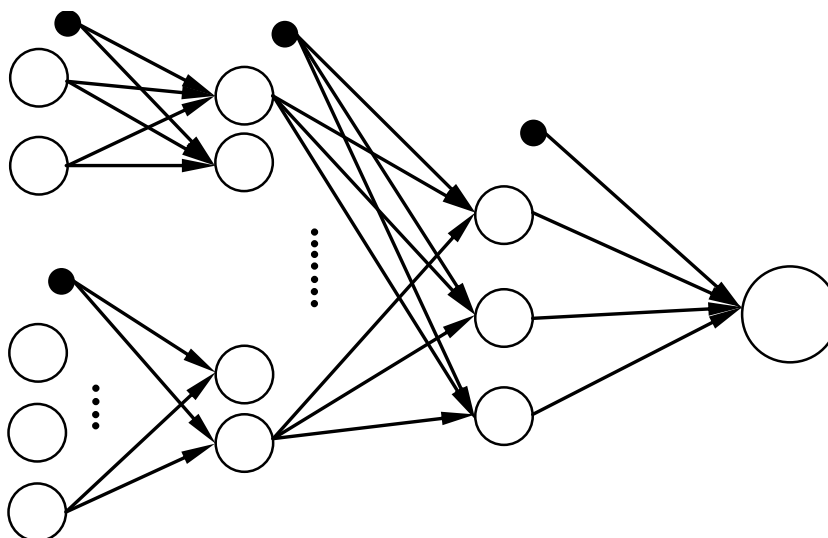
Figure 2.12. A partially connected modular neural network architecture.

2.4.9. Additional Materials for Neural Networks and Statistical Parameters

2.4.9.1. k-fold Cross Validation. Testing is very important for neural networks since it gives the degree of generalization accuracy. k-fold cross validation method starts with the division of the whole data set into k subsets randomly, and then the network is trained using k-1 subsets while the remaining one subset is used as the test data. This procedure is repeated until all the subsets are tested with the rest of the data. Finally, the k pieces of the test data are merged and the calculation of the RMSE within the entire test data eventually provides the testing error, or in other words the generalization accuracy of the neural network [51, 53].

2.4.9.2. Input Significance. The test of significance for the continuous input parameters can be done by partial differentiation of the fully trained neural network, which was reported as a very successful tool for the identification of primary factors controlling catalytic activity [47, 48]. The procedure of the method is as follows: First, the local importance of the ith input variable on k^{th} catalyst $(LRI_{i,k})$ is calculated as shown in Equation 2.37, where f is the output function and x_i is the ith input data. Then, the global relative importance of the ith input variable (GRI_i) is calculated (Equation 2.38) [47, 48]. The partial differentiation

operation can be done numerically by using higher order approximations such as using five point method as shown in Equation 2.39 (in this work with $h=10^{-10}$).

$$LRI_{i,k} = \frac{\left(\partial f / \partial x_i\right)_k}{\sum_i \left(\partial f / \partial x_i\right)_k} \tag{2.37}$$

$$GRI_i = \frac{\sum_k \left|\left(\partial f / \partial x_i\right)_k\right|}{\sum_i \sum_k \left|\left(\partial f / \partial x_i\right)_k\right|} \tag{2.38}$$

$$f'(x) \approx \frac{f(x-2h)-8f(x-h)+8f(x+h)-f(x+2h)}{12h} \tag{2.39}$$

The procedure of the change of root mean square technique, which can be applied to find the relative significances for both continuous and categorical variables, begins with the removal of one of the input variables. Then the network is trained with the remaining input groups. After the training is complete, the RMSE value of the model calculated in the absence of this variable is compared with the value obtained in the presence of all inputs, and the difference is used as the indicator of the significance of this variable [51, 53].

2.4.9.3. Pearson Correlation and Coefficient of Determination. Pearson correlation coefficient, r, is the measure of multicollinearity, which indicates whether there is a directly proportional or inversely proportional relationship between two variables. The Pearson correlation is calculated by Equation 2.40; where N is the number of data points, x and y are the data pairs the correlation of which is to be found. The value of r is between -1 and +1. While negative values of r indicate inverse proportionality, positive values indicate direct proportionality [56]. The degrees of correlation for different values of "r" are shown in Table 2.1.

$$r = \frac{N\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{\sqrt{\left[N\sum x_i^2 - \left(\sum x_i\right)^2\right]\left[N\sum y_i^2 - \left(\sum y_i\right)^2\right]}} \tag{2.40}$$

Table 2.1. Rough rule for determining the presence of correlation.

r value	Correlation
greater than 0.7	positively correlated
0.33 to 0.7	mildly positively correlated
-0.33 to 0.33	not correlated
-0.7 to -0.33	mildly negatively correlated
less than -0.7	negatively correlated

The coefficient of determination, r^2 or R^2, represents how well a regression approximation fits to the experimental data. The value of R^2 can be between 0 and 1, and a value of 1 indicates a perfect fit [56]. The mathematical formula used to calculate the coefficient of determination is shown in Equation 2.41; where, y_i represent the ith experimental data, y_i is the estimated value of the corresponding experiment and $\overline{y}$ is the mean value of the experimental data.

$$R^2 = 1 - \frac{\sum\left(y_i - y_i\right)^2}{\sum\left(y_i - \overline{y}\right)^2} \tag{2.41}$$

2.5. Knowledge Extraction from Catalysis by Using Data Mining Methods

In the last two decades, the applications of the data mining methods have become quite widespread due to the great development in computer hardware and software. Even the methods demanding a massive number of computational calculations are now performed quite easily by the high capacity computer processors. Thus, data mining is now easily applicable in almost every field of science for the discovery of useful information, such as correlations, trends or patterns. Similarly, the application of data mining techniques in the field of catalysis has also become prevalent in the recent years. Multiple regression, decision trees, artificial neural networks are the most commonly applied methods, which help the researcher to analyze their data in a more systematic manner and to extract the essential knowledge to improve the catalyst design and testing conditions.

2.5.1. Multiple Regression Modeling

Istadi and Amin studied the catalytic conversion of methane to ethane and ethylene using carbon dioxide as an oxidant (carbon dioxide oxidative coupling of methane). Central composite design was employed to create the experimental data set and response surface methodology was applied to analyze the results [32]. A quadratic polynomial regression was constructed to model the CH_4 conversion, C_2 selectivity and C_2 yield by using two catalyst preparation variables (CaO and MnO amounts) and two operating variables (reaction temperature, CO_2/CH_4 ratio in the feed). In order to find the optimum parameters leading to maximum CH_4 conversion, C_2 selectivity and C_2 yield, a simultaneous multiple output optimization was performed on the regression models using weighted sum of squared objective functions method. It was found that using CO_2/CH_4 ratio of 1.99, reactor temperature of 1127 K, 12.78 wt.% CaO and 6.39 wt.% MnO leaded to the optimum values of C_2 selectivity and yield, which were 76.56% and 3.74%, respectively [32].

Du *et al.* studied the catalytic partial oxidation of methane to formaldehyde and modeled the formaldehyde selectivity, methane conversion and the yield of formaldehyde from five continuous variables: reaction temperature, vanadium loading, gas hourly space velocity, reactant ratio (CH_4/O_2) and reaction pressure by using full quadratic multiple regression [33]. The data set was prepared by applying a statistical design method covering the whole range of variables in a well-proportioned manner. Multiple regression models were created and t statistics were performed to eliminate the regression parameters having smaller level of significance. It was reported that temperature, CH_4/O_2 ratio, and pressure are the three most significant factors affecting the selectivity of formaldehyde. The optimal formaldehyde selectivity was achieved at high temperature with low vanadium loading or at low temperature with high vanadium loading. Contrary to that, methane conversion was favored at high temperature, high vanadium loading and high pressure, but low gas hourly space velocity and low CH_4/O_2 ratio. The optimum value of formaldehyde space time yield, which is the combined effect of the formaldehyde selectivity and methane conversion, was achieved at the highest levels of all the five input variables. The regression model successfully predicted all these trends successfully. It was concluded that statistically designed data can improve the success of the multiple regression models [33].

Günay and Yıldırım employed multiple regression and artificial neural networks to model the activity of Pt based CO oxidation catalysts by using the catalyst preparation variables (Pt, Co, Ce amounts; calcination temperature and time) [34]. The prediction accuracies of the two methods were compared and multiple regression modeling was found to be less accurate. Hence, additional analyses on the data were performed by using neural network models.

Olutoye and Hameed studied the transesterification of palm oil to produce biodiesel by using heterogeneous catalysts composed of the mixed oxides MgO, ZnO and K [35]. Central composite design was used to prepare the experiments and response surface methodology was applied to analyze the results. Multiple regression modeling was employed to analyze the effects of Mg/Zn ratio, calcination temperature and time on the conversion of palm oil (triglyceride) to methyl ester, which was used as the degree of catalyst efficiency. It was reported that high triglyceride conversion was found to be favored at low Mg/Zn ratio and calcination temperature. The optimum Mg/Zn ratio and calcination temperature were found as 1/4.81 and 460 °C, which leaded to 73% conversion. The high conversion value was attributed to the synergetic effect of mixed oxides MgO, ZnO and K [35].

Chen *et al.* worked on the oxidation of NO by using MnO_x catalysts supported on rice husk ash [36]. The experiments were prepared by using quadratic regression rotatable orthogonal design. The catalytic activity was modeled by multiple regression from catalyst preparation variables (calcination temperature and time, amount of MnO_x and incineration temperature of rice husk). According to the regression results, the significances of the variables were found from the highest to the lowest one as calcination temperature, amount of MnO_x, incineration temperature of rice husk and calcination time. As the calcination temperature increased CO conversion was observed to decrease, which was attributed to the increase of MnO_x size on the surface of the catalyst at high calcination temperatures. Increasing the amount of MnOx in the catalyst or the incineration temperature of the rice husk was reported to improve the catalytic activity, and all these trends were observed by the multiple regression results [36].

2.5.2. Decision Tree Modeling

Using decision trees or regression trees are very recent approaches in the field of catalysts and the amount of studies are limited. Corma *et al.* applied decision trees on the high throughput catalytic data (epoxidation catalysts based on mesoporous titanium silicate materials) together with several different data mining techniques [65]. First dimension reduction, using an unsupervised method (Kohonen networks), was employed on the experimental database, and then the data was classified by decision trees. The object was to determine the yield of cyclohexene epoxide for the optimal material synthesis parameters, which were the molar concentrations of the components of starting gel. The yield of epoxide was classified into five categories from very bad to very good. Decision tree technique applied alone was reported to give 89.8% accuracy, however applying decision tree after clustering with Kohonen networks resulted 100% classification accuracy [65].

Cukic *et al.* studied the influence of the catalyst preparation variables on the performance of Pd/Al_2O_3 catalyst for the hydrogenation of 1,3-butadiene by using regression trees [67]. The complete set of data was split iteratively into meaningful subsets, creating a non-linear regression model. The model was represented by a sequence of questions that can be answered with yes or no, combined with a set of fitted response values. The best non-linear regression-tree model, shown in Figure 2.13, was chosen with respect to the smallest generalizing error, which was estimated by cross-validation method (to avoid over-fitting of the model). The stirring speed (n), the pH value of the solution, impregnation time (t_i), the ratio between volume of solution and pore volume of support (V_i/V_p), heating rate of calcination (s_c), heating rate of drying (s_d), temperature of drying (T_d), drying time (t_d), temperature of calcination (t_c) and the time of calcination (t_c) were the ten catalyst preparation conditions affecting the conversion of butadiene. The tree model in this study provided the prediction of the catalytic activity from various preparation variables, and the tree was proposed as a guide to determine the catalytic variables for a desired activity [67].
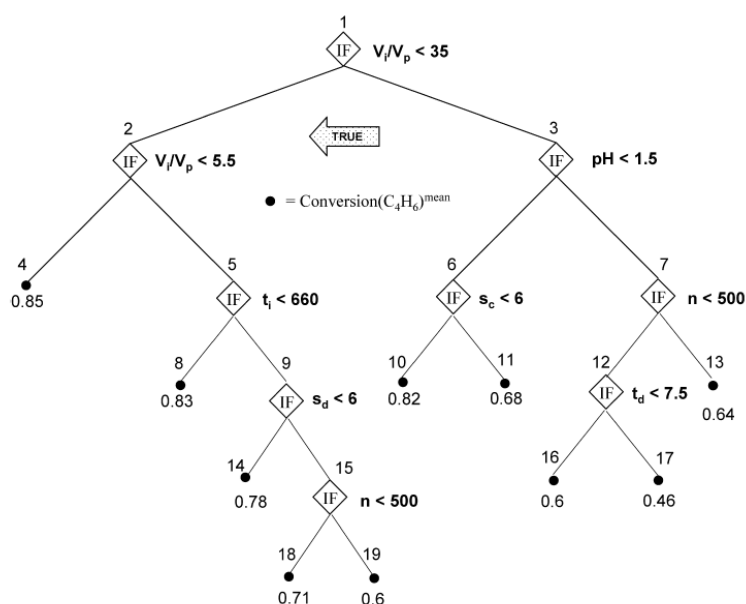
Figure 2.13. Regression tree for predicting butadiene conversion [67].
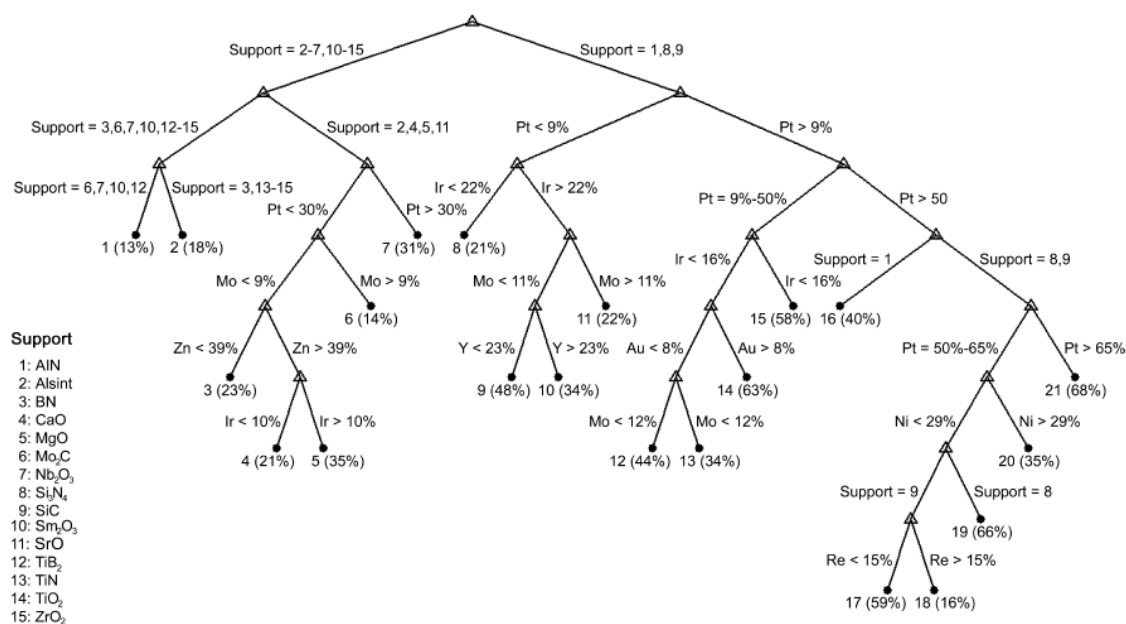
Figure 2.14. Regression tree modeling HCN yield [68].

Another regression tree study is on the catalytic conversion of methane and ammonia to hydrocyanic acid [68]. HCN yield was modeled by regression trees from the catalytic variables (the support and the relative amounts of the individual metal additives). Support was a categorical variable consisting of fifteen different options while the metal additives were numerical variables. The choice of the most appropriate tree size was based

on the generalization error of the tree, estimated by means of k-fold cross validation. According to the results of the best decision tree shown in Figure 2.14, it was reported that using support materials Si_3N_4, SiC and AlN with a Pt loading higher than 50% gives promising catalytic activity together with the catalysts composed of the metal additives, Ni (<30%), Re (<15%), Ir (<16%) and Au (>8%) [68].

2.5.3. Artificial Neural Network Modeling and Optimization of the Model

Artificial neural networks, which have superior ability on nonlinear regression and data approximation, can improve the catalytic experimental studies by helping the experimenter to get useful trends, patterns and valuable information that cannot be easily interpreted with simple observations. Indeed numerous researches on this subject have been published in the literature. An extensive list of these works is given in Table 2.2 and summarized below. The published works derived from this dissertation (as discussed in the later sections) was also added to the end of the table for completeness.

Hattori and Kito are the first researchers who introduced the use of artificial neural networks in the field of catalysis. Two of their early works, which pioneered further such studies, were the oxidative dehydrogenation of ethylbenzene [119] and catalytic activity of lanthanide oxides in the oxidation of butane published in 1994 and 1995, respectively [41]. These studies indicated that neural networks can successfully be used for catalyst development in search for new catalysts or additives, which reduce the time spent on conventional trial and error experiments.

In 1997, Hou *et al.* studied propone ammoxidation catalysts supported over Al_2O_3/SiO_2 to produce acrylonitrile [120]. The conversion and selectivity of propane were modeled by using artificial neural networks from the catalyst components: P, K, Cr, Mo, V contents and the weight ratio of Al_2O_3/SiO_2. The neural network model was used to find the optimum catalyst ingredients giving the highest acrylonitrile yield [120].

Cundari *et al.*, in 2001, investigated propane ammoxidation catalysts and modeled the catalytic data by using neural networks [121]. Several network topologies were tested to establish the relationship between P, K, Cr, Mo, V contents and the weight ratio of

Al_2O_3/SiO_2 with the conversion and selectivity of propane. Then, the best 14 neural networks with the highest performance were linearly combined to model the acrylonitrile yield. It was stated that combining the trained networks like this way can help to integrate the knowledge acquired by single network, and produces higher model accuracy. The optimum catalyst component leading to the maximum acrylonitrile yield was searched by applying genetic algorithm to the combined neural network providing an acrylonitrile yield of 79%, which was reported to be higher than those of the similar studies [121].

In 2001, Huang *et al.* experimentally studied the catalytic oxidative coupling of methane to produce C_2 hydrocarbons, and for this purpose, prepared twenty five catalysts designed by the orthographical method [122]. Neural network modeling was applied to model the catalytic performance (C_2 selectivity and CH_4 conversion), from the catalyst components: Na, Si, W, P, Mn and Zr (Figure 2.15). SWIFT method was applied to the trained neural network to find the optimum catalyst ingredients. CH_4 conversion and C_2 selectivity were found as 27.54% and 75.40%, respectively; both results being higher than any of the experimental results that had been reported [122].
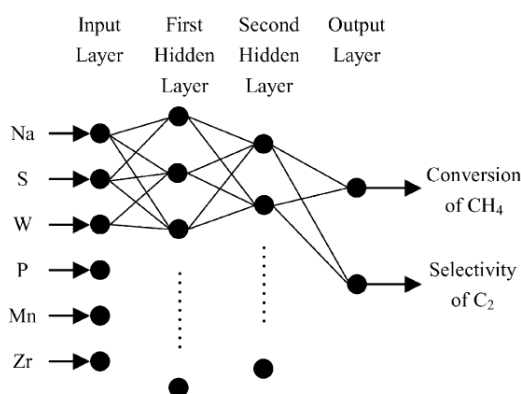
Figure 2.15. Neural network modeling C_2 selectivity and CH_4 conversion [122].

In 2001, Liu *et al.* applied artificial neural networks for the design of CO_2 hydrogenation catalyst for a database composed of 73 samples achieved from literature [123]. Five output variables, namely; CO_2 conversion, CO selectivity, CH_4 selectivity, CH_3OH selectivity and light alkenes (C_nH_m) selectivity were modeled from the catalyst composition (base metal, promoter A, promoter B) and the operating conditions (support, temperature, pressure, H_2/CO_2 ratio and the gas hourly space velocity). The prediction

accuracy of the neural network modeling for the CO_2 hydrogenation activity was analyzed and it was concluded that neural network modeling is a very powerful tool for the design of CO_2 hydrogenation catalysts [123].

Holena and Baerns, in 2003, studied the oxidative dehydrogenation of propane to propene [43]. They prepared 226 different catalyst compositions by high throughput experimentation and modeled the catalytic performance of the catalysts (selectivity and yield) by using neural networks (Figure 2.16). First, some concepts were designated such as the phenomenon of over-fitting due to the overtraining of the neural network and the use of early stopping to improve the neural network generalization accuracy. Moreover, the use of neural networks to find the optimum catalyst composition, and extracting logical rules between input and output variables was described in detail. Then, neural network modeling was applied for the oxidative dehydrogenation of propane to propene, with the oxides of B, Fe, Ga, La, Mg, Mn, Mo and V in the catalyst as the input variables, whereas the propane yield as the output variable. The composition of the catalyst was then optimized to find the maximum yield by the method of sequential quadratic programming [43].
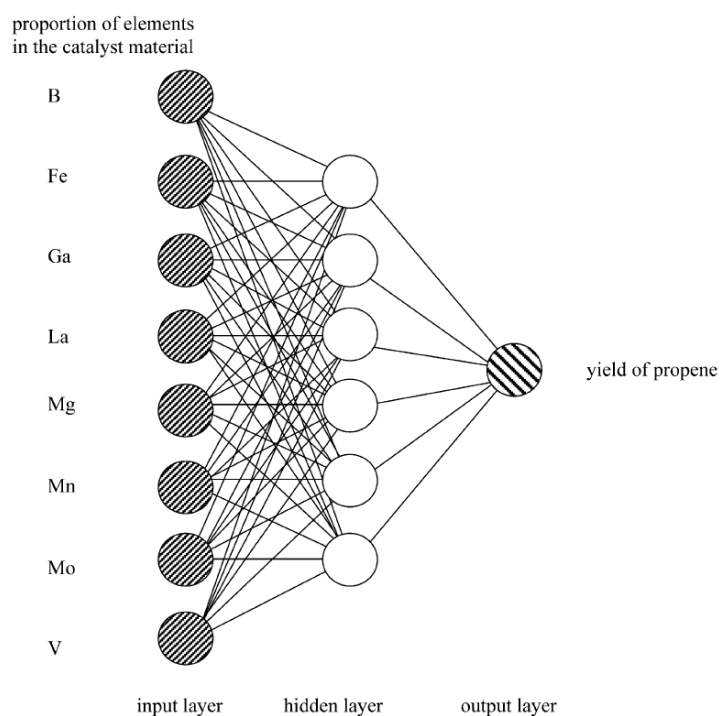
Figure 2.16. Neural network topology modeling propane yield [43].

In 2003, Serra *et al.* studied the n-octane isomerization catalysts prepared by high throughput experimentation and applied neural network modeling for the prediction of *n*-paraffin conversion, mono-branched and di-branched yields from the input variables of n-octane partial pressure, hydrogen partial pressure, reactor temperature and contact time [124]. The network size, activation functions, learning rates of the network were changed and their effects on the estimation performance were analyzed. It was found that when the neural network trained with n-octane samples were retrained with n-hexane samples, only a small number of samples were sufficient to obtain good quality predictions. Hence, it was concluded that having a library of pre-trained neural networks for several reaction schemes can be used as black box models to provide the prediction of new experimental data [124].

Huang *et al.*, in 2003, modeled the catalytic data related to the oxidative dehydrogenation of propane to propene by artificial neural networks and developed a hybrid genetic algorithm optimization technique to optimize the C_2 selectivity, CH_4 conversion and catalyst yield [125] that complements their previous work published in 2001 [122]. Genetic algorithm optimization was used as a global optimization method that provided the determination of the optimum catalyst composition leading to maximum catalyst performance, (C_2 selectivity higher than 70%, CH_4 conversion exceeding 35% and catalyst yield reaching 27.78%) [125]. This catalytic performance was stated to be higher than the previously reported ones [122].

Rodemerck *et al.*, in 2004, used neural networks for the modeling of the catalytic data of oxidative dehydrogenation of propane to propene that were prepared by high throughput experimentation [44]. They described the use of genetic algorithms to optimize the catalytic activity. The best catalyst composition was found from the candidate catalysts, which evolve through several generations by applying crossover (catalyst components are exchanged between two materials of the former generation), qualitative mutation (one or more chemical elements or compounds are replaced by others), quantitative mutation (the concentration of one or more compounds in the inorganic material is changed). A neural network trained by 328 data points were used to generate virtual experimental data to predict propene yield for given catalytic compositions. It was reported that genetic algorithm applied on the trained neural network resulted a maximum propene yield of 9.1% with the catalyst composition of $Ga_{0.21}Mg_{0.40}Mn_{0.21}V_{0.18}O_x$ [44].

Kito *et al.*, in 2004, investigated the methanol conversion over modified mordenite catalysts and modeled the catalyst deactivation rate by the help of artificial neural networks [126]. Si/Al_2 ratio, the degree of Ba ion exchange, the degree of La ion exchange and the amount of strong acid sites of the catalyst were used as the input variables for the neural network to estimate the deactivation rate constant (Figure 2.17). Another model was constructed to extrapolate the hydrocarbon yield (at 4h and 5h) by using the same input variables together with the initial decay data (yield at 5min, 1h, 2h, 3h). This provided to extrapolate the degree of catalyst deactivation. The accuracy of the modeling results were found to be high, even in the case where the experimental data seemed to contain some experimental error. It was then concluded that neural network modeling is a powerful tool for the estimation of catalyst life through the deactivation rate constant and for the extrapolation of the catalyst decay curve [126].
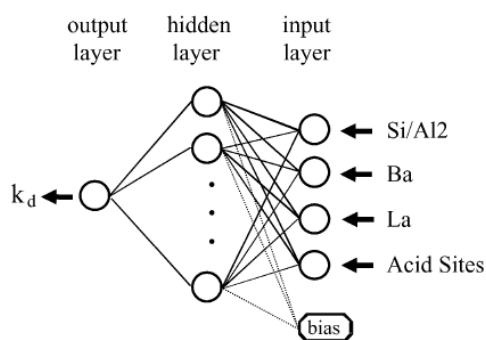
Figure 2.17. Neural network topology modeling catalyst deactivation rate [126].

Nandi *et al.* presented a detailed instruction for support vector regression modeling and described the adaptation of this approach to the catalytic data in 2004 [127]. The experimental data (42 points) related to the benzene isopropylation using Hbeta catalyst was modeled by using support vector regression and artificial neural networks with the input variables of temperature, pressure, benzene to isopropyl alcohol mole ratio and weight hourly space velocity. The models were used further for the genetic algorithm based optimization to find the maximum cumene yield and selectivity, and the performances of both strategies were compared. Although the prediction accuracy of support vector regression was found to be somewhat lower than that of the neural networks, it was still in the acceptable region. It was reported that neural networks-genetic algorithm strategy provided an optimum cumene yield of 24.88% and selectivity of

99.04% while support vector regression-genetic algorithm strategy provided a yield of 24.80% and selectivity of 95.76%. The neural network-genetic algorithm optimum point matched almost perfectly to the verification experiment and as a result, the cumene yield and selectivity were experimentally improved by 2.59 and 5.10%, respectively [127].

Watanabe *et al.*, in 2004, applied high throughput experimentation to study Cu-Zn-Al-Sc-B-Zr oxide catalyst for low-pressure methanol synthesis in a high throughput reactor system, which tested 96 samples simultaneously [128]. A total of 190 random catalysts were analyzed in the reactor in two runs and the data achieved were trained by artificial neural networks with both preparation and operating conditions as the input variables and the catalytic activity as the output variable (space time yield). Then, the well trained neural network was optimized by using genetic algorithms. However, the neural network predicted optimum result was quite different than the result of the verification experiment, which was attributed to the insufficient experimental information around the optimum point. Hence, additional 44 catalysts were collected experimentally and the neural network-genetic algorithm strategy was applied to the new database composed of 234 points, which provided an optimum point with a quite high accuracy. It was then reported that catalyst with Cu/Zn/Al/Sc/B/Zr ratio of 43/17/23/11/0/6 prepared using 2.2 times of oxalic acid and calcined at 605 K was the optimum catalyst giving a space time yield of 427 g MeOH/kg-cat./h, which was almost two times that of an industrial catalyst [128].

Corma *et al.*, in 2005, developed a neural network-genetic algorithm hybrid method to model the olefin epoxidation Ti catalysts in order to maximize epoxide yield [46]. The proposed procedure started with the generation of a random experimental space by using high throughput experimentation, and the modeling of this database by using neural networks. Then, genetic algorithm sought for the catalysts giving higher yield by simulating the input variables through the neural network model, which created the next generation. After the experimental evaluation of the new generation, the neural network was retrained with the new available data. It was observed that a progressive improvement of the epoxidation catalyst was achieved through each generation giving a higher epoxide yield, and at the third generation the activity of the catalyst was significantly higher than the reference catalyst. It was then concluded that the proposed hybrid system based on high throughput experimentation, neural network modeling and genetic algorithm optimization

can be adopted for the intelligent discovery of new catalytic materials also for different catalytic systems [46].

Tompos *et al.*, in 2003, described a new powerful optimization method, holographic research strategy, for the design and testing of the catalyst libraries [129]. It was proposed that holographic research can provide the visualization of a continuous multidimensional experimental space by two-dimensional presentation (Figure 2.18); hence, it can be used in the study of virtual preparation and testing of relatively large number of multi-component catalysts. It was also stated that when compared to stochastic genetic algorithms, this method has the advantage of being deterministic. The gray scale colors in Figure 2.18 represent the output results of the variables A, B, C, D, E, F that change periodically. The closer the data to the optimum point the darker the color of the given point is. The method continues with selecting the darkest region and applying a more detailed holographic search into that area. This procedure is repeated until the global optimum is found [129].
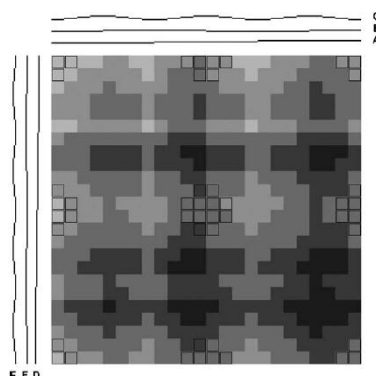
Figure 2.18. Six parameter holographic search [129].

Tompos *et al.*, in 2003, applied artificial neural network modeling combined with holographic research strategy to design a catalyst library for low temperature propane oxidation, the experimental data of which was collected by high throughput experimentation [130]. Propane conversion was modeled from the catalyst components, Pt, Pd, Rh, Ru, Au, Pd, Ag and Mn wt.% by artificial neural networks and the maximum propane conversion was searched by the holographic method. The optimum catalyst components provided by the neural network-holographic research strategy was found to be 2.4 wt.% Ru and 0.4 wt.% Mn, resulting a maximum propane conversion of 89% at 150°C, which was stated as the highest conversion ever reported [130]. It was then concluded that

application of artificial neural networks together with the holographic research strategy can be considered as a powerful tool for data mining in high throughput experimentations.

The total oxidation of methane was proposed as a case study by Tompos *et al.*, in 2005, for the application of artificial neural network modeling and holographic research strategy [45]. The concentration of catalyst components (Ce, Co, Zr, Cr, La, Cu, Pt, Pd and Au) were used as the catalyst variables while methane conversion was used as a degree of the catalytic activity. Holographic mapping of the neural network outputs translated black box nature of neural networks to a more comprehendible way and provided the detailed analysis of the catalytic components, such as which component affects the conversion positively or negatively. It was reported by the help of this hybrid method that almost complete methane conversion was observed at 350°C [45].

Tompos *et al.*, in 2006, compared the advantages and disadvantages of genetic algorithm and holographic research strategy for the catalyst library optimization [131]. For this purpose three different catalytic data (methane oxidation, propane oxidation and methane oxidative coupling) were modeled by neural networks and optimized by these two methods. It was reported that holographic research reduces the number of necessary catalytic tests while reaching the optimum catalyst composition. In contrast, genetic algorithms test a numerous compositions of catalysts, even the catalytically poor materials, and this causes a waste of time. Hence, holographic research technique found the optimum points faster than the genetic algorithm in all three cases. Moreover, it was stated that qualitative relationships between compositions and activities can be recognized more easily by visual holographic mapping. However, due to the fact that the experimental points are arranged randomly in genetic algorithms, no such conclusions can be drawn with respect to the activity-composition relationship [131].

Tompos *et al.*, in 2010, modeled water gas shift reaction catalysts by using artificial neural networks, and applied genetic algorithms together with the holographic research technique to find the optimum catalyst composition giving the maximum CO_2 yield [132]. Several neural networks were trained and the best 100 networks with the highest performance were linearly combined as described by Cundari *et al.* [121]. The catalytic addives Co, Eu, Fe, Ge, Mo, Ti, Sb, Pt, Ru, Sn, V and Zr at three levels (531,441 data

points) were virtually observed by the holographic technique. It was then found that, addition of Pt, Eu and Fe was quite crucial for the water gas shift reaction catalysts tested in this study and V had also promoting effect, whereas Mo, Ru, Sb, Co and Ge had negative effect on the CO_2 yield. It was then concluded that analysis by holographic maps provides the researcher to focus on only the promising experimental regions, instead of the areas resulting poor performance [132].

Hattori and Kito, in 2006, studied the oxidation of propene over oxide catalysts, oxidation of butane over lanthanide oxides, the decomposition of formic acid on metals and oxidation of methane on lanthanide oxides in order to estimate the importance of each input factors controlling the catalytic activity by using the weights of the trained neural network [133]. They named this new approach the "connecting weight method". Three equations (Equation 2.42-44) were proposed to calculate the contribution of ith input variable to the kth output variable (C_{ik}), where w_{ij} represents the weights between ith input variable and jth hidden unit while w_{jk} are for the weights between jth hidden unit and kth output variable. This approach was reported to give reliable results in finding the relative significances of the factors controlling the activity for all the four catalytic cases [133].

$$C_{ik} = \left| \sum_j w_{ij} \times w_{jk} \right| \tag{2.42}$$

$$C_{ik} = \sum_j \left(\left| w_{ij} \right| + \left| w_{jk} \right| \right) \tag{2.43}$$

$$C_{ik} = \sum_j \left(\left| w_{ij} \right| \right) \tag{2.44}$$

Hattori and Kito, in 2007, described the use of partial differentiation of the neural networks to find the relative importance of the input variables to identify which one of them affect the output most [47, 48]. The procedure is described in Equation 2.37-39 in Section 2.4.9.2. A preliminary study was presented to analyze the primary factors controlling the activity of catalytic oxidations of butane and methane on lanthanide oxides [47]. Then, a more detailed application of this method was presented for four cases; oxidation of propene over oxide catalysts, oxidation of butane over lanthanide oxides, the decomposition of formic acid on metals and oxidation of methane on lanthanide oxides

[48], and the results were compared with the ones that had been found by the connecting weight method applied in one of the previous works [133]. In addition to these, support and additive effect on the activity of Pt catalyst in the combustion of propane was also examined. The partial differentiation of the neural networks was reported to give superior results compared to the ones found by the connecting weight method [48]. It was reported that, this method provides additional knowledge such that if the partial derivative with respect to a variable is positive, the variable positively affects the catalytic activity or vice versa. Thus, the method was proposed to be used to improve catalytic studies by changing the additives according to the numerical values of the partial derivatives [48].

Cobalt catalysts supported on $SrCO_3$ for selective CO oxidation was investigated by Omata *et al.*, in 2005 [134]. The experimental database was prepared by full factorial design, CO and O_2 conversions were modeled by radial basis function neural networks from the catalyst preparation variables (Co content and pretreatment temperature), and the optimum variables to maximize the performance of the catalysts were found by applying grid search [134]. Omata *et al.* investigated the same catalytic system by the same data mining approach in a more detailed way in 2006 [135]. First, catalytic CO conversion in the H_2O and CO_2 free stream were modeled and optimized. Then, the effects of H_2O and CO_2 were determined. In order to improve the catalytic activity in the presence of H_2O and CO_2, several additives to $Co/SrCO_3$ were investigated. Several physicochemical properties; such as atomic number, atomic weight, melting point, boiling point, heat of vaporization etc.; of ten representative elements (B, K, Sc, Mn, Zn, Nb, Ag, Nd, Re, and Tl) were selected as input variables for the radial basis neural networks. This network, which correlates physicochemical properties with catalytic activities, was then used to predict the performance of new additives to the $Co/SrCO_3$. The virtual screening was successfully performed by the neural networks, and for this catalytic system, the elements such as Bi, Ga, and In were predicted to be promising additives [135].

As a supplementary study to their previous studies, Omata *et al.*, in 2007, investigated the performance of new metal additives to the $Co/SrCO_3$ catalysts for preferential CO oxidation by using artificial neural networks [136]. Similar to the previous study, the physicochemical properties of ten different elements (B, K, Sc, Mn, Zn, Nb, Ag, Nd, Re, and Tl) were selected as input variables for the radial basis neural networks to

model the CO and O_2 conversions. Then, in order to predict the promoting effect of a new metal additive (unseen by the neural network before), the physicochemical properties of that element was introduced to the trained network, and the network was forced to find the corresponding outputs. It was reported that the performance of the elements such as Ga, In, Bi and Yb were predicted to be promising, but only the promoting performance of Yb and Bi were experimentally verified. Finally, it was concluded that neural network modeling can help the experimenter to discover new additives by virtually screening their effects [136].
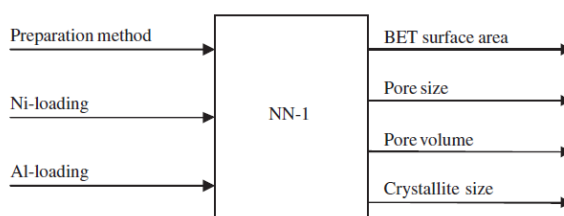
Figure 2.19. Neural network approach to model some catalytic characteristics [49].

Song *et al.*, in 2007, used artificial neural network approach to design an optimum Ni/Al_2O_3 catalyst for the production of hydrogen by the catalytic reforming of crude ethanol [49]. For this purpose, three neural network models were constructed: The input variables for the first network was the preparation method (co-precipitation, precipitation and impregnation), Ni loading and Al loading while the BET surface area, pore volume, pore size and crystallite size were the outputs as shown in Figure 2.19. The second network was constructed to model the ethanol conversion, H_2 selectivity and H_2 yield from the same input variables. Finally, the third one was used to predict the preparation method, Ni and Al loadings by using the ethanol conversion, H_2 selectivity and H_2 yield as the input variables. Then, the second network was optimized to find the optimum preparation method, Ni and Al loadings giving the maximum H_2 yield. The optimization was performed by starting from the minimum values of the variables and adding small delta units to them until the entire experimental range was virtually searched. The optimal catalyst was found as the one prepared by the co-precipitation method with the nickel loading of 12.4 wt.% and an aluminum composition of 42.5 wt.% giving hydrogen yield of 4.4 mol%, crude-ethanol conversion 79.6 mol% and H_2 selectivity of 91.4 mol% [49].

Umegaki *et al.*, in 2008, studied the Cu based catalysts for the oxidative steam reforming of methanol by applying L_9 orthogonal array as the experimental design, neural networks to model the data and grid search to find the optimum catalyst [137]. Radial basis function neural networks were used to model the methanol conversion, CO concentration in the reformed gas and H_2 selectivity as a function of the physicochemical properties of several representative additive elements (Mn, Ni, Fe, Co, Ga, Lu, Li, B, Ca, Zr, Mg, Pr, Zn, Ce). Then, grid search was applied to the trained neural network to virtually screen the performance of several metal additives to the Cu based catalysts, and Cu-Ca catalyst was obtained as the highest performance catalyst among the other binary catalysts. In the next step, the effects of second metal additive to the Cu-Ca, Cu-Ce, and Cu-Pr catalysts were observed through the neural network models. A second metal additive to the Cu-Ca catalyst was reported to decrease the catalytic performance; in contrast Mn addition to the Cu-Ce catalyst improved the performance. However, Cu-Pr-Ti catalyst was found as the catalyst with the highest performance. In the next step, catalyst composition and preparation conditions (amounts of Cu, Pr, Ti, calcination temperature, total metal salt concentration) were optimized to improve the performance of the Cu-Pr-Ti catalyst further. It was concluded that simultaneous application of an experimental design method, neural networks and grid search can provide the discovery of the optimum performance catalysts in a very short time with a very few number of experimental runs [137].

Günay and Yıldırım, in 2008, investigated the Pt-Co-Ce/Al_2O_3 catalysts for selective CO oxidation [34]. The experimental data was created by the response surface methodology, and CO conversion was modeled by using artificial neural networks as a function of the input variables: amount of base metal (Pt wt.%), the promoters (Co and Ce wt.%), calcination temperature and calcination time. The "Change of RMSE" method was applied to find the relative significances of the input variables, and Pt amount was reported to be the most significant one, followed by Ce and Co amounts. The effects of all the five input variables were virtually observed at various levels of the other variables, and the experimental values were in agreement with the neural network predictions. It was finally concluded that neural network modeling can be very helpful to improve the experimental works in catalyst design and it may be combined with the statistical experimental design techniques, which provide high modeling accuracy even with small number of experimental data points [34].

Table 2.2. Publications on neural network modeling in the field of catalysis.

Reference	Year	Catalyst	Input variables	Output variables	Novelty
Hattori and Kito [41]	1995	oxidation of butane	various physicochemical properties	CO_2 yield	one of the first neural network applications in the field of catalysis
Hou *et al.* [120]	1997	propone ammoxidation catalysts supported over Al_2O_3/SiO_2	P, K, Cr, Mo, V contents, weight ratio of Al_2O_3/SiO_2	conversion and selectivity of propane	optimum catalyst ingredients giving the highest acrylonitrile yield were found by the neural network
Cundari *et al.* [121]	2001	propone ammoxidation catalysts supported over Al_2O_3/SiO_2	P, K, Cr, Mo, V contents, weight ratio of Al_2O_3/SiO_2	conversion and selectivity of propane	combining the trained networks, optimum catalyst was found by applying genetic algorithm to the trained neural network
Huang *et al.* [122]	2001	oxidative coupling of methane to produce C_2 hydrocarbons	Na, Si, W, P, Mn, Zr contents	C_2 selectivity, CH_4 conversion	SWIFT method was applied to the trained neural network to find the optimum catalyst ingredients
Liu *et al.* [123]	2001	CO_2 hydrogenation	base metal, promoter A, promoter B, support, temperature, pressure, H_2/CO_2 ratio and the gas hourly space velocity	CO_2 conversion, CO selectivity, CH_4 selectivity, CH_3OH selectivity, light alkenes selectivity	data from the literature was modeled
Holena and Baerns [43]	2003	oxidative dehydrogenation of propane to propene	B, Fe, Ga, La, Mg, Mn, Mo, V contents	propene yield	sequential quadratic programming was applied to the trained neural network to find the optimum catalyst ingredients
Serra *et al.* [124]	2003	n-octane isomerization catalysts	n-octane partial pressure, hydrogen partial pressure, reactor temperature and contact time	*n*-paraffin conversion, mono-branched and di-branched yields	pre-trained neural networks were used as black box models to provide the prediction of new experimental data
Huang *et al.* [125]	2003	oxidative dehydrogenation of propane to propene	Na, Si, W, P, Mn, Zr contents	C_2 selectivity, CH_4 conversion, catalyst yield	optimum catalyst was found by applying genetic algorithm to the trained neural network
Rodermeck *et al.* [44]	2004	oxidative dehydrogenation of propane to propene	Ga, Mg, Mo, V contents	propene yield	optimum catalyst was found by applying genetic algorithm to the trained neural network
Kito *et al.* [126]	2004	methanol conversion into hydrocarbons	Si/Al_2 ratio, the degree of Ba ion exchange, the degree of La ion exchange, the amount of strong acid sites of the catalyst	deactivation rate constant	neural network model was used to extrapolate the catalyst decay curve
Nandi *et al.* [127]	2004	benzene isopropylation	temperature, pressure, benzene to isopropyl alcohol mole ratio, weight hourly space velocity	cumene yield and selectivity	the performances of the neural networks were compared with the support vector regression, optimum catalyst was found by genetic algorithm

Table 2.2. Publications on neural network modeling in the field of catalysis (cont.).

Reference	Year	Catalyst	Input variables	Output variables	Novelty
Watanabe *et al.* [128]	2004	low-pressure methanol synthesis from syngas	Cu, Zn, Al, Sc, B, Zr contents, calcination temperature, oxalic acid amount	space time yield	optimum catalyst was found by applying genetic algorithm to the trained neural network
Corma *et al.* [46]	2005	Ti-silicate-based catalysts for olefin epoxidation	OH^-, titanium, surfactant concentration	epoxide yield	a hybrid system based on high throughput experimentation, neural network modeling and genetic algorithm optimization was proposed
Tompos *et al.* [130]	2003	low temperature propane oxidation	Pt, Pd, Rh, Ru, Au, Pd, Ag, Mn contents	propane conversion	optimum catalyst was found by holographic research strategy
Tompos *et al.* [45]	2005	total oxidation of methane	Ce, Co, Zr, Cr, La, Cu, Pt, Pd, Au contents	methane conversion	optimum catalyst was found by holographic research strategy
Tompos *et al.* [131]	2006	methane oxidation, propane oxidation, methane oxidative coupling	Same as refs: [45, 122]	Same as refs: [45, 122]	The performance of genetic algorithms and holographic research strategy to find the optimum catalyst were compared
Tompos *et al.* [132]	2010	water gas shift reaction	Co, Eu, Fe, Ge, Mo, Ti, Sb, Pt, Ru, Sn, V, Zr contents	CO_2 yield	applied genetic algorithms together with the holographic research technique to find the optimum catalyst composition
Hattori and Kito [133]	2006	oxidation of propene, oxidation of butane, decomposition of formic acid, oxidation of methane	various physicochemical properties	various degrees for catalytic activity such as reaction rate, CO_2 yield, CH_4 conversion	introduced the "connecting weight method" for the calculation of input significance
Kito and Hattori [47]	2007	oxidations of butane and methane	various physicochemical properties	for oxidation of butane: CO_2 yield; *for* oxidation of CH_4 conversion	introduced the "partial differentiation of the neural network" for the calculation of input significance
Hattori and Kito [48]	2007	oxidation of propene, oxidation of butane, decomposition of formic acid, oxidation of methane, combustion of propane	various physicochemical properties	various degrees for catalytic activity such as reaction rate, CO_2 yield, CH_4 conversion	compared the "connecting weight method" with the "partial differentiation of the neural network" for the calculation of input significance
Omata *et al.* [134]	2005	preferential CO oxidation	Co content and pretreatment temperature	CO and O_2 conversions	optimum catalyst was found by applying grid search to the trained neural network
Omata *et al.* [135]	2006	preferential CO oxidation	various physicochemical properties of B, K, Sc, Mn, Zn, Nb, Ag, Nd, Re Tl	CO and O_2 conversions	virtually screening of the effects of new metal additives
Omata *et al.* [136]	2007	preferential CO oxidation	various physicochemical properties of B, K, Sc, Mn, Zn, Nb, Ag, Nd, Re Tl	CO and O_2 conversions	virtually screening of the effects of new metal additives

Table 2.2. Publications on neural network modeling in the field of catalysis (cont.).

Reference	Year	Catalyst	Input variables	Output variables	Novelty
Song *et al.* [49]		hydrogen production by catalytic reforming of crude ethanol	preparation method, Ni and Al loadings	BET surface area, pore volume, pore size and crystallite size or ethanol conversion, H_2 selectivity, H_2 yield	used preparation method as an input variable, optimum catalyst was found by applying grid search to the trained neural network
Umegaki *et al.* [137]	2008	oxidative steam reforming of methanol	various physicochemical properties and preparation conditions	methanol conversion, CO concentration in the reformed gas and H_2 selectivity	simultaneous application of an experimental design method, neural networks and grid search to find the optimum catalyst
Günay and Yıldırım [34]	2008	preferential CO oxidation	Pt, Co, Ce amounts; calcination temperature, calcination time, time on stream	CO conversion	data created by an experimental design method, input significance was found by the "Change of RMSE" method.
Günay and Yıldırım [51]	2010	preferential CO oxidation	Pt, Co, Ce, Mg, Fe, Mn, Zr, K, Ni amounts; reaction temperature; CO_2 and H_2O contents; time on stream	CO conversion	modular approach to treat the preparation and the operating variables differently, "Change of RMSE" and "partial differentiation of the neural network" was used for the input significance
Günay *et al.* [50]	2011	water gas shift reaction	Ce, K, Ni amounts; reaction temperature; CO_2 and H_2 contents; H_2O/CO ratio; time on stream	CO conversion	modular approach to treat the preparation and the operating variables differently, "Change of RMSE" and "partial differentiation of the neural network" was used for the input significance
Günay and Yıldırım [53]	2011	preferential CO oxidation	various preparation and operating variables	CO conversion	By using the data of a set of published papers, neural networks were applied to predict the experimental correlations and trends reported in the other publications

3. EXPERIMENTAL AND COMPUTATIONAL DATA

First, the details of the experimental data generated in our laboratories are presented in three consecutive subsections: "selective CO oxidation over promoted Pt/Al_2O_3 catalysts", "selective CO oxidation over promoted Au/Al_2O_3 catalysts" and "water gas-shift activity of promoted Pt/Al_2O_3 catalysts". Then, the computational data related to CO and O_2 adsorption over gold nanoparticles using DFT is described briefly. Finally, the details of the experimental data extracted from published papers in the literature are presented in the following two subsections: "selective CO oxidation over Cu-based catalysts from published data" and "selective CO oxidation over noble metal-gold based catalysts".

3.1. Selective CO Oxidation over Promoted Pt/Al_2O_3 Catalysts

Preferential CO oxidation over $Pt-M^1O_x-M^2O_x/Al_2O_3$ catalysts with various promoters ($M^1=$ Co, K, Ni; $M^2 =$ Ce, Mg, Fe, Mn and Zr) was experimentally studied in hydrogen rich streams by Uguz and Yıldırım [138]. The catalysts were prepared using incipient to wetness impregnation technique and the catalytic activities were measured in a microreactor flow system with the feed composition of 1% CO, 1%, O_2, 60% H_2 and He as inert balance. F/W (inlet gas flowrate/catalyst weight) was also constant at 24000 $cm^3/(g\ h)$. The type and range of the input variables are presented in Table 3.1 while the details of the experimental work were presented and discussed elsewhere [34, 138].

Table 3.1. Range of variables for Pt/Al_2O_3 catalyst for selective CO oxidation.

Input Variables	Range	Input Group
Pt (wt.%)	0.7 – 1.4	Group 1
promoter type	Co, Ce, Mg, Fe, Mn, Zr, K, Ni	Group 2
reaction temperature (^{o}C)	80-130	Group 3
CO_2 content (%)	0-25	
water content (%)	0-10	
TOS (min)	0-120	Group 4

3.2. Selective CO Oxidation over Promoted Au/Al_2O_3 Catalysts

Preferential CO oxidation over $Au/MO_x/Al_2O_3$ was experimentally studied for various operating conditions with six different promoters (M=Mg, Mn, Co, Ce, Fe, Ni) by Davran-Candan *et al.* [139]. The promoters were loaded to the metal oxide supports by incipient to wetness impregnation method while Au was loaded over MO_x/Al_2O_3 by the homogeneous deposition precipitation method to attain a catalyst composition of 1 wt.% Au and 1.25 wt.% MO_x. The performances of the catalysts were tested in a microflow reaction system at 1% CO, 1% O_2, 60% H_2, He as balance with a constant F/W (inlet gas flowrate/catalyst weight) at 24000 $cm^3/(g.h)$. The type and range of the preparation and the operating variables are shown in Table 3.2 while the details of the experimental work were presented in another communication [139].

Table 3.2. Range of variables for Au/Al_2O_3 catalyst for selective CO oxidation.

Input variables	Range	Input group
promoter type (MO_x)	M = Mg, Mn, Co, Ce, Fe, Ni	preparation variables
reaction temperature (^{o}C)	50-150	operating variables
CO_2 content (%)	0-25	
H_2O content (%)	0-10	

3.3. Water Gas-Shift Activity of Promoted Pt/Al_2O_3 Catalysts

WGS activity of $Pt-Ce/Al_2O_3$ (1 wt.% Pt, 1.25 wt.% Ce) catalysts in the absence and presence of a 1.25 wt.% second promoter (K, Ni, Co) was experimentally studied by Akpınar [140]. The catalysts were prepared using the incipient to wetness co-impregnation technique. The activities of the catalysts were measured using a micro-reactor flow system. The feed flow rate was kept constant at 100 cm^3/min. The 0.25 g catalyst samples used in the experiments were reduced *in situ* at 400°C in a hydrogen environment for 3 hours before each activity measurement. All the experiments were performed at atmospheric pressure. The range of the preparation and the operating variables are shown in Table 3.3 while the additional details can be found elsewhere [50].

Table 3.3. Range of variables for Pt/Al_2O_3 catalyst for WGS reaction.

Input Variables	Range	Input Group
promoter type	Ce, K, Ni	preparation variables
reaction temperature (^{o}C)	250-300	operating variables
CO_2 content (%)	0-10	
H_2 content (%)	0-40	
H_2O / CO ratio	2-3	
TOS (min)	30-180	

3.4. CO and O_2 Adsorption over Gold Nanoparticles Using DFT

CO and O_2 adsorption on gold nanoparticles (Figure 3.1) were studied by Davran-Candan using the Density Functional Theory [141]. As described in her dissertation, the optimum geometries and the structural properties of various gold nanoclusters were determined first. Then, the adsorption energies of CO and O_2 were investigated by bonding this molecule to all possible sites of the cluster, and optimizing the geometry of the cluster-CO complex. As a result, a data set containing the structural properties (such as binding energy, HOMO-LUMO gap and ionization potential) and CO and O_2 adsorption energies over 97 possible adsorption sites was formed [52]. In the current work, this data set was used to extract knowledge about the effects of user defined variables (such as cluster size, coordination number and charge) and structural properties on CO and O_2 adsorption energies using artifical neural networks and logistic regression.
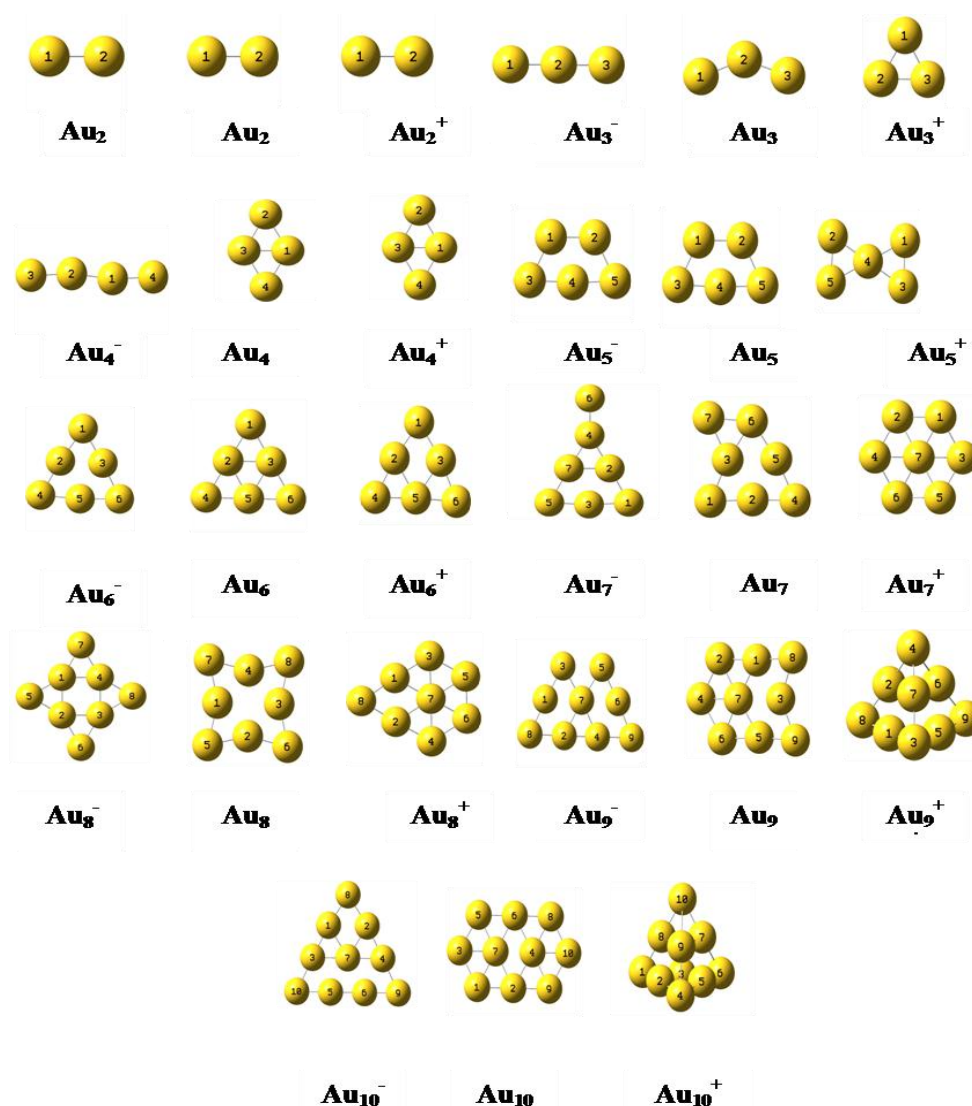
Figure 3.1. Equilibrium geometries of clusters studied [141].

3.5. Selective CO Oxidation over Cu-Based Catalysts from Published Data in the Literature

Twenty research publications on selective CO oxidation over copper-based catalysts were reviewed, and the experimental data presented in these works were extracted [53]. The reactor details and catalyst preparation and operating variables studied in these publications are presented in Table 3.4, together with the number of data points extracted from each work. Each research publication reported the effects of at least one input variable on the steady state CO conversion, and the reaction temperature was the common input variable for all of the works (Table 3.5). All of the data were merged into a large database composed of 1337 data points.

Table 3.4. Publications used for database construction, reactor details, variables analyzed and number of data extracted for Cu based catalysts.

Reference Publication	Reactor type	Reactor Inside Diameter (mm)	Particle Diameter (µm)	Variables Analyzed Using the Data in the Corresponding Publication	Number of Data Points Extracted
Zou *et al.* [28]	fixed-bed quartz reactor	-	250-400	H_2O%, react. T	17
Zhu *et al.* [29]	fixed-bed quartz reactor	4	177-250	Cu content, react. T	36
Liu *et al.* [30]	quartz tubular reactor	10	50-150	O_2%, H_2O%, CO_2%, react. T, F/W	80
Avgouropoulos *et al.* [142]	fixed-bed reactor	-	90-160	H_2O%, CO_2%, react. T, F/W	35
Avgouropoulos & Ioannides [85]	fixed-bed reactor	4	90-160	Cu content, H_2O%, CO_2%, react. T, F/W	146
Kim & Cha [88]	quartz tubular reactor	4	150-180	O_2%, CO%, H_2O%, CO_2%, react. T, F/W	63
Jung *et al.* [143]	fixed-bed reactor	-	-	calcination temperature, react. T	34
Park *et al.* [144]	glass tubular reactor	4	125-149	Cu content, second metal additive, O_2%, H_2O%, CO_2%, react. T	92
Marino *et al.* [145]	glass fixed-bed reactor	8	-	Cu content, support type, O_2%, react. T	98
Martinez-Arias *et al.* [86]	glass tubular catalytic reactor	-	-	react. T	28
Kosmambetova *et al.* [146]	quartz tubular reactor	10	-	support type, O_2%, react. T	37
Firsova *et al.* [147]	tubular reactor	3	-	second metal additive, react. T	28
Gomez-Cortés *et al.* [148]	fixed-bed reactor	-	-	Cu content, H_2O%, CO_2%, react. T	54
Ribeiro *et al.* [149]	quartz tubular reactor	-	-	support type, react. T	21
Polster *et al.* [150]	U-shaped quartz reactor	10	75-125	Cu content, catalyst preparation technique, O_2%, CO%, react. T, time on stream, F/W	30
Ayastuy *et al.* [151]	plug flow reactor	-	-	Cu content, H_2%, H_2O%, CO_2%, react. T	296
Wu *et al.* [152]	U-shaped quartz reactor	-	-	Cu content, catalyst preparation technique, H_2O%, CO_2%, react. T	95
Li *et al.* [153]	fixed-bed quartz reactor	6	177-250	second metal additive, react. T	40
Razeghi *et al.* [154]	U-shaped quartz reactor	4	125-297	H_2O%, CO_2%, react. T	62
Lendzion-Bielun *et al.* [155]	fixed-bed quartz reactor	-	-	Cu content, second metal additive, react. T	45

Table 3.5. Input variables and their ranges for Cu based catalysts.

Input Variable	Type	Range (for Continuous Variables) or Identity (for Categorical Variables)
Cu wt.%	continuous	0-20
preparation method	categorical	incipient to wetness impregnation (IWI), wet impregnation (WI), sequential impregnation (SI), co-precipitation (CP), sol-gel precipitation (SGP), urea nitrate combustion (UNC), urea gelation (UG), hydrothermal method (HM)
calcination temperature (^{o}C)	continuous	300-900
calcination time (h)	continuous	2-5
support type	categorical	CeO_2, Al_2O_3, ZrO_2, MgO, La_2O_3, TiO_2, Nb_2O_5, MnO
second metal type and wt.%	categorical/continuous	Ce(0-20), Co(0-5), Fe(0-2.5), Mn(0-1.7), Ni(0-2.5)*
operation temperature (^{o}C)	continuous	20-305
H_2 vol.%	continuous	0-98
O_2 vol.%	continuous	0.5-4.0
CO vol.%	continuous	0.8-2.0
H_2O vol.%	continuous	0-20
CO_2 vol.%	continuous	0-25
time on stream (min)	continuous	0-420
F/W (cm^3/(h.g))	continuous	12000-240000

* Numbers in parentheses represent the range of metal weight percentages.

3.6. Selective CO Oxidation Over Noble Metal-Gold Based Catalysts from Published Data in the Literature

A large database containing 5008 experimental data points was extracted from 71 publications on selective CO oxidation over noble metal and gold based catalysts. The variables analyzed in these publications are presented in Table 3.6, and their ranges are shown in Table 3.7. The database contained seven base metals (Pt, Au, Ru, Rh, Ir, Cu and Pd), nine preparation methods (incipient to wetness impregnation, wet impregnation, co-impregnation, sequential impregnation, sol-gel precipitation, co-precipitation, deposition precipitation, sequential precipitation and homogenous precipitation), fifteen supports (Al_2O_3, MgO, CeO_2, CuO, TiO_2, ZnO, MnO, FeO, Nb_2O_5, SiO_2, ZrO_2, La_2O_3, Co_3O_4, zeolite and activated carbon) and nineteen promoters (Ce, Co, Mg, Fe, Mn, Zr, K, Ni, Sn, Li, Na, Rb, Cs, Nb, La, Ba, Pb, Sm and Zn). Reaction temperature; H_2, O_2, CO, H_2O, CO_2 percentages in the feed stream; time on stream and F/W were the operating variables.

Table 3.6. Details for the extracted data for noble metal and Au based catalysts.

Reference Publication	Reactor Type	Variables Analyzed Using the Data in the Corresponding Publication	Unique Catalysts	Number of Data Points Extracted	Number of Data Points to be Tested
Bethke and Kung [156]	tubular fused silica reactor	Au content (base metal), reaction temperature		13	13
Ito et al. [157]	continuous flow reactor	Rh content (base metal); Mn, Nb contents (promoters); support type; calcination temperature and time; reaction temperature	Rh/Nb_2O_5*, Rh/SiO_2**	39	0
Grisel and Nieuwenhuys [158]	continuous flow reactor	Mg, Mn contents (promoters); CO, O_2, H_2 percentages; reaction temperature		132	132
Son and Lane [159]	continuous flow reactor	Ce content (promoter); O_2, H_2 percentages; reaction temperature		22	22
Manasilp and Gulari [160]	glass tubular reactor	Pt content (base metal); H_2O, CO_2, CO, O_2 percentages; reaction temperature; F/W		51	51
Bulushev et al. [161]	plug-flow reactor	Au content (base metal), support type, TOS	Au/activated C**	27	18
Avgouropoulos et al. [142]	fixed bed flow reactor	base metal type; support type; preparation method; H_2O, CO_2 percentages; reaction temperature; F/W		73	73
Zhang et al. [162]	fixed bed flow reactor	Base metal type, calcination temperature, reaction temperature, TOS		38	38
Epling et al. [163]	quartz tubular reactor	Pt content (base metal); Co content (promoter); H_2O, CO_2, O_2, H_2 percentages; reaction temperature	Pt/TiO_2*	62	0
Tanaka et al. [164]	fixed bed flow reactor	K content (promoter), reaction temperature	Rh/SiO_2**	20	0
Özdemir et al. [165]	fixed bed flow reactor	CO_2, O_2 percentages; reaction temperature; TOS		40	40
Han et al. [80]	quartz tubular reactor	Base metal type, support type, reaction temperature	Rh/MgO*	31	20
Luengnaruemitchai et al. [24]	U shaped glass tubular reactor	Preparation method; H_2O, CO_2, O_2 percentages; reaction temperature		62	62
Tibiletti et al. [166]	plug-flow reactor	Pt content (base metal), H_2 percentage, reaction temperature		30	30
Wootsch et al. [167]	continuous flow reactor	Pt content (base metal), support type, O_2 percentage, reaction temperature, F/W		46	46
Marino et al. [168]	continuous flow reactor	base metal type, support type, reaction temperature	Pt/MgO*, Pt/La_2O_3*	41	29
Yan et al. [169]	fixed bed flow reactor	Pt content (base metal); Co content (promoter); H_2O, CO_2, O_2, H_2 percentages; reaction temperature, F/W		65	65

Table 3.6. Details for the extracted data for noble metal and Au based catalysts (cont.).

Reference Publication	Reactor Type	Variables Analyzed Using the Data in the Corresponding Publication	Unique Catalysts	Number of Data Points Extracted	Number of Data Points to be Tested
Ince et al. [170]	fixed bed flow reactor	Pt content (base metal); Ce, Co contents (promoters); calcination temperature and time; O_2 percentage; reaction temperature; F/W; TOS		140	140
Luengnaruemitchai et al. [22]	fixed bed flow reactor	Au content (base metal); support type; calcination temperature; H_2O, CO_2 percentages; reaction temperature; TOS	Au/MnO*	28	14
Suh et al. [171]	fixed bed flow reactor	base metal type; Co, Mn, Ni contents (promoters); support type; reaction temperature	Pd/Al$_2$O$_3$*, Pt/Act.C*	120	96
Zhou et al. [172]	fixed bed flow reactor	base metal type, reaction temperature	Co/Act.C*, Ni/Act.C*	24	0
Chin et al. [173]	fixed bed flow reactor	support type, reaction temperature	Ru/SiO$_2$*	30	15
Pozdnyakova et al. [174]	glass tubular reactor	O_2, H_2 percentages; reaction temperature		96	96
Pozdnyakova et al. [175]	glass tubular reactor	O_2, H_2 percentages; reaction temperature	Pd/CeO$_2$*	104	0
Ko et al. [10]	fixed bed flow reactor	base metal type, support type; preparation method; Co, Ni contents (promoters); CO_2, H_2 percentages; reaction temperature	Au/CuO*	113	113
Ko et al. [18]	fixed bed flow reactor	Ni content (promoter); preparation method; CO_2, H_2 percentages; reaction temperature		40	40
Ko et al. [176]	fixed bed flow reactor	Ce, Co, Cu, Fe, Mn, Zr, Ni, Zn contents (promoters); H_2 percentage, reaction temperature		150	150
Marques et al. [177]	fixed bed flow reactor	Sn content (promoter), support type, H_2 percentage; reaction temperature	Pt/Nb$_2$O$_5$*	41	21
Cho et al. [14]	fixed bed flow reactor	Mg content (promoter); H_2O, CO_2, H_2 percentages; reaction temperature		163	163
Son [13]	quartz tubular reactor	Pt content (base metal); Ce content (promoter); H_2O, CO_2, O_2, H_2 percentages; reaction temperature; F/W		72	72
Minemura et al. [19]	fixed bed flow reactor	K, Li, Na, Rb, Cs contents (promoters)		23	23
Parinyaswan et al. [178]	U shaped glass tubular reactor	Base metal type; H_2O, CO_2, O_2 percentages; reaction temperature		88	88
Iwasa et al. [179]	continuous flow reactor	K, Li, Na, Rb, Cs contents (promoters)	Pd/ZnO**	11	0

Table 3.6. Details for the extracted data for noble metal and Au based catalysts (cont.).

Reference Publication	Reactor Type	Variables Analyzed Using the Data in the Corresponding Publication	Unique Catalysts	Number of Data Points Extracted	Number of Data Points to be Tested
Wang et al. [180]	fixed bed flow reactor	Base metal type		17	17
Wootsch et al. [78]	tubular glass reactor	reaction temperature, F/W		39	39
Monyanon et al. [181]	U shaped tubular reactor	base metal type; preparation method; H_2O, CO_2 percentages; reaction temperature		120	120
Chin et al. [182]	fixed bed flow reactor	base metal type, preparation method, reaction temperature		26	26
Ayastuy et al. [15]	plug-flow reactor	Pt content (base metal); Mn content (promoter); preparation method; O_2, H_2 percentages; reaction temperature		200	200
Souza et al. [183]	fixed bed flow reactor	support type, H_2 percentage, reaction temperature	Conversion too low compared to similar studies	60	0
Chen et al. [184]	fixed bed flow reactor	K, La contents (promoters); reaction temperature		36	36
Naknam et al. [185]	U shaped fixed bed reactor	base metal type; preparation method; H_2O, CO_2 percentages; reaction temperature	Pt-Au/zeolite*	81	0
Gluhoi and Nieuwenhuys [27]	fixed bed flow reactor	Ce, Li, Rb, Ba contents (promoters); reaction temperature		172	172
Huang et al. [186]	U shaped quartz reactor	Ir content (base metal), support type, reaction temperature	Ir/MgO*, Ir/TiO$_2$*, Ir/Al$_2$O$_3$**	49	35
Teschner et al. [187]	continuous flow reactor	CO, O_2, H_2 percentages; reaction temperature; F/W		35	35
Chang et al. [188]	fixed bed flow reactor	Mn content (promoter), reaction temperature		26	26
Wang and Lu [189]	fixed bed flow reactor	K content (promoter), reaction temperature	Au/activated C**	55	0
Galetti et al. [190]	fixed bed flow reactor	Rh content (base metal), support type, reaction temperature	Rh/CeO$_2$*, Rh/TiO$_2$*, Rh/zeolite*	51	23
Huang et al. [191]	fixed bed flow reactor	Ir content (base metal); preparation method; H_2O, CO_2, H_2 percentages; reaction temperature		68	68

Table 3.6. Details for the extracted data for noble metal and Au based catalysts (cont.).

Reference Publication	Reactor Type	Variables Analyzed Using the Data in the Corresponding Publication	Unique Catalysts	Number of Data Points Extracted	Number of Data Points to be Tested
Ribeiro *et al.* [192]	U shaped fixed bed reactor	base metal type, support type, preparation method, reaction temperature	Au/ZrO$_2$*	51	24
Scire *et al.* [193]	continuous flow reactor	Au content (base metal), preparation method, calcination temperature, reaction temperature		32	32
Luengnaruemitchai *et al.* [194]	U shaped tubular reactor	support type; H$_2$O, CO$_2$ percentages; reaction temperature	Pt/zeolite**	55	11
Iwasa *et al.* [83]	fixed bed flow reactor	Pd content (base metal); Cs content (promoter); CO, O$_2$ percentages; reaction temperature	Pd/ZnO**	18	0
Tanaka *et al.* [195]	fixed bed flow reactor	support type; K, Li, Na, Rb, Cs contents (promoters); CO, O$_2$, H$_2$ percentages; reaction temperature		126	126
Tompos *et al.* [196]	continuous flow reactor	Au content (base metal); La, Pb, Sm contents (promoters); reaction temperature		64	64
Qiao *et al.* [197]	fixed bed flow reactor	Au content (base metal), O$_2$ percentage, reaction temperature		42	42
Avgouropoulos *et al.* [198]	fixed bed flow reactor	La, Sm, Zn contents (promoters); H$_2$O, CO$_2$ percentages; reaction temperature; F/W		84	84
Zhao *et al.* [199]	U shaped quartz reactor	support type, reaction temperature		40	40
Zhang *et al.* [200]	fixed bed flow reactor	support type, preparation method, Fe content (promoter), reaction temperature	Ir/FeO*, Ir/Al$_2$O$_3$**	77	65
Yung *et al.* [201]	fixed bed flow reactor	CO, O$_2$, H$_2$ percentages; reaction temperature; F/W; TOS		71	71
Imai *et al.* [202]	fixed bed flow reactor	Au content (base metal), support type, preparation method, reaction temperature		57	57
Kim *et al.* [203]	fixed bed flow reactor	Base metal type, reaction temperature, F/W	commercial catalyst***	63	0
Neri *et al.* [204]	U shaped quartz glass reactor	support type, reaction temperature	Pt/zeolite**	12	6
Mozer *et al.* [205]	continuous flow reactor	base metal type; H$_2$O, CO$_2$, O$_2$, H$_2$ percentages; reaction temperature	Au-Cu/Al$_2$O$_3$*	70	7
Padilla *et al.* [206]	PID micro-activity Reactor	support type; reaction temperature		48	48

Table 3.6. Details for the extracted data for noble metal and Au based catalysts (cont.).

Reference Publication	Reactor Type	Variables Analyzed Using the Data in the Corresponding Publication	Unique Catalysts	Number of Data Points Extracted	Number of Data Points to be Tested
Sangeetha and Chen [207]	fixed bed flow reactor	Co, La contents (promoters); reaction temperature		18	18
Uguz and Yildirim] [138]	fixed bed flow reactor	Pt content (base metal); Ce, Co, Mg, Fe, Mn, Zr, K, Ni contents (promoters); H_2O, CO_2 percentages; reaction temperature; TOS		520	520
Woods et al. [208]	Not specified	Co content (base metal); O_2, H_2 percentages; reaction temperature; F/W;	Co/CeO_2**	38	0
Liotta et al. [209]	U shaped quartz glass reactor	Au content (base metal), support type, H_2 percentage, reaction temperature	Au/Co_3O_4*	82	50
Zhang et al. [210]	fixed bed flow reactor	preparation method, Fe content (promoter), CO_2 percentage, reaction temperature	Ir/SiO_2*	32	0
Tabakova et al. [211]	fixed bed flow reactor	Au content (base metal), support type, H_2O, CO_2 percentages, reaction temperature		86	86
Davran-Candan et al. [139]	fixed bed flow reactor	Ce, Co, Mg, Fe, Mn, Ni contents (promoters); H_2O, CO_2 percentages; reaction temperature		252	252

* Catalyst studied in only the current publication, ** catalyst studied in only one more publication, *** preparation details unknown.

Table 3.7. Input variables and their ranges for noble metal and Au based catalysts.

Input Variable	Type	Range (for continuous variables) or Identity (for categorical variables)
base metal type and weight percentage	categorical/continuous	Pt (0-5), Au (0-10), Ru (0-5), Rh (0-5), Ir (0-5), Cu (0-3.3), Pd (0-10)[*]
preparation method	categorical	incipient to wetness impregnation (IWI), wet impregnation (WI), co-impregnation (CI), sequential impregnation (SI), sol-gel precipitation (SGP), co-precipitation (CP), deposition precipitation (DP), sequential precipitation (SP), homogenous precipitation (HP)
calcination temperature (^{o}C)	continuous	110-900
calcination time (h)	continuous	0.5-13
support type	categorical	Al_2O_3, MgO, CeO_2, CuO, TiO_2, ZnO, MnO, FeO, Nb_2O_5, SiO_2, ZrO_2, La_2O_3, Co_3O_4, zeolite, activated carbon
second metal (promoter) type and wt.%	categorical/continuous	Ce (0-7.2), Co (0-10), Mg (0-3.1), Fe (0-4.4), Mn (0-15), Zr (0-4.7), K (0-20), Ni (0-3), Sn (0-3), Li (0-2), Na (0-3.5), Rb (0-16.3), Cs(0-13.6), Nb (0-5), La (0-12), Ba(0-6), Pb(0-3.5), Sm(0-5.4) Zn(0-3.4)[*]
operating temperature (^{o}C)	continuous	(-20) - (+380)
H_2 vol.%	continuous	0-99.3
O_2 vol.%	continuous	0.2-5
CO vol.%	continuous	0.2-5
H_2O vol.%	continuous	0-10.7
CO_2 vol.%	continuous	0-30
time on stream (min)	continuous	0-600
F/W (cm^3/(h.g))	continuous	3000-270000

[*]Numbers in parentheses represent the range of metal weight percentages

4. COMPUTATIONAL DETAILS

4.1. Artificial Neural Network Modeling

Artificial neural networks were created by using computer codes written in MATLAB environment. For the studies given under Section 5.1, modular networks trained by backpropagation algorithm (with delta rule of error correction) were used [37]. In order to prevent any variable dominate the others, all the input variables were normalized (before beginning the training process) in the range of 0.1-0.9 as shown in Equation 2.17 (min-max normalization). The random initialized weights between -0.1 to +0.1 were created and the training was done accordingly. In order to eliminate the possible extreme conditions occurring due to the random initialization of the weights, each network was trained three times and the best solutions were kept. The logistic sigmoid function (Equation 4.1) was employed on the entire network as the main activation function to determine the optimal neural network structure. Then, the best network was further optimized by testing different activation functions (hyperbolic tangent function and identity function) for each input group, so that the advantage of using modular network structure can be fully utilized [115]. The other activation functions with their derivatives are given in Equation 4.2 (hyperbolic tangent function) and Equation 4.3 (identity function) [115], where x is the input to a particular neuron, $\sigma(x)$ is the output of that neuron, and $\sigma'(x)$ is the corresponding derivative of the activation function.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$
$$\sigma'(x) = \sigma(x)\big(1-\sigma(x)\big)$$

(4.1)

$$\sigma(x) = \tanh(x)$$
$$\sigma'(x) = \big(1-\sigma^2(x)\big)$$

(4.2)

$$\sigma(x) = x$$
$$\sigma'(x) = 1$$

(4.3)

The artificial neural network models presented in Section 5.2.1 had similar parameters, except they were in monolithic structure (fully connected), and logistic sigmoid function was always used as the transfer function.

The models presented in Section 5.3 were trained by the Levenberg-Marquardt method, which is a blend of gradient descent and Gauss-Newton iteration [49, 117, 118]. The hyperbolic tangent function was used as the activation function in the hidden layers, and the mean square error (MSE) was employed as the measure of fitness. Each network topology was trained 10 times to compensate the effects of random initialization of the neural network weights, and the best performances were recorded. To prevent overlearning of the neural networks, the early-stopping technique was applied during the training of the neural networks by using some random data points among the training set as the validation data.

Independent of the neural network training method, the optimal network topology was always determined by testing the performances of several networks according to their generalization accuracies (the ability to predict the data unseen during the training process) [212]. The root mean square error (RMSE) of testing, which indicates the degree of the generalization accuracy [121], was estimated by applying the k-fold cross validation technique [43, 51]. The entire database was divided randomly into k subsets, and then the data acquired from k-1 sets were used to train the network to predict the outcome of the remaining one set. The errors between these predictions and the corresponding experimental results were recorded. This procedure was repeated k times covering all the data points, and the RMSE (Equation 4.4) calculated through the entire database was used as the indicator to determine the optimal neural network topology.

$$RMSE \;=\; \sqrt{\frac{1}{n}\sum_{1}^{n}\left(p_i - t_i\right)^2} \qquad\qquad (4.4)$$

Where; p_i is the predicted, t_i is the target value and n is the total number of experiments.

The coefficient of determination (R^2) was also calculated to indicate the fitness of the models. The RMSE of training, which is the error obtained between the experimental data points and the model predictions when all the data are used for training, was also calculated, together with the corresponding R^2 values, to give some idea of the suitability of the database to construct a single model, although it was not used in decision making.

The test of significance for the continuous input variables was done by partial differentiation of the fully trained neural network, which was reported as a very successful tool for the identification of primary factors controlling catalytic activity [47, 48]. On the other hand, "the change of root mean square error" technique [51, 213] was used to test the significance of the categorical inputs (promoter type) since the method of partial differentiation cannot be applied to discrete data.

4.2. Decision Tree Modeling

Decision trees were employed for the study in Section 5.1.2. The program codes were created by writing computer codes in MATLAB environment. The error rate for a leaf node, which is the fraction of wrong classifications to the total records in that node, was minimized by trying several candidate splits and choosing the best separation. The total error rate of the entire tree, which is the weighted average of the individual leaf error rates, should be small for a successful classification. Even though large trees may have very small error rates, they may cause the overlearning of the data, and the rules they deduce may be too complex for practical purposes. For this reason, the tree size was optimized to have an acceptable error rate for training (error from the data used to construct the tree) and the smallest error for testing (error from the data not seen by the model before), which is the indicator of the generalization ability of the model [67, 68]. The data was randomly divided into two equal sets: one set was used for training and constructing the decision tree model while the other was used for testing its generalization ability.

4.3. Multiple Logistic Regression Modeling

The logistic regression models were also created using computer codes written in MATLAB environment and applied for the work presented in Section 5.2.2. The significance of the entire model was examined by the G test (Equation 2.12). The significance of each predictor variable was calculated by Wald test (Equation 2.15). Since the details of the method have already been given in Section 2.4.3, they are omitted here.

4.4. Genetic Algorithm Based Clustering

This method was employed for pre-processing of the data presented in Section 5.3.2. A population size of 20 strings (candidate solutions of the problem) each containing all the centers of the clusters. Roulette wheel method was implemented for selection of the strings to reproduce (with a crossover probability of 0.8) and one point crossover was employed between these strings. A mutation probability of 0.01 was applied for each string; in addition, strings containing empty clusters were eliminated from the population and they were replaced by random new strings (immigration) in each generation while the solution was evolving towards the optimum cluster centers. The full procedure of genetic algorithm based clustering has been given in Section 2.4.6.

5. RESULTS AND DISCUSSION

Results and Discussion are presented in three sections. First, knowledge extraction from the experimental data generated in our laboratories are presented in three subsections: "modeling promoted Pt/Al_2O_3 catalysts for selective CO oxidation using modular neural networks", "modeling promoted Au/Al_2O_3 catalysts for selective CO oxidation using decision trees and modular neural networks" and "modeling promoted Pt/Al_2O_3 catalysts for water gas shift activity by modular neural networks". Then, the results of the DFT generated data modeled by two different methods are discussed in two subsections: "structure and activity relationship for CO and O_2 adsorption over gold nanoparticles using DFT and artificial neural networks" and "analysis of O_2 adsorption stability and strength over gold nanoparticles using DFT and logistic regression". Finally, the modeling results of the data collected from the publications in the literature are discussed in two subsections: "neural network analysis of selective CO oxidation over copper-based catalysts for knowledge extraction from published data" and "analysis of selective CO oxidation over noble metal-gold based catalysts for knowledge extraction from published data by clustering and artificial neural networks".

5.1. Knowledge Extraction from Experimental Data

5.1.1. Modeling Promoted Pt/Al_2O_3 Catalysts for Selective CO Oxidation Using Modular Neural Networks

CO conversion, which is the main indicator of the performance of the catalytic system studied, was used as the output variable (with 520 different data points). The input variables were first classified in two main groups as preparation and operating variables. Then, the preparation variables were divided further into two subgroups as Pt amount (group 1) and promoter type (group 2), which were continuous and categorical type of variables, respectively as shown in Section 3.1 Eight input neurons, each assigned to a different promoter type, were used for categorical variables indicating whether that particular promoter exists (indicated as 1) or not (indicated as 0). Likewise, the operating

variables were also divided into two subgroups as group 3 (reaction temperature, CO_2 content and H_2O content in the feed) and group 4 as time on stream (TOS) considering that the effect of TOS on catalytic performance is related to the "stability" of catalysts while the others primarily affect the "activity".

First, the neural network structure having sufficiently high prediction and generalization accuracies in representing the experimental data was searched. Then, the optimal neural network was used to analyze the significances of input variables and their effects on CO conversion [51].

5.1.1.1. Determining the Optimal Neural Network Topology. Fourteen neural networks with 13 input variables and one output variable (percent CO conversion) were constructed in different architectures with different number of hidden layer neurons. The notation of "(b-c-d-e)-f" is used to label the neural networks, where "b, c, d, e" represent the first hidden layer neurons connected to four different input groups (with the order of Pt wt.%, promoter type, reaction conditions and time on stream) and "f" shows the number of neurons in the second hidden layer.

The RMSE of training was calculated by the error obtained between the experimental data points and the model predictions when all the data was used for training. On the other hand, the RMSE of testing, which was calculated by 5-fold cross validation analysis, was applied to test the generalization accuracy (prediction ability for the unseen data). Since the change of the "time on stream" does not change the main character of the catalyst like the other variables, the data of a particular catalyst at all time on stream values were excluded simultaneously while preparing the cross validation subsets. Otherwise, the exclusion of one data point at one time on stream value would be compensated by the others, and the validation results would demonstrate the network more successful than it really was [34, 51].

The training and testing errors of fourteen networks are compared in Figure 5.1 with the order of increasing number of weights. The grey bars in Figure 5.1 represent the training error (indicating the prediction accuracy) of that particular network while the black bars represent the validation error (indicating the generalization accuracy).

Figure 5.1 shows that there is an obvious reduction of the training RMSE (increase of prediction accuracy) at the beginning with increasing network size. As the network gets larger, the modeling error continues to decline; on the other hand the experimental error, which is independent of the network size, keeps unchanged. Therefore, the decrease in the training RMSE, which is composed of both the modeling and the experimental error, slows down as the network gets larger.
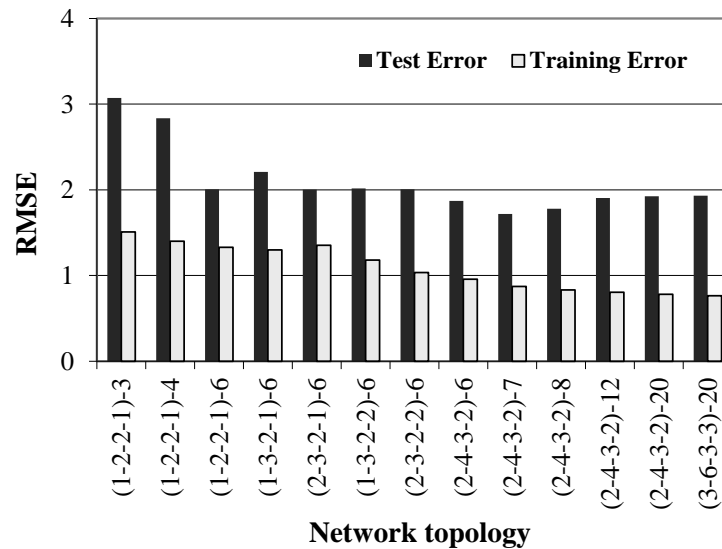
Figure 5.1. Training and testing errors of various neural network topologies [51].

The testing error, as indicated by the black bars, decreases first with increasing network size and starts to increase (probably due to some degree of over-fitting) after reaching a minimum. This increase in the validation error for larger networks is quite natural, since the number of data points should be several times higher than the number of weights of the network for a sufficient degree of generalization [214].

It should be noted that the generalization ability of a network, which measures the success of estimating the unseen data, is more valuable than the prediction ability of the training data (involves the prediction of the data already used for training). The network structure 1(2-4-3-2)-7 exhibits the minimum RMSE of testing (1.718) with a considerably low RMSE of training (0.872), and thereby it was used in the remaining part of the work as the one that best represents the experimental data among the networks tested. The schematic representation of this network is shown in Figure 5.2, with black circles as the

bias neurons having a constant value of 1. Since the number of connections among the layers is very large, only some representative lines were drawn in figure for simplicity while the exact number of neurons was used.
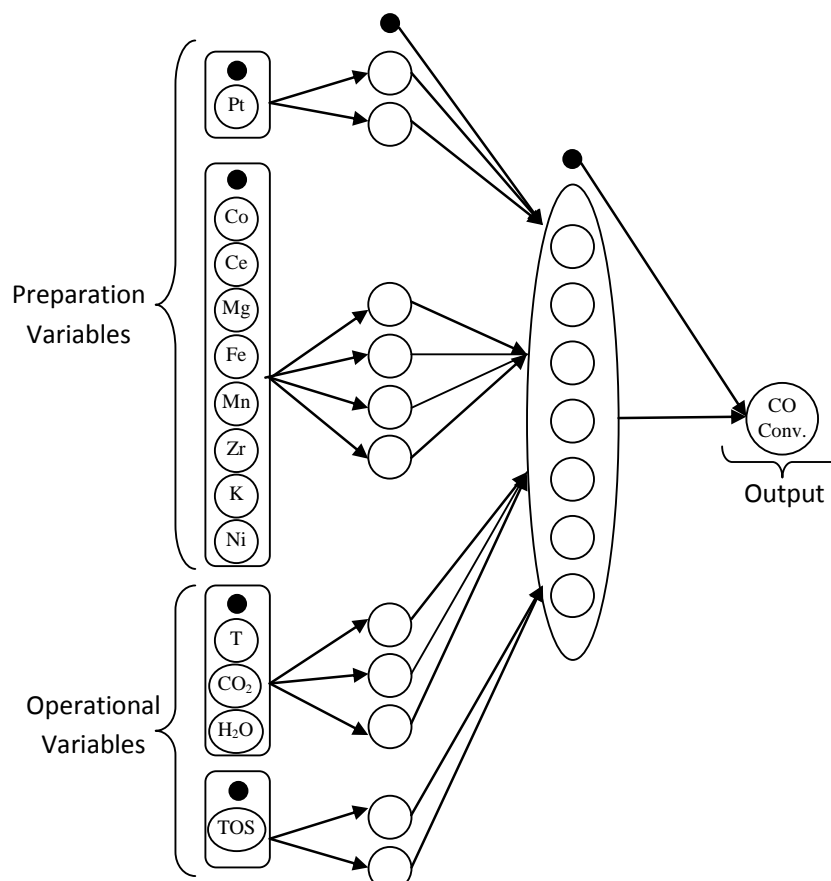
Figure 5.2. Optimal network topology for selective CO oxidation over Pt/Al_2O_3.

The statistical performance of (2-4-3-2)-7 neural network, which has 148 weights (about 2/7 of 520 experimental data points), was also compared with a monolithic neural network (7 neurons in the first hidden layer, 5 neurons in the second) having almost the same number of connections (144). Although the training RMSE value of the modular network (0.872) was a little higher than the training RMSE of the monolithic network (0.779), its RMSE of testing value (1.718) was much lower than that of the monolithic network (1.969). This is an expected result, since properly arranged partially connected networks have less unnecessary connections showing usually higher generalization accuracy than fully connected neural networks [54, 55]. Besides, the performance of the modular networks can further be improved using different activation functions for different

modules considering that an activation function suitable for a particular problem may not be appropriate for another type [115]. This flexibility does not exist for the monolithic networks. Hence, different activation functions on each input group were tested in the first layer of the modular network (2-4-3-2)-7 in various combinations while the logistic sigmoid function was kept in the remaining layers. The results of some combinations are shown in Figure 5.3. The performance of the network discussed above (fully activated by the logistic sigmoid function in all modules) was also added to the same figure for statistical comparison. The bars in the figure are labeled using the abbreviation of activation functions in the order of input groups. For example, "lsf-tanh-ide-lsf" indicates the use of "logistic sigmoid function (lsf)" for group 1, "hyperbolic tangent function (tanh)" for group 2, "identity function (ide)" for group 3 and "logistic sigmoid function (lsf)" for group 4.
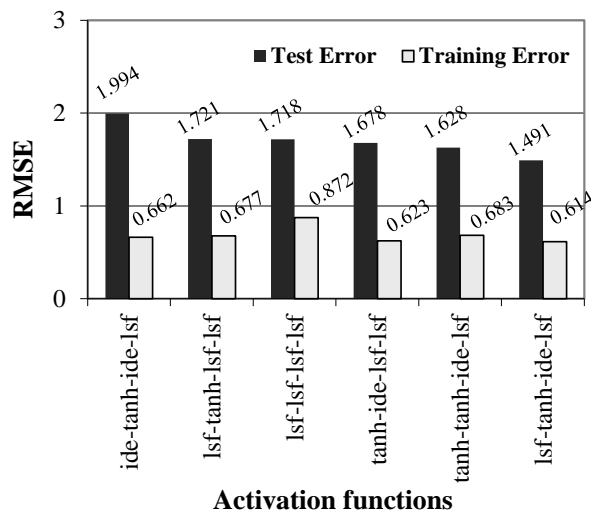
Figure 5.3. Training and testing errors of various neural network topologies with different activation functions in the first hidden layer [51].

Figure 5.3 indicates that the network activated with the combination of "lsf-tanh-ide-lsf" has the highest generalization accuracy. The training and the testing errors of 0.614 and 1.491 in this combination are much lower than those of the modular network fully activated by the logistic sigmoid function (0.872 and 1.718, respectively). The superiority of the modular network for this type of problems becomes more apparent after utilization of this flexibility since the training and validation errors of the corresponding monolithic network was 0.779 and 1.969 respectively. As a result, the modular network with (2-4-3-

2)-7 structure and with the "lsf-tanh-ide-lsf" activation function set were used for the remaining parts of the work to extract additional knowledge. The experimental versus predicted CO conversion plots for both training and the testing are given in Figure 5.4.
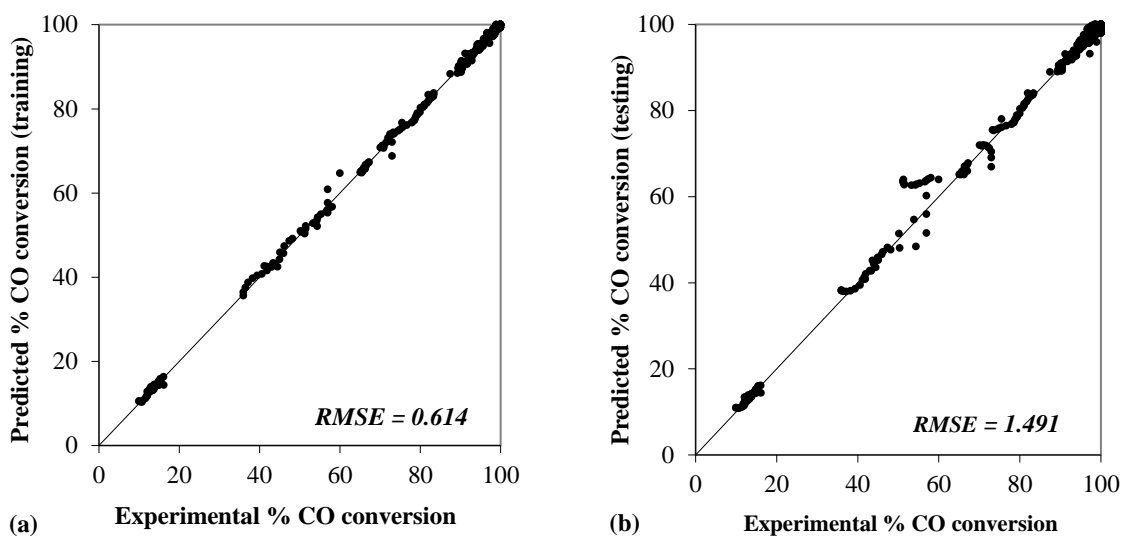
Figure 5.4. Experimental versus predicted CO conversion for: (a) training, (b) testing data by the optimal neural network topology [51].

5.1.1.2. Analyzing the Input Significance. The relative group importance was calculated by the method of change of mean square error, and the results are in the first column of Table 5.1. First, the preparation variables were deactivated and the model was trained with the rest of the data using only the operating variables; then, the same procedure was reversed. The increase of RMSE in each trial was noted and the relative group importance was found accordingly. The preparation variables were found to have 69.49% relative significance while the operating variables had 30.51% total significance.

Then, the relative significances of the preparation variables with respect to each other were tested. The change of RMSE was used for this purpose due to the categorical nature of the first and the second promoter types. The same method was applied for Pt wt.% for comparison, although this parameter is a continuous variable for which the partial differentiation is a better method.

Table 5.1. Relative significances for Pt/Al_2O_3 catalyst for selective CO oxidation.

Input Variables	% Group Significance (change of RMSE)	% Global Relative Significance	
		(change of RMSE)	(partial differentiation)
Pt wt.%		17.12	-
Co, K, Ni	69.49	45.37	-
Mg, Ce, Fe, Mn, Zr		37.51	-
Rxn. Temp		-	57.51
H_2O%	30.51	-	14.66
CO_2%		-	16.59
TOS		-	11.25

The first promoter type was found to have a relatively higher significance than the second promoter type as expected considering that Co catalyzes CO oxidation itself [11, 12, 138], although the metals used as the second promoter contribute through O_2 adsorption. Pt wt.% was found to have relatively lower significance value compared to those of the first and the second promoters. However, this does not mean that Pt, which is the primary constituent of the catalysts, has less impact on the catalyst. The value of 17.12% of relative importance in Table 5.1 indicates the sensitivity of CO conversion to the change of Pt wt.% between 0.7-1.4, while the percent significances of the promoters indicate the sensitivity of CO conversion to their presence or absence.

The method of partial differentiation was applied to the operating variables, all of which are continuous. The most important process parameter was found as the reaction temperature, which is also quite reasonable considering low temperature CO oxidation reactions are too sensitive to reaction temperature changes [15, 215]. This is followed by CO_2 vol.%, H_2O vol.% and TOS with the significances in decreasing order.

5.1.1.3. Analyzing the Effects of the Catalyst Variables. The optimal neural network model proposed above was used for analyzing the effects of Pt wt.%, reaction temperature and second promoters (excluding Zr) on the catalysts with Co as the first promoter (Figure 5.5). The results for the catalysts containing K and Ni as the first promoter and Zr as the second promoter are not presented here because these catalysts were not as much active as those discussed here.
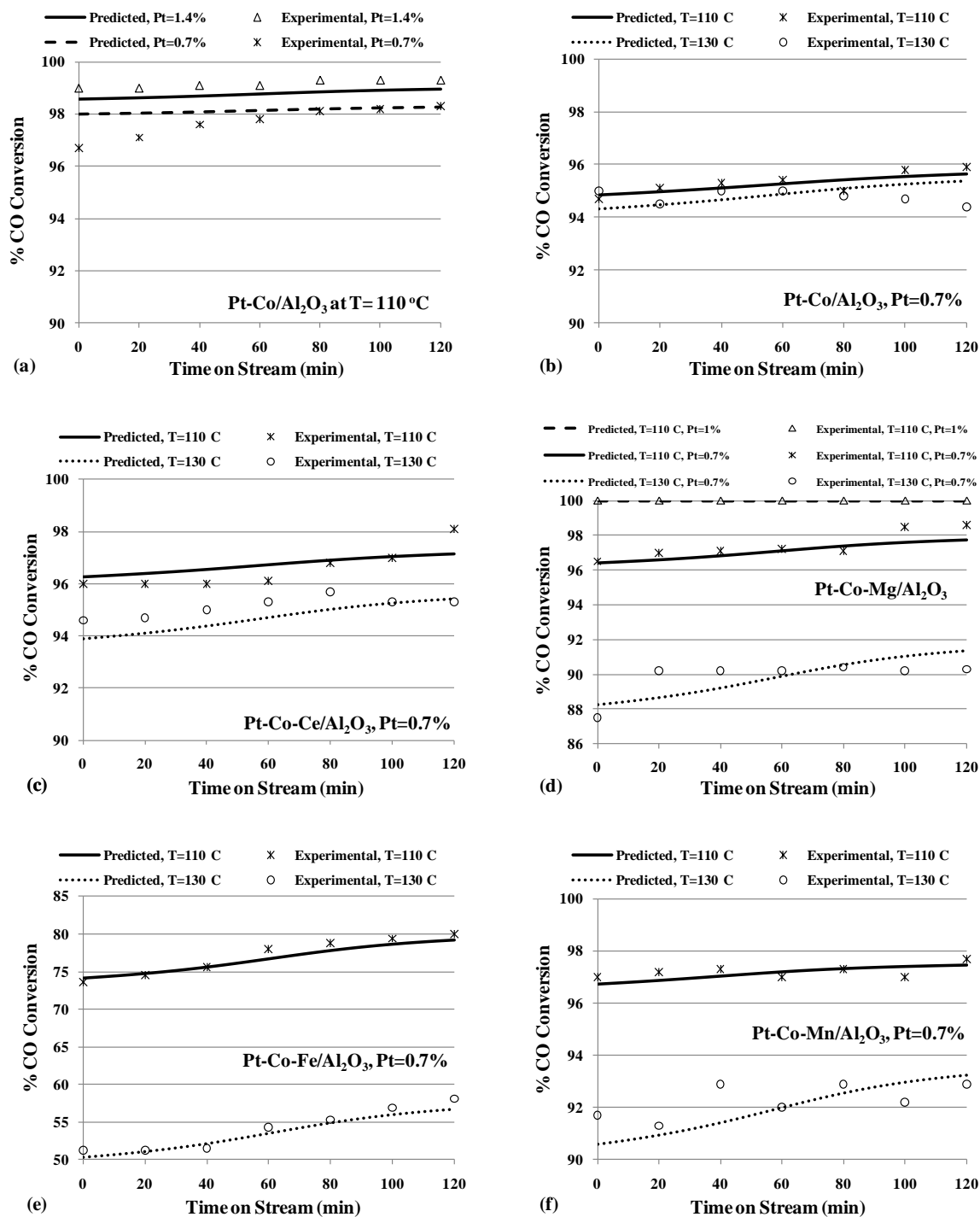
Figure 5.5. Effects of variables on CO conversion: (a) effect of Pt wt.% in the absence of H_2O and CO_2; (b-f) effects of reaction temperature ($H_2O = 10\%$, $CO_2 = 25\%$) [51].

Figure 5.5a shows the effect of Pt wt.% on CO conversion for the Pt-Co catalyst without any secondary promoter in the absence of H_2O and CO_2, together with the experimental values (triangles for 1.4 wt.% Pt, stars for 0.7 wt.% Pt). The higher Pt content

provides higher CO conversion as expected and the model predictions are in quite good agreement with the experimental values [138, 203]. The discrepancies between model predictions and experimental data at low TOS for the catalyst containing 0.7 wt.% Pt seems to be visually large due to the scale of the graph although the actual deviation is only about 2% indicating a very good fit. The model is even more successful for 1.4 wt.% Pt containing catalyst.

The CO conversion over 1.4 wt.% Pt catalyst in the presence of a second promoter is nearly 100% making to analyze the effects of the variables on the activity of the catalyst almost impossible [138]. Thus, the remaining part of the analysis was performed using the experimental data obtained over 0.7 wt.% Pt containing catalyst (Figure 5.5b-f).

Figure 5.5b compares the experimental and the predicted temperature effects on CO conversion over 0.7 wt.% Pt-Co catalysts while Figure 5.5c-f present the same comparison in the presence of various second promoters. The model predictions are quite reasonable for all cases indicating that the CO conversion decreased significantly with increasing temperature from 110 ^{o}C to 130 ^{o}C over all the catalysts having a second promoter. This is explained by the fact that CO has the ability to cover the whole surface of the Pt catalyst providing the CO oxidation to occur as the main reaction at low temperatures; however, as the temperature increases H_2 competes with CO to be adsorbed on the catalyst surface causing a decrease in the CO oxidation activity [15]. This decrease was minimum for Pt-Co/Al_2O_3 in the absence of a second promoter.

Uguz and Yildirim [138] reported that the addition of Mg increased the activity of Pt-Co/Al_2O_3 most while Mn and Ce also have some positive effects. The addition of Fe on the other hand, has a significant negative effect on CO conversion. Hence, they repeated the experiments using 1 wt.% Pt containing Pt-Co-Mg/Al_2O_3 catalyst and they found 100% CO conversion at 110 ^{o}C for all values of TOS as correctly predicted by our model in Figure 5.5d.

The CO conversions in the absence and presence of 25% CO_2 and 10% H_2O were also compared in the experimental work considering that these two components exist in a realistic gas mixture from the fuel processor at these concentrations, and the catalyst to be

used for the selective CO oxidation should be active in their presence [138]. The model predictions are in a good agreement with the available experimental data points as it is shown in Figure 5.6. Although only two experimental data points can be normally considered as insufficient to test the success of the model, the general trend predicted by the model is in accordance with the literature indicating that the presence of CO_2 has detrimental influence while H_2O enhance the catalytic activity. It was reported that the competitive adsorption of CO_2 on Pt surface decreases the rate of CO adsorption, and CO_2 in the feed can cause reverse water gas shift reaction producing CO and consuming H_2 [160]. The positive effect of H_2O, on the other hand, was attributed to the generation of additional free Pt sites by hydrated Al_2O_3 [84], to the possible water gas shift activity [160, 170] and to formation of hydroxyl groups (from the dissociation of H_2O) assisting to the interaction of CO and O_2 [216].
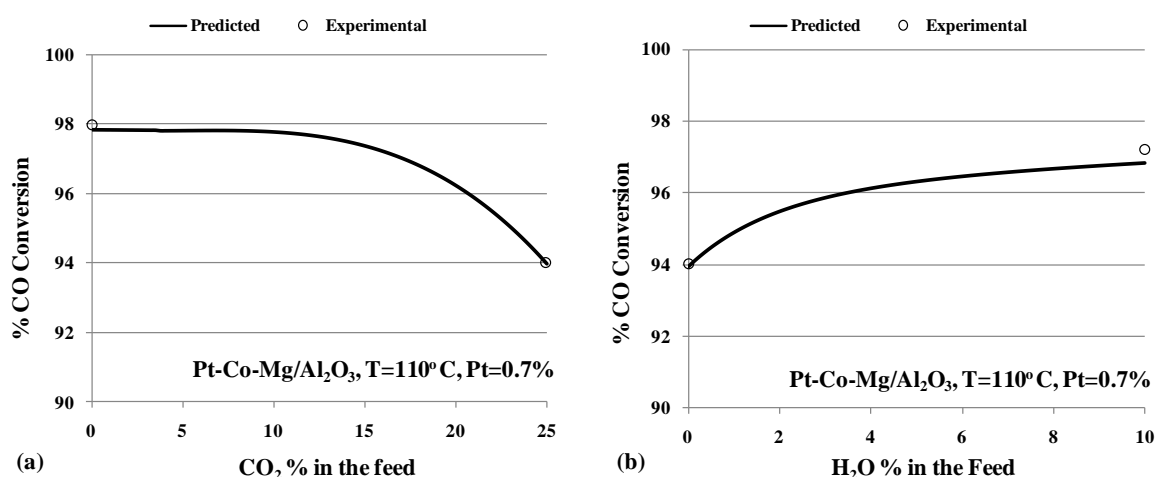
Figure 5.6. Effects of CO_2 (a) and H_2O (b) in the feed stream on CO conversion: (a) Pt = 0.7 wt.%, T=110°C, TOS=120 min, H_2O = 0%; (b) Pt = 0.7 wt.%, T=110°C, TOS=120 min, CO_2 = 25% [51].

5.1.2. Modeling Promoted Au/Al_2O_3 Catalysts for Selective CO Oxidation Using Decision Trees and Modular Neural Networks

The procedure used to analyze the CO oxidation data (given in Section 3.2), included two steps as shown in Figure 5.7. The first step was to classify the entire catalytic data composed of 252 points by decision trees to derive rules to identify the variables and their ranges that lead high catalytic performance. The data set was then reduced by eliminating

the data belonging to the conditions that were not favoring high catalytic activity. As the second step, the new reduced dataset was analyzed using modular neural networks to extract additional knowledge about the effects of variables and their relative significances.
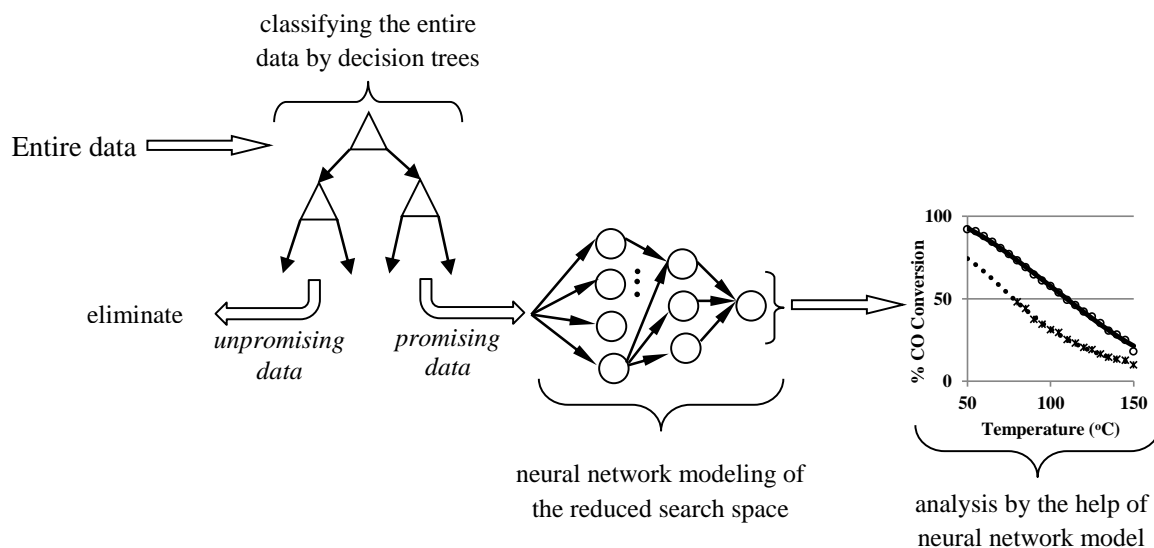
Figure 5.7. Conceptual approach of knowledge extraction from experimental data.

5.1.2.1. Classification of the Data Using Decision Trees. The experimental data of Au/Al_2O_3 catalysts were classified according to their CO conversion levels as very high (90-100), high (80-90), moderate (50-80), low (25-50) and very low (0-25) by using decision trees Half of the data (randomly selected) was used to construct the decision tree and to train it; the other half was used for testing the generalization ability of the decision tree. The optimal decision tree was found by first starting with a large tree and then pruning it from the lower branches until the minimum test error was found. The training and the testing error rates for various tree sizes are shown in Figure 5.8. The training error decreased as the tree size increased while the testing error decreased first and then increased at large trees due to the overlearning of the data, as expected. The tree with the minimum testing error (with maximum generalization ability) was found to be the one with 14 terminal nodes (Figure 5.8) having a total training and testing errors of 11.9% and 18.3% respectively. Figure 5.9 shows the optimal decision tree. The variables that were used to separate the data are given on the branches of the tree while the classes of the data as well as the number of the erroneous classifications during training and testing are

presented in terminal nodes. The distribution of training and testing errors among the conversion levels for the optimum tree structure are presented in Table 5.2 and Table 5.3.
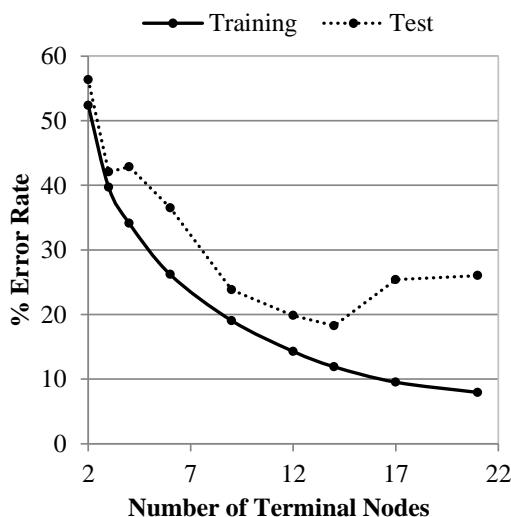
Figure 5.8. Change of error rates with size of decision tree.

The classification by the decision tree was quite successful considering that 111 out of 126 data points were correctly classified during training, and all of the wrong classifications took their places in one level higher or lower categories they were actually belonged to. For example, a moderate performing catalyst, which was classified wrongly, was placed into either high or low performance levels but it was never labeled as a very high or a very low performing catalyst.

The success of the tree is also apparent from the results of testing (Table 5.3); 103 out of 126 catalysts (not seen by the tree during training) were correctly placed into their performance levels. Even the predictions for the remaining 23 catalyst were not completely wrong considering that they were placed into the neighboring performance levels similar to the erroneous placement for training.

Figure 5.9 reveals that the data were first divided at the top of the tree according to the reaction temperature indicating that this variable was the most significant factor determining the CO conversion level. Low temperatures (T<67.5 ^{o}C) leaded to higher CO conversions while higher temperatures (T>107.5 ^{o}C) usually resulted lower CO conversions for all the catalysts tested. At intermediate temperatures; on the other hand;

low, moderate or high conversion levels were possible depending on the promoter type and/or the presence of H_2O and CO_2, which were the other classifiers.
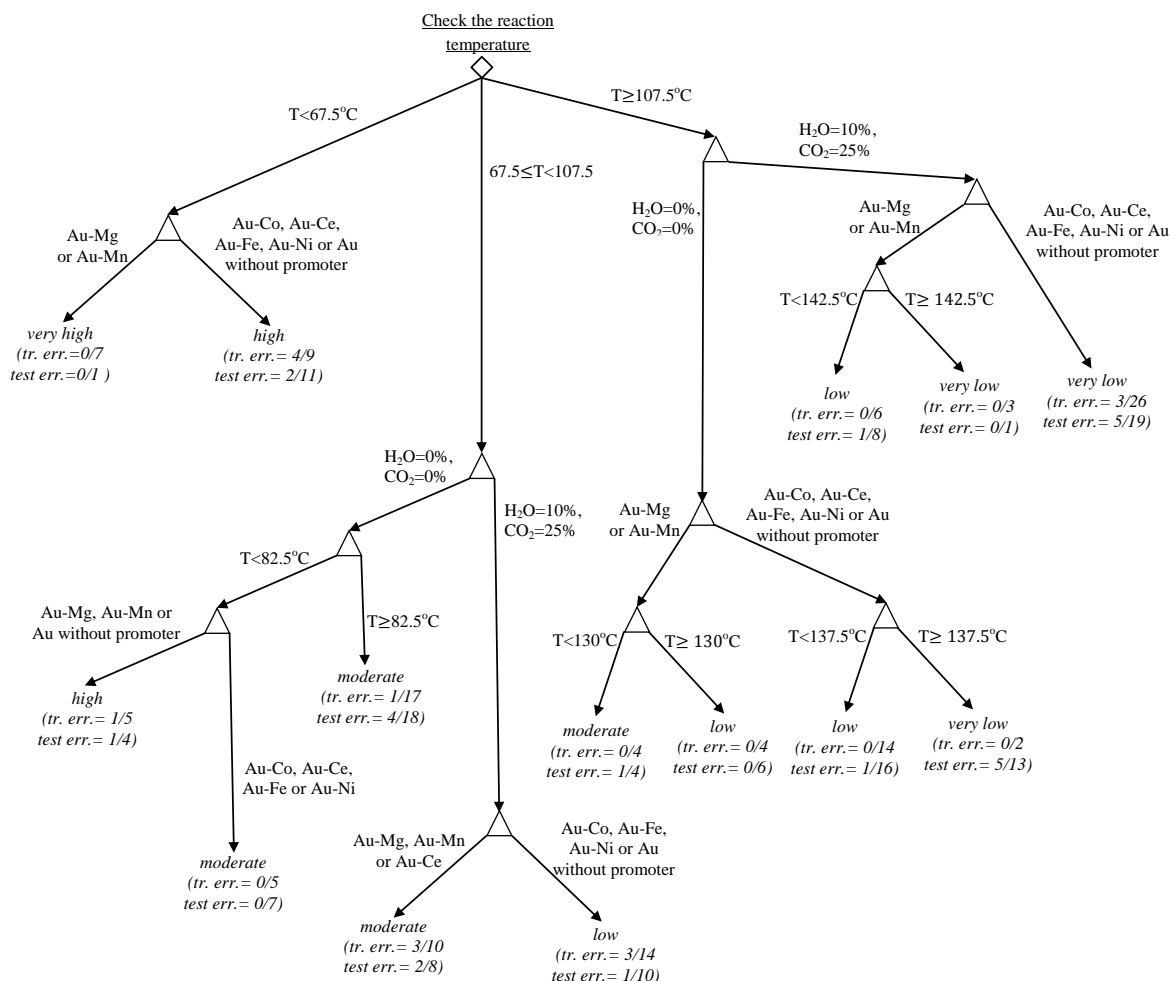
Figure 5.9. Optimal decision tree structure ("tr. err." refer to "training error", "test err." refer to "test error").

The catalysts promoted by Mg or Mn exhibited higher CO conversion compared to the catalysts with the other promoters as observed in the branches of the optimal tree. It is believed that Au particle size directly affects CO oxidation on gold catalysts: smaller the Au particles dispersed on a support, higher the catalytic activity is [26, 217]. Mg as a promoter stabilizes the small Au particles on the support improving CO oxidation activity of the catalyst [26, 158]. The addition of Mn, on the other hand, was reported to increase CO oxidation rate by providing additional active oxygen for the reaction [158, 218]. Although, the addition of Ce might also act as an oxygen supplier, it was reported to be

less effective promoter for Au catalysts for low temperature CO oxidation [27]. The addition of Co, Fe or Ni were also known to be less effective [218]. All these trends were successfully indicated by the tree.

Table 5.2. Prediction errors for individual performance levels for the training data.

Experimental data		Predictions for training					
Conversion	number of data	very high	high	moderate	low	very low	% error
very high	9	7	2				22.2
high	11		9	2			18.2
moderate	36		3	32	1		11.1
low	40			2	35	3	12.5
very low	30				2	28	6.7

Table 5.3. Prediction errors for individual performance levels for the testing data.

Experimental data		Predictions for testing					
Conversion	number of data	very high	high	moderate	low	very low	% error
very high	1	1					0
high	12		12				0
moderate	35		3	30	2		14.3
low	54			7	37	10	31.5
very low	24				1	23	4.2

5.1.2.2. Determining the Optimal Neural Network Topology. Since the catalysts promoted by Co, Ce, Fe and Ni had lesser performance under the conditions studied, the data related to these catalysts were removed from the database. Then, the neural network models were constructed using only the reduced data set containing 108 data points (belonging to Au/Al$_2$O$_3$, Au-Mg/Al$_2$O$_3$ and Au-Mn/Al$_2$O$_3$ catalysts). Although the conversion was also low at high temperatures, no data was eliminated based on this variable to be able to analyze the effects of the reaction temperature on CO conversion in a wider range.

Eight different modular neural network topologies were constructed. The optimal neural network topology with the minimum testing error was found to be 5-(2-2)-3-1 (indicating five input neurons, two modules each containing two neurons in the first hidden layer, three neurons for the second hidden layer and one output). This network was found to have a training RMSE of 1.05 and a testing RMSE of 3.69. The corresponding topology is shown in Figure 5.10.
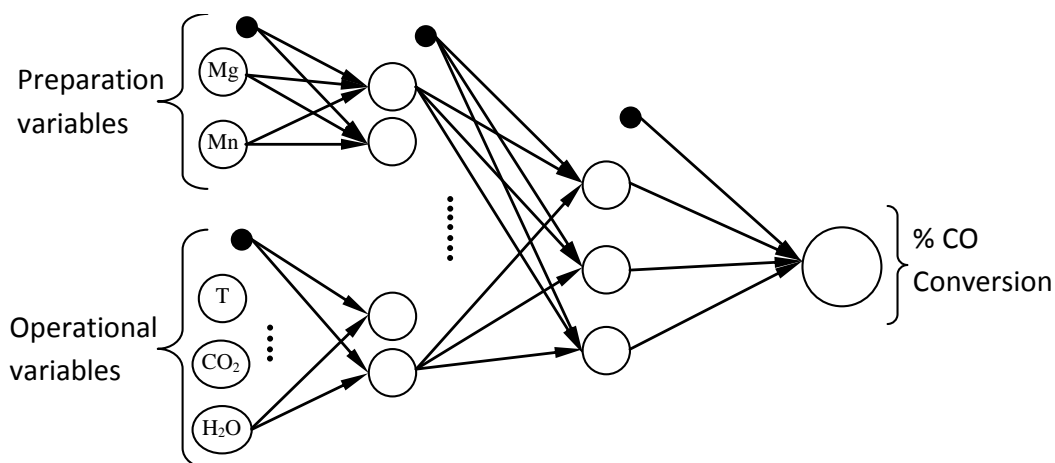
Figure 5.10. Optimal network topology for selective CO oxidation over Au/Al_2O_3.

The statistical performance of the optimal network topology was further improved using different activation functions for the preparation and the operating variables; this is one of the advantages of using modular neural networks. Logistic sigmoid function (lsf), hyperbolic tangent function (tanh) and identity function (ide) were tested for each module in the first hidden layer while logistic sigmoid function was kept in the second hidden layer all the times. Figure 5.11 shows training and testing errors of all the combinations of activation functions sorted according to the testing error. The labeling of the bars in the figure was done by the activation functions used for the preparation variables (first module) and the operating variables (second module), respectively.

According to Figure 5.11, the generalization accuracy of the optimal neural network was improved most with the use of hyperbolic tangent function for the preparation variables and identity function for the operating variables, providing a testing RMSE of 1.31. The plots of predicted versus experimental CO conversion values for this network are shown in Figure 5.12 indicating that the model exhibits excellent fits for both training and testing data.
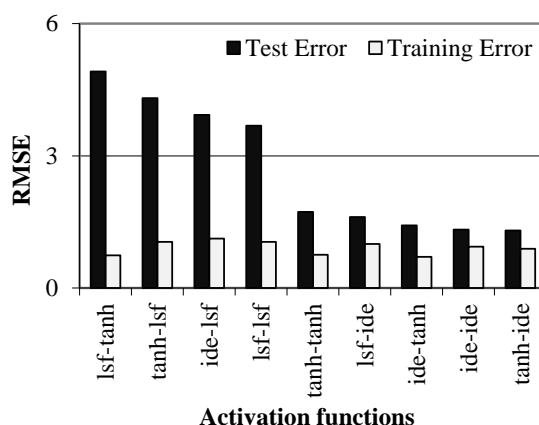
Figure 5.11. Testing and training errors of optimal neural networks with various combinations of activation functions for the first hidden layer.
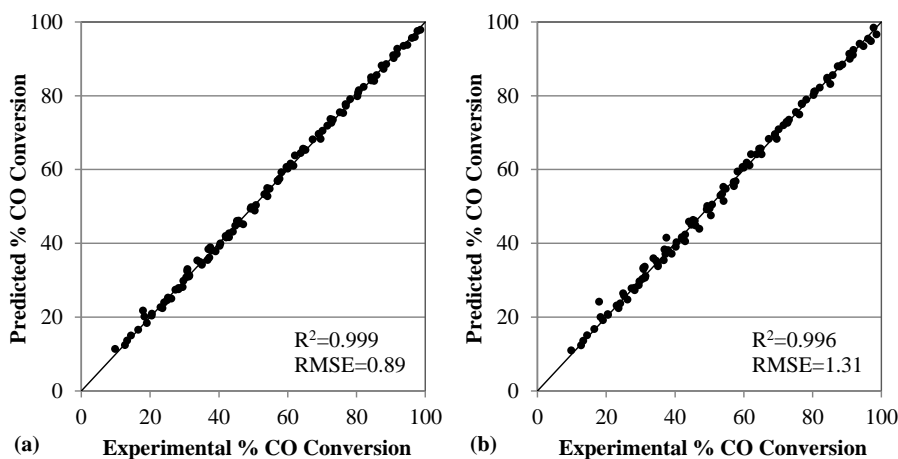
Figure 5.12. Experimental versus predicted CO conversion obtained using the optimal network (a) training, (b) testing.

5.1.2.3. Analyzing the Effects of Catalyst Variables and Their Relative Significances. The optimal neural network topology, proposed above, was used to screen the effects of input variables on CO conversion. Figure 5.13 shows CO conversion versus temperature in the absence and presence of H_2O and CO_2 for all the three catalysts analyzed (Au/Al$_2$O$_3$, Au-Mg/Al$_2$O$_3$ and Au-Mn/Al$_2$O$_3$). Apparently, the reaction temperature was the most influential variable; for example, CO conversion drops from 99% to about 35% in the absence of H_2O and CO_2 for the Au-Mg/Al$_2$O$_3$. For the same feed composition, the performances of the catalysts containing Mg or Mn were found to be higher than the catalyst without promoter indicating the positive effects of these promoters.

It is generally believed that the presence of H_2O positively affects CO conversion by creating surface OH groups assisting the activation of O_2 [148]. The detrimental effect of CO_2 is, on the other hand, attributed to the reverse WGS reaction, formation of carbonates and adsorbed carboxylates [23, 24, 142]. The negative effect of CO_2 seems to be stronger than the positive effect of H_2O under the conditions studied, considering that CO conversions were lower when both were present than the conversions obtained in their absence (Figure 5.13). Both Mg and Mn provided some resistance to the negative effects of CO_2 compared to the catalyst without promoter while Mg appeared to be significantly better.
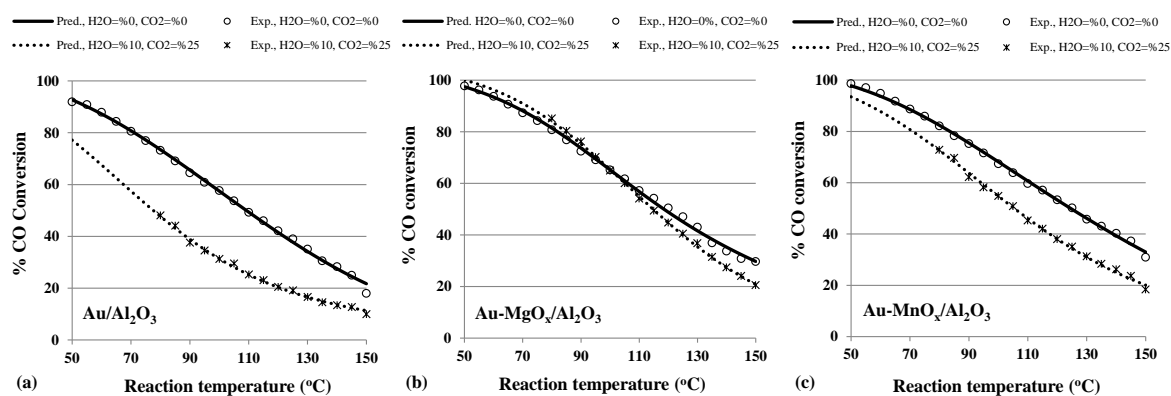
Figure 5.13. Experimental and predicted CO conversions versus temperature for Au catalysts in the absence and presence of CO_2 and H_2O.

Table 5.4 shows the relative significances of the catalyst preparation and operating variables calculated using the change of RMSE method (for promoters) and partial differentiation method (continuous variables). The relative significances of the promoter addition represent the influence of each promoter compared to the catalyst with no promoter. Mg was found to have higher significance than Mn, which supports the results discussed above. Among the operating conditions, the most significant variable was the reaction temperature as expected. The group significance of the operating variables was found to be 76.5% while the preparation variables had 23.5% relative significance. This is quite reasonable since the operating variables (especially temperature) influenced the CO conversion more strongly compared to the addition of promoters. However, it should be noted that the relative significance of the variables calculated here are valid in their ranges

used in the experiments; otherwise, using different ranges may result different relative significances.

Table 5.4. Relative significances for Au/Al_2O_3 catalyst for selective CO oxidation.

	Variables	RMSE without Variable	RMSE Difference*	% Relative Significance (change of RMSE)	% Relative Significance (partial differentiation)	Group Significance (change of RMSE)
preparation variables	Mg	7.6	6.7	56.7		23.5
	Mn	6.0	5.1	43.3		
operating variables	react. T				84.3	76.5
	H_2O				7.4	
	CO_2				8.3	

*difference between "RMSE without the variable" - "RMSE of the original model (0.89)"

5.1.3. Modeling Promoted Pt/Al_2O_3 Catalysts for Water Gas Shift Activity by Modular Neural Networks

First, the optimal network topology that represents the experimental data (given in in Section 3.3) will be presented. Then, the effects of catalyst preparation and operating variables on CO conversion will be discussed in detail followed by the analysis of relative significances of these variables [50].

5.1.3.1. Determining the Optimal Neural Network Topology. In order to find the optimal neural network topology that may best represent the experimental data, nine modular neural networks with different topologies were constructed. The notation of "(a-b)-c" is used to label the neural networks, where "a" represents the number of the first hidden layer neurons connected to the preparation variables (promoter type), similarly "b" represents the first hidden layer neurons connected to the operating variables, c is the number of neurons in the second hidden layer. There are eight input variables and one output variable, which is the CO conversion.

The training errors were calculated by using the entire data set for training. On the other hand, the testing errors were calculated by applying 4-fold cross validation. Among the networks analyzed, the network with the topology of (2-3)-3 (Figure 5.14) was chosen

as the topology (testing RMSE=5.22) with the highest generalization ability, which is the indicator of network's ability to predict the data not seen before.
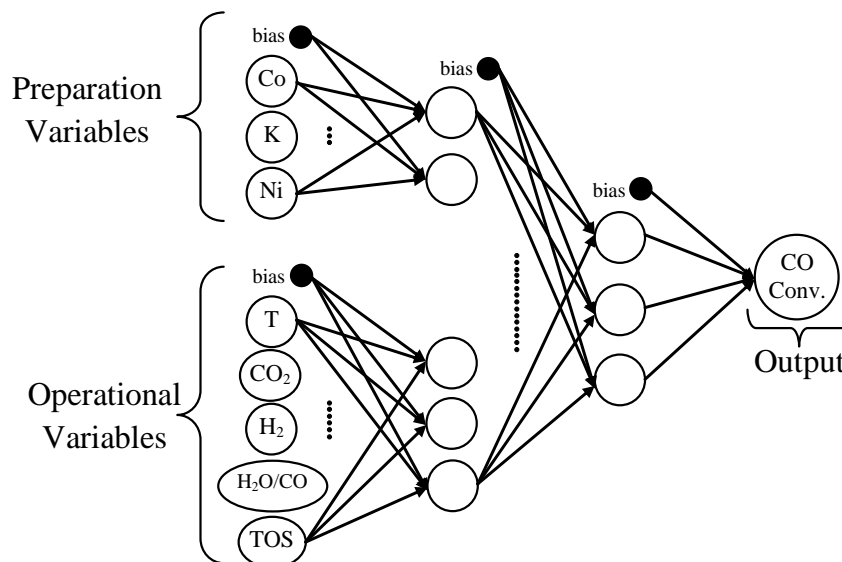
Figure 5.14. Optimal network topology for WGS reaction over Pt/Al_2O_3 [50].

This network was improved further by using different activation functions in the modules of the first hidden layer. All the nine combinations of logistic sigmoid function (lsf), hyberbolic tangent function (tanh) and identity function (ide) [115] were tested for the modules of catalyst preparation and operating variables while the logistic sigmoid function was kept in the second hidden layer (Figure 5.15).

It was found that the network with the topology (2-3)-3 was improved significantly with the use of identity function for the preparation variables and hyperbolic tangent function for the operating variables. The training and the testing RMSE of this network were calculated as 1.99 and 3.36 respectively. The experimental versus predicted CO conversion plots for both training and testing data are given in Figure 5.16. The data are distributed around the y=x lines with very high R^2 values for both plots indicating that the neural network model constructed is quite successful in predicting the data. Especially, the fitness of the testing plot (Figure 5.16b) is quite remarkable considering that each predicted CO conversion in this plot was made by the network which did not see that point during training. The optimal neural network topology proposed here was used for the remaining part of the analysis.
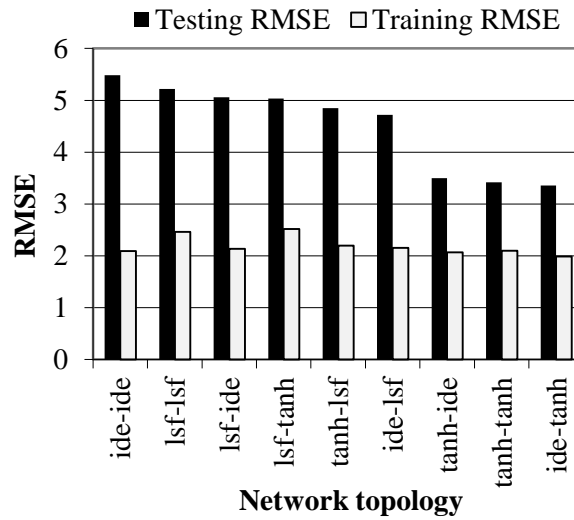
Figure 5.15. Testing and training errors for (2-3)-3 topology with different activation functions for the preparation and the operating variables [50].
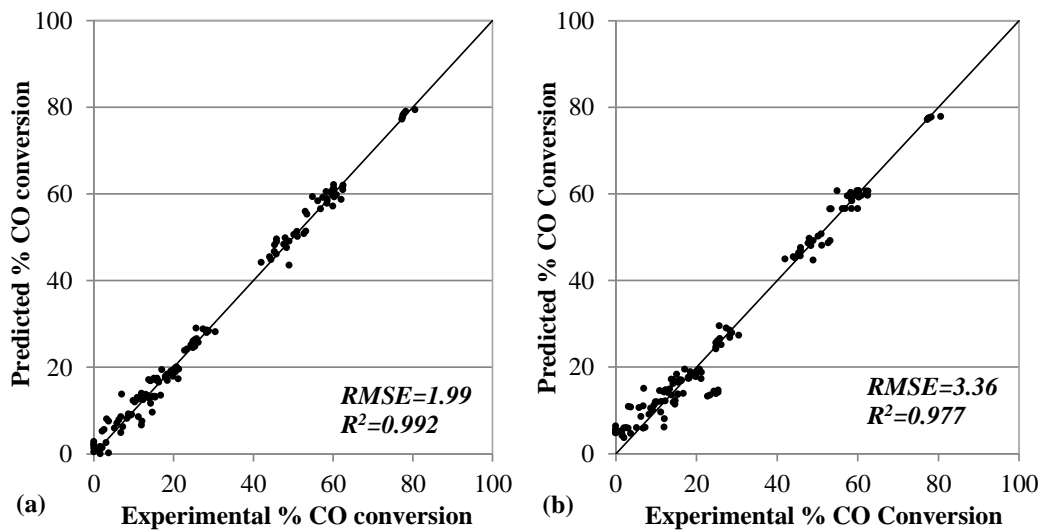
Figure 5.16. Experimental vs. predicted CO conversion values (a) training, (b) testing [50].

5.1.3.2. Analyzing the Effects of the Catalyst Variables. Figure 5.17 compares the predicted and the experimental CO conversions versus TOS for all the catalysts in the absence and presence of H_2 or/and CO_2. The predictions are generally very close to the experimental values, which is a further indicator for the success of the model. Although neural networks have superior ability in predicting the nonlinear effects of variables, the model predictions in this case are almost linear because the experimental CO conversions are not significantly influenced by the change of TOS. The prediction of slight nonlinearity could be improved by employing a larger network topology; however the use of the

smallest network capturing the general trends was preferred to keep the generalization ability higher. It is quite apparent from Figure 5.17 that the presence of H_2 decreases CO conversion significantly for all the catalysts due to the reverse WGS reaction, which consumes CO_2 and H_2 producing CO and H_2O. For the same reason, the presence of CO_2 also decreases the WGS activity though its effect seems to be smaller in the present case, actually due to the small feed concentration of this component compared to hydrogen. When both CO_2 and H_2 are present in the reaction stream (realistic stream composition exiting a reformer), the catalytic activity reduces even more. It is also apparent from Figure 5.17 that the activities of all catalysts slightly decrease with increasing TOS which may be an evidence for possible deactivation.

In the absence of CO_2 and H_2, the catalyst with the best performance is the Pt-Ce catalyst (without any second promoter), which was proposed as a very good catalyst for WGS reaction due to the high oxygen storage capacity of ceria and its ability of stabilizing the support and the noble metal [6]. On the other hand, the addition of any of the promoters Co, K, Ni had a detrimental effect in the absence of CO_2 and H_2. However, a good water gas shift reaction catalyst has to work in the presence of significant levels of CO_2 and especially H_2 [28, 151]. Hence, all the catalysts were also tested in the presence of these components, and it was found that the Pt-Ce-K catalyst was less affected from CO_2 and/or H_2 in contrast to the Pt-Ce-Co and Pt-Ce-Ni catalysts. When only CO_2 was present in the reaction stream, Pt-Ce-K had significantly higher activity even compared to Pt-Ce. The performance of the Pt-Ce-K catalyst was almost the same (slightly better) as the Pt-Ce catalyst when both CO_2 and H_2 were present. The Pt-Ce-K catalyst can be analyzed further in future studies, since it may be superior to the Pt-Ce catalyst under some other operating conditions.

The predicted and the experimental performances of Pt-Ce, Pt-Co-Ce and Pt-K-Ce catalysts at different temperatures under realistic conditions are given in Figure 5.18, which shows that decreasing the reaction temperature has a negative effect on CO conversion. This is expected because, although WGS is an exothermic reaction (equilibrium constant increases with decreasing temperature), it also becomes kinetically controlled at lower temperatures which prevents the approach to thermodynamic limits

[219, 220]. The effects of temperature on all the catalysts are quite successfully predicted by the neural network model.
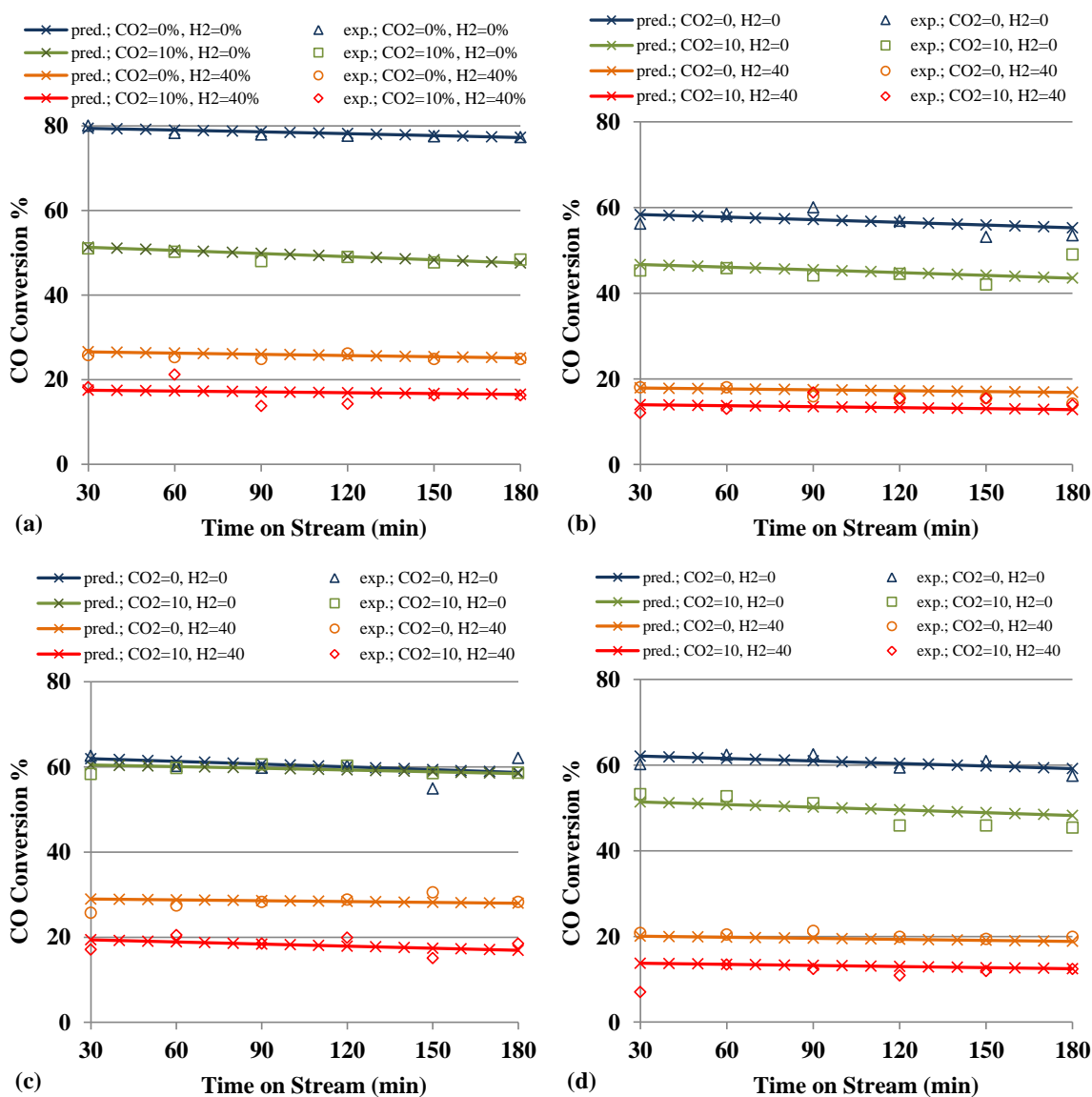
Figure 5.17. CO conversion versus TOS at 300°C in various reaction streams for: (a) Pt-Ce, (b) Pt-Co-Ce, (c) Pt-K-Ce, (d) Pt-Ni-Ce [50].

Finally, the effect of H_2O/CO ratio in the feed on CO conversion over the Pt-Ce catalyst at 275°C is shown in Figure 5.19 as a sample case. The increase in the H_2O/CO ratio positively affects the catalytic activity as also reported by various previous studies [221, 222]. The promotion of WGS is due to the shifting of reaction equilibrium to the product side by the presence of excess water as also successfully predicted by the model.
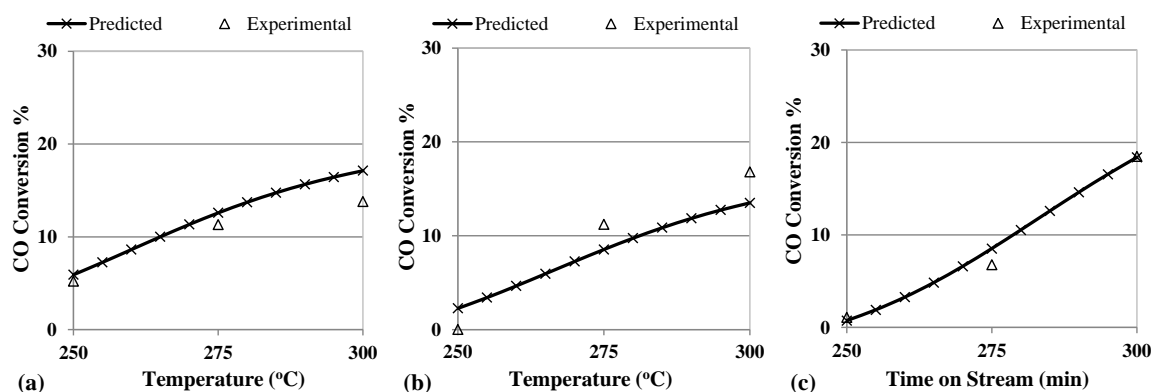
Figure 5.18. CO conversion versus temperature in the presence of 10% CO_2 and 40% H_2 at 90 min of TOS for: (a) Pt-Ce, (b) Pt-Co-Ce, (c) Pt-K-Ce [50].
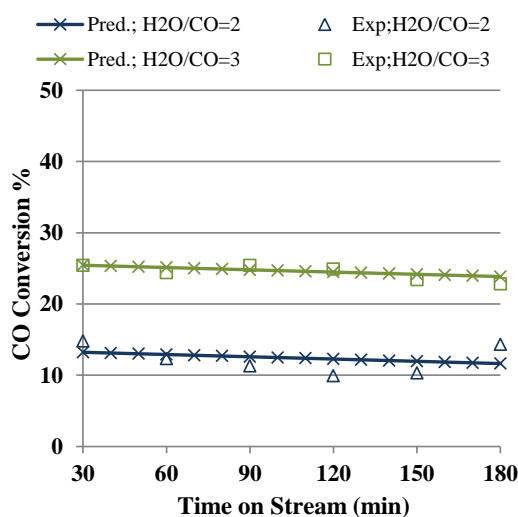
Figure 5.19. Effect of H_2O/CO on catalyst activity in the presence of 10% CO_2 and 40% H_2 at 275°C for Pt-Ce [50].

Normally, the neural network can be optimized to find the input variables maximizing CO conversion. However, it was not necessary in the present case because it is evident that higher temperatures and H_2O/CO ratios would increase CO conversion. On the other hand, the lower values of these variables are more desirable for the heat and water management in fuel cell systems. Hence, the optimization of the neural network was omitted for this case.

5.1.3.3. Analyzing the Input Significance. Table 5.5 shows the relative significances of the catalyst preparation and the operating variables. It is to be noted that these significances

should be treated within the range of conditions studied; otherwise, they may produce misleading conclusions. For example, the relative significance of the temperature indicates its relative influence over CO conversion between the lowest to highest values selected in the study; if the reaction temperature were set to very low values (like room temperature), no CO conversion would be obtained and the temperature would have a much higher relative significance. This is also true for all the other variables.

Table 5.5. Relative significances for Pt/Al$_2$O$_3$ catalyst for WGS reaction.

	Variables	RMSE without the Variable	RMSE Difference*	% Relative Significance (change of RMSE method)	% Relative Significance (partial differentiation)	Group Significance (change of RMSE method)
preparation variables	Co	4.11	2.13	40.0		
	K	3.79	1.81	34.0		18.3
	Ni	3.37	1.38	26.0		
operating variables	React T				9.4	
	CO$_2$				10.7	
	H$_2$				43.7	81.7
	H$_2$O/CO				33.3	
	TOS				2.9	

*difference between "RMSE without the variable" - "RMSE of the original model (1.99)"

Under the conditions studied, the group significance of the operating variables was found to be 81.7% while the second promoter addition was 18.3%. This is quite reasonable since the Pt-Ce catalyst has a certain level of performance, and the second promoter has a minor impact. On the other hand, the chances of temperature and feed composition to influence CO conversion are significant. Among the operating variables, H$_2$% was the most significant because it increased reverse WGS reaction decreasing the CO conversion severely. The presence of CO$_2$ had the same effect with lesser degrees due to its smaller percentage in the feed.

The H$_2$O/CO ratio was the second most significant operating variable; increasing this ratio was expected to increase the water gas shift catalytic activity considerably [221, 222]. The influence of TOS was not found significant because it had only some minor impact on CO conversion due a very slight activity loss. This is also apparent in Figure 5.17, which shows that CO conversions do not decrease significantly even after 3 hours of operation for all the catalysts. The low relative significance value (9.4%) for the operation temperature

can be attributed to the fact that experimental analysis was conducted in a narrow temperature range compatible with actual operating conditions used in fuel processors.

5.2. Knowledge Extraction from DFT Generated Data

5.2.1. Structure and Activity Relationship for CO and O_2 Adsorption over Gold Nanoparticles Using DFT and Artificial Neural Networks

The data generated by Davran-Candan [141] for CO and O_2 adsorption over various gold nanoparticles was organized and modeled as follows (Figure 5.20):

- CO and O_2 adsorption energies for 97 possible sites over 27 clusters to be used as the response variables for the models.
- The values of user defined descriptors (such as particle size, unpaired electron, charge and coordination number) as the input variables considering that the adsorption energies will depend on them.
- Some structural properties of clusters (such as HOMO-LUMO gap, binding energy and ionization potential) as the intermediate variables that can be predicted from the input variables and to be used to predict the adsorption energies.

The neural network models for the descriptors versus adsorption energies were constructed first to see the general picture. All the descriptors (size, charge, coordination number and unpaired electron) were used for this purpose. However, the linear dependence among these variables was checked first to avoid multicolinearity [56]. No pair of correlated variables was used together in the same network. Then, the neural network models for the descriptors versus the structural properties (HOMO-LUMO gap, binding energy per atom, ionization potential and electron affinity) were also built similar to the modeling of adsorption energies. Finally, the adsorption energies were tried to be modeled using the structural properties to see their relative significances for the CO and O_2 the adsorption. Each of the descriptors was treated according to its characteristics when the neural networks were constructed. The size, for instance, is taken as a continuous variable while the coordination number was treated as categorical, since it is mainly used to distinguish the position of the adsorbed atom on the cluster. The charge, which has only

three distinct values and unpaired electron with two possible states were also used as categorical inputs. All of the structural properties, on the other hand, were taken as continuous variables when they were used as the inputs of the networks.
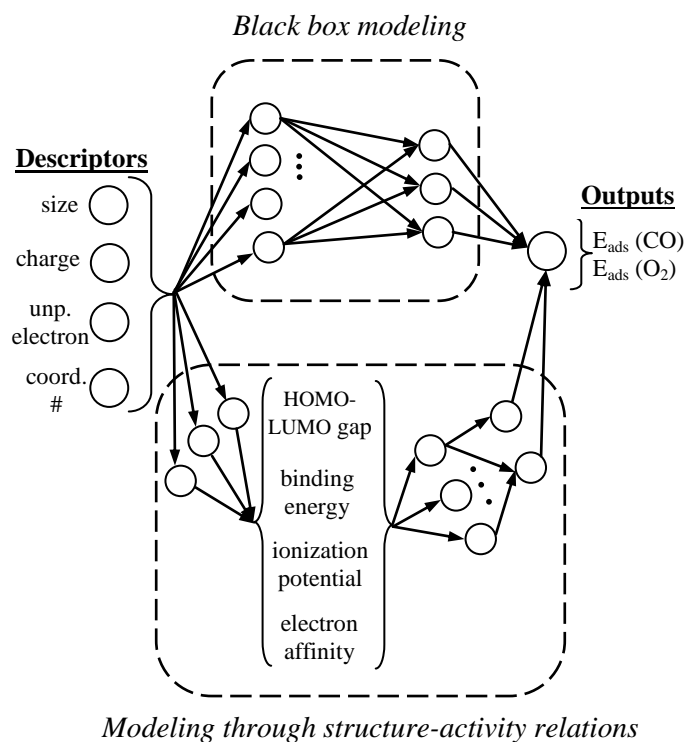
Figure 5.20. Conceptual approach to model CO and O_2 adsorptions over Au nanoparticles.

The results of the CO adsorption models are presented in full detail in the following sections, and then the O_2 adsorption work is summarized to avoid repetition. The structure-activity relationship was established using the simplified model shown in Figure 5.20 based on the assumption that any changes in the user defined descriptors influence the adsorption energies by changing some structural properties of the catalyst or catalyst-reactant complex [52].

5.2.1.1. Descriptors Versus CO Adsorption Energy. In order to find the optimal neural network that represents the CO adsorption data, various neural networks having different structures and neuron numbers were constructed with the descriptors of size, charge, unpaired electron and coordination number of the Au atom to which CO was bonded while the CO adsorption energy was the output.

The statistical analysis for training and validation was performed and presented in Figure 5.21 with the order of increasing number of weights. The "a-b-c-d" notation was used to label the neural networks, where "a" represents the number of input variables (4 in this case), "b" and "c" represent the number of neurons in the first and second hidden layers, respectively, and d (1 in this case) is the output number. The numbers in parentheses on the x axis represents the number of weights (connections) of the networks.
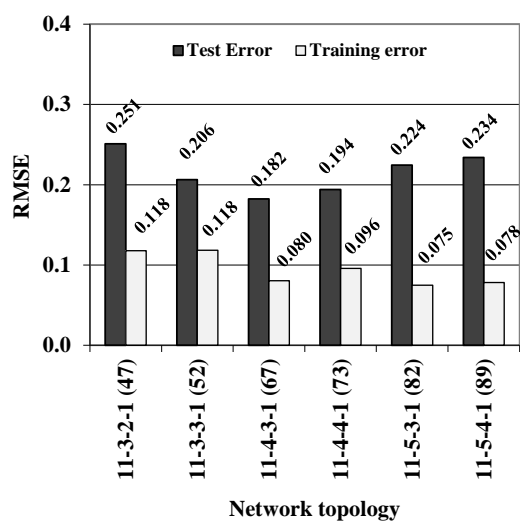
Figure 5.21. Training and testing RMSE values of various neural networks modeling the CO adsorption energy from the user defined descriptors [52].

The grey bars in Figure 5.21 represent the training error, which is calculated by using the entire data set. The black bars, on the other hand, show the testing error (indicating generalization accuracy), which was obtained by employing 4-fold cross validation. The network structure 11-4-3-1 has the smallest RMSE of testing (0.182) and considerably low RMSE of training (0.080) when compared with the other networks. Therefore, it was used as the optimal network to model the CO adsorption energy from the descriptors. The schematic representation of this network is shown in Figure 5.22, with black circles as the bias neurons having a constant value of 1. Since the number of connections among the layers is very large, only some representative lines were drawn in the figure for simplicity while the exact number of neurons was used.
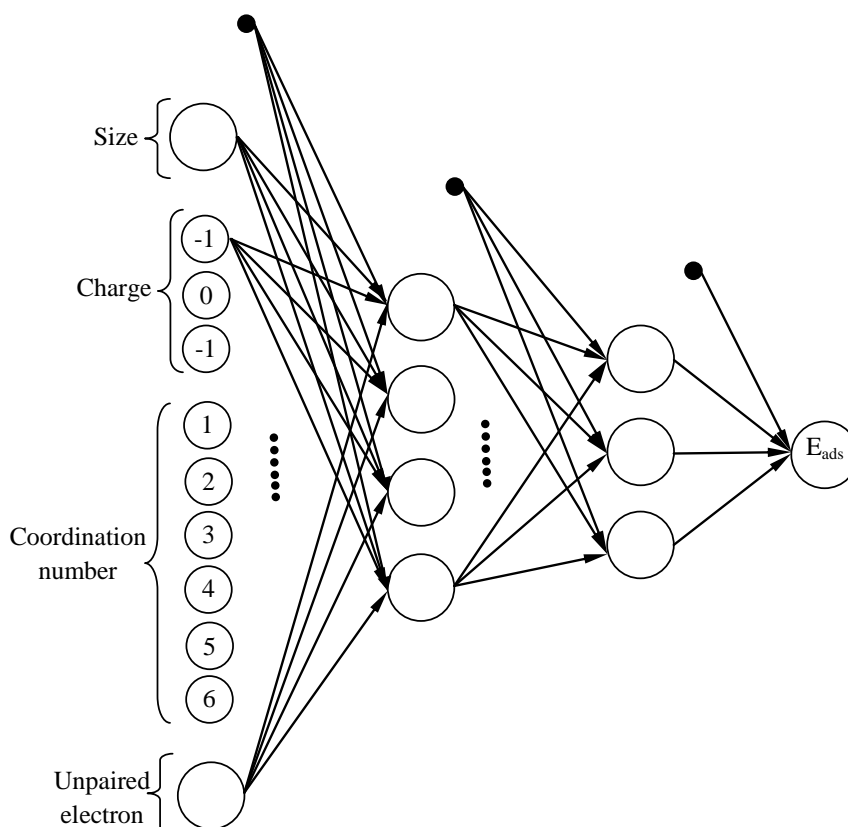
Figure 5.22. Optimal neural network topology for CO adsorption energy from user defined descriptors [52].

In Figure 5.23a, the neural network predicted CO adsorption energies for the training data were plotted against DFT calculated values. The data points lie quite close to y=x line, indicating that the model is quite successful to represent the CO adsorption data. This is also supported by the R^2 value of 0.956, which is quite high. Besides, the statistical results of the 4-fold cross validation analysis, which indicates the prediction power of the neural network for the data not included in the training set (not seen by the network before), is also quite satisfactory with an R^2 value of 0.764 (Figure 5.23b). In contrast to the neural network training R^2 value of 0.956 and the RMSE value of 0.080, the same data modeled by full quadratic multiple linear regression provided a training R^2 value of 0.653 and a RMSE value of 0.225, indicating the superiority of neural networks against multiple linear regression.
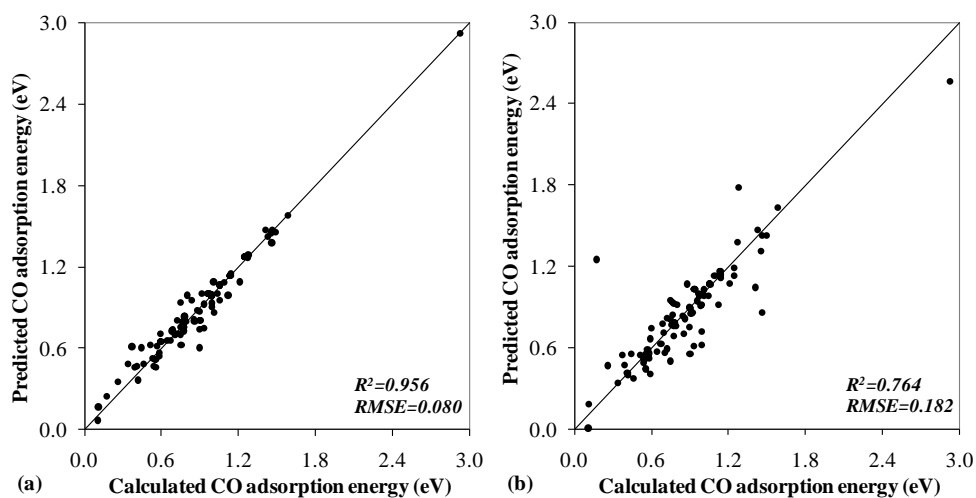
Figure 5.23. Predicted (from descriptors) vs. calculated CO adsorption energy; (a) training data, (b) test data [52].

The relative significance of each variable was calculated using the change of RMSE method and the results are presented in Table 5.6. The charge of the cluster seems to be the most significant variable for CO adsorption energy followed by the coordination number and the size of the cluster. The unpaired electron has relatively less significance.

Table 5.6. Relative significances of descriptors for the CO adsorption.

Descriptors	RMSE without the Variable	RMSE Difference*	Relative Significance
size	0.151	0.071	20.4
charge	0.258	0.177	51.1
coordination number	0.162	0.082	23.6
unpaired electron	0.097	0.017	4.9

*the difference between "RMSE without the variable" - "RMSE of the original model (0.080)"

The plots of the maximum adsorption energies versus the cluster size are given in Figure 5.24, indicating a remarkable agreement between the quantum mechanically computed and neural network predicted energies. Both our DFT calculations and previous findings [97] indicated that the adsorption strength of CO generally followed the order: cationic > neutral > anionic, which may be explained with the electron donating nature of CO during adsorption [223]. No clear correlation between the size and CO adsorption strength was found for the neutral and anionic clusters, while a decreasing trend with increasing cluster size was observed for the cationic clusters. This may be attributed to the

fact that the charge is diluted as the number of atoms involved in the cluster got higher, similar to the previous findings in the literature [97].
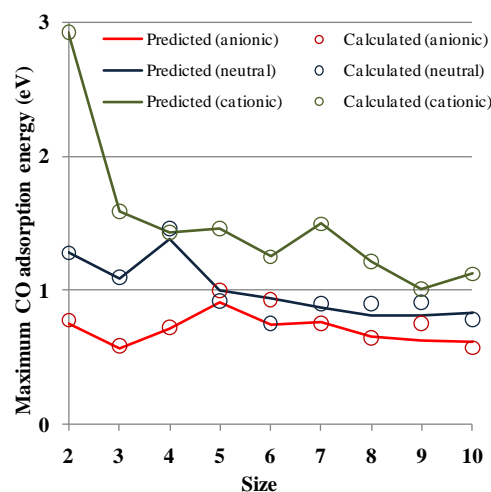
Figure 5.24. Effect of size on maximum CO adsorption energy [52].
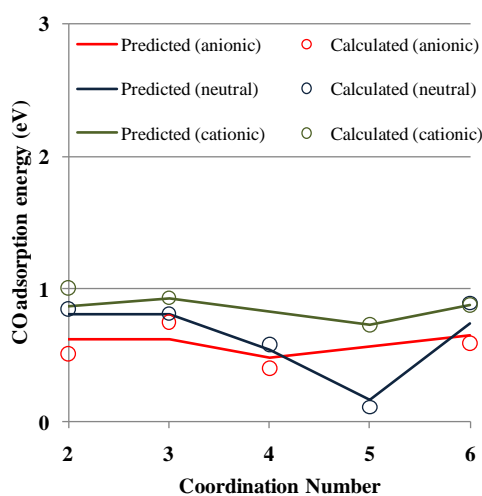
Figure 5.25. Effect of coordination number on CO adsorption energy for n=9 [52].

The effect of the coordination number on the adsorption energy is presented in Figure 5.25 for the cluster size of 9 as an example. Although the presence of low-coordinated Au atoms was reported to be crucial in determination of the adsorption strength [106, 107], the decreasing trend of adsorption energy with increasing coordination number was not so evident in our DFT results and therefore in our model predictions. This is probably due to the fact that the Au sites involved in this study are already under-

coordinated (max 6), so that the effect of the coordination number was not so deterministic on CO adsorption strength.

Finally, the presence of the unpaired electron was found to have relatively less effect on the CO adsorption energy, which is expected since there is no report in the literature showing such a correlation as far as we know.

5.2.1.2. Descriptors Versus Structural Properties of Clusters. Then, the structural properties of the clusters were modeled as a function of descriptors. All four structural properties were predicted successfully as it can be seen from the plots of neural network predicted versus quantum mechanically calculated properties and their high training R^2 values in Figure 5.26. R^2 values for validation (predicting the data not included to training set) were also quite high as 0.806, 0.900, 0.926 and 0.931 for the HOMO-LUMO gap, the binding energy, the ionization potential and the electron affinity, respectively.
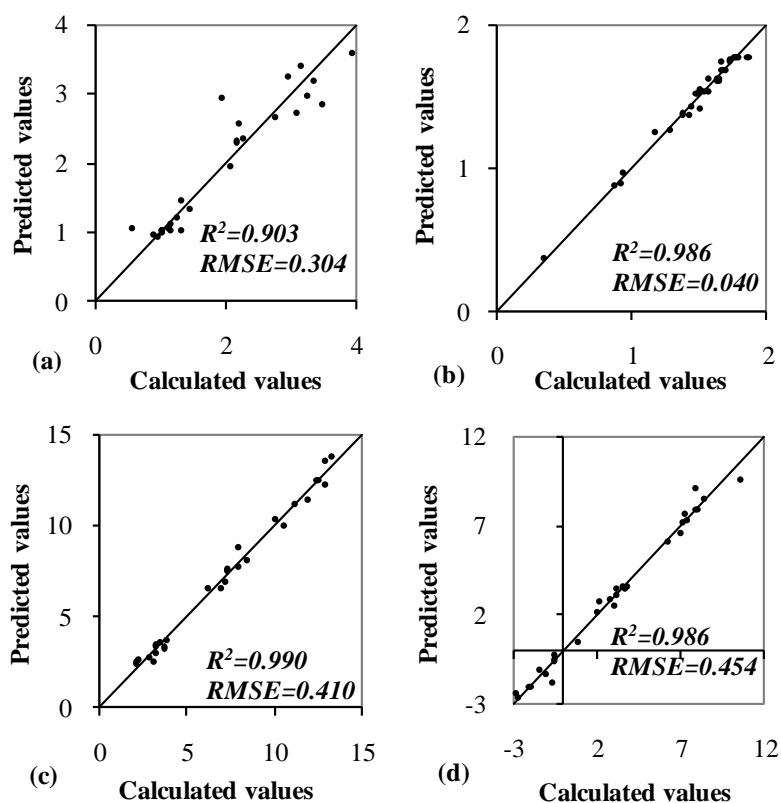
Figure 5.26. Predicted (from descriptors) vs. calculated structural properties using training data. (a) HOMO-LUMO gap (eV), (b) binding energy (eV/atom), (c) ionization potential (eV), (d) electron affinity (eV) [52].

The significance analysis (Table 5.7) indicates that the size of the cluster was found to be the most important variable for the binding energy, which generally increases with increasing cluster size in agreement with the literature (Figure 5.27) [108, 111]. This may be explained as follows: As the cluster becomes larger, the atomic interactions increase due to the increase in the number of neighboring atoms [111]. The effect of charge was also remarkable. The strength of the binding energy generally followed the order: cationic > anionic > neutral, similar to the results obtained by Fernandez *et al.* [224]. The unpaired electron, on the other hand, had relatively small impact on the binding energy as evident from the weak even-odd oscillations (Figure 5.27), which was also observed in some previous studies [108, 111].

Table 5.7. Relative significances of descriptors for the structural properties.

Descriptors	Relative Significances			
	HOMO-LUMO Gap	Binding Energy	Ionization Potential	Electron Affinity
size	7.2	63.2	14.4	9.6
charge	14.2	28.2	79.7	84.3
unpaired electron	78.6	8.6	5.9	6.1
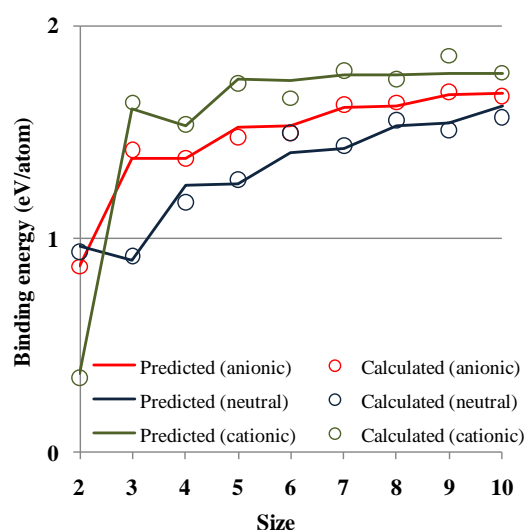
Figure 5.27. Effect of cluster size on binding energy [52].

For the HOMO-LUMO gap, on the other hand, the most significant variable was found to be the unpaired electron while the charge and the size of the cluster had relatively small effects. When the HOMO-LUMO gap was plotted as a function of size, the even-odd oscillation pattern (Figure 5.28) was evident similar to the previous findings [52, 97, 110].

The clusters with even number of electrons have higher HOMO-LUMO gaps compared to the neighboring clusters having odd-number of electrons, which was attributed to the electron pairing effect [110].
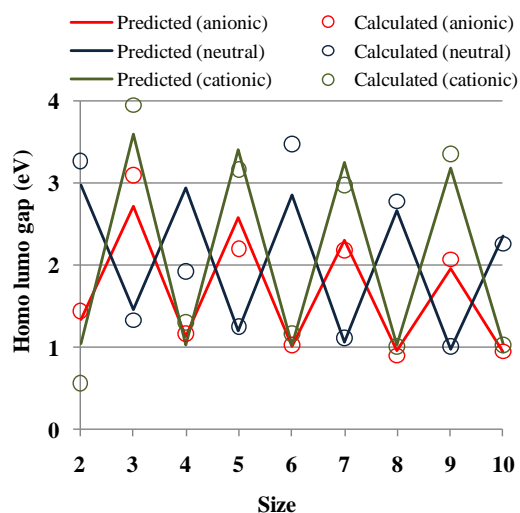
Figure 5.28. Effect of cluster size on HOMO-LUMO gap [52].

The agreement between the neural network predicted and the quantum mechanically computed values of the ionization potential and the electron affinity was much higher with the R^2 values of 0.990 and 0.986, respectively (Figures 5.26c and 5.26d). The analysis of significance reveals that the charge of the cluster is the most dominating factor for both the ionization potential and the electron affinity, of which their values followed the order: cationic > neutral > anionic. This is an expected result for the ionization potential because it is obviously more difficult to remove an electron from a cationic cluster compared to neutral and anionic ones. The electron affinity follows the same trend for similar reasons. This is clearly seen from the plots in Figure 5.29 and Figure 5.30, which also show the size dependencies of two properties. On the other hand, even though the relative significance of the size is found to be relatively small compared to the charge, its effect on the ionization potential and the electron affinity is still significant for cationic and anionic clusters.

Finally, despite the low significance of the unpaired electron, ionization potential was generally higher for the clusters having even number of electrons (when compared to the neighboring clusters with odd number of electrons) while electron affinity was generally higher for the clusters having odd number of electrons. This is due to the fact that

it is more difficult to remove an electron from the valence shell of a cluster having even number of electrons.
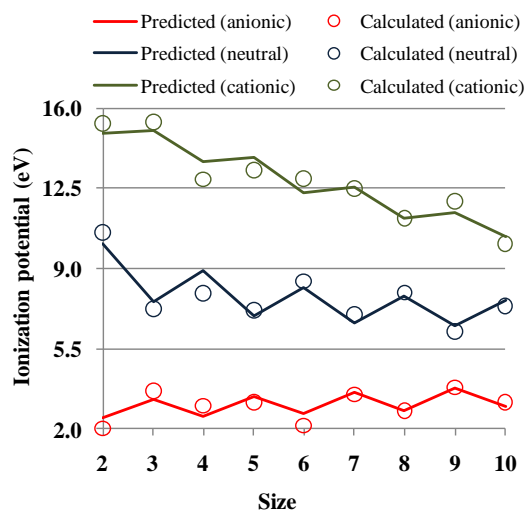
Figure 5.29. Effect of cluster size on ionization potential [52].
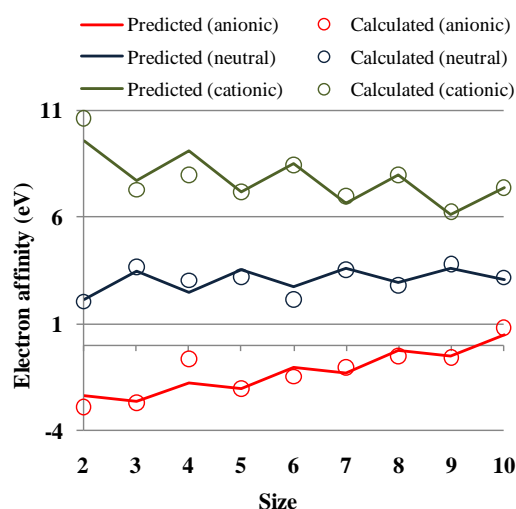
Figure 5.30. Effect of cluster size on electron affinity [52].

5.2.1.3. Structural Properties Versus CO Adsorption Energy. As the last step for CO adsorption, various neural networks were constructed to model adsorption energy from the structural properties and the optimal network topology was chosen. The coordination number was also added as an input list to identify the location of the adsorption although it is not an actual structural property. The ionization potential and the electron affinity were not used in the same network because they are not independent. The best network results

were obtained if the ionization potential was used together with the binding energy, HOMO-LUMO gap and coordination number. The predicted versus calculated adsorption energies shown in Figure 5.31 was impressive with R^2 values of 0.962 and 0.806 for the training and the validation data, respectively. The relative significances of the properties are given in Table 5.8.
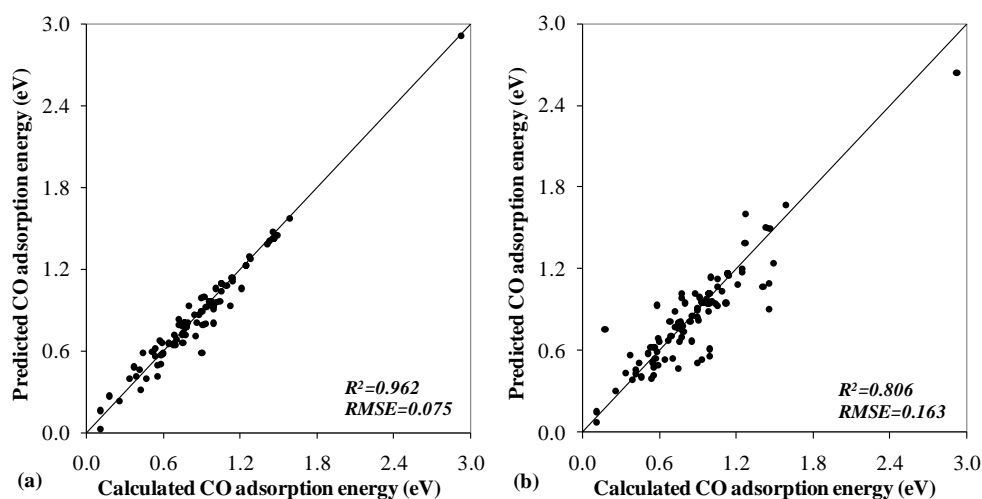
Figure 5.31. Predicted (from structural properties) vs. calculated CO adsorption energy; (a) training data, (b) test data [52].

Table 5.8. Relative significances of structural properties for CO adsorption.

Structural Properties	RMSE without the Variable	RMSE Difference*	Relative Significance
HOMO-LUMO gap	0.131	0.056	21.7
binding energy	0.105	0.030	11.6
ionization potential	0.146	0.071	27.4
coordination number	0.176	0.102	39.3

*the difference between "RMSE without the variable" - "RMSE of the original model (0.075)"

Replacing ionization potential with electron affinity produced almost the same result with the training R^2 of 0.951 as expected since these two variables are not independent. It is interesting to note that the removal of the ionization potential (or electron affinity) from the model (using only HOMO-LUMO gap, binding energy and coordination number) does not worsen the results much either as evidence from the training R^2 value of 0.904. Instead, the contributions of the remaining three variables increase; resulting a quite good agreement between model prediction and quantum mechanically computed CO adsorption

energies. Apparently, the HOMO-LUMO gap and the coordination number together with any one of the remaining three (binding energy, electron affinity or ionization potential) carry the significant part of the necessary information to predict the CO adsorption energy although the nature of this information is not clear at this stage.

5.2.1.4. Descriptors Versus O_2 Adsorption Energy. A similar analysis was also performed for O_2 adsorption energy since most of the mechanisms proposed for the CO oxidation over Au based catalysts involve the adsorption of O_2 as well [100, 225]. The details of the analysis are quite similar to CO adsorption case; hence, they will not be given here to avoid repetition. Instead, the results will be summarized with the discussion of the issues unique to O_2 adsorption.

The most significant difference between the O_2 and CO adsroption data is the fact that adsorption of O_2 over some sites is endothermic. Although these results are not physically meaningful, the degree of endothermicity is an indicator of instability of the adsorption complex. Hence, these data points were not excluded in the models because they still carry some information even though they are not useful by themselves. The positive sign was used for the stable adsorption sites while the negative values represent the unstable complexes during the model construction.

The optimal network topology for the O_2 adsorption energy was searched, and it was found to be the same as the one used for the CO adsorption. The plot of quantum mechanically calculated versus predicted O_2 adsorption is given in Figure 5.32. The R^2 values of 0.967 and 0.873 for the training and the testing are quite good.

The input significance analysis indicated that the unpaired electron is the most significant descriptor affecting the O_2 adsorption energy (Table 5.9) in contrast to the case of the CO adsorption (Table 5.6). This makes sense, since it is already known that the existence of the unpaired electron has a deterministic effect on O_2 adsorption stemming from the interaction mechanism of O_2 with Au clusters. The adsorption mechanism for anionic and neutral clusters was argued to take place through electron transfer from Au to the π^* orbital of O_2 [110, 226]. For anionic clusters, this transfer is accompanied by the activation of O-O bond to a superoxo state through stretching of the bond [100]. No such

activation of O_2 was reported for neutral and cationic clusters, even though electron transfer from Au to O_2 was observed to some extent for neutral clusters [100, 226]. Hence, the electron transfer required for the adsorption of O_2 is easily achieved by the clusters having unpaired electrons and adsorption of O_2 is stronger for clusters having odd number of electrons. Parallel to these findings, we found that Au clusters having odd number of electrons (whatever the total charge is) adsorbed O_2, while clusters with even number of electrons showed no activity towards O_2.
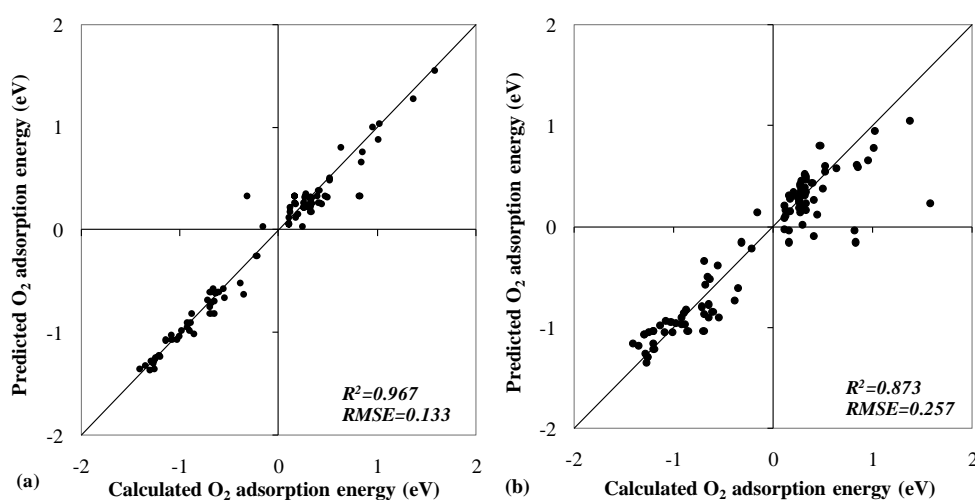

Figure 5.32. Predicted (from descriptors) vs. calculated O_2 adsorption energy; (a) training data, (b) test data [52].

Table 5.9. Relative significances of descriptors for O_2 adsorption.

Descriptors	RMSE without the Variable	RMSE Difference*	Relative Significance
size	0.176	0.043	10.0
charge	0.231	0.097	22.5
coordination number	0.166	0.033	7.5
unpaired electron	0.393	0.260	60.0

*the difference between "RMSE without the variable" - "RMSE of the original model (0.133)"

The input significance analysis showed that the effect of charge was also significant, which has already been shown by various groups [100, 227]. Although early experimental findings indicated that neutral and cationic clusters showed no activity towards O_2 molecule except for Au_{10}^+ cluster [227], the general consensus reached in the field is that both neutral and cationic clusters can adsorb O_2 but adsorption strength is the greatest for

the anionic ones [99, 100], which is also supported by our calculations. For the anionic and neutral clusters, the adsorption of O_2 was accompanied with the electron transfer from Au to O_2 molecule, in agreement with the previous findings [100, 226]. For the cationic clusters, however, an electron transfer in the other direction (from O_2 to Au) was observed, which may be explained with the high electron affinity values calculated for cationic clusters (Figure 5.30). The amount of charge transferred from O_2 to Au decreased with increasing cluster size, parallel to the slight decrease in electron affinity of cationic clusters with increasing size (Figure 5.30) and at n=10, the transfer changes direction and electron transfer from Au to O_2 is observed.

Finally, although the adsorption strength of O_2 was argued to increase with decreasing coordination number [106], this effect was found to be small compared to the other descriptors in our input significance analysis as in the case of CO adsorption probably for the same reason (absence of high coordinated Au atoms in our clusters).
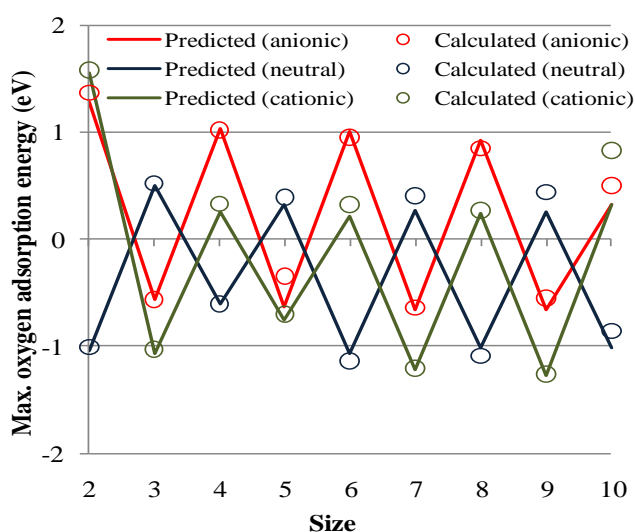
Figure 5.33. Effect of cluster size on maximum O_2 adsorption energy [52].

The effects of cluster size, charge and unpaired electron can be seen together in the plots in Figure 5.33, which shows the success of the model clearly. The significance of the unpaired electron is reflected itself as the even-odd alternating pattern in energy as opposite to the CO adsorption energy.

5.2.1.5. Structural Properties Versus O_2 Adsorption Energy. Then, the O_2 adsorption energies were modeled using the same structural properties (HOMO-LUMO gap, binding energy, ionization potential together with the coordination number) as the input set. The plots of calculated versus predicted O_2 adsorption energies are given in Figure 5.34 with the R^2 values of 0.963 and 0.827 for the training and the validation respectively. The input significance analysis reveals that the HOMO-LUMO gap and the ionization potential are the most effective variables in determination of the adsorption strength of O_2 (Table 5.10). It was observed that O_2 adsorption was stronger for the clusters having low ionization potentials, which can be attributed to the interaction mechanism of O_2 with the cluster as explained in the previous section. The effect of HOMO-LUMO gap, on the other hand, seems to be through the unpaired electron, since this descriptor was found to be significant for both HOMO-LUMO gap and adsorption energy of O_2.
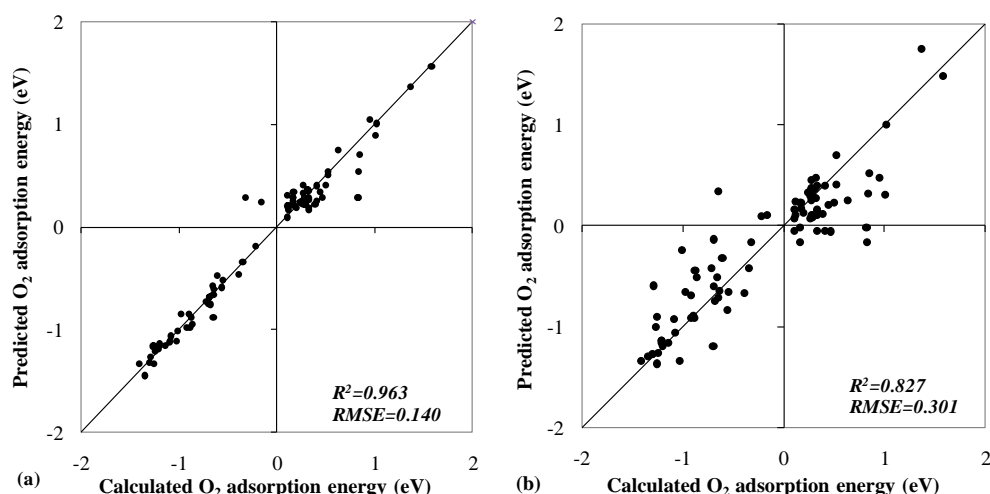
Figure 5.34. Predicted (from structural properties) vs. calculated O_2 adsorption energy; (a) training data, (b) test data [52].

Table 5.10. Relative significances of structural properties for O_2 adsorption.

Structural Properties	RMSE without the Variable	RMSE Difference*	Relative Significance
HOMO-LUMO gap	0.204	0.063	36.6
binding energy	0.150	0.010	5.7
ionization potential	0.215	0.074	43.0
coordination number	0.166	0.025	14.7

*the difference between "RMSE without the variable" - "RMSE of the original model (0.140)"

When the ionization potential was replaced by the electron affinity, the training R^2 was almost the same (0.961) indicating two variables can be interchanged as in the case of CO adsorption. However, removing the ionization potential (or electron affinity) from the input set and using only the HOMO-LUMO gap, binding energy and coordination number lowered the training R^2 value to only 0.920, which is still meaningful.

5.2.2. Analysis of O_2 Adsorption Stability and Strength over Gold Nanoparticles Using DFT and Logistic Regression

As a complementary study to artificial neural network modeling of the DFT data, multiple logistic regression modeling was also applied to understand the role of user defined descriptors and DFT computed structural properties on the stability and strength of O_2. The results are presented and discussed in two parts. In part one, the data set was divided into two categories, namely "stable" and "unstable" adsorption depending on the sign of adsorption energies, where the computed positive adsorption energies indicated instability. In part two, the stable adsorption values were analyzed in terms of strength, dividing the data set as "strong" (if the value of the adsorption energy was smaller than -0.5 eV) and "weak" (if the value of the adsorption energy was higher than -0.5 eV). It should be noted that the -0.5 eV border separating strong and weak adsorption has no physical significance. It was selected by trial and error such that the data points can be classified effectively based on the values of the descriptors and the structural properties, which is the main purpose of the work [228].

5.2.2.1. Descriptors Versus O_2 Adsorption Stability. The stability analysis was performed by using both the descriptors and the structural properties. However, the "stable" or "unstable" classification based on the descriptors can be treated as the demonstration and verification of our approach, otherwise one needs no special tool to make this classification. It is very clear from the visual inspection of the data that the stability of adsorption depends on the absence or the presence of the unpaired electron. If an unpaired electron was present in the cluster, the adsorption was found to be stable except for a few cases. However, the implementation of the logistic regression for this "known" problem seemed to form a good demonstration. Hence, we first discussed this model and then continued with the analysis of stability through the structural properties.

The descriptors of the clusters were used as the predictor variables of the probability of the stable adsorption, and the results of the logistic regression model developed are presented in Table 5.11. Since the G value (106.4) of the model is much greater than the critical chi-square value of 9.49 (for four degrees of freedom), the null hypothesis that all the coefficients are zero (meaning no significant pattern or correlation) was rejected and the model was concluded as being significant. As indicated from the p values in the table, the stability of O_2 adsorption depends only on the presence of unpaired electron. This is also evident from the upper and lower bounds of the regression coefficients (β). Except unpaired electron, the lower and upper limits of the regression coefficients for all the descriptors have values with opposite signs, which indicate that their coefficients can have the value of zero within 95% confidence interval. Hence, these variables cannot be deterministic for the stable-unstable classification. On the other hand, lower (1.89) and upper (7.04) limits of the regression coefficient for unpaired electron have the same sign (no possibility to have the value of zero), indicating that it is a significant variable. The positive value of the regression coefficient also indicates that the presence of unpaired electron affects the stability positively, i.e. clusters having odd number of electrons adsorb O_2, while clusters with even number of electrons show small or zero reactivity towards O_2, parallel to the previous findings [52, 100]. This result clearly demonstrates that logistic regression modeling can be easily and effectively employed in this subject.

Table 5.11. Statistics of descriptors versus O_2 adsorption stability.

Descriptors	β (regression coefficient)	SE(β) (standard error of regression coefficient)	Z_{Wald} (β /SE(β))	p value of Z_{Wald}	β_{lower}	β_{upper}
constant term	0.65	2.47	0.26	0.793	-4.19	5.49
size	-0.13	0.36	-0.36	0.722	-0.83	0.57
charge	-0.57	0.85	-0.67	0.503	-2.24	1.10
coordination number	-0.30	0.45	-0.66	0.510	-1.18	0.59
unpaired electron	4.46	1.31	3.40	0.001	1.89	7.04
$G = 106.4$			$\chi^2(0.95, 4) = 9.49$			

The logistic curve obtained for the unpaired electron versus the probability of having stable adsorption is shown in Figure 5.35. The red symbols indicate the status of stability for the quantum mechanically computed values; they have 100% or 0% probabilities,

because O_2 either adsorb or do not adsorb on a given cluster. The blue symbols, on the other hand, show the corresponding model predictions, which are quite close to the computed data. Since x axis has only two values (two categories), the separation between the data points in the same category has no physical meaning; they were just created to prevent overlapping and make the visual inspection easier.

The test of extrapolation ability of the model was not necessary in this section because the size was found to be not significant as a variable. The model can predict the adsorption stability of any cluster if the absence or presence of unpaired electron, which is the only variable determining the stability, is known.
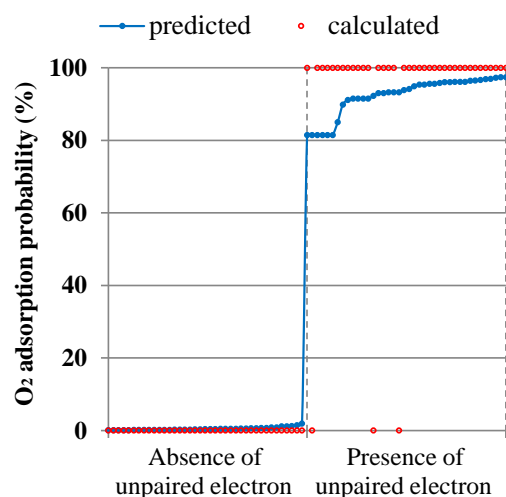
Figure 5.35. Effect of unpaired electron on O_2 adsorption probability [228].

5.2.2.2. Structural Properties Versus O_2 Adsorption. Table 5.12 shows the results of the logistic regression when the structural properties were used as the predictor variables against the stability of adsorption. Since the ionization potential and the electron affinity were not linearly independent, only the ionization potential was used here as one of the structural properties. On the other hand, the coordination number was also added as a predictor variable to identify the location of the adsorption, although it is not an actual structural property of the cluster. Like the previous case, the G value (97.4) of the model is much greater than the critical chi-square (9.49) indicating that the model is significant. The Wald test demonstrates that all the structural properties had p values quite higher than 0.05, except the HOMO-LUMO gap. This means that among the structural properties

considered, HOMO-LUMO gap is the only structural property effective on the stability of O_2 adsorption.

Table 5.12. Statistics of structural properties versus O_2 adsorption stability.

Descriptors	β (regression coefficient)	SE(β) (standard error of regression coefficient)	Z_{Wald} (β /SE(β))	p value of Z_{Wald}	β_{lower}	β_{upper}
constant term	6.26	2.91	2.15	0.032	0.55	11.97
HOMO-LUMO gap	-4.55	0.99	-4.59	0.0001	-6.49	-2.60
binding energy	1.08	2.01	0.54	0.592	-2.86	5.01
ionization potential	0.10	0.14	0.73	0.466	-0.17	0.38
coordination number	-0.40	0.35	-1.15	0.252	-1.09	0.29
$G = 97.4$			$\chi^2(0.95,4) = 9.49$			

Next, a new reduced model was constructed based only on HOMO-LUMO gap so that the results could be further verified and visually presented in a two dimensional plot. As it is shown in Table 5.13, the G value (96.2) of the new model is very close to the G value (97.4) of the full model verifying the significance of HOMO-LUMO gap. This effect seems to be through the presence of unpaired electron, which is the most significant variable for HOMO-LUMO gap [52].

Table 5.13. Statistics of the reduced model.

Structural Properties	β (regression coefficient)	SE(β) (standard error of regression coefficient)	Z_{Wald} (β /SE(β))	p value of Z_{Wald}	β_{lower}	β_{upper}
constant term	7.06	1.30	5.44	0.0001	4.52	9.60
HOMO-LUMO gap	-4.42	0.90	-4.94	0.0001	-6.18	-2.67
$G = 96.2$			$\chi^2(0.95,1) = 3.84$			

The negative value of the coefficient of HOMO-LUMO gap indicates that the probability of stable O_2 adsorption is negatively affected by the increase of HOMO-LUMO gap, the cut off value being 1.44 eV. This makes sense when the even-odd pattern observed in the change of HOMO-LUMO gap with size is considered. The lower values of the gap correspond to the clusters with odd number of electrons, which were shown to adsorb O_2 [52, 100]. The plot of HOMO-LUMO gap versus the probability of O_2 adsorption stability

shows a logistic curve supporting these findings (Figure 5.36). Except three data points, all the predicted O_2 adsorption probabilities lie quite close to the actual situations.
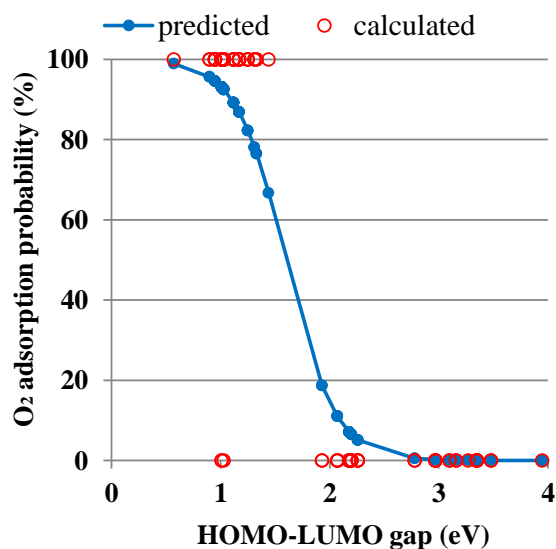
Figure 5.36. Effects of HOMO-LUMO gap on O_2 adsorption probability [228].

5.2.2.3. Descriptors Versus O_2 Adsorption Strength. In the second part of the work, the stable adsorption data (50 data points) were divided into two categories as explained at the beginning of Results and Discussion: strong adsorption if $E_{ads} < -0.50$ eV and weak adsorption if $E_{ads} > -0.50$ eV. The effects of descriptors and structural properties on the adsorption strength are presented below.

Since all the clusters with unpaired electron already had the ability to adsorb O_2, the clusters without unpaired electron were excluded; therefore, this descriptor was not treated as a predictor variable anymore. Hence; size, charge and coordination number were used as the descriptors of this model. The results are summarized in Table 5.14. Although the G value of 16.4 is not as high as the G values of the models discussed above for the adsorption stability, it is still well above the critical chi-square value (7.81). Thus, this model should also be accepted as significant.

According to the findings summarized in Table 5.14, the strength of O_2 adsorption depends on the size and the charge of the cluster (p values are 0.018 and 0.010 respectively). The coordination number seems to have no significant effect.

Table 5.14. Statistics of descriptors versus O_2 adsorption strength.

Descriptors	β (regression coefficient)	SE(β) (standard error of regression coefficient)	Z_{Wald} ($\beta /SE(\beta)$)	p value of Z_{Wald}	β_{lower}	β_{upper}
constant term	0.82	68.5	0.01	0.99	-133.4	135.1
size	-0.52	0.22	-2.36	0.018	-0.96	-0.09
charge	-1.53	0.59	-2.59	0.010	-2.68	-0.37
coordination number	0.27	0.35	0.78	0.44	-0.41	0.96
$G = 16.4$			$\chi^2 (0.95,3) = 7.81$			

The coefficient of size and charge are both negative as shown in Table 5.14, indicating that O_2 adsorption becomes weaker as the cluster size increases or the charge changes in the order of anionic-neutral-cationic (Figure 5.37a). Indeed, the effect of charge on the adsorption strength of O_2 over Au clusters was shown by previous studies. Despite of the early claims that cationic and neutral clusters show no reactivity towards O_2 except for Au_{10}^+, the cationic and neutral clusters were also shown to adsorb O_2 if they possess unpaired electrons [52, 228]. Nonetheless, the strength of O_2 adsorption on anionic Au clusters was always demonstrated to be greater than the adsorption strength on cationic and neutral clusters [52, 100, 228].

The decrease of strong adsorption probability with increasing size is also clear from Figure 5.37a. It should be noted that, the figure indicates the mean probabilities for strong adsorption, since the percent probability values shown in the figure were calculated by averaging the values over all the possible adsorption sites for that cluster, rather than considering only the maximum adsorption energies. Some cluster sizes do not appear in the figure (for example the anionic cluster with the size of three), because they do not result any stable adsorption for that charge; hence they were eliminated from the data set in this section.

The decrease of the mean probabilities for strong adsorption with the increasing size can be attributed to the increase in the number of inner atoms (atoms located on the flat surfaces of the clusters) with high coordination number parallel to the arguments in the literature that the strength of adsorption decreases with increasing coordination number [107]. The reason why coordination number was found to be insignificant in this work may

be related to its mild correlation with the size (Pearson correlation coefficient of 0.57). On the other hand, it was also argued that the underlying factor determining the adsorption ability of a cluster was the shape of the frontier orbitals (HOMO, in this case) rather than the coordination number of the adsorption site [229, 230]. In other words, O_2, being an electron acceptor, was argued to adsorb on the sites where HOMO protruded furthest in the vacuum [230]. Since the lobes of HOMO over the inner atoms are generally smaller than those on the corners/edges, and the number of inner atoms increases with increasing the cluster size, the orbital shape seems to reflect its effect through the cluster size.
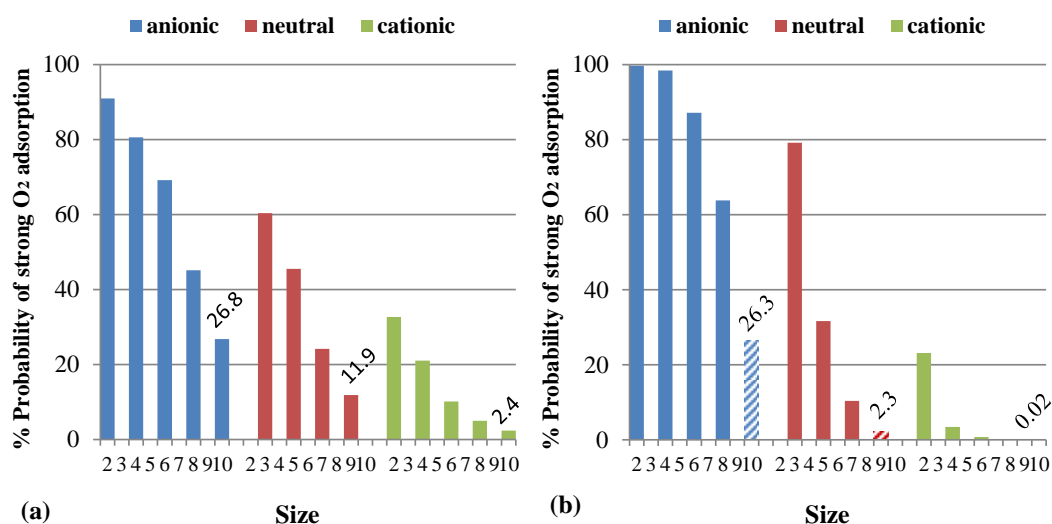
Figure 5.37. Dependency of strong adsorption probability on size and charge for: (a) full model, (b) using Au_2-Au_8 clusters to predict the results of Au_9-Au_{10} [228].

The extrapolation ability of logistic regression was tested by constructing a model from the data belonging to Au_2-Au_8 clusters, and this model was used to predict the probability of strong adsorption over Au_9-Au_{10} clusters (data unseen by the model). The new model was found to be also significant with its G value of 27.7, which is significantly higher than the critical chi-square value (7.81). Figure 5.37a and Figure 5.37b compares the probabilities of models generated from Au_2-Au_{10} and Au_2-Au_8 respectively; the numerical values of predicted probabilities for Au_9-Au_{10} were also given on the top of the corresponding bars. The predictions of both models for Au_{10}^- clusters are quite close with each other (26.8%-26.3%) as well as with the DFT computed value (16.7%). The probability of strong adsorption for Au_9 is predicted better by the new model (2.3% compared to 11.9% of the original model) since the DFT computed value is 0%. The

probability of the Au_{10}^+ cluster changes from 2.4% to about 0%, both being lower than the actual probability of 20% found by DFT calculations; however the model still successfully captures the low probability of strong adsorption for this cluster.

5.2.2.4. Structural Properties Versus O_2 Adsorption Strength. Finally, the structural properties were used to model the strength of O_2 adsorption (Table 5.15). The G value (14.3) of the model is still above the critical chi-square value (9.49) indicating that the model is still significant. The only significant variable for the adsorption strength, which has a p value below 0.05, is the ionization potential.

Table 5.15. Statistics of structural properties versus O_2 adsorption strength.

Structural Properties	β (regression coefficient)	SE(β) (standard error of regression coefficient)	Z_{Wald} (β /SE(β))	p value of Z_{Wald}	β_{lower}	β_{upper}
constant term	3.84	3.65	1.05	0.294	-3.31	11.0
HOMO-LUMO gap	0.37	2.36	0.15	0.877	-4.27	5.00
binding energy	-2.50	1.55	-1.61	0.107	-5.54	0.54
ionization potential	-0.25	0.12	-2.11	0.035	-0.48	-0.02
coordination number	0.03	0.30	0.10	0.924	-0.56	0.61
$G = 14.3$			$\chi^2(0.95,4) = 9.49$			

The ionization potential has a negative coefficient meaning that the strength of adsorption is negatively affected by this variable. This is something expectable for the anionic and neutral clusters since the adsorption of O_2 on these clusters are known to be accompanied by an electron transfer from Au to the π^* orbital of O_2 [100, 226, 229]. Hence, the lower the value of ionization potential, which is defined as the amount of energy required to pull off an electron, the easier to make the electron transfer required for the adsorption. On the other hand, an electron transfer in the opposite direction (from O_2 to Au) is observed for the cationic clusters during O_2 adsorption [52]. This means that for the cationic clusters, adsorption of O_2 need not be stronger as ionization potential decreases, which is contradictory to the suggestion of the model. This explains the relatively lower G value (lower prediction ability) for this model compared to the previous models above.

5.3. Knowledge Extraction from Published Data

5.3.1. Neural Network Analysis of Selective CO Oxidation over Copper-Based Catalysts for Knowledge Extraction from Published Data

The database for Cu based catalysts was constructed from the experimental data reported in 20 selected publications (as described in Section 3.5). Some of the variables in these works (as shown in Table 3.4) are continuous in their nature (Cu weight percentage, metal additive weight percentage, calcination temperature and time, reaction temperature, feed compositions, and F/W), whereas others are categorical (catalyst preparation method, metal additive type, support type).

Each continuous variable was treated as one input to the neural network in the range between the minimum and maximum values reported in the 20 publications. Each option for the categorical variables, such as support type and preparation method, was treated as an individual input variable having the value of 1 or 0 depending whether it was used or not, respectively. For example, the catalysts in the database were prepared by one of the eight preparation methods described in Table 3.5. This variable was introduced into the neural network through eight input neurons, corresponding to one neuron for each method. For the data points taken from each publication, the input neuron corresponding to the preparation method used in that publication had the value of 1, and the input neurons of the other seven methods had the value of 0. Similarly, each second metal additive option was introduced through a neuron that had the value of 0 if that metal was not used for that data point. On the other hand, if the metal was used, the percentage loading of that metal was taken as the value of that variable (continuous), allowing the network to predict the effects of both metal type and loading. As a result, 14 variables (Table 3.5, column 1) treated through 32 input entries (Table 3.5, column 3), were used for the neural network models. The ranges of the continuous variables are also as given in Table 3.5.

Figure 5.38 shows a simple representation of the conceptual approach to the problem. First, the optimal neural network topology representing this database was searched by analyzing the performances of several networks according to their generalization accuracies [212], which should be as high as possible if the network is to be

used for knowledge extraction from the literature (a generic neural network structure is shown in Figure 5.39). Then, the optimal network topology was trained again using the data from 19 publications and used to predict the results in the remaining publication to test the ability of the network to predict the unseen data. This procedure was repeated until all of the publications had been tested to ensure that the optimal network structure could predict the unseen data over the widest possible range of experimental conditions. The differences between the predictions and the corresponding experimental results were used to determine the mean absolute error of testing for each publication (dividing the sum of absolute errors by the number of data points used in that publication). These error values indicate whether the results of the corresponding publication could be predicted by a neural network instructed by the experimental results of the other works [53].

As the next step, if necessary, the unsuccessful predictions could be used to determine the limits and exclusions of the network and to reconstruct the database and/or the optimal network topology. In the present case, however, the exclusions and limits identified did not require any reconstruction.
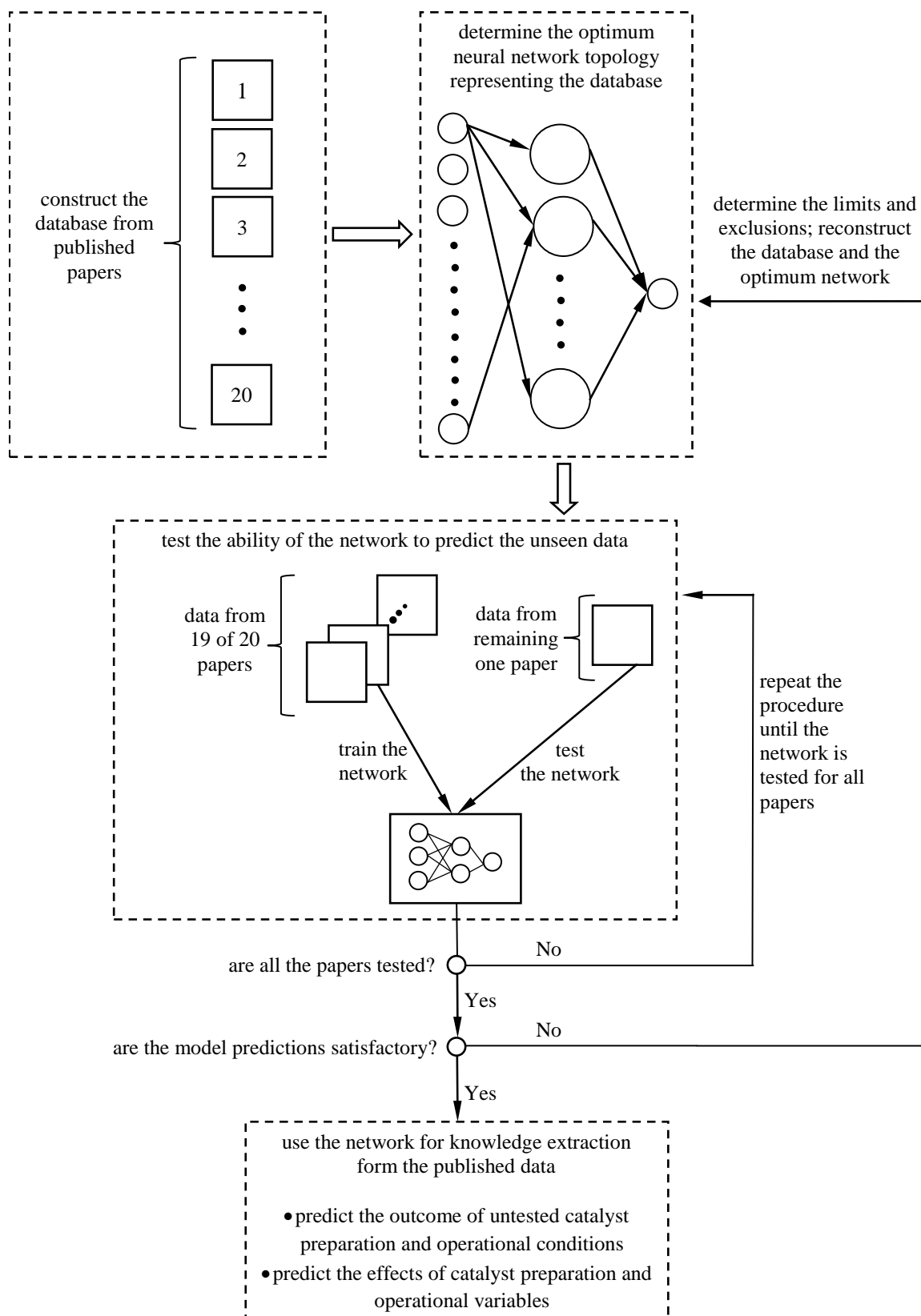
Figure 5.38. Conceptual approach for knowledge extraction from published data for Cu-based catalysts [53].
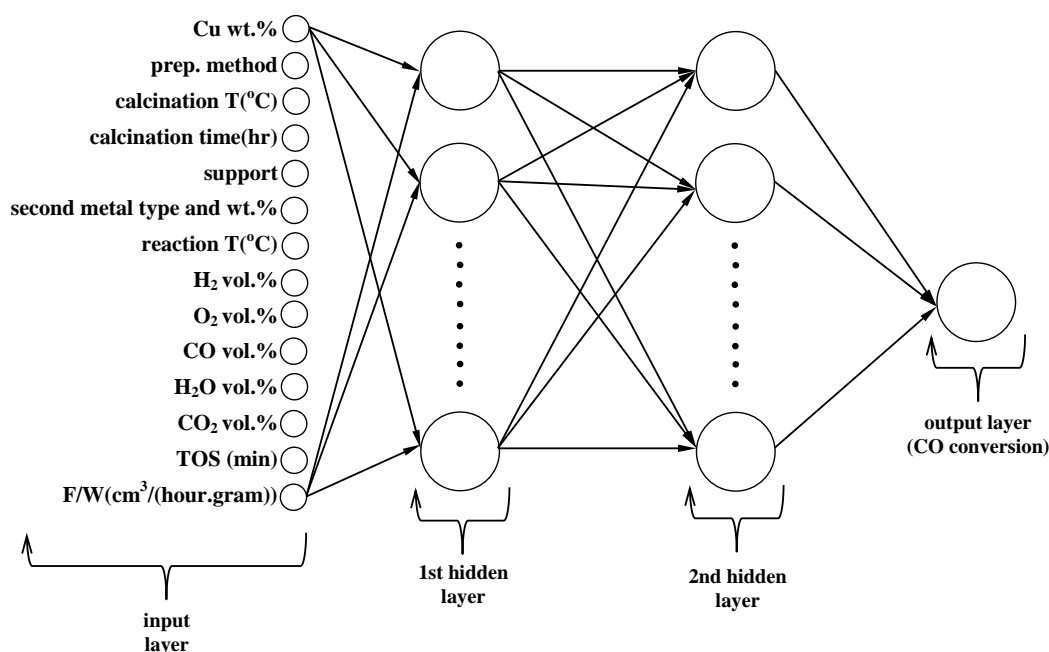
Figure 5.39. A generic neural network topology modeling catalytic performance [53].

5.3.1.1. Constructing the Optimal Neural Network Model. To find the best neural network model representing the database, several network topologies were tested and compared according to their RMSEs for testing and training (Figure 5.40). Only the results of two-hidden-layer networks are presented here because they produced better fits than single-hidden-layer networks. The notation a-b is used to label the networks, where a and b refer to the numbers of neurons in the first and second hidden layers, respectively.
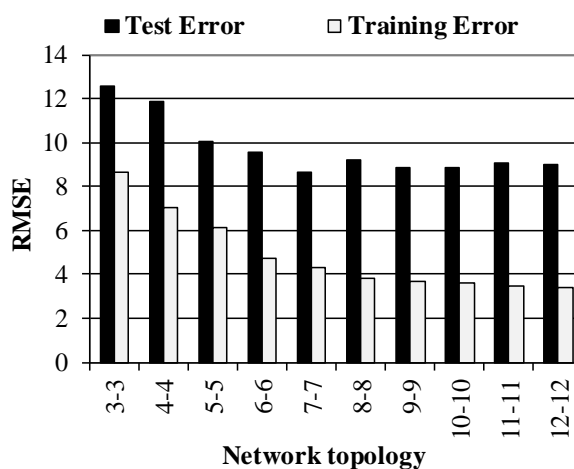
Figure 5.40. Comparison of errors for various neural network topologies [53].

As the network size increases, the training error generally decreases because of the use of more weights [118, 212]. The testing error, on the other hand, has the tendency to decrease first and then to increase as the network gets larger, because of the increase in the degree of overlearning [53]. Because the main purpose was to find a successful knowledge-extraction model with the ability to predict the results of a new set of conditions using existing data, the testing error should be as small as possible. Hence, the network architecture (7-7) with the lowest RMSE of testing (8.69) was chosen as optimal and used in the remaining parts of the work. It should also be noted from Figure 5.40 that the increase in the RMSE for testing in large networks was not as sharp as the decrease in the small sizes. This is a good indicator that the overlearning in the large networks was apparently prevented because of the use of the early-stopping technique. Plots of experimental versus predicted CO conversions for training and testing are shown in Figure 5.41, which indicates considerably successful fittings.
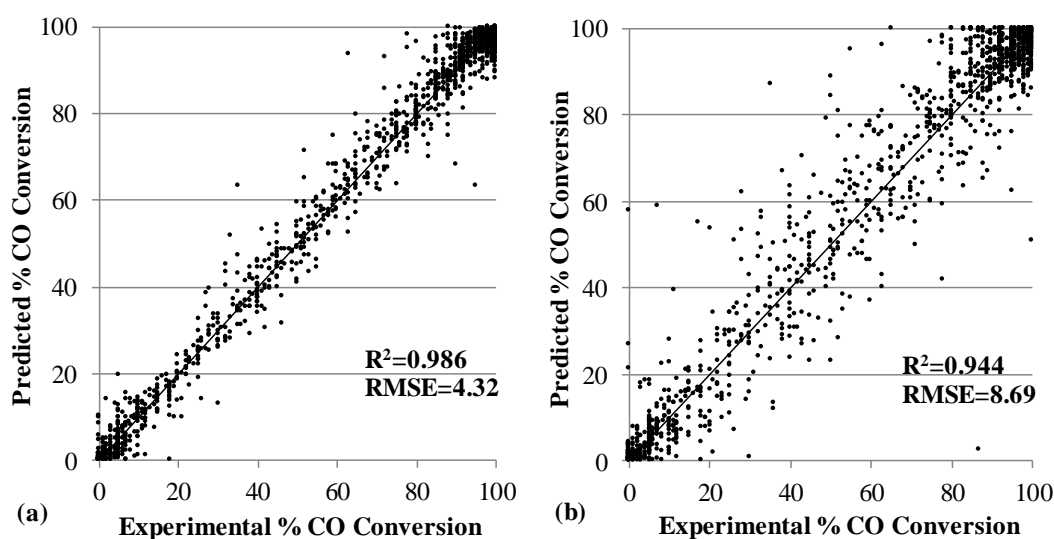
Figure 5.41. Experimental vs. predicted CO conversion values for the entire database: for (a) training, (b) testing [53].

5.3.1.2. Testing the Optimal Network to Predict the Unseen Data. The 7-7 network structure was used to predict the results obtained in each publication after being retrained with the remaining 19 publications. This procedure was repeated for 14 publications, so that the ability of the network to predict the unseen data could be seen over a wide range of experimental conditions. The remaining six publications were left out of testing because they had some unique variables that did not exist in the other publications. For example,

Kosmambetova *et al.* [146] tested the effects of TiO_2 and MnO_2 as supports, whereas the other publications mostly used CeO_2 as the support. Similarly, Ribeiro *et al.* [149] studied an uncommon support (Nb_2O_5) that was not repeated in any other work. Naturally, when the data related to these unique experiments is not fed to the neural network during the training process, there is no way for a model to learn the effects of these supports and accurately predict the corresponding CO conversions. Indeed all those supports were reported to provide poor CO conversions compared to the use of CeO_2 [146, 149]; and there were no such data in the other publications that would have taught this information to the network. Although these six publications were excluded in the testing stage, the database or optimal network did not need to be reconstructed. These data were still beneficial in enriching the model because of their common variables such as temperature or Cu amount, and the unique variables did not participate in the predictions, so that they did not have any negative effects.

Table 5.16 reports the performances of the neural networks in predicting the outcomes of the publications in order from the highest R^2_{test} values to the lowest. The predictions for 11 publications of the 14 tested were quite successful, as indicated by their mean absolute errors (lower than 15%). For example, the mean absolute error of testing for the work of Gomez-Cortés *et al.* [148] was 6.06%, meaning that the corresponding conversion values could be estimated with an average accuracy of ±6.06%. This can be considered as a quite successful prediction, as is also evident from the R^2 test value of 0.946.

Plots of predicted versus experimental CO conversions for the most successfully predicted six publications are presented in Figure 5.42. The symbols represent the CO conversion results of the corresponding publication predicted by the neural network model constructed with the remaining 19 publications, plotted against the experimental values from that publication. Although the deviation of the data points from the y=x line seems to be significant in some of the plots, the predictions are actually much better than they appear. Most of the experimental data points are near 100% CO conversion level leading to the fact that the neural networks are better tuned to this region (because the data around this conversion level had more effect on determining the weights of neural networks). Consequently, the model predictions are better for the conversions near 100%. These well-

predicted data points overlap with each other, whereas the data points at lower conversion values are mostly scattered as individual points, giving a false impression of the fitness of the models. As a result, the models should be evaluated and compared using the R^2 values rather than their appearances.

Table 5.16. Prediction errors of individual publications for Cu based catalysts.

Reference	Mean Absolute Error of Testing	R^2_{test}	Possible Explanation	
Gomez-Cortés et al. [148]	6.06	0.946	predicted successfully	
Li et al.30 [153]	4.55	0.935		
Razeghi et al. [154]	7.63	0.928		
Zhu et al. [29]	3.70	0.910		
Jung et al. [143]	6.09	0.897		
Wu et al. [152]	8.66	0.882		
Liu et al. [30]	9.87	0.855		
Avgouropoulos & Ioannides [85]	11.01	0.798		
Kim & Cha [88]	11.98	0.791		
Avgouropoulos et al. [142]	10.72	0.725		
Ayastuy et al. [151]	13.48	0.694		
Zou et al. [28]	16.56	0.373	conversion is too low compared to other studies	could not be predicted successfully probably due to these reasons
Martinez-Arias et al. [86]	20.61	0.309	significant activity loss at high temperatures	
Lendzion-Bielun et al. [155]	15.47	0.142	conversion is too low compared to other studies	
Park et al. [144]			one of the two papers with Al_2O_3 as support*	these variables are unique to these papers
Polster et al. [150]			use of urea gelation as preparation method	
Marino et al. [145]			use of MgO, La_2O_3 as support	
Firsova et al. [147]			use of Fe & Ni as second metal additive	
Kosmambetova et al. [146]			use of TiO_2, Al_2O_3, MnO_2 as support	
Ribeiro et al. [149]			use of ZrO_2, Nb_2O_5 as support	

* the other publication containing Al_2O_3 (Kosmambetova et al [146]) did not have sufficient number of data points for prediction of its results.
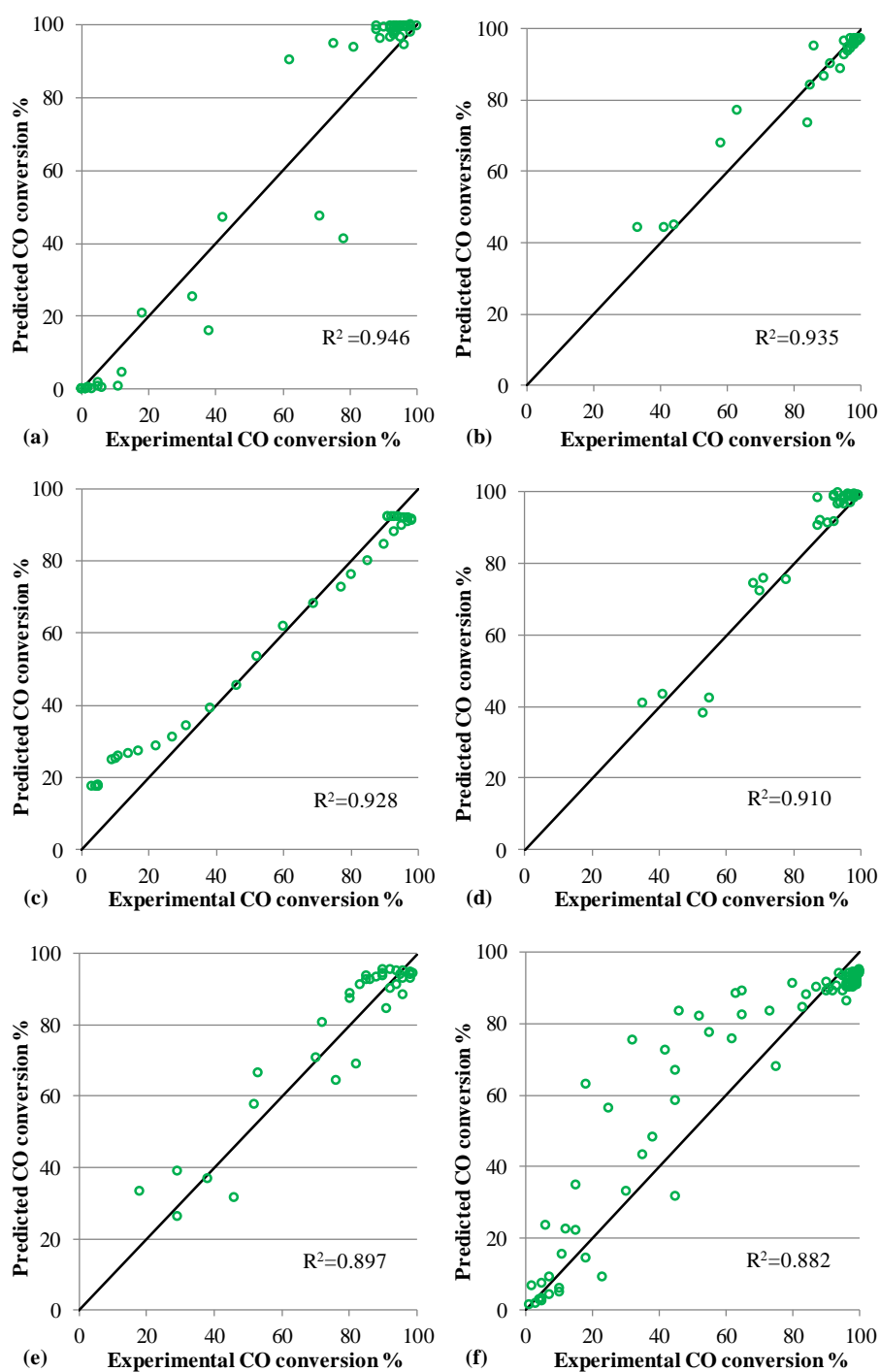
Figure 5.42. Experimental vs. predicted CO conversion for: (a) Gomez-Cortés *et al.* [148], (b) Li *et al.* [153], (c) Razeghi *et al.* [154], (d) Zhu *et al.* [29], (e) Jung *et al.* [143], (f) Wu *et al.* [152], (g) Liu *et al.* [30], (h) Avgouropoulos and Ioannides [85].

The data from publications that were successfully predicted usually contained the most commonly studied variables such as temperature, Cu content, and feed composition (as opposed to the six publications left out because of their unique variables), whereas the

three publications that could not be predicted successfully had unexpectedly low or high CO conversions. This suggests that, if a sufficient number of data points around a set of conditions are available in the literature on catalytic CO oxidation, the neural networks can indeed be used to predict the outcome of the experimental conditions that have not yet been studied.

5.3.1.3. Predicting the Effects of Catalyst Variables. Experimental works in the field of catalysis are performed not just to find the optimum conditions, but also to analyze the effects of catalyst preparation or operating variables so that the catalytic processes can be understood better. These effects should also be part of the knowledge extraction from the literature. Therefore, the neural network was also tested whether it could predict the trends reported in the literature. In this section, some sample plots of experimental versus predicted CO conversions for the most commonly studied variables are present. The symbols in Figures 5.43-5.49 show the experimental data reported in the corresponding publication, and the lines are the model predictions.

In Figure 5.43, the effect of Cu content, which is one of the most important catalyst preparation variables for this catalytic system, on CO conversion is analyzed through the experimental results reported by Wu *et al.* [152], Zhu *et al.* [29], Gomez-Cortés *et al.* [148] and Ayastuy *et al.* [151]. The CO conversion was found to increase with increasing temperature in all four works, although some activity loss with a further increase in temperature was also observed in the studies of Wu *et al.* [152] and Gomez-Cortes *et al.* [148]. As shown in Figure 5.43, the model predictions for all four publications were quite satisfactory.

It should be noted that the experimental data in these four publications were obtained for different values of other input variables, as detailed in the caption of Figure 5.43. Therefore, the plots presented in this figure, as well as the others in the remainder of this section should be used to compare the experimental and predicted CO conversion values but not the experimental works among themselves.
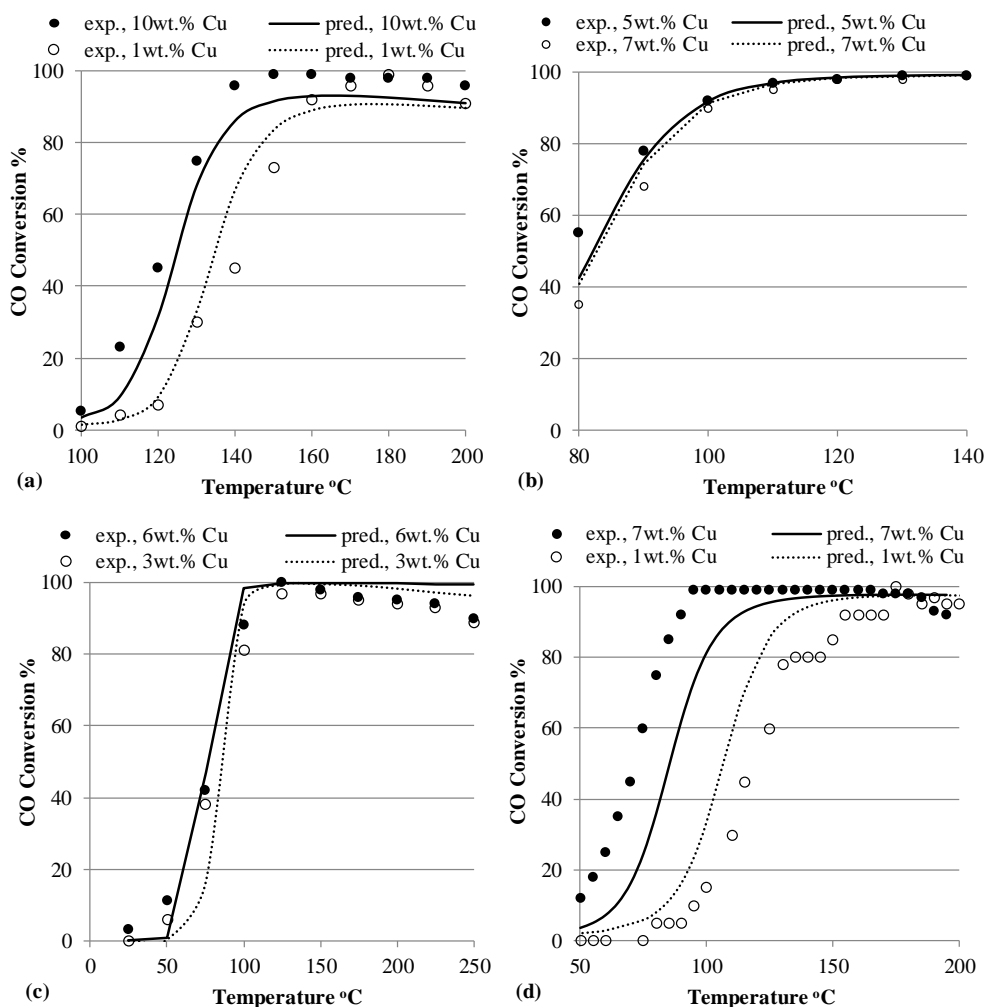
Figure 5.43. Effect of Cu content on CO conversion for: (a) 1% CO, 1% O_2, 50% H_2, 15% CO_2, 10% H_2O [152], (b) 1% CO, 1% O_2, 50% H_2 [29], (c) 1% CO, 1% O_2, 50% H_2 [148], (d) 1% CO, 1% O_2, 60% H_2 [151].

The type and loading of metal oxide additive are also significant for the activity of the CO oxidation catalysts. Li *et al.* [153] reported that, when a small amount of manganese oxide was added to the catalyst, a more stable solid with a larger catalyst surface area was formed, and more lattice oxygen was supplied, thereby increasing the catalytic activity. The positive effect of Mn addition was more apparent at lower temperatures, as shown in Figure 5.44. The neural network model also successfully predicted the general trends, although the predictions were slightly lower than the corresponding experimental values.
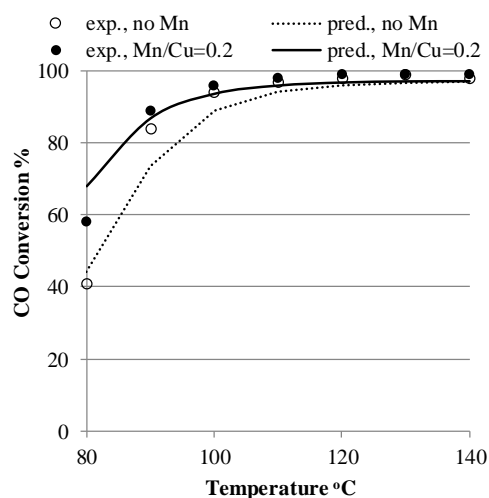
Figure 5.44. Effect of Mn/Cu molar ratio on CO conversion for 5 wt.% Cu over CeO_2 (1% CO, 1% O_2, 50% H_2, F/W=60000 cm^3/g·h) [153].
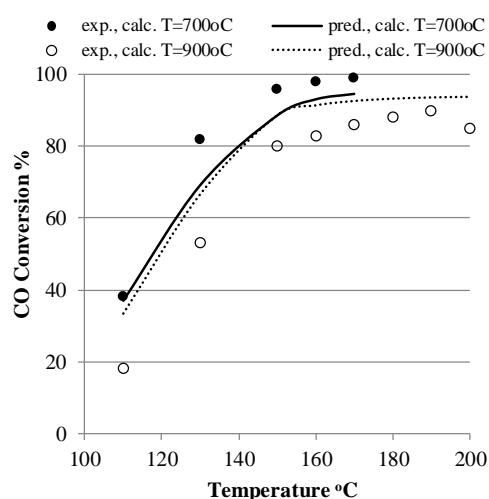
Figure 5.45. Effect of calcination temperature on CO conversion for 5.1 wt.% Cu over CeO_2 (0.8% CO, 0.8% O_2, 71.9% H_2, 23.5% CO_2, F/W=37440 cm^3/g·h) [143].

The effects of support type and catalyst preparation method could not be analyzed by the neural network model, because each publication used only one support and preparation method. Comparison of the data from two or more publications in this respect was also not possible because their experiments were carried out under different conditions. The only remaining catalyst preparation variable that could be analyzed is the calcination temperature. Although only Jung *et al.* [143] explicitly studied the effect of calcination temperature by comparing the activities of the catalysts calcined at various temperatures, other researchers also calcined their catalysts, providing some additional knowledge in this

area. The experimental and predicted CO conversions obtained at calcination temperatures of 700 and 900 ^oC are compared in Figure 5.45 for the study of Jung *et al.* [143]. The general trend was successfully predicted by the neural network model, although the predictions were not as good as for the Cu and Mn loadings.

Next, the effects of operating variables on CO conversion starting from H_2 in the feed were analyzed. The performance of the CO oxidation catalysts in a hydrogen-rich environment is important, as the catalysts should be able to oxidize carbon monoxide but not hydrogen [1]. Ayastuy *et al.* [151] tested their 7 wt.% Cu-containing catalyst, which exhibited the highest activity under the experimental conditions studied, in the presence and absence of H_2. As shown in Figure 5.46, the conversion was found to increase with increasing temperature and reached 100% at about 130 ^oC in the absence of H_2. The presence of 60% H_2 lowered the temperature range required for high activity [151]. Again, the effect of this variable was also predicted with an acceptable accuracy by the neural network model.
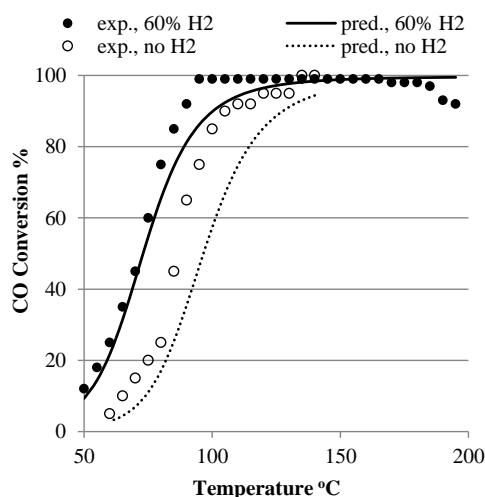
Figure 5.46. Effect of H_2 on CO conversion for 7 wt.% Cu over CeO_2 (1% CO, 1% O_2, F/W=120000 cm^3/g·h) [151].

In Figure 5.47, the effect of the amount of O_2 in the feed on the catalytic activity is analyzed through the results of two works that were performed with 1% CO in the feed [30, 88]. As indicated by Figure 5.47, the use of 0.5% (stoichiometric) or 1% O_2 did not make much difference at low temperatures, whereas the use of 1% O_2 provided slightly

better performance at high temperatures. The model predictions were again quite reasonable at low and moderate temperatures. The discrepancy between the experimental and predicted conversions at high temperatures (for $0.5\%O_2$) in Figure 5.47a [88] can be attributed to the fact that the effect of O_2 on the catalytic activity was studied in only these two publications. As a result, the decrease in CO conversion at high temperatures in the work of Kim and Cha [88] probably caused a false prediction of the same trend for the results of Liu *et al.* [30] and vice versa.
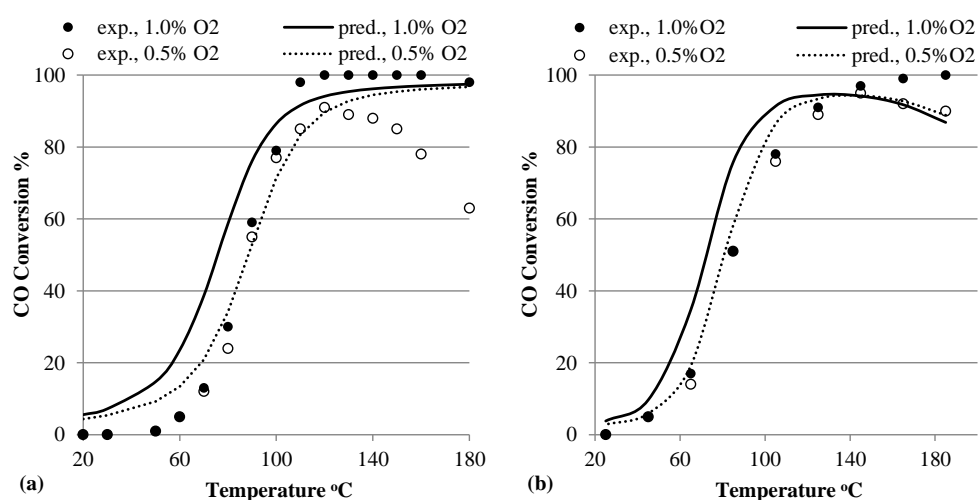
Figure 5.47. Effect of O_2 on CO conversion (a) 20 wt.% Cu over CeO_2 (1%CO, 50%H_2, F/W=120000 cm^3/g·h) [88], (b) 10 at.% Cu over CeO_2 (1%CO, 50%H_2, F/W=80000 cm^3/g·h) [30].

For small or medium-size fuel-cell applications such as houses and transportation vehicles, it has been proposed that H_2 be generated by steam reforming with the use of hydrocarbons or alcohols instead of being stored in gaseous form [53]. The hydrogen produced by this method contains CO_2 and H_2O; hence, any CO oxidation catalyst that can be used for this purpose has to work in the presence of some significant levels of CO_2 and H_2O [28, 142]. The effects of CO_2 and H_2O are analyzed in Figure 5.48 through the experimental results reported by Avgouropoulos and Ioannides [85], Liu *et al.* [30], Gomez-Cortés *et al.* [148] and Razeghi *et al.* [154].
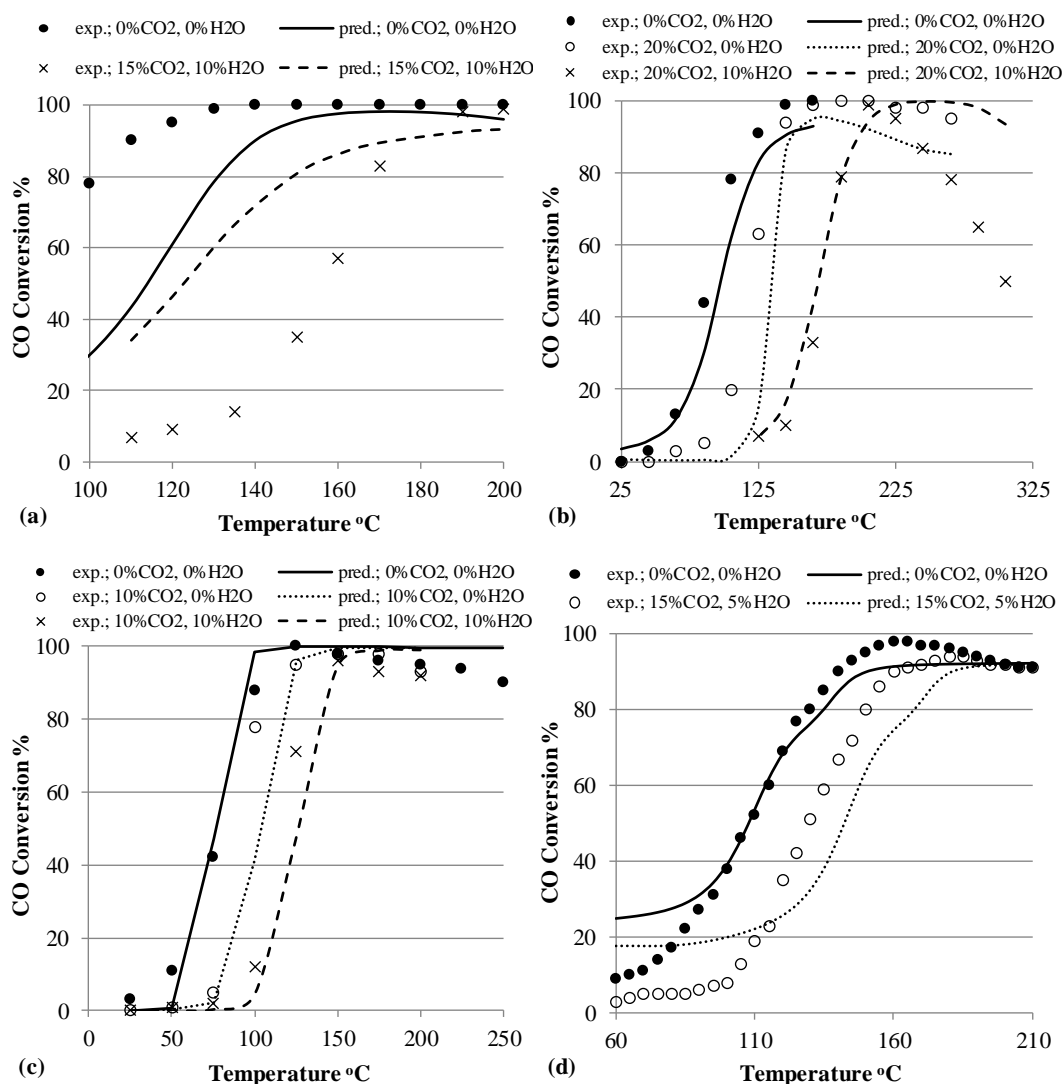
Figure 5.48. Effects of CO_2 and H_2O on CO conversion for: (a) Cu/(Cu+Ce) molar ratio of 0.15 over CeO_2 [85], (b) 10 at.% Cu over CeO_2 [30], (c) 6 wt.% Cu over CeO_2 [148], (d) 5 wt.% Cu over CeO_2 [154].

Even though the experiments were carried out under different catalytic preparation and operating conditions, the observations mostly coincide: The addition of both CO_2 and H_2O decreases CO conversion and increases the temperature required for high catalytic activity. Competitive adsorption of CO and CO_2 [30, 85, 148] and carbonate formation over the catalyst surface caused by the adsorbed CO_2 [154] were argued to be the possible reasons for the negative effects of CO_2. On the other hand, H_2O is believed to block the active sites of the catalyst [30, 85, 148] and form hydroxyl groups, decreasing the CO adsorption [151]. The trends reported by each publication were again successfully predicted by the neural networks.

Finally, the effect of F/W (W/F is used in some of the works) on the catalytic activity is analyzed through the results of Avgouropoulos and Ioannides [85] and Kim and Cha [88] in Figure 5.49. Two levels of F/W are compared in the figure, and it can be observed that, at high F/W, the catalytic activity is lower because of the use of a rapid gas feed per amount of catalyst, as expected. Although there are some discrepancies between the experimental and the neural-network-estimated CO conversion values, the predictions are still in accordance with the experimental trends.
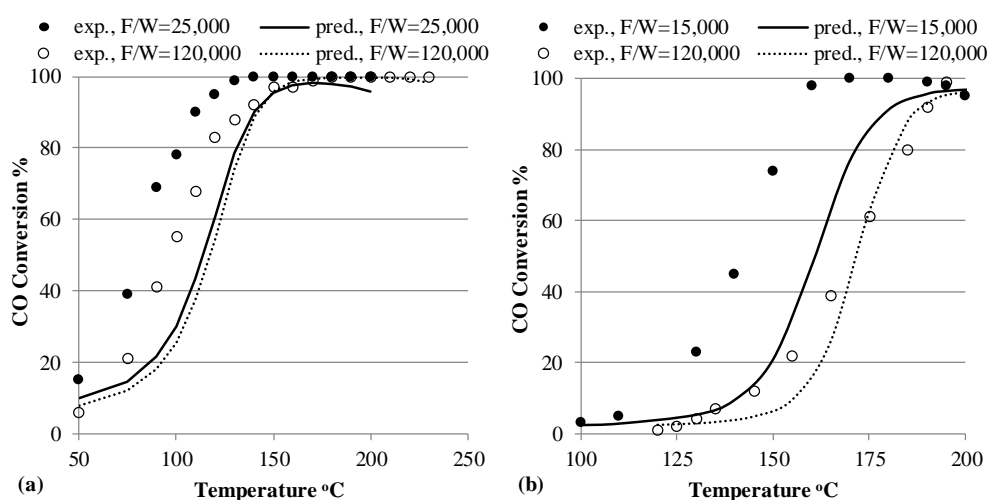
Figure 5.49. Effect of F/W on CO conversion for: (a) Cu/(Cu+Ce) molar ratio of 0.15 over CeO_2 (1% CO, 1.25% O_2, 50% H_2) [85], (b) 20 wt.% Cu over CeO_2 (1% CO, 1% O_2 50% H_2, 13.5% CO_2, 20% H_2O) [88].

5.3.1.4. Significance Analysis for Catalyst Variables. In this section, the relative significances of the catalyst preparation and operating variables on CO conversion are analyzed. This is an important piece of knowledge because an effective search for the best catalyst requires the manipulation of the most influential preparation and operating variables. Table 5.17 reports the relative significances of the preparation and operating variables. The reaction temperature was found to be the most influential variable (relative significance of 54.3%); its exclusion increased the RMSE of the model more than the exclusion of any other variable. This is an expected result from both mathematical (conversion goes from nearly 0% to 100% in the temperature range studied) and theoretical (temperature is usually important for most reactions) points of view. This also indicates

that the current approach can be used to find the most influential variable if it is not already known [53].

Table 5.17. Relative input significances for Cu based catalysts.

Variable	RMSE without the Variable	RMSE Difference*	Relative Significance (%)	Relative group Significance (%)
base metal (Cu wt.%)	11.39	7.07	14.0	
support	8.06	3.74	7.4	
preparation method	5.36	1.04	2.1	26.5
second metal additive	5.10	0.78	1.5	
calcination conditions	4.72	0.40	0.8	
reaction temperature	31.74	27.42	54.3	
H_2O vol.%	7.61	3.29	6.5	
CO_2 vol.%	6.71	2.39	4.7	
F/W	5.54	1.22	2.4	
H_2 vol.%	5.50	1.18	2.3	73.5
O_2 vol.%	5.41	1.09	2.2	
CO vol.%	4.82	0.50	1.0	
time on stream	4.75	0.43	0.8	

*the difference between RMSE without the variable or variables and RMSE of the original model (4.32)

Cu weight percentage was found to be the second most significant variable, with 14.0% relative significance. This is also reasonable, as one might expect the significance of Cu to be high because it is the active ingredient of the catalyst. However, all of the catalysts examined by the neural network models contained quite similar amounts of Cu as a result of experience accumulated over many years. In fact, the relative significance of 14.0% measured how the CO conversion was affected by the change of Cu weight percentage within the experimental range used in these publications. Otherwise, dramatic changes in Cu content (for example, setting the Cu weight percentage to 0) would have resulted in a higher relative significance. Similarly, if the temperature range had been set to a narrower interval (for example, 90-100 ^{o}C), the significance of the reaction temperature would not have been as great as found here, because a 10 ^{o}C change would not have produced a significant effect on CO conversion. Therefore, the relative significances for all of the variables in Table 5.17 should be considered strictly within the range of

experimental conditions investigated. This is exactly the information needed to be able to manipulate the most significant variable for future works by keeping all other variables in commonly reported ranges.

5.3.2. Analysis of Selective CO Oxidation over Noble Metal-Gold Based Catalysts for Knowledge Extraction from Published Data by Clustering and Artificial Neural Networks

The database for noble metal-gold based catalysts were constructed from the experimental data reported in 71 publications (as explained in Section 3.6), and the variables were treated similar to the work described above for the Cu based catalysts. The metal amount, calcination temperature and time, promoter amount as well as all the operating conditions were taken as continuous variables; whereas catalyst preparation method, base metal, support and promoter types were categorical. Each continuous variable was treated as one input to the neural network in the range between the minimum and maximum values reported in the entire set of publications. Each option for the categorical variables was treated as an individual input variable having the value of 1 or 0 depending whether it was used or not, respectively. For example fifteen different supports were used to prepare the catalysts in the database. This variable was introduced into the neural network through fifteen input neurons, corresponding to one neuron for each support type. The input neuron referring to the support used in a publication had the value of 1 while the input neurons representing the other fourteen supports had the value of 0. Base metal type and amount as well as the second metal type and amount carry categorical and continuous nature at the same time; thus, each base metal and second metal additive option was introduced through an input entry that had the value of 0 if that metal was not used. On the other hand, if the metal was used, the percentage loading of that metal was taken as the value of that variable (continuous), allowing the model to predict the effects of both metal types and their loadings. As a result, 14 variables (Table 3.7, column 1) treated through 60 input entries (Table 3.7, column 3), were used for modeling purposes.

Since it was impossible to predict the effects of unique variables on the catalytic activity, the data to be tested was reduced by eliminating these unique variables from the datasets extracted from each publication. If a base metal-support combination was not

studied in at least three publications those data were treated as unique; thus, the data of some publications had to be completely removed or only some parts of them (that were unique) were discarded. As a result, the reduced dataset (to be used for testing) contained 4070 data points from 58 publications as shown in the last column of Table 3.6.

First, the direct neural network approach was employed by finding the optimal neural network model that best represents the data. Then, the two-step approach composed of clustering and neural network modeling was implemented and the performances of both approaches on predicting the unseen data were compared. Finally, the effects of catalyst preparation and operating variables as well as their significances were determined by the predictions of the two-step approach.

The two-step approach to the problem is shown in Figure 5.50. The entire database (containing 5008 data points) was constructed from 71 publications and after eliminating the unique variables, the reduced dataset (containing 4070 data points) from 58 publications was formed. First, the direct neural network approach was employed by calculating the training errors of the neural networks from the full database and the testing errors from the reduced database. The optimal neural network topology representing this database was searched by analyzing the performances of several networks according to their generalization accuracies. After finding the optimal neural network topology it was used for predicting the unseen data similar to the procedure implemented for the data of Cu-based catalysts extracted from the literature (Section 5.3.1). In addition, the optimal topology was used to calculate the relative significances of each input variable.

The two-step approach procedure started by genetic algorithm based clustering of the reduced data. In order to find the optimum number of clusters separating the reduced dataset, several numbers of clusters were tested by a procedure similar to the k-fold cross validation method used for neural network modeling. Since the main purpose was to predict the experimental results of a publication by using the data of the other publications, the data of 57 out of 58 publications were used to construct the clusters; whereas the data of the remaining one publication was allocated to those clusters and the SSE value (calculated by the distance of each data point to the center of the cluster it belongs to) was recorded. The procedure was repeated 58 times each time with different publication to be

tested and different publications to construct the clusters [231]. The total SSE value calculated through 58 publications was recorded as the SSE of testing. The optimum number of clusters was chosen as the one that leads to the smallest SSE value.
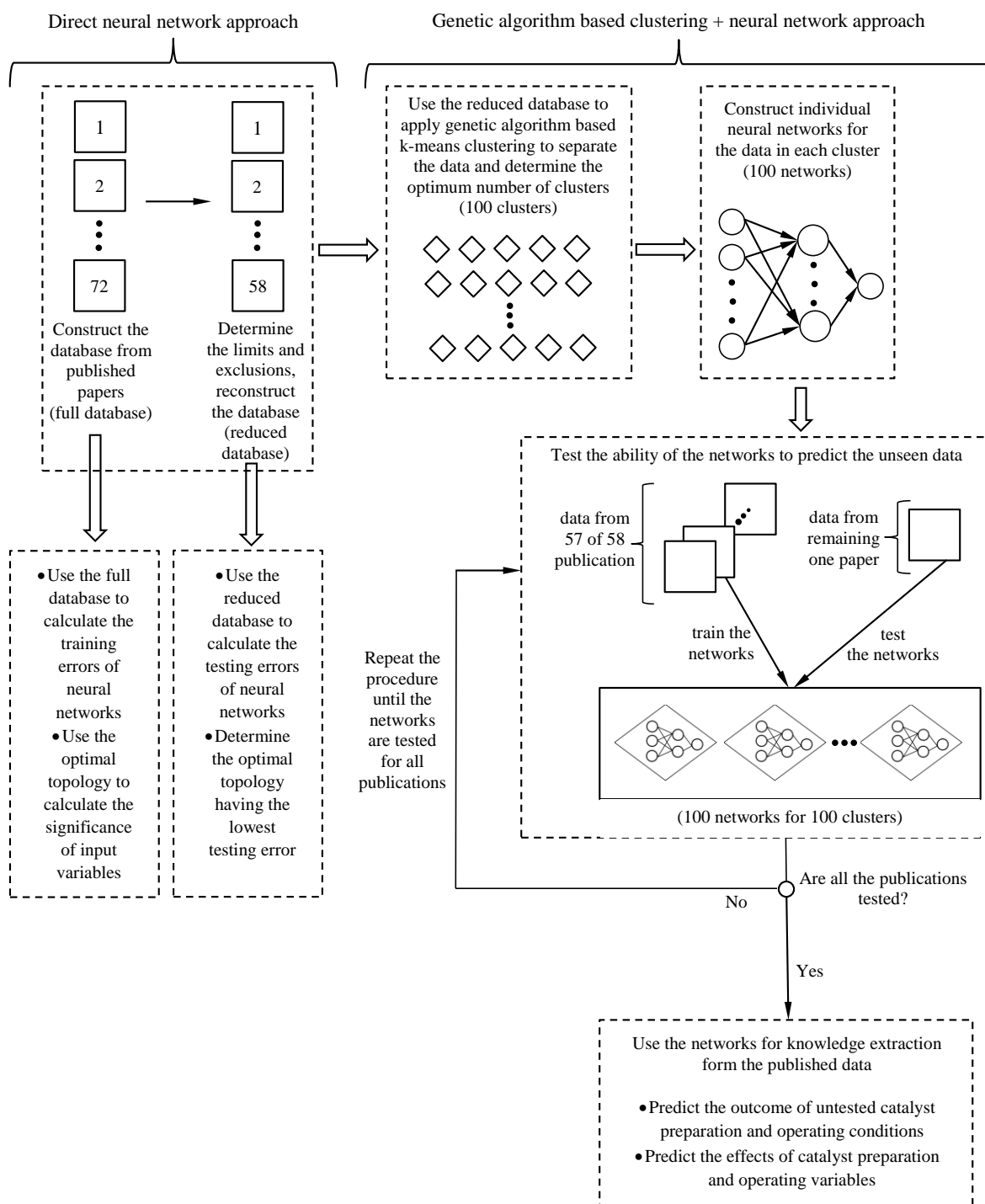
Figure 5.50. Two-step approach for knowledge extraction from published data for noble metal-Au based catalysts.

After finding the optimum number of clusters to separate the data, individual one hidden layer neural networks were constructed for the data in each cluster. These networks were trained using the data from 58 publications and used to predict the results in the remaining publication to test the ability of the networks to predict the unseen data. This procedure was repeated until all of the publications were tested to ensure that the optimal network structure could predict the unseen data over the widest possible range of experimental conditions. The differences between the predictions and the corresponding experimental results were used to determine the mean absolute error of testing and the R^2 values for each publication. These parameters indicate whether the results of the corresponding publication could be predicted by a neural network instructed by the experimental results of the other works.

5.3.2.1. Modeling by the Direct Neural Network Approach. Several two hidden layer network topologies with 60 input variables and one output variable (CO conversion) were constructed, and the performances of the networks were compared according to their RMSEs of training and testing as shown in Figure 5.51. The entire dataset (containing 5008 data points from 71 publications) was used to find the training errors while the reduced dataset (containing 4070 data points from 58 publications without any unique catalysts) was used to determine the testing errors of the neural networks. The notation "a-b" is used to label the networks, where "a" and "b" refer to the numbers of neurons in the first and second hidden layers, respectively.
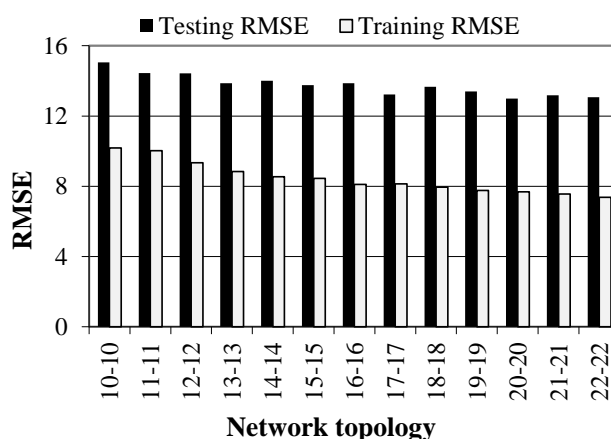
Figure 5.51. Comparison of errors for various neural network topologies (x-axis shows the number of neurons in the first and second hidden layers, respectively).

R²=0.952, RMSE=7.67   R²=0.856, RMSE=13.00
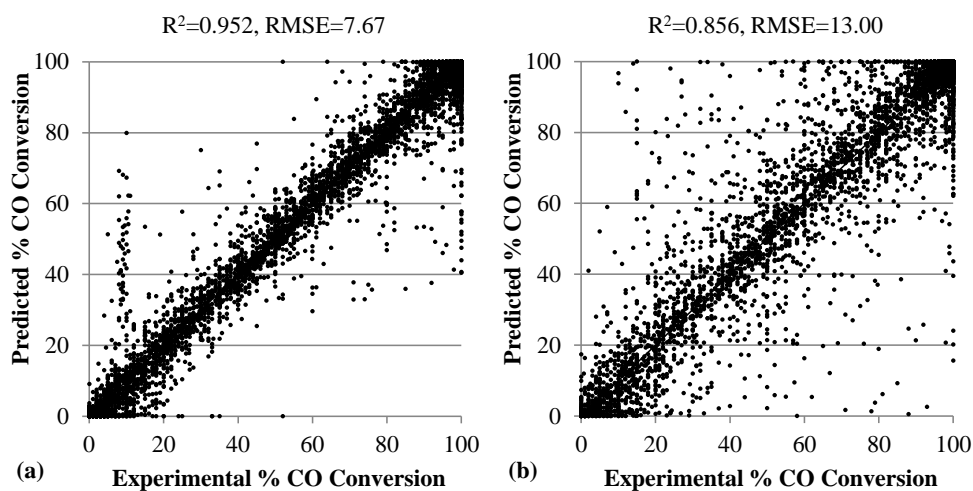
Figure 5.52. Experimental versus predicted conversion values (a) for training the entire database (b) for testing (4 fold cross validation) the reduced data.

Similar to the previous cases, the training error decreases with the increase of the network size [51, 53]. On the other hand, the testing error decreases with increasing network size and tends to stay constant (probably due to the use of early stopping to prevent overlearning). The neural network topology (20-20) with the minimum RMSE of testing (13.00) was chosen as the network that best represents the experimental data. Plots of experimental versus predicted CO conversions of training and testing for this network are shown in Figure 5.52, which indicates considerably successful fittings. The statistical fitness of the model for the individual papers is also given in Table 5.18 to compare with the fitness of the two-step model presented in the next section.

5.3.2.2. Modeling by the Two-Step Approach. After finding the optimal neural network topology for the direct neural network approach, the two-step approach was implemented by first finding the optimal number of clusters to separate the data. If the database was divided to small number of clusters, the data belonging to each cluster would be too dissimilar and the SSE value will naturally be too high [231]. On the other hand, if the number of clusters was large, each cluster may contain very little number of data points being unique to that cluster (lack of generalization ability); in addition, it would be quite difficult to construct neural network models with such datasets too small in size. Hence, the optimum number of clusters with the smallest SSE value (with the highest

generalization ability) was searched. In order to find the optimum number of clusters, 4-fold cross validation method was applied on the reduced dataset.

Table 5.18. Prediction errors of individual publications for noble metal-Au based catalysts by direct neural network approach.

Reference	R^2_{test}	Mean Absolute Error of Testing	Reference	R^2_{test}	Mean Absolute Error of Testing
Neri *et al.* [204]	0.77	14.8	Uguz and Yildirim [138]	0.18	18.2
Luengnaruemitchai *et al.* [194]	0.62	14.0	Manasilp and Gulari [160]	0.16	20.5
Liotta *et al.* [209]	0.62	21.2	Son [13]	0.15	30.8
Wootsch *et al.* [78]	0.55	21.3	Tibiletti *et al.* [166]	0.15	24.2
Ko *et al.* [18]	0.52	20.3	Avgouropoulos *et al.* [198]	0.14	10.9
Chin *et al.* [173]	0.46	30.0	Imai *et al.* [202]	0.14	25.3
Suh *et al.* [171]	0.45	19.9	Son and Lane [159]	0.13	19.4
Qiao *et al.* [197]	0.43	17.7	Bulushev *et al.* [161]	0.13	11.5
Ribeiro *et al.* [192]	0.41	23.0	Luengnaruemitchai *et al.* [22]	0.13	10.7
Yung *et al.* [201]	0.39	24.5	Chang *et al.* [188]	0.11	21.1
Cho *et al.* [14]	0.34	28.2	Chen *et al.* [184]	0.10	14.8
Minemura *et al.* [19]	0.33	20.7	Ince *et al.* [170]	0.10	26.1
Ayastuy *et al.* [15]	0.32	22.7	Bethke and Kung [156]	0.08	10.5
Özdemir *et al.* [165]	0.30	18.9	Galetti *et al.* [190]	0.06	24.4
Yan *et al.* [169]	0.28	16.9	Zhang *et al.* [200]	0.05	24.8
Padilla *et al.* [206]	0.27	31.3	Tabakova et al. [211]	0.05	18.9
Ko *et al.* [176]	0.27	23.2	Monyanon *et al.* [181]	0.03	21.2
Han *et al.* [80]	0.26	21.3	Pozdnyakova *et al.* [174]	0.01	21.7
Gluhoi and Nieuwenhuys [27]	0.25	9.2	Mozer *et al.* [205]	0.00	16.7
Teschner *et al.* [187]	0.25	16.7	Wang *et al.* [180]	0.00	1.5
Parinyaswan *et al.* [178]	0.24	15.7	Grisel and Nieuwenhuys [158]	0.00	32.7
Marques *et al.* [177]	0.24	33.0	Tompos *et al.* [196]	0.00	14.2
Tanaka *et al.* [195]	0.23	23.4	Wootsch *et al.* [167]	0.00	22.9
Luengnaruemitchai *et al.* [24]	0.22	30.0	Davran-Candan *et al.* [139]	0.00	22.3
Scire *et al.* [193]	0.21	18.1	Sangeetha and Chen [207]	0.00	17.5
Avgouropoulos *et al.* [142]	0.20	27.1	Huang *et al.* [186]	0.00	18.2
Chin *et al.* [17]	0.19	35.9	Marino *et al.* [168]	0.00	19.2
Ko *et al.* [10]	0.19	18.3	Zhao *et al.* [199]	0.00	34.4
Huang *et al.* [191]	0.19	19.7	Zhang *et al.* [162]	0.00	2.4
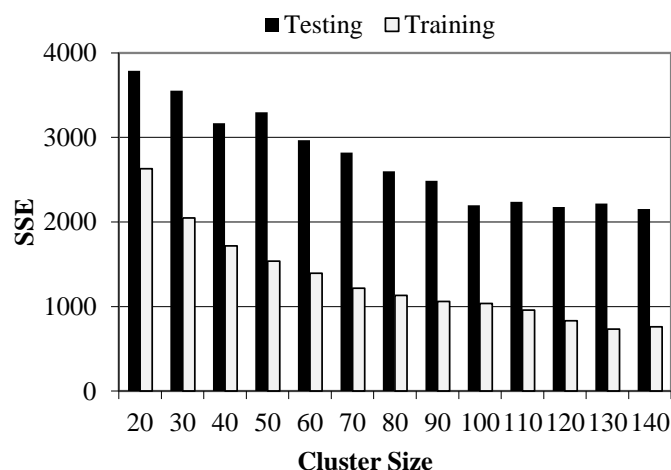
Figure 5.53. Testing and training SSE values for different number of clusters.

In Figure 5.53, testing SSEs for different number of clusters are shown by black bars. The white bars on the same figure correspond to the total SSE of all the clusters when all the data (58 out of 58 publications) were used for training (for constructing the clusters). As it was expected, the SSE of training decreases as the number of clusters increases (would become zero if the number of clusters were equal to the number of data points). On the other hand, the SSE of testing first decreases then tends to stay constant. It seems that separating the data to more than 100 clusters provides no additional gain in terms of generalization accuracy; hence, the optimum number of clusters was chosen as 100.

The power of the two-step approach was demonstrated on the training data of the entire database (71 publications) by first separating them into 100 clusters and then modeling the data belonging to each cluster by different small size neural networks (with one hidden layer). The improvement of the model can be observed by comparing Figure 5.52a ($R^2=0.952$) and Figure 5.54 ($R^2=0.978$), where the former is the predicted CO conversions versus the experimental ones obtained by direct neural network approach and the latter is the same analysis for the two-step approach.

Next, the testing performance of two-step approach based on clustering and neural network modeling was analyzed. The experimental results of 32/58 publications were predicted with R^2_{test} values higher than 0.5 by this approach (Table 5.19 and Table 5.20) while 5/58 publications had been predicted with R^2_{test} values higher than 0.5 by the direct neural network approach (Table 5.18). The two-step approach was found to be superior in

predicting the outcomes of the individual publications and hence, all further analyses in the rest of this study were done by this method.
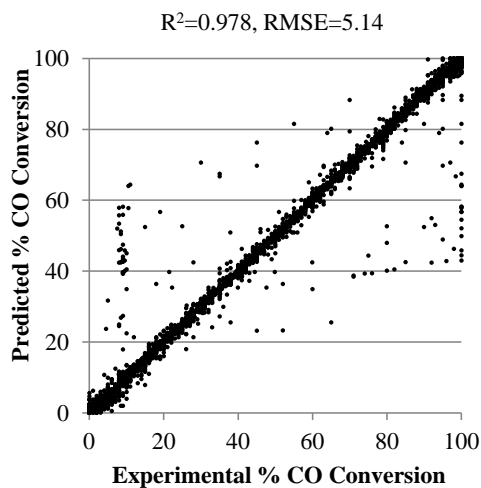
Figure 5.54. Experimental versus predicted CO conversion values for training data (71 publications) by genetic algorithm based clustering and neural network approach.

The results of publications predicted with R^2_{test} values higher than 0.25 are shown in Table 5.19 while those with R^2_{test} values lower than 0.25 are shown in Table 5.20 with possible explanations or comments for the unsuccessful predictions. The most common reason for unsuccessful predictions in Table 5.20 was that those publications investigated the effects of variables that had been rarely studied in the other publications such as using Sm or Pb as a promoter as in the work of Tompos *et al.* [196] or using Zn as a promoter as in the work of Avgouropoulos *et al.* [198]. If there are no sufficient data for a particular variable, the effects related to the presence or absence of that variable is quite hard to be determined. In addition, some of the unsuccessfully predicted publications contained variables with values being the maximum in the range of all publications, such as using F/W of 270000 cm^3/(h.g) as in the work of Bethke and Kung [156]. When the data at the limit of the range of a variable was excluded from the training data, the range of that variable becomes narrower and the data to be tested becomes out of the range of the model; hence, the model might fail to predict the corresponding results.

Table 5.19. Prediction errors of publications for noble metal-Au based catalysts with R^2_{test} higher than 0.25 by clustering and neural network approach.

Reference	R^2_{test}	Mean Absolute Error of Testing	Reference	R^2_{test}	Mean Absolute Error of Testing
Chin et al. [17]	0.95	6.21	Özdemir et al. [165]	0.65	13.57
Padilla et al. [206]	0.95	6.52	Ayastuy et al. [15]	0.64	15.21
Luengnaruemitchai et al. [194]	0.91	6.22	Huang et al. [186]	0.63	8.09
Neri et al. [204]	0.87	8.30	Zhao et al. [199]	0.60	11.30
Chin et al. [173]	0.85	9.68	Gluhoi and Nieuwenhuys [27]	0.58	5.84
Bulushev et al. [161]	0.85	3.48	Teschner et al. [187]	0.57	12.12
Koe et al. [10]	0.83	7.54	Avgouropoulos et al. [142]	0.54	18.67
Yung et al. [201]	0.81	10.93	Galetti et al. [190]	0.53	11.21
Pozdnyakova et al. [174]	0.81	7.93	Scire et al. [193]	0.53	14.13
Wang et al. [180]	0.78	0.67	Zhang et al. [200]	0.49	15.76
Luengnaruemitchai et al. [24]	0.76	14.57	Cho et al. [14]	0.48	21.89
Uguz and Yildirim [138]	0.76	6.61	Sangeetha and Chen [207]	0.48	10.88
Ko et al. [18]	0.73	12.51	Zhang et al. [162]	0.48	0.71
Davran-Candan et al. [139]	0.73	9.73	Ribeiro et al. [192]	0.47	19.66
Marques et al. [177]	0.73	18.49	Wootsch et al. [78]	0.46	24.74
Qiao et al. [197]	0.71	10.98	Imai et al. [202]	0.40	21.11
Son [13]	0.70	13.33	Manasilp and Gulari [160]	0.40	17.38
Suh et al. [171]	0.70	13.81	Tanaka et al. [195]	0.38	23.79
Marino et al. [168]	0.69	7.80	Yan et al. [169]	0.34	16.77
Liotta et al. [209]	0.69	16.32	Mozer et al. [205]	0.33	13.44
Tibiletti et al. [166]	0.67	13.48	Ko et al. [176]	0.29	20.06
Tabakova et al. [211]	0.66	11.17	Ince et al. [170]	0.29	20.89
Grisel and Nieuwenhuys [158]	0.66	16.92	Monyanon et al. [181]	0.28	18.37

Table 5.20. Prediction errors of publications for noble metal-Au based catalysts with R^2_{test} lower than 0.25 by clustering and neural network approach.

Reference	R^2_{test}	Mean Absolute Error of Testing	Comments for Unsuccessful Predictions
Avgouropoulos et al. [198]	0.24	10.54	Au-Zn/CeO$_2$ catalyst was investigated. Zn was used in one more publication as a promoter (7 data points) [176]. Although the predictions were close, the trends related to the presence of Zn could not be predicted.
Wootsch et al. [167]	0.20	17.75	5% CO (maximum in the range of all publications) was used. The predicted conversions were somewhat lower than the experimental values.
Han et al. [80]	0.18	23.88	Some catalyst preparation details were not clear in this publication. Predictions were not satisfactory probably due to the missing data.
Parinyaswan et al. [178]	0.17	14.57	Bimetallic Pt-Pd/CeO$_2$ catalyst was investigated. Although the effects related to the change of Pt could be predicted to some extent those of Pd could not, probably due to Pd being only in this publication
Tompos et al. [196]	0.15	11.52	Au-Pb/MgO and Au-Sm/MgO catalysts were investigated; the use of Pb was unique to this publication and the use of Sm was done in one more publication (18 data points) [198]. Although the predictions were close, the trends related to Pb and Sm could not be predicted.
Son and Lane [159]	0.13	14.91	The predicted conversion values were close but somewhat higher than the experimental values.
Minemura et al. [19]	0.06	23.70	Effects of alkali promoters were studied at T=100 ^{o}C. Effects of K and Na were successfully predicted while those of Li, Rb and Cs were not.
Huang et al. [191]	0.03	19.44	Ir/CeO$_2$ was used in only 2 more publications, the data of which (33 points) were insufficient to predict the results (70 points) in this work.
Chang et al. [188]	0.03	22.08	Au/TiO$_2$ was used in only 2 more publications, the data of which (33 points) were insufficient to predict the results (26 points) in this work.
Chen et al. [184]	0.00	10.86	Predictions were quite close to the experimental results but no trend was achieved.
Bethke and Kung [156]	0.00	10.85	Conversions were too low due to high F/W (maximum in the range of all publications); predictions were close but no trend was achieved.
Luengnaruemitchai et al. [22]	0.00	10.56	Data accumulated in a small range near 100%; predictions were close but no trend was achieved

The experimental CO conversion values from each publication was plotted against the predicted conversions, which were achieved by using the data of the remaining 57 publications, and the most successfully predicted nine publications with the highest R^2_{test} values are presented in Figure 5.55. In some of the plots the experimental data points are accumulated near 100% or 0% CO conversion levels; hence, these points overlap each other in the plots. Although some plots indicate some discrepancies between the predicted and experimental values, in general good data distributions around the y=x line can be observed for all the plots.
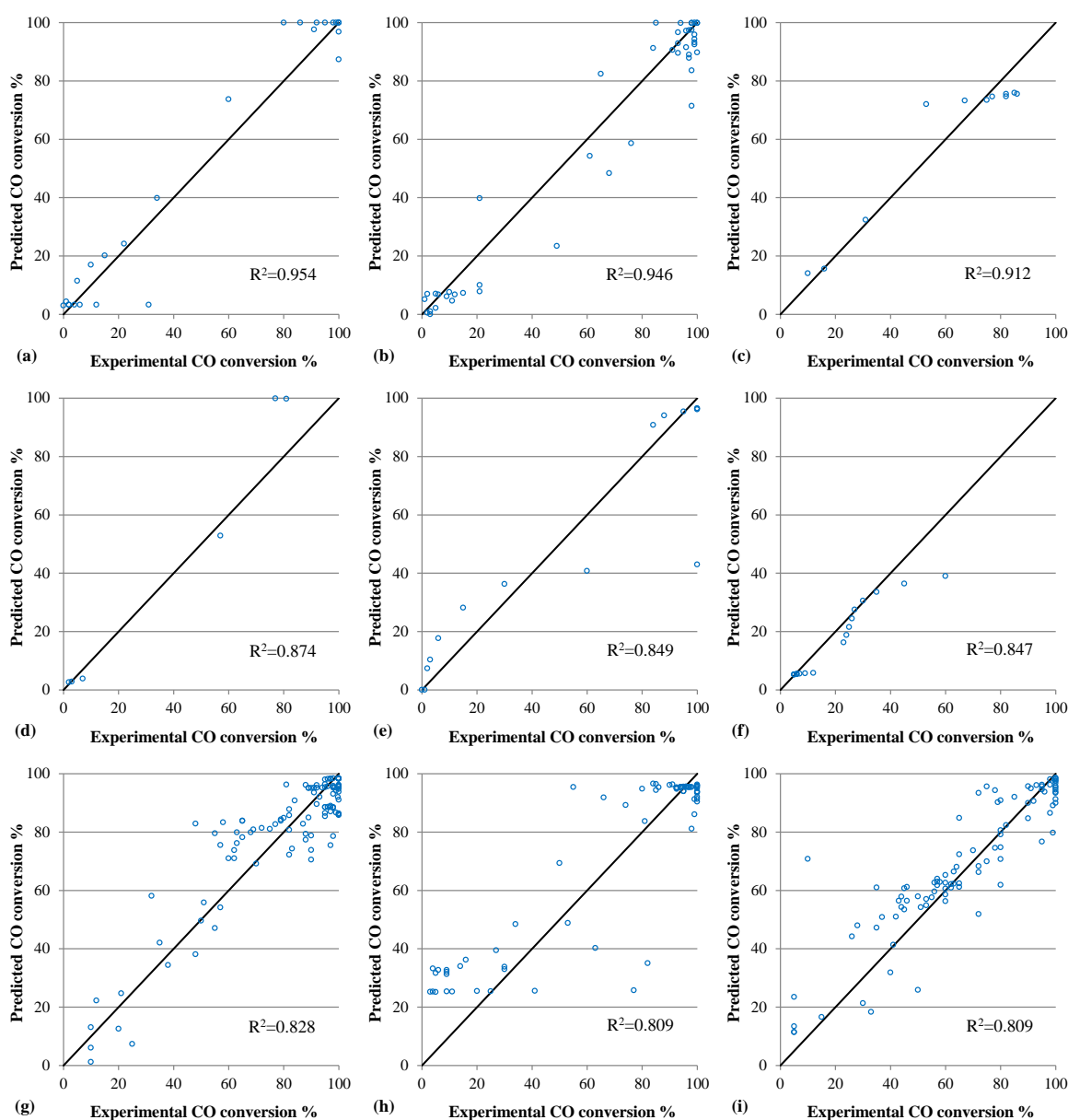
Figure 5.55. Experimental versus predicted CO conversion for references (a) [182], (b) [206], (c) [194], (d) [204], (e) [173], (f) [161], (g) [10], (h) [201], (i) [174].

5.3.2.3. Predicting the Effects of Catalyst Variables. The two-step approach was used to predict the effects of catalyst variables on the catalytic activity. Revealing this information accumulated within the publications is a very important piece of knowledge because an effective search for the best catalyst requires the analysis of each variable on the catalytic activity. The effects of preparation method; calcination conditions; base metal, promoter and support types as well as various operating conditions on catalytic activity was investigated by some sample plots of experimental versus predicted CO conversions. The

symbols in Figures 5.56-5.65 show the experimental data reported in the corresponding publication, and the lines are the model predictions.

The effect of preparation method was investigated through the works of Luengnaruemitchai *et al.* [24] and Ko *et al.* [10] for Au/CeO_2 catalyst (Figure 5.56). Luengnaruemitchai *et al.* compared the activity of Au/CeO_2 catalyst prepared by sol-gel precipitation (SGP) and co-precipitation (CP), and found that the CP-catalyst had higher performance compared to the SGP-catalyst. This was attributed to the fact that catalysts prepared by CP had much smaller crystallite size leading to a higher activity [24]. Similarly, Ko *et al.* compared the performance of the catalysts prepared by CP and deposition precipitation (DP), and reported that the DP-catalyst was more active for a much wider range of temperature due to its relatively smaller crystallite size [10]. All these trends were predicted almost perfectly by the neural networks.
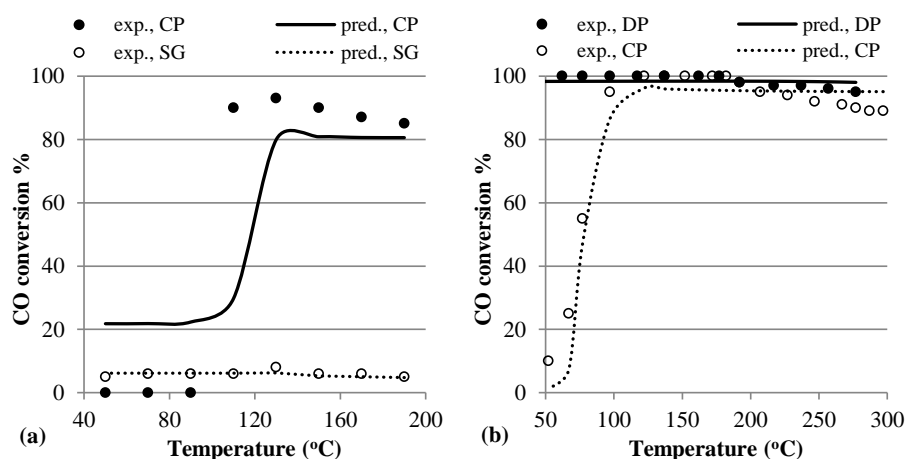
Figure 5.56. Effect of preparation method on CO conversion for Au/CeO_2 (a) 1% CO, 1% O_2, 40% H_2, 2.6% H_2O, 2% CO_2, F/W=30000 [24]; (b) 1% CO, 1% O_2, 10% H_2, 2% H_2O, 0% CO_2, F/W=60000 [10].

The effect of calcination temperature on CO conversion was investigated for Au/FeO catalyst under 50% H_2 containing stream through the work of Qiao *et al.* (Figure 5.57) [197]. They reported that calcination at high temperatures was detrimental for the catalytic activity for this catalyst. The catalyst calcined at $200^{o}C$ were more active in lower temperatures, and in order for the catalyst calcined at $400^{o}C$ to reach to the same activity a

much higher operating temperature was needed. This trend was also successfully predicted by the neural network models.
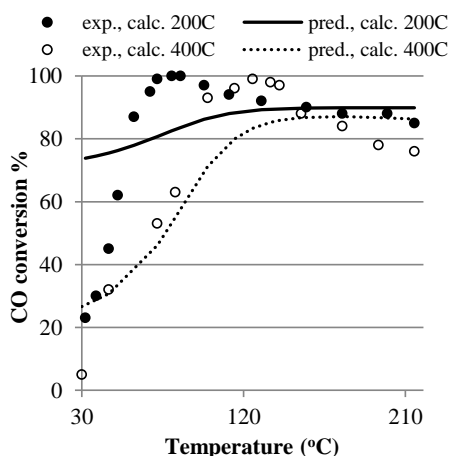
Figure 5.57. Effect of calcination temperature on CO conversion for Au/FeO (1% CO, 4% O_2, 50% H_2, 0% H_2O, 0% CO_2, F/W=20000 [197].

The effect of promoter type on CO conversion was analyzed in Figure 5.58, through the experimental results of 1 wt.% Pt over Al_2O_3 (1% CO, 1% O_2, 80% H_2, 2% H_2O, 0% CO_2, F/W=60000) [10], 1.8 wt.% Pt over Al_2O_3 (0.7% CO, 0.7% O_2, 64% H_2, 10.7% H_2O, 21% CO_2, F/W=10000) [171], 1 wt.% Au over Al_2O_3 (1% CO, 1% O_2, 60% H_2, 0% H_2O, 0% CO_2, F/W=24000) [139] and 0.6 wt.% Au over Al_2O_3 (1% CO, 5% O_2, 0% H_2, 0% H_2O, 0% CO_2, F/W=25500) [161]. Ko *et al.* reported that the addition of Ni or Co metal oxide promoters to the Pt/Al_2O_3 catalyst enhanced the catalytic activity by forming highly dispersed bimetallic phases on the surface [10]. Almost the same conclusion was achieved by Suh *et al.* for the same catalyst and the positive effect of Co was attributed to the fact that Co creates a strong synergetic phase between platinum that also decreases the interaction between Pt and Al_2O_3 [171].

In Figure 5.58c and Figure 5.58d the experimental and the predicted activities of Au/Al_2O_3 catalysts promoted by Mn and Fe are compared with the activities of the un-promoted catalysts. The addition of Mn was reported to increase CO oxidation rate by providing additional active oxygen for the reaction [139]; whereas Fe was reported to enhance the activity by increasing the Au dispersion on the support [161]. Both of these effects were successfully predicted by the neural network models.
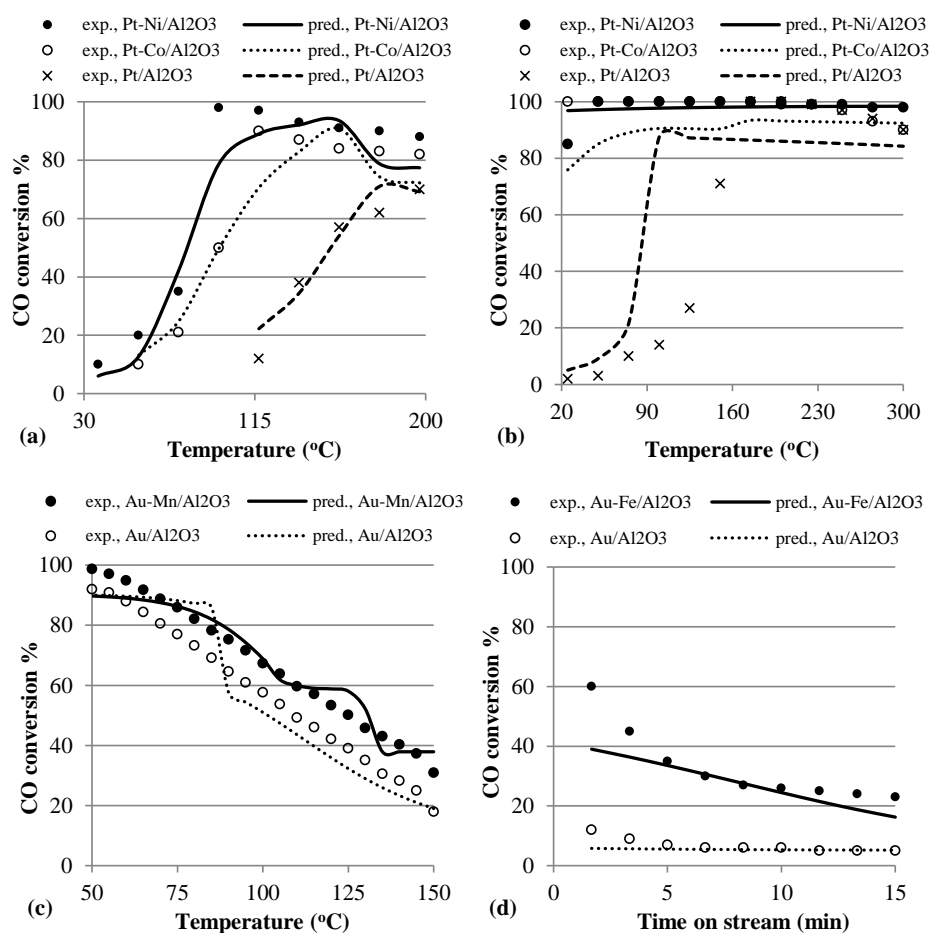
Figure 5.58. Effect of promoter type on CO conversion for (a) 1 wt.% Pt over Al_2O_3 [10]; (b) 1.8 wt.% Pt over Al_2O_3 [171]; (c) 1 wt.% Au over Al_2O_3 [139]; (d) 0.6 wt.% Au over Al_2O_3 at $30^{o}C$ [161].

The catalytic activity of a selective CO oxidation catalyst can further be improved by the addition of a second promoter. In Figure 5.59a the addition of Mg, Fe or Zr to the Pt-Co/Al_2O_3 catalyst (1% CO, 1% O_2, 60% H_2, 0% H_2O, 0% CO_2, F/W=24000) was investigated [138], and in Figure 5.59b the addition of Ce to the Au-Rb/Al_2O_3 catalyst (2.6% CO, 1.3% O_2, 0% H_2, 0% H_2O, 0% CO_2, F/W=12000) was observed [27]. Under the studied operating conditions, the enhancement of the activity of Pt-Co/Al_2O_3 catalyst by the addition of Mg was experimentally found to be superior to the additions of Fe or Zr [138]. On the other hand, the addition of Ce to the Au-Rb/Al_2O_3 catalyst was found to be detrimental in spite of the oxygen supplying property of Ce [27]. The experimental results in Figure 5.59 are quite close to the neural network models, which is a further success of the modeling approach.
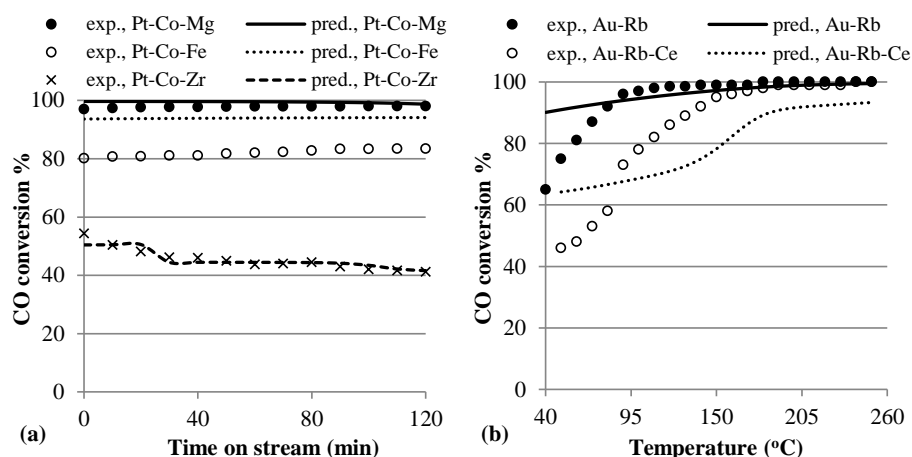
Figure 5.59. Effect of second promoter type on CO conversion for (a) 0.7 wt.% Pt and 1.25 wt.% Co over Al_2O_3 at 110 ^{o}C [138],; (b) 5 wt.% Au and 5 wt.% Rb over Al_2O_3 [27].
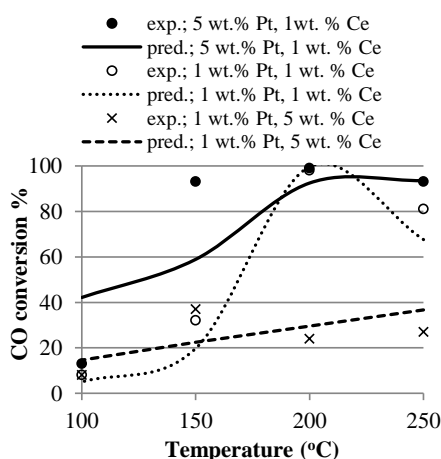
Figure 5.60. Effect of base metal and promoter amounts on CO conversion for Pt-Ce/Al_2O_3 (1% CO, 1% O_2, 98% H_2, 0% H_2O, 0% CO_2, F/W=60000) [13].

Besides the analysis of catalytic activity according to the presence or absence of a base metal or a promoter, the neural network models could also predict the effect of base metal and promoter amount on CO conversion as shown in Figure 5.60 [13]. For Pt-Ce/Al_2O_3 catalyst, two different levels of Pt and Ce in three combinations were compared and it was found that increasing the Pt amount had positive effect on CO conversion, whereas, increasing the Ce amount had detrimental effect, which was attributed to the fact that excess amount of Ce leaded to higher H_2 oxidation decreasing the O_2 amount. The optimum base metal and promoter amounts were reported to be 5 wt.% Pt and 1 wt.% Ce as also successfully predicted by the neural network models.

Among the preparation variables the last to be analyzed is the effect of support type on CO conversion through some example studies. Figure 5.61a compares the experimental and predicted activities of Pt/Al_2O_3 with Pt/CeO_2 catalysts (2% CO, 1% O_2, 70% H_2, 0% H_2O, 0% CO_2, F/W=60000) [168], similarly Figure 5.61b compares Pt/Al_2O_3 with $Pt/SiO_2/Al_2O_3$ (1% CO, 1% O_2, 41%H_2, 0% H_2O, 0% CO_2, F/W=40000) [206], Figure 5.61c compares Au/TiO_2 with Au/Al_2O_3 (1% CO, 1% O_2, 0% H_2, 0% H_2O, 0% CO_2, F/W=60000) [209], Figure 5.61d compares Au/CeO_2 with Au/Fe_2O_3 (1% CO, 1.3% O_2, 50% H_2, 10% H_2O, 15% CO_2, F/W=120000) [211], Figure 5.61e compares Ir/CeO_2 with Ir/Al_2O_3 (2% CO, 1% O_2, 40% H_2, 0% H_2O, 0% CO_2, F/W=40000) [186] and Figure 5.61f compares Co/ZrO_2 with Co/CeO_2 (1% CO, 1% O_2, 60% H_2, 0% H_2O, 0% CO_2, F/W=30000) [199]. Depending on the preparation and operating variables, each support leads to a different catalytic activity as almost perfectly predicted by the neural network models.

Next, the effects of the operating variables are analyzed through some example studies starting from the effect of the presence of H_2 on CO conversion. Pozdnyakova *et al.* investigated this effect on Pt/CeO_2 catalyst (1% CO, 1% O_2, 0% H_2O, 0% CO_2, F/W=73000), and found for the H_2 containing stream that the conversion decreased as the temperature increased due to the consumption of O_2 with H_2 to form H_2O at high temperatures (Figure 5.62a) [174]. A similar trend was observed (Figure 5.62b) by Tibiletti *et al.* for Pt/CeO_2 catalyst (1% CO, 2% O_2, 0% H_2O, 0% CO_2, F/W=22000), and they attributed the negative effect of H_2 to its competitive adsorption with CO to the surface and its direct interaction with ceria lattice oxygen enhancing the H_2O formation [166]. On the other hand, Marques *et al.* investigated the H_2 effect on $Pt-Sn/Al_2O_3$ catalyst (5% CO, 5% O_2, 0% H_2O, 0% CO_2, F/W=34000), and found that the presence of H_2 enhances the reaction by lowering the temperature range required for high activity (Figure 5.62c) [177]. The predictions in Figure 5.62 lie quite close to the experimental values.
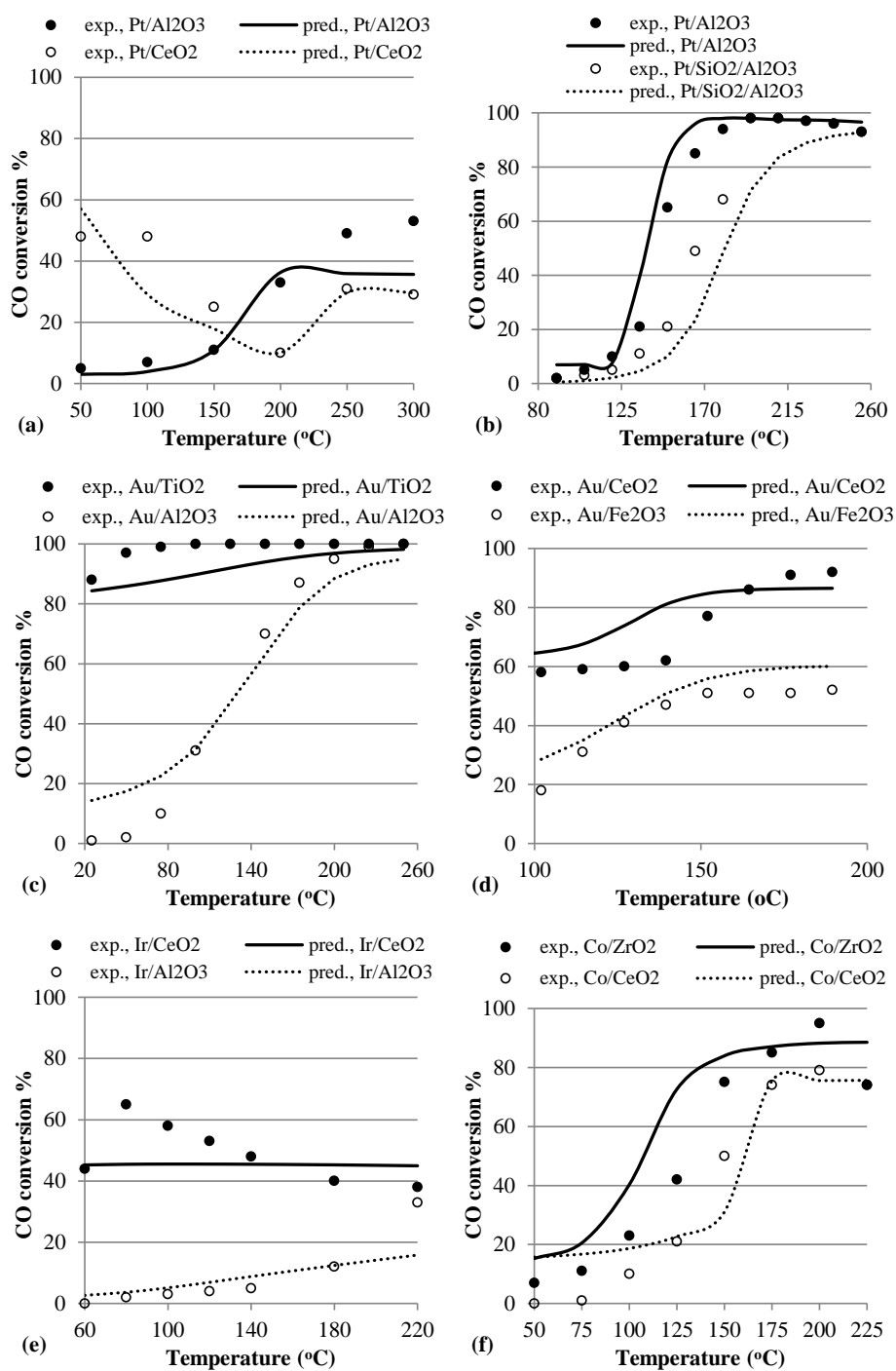
Figure 5.61. Effect of support on CO conversion for (a) Marino *et al* [168], (b) Padilla *et al.* [206], (c) Liotta *et al.* [209], (d) Tabakova et al. [211], (e) Huang *et al.* [186], (f) Zhao *et al.* [199].
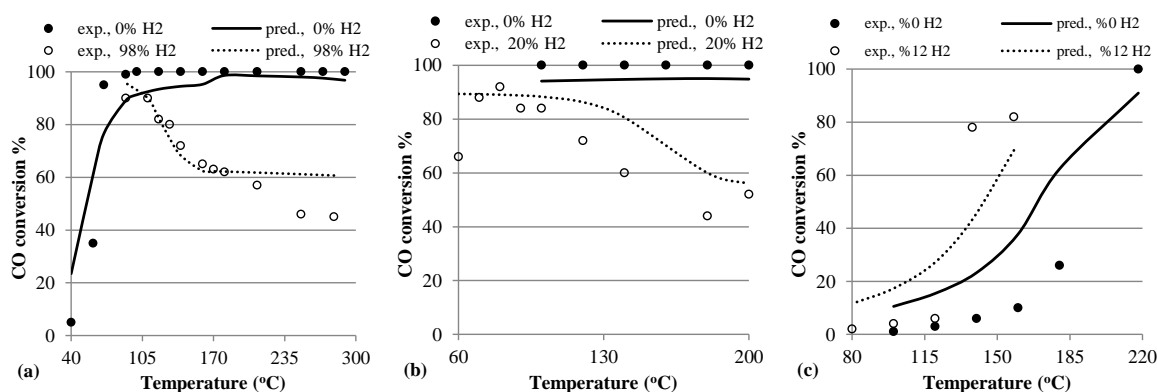
Figure 5.62. Effect of H_2 on CO conversion for (a) 1 wt.% Pt over CeO_2 [174], (b) 4.75 wt.% Pt over CeO_2 [166], (c) 1 wt.% Pt and 1 wt.% Sn over Al_2O_3 [177].

The effect of O_2 on CO conversion is shown in Figure 5.63a [174] and Figure 5.63b [13] for Pt/CeO_2 and Pt/Al_2O_3, respectively. Both plots compare the experimental and predicted conversions attained by 1% O_2 and 0.5% O_2 in an identical reaction stream, and the same conclusions can be predicted: higher O_2 leads to higher conversion, which are in accordance with the experimental results.
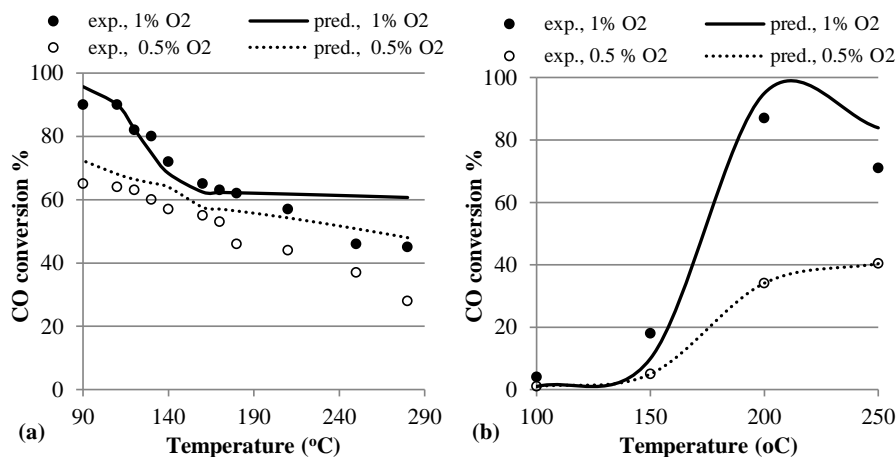
Figure 5.63. Effect of O_2 on CO conversion for (a) 1 wt.% Pt over CeO_2 (1% CO, balance H_2, F/W=73000) [174]; (b) 1 wt.% Pt over Al_2O_3 (1% CO, balance H_2, F/W=60000) [13].

Since, the reaction stream used for selective CO oxidation also contains some amount of H_2O and CO_2, the effects of these variables on CO conversion are frequently analyzed in the literature. The detrimental effect of CO_2 on CO conversion is observed in Figure 5.64a for $Pt-Co/Al_2O_3$ [138] and in Figure 5.64b for Au/CeO_2 [24]. The negative effect of CO_2 was attributed to the competitive adsorption of CO_2 on Pt surface decreasing

the rate of CO adsorption, and to the increase of the reverse water gas shift reaction rate due to the presence of CO_2 [138]. The neural network predictions shown in Figure 5.64 successfully coincide with the experimental findings.
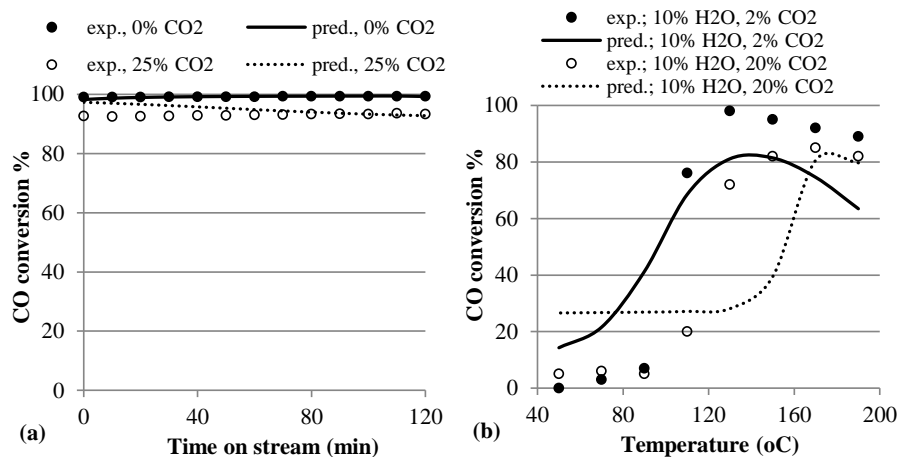
Figure 5.64. Effect of H_2O and CO_2 on CO conversion for (a) 0.7 wt.% Pt and 1.25 wt.% Co over Al_2O_3 at 110 ^oC (1% CO, 1% O_2, 60% H_2, F/W=24000) [138]; (b) 1 wt.% Au over CeO_2 (1% CO, 1% O_2, 40% H_2, F/W=30000) [24].
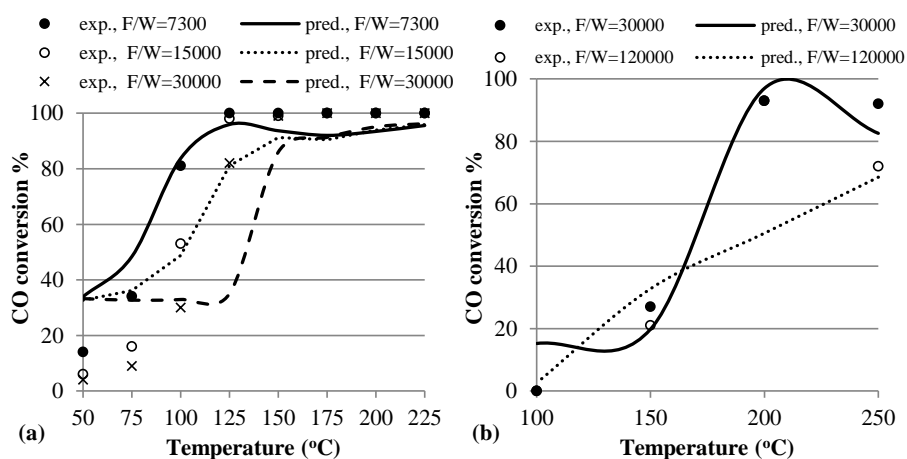
Figure 5.65. Effect of F/W on CO conversion for (a) 10 wt.% Co over ZrO_2 (0.5% CO, 0.5% O_2, 5% H_2, 0% H_2O, 0% CO_2) [201]; (b) 1 wt.% Pt over Al_2O_3 (1% CO, 1% O_2, 98% H_2, 0% H_2O, 0% CO_2) [13].

Among the operating variables the last to be analyzed is the effect of F/W on CO conversion. In Figure 5.65a three different levels of F/W are compared for Co/ZrO_2 [201] and in Figure 5.65b two different levels of F/W are compared for Pt/Al_2O_3 [13]. As the

ratio of the feed flow rate to the catalyst weight increases, the CO conversion decreases expectedly.

Table 5.21. Relative input significances for noble metal-gold based catalysts.

Variable	RMSE without the Variable	RMSE Difference*	% Relative Significance (change of RMSE)	% Relative Significance (partial differentiation)	(%) Group Significance (change of RMSE)
promoter type and amount	13.73	6.05	43.5		
base metal type and amount	10.82	3.14	22.6		
support type	10.74	3.06	22.0		43.4
preparation method	8.93	1.26	9.0		
calcination conditions	8.08	0.41	2.9		
reaction temperature				20.3	
F/W				14.7	
O_2 vol.%				14.3	
CO vol.%				14.0	
H_2 vol.%				13.5	56.6
CO_2 vol.%				8.5	
Time on stream				8.4	
H_2O vol.%				6.3	

*the difference between RMSE without the variable or variables and RMSE of the original model (7.67)

5.3.2.4. Significance Analysis for Catalyst Variables. In this section, the relative significances of the preparation and operating variables on CO conversion are discussed. The significances of the preparation variables from the most significant to the lowest one were determined as the promoter type and amount, base metal type and amount, support type, preparation method and the calcination conditions as shown in Table 5.21. It should be noted that the calculated relative significance values indicate the CO conversion influenced by the change of a variable from its minimum to maximum value. This is the reason why the base metal type and amount seems to have lower significance level than the promoter type and amount. The relative significance of the promoter type includes the effect of the presence or absence of one promoter or two promoters at the same time, whereas the significance of the base metal is the sensitivity of CO conversion to the change of base metals between their minimum and maximum. The relative significances of the

operating variables from the most significant to the lowest one were determined as the reaction temperature, F/W, amounts of O_2, CO and H_2 in the feed. Each of the preparation and operating variables has some part of contribution to the catalytic activity and according to the group significance analysis both preparation and operating variables were found to be influential on the activity of the catalysts.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

The objective of this dissertation was to apply several data mining techniques on catalysis data for knowledge extraction to improve the catalyst design and testing conditions. This thesis consists of three main studies: analysis of the experimental data produced in our laboratories, the data generated by the density functional theory and the data mined out from the publications in the literature. The major conclusions from these works were summarized below.

The experimental data for selective CO oxidation over promoted Pt-Co/Al_2O_3 catalyst were analyzed using modular neural networks. The optimal network topology was used to analyze the relative significances of the preparation and the operating variables, and their effects on CO conversion. The network predictions were in a perfect agreement with the experimental results as well as the findings in the literature as summarized below:

- Higher Pt content provided higher CO conversion.
- CO conversion decreased significantly with increasing temperature from 110 ^{o}C to 130 ^{o}C over all the catalysts having a second promoter.
- The addition of Mg increased the activity of Pt-Co/Al_2O_3 most while Mn and Ce also had some positive effects. The addition of Fe on the other hand, had a significant negative effect on CO conversion.
- The presence of CO_2 had detrimental influence while H_2O enhanced the catalytic activity.

The experimental data of CO oxidation over promoted Au/Al_2O_3 was analyzed by decision tree classification and neural network modeling. It was found from the decision tree analysis that the catalysts promoted by Mg or Mn always exhibited higher CO conversions compared to the catalysts with the other promoters, and low temperatures leaded to higher CO conversions. Then, the reduced dataset containing only the promising

data (excluding the data related to the promoters: Ce, Co, Fe and Ni) were modeled using modular neural networks. The effects of catalyst preparation and operating variables over CO conversion were successfully predicted as given below:

- The addition of both Mg and Mn improved the catalytic activity while the positive effect of Mg was higher especially in the presence of H_2O and CO_2.
- The operating variables (especially reaction temperature) were found to be relatively more significant for catalytic activity than the catalyst preparation variables.

The effects of second promoter (K, Co and Ni) on the water gas shift activity of Pt-CeO_2/Al_2O_3 catalyst were analyzed computationally using modular neural networks. The neural network model successfully predicted the effects of promoter type and operating variables and the conclusions can be summarized as follows:

- K had some beneficial effects under product-containing feed compositions.
- Co and Ni promoters worsened the catalyst performance.
- The reaction temperature and feed H_2O/CO ratio positively affected the catalytic activity
- CO_2 and H_2 addition to the feed decreased CO conversion.
- H_2 concentration and H_2O/CO ratio in the feed were the most significant variables.

These results suggested that modular neural networks may be used to analyze similar systems to capture the effects and significances of reaction variables, and they may improve and expedite experimental studies. Moreover, the decision trees and neural networks can complement each other to extract easily comprehensible information from the experimental data.

The structure and activity relationship was analyzed for CO and O_2 adsorptions over Au_2-Au_{10} clusters using DFT and artificial neural networks. It was demonstrated that the structural properties of the catalyst clusters can be successfully predicted from the user defined descriptors by using artificial neural networks trained by the data generated with DFT. Then, these properties can be used to predict the adsorption energies of CO and O_2.

The results showed that the use of only a few structural properties is sufficient to predict the adsorption energies. The conclusions can be summarized as follows:

- The adsorption strength of CO followed the order: cationic > neutral > anionic. The presence of the unpaired electron was found to have relatively less effect on the CO adsorption energy.
- The size of the cluster was found to be the most important variable for the binding energy and the degree of the binding energy followed the order: cationic > anionic > neutral.
- For the HOMO-LUMO gap, the most significant variable was found to be the unpaired electron while the charge and the size of the cluster had relatively small effects.
- The charge of the cluster was the most dominating factor for both the ionization potential and the electron affinity.
- The HOMO-LUMO gap and the coordination number together with any one of the binding energy, electron affinity or ionization potential carried the significant part of the necessary information to predict the CO adsorption energy.
- Unpaired electron was the most significant descriptor affecting the O_2 adsorption energy.
- The HOMO-LUMO gap and the ionization potential were the most effective structural properties in determination of the adsorption strength of O_2

The effects of user defined descriptors and structural properties of the Au clusters on the stability and strength of the O_2 adsorption were also analyzed by classifying the DFT computed adsorption data using multiple logistic regression. It was found that:

- Among the user defined descriptors, unpaired electron was the most significant variable for the stability of O_2 adsorption, whereas the size and charge of the cluster determined the adsorption strength.
- The probability of finding strong adsorption decreased with increasing cluster size and changing the charge of the cluster in the order of anionic-neutral-cationic.

- For the structural properties, the stability of the O_2 adsorption depended on HOMO-LUMO gap; probability of stable adsorption being smaller for the clusters with high HOMO-LUMO gaps.
- The strength of the adsorption, on the other hand, was mostly determined by the ionization potential.

These results suggested that multiple logistic regression modeling can successfully classify the DFT computed data, and the combination of neural networks with DFT may improve the computational studies in catalysis by reducing the computational time and efforts as well as by helping to extract useful knowledge to understand the real catalytic systems.

Finally, two databases were constructed from the data mined out from the publications on selective CO oxidation: Cu based catalysts and noble metal-gold based catalysts. The database of Cu based catalysts was formed from twenty publications in the literature (containing 1337 data points), and the experimental CO conversions reported in each publication were successfully predicted by a neural network trained using the data from the remaining nineteen papers. It was found that neural network models created by this way were quite successful in predicting the experimental results of a new set of experimental conditions. The models were more effective to predict the results of the publications that contained the most commonly studied variables within the commonly studied ranges. However, the predictions were also quite acceptable even for the rarely studied variables unless an experimental dataset held unique information. The relative significances of the main catalyst preparation and operating variables as well as the major trends associated with these variables were also successfully estimated.

The database of noble metal-gold based catalysts was constructed from seventy one publications in the literature (containing 5008 data points), and genetic algorithm based clustering together with artificial neural network modeling was implemented for the prediction and the analysis of the data. First, a reduced dataset (containing 4070 data points from 58 publications) was formed by eliminating the unique data from the database. Then, a two-step approach, composed of clustering the data and then neural network modeling for each cluster, was applied for knowledge extraction. In contrast to the direct neural

network modeling, the two-step approach was superior in predicting the outcomes of the individual publications. Only the results of 5/58 publications were predicted with R^2_{test} values higher than 0.5 by the direct neural network approach while 32/58 publications were predicted with R^2_{test} values higher than 0.5 by the two-step approach. Several major trends associated with the catalytic variables were extracted by the two-step model, and they were in a very good accordance with the results in the publications. According to the input significance analysis, both the preparation and the operating variables were found to be influential on the activity of the catalysts. The significances of the preparation variables from the most significant to the lowest one were determined as the promoter type and amount, base metal type and amount, support type, preparation method and the calcination conditions while those of the operating variables were in the order of reaction temperature, F/W, amounts of O_2, CO and H_2 in the feed.

As a result of these findings, it can be concluded that genetic algorithm based clustering together with neural networks can be used in catalytic systems to extract the essential knowledge and experience accumulated in the published experimental data in a very accurate manner, which may help the researchers in planning their future experimental work more effectively.

6.2. Recommendations

Based on the results obtained in this study the following recommendations will be beneficial for future studies aiming for knowledge extraction from catalytic data:

- Different modeling techniques may be employed such as support vectors or genetic programming for knowledge extraction from the catalytic data, and their performances may be compared with the methods presented in this work.
- Analysis of the published data by using data mining techniques is a very promising study. This may also be implemented for the data of similar catalytic reactions (i.e. water gas shift reaction) or for other areas such as to model the reaction rate from the operating variables to acquire the reaction mechanisms and the kinetic parameters.

REFERENCES

1. Trimm, D. L. and Z. I. Onsan, "Onboard Fuel Conversion for Hydrogen-Fuel-Cell-Driven Vehicles", *Catalysis Reviews*, Vol. 43, No. 1-2, pp. 31-84, 2001.

2. Choudhary, T. V. and D. W. Goodman, "CO-free fuel processing for fuel cell applications", *Catalysis Today*, Vol. 77, No. 1-2, pp. 65-78, 2002.

3. Panagiotopoulou, P. and D. I. Kondarides, "Effect of the nature of the support on the catalytic performance of noble metal catalysts for the water–gas shift reaction", *Catalysis Today*, Vol. 112, No. 1-4, pp. 49-52, 2006.

4. Lim, S., J. Bae and K. Kim, "Study of activity and effectiveness factor of noble metal catalysts for water–gas shift reaction", *International Journal of Hydrogen Energy*, Vol. 34, No. 2, pp. 870-876, 2009.

5. Yeragi, D. C., N. C. Pradhan and A. K. Dala, "Low-temperature water-gas shift reaction over Mn-promoted Cu/Al_2O_3 catalysts", *Catalysis Letters*, Vol. 112, No. 3-4, pp. 139-148, 2006.

6. Li, Y., Q. Fu and M. Flytzani-Stephanopoulos, "Low-temperature water-gas shift reaction over Cu- and Ni- loaded cerium oxide catalysts", Vol. 27, No. 3, pp. 179-191, 2000.

7. Kim, S. H, J. H. Chung, Y. T. Kim, J. Han, S. P. Yoon, S. W. Nam, T. H. Lim and H. I. Lee, "SiO_2/Ni and CeO_2/Ni catalysts for single-stage water gas shift reaction", *International Journal of Hydrogen Energy*, Vol. 35, No. 7, pp. 3136-3140, 2010.

8. Tabakova, T., V. Idakiev, D. Andreeva and I. Mitov, "Influence of the microscopic properties of the support on the catalytic activity of Au/ZnO, Au/ZrO$_2$, Au/Fe$_2$O$_3$, Au/Fe$_2$O$_3$–ZnO, Au/Fe$_2$O$_3$–ZrO$_2$ catalysts for the WGS reaction", *Applied Catalysis A*, Vol. 202, No. 1, pp. 91-97, 2000.

9. Yahiro, H., K. Murawaki, K. Saiki, T. Yamamoto and H. Yamaura, "Study on the supported Cu-based catalysts for the low-temperature water–gas shift reaction", *Catalysis Today*, Vol. 126, No. 3-4, pp. 436-440, 2007.

10. Ko, E. Y., E. D. Park, K. W. Seo, H. C. Lee, D. Lee and S. Kim, "A comparative study of catalysts for the preferential CO oxidation in excess hydrogen", *Catalyis Today*, Vol. 116, No. 3, pp. 377-383, 2006.

11. Thormahlen, P., M. Skoglundh, E. Fridell and B. Andersson, "Low-Temperature CO Oxidation over Platinum and Cobalt Oxide Catalysts", *Journal of Catalysis*, Vol. 188, No. 2, pp. 300-310, 1999.

12. Törncrona, A., M. Skoglundh, P. Thormahlen, E. Fridell and B. Andersson, "Low temperature catalytic activity of cobalt oxide and ceria promoted Pt and Pd: influence of pretreatment and gas composition", *Applied Catalysis B*, Vol. 14, No. 1-2, pp. 131-145, 1997.

13. Son, I. H., "Study of Ce-Pt/γ-Al$_2$O$_3$ for the selective oxidation of CO in H$_2$ for application to PEFCs: Effect of gases", *Journal of Power Sources*, Vol. 159, No. 3, pp. 1266-1273, 2006.

14. Cho, S. H., J. S. Park, S. H. Choi and S. H. Kim, "Effect of magnesium on preferential oxidation of carbon monoxide on platinum catalyst in hydrogen-rich stream", *Journal of Power Sources*, Vol. 156, No. 2, pp. 260-266, 2006.

15. Ayastuy, J. L., M. P. Gonzalez-Marcos, J. R. Gonzalez-Velasco and M. A. Gutierrez-Ortiz, "$MnO_x/Pt/Al_2O_3$ catalysts for CO oxidation in H_2-rich streams", *Applied Catalysis B*, Vol. 70, No. 1-4, pp. 532-541, 2007.

16. Sirijaruphan, A., J. G. Goodwin Jr., R. W. Rice, D. Wei, K. R. Butcher, G. W. Roberts and J. J. Spivey, "Effect of metal foam supports on the selective oxidation of CO on Fe-promoted Pt/γ-Al_2O_3", *Applied Catalysis A*, Vol. 281, No. 1-2, pp. 11-18, 2005.

17. Chin, P., X. Sun, G. W. Roberts and J. J. Spivey, "Preferential oxidation of carbon monoxide with iron-promoted platinum catalysts supported on metal foams", *Applied Catalysis A*, Vol. 302, No. 1, pp. 22-31, 2006.

18. Ko, E. Y., E. D. Park, K. W. Seo, H. C. Lee, D. Lee and S. Kim, "Pt–Ni/γ-Al_2O_3 catalyst for the preferential CO oxidation in the hydrogen stream", *Catalysis Letters*, Vol. 110, No. 3-4, pp. 275-279, 2006.

19. Minemura, Y., M. Kuriyama, S. Ito, K. Tomishige and K. Kunimori, "Additive effect of alkali metal ions on preferential CO oxidation over Pt/Al_2O_3", *Catalysis Communications*, Vol. 7, No. 9, pp. 623-626, 2006.

20. Ayastuy, J. L., M. P. Gonzalez-Marcos, A. Gil-Rodriguez, J. R. Gonzalez-Velasco and M. A. Gutierrez-Ortiz, "Selective CO oxidation over $Ce_XZr_{1-X}O_2$-supported Pt catalysts", *Catalysis Today*, Vol. 116, No. 3, pp. 391-399, 2006.

21. Schubert, M. M., A. Venugopal, M. J. Kahlich, V. Plzak and R. J. Behm, "Influence of H_2O and CO_2 on the selective CO oxidation in H_2-rich gases over Au/α-Fe_2O_3", *Journal of Catalysis*, Vol. 222, No. 1, pp. 32-40, 2004.

22. Luengnaruemitchai, A., D. T. K. Thoa, S. Osuwan and E. Gulari, " A comparative study of Au/MnO$_x$ and Au/FeO$_x$ catalysts for the catalytic oxidation of CO in hydrogen rich stream ", *International Journal of Hydrogen Energy*, Vol. 30, No. 9, pp. 981-987, 2005.

23. Schubert, M. M., S. Hackenberg, A. C. van Veen, M. Muhler, V. Plzak and R.J Behm, "CO Oxidation over Supported Gold Catalysts "Inert" and "Active" Support Materials and Their Role for the Oxygen Supply during Reaction", *Journal of Catalysis*, Vol. 197, No. 1, pp. 113-122, 2001.

24. Luengnaruemitchai, A., S. Osuwan and E. Gulari, "Selective catalytic oxidation of CO in the presence of H$_2$ over gold catalyst", *International Journal of Hydrogen Energy*, Vol. 29, No. 4, pp. 429-435, 2004.

25. Tomska-Foralewska, I., W. Przystajko, M. Pietrowski, M. Zielinski and M. Wojciechowska, "Effect of MgO content in the support of Au/MgF$_2$–MgO catalysts on CO oxidation", *Reaction Kinetics Mechanisms and Catalysis*, Vol. 100, No. 1, pp. 111-121, 2010.

26. Grisel, R. J. H., C. J. Weststrate, A. Goossens, M. W. J Crajé, A. M. van der Kraan and B. E. Nieuwenhuys, "Oxidation of CO over Au/MO$_x$/Al$_2$O$_3$ multi-component catalysts in a hydrogen-rich environment", *Catalysis Today*, Vol. 72, No. 1-2, pp. 123-132, 2002.

27. Gluhoi, A. C. and B. E. Nieuwenhuys, "Structural and chemical promoter effects of alkali (earth) and cerium oxides in CO oxidation on supported gold", *Catalysis Today*, Vol. 122, No. 3-4, pp. 226-232, 2007.

28. Zou, H., Z. Dong and W. Lin, "Selective CO oxidation in hydrogen-rich gas over CuO/CeO$_2$ catalysts", *Applied Surface Science*, Vol. 253, No. 5, pp. 2893-2898, 2006.

29. Zhu, P., J. Li, S. Zuo and R. Zhou, "Preferential oxidation properties of CO in excess hydrogen over CuO-CeO$_2$ catalyst prepared by hydrothermal method", *Applied Surface Science*, Vol. 225, No. 5, pp. 2903-2909, 2008.

30. Liu, Y., Q. Fu and M. F. Stephanopoulos, "Preferential oxidation of CO in H$_2$ over CuO-CeO$_2$ catalysts", *Catalysis Today*, Vol. 93, pp. 241-246, 2004.

31. Larose, D. T., *Discovering Knowledge in Data*, Wiley, New Jersey, 2005.

32. Istadi, and N. A. S. Amin, "A hybrid numerical approach for multi-responses optimization of process parameters and catalyst compositions in CO$_2$ OCM process over CaO-MnO/CeO$_2$ catalyst", *Chemical Engineering Journal*, Vol. 106, No. 3, pp. 213-227, 2005.

33. Du, G., Y. Yang, W. Qiu, S. Lim, L. Pfefferle and G. L. Haller, "Statistical design and modeling of the process of methane partial oxidation using V-MCM-41 catalysts and the prediction of the formaldehyde production", *Applied Catalysis A*, Vol. 313, No. 1, pp. 1-13, 2006.

34. Günay, M. E. and R. Yildirim, "Neural network aided design of Pt-Co-Ce/Al$_2$O$_3$ catalyst for selective CO oxidation in hydrogen-rich streams", *Chemical Engineering Journal*, Vol. 140, No. 1-3, pp. 324-331, 2008.

35. Olutoye, M. A. and B. H., Hameed, "Transesterification of palm oil on K$_y$Mg$_{1-x}$Zn$_{1+x}$O$_3$ catalyst: Effect of Mg–Zn interaction", *Fuel Processing Technology*, Vol. 91, No. 6, pp. 653–659, 2010.

36. Chen, G., J. Gao, L. Xu, X. Fu, Y. Yin, S. Wu and Y. Qin, *Optimizing conditions for preparation of MnO$_x$/RHA catalyst particle for the catalytic oxidation of NO*, 2011 http://dx.doi.org/doi:10.1016/j.apt.2011.03.003, accessed at January 2012.

37. Callan, R., *The Essence of Neural Networks*, Prentice Hall, Hertfordshire, 1999.

38. Burns, J. A. and G. M. Whitesides, "Feed-Forward Neural Networks in Chemistry: Mathematical Systems for Classification and Pattern Recognition", *Chemical Reviews*, Vol. 93, No. 8, pp. 2583-2599, 1993.

39. Himmelblau, D. M., "Applications of Artificial Neural Networks in Chemical Engineering", *Korean Journal of Chemical Engineering*, Vol. 17, No. 4, pp. 373-392, 2000.

40. Baerns, M. and M. Holena, *Combinatorial Development of Solid Catalytic Materials,* Imperial College Press, London, 2009.

41. Hattori, T. and S. Kito, "Neural network as a tool for catalyst development", *Catalysis Today*, Vol. 23, No. 4, pp. 347-355, 1995.

42. Kobayashi, Y.,Omata, K. and M. Yamada, "Screening of Additives to a $Co/SrCO_3$ Catalyst by Artificial Neural Network for Preferential Oxidation of CO in Excess H_2", *Industrial & Engineering Chemistry Research*, Vol. 49, No. 4, pp. 1541-1549, 2010.

43. Holena, M. and M. Baerns, "Feedforward neural networks in catalysis A tool for the approximation of the dependency of yield on catalyst composition, and for knowledge extraction", *Catalysis Today*, Vol. 81, No. 3, pp. 485-494, 2003.

44. Rodermerck, U., M. Baerns, M. Holena and D. Wolf, "Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials", *Applied Surface Science*, Vol. 223, No. 1-3, pp. 168-174, 2004.

45. Tompos, A., J. L. Margitfalvi, E. Tfirst, L. Végvari, M. A. Jaloull, H. A. Khalfalla and M. M. Elgarni, "Development of catalyst libraries for total oxidation of methane A case study for combined application of holographic research strategy and artificial neural networks in catalyst library design", *Applied Cataysis A*, Vol. 285, No. 1-2, pp. 65-78, 2005.

46. Corma, A., J. M. Serra, P. Serna, S. Valero, E. Argente and V. Botti, "Optimization of olefin epoxidation catalysts with the application of high-throughput and genetic algorithms assisted by artificial neural networks (softcomputing techniques)", *Journal of Catalysis,* Vol. 229, No. 2, pp. 513-524, 2005.

47. Kito, S. and T. Hattori, "Analysis of catalytic performance by partial differentiation of neural network pattern", *Chemical Engineering Science*, Vol. 62, No. 18-20, pp. 5575-5578, 2007.

48. Hattori, T. and S. Kito, "Partial differentiation of neural network for the analysis of factors controlling catalytic activity", *Applied Catalysis A*, Vol. 327, No. 2, pp. 157-163, 2007.

49. Song, S., A. J. Akande, R. O. Idem and N. Mahinpey, "Inter-relationship between preparation methods, nickel loading, characteristics and performance in the reforming of crude ethanol over Ni/Al_2O_3 catalysts: A neural network approach", *Engineering Applications of Artificial Intelligence*, Vol. 20, No. 2, pp. 261-271, 2007.

50. Günay, M. E., F. Akpınar, Z. I. Önsan and R. Yıldırım, "Investigation of water-gas shift activity of $Pt-MO_x-CeO_2/Al_2O_3$ (M=K, Ni, Co) using modular artificial neural networks", *International Journal of Hydrogen Energy*, Vol. 37, No. 3, pp. 2094-2102, 2012.

51. Günay, M. E. and R. Yildirim, "Analysis of selective CO oxidation over promoted Pt/Al$_2$O$_3$ catalysts using modular neural networks: Combining preparation and operational variables", *Applied Catalysis A*, Vol. 377, No. 1-2, pp. 174-180, 2010.

52. Davran-Candan, T., M. E. Günay and R. Yıldırım, "Structure and activity relationship for CO and O$_2$ adsorption over gold nanoparticles using density functional theory and artificial neural networks", *Journal of Chemical Physics*, Vol. 132 , pp. 174113-1-9, 2010.

53. Günay, M. E. and R. Yıldırım, "Neural network Analysis of Selective CO Oxidation over Copper-Based Catalysts for Knowledge Extraction from Published Data in the Literature", *Industrial & Engineering Chemistry Research*, Vol. 50, pp. 12488-12500, 2011.

54. Melin, P., A. Mancilla, M. Lopez and O. Mendoza, "A hybrid modular neural network architecture with fuzzy Sugeno integration for time series forecasting", *Applied Soft Computing*, Vol. 7, No. 4, pp. 1217-1226, 2007.

55. Kang, S. and C. Işık, "Partially Connected Feedforward Neural Networks Structured by Input Types", *IEEE Transactions on Neural Networks*, Vol. 16, No. 1, pp. 175-184, 2005.

56. Larose, D. T., *Data Mining Methods and Models*, Wiley, New Jersey, 2006.

57. Kutner, M. H., C. J. Nachtsheim and J. Neter, *Applied Linear Regression Models*, McGraw Hill, New York, 2004.

58. Hosmer, D. W. and S. Lemeshow, *Applied Logistic Regression*, Wiley, New York, 2000.

59. Salter, M. A., D. A. Ratkowsky, T. Ross and T. A. McMeekin, "Modelling the combined temperature and salt (NaCl) limits for growth of a pathogenic Escherichia coli strain using nonlinear logistic regression", *International Journal of Food Microbiology*, Vol. 61, No. 2-3, pp. 159-167, 2000.

60. Gysemans, K. P. M., K. Bernaerts, A. Vermeulen, A. H. Geeraerd, J. Debevere, F. Devlieghere and J. F. Van Impe, "Exploring the performance of logistic regression model types on growth/no growth data of Listeria monocytogenes", *International Journal of Food Microbiology*, Vol. 114, No. 3, pp. 316-331, 2007.

61. Chang, C. L. and M. Y. Hsu, "The study that applies artificial intelligence and logistic regression for assistance in differential diagnostic of pancreatic cancer", *Expert Systems With Applications*, Vol. 36, No. 7, pp. 10663-10672, 2009.

62. Kurt, I., M. Ture and A. T. Kurum, "Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease", *Expert System With Applications*, Vol. 34, No. 1, pp. 366-374, 2008.

63. Jalkanen, A. and U. Mattila, "Logistic regression models for wind and snow damage in northern Finland based on the National Forest Inventory data", *Forest Ecology and Management*, Vol. 135, No. 1-3, pp. 315-330, 2000.

64. Papritz, A. and P. U. Reichard, "Modelling the risk of Pb and PAH intervention value exceedance in allotment soils by robust logistic regression", *Environmental Pollution*, No. 7, Vol. 157, pp. 2019-2022, 2009.

65. Corma, A., J. M. Serra, P. Serna and M. Moliner, "Integrating high-throughput characterization into combinatorial heterogeneous catalysis: unsupervised construction of quantitative structure/property relationship models", *Journal of Catalysis*, Vol. 232, No. 2, pp. 335-341, 2005.

66. Breiman, L., *Classification and regression trees*, Chapman & Hall, New York, 1993.

67. Cukic, T., R. Kraehnert, M. Holena, D. Herein, D. Linke and U. Dingerdissen, "The influence of preparation variables on the performance of Pd/Al_2O_3 catalyst in the hydrogenation of 1,3-butadiene: Building a basis for reproducible catalyst synthesis", *Applied Catalysis A*, Vol. 323, pp. 25-37, 2007.

68. Moehmel, S., N. Steinfeldt, S. Engelschalt, M. Holena, S. Kolf, M. Baerns, U. Dingerdissen, D. Wolf, R. Weber and M. Bewersdorf, "New catalytic materials for the high-temperature synthesis of hydrocyanic acid from methane and ammonia by high-throughput approach", *Applied Catalysis A*, Vol. 334, No. 1-2, pp. 73-83, 2008.

69. Maulik, U and S Bandyopadhyay, "Genetic Algorithm Based Clustering Technique", *Pattern Recognition*, Vol. 33, No. 9, pp. 1455-1465, 2000.

70. Chang, D. X, X. D. Zhang and C. W. Zheng, "A genetic algorithm with gene rearrangement for K-means clustering", *Pattern Recognition*, Vol. 42, No. 7, pp. 1210-1222, 2009.

71. Docter, A. and A. Lamm, "Gasoline fuel cell systems", *Journal of Power Sources*, Vol. 84, No. 2, pp. 194-200, 1999.

72. Ruettinger, W., Liu X. and R. J. Farrauto, "Mechanism of aging for a Pt/CeO_2-ZrO_2 water gas shift catalyst", *Applied Catalysis B*, Vol. 65, No. 1-2, pp. 135-141, 2006.

73. Zalc, J. M. and D. G. Loffler, "Fuel processing for PEM fuel cells: transport and kinetic issues of system design", *Journal of Power Sources*, Vol. 111, No. 1, pp. 58-64, 2002.

74. Wilhelm, D. J., D. R. Simbeck, A. D. Karp and R. L. Dickenson, "Syngas production for gas-to-liquids applications: technologies, issues and outlook", *Fuel Processing Technology*, Vol. 71, No. 1-3, pp. 139-148, 2001.

75. Jacobs, G., M. P. Patterson, L. Williams, E. Chenu, D. Sparks, G. Thomas and B. H. Davis, "Water-gas shift: in situ spectroscopic studies of noble metal promoted ceria catalysts for CO removal in fuel cell reformers and mechanistic implications", *Applied Catalysis A*, Vol. 262, No. 2, pp. 177-187, 2004.

76. Park, E. D., D. Lee and H.C. Lee, "Recent progress in selective CO removal in a H_2-rich stream", *Catalysis Today*, Vol. 139, No. 4, pp. 280-290, 2009.

77. Onsan, Z. I., "Catalytic processes for clean hydrogen production from hydrocarbons", *Turkish Journal of Chemistry*, Vol. 31, pp. 531-550, 2007.

78. Wootsch, A., C. Descorme, S. Rousselet, D. Duprez and C. Templier, "Carbon monoxide oxidation over well-defined Pt/ZrO_2 model catalysts: Bridging the material gap", *Applied Surface Science*, Vol. 253, No. 3, pp. 1310-1322, 2006.

79. H., Igarashi., H. Uchida, M. Suzuki, Y. Sasaki and M. Watanabe, "Removal of carbon monoxide from hydrogen-rich fuels by selective oxidation over platinum catalyst supported on zeolite", *Applied Catalysis A*, Vol. 159, No. 1-2, pp. 159-169, 1997.

80. Han, Y. F., M. J. Kahlich, M. Kinne and R. J. Behm, "CO removal from realistic methanol reformate via preferential oxidation-performance of a Rh/MgO catalyst and comparison to $Ru/\gamma-Al_2O_3$, and $Pt/\gamma-Al_2O_3$", *Applied Catalysis B*, Vol. 50, No. 4, pp. 209-218, 2004.

81. Choi, J., C. B. Shin and D. J. Suh, "Co-promoted Pt catalysts supported on silica aerogel for preferential oxidation of CO", *Catalysis Communications*, Vol. 9, No. 5, pp. 880-885, 2008.

82. Rosso, I., C. Galletti, G. Saracco, E. Garrone and V. Specchia, "Development of A zeolites-supported noble-metal catalysts for CO preferential oxidation: H_2 gas purification for fuel cell", *Applied Catalysis B*, Vol. 48, No. 3, pp. 195-203, 2004.

83. Iwasa, N., S. Arai and M. Arai, "Selective oxidation of CO with modified Pd/ZnO catalysts in the presence of H_2: Effects of additives and preparation variables", *Applied Catalysis B*, Vol. 79, No. 2, pp. 132-141, 2008.

84. Kim, D.H. and M. S. Lim, "Kinetics of selective CO oxidation in hydrogen-rich mixtures on Pt/alumina catalysts", *Applied Catalysis A*, Vol. 224, No. 1-2, pp. 27-38, 2002.

85. Avgouropoulos, G. and T. Ioannides, "Selective CO oxidation over CuO-CeO$_2$ catalysts prepared via the urea–nitrate combustion method", *Applied Catalysis A*, Vol. 244, No. 1, pp. 155-167, 2003.

86. Martinez-Arias, A., A. B. Hungria, G. Munuera and D. Gamarra, "Preferential oxidation of CO in rich H_2 over CuO/CeO$_2$: Details of selectivity and deactivation under the reactant stream", *Applied Catalysis B*, Vol. 65, No. 3-4, pp. 207-216, 2006.

87. Wang, J. B., S. C. Lin and T. J. Huang, "Selective CO oxidation in rich hydrogen over CuO/samaria-doped ceria", *Applied Catalysis A*, Vol. 232, No. 1-2, pp. 107-120, 2002.

88. Kim, D. H. and J. E. Cha, "A CuO-CeO$_2$ mixed-oxide catalyst for CO clean-up by selective oxidation in hydrogen-rich mixtures", *Catalysis Letters*, Vol. 86, No. 1-3, pp. 107-112, 2003.

89. Haruta, M., "Novel catalysis of gold deposited on metal oxides", *Catalysis Surveys From Japan*, Vol. 1, No. 1, pp. 61-73, 1997.

90. Haruta, M., "Size- and support-dependency in the catalysis of gold", *Catalysis Today*, Vol. 36, No. 1, pp. 153-166, 1997.

91. Wallace, W. T. and R. L. Whetten, "Coadsorption of CO and O_2 on Selected Gold Clusters: Evidence for Efficient Room-Temperature CO_2 Generation", *Journal of American Chemical Society*, Vol. 124, No. 25, pp. 7499-7505, 2002.

92. Sanchez, A., S. Abbet, U. Heiz, W. D. Schneider, H. Häkkinen, R. N. Barnett and U. Landman, "When Gold Is Not Noble: Nanoscale Gold Catalysts", *Journal of Physical Chemistry A*, Vol. 103, No. 48, pp. 9573–9578, 1999.

93. Iizuka, Y., T. Tode, T. Takao, K. Yatsu, T. Takeuchi, S. Tsubota and M. Haruta, "A Kinetic and Adsorption Study of CO Oxidation over Unsupported Fine Gold Powder and over Gold Supported on Titanium Dioxide", *Journal of Catalysis*, Vol. 187, No. 1, pp. 50-58, 1999.

94. Valden, M., X. Lai and D. W. Goodman, "Onset of Catalytic Activity of Gold Clusters on Titania with the Appearance of Nonmetallic Properties", *Science*, Vol. 281, pp. 1647-1650, 1998.

95. Haruta, M., T. Kobayashi, H. Sano and N. Yamada, "Novel Gold Catalysts for the Oxidation of Carbon Monoxide at a Temperature far Below 0 °C", *Chemistry Letters*, Vol. 16, No. 2, pp. 405-408, 1987.

96. Rossignol, C., S. Arrii, F. Morfin, L. Piccolo, V. Caps and J. L. Rousset, "Selective oxidation of CO over model gold-based catalysts in the presence of H_2", *Journal of Catalysis* , Vol. 230, No. 2, pp. 476-483, 2005.

97. Wu, X., L. Senapati, S. K. Nayak, A. Selloni and M. Hajaligol, "A density functional study of carbon monoxide adsorption on small cationic, neutral, and anionic gold clusters", *Journal of Chemical Physics*, Vol. 117, No. 8, pp. 4010-4015, 2002.

98. Yuan, D. W. and Z. Zeng, "Saturated adsorption of CO and coadsorption of CO and O_2 on Au_N^- (N = 2–7) clusters", *Journal of Chemical Physics*, Vol. 120, No. 14, pp. 6574-6584, 2004.

99. Mills, G., M. S. Gordon and H. Metiu, "The adsorption of molecular oxygen on neutral and negative Au_n clusters (n=2–5)", *Chemical Physics Letters*, Vol. 359, No. 5-6, pp. 493-499, 2002.

100. Yoon, B., H. Häkkinen and U. Landman, "Interaction of O_2 with Gold Clusters: Molecular and Dissociative Adsorption", *Journal of Physical Chemistry A*, Vol. 107, No. 20, pp. 4066-4071, 2003.

101. Davran-Candan, T., E. A. Aksoylu and R. Yıldırım, "Reaction pathway analysis for CO oxidation over anionic gold hexamers using DFT", *Journal of Molecular Catalysis A*, Vol. 306, No. 1-2, pp. 118-122, 2009.

102. Molina, L. M. and B. Hammer, "Active Role of Oxide Support during CO Oxidation at Au/MgO", *Physical Review Letters*, Vol. 90, No. 20, pp. 206102-1-4, 2003.

103. Molina, L. M. and B. Hammer, "Theoretical study of CO oxidation on Au nanoparticles supported by MgO(100)", *Physical Review B*, Vol. 69, No. 15, pp. 155424-1-22, 2004.

104. Molina, L. M., M. D. Rasmussen and B. Hammer, "Adsorption of O_2 and oxidation of CO at Au nanoparticles supported by TiO_2(110)", *Journal of Chemical Physics*, Vol. 120, No. 16, pp. 7673-7680, 2004.

105. Franceschetti, A., S. J. Pennycook and S. T. Pantelides, "Oxygen chemisorption on Au nanoparticles", *Chemical Physics Letters*, Vol. 374, No. 5-6, pp. 471-475, 2003.

106. Hvolbæk, B., T. V. W. Janssens, B. S. Clausen, H. Falsig, C. H. Christensen and J. K. Norskov, "Catalytic activity of Au nanoparticles", *Nano Today*, Vol. 2, No. 4, pp. 14-18, 2007.

107. Lopez, N. and J. K. Norskov, "Catalytic CO Oxidation by a Gold Nanoparticle: A Density Functional Study", *Journal of American Chemical Society*, Vol. 124, No. 38, pp. 11262-11263, 2002.

108. Xiao, L., B. Tollberg, X. Hu and L. Wang, "Structural study of gold clusters", *Journal of Chemical Physics*, Vol. 124, No. 11, pp. 114309-1-10, 2006.

109. H., Häkkinen, B. Yoon, U. Landman, X. Li, H. J. Zhai and L. S. Wang, "On the Electronic and Atomic Structures of Small Au_N (N = 4−14) Clusters: A Photoelectron Spectroscopy and Density-Functional Study", *Journal of Physical Chemistry A*, Vol. 107, No. 32, pp. 6168-6175, 2003.

110. Wang, J., G. Wang and J. Zhao, "Density-functional study of Au_n (n=2−20) clusters: Lowest-energy structures and electronic properties", *Physical Review B*, Vol. 66, No. 3, pp. 035418-1-6, 2002.

111. Jain, P.K., "A DFT-Based Study of the Low-Energy Electronic Structures and Properties of Small Gold Clusters", *Structural Chemistry*, Vol. 16, No. 4, pp. 421-426, 2005.

112. Hill, T. and P. Lewicki, *STATISTICS Methods and Applications*, StatSoft, Tulsa, 2007.

113. Tominaga, Y., "Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN", *Chemometrics and Intelligent Laboratory Systems*, Vol. 49, No. 1, pp. 105-115, 1999.

114. Edgar, T. F., D. M. Himmelblau and L. S. Lasdon, *Optimization of Chemical Processes*. Singapore: McGraw Hill, 2001.

115. Mandic, D. P. and J. A. Chambers, *Recurrent Neural Networks for Prediction*. London: Wiley , 2001.

116. Williams, R. W. and K. Herrup, "The Control of Neuron Number", *Annual Review of Neuroscience*, Vol. 11, pp. 423-453, 1988.

117. Wilamowski, B. M. and Y. Chen, "Efficient Algorithm for Training Neural Networks with one Hidden Layer", *International Joint Conference on Neural Networks*, Vol. 3, pp. 1725-1728, 1999.

118. Hagan, M. T. and M. B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm", *IEEE Transactions On Neural Networks*, Vol. 5, No. 6, pp. 989-993, 1994.

119. Kite, S., T. Hattori and Y. Murakami, "Estimation of catalytic performance by neural network-product distribution in oxidative dehydrogenation of ethylbenzene", *Applied Catalysis A*, Vol. 114, No. 2, pp. 173-178, 1994.

120. Hou, Z. Y., Q. Dai, X. Q. Wu and Chen G. T., "Artificial neural network aided design of catalyst for propane ammoxidation", *Applied Catalysis A*, Vol. 161, No. 1-2, pp. 183-190, 1997.

121. Cundari, T. R., J. Deng and Y. Zhao, "Design of a Propane Ammoxidation Catalyst Using Artificial Neural Networks and Genetic Algorithms", *Industrial & Engineering Chemistry Research*, Vol. 40, No. 23, pp. 5475-5480, 2001.

122. Huang, K., F. Q. Chen and D. W. Lü, "Artificial neural network-aided design of a multi-component catalyst for methane oxidative coupling", *Applied Catalysis A*, Vol. 219, No. 1-2, pp. 61-68, 2001.

123. Liu, Y., Y. Liu, D. Liu, T. Cao, S. Han and G. Xu, "Design of CO_2 hydrogenation catalyst by an artificial neural network", *Computers and Chemical Engineering*, Vol. 25, No. 11-12, pp. 1711-1714, 2001.

124. Serra, J. M., A. Corma, A. Chica, E. Argente and V. Botti, "Can artificial neural networks help the experimentation in catalysis?", *Catalysis Today*, Vol. 81, No. 3, pp. 393-403, 2003.

125. Huang, K., X. L. Zhan, F. Q. Chen and D. W. Lü, "Catalyst design for methane oxidative coupling by using artificial neural network and hybrid genetic algorithm", *Chemical Engineering Science*, Vol. 58, No. 1, pp. 81-87, 2003.

126. Kito, S., A. Satsuma, T. Ishikura, M. Niwa, Y. Murakami and T. Hattori, "Application of neural network to estimation of catalyst deactivation in methanol conversion", *Catalysis Today*, Vol. 97, No. 1, pp. 41-47, 2004.

127. Nandi, S., Y. Badhe, J. Lonari, U. Sridevi, B.S. Rao, S. S. Tambe and B. D. Kulkarni, "Hybrid process modeling and optimization strategies integrating neural networks/support vector regression and genetic algorithms: study of benzene isopropylation on Hbeta catalyst", *Chemical Engineering Journal*, Vol. 97, No. 2-3, pp. 115-129, 2004.

128. Watanabe, Y., T. Umegaki, M. Hashimoto, K. Omata and M., Yamada, "Optimization of Cu oxide catalysts for methanol synthesis by combinatorial tools using 96 well microplates, artificial neural network and genetic algorithm", *Catalysis Today*, Vol. 89, No. 4, pp. 455-464, 2004.

129. Végvári, L., A. Tompos, S. Gobölös and J. Margitfalvi, "Holographic research strategy for catalyst library design. Description of a new powerful optimisation method", *Catalysis Today*, Vol. 81, No. 3, pp. 517-527, 2003.

130. Tompos, A., J. L. Margitfalvi, E. Tfirst and L. Végvári, "Information mining using artificial neural networks and holographic research strategy", *Applied Catalysis A*, Vol. 254, No. 1, pp. 161-168, 2003.

131. Tompos, A., J. L. Margitfalvi, E. Tfirst and L. Végvári, "Evaluation of catalyst library optimization algorithms: Comparison of the Holographic Research Strategy and the Genetic Algorithm in virtual catalytic experiments", *Applied Catalysis A*, Vol. 303, No. 1, pp. 72-80, 2006.

132. Tompos, A., J. L. Margitfalvi, L. Vegvari, A. Hagemeyer, T. Volpe and C. J. Brooks, "Visualization of Large Experimental Space Using Holographic Mapping and Artificial Neural Networks. Benchmark Analysis of Multicomponent Catalysts for the Water Gas Shift Reaction", *Topics in Catalysis*, Vol. 53, No. 1-2, pp. 100-107, 2010.

133. Hattori, T. and S. Kito, "Analysis of factors controlling catalytic activity by neural network", *Catalysis Today*, Vol. 111, No. 3-4, pp. 328-332, 2006.

134. Omata, K., Y. Kobayashi and M. Yamada, "Artificial neural network-aided development of supported Co catalyst for preferential oxidation of CO in excess hydrogen", *Catalysis Communications*, Vol. 6, No. 8, pp. 563-567, 2005.

135. Omata, K., Y. Kobayashi and M., Yamada, "Artificial neural network-aided design of Co/SrCO$_3$ catalyst for preferential oxidation of CO in excess hydrogen", *Catalysis Today*, Vol. 117, No. 1-3, pp. 311-315, 2006.

136. Omata, K., Y. Kobayashi and M. Yamada, "Artificial neural network aided virtual screening of additives to a Co/SrCO$_3$ catalyst for preferential oxidation of CO in excess hydrogen", *Catalysis Communications*, Vol. 8, No. 1, pp. 1-5, 2007.

137. Umegaki, T., A. Masuda, K. Omata and M. Yamada, "Development of a high performance Cu-based ternary oxide catalyst for oxidative steam reforming of methanol using an artificial neural network", *Applied Catalysis A*, Vol. 351, No. 2, pp. 210-216, 2008.

138. Uguz, K. E. and R. Yildirim, "Comparative study of selective CO oxidation over Pt-Co-M/Al$_2$O$_3$ catalysts (M=Ce, Mg, Mn, Zr, Fe) in hydrogen-rich streams: effects of a second promoter", *Turkish Journal of Chemistry*, Vol. 33, pp. 545-553, 2009.

139. Davran-Candan, T., S. T. Tezcanlı and R. Yıldırım, "Selective CO oxidation over promoted Au/γ-Al$_2$O$_3$ catalysts in the presence of CO$_2$ and H$_2$O in the feed", *Catalysis Communications*, Vol. 12, No. 12, pp. 1149-1152, 2011.

140. Akpınar, F., *Low temperature water gas shift reaction over promoted Pt-based catalysts*, M.S. Thesis, Boğaziçi University, 2009.

141. Davran-Candan, T., *Experimental and Computational Study of Selective CO oxidation Over Au/Al$_2$O$_3$ Catalyst*, Ph.D. Thesis, Boğaziçi University, 2011.

142. Avgouropoulos, G., T. Ioannides, C. Papadopoulou, J. Batista, S. Hocevar and H. K. Matralis, "A comparative study of Pt/γ-Al$_2$O$_3$, Au/α-Fe$_2$O$_3$ and CuO–CeO$_2$ catalysts for the selective oxidation of carbon monoxide in excess hydrogen", *Catalysis Today*, Vol. 75, No. 1-4, pp. 157-167, 2002.

143. Jung, C. R., J. Han, S. W. Nam, T. H. Lim, S. A. Hong and H. I. Lee, "Selective oxidation of CO over CuO-CeO$_2$ catalyst: effect of calcination temperature", *Catalysis Today*, Vol. 93, pp. 183-190, 2004.

144. Park, J. W et al., "Activity and characterization of the Co-promoted CuO-CeO$_2$/γ-Al$_2$O$_3$ catalyst for the selective oxidation of CO in excess hydrogen", *Applied Catalysis A*, Vol. 274, No. 1-2, pp. 25-32, 2004.

145. Marino, F., C. Descorme and D. Duprez, "Supported base metal catalysts for the preferential oxidation of carbon monoxide in the presence of excess hydrogen (PROX)", *Applied Catalysis B*, Vol. 58, No. 3-4, pp. 175-183, 2005.

146. Kosmambetova, G.R., V. I. Gritsenko, P. E. Strizhak and A. M. Korduban, "Effect of the nature of the support for copper-cerium oxide catalysts on selective oxidation of CO in hydrogen-rich mixtures", *Theoretical and Experimental Chemistry*, Vol. 42, No. 2, pp. 133-138, 2006.

147. Firsova, A. A., A. N. Ilichev, T. I. Khomenko, L.V. Gorobinskii, Yu. V. Maksimov, I. P. Suzdalev and V. N. Korchak, "Selective Oxidation of CO in the Presence of Hydrogen on CuO/CeO$_2$ Catalysts Modified with Fe and Ni Oxides", *Kinetics and Catalysis*, Vol. 48, No. 2, pp. 282-292, 2007.

148. Gomez-Cortés, A., Y. Marquez, J. Arenas-Alatorre and G. Diaz, "Selective CO oxidation in excess of H$_2$ over high-surface area CuO/CeO$_2$ catalysts", *Catalysis Today*, Vol. 133, pp. 743-749, 2008.

149. Ribeiro, N. F. P., M. M. V. M. Souza and M. Schmal, "Combustion synthesis of copper catalysts for selective CO oxidation", *Journal of Power Sources*, Vol. 179, No. 1, pp. 329-334, 2008.

150. Polster, C. S., H. Nair and C. D. Baertsch, "Study of active sites and mechanism responsible for highly selective CO oxidation in H_2 rich atmospheres on a mixed Cu and Ce oxide catalyst", *Journal of Catalysis*, Vol. 266, No. 2, pp. 308-319, 2009.

151. Ayastuy, J. L., A. Gurbani, M. P. Gonzalez-Marcos and M. A. Gutiérrez-Ortiz, "Effect of copper loading on copper-ceria catalysts performance in CO selective oxidation for fuel cell applications", *International Journal of Hydrogen Energy*, Vol. 35, No. 3, pp. 1232-1244, 2010.

152. Wu, Z., H. Zhu, Z. Qin, H. Wang, J., Huang, L. Ding and J. J. Wang, *CO preferential oxidation in H_2-rich stream over a CuO/CeO_2 catalyst with high H_2O and CO_2 tolerance*, 2010, http://dx.doi.org/doi:10.1016/j.fuel.2010.03.001, accessed at January 2012.

153. Li, J., P. Zhu, S. Zuo, Q. Huang and R. Zhoua, "Influence of Mn doping on the performance of CuO-CeO_2 catalysts for selective oxidation of CO in hydrogen-rich streams", *Applied Catalysis A*, Vol. 381, No. 1-3, pp. 261-266, 2010.

154. Razeghi, A., A. Khodadadi, H. Ziaei-Azad and Y. Mortazavi, "Activity enhancement of Cu-doped ceria by reductive regeneration of CuO–CeO_2 catalyst for preferential oxidation of CO in H_2-rich streams", *Chemical Engineering Journal*, Vol. 164, No. 1, pp. 214-220, 2010.

155. Lendzion-Bielun, Z., M. M. Bettahar, S. Monteverdi, D. Moszynski and U. Narkiewicz, "Effect of Cobalt on the Activity of CuO/CeO_2 Catalyst for the Selective Oxidation of CO", *Catalysis Letters*, Vol. 134, No. 3-4, pp. 196-203, 2010.

156. Bethke, G. K. and H. H. Kung, "Selective CO oxidation in a hydrogen-rich stream over Au/γ-Al_2O_3 catalysts", *Applied Catalysis A*, Vol. 194, pp. 43-53, 2000.

157. Ito, S. I., T. Fujimori, K. Nagashima, K. Yuzaki and K. Kunimori, "Strong rhodium–niobia interaction in Rh/Nb$_2$O$_5$, Nb$_2$O$_5$–Rh/SiO$_2$ and RhNbO$_4$/SiO$_2$ catalysts", *Catalysis Today*, Vol. 57, No. 3-4, pp. 247-254, 2000.

158. Grisel, R. J. H. and B. E. Nieuwenhuys, "Selective Oxidation of CO over Supported Au Catalysts", *Journal of Catalysis*, Vol. 199, No. 1, pp. 48-59, 2001.

159. Son, I. H. and A. M. Lane, "Promotion of Pt/γ-Al$_2$O$_3$ by Ce for preferential oxidation of CO in H$_2$", *Catalysis Letters*, Vol. 76, No. 3-4, pp. 151-154, 2001.

160. Manasilp, A. and E. Gulari, "Selective CO oxidation over Pt/alumina catalysts for fuel cell applications", *Applied Catalysis B*, Vol. 37, No. 1, pp. 17-25, 2002.

161. Bulushev, D. A., L. Kiwi-Minsker, I. Yuranov, E. I. Suvorova, P. A. Buffat and A. Renken, "Structured Au/FeO$_x$/C Catalysts for Low-Temperature CO Oxidation", *Journal of Catalysis*, Vol. 210, No. 1, pp. 149-159, 2002.

162. Zhang, J., Y. Wang, B. Chen, C. Li, D. Wu and X. Wang, "Selective oxidation of CO in hydrogen rich gas over platinum–gold catalyst supported on zinc oxide for potential application in fuel cell", *Energy Conversion and Management*, Vol. 44, No. 11, pp. 1805-1815, 2003.

163. Epling, W. S., P. K. Cheekatamarla and A. M. Lane, "Reaction and surface characterization studies of titania-supported Co, Pt and Co/Pt catalysts for the selective oxidation of CO in H$_2$-containing streams", *Chemical Engineering Journal,* Vol. 93, No. 1, pp. 61-68, 2003.

164. Tanaka, H., S. I. Ito, S. Kameoka, K. Tomishige and K. Kunimori, "Promoting effect of potassium in selective oxidation of CO in hydrogen-rich stream on Rh catalysts", *Catalysis Communications*, Vol. 4, No. 1, pp. 1-4, 2003.

165. Özdemir, C., A. N. Akın and R. Yıldırım, "Low temperature CO oxidation in hydrogen rich streams on Pt-SnO$_2$/Al$_2$O$_3$ catalyst using Taguchi method", *Applied Catalysis A*, Vol. 258, No. 2, pp. 145-152, 2004.

166. Tibiletti, D., E. A. B. de Graaf, S. P. Teh, G. Rothenberg, D. Farrusseng and C. Mirodatos, "Selective CO oxidation in the presence of hydrogen: fast parallel screening and mechanistic studies on ceria-based catalysts", *Journal of Catalysis,* Vol. 225, No. 2, pp. 489-497, 2004.

167. Wootsch, A., C. Descorme and D. Duprez, "Preferential oxidation of carbon monoxide in the presence of hydrogen (PROX) over ceria–zirconia and alumina-supported Pt catalysts", *Journal of Catalysis*, Vol. 225, No. 2, pp. 259-266, 2004.

168. Mariño, F., C. Descorme and D. Duprez, "Noble metal catalysts for the preferential oxidation of carbon monoxide in the presence of hydrogen (PROX)", *Applied Catalysis B*, Vol. 54, No. 1, pp. 59-66, 2004.

169. Yan, J., J. Ma, P. Cao and P. Li, "Preferential oxidation of CO in H$_2$-rich gases over Co-promoted Pt-Al$_2$O$_3$ catalyst", *Catalysis Letters*, Vol. 93, No. 1-2, pp. 55-60, 2004.

170. Ince, T., G. Uysal, Akın A. N. and R. Yıldırım, "Selective low-temperature CO oxidation over Pt-Co-Ce/Al$_2$O$_3$ in hydrogen-rich streams", *Applied Catalysis A*, Vol. 292, pp. 171-176, 2005.

171. Suh, D. J., C. Kwak, J. H. Kim, S. M. Kwon and T. J. Park, "Removal of carbon monoxide from hydrogen-rich fuels by selective low-temperature oxidation over base metal added platinum catalysts", *Journal of Power Sources*, Vol. 142, No. 1-2, pp. 70-74, 2005.

172. Zhou, G., Y. Jiang, H. Xie and F. Qiu, "Non-noble metal catalyst for carbon monoxide selective oxidation in excess hydrogen", *Chemical Engineering Journal*, Vol. 109, No. 1-3, pp. 141-145, 2005.

173. Chin, S.Y., O. S. Alexeev and M. D. Amiridis, "Preferential oxidation of CO under excess H_2 conditions over Ru catalysts", *Applied Catalysis A*, Vol. 286, No. 2, pp. 157-166, 2005.

174. Pozdnyakova, O. et al., "Preferential CO oxidation in hydrogen (PROX) on ceria-supported catalysts, part I: Oxidation state and surface species on Pt/CeO_2 under reaction conditions", *Journal of Catalysis*, Vol. 237, No. 1, pp. 1-16, 2006.

175. Pozdnyakova, O. et al., "Preferential CO oxidation in hydrogen (PROX) on ceria-supported catalysts, part II: Oxidation states and surface species on Pd/CeO_2 under reaction conditions, suggested reaction mechanism", *Journal of Catalysis*, Vol. 237, No. 1, pp. 17-28, 2006.

176. Ko, E.Y., E. D. Park, K. W. Seo, H. C. Lee, D. Lee and S. Kim, "Selective CO oxidation in the presence of hydrogen over supported Pt catalysts promoted with transition metals", *Korean Journal of Chemical Engineering*, Vol. 23, No. 2, pp. 182-187, 2006.

177. Marques, P., N. F. P. Ribeiro, M. Schmal, A. G. D. Aranda and M. M. V. M. Souza, "Selective CO oxidation in the presence of H_2 over Pt and Pt-Sn catalysts supported on niobia", *Journal of Power Sources*, Vol. 158, No. 1, pp. 504-508, 2006.

178. Parinyaswan, A., S. Pongstabodee and A. Luengnaruemitchai, "Catalytic performances of $Pt–Pd/_{CeO2}$ catalysts for selective CO oxidation", *International Journal of Hydrogen Energy*, Vol. 31, No. 13, pp. 1942-1949, 2006.

179. Iwasa, N., S. Arai and M. Arai, "Effect of Cs promoter on the activity of Pd/ZnO catalyst for selective oxidation of CO in H_2-rich gas", *Catalysis Communications*, Vol. 7, No. 11, pp. 839-842, 2006.

180. Wang, Y. H., J. L. Zhu, J. C. Zhang, L. F. Song, J. Y. Hu, S. L. Ong and W. J. Ng, "Selective oxidation of CO in hydrogen-rich mixtures and kinetics investigation on platinum-gold supported on zinc oxide catalyst", *Journal of Power Sources*, Vol. 155, No. 2, pp. 440-446, 2006.

181. Monyanon, S., S. Pongstabodee and A. Luengnaruemitchai, "Catalytic activity of Pt–Au/CeO$_2$ catalyst for the preferential oxidation of CO in H_2-rich stream", *Journal of Power Sources*, Vol. 163, No. 1, pp. 547-554, 2006.

182. Chin, S. Y., O. S. Alexeev and M. D. Amiridis, "Structure and reactivity of Pt–Ru/SiO$_2$ catalysts for the preferential oxidation of CO under excess H_2", *Journal of Catalysis*, Vol. 243, No. 2, pp. 329-339, 2006.

183. Souzaa, M. M. V. M., N. F. P. Ribeiro and M. Schmal, "Influence of the support in selective CO oxidation on Pt catalysts for fuel cell applications", *International Journal of Hydrogen Energy*, Vol. 32, No. 3, pp. 425-429, 2007.

184. Chen, X., H. Zou, S. Chen, X. Dong and W. Lin, "Selective Oxidation of CO in Excess H_2 over Ru/Al$_2$O$_3$ Catalysts Modified with Metal Oxide", *Journal of Natural Gas Chemistry*, Vol. 16, No. 4, pp. 409-414, 2007.

185. Naknam, P., A. Luengnaruemitchai, S. Wongkasemjit and S. Osuwan, "Preferential catalytic oxidation of carbon monoxide in presence of hydrogen over bimetallic AuPt supported on zeolite catalysts", *Journal of Power Sources,* Vol. 165, No. 1, pp. 353-358, 2007.

186. Huang, Y., A. Wang, X. Wang and T. Zhang, "Preferential oxidation of CO under excess H_2 conditions over iridium catalysts", *International Journal of Hydrogen Energy*, Vol. 32, No. 16, pp. 3880-3886, 2007.

187. Teschner, D. et al., "Partial pressure dependent in situ spectroscopic study on the preferential CO oxidation in hydrogen (PROX) over Pt/ceria catalysts", *Journal of Catalysis*, Vol. 249, No. 2, pp. 318-327, 2007.

188. Chang, L. H., N. Sasirekha and Y. W., Chen, "Au/MnO_2–TiO_2 catalyst for preferential oxidation of carbon monoxide in hydrogen stream", *Catalysis Communications*, Vol. 8, No. 11, pp. 1702-1710, 2007.

189. Wang, F. and G. Lu, "The effect of K addition on Au/activated carbon for CO selective oxidation in hydrogen-rich gas", *Catalysis Letters*, Vol. 115, No. 1-2, pp. 46-50, 2007.

190. Galletti, C., S. Specchia, G. Saracco and V. Specchia, "CO preferential oxidation in H_2-rich gas for fuel cell applications: Microchannel reactor performance with Rh-based catalyst", *International Journal of Hydrogen Energy*, Vol. 33, No. 12, pp. 3045-3048, 2008.

191. Huang, Y., A. Wang, L. Li, X. Wang, D. Su and T. Zhang, "Ir-in-ceria: A highly selective catalyst for preferential CO oxidation", *Journal of Catalysis*, Vol. 255, No. 2, pp. 144-152, 2008.

192. Ribeiro, N. F. P., F. M. T. Mendes, C. A. C. Perez, M. M. V. M. Souza and M. Schmal, "Selective CO oxidation with nano gold particles-based catalysts over Al_2O_3 and ZrO_2", *Applied Catalysis A*, Vol. 347, No. 1, pp. 62-71, 2008.

193. Scire, S., C. Crisafulli, S. Minico, G. G. Condorelli and A. D. Mauro, "Selective oxidation of CO in H_2-rich stream over gold/iron oxide: An insight on the effect of catalyst pretreatment", *Journal of Molecular Catalysis A*, Vol. 284, No. 1-2, pp. 24-32, 2008.

194. Luengnaruemitchai, A., M. Nimsuk, P. Naknam, S. Wongkasemjit and S. Osuwan, "A comparative study of synthesized and commercial A-type zeolite-supported Pt catalysts for selective CO oxidation in H_2-rich stream", *International Journal of Hydrogen Energy*, Vol. 33, No. 1, pp. 206-213, 2008.

195. Tanaka, H., M. Kuriyama, Y. Ishida, S. Ito, K. Tomishige and K. Kunimori, "Preferential CO oxidation in hydrogen-rich stream over Pt catalysts modified with alkali metals Part I. Catalytic performance", *Applied Catalysis A*, Vol. 343, No. 1-2, pp. 117-124, 2008.

196. Tompos, A., M. Hegedus, J. L. Margitfalvi, E. G. Szabo and L. Vegvari, "Multicomponent Au/MgO catalysts designed for selective oxidation of carbon monoxide Application of a combinatorial approach", *Applied Catalysis A*, Vol. 334, No. 1-2, pp. 348-356, 2008.

197. Qiao, B., J. Zhang, L. Liu and Y., Deng, "Low-temperature prepared highly effective ferric hydroxide supported gold catalysts for carbon monoxide selective oxidation in the presence of hydrogen", *Applied Catalysis A*, Vol. 340, No. 2, pp. 220-228, 2008.

198. Avgouropoulos, G., M. Manzoli, F. Boccuzzi, T. Tabakova, J. Papavasiliou, T. Ioannides and V. Idakiev, "Catalytic performance and characterization of Au/doped-ceria catalysts for the preferential CO oxidation reaction", *Journal of Catalysis*, Vol. 256, No. 2, pp. 237-247, 2008.

199. Zhao, Z., M. M. Yung and U. S. Ozkan, "Effect of support on the preferential oxidation of CO over cobalt catalysts", *Catalysis Communications*, Vol. 9, No. 6, pp. 1465-1471, 2008.

200. Zhang, W., A. Wang, L. Li, X. Wang and T. Zhang, "Design of a novel bifunctional catalyst IrFe/Al$_2$O$_3$ for preferential CO oxidation", *Catalysis Today*, Vol. 131, No. 1-4, pp. 457-463, 2008.

201. Yung, M. M., Z. Zhao, M. P. Woods and Ozkan U. S., "Preferential oxidation of carbon monoxide on CoO$_x$/ZrO$_2$", *Journal of Molecular Catalysis A*, Vol. 279, No. 1, pp. 1-9, 2008.

202. Imai, H., M. Date and S. Tsubota, "Preferential Oxidation of CO in H$_2$-Rich Gas at Low Temperatures over Au Nanoparticles Supported on Metal Oxides", *Catalysis Letters*, Vol. 124, No. 1-2, pp. 68-73, 2008.

203. Kim, Y. H., E. D. Park, H. C. Lee, D. Lee and K. H. Lee, "Preferential CO oxidation over supported noble metal catalysts", *Catalysis Today*, Vol. 146, No. 1-2, pp. 253-259, 2009.

204. Neri, G., G. Rizzo, F. Corigliano, I. Arrigo, M. Caprì, L. De Luca, V. Modafferi and A. Donato, "A novel Pt/zeolite-based honeycomb catalyst for selective CO oxidation in a H$_2$-rich mixture", *Catalysis Today*, Vol. 147, pp. 210-214, 2009.

205. Mozer, T. S., D. A. Dziuba, C.T. P. Vieira and F. B. Passos, "The effect of copper on the selective carbon monoxide oxidation over alumina supported gold catalysts", *Journal of Power Sources*, Vol. 187, No. 1, pp. 209-215, 2009.

206. Padilla, R., M. Benito, L. Rodríguez, A. Serrano-Lotin and L. Daza, "Platinum supported catalysts for carbon monoxide preferential oxidation: Study of support influence", *Journal of Power Sources*, Vol. 192, No. 1, pp. 114-119, 2009.

207. Sangeetha, P. and Y. W. Chen, "Preferential oxidation of CO in H_2 stream on $Au/CeO_2–TiO_2$ catalysts", *International Journal of Hydrogen Energy*, Vol. 34, No. 17, pp. 7342-7347, 2009.

208. Woods, M. P., P. Gawade, B. Tan and U. S. Ozkan, "Preferential oxidation of carbon monoxide on Co/CeO_2 nanoparticles", *Applied Catalysis B*, Vol. 97, No. 1-2, pp. 28-35, 2010.

209. Liotta, L. F., G. D. Carlo, G. Pantaleo and A. M. Venezia, "Supported gold catalysts for CO oxidation and preferential oxidation of CO in H_2 stream: Support effect", *Catalysis Today*, Vol. 158, No. 1-2, pp. 56-62, 2010.

210. Zhang, W., Y. Huang, J. Wang, K. Liu, X. Wang, A. Wang and T. Zhang, "$IrFeO_x/SiO_2$- A highly active catalyst for preferential CO oxidation in H_2", *International Journal of Hydrogen Energy*, Vol. 35, No. 7, pp. 3065-3071, 2010.

211. Tabakova, T., M. Manzoli, D. Paneva, F. Boccuzzi, V. Idakiev and I. Mitov, "CO-free hydrogen production over $Au/CeO_2–Fe_2O_3$ catalysts: Part 2. Impact of the support composition on the performance in the water-gas shift reaction", *Applied Catalysis B*, Vol. 101, No. 3-4, pp. 266-274, 2011.

212. Khajeh-Hosseini-Dalasm, N., S. Ahadian, K. Fushinobu, K. Okazaki and Y. Kawazoe, "Prediction and analysis of the cathode catalyst layer performance of proton exchange membrane fuel cells using artificial neural network and statistical methods", *Journal of Power Sources*, Vol. 196, No. 8, pp. 3750-3756, 2011.

213. Sung, A. H., "Ranking importance of input parameters of neural networks," *Expert Systems with Applications*, Vol. 15, No. 3-4, pp. 405-411, 1998.

214. Molga, E. J., "Neural network approach to support modelling of chemical reactors: problems, resolutions, criteria of application", *Chemical Engineering and Processing*, Vol. 42, No. 8-9, pp. 675-695, 2003.

215. Atalik, B. and D. Uner, "Structure sensitivity of selective CO oxidation over Pt/γ-Al_2O_3", *Journal of Catalysis*, Vol. 241, No. 2, pp. 268-275, 2006.

216. Li, S., G. Liu, H. Lian, M. Jia, G. Zhao, D. Jiang and W. Zhang, "Low-temperature CO oxidation over supported Pt catalysts prepared by colloid-deposition method", *Catalysis Communications*, Vol. 9, No. 6, pp. 1045-1049, 2008.

217. Bond, G. C and D. T. Thompson, "Catalysis by Gold", *Catalysis Reviews Science and Engineering*, Vol. 41, No. 3-4, pp. 319-388, 1999.

218. Grisel, R. J. H., Nieuwenhuys, B. E., "A comparative study of the oxidation of CO and CH_4 over $Au/MO_x/Al_2O_3$ catalysts", *Catalysis Today*, Vol. 64, No. 1-2, pp. 69-81, 2001.

219. Mhadeshwar, A. B. and D. G. Vlachos, "Is the water–gas shift reaction on Pt simple? Computer-aided microkinetic model reduction, lumped rate expression, and rate-determining step", *Catalysis Today*, Vol. 105, No. 1, pp. 162-172, 2005.

220. Thinon, O., F. Diehl, P Avenier and Y. Schuurman, "Screening of bifunctional water-gas shift catalysts", *Catalysis Today*, Vol. 137, No. 1, pp. 29-35, 2008.

221. Caglayan, B. S. and A. E. Aksoylu, "Water-Gas Shift Reaction over Bimetallic Pt-Ni/Al_2O_3 Catalysts", *Turkish Journal of Chemistry*, Vol. 33, pp. 249-256, 2009.

222. Figueiredo, R. T., A. L. D. Ramos, H. M. C. de Andrade and J. L. G. Fierro, "Effect of low steam/carbon ratio on water gas shift reaction", *Catalysis Today*, Vol. 107, pp. 671-675, 2005.

223. Zhai, H. J. and L. S. Wang, "Chemisorption sites of CO on small gold clusters and transitions from chemisorption to physisorption", *Journal of Chemical Physics*, Vol. 122, No. 5, pp. 051101-1-4, 2005.

224. Fernández, E. M., J. M. Soler, I. L. Garzon and L. C. Balbás, "Trends in the structure and bonding of noble metal clusters", *Physical Review B*, Vol. 70, No. 16, pp. 1654031-14, 2004.

225. Sun, Q., P. Jena, Y. D. Kim, M. Fischer and G. Ganteför, "Interaction of Auanions with oxygen", *Journal of Chemical Physics*, Vol. 120, No. 14, pp. 6510-6515, 2004.

226. Ding, X., Z. Li, J. Yang, J. G. Hou and Q. Zhu, "Adsorption energies of molecular oxygen on Au clusters", *Journal of Chemical Physics*, Vol. 120, No. 20, pp. 9594-9600, 2004.

227. Cox, D. M., R. Brickman, K. Creegan and A. Kaldor, "Gold clusters reactions and deuterium uptake", *Zeitschrift Für Physik D*, Vol. 19, No. 1-4, pp. 353-355, 1991.

228. Günay, M. E., T. Davran-Candan and R. Yıldırım, *Analysis of O_2 Adsorption Stability and Strength Over Gold Clusters Using DFT and Logistic Regression*, 2012, http://dx.doi.org/doi:10.1007/s10876-011-0422-2, accessed at January 2012.

229. Mills, G., M. S. Gordon and H. Metiu, "Oxygen adsorption on Au clusters and a rough Au 111 surface: The role of surface flatness, electron confinement, excess electrons, and band gap", *Journal of Chemical Physics*, Vol. 118, No. 9, pp. 4198-4205, 2003.

230. Chretien, S., S. K. Buratto and H. Metiu, "Catalysis by very small Au clusters", *Current Opinion in Solid State and Materials Science*, Vol. 11, pp. 62-75, 2008.

231. Mirkin, B., *Clustering for Data Mining*, Chapman and Hall, New York, 2005.