T.C. HASAN KALYONCU UNIVERSITY GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

A COMPARATIVE STUDY OF DEEP LEARNING METHODS ON FLEXURAL BUCKLING LOAD PREDICTION OF ALUMINUM ALLOY COLUMNS

M.Sc. THESIS IN ELECTRONICS AND COMPUTER ENGINEERING

> Supervisor Assist. Prof. Dr. Bülent HAZNEDAR

> > Zeliha Begüm KILINÇ

January 2020

A COMPARATIVE STUDY OF DEEP LEARNING METHODS ON FLEXURAL BUCKLING LOAD PREDICTION OF ALUMINUM ALLOY COLUMNS

M.Sc. Thesis In Electronics and Computer Engineering Hasan Kalyoncu University

Supervisor Asst. Prof. Dr. Bülent HAZNEDAR

> Zeliha Begüm KILINÇ January 2020

© 2020 [ZELİHA BEGÜM KILINÇ]

GRADUATE SCHOOL OF NATURAL & APPLIED SCIENCES INSTITUTE MSc ACCEPTANCE AND APPROVAL FORM

Electronics-Computer Engineering M.Sc. (Master of Science) programme student Electronics-Computer Engineering M.S.C. (Master of Security programs)

Zeliha Begüm KILINÇ prepared and submitted the thesis titled "A comparative study of Zeliha Begüm KILINÇ prepared and submitted the thesis filed 11 complete. deep learning methods on the surface buckling load prediction of and material many
defended successfully on the date of 15/01/2020 and accepted by the jury as an M.Sc. thesis.

Prof. Dr. Mehmet KARPUZCU

Document no: ENS.FR.32 Yayın Tarihi: 26.03.2018 Rev no/ Date: 00/--

ABSTRACT

A COMPARATIVE STUDY OF DEEP LEARNING METHODS ON FLEXURAL BUCKLING LOAD PREDICTION OF ALLIMINUM ALLOY COLUMNS

KILINÇ, Zeliha Begüm

M.Sc. in Electronic - Computer Engineering Supervisor: Asst. Prof. Dr. Bülent HAZNEDAR

January 2020

In recent years, aluminum alloy columns have been widely used in construction fields. This is due to the light weight of aluminum alloys, high corrosion resistance, long life, low maintenance costs, the possibility of recovery, versatility of the metal and the possibility to obtain endless variety of profiles has many advantages. The calculation of the critical buckling loads of the columns is the most important issue. However, it is known that heat treated aluminum alloys have higher proof stress yield strength than non-heat treated aluminum alloys. In this study, buckling load estimation of heat treated aluminum alloy columns is made by using deep learning method and soft computing techniques and laboratory test results are compared. Sequential Model is used while using deep learning method. Adam, Adamax, Nadam, Adadelta, Adagrad, RMSProp and SGD and the optimizer functions of deep learning are evaluated separately. In addition, the results are evaluated using both MAE and MSE Loss functions for each optimizer. As a result of the study, it is understood that the optimizer and loss functions used together are more successful when estimating value for the dataset using deep learning model.

Keywords: Prediction of buckling load, Aluminum alloy columns, Soft computing, Deep learning, Artificial neural network

ÖZET

ALÜMİNYUM ALAŞIMLI KOLONLARIN EĞİLME BURKULMA YÜKÜ TAHMİNİ ÜZERİNE DERİN

ÖĞRENME YÖNTEMLERİNİN KARŞILAŞTIRMALI BİR ÇALIŞMASI

KILINÇ, Zeliha Begüm

Yüksek Lisans, Elektronik Bilgisayar Mühendisliği Tez Danışmanı: Dr. Öğr. Üyesi Bülent HAZNEDAR

Ocak 2020

Son yıllarda, Alüminyum alaşımlı sütunlar inşaat alanlarında çok yaygın kullanılmaktadır. Bunun sebebi alüminyum alaşımlarının hafifliği, yüksek korozyon direnci, uzun ömürlülüğü, düşük bakım maliyetleri, geri kazanma imkanı, metalin çok yönlülüğü ve sonsuz değişik şekilde profil elde edebilme olanağı gibi pek çok avantaja sahip olmasıdır. Kolonların kritik burkulma yüklerinin hesaplanması ise buradaki en önemli konudur. Bu problem çeşitli hesaplama teknikleri ve bilinen yaklaşımlar kullanılarak giderilmeye çalışılsa da doğrusal olmayan sonuçları içermektedir. Bu da kritik burkulma yükü tahminini zorlaştırmaktadır. Buna rağmen, Isıl işlem görmüş alüminyum alaşımlarının ısıl işlem görmemiş alüminyum alaşımlarına göre daha yüksek prova gerilimi verim mukavemetine sahip olduğu bilinmektedir. Bu çalışmada, ısıl işlem görmüş alüminyum alaşımlı kolonların derin öğrenme yöntemi kullanılarak burkulma yükü tahmini yapılmış softcomputing teknikleri ve laboratuvar ortamından alınan test sonuçları ile karşılaştırılmıştır. Derin öğrenme yöntemi kullanılırken Sequential Modelinden yararlanılmıştır. Derin öğrenmenin optimizer fonksiyonlarından olan Adam, Adamax, Nadam, Adadelta, Adagrad, RMSProp ve SGD ve kullanılarak ayrı ayrı değerlendirilmiştir. Ayrıca, her bir optimizer için hem MAE hem de MSE Loss fonksiyonları kullanılarak sonuçlar değerlendirilmiştir. Çalışmanın sonucunda, derin öğrenme modelinin kullanıldığı veriseti için değer tahmini yaparken birlikte kullanılan hangi Optimizer ve Loss fonksiyonları ikilisinin daha başarılı olduğu anlaşılmıştır.

Anahtar Kelimeler: Burkulma tahmini, Alüminyum alaşımlı kolonlar, Yapay sinir ağları, Derin Öğrenme.

ACKNOWLEDGEMENT

I would like to thank my thesis advisor Bülent Haznedar for his support and helpfulness in sharing his knowledge and experience with me during the selection, writing and preparation process of my thesis.I would like to express my gratitude to Abulkadir Çevik for sharing his sincerity with me and the owner of the dataset used in my thesis.I would also like to thank my husband Çağan Kılınç, who encouraged me in this process, supported me financially and did not spare his help and labor, and all my family who raised my motivation and encouraged me when I gave up ...

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

CHAPTER 1 INTRODUCTION

1.1 General Overview

Aluminum alloys are non-linear metallic materials with round stress-strain curves not well represented by simplified elastic-excellent plastic material. Some aluminum alloys have tensile strength up to 600 MPa. In general, some aluminum alloys which are not very suitable for high temperatures can easily be used up to temperatures up to 300° C. Aluminum; It has high corrosion resistance in many chemical media including air, water, saline and petrochemicals. Maintaining its mechanical properties at low temperatures is another outstanding feature of aluminum. In general, the welding properties of high strength aluminum alloys are not good. For this reason, aircraft are produced using riveted joints. However, there is a high-strength aluminum alloy that can be welded, albeit in a small number.

Heat treatment for aluminum alloys is generally limited to processes applied to increase the strength and hardness of machined and cast alloys which may exhibit precipitation hardening. Heat treatment to increase the strength of aluminum alloys; solution; dissolution of the phases (formation of solid solution), quenching; formation of supersaturated structure, aging; precipitation of dissolved atoms at room temperature or higher (precipitation hardening).

27% of the aluminum alloys produced annually in the world are used in the construction sector. It is used in the Both of non-structural elements (doors, windows, facade cladding ...) and structural elements (bridge decks, wide span roofing, panels and shell structures). Aluminum alloys are also used as formwork for on-site concrete casting. By producing composite elements with fiber-reinforced polymer and carbon-reinforced polymer materials, the relatively low elasticity of aluminum can be removed. Nowadays, aluminum has become the most widely used metal after steel. However, the buckling estimates of the aluminum alloy are a major problem for aluminum alloy columns.

Recently, deep learning methods have been used in many fields and they are improving day by day. It is known that the amount of dataset is important when training the network in deep learning methods. The more training and test data the success of the model is increasing. However, even when there is not enough data, it is possible to achieve successful results even in small data sets by changing the variables such as epoch, mini-batch, layer number and optimizer method correctly. Since the dataset used in this study is not very large, it is utilized in the variables of hyper parameters.

The objective of this paper is to analyze the application of deep learning modelling for strenght prediction of aluminum alloy columns failing by flexural buckling. Sequential Model (SM) is presented as deep learning techniques used in the study. Adam, Adamax, Nadam, Adadelta, Adagrad, SGD and RMSProp optimizers which are the functions to SM is used. Also MAE,MSE are used as loss functions.The training and test sets for deep learning model are obtained from Cevik et al., (2009) related with the subject of 'Flexural buckling load prediction of aluminum alloy columns using soft computing techniques'. This paper also deals with comparative of the results of the soft computing techniques and deep learning modelling. These results are the statistical methods of *RMSE, MSE, MAD, MAE, MAPE, MEAN, STD* respectively.The proposed deep learning model is presented in explicit form to be used in practical applications.

The other chapters are explained below:

- Chapter 2, is a literature survey which is related on previous studies on aluminum alloy columns, deep learning and analyses according to GEP (Gene-Expression Programming) and NN (Neural Networks) accuracy.

- Chapter 3, general the background of the study about flexural buckling load prediction of the aluminum alloy columns datasets utilize in deep learning.

- Chapter 4, illustrates the broad scope of the necessary process and analyses that can be accomplished with deep learning.

- Chapter 5, discusses the results obtained from processing and analysis of the study.

- Chapter 6, related with the overall results.

CHAPTER 2 BACKGROUND & LITERATURE REVIEW

2.1 Aluminum Alloy Columns

Yuanqing et al. (2017), experimental and numerical and numerical investigations of the behavior of extruded aluminum alloy I section columns with fixed pin end conditions against compression buckling have been conducted. A total of 11 column tests, including two pieces of heat-treated aluminum 6061-T6 and 6063-T5, are performed to obtain the results of compressed buckling strengths. Before the column tests are performed, the material properties of the two aluminum alloys are decided and determined from tensile coupon tests, while the first local and global geometric defects are measured differently by experimental techniques to reach different values. Using the ABAQUS software system in a comfortable way, finite element (FE) models have been developed to help explain the linearity of the material and the defects found in the initial shape, ie geometry. FE models also provide reliable simulation of the tightened buckling behavior with the fixed connection end conditions of the columns verified, verified and tested according to the test results given as a result of the entered data. Based on its test and numerical results, all calculation methods in current design standards, including European, Chinese, American and Australian / New Zealand specifications, are evaluated in detail. It has been shown that the design provisions of all four standards provide relatively conservative strength estimates, especially for aluminum alloys with a tensile hardening capacity that makes them more specific.

Studies have been made to develop a finite element model to investigate the buckling movements of some aluminum columns exposed to fire. The results of this model were verified and validated in the results of the experiments. Considering a parametric study with the help of the finite element model, it was concluded in EN 1999-1-2 that as a result of the bending movements of the aluminum columns exposed to fire, the simple calculation model could not give accurate estimation of the resistance against buckling in fire. This paper proposes a different way of suggesting an alternative design model by emphasizing the stressstrain relationship of aluminum alloys at high temperatures. Estimates of this model give approximate values to the results found in the finite element model (Maljhaars, 2009).

Wang et al. (2017) showed that the failure behavior of the aluminum alloy elements which are faced with the axial compression load and failed in the general buckling mode is moved and focused on the finite element simulation. The column elements play a major role in extrusion production using approximately a new heat-treated aluminum alloy 6082-T6. Thus, it is not difficult to say that it is produced. We have tried to develop an accurate finite element model for the simulation of aluminum alloy columns and have been quite successful. The developed finite element model shows that it can predict the damaged shapes and final loads of the columns tested according to the experimental results, but some notes have been entered. A parametric study has been carried out on the stability of aluminum alloy column elements which fail with general alloys. A suitable column curve is recommended for the aluminum alloy design; it is slightly more conservative than the current Chinese curve in GB50429.

Rasmussen et al. al (2000) forms a column curve formulation that can produce accurate strength for extruded aluminum alloys that fail to bend bending. The formulation uses a simple extension of the Perry curve and checked its validity for all alloys used in practice. It is believed that the material properties are typically expressed in terms of Ramberg-Osgood parameters obtained from a stud column test of the finished product. First, the formulation has been shown to be able to produce ECCS a, b- and c-column curves closely for aluminum alloys. Secondly, a better agreement is reached by accepting ECCS's recommendations based on the column curve selection based on the alloy type (heat treated or not heat treated). In addition, the tests can be obtained using the column curve formulation proposed on paper compared to the column curves of ISO.

2.2 Deep Learning

Deep learning is a machine learning type that try to teaches computers to do what comes naturally to humans; learn by using example. Deep learning is the clef technology behind driverless cars, unmanned aerial vehicles, enabling them to get to know a stop sign or to seperate a pedestrian from a lamppost. It is the clef to voice control in user devices like phones, tablets, wireless headset, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It is achieving results that are not possible before.

Deep learning has attracted a lot of attention in recent years and for good reasons. Results not previously possible. Recent advances in deep learning have improved to the point where deep learning performs better than humans in some tasks, such as classifying objects in images. While deep learning is first theoretical in the 1980s, there are two main reasons why it has been useful recently; Deep learning must have a large amount of tagged data in the system. For example, the development of vehicles without drivers requires millions of photos, images, pictures and thousands of hours of video. Deep learning requires significant computing power; High-performance GPUs have a parallel architecture that is efficient for deep learning. When combined with clusters or cloud computing, this allows development teams to reduce training time from weeks to hours or less for a deep learning network (Abdulkader, 2019).

Shalev-Shwartz and Ben-David (2014) explained that a machine can as well memorize emails which are labelled as spam and non-spam in order to establish relations and differentiate the types from each other in future encounters, which in the end makes a perfect example for "Machine Learning". The term "Machine Learning" became more popular in the past decades and a conceptually well-known approach among many individuals.

The methodologies introduced and developed in the style of machine learning enable a large number of automation capabilities to be used to solve many problems in different research areas. Again, different types of algorithms are used to find different solutions to problems related to machine learning and give different perspectives to research or problem area. (İnce, 2019).

2.3 Overview Studies

Çevik et al. (2009) tried to apply soft calculation techniques for the calculation of strength, ie strength, of heat-treated rolled aluminum alloy columns that failed with bending buckling. Neural networks (NN) and genetic programming (GP) are used as computational techniques on experimental data obtained from the literature. It has been found that the proposed NN and GEP models are more accurate than the existing models and related codes (EC9 and ISO), which are preferred and proposed by Rasmussen from the previous analytical terms and codes applied and used.

According to Üner (2019), the study is investigate whether more information can be extracted from this remaining set of experiments with a deep learning-based approach. They experiment with 6 different deep learning architectures that use gene expression data from the LINCS L1000 project, chemical structure fingerprints of drugs, SMILES string representation of drug structure, and the atomic structure of the drug molecules. The multilayer perceptron (MLP) based model which uses chemical structures and gene expression features achieve 88% microAUC and 79% macro-AUC, thus offering better performance in comparison to the stateof-the-art studies on side effect prediction. They observe that the chemical structure is more predictive than the gene expression profiles despite the fact that the features are extracted with different deep learning models. Finally, the convolutional neural network-based model that uses only SMILES strings of the drugs provides 82% macro-AUC, and 88%micro-AUC improvements, better performing than the models that use gene expression and chemical structure features simultaneously.

Abdulkader (2019), the main purpose of this study is to define and create new VO2max prediction models with the help of deep learning (DL). Data set was tried to be separated into training and test data by using 70-30%, 80-20% split rate and 10 times cross validation. When the comparison was applied, the aim of the study was to develop the VO2max estimation models based on Multi-Layer Sensor (MLP), Support Vector Machine (SVM) and Single Decision Tree (SDT). The performance of the estimation models was evaluated using the Standard Estimation Error (SEE) and Multiple Correlation Coefficient (R). As a result, DL VO2max can be used easily and safely in the desired branch.

Yetiş (2019), studied within the scope of this thesis, 2D floor plans and facade drawings collected and / or produced from scratch are used as data sets. These data are semantically divided into three different Convolutionary Neural Networks to obtain relevant architectural information, because with the widespread use of Deep Learning, it shows promising success in solving a wide range of problems. Semantically segmented drawings are then converted to 3D models using Digital Geometry Processing methods. Finally, a web application was defined on the system in order to allow any user to choose, buy and use 3D models comfortably and easily. In 2D, the results of two different case studies are evaluated for the results given by semantic segmentation and in 3D, and different comparisons are made with different measurement methods to represent the accuracy of the process to the system. As a

result, this work has been done and made possible by recommending and applying the most advanced methods, recommending an automated process for the reconstruction of 3D models and making this system easily and easily available even for a person without technical knowledge.

Paker (2019) delineated nowadays with the increase in biological knowledge, the use of deep learning in bioinformatics and computational biology has increased. Deep learning is widely used to classify and analyze biological sequences. In recent years, deep neural network architectures such as Convolutional and Recurrent Neural Networks have been developed to achieve more successful results compared to classical machine learning algorithms. The problem discussed in this thesis is a bioinformatics problem. Therefore, it is discussed whether the given microRNA molecule binds to the mRNA molecule. MicroRNAs (miRNAs) are \sim 21-23 base length non-coding RNA molecules that play an important role in gene expression. After transcription, they target mRNAs and cause mRNA degradation or translation inhibition. Rapid and efficient determination of the binding sites of miRNAs is a major problem in bioinformatics. In this thesis, a deep learning approach based on Long Short Term Memory (LSTM) has been developed with the help of an existing duplex sequence model. The study provides a comparative approach based on different data sets and configurations. In addition, a web interface has been developed to efficiently and quickly identify human miRNA target regions and provide a visual interface to the end user. Compared to the six classical machine learning methods, the proposed LSTM model gives better results in terms of some evaluation criteria.

Pala (2019) has sought to systematically compare systematic videos using challenging conditions and conditions using some of the deeper learning-based methods recently developed to assist the facial recognition system and make it work better and more comfortably. Openface, VGGFace2 and Arcface tried to make evaluations in 3 different deep learning models and succeeded. The results of the UvA-NEMO video dataset for Pala make a difference in their work by showing that even the most successful deep learning-based face recognition methods provide poor performance in difficult distortions such as noise, blur and contrast.

Tanimu (2019) examined the identification of cracks in infrastructure such as roads and bridges that may cause future problems. The methods of detecting cracks are examined in two groups as destructive and non-destructive. This study aims to analyze the cracks in the materials using ground radar analysis due to its many advantages compared to other methods. In this study, a laboratory environment is created and ground radar and thermal image measurements of various blocks with and without fractures are made in different shapes and materials. After visual and thermal analysis, GPR raw data are analyzed by wavelet transform and entropy methods. Finally, the continuous wavelet transform coefficients are classified, separated and divided into different forms and branches with the help of the possibilities provided by the deep learning methods, especially with convolutional neural network. In addition, they have tried and used a case study of a bridge used for the aforementioned previous research studies to test the methods on a larger scale. The results show that the proposed Wavelet-CNN crack detection method is more useful and better than direct detection of cracks from raw data or b-scan signals. In this way, this system is preferred.

Munir (2019) examined to classify a single handwriting character so that the handwriting text can be translated into a digital form. To classify a complete word or text, the first step is to correctly recognize the lines of text. A text line detection system has been developed that can detect all text lines according to the curvature angle of the text lines by dividing the original image of an A4 scanned document. Each letter image in the text lines is identified and given as an introduction to the deep learning network for later recognition. A set of data from our own handwriting, which includes 2200 images of each letter, is also used, together with a public data set for the educational disruption of the deep learning network. A total of 26 x 7800 = 202, 800 pictures are used for the training of the artificial neural network. A GUI system has been designed that can retrieve a scanned document in image science using MATLAB and provide text line detection, letter detection and letter recognition results.

Süberk (2019) conducted point estimation using the help of deep learning methods of building density in some images of remote sensing optics. The aim of the project is to minimize the mean square error of the estimated intensity by introducing and applying architectural changes to the deep learning network and using and opting for increased educational data. The results obtained or reached show that it is possible to easily and quickly predict the density of the building using vanilla convolutional neural networks (CNN) aids. Adding the

sigmoid layer to the system, reducing the network to a small data set and data amplification for easy amplification significantly increases its accuracy in regression. Data amplification is the most accurate or preferred method to reduce the mean square root error in this thesis.

Soydaş (2019) studied the subject of aircraft detection from satellite imagery is discussed, and different neural network architectures based on deep learning technique with traditional methods have been trained and tested. For the test, labeling is performed manually on the images containing the airport regions. An algorithm has been developed for rapid detection and high performance in large scale images, and the use of different architectures and the effects of training methods on success have been examined. In the study, firstly the literature is searched and different approaches are examined and then information about machine learning basics is shared. Deep learning, which is examined as a subtitle of machine learning, is also studied and mentioned. Convolutional neural networks, which constitute the most important point of the backbone of the study, have been introduced into the system and introduced to the system and the basic concepts have been emphasized. Many of the object detection models that work with the help of deep learning techniques use convective neural networks as feature extractors. Therefore, in the methodology section of the study, the sections where CNN and object detection architectures intersect are examined and discussed from different angles and the latest technology detection architectures are examined. The sliding window method and maximum printing algorithms are used safely and easily for the fast and high performance detection of large scale satellite images.

Rasheed (2019) examined the deep learning from different frameworks. Our study then focuses on object detection with a deep learning technique and compares which methods are suitable for the detection of one or more dogs. For object detection, firstly, classification and placement of objects in the image are provided. Object detection is a difficult issue involving computer vision and machine learning. In this study, the most impressive results are achieved by applying various deep learning techniques with the help of machine learning and computer vision, because deep learning has become one of the most popular and widely used methods in research communities. The convolutional neural network is a subtype that easily overcomes some competitions in its views on the computer system in the deep neural network, because good results are easily achieved and are well suited for use in object detection tasks. Based on the latest research of the discoveries in recent research, the technique appropriate for the

studies to be made or to be conducted is considered to be the Faster R-CNN (Faster R-CNN) preferred for classifying images from multiple dog species. R-CNN, which is faster for object detection and classification, passes several different and distant paths, especially for detection. The R-CNN method, which is faster especially in terms of accuracy and speed, is generally preferred. This provides a faster R-CNN trained to classify and detect dog images.

According to Çakır (2019), study analyzed the effectiveness of supervised and unattended deep learning algorithms in detecting zero-day attacks. We conducted analyzes on different deep learning models and compared the performance of these algorithms with each other. The models tested are deep learning algorithms, forward neural networks, recursive neural networks and convolutional neural networks. Test results showed that unguided deep learning methods are successful in detecting attacks with an accuracy score of 95.3% and a f1-score of 97%. In addition to the test sets produced in the same environment as the training set, test sets created in different environments from the training set are used to test the deep learning methods. These tests showed that although deep learning methods cannot detect every attack in sets created in different environments, some attacks can detect low false positive rates.

Özgenel (2019) explained data design term as introduced to describe end to end process of problem solving with deep learning algorithms which is suitable for broad range of applications including problems in architecture. Data design defined as a holistic approach embracing the process from problem (re)formulation to evaluation of the results considering the interrelations of decisions made throughout the process. In this context, data design in architecture is exemplified with the task of crack detection in buildings in order to minimize subjectivity in the course of evaluating the results. For this purpose, the relation between data and deep learning framework, case specific evaluation requirements and strategies for enhancing the performance are inspected through image classification and semantic segmentation applications for crack detection. Concordantly, this study contributes to the literature not only with the introduction and framing of data design but also with the proposal of crack detection specific evaluation metrics for both image classification and segmentation applications and a novel method is proposed employing quad tree and deep learning algorithms in conjunction for semantic segmentation of objects with limited visual features.

As a result, data design and respective consequences are discussed in depth and demonstrated regarding the case dependency, decisions taken in the course of implementation and their influences to both process and the results.

Tienin (2019) examined the main purpose of this study is to compare traditional image processing methods with deep learning techniques. First, a two-class classification study is conducted. The aim of this study is to differentiate between cloud images and non-cloud images. Convolutional Neural Networks and SoftmaxWithLoss function are used for estimation and classification. In the training phase, 92% accuracy is achieved after fine tuning of the model. In the next step, detection and segmentation is performed on cloud images using image processing techniques. At this stage, edge detection methods and watershed algorithm are applied. In the third stage, Full Convolutional Networks (THA, FCN) and U-NET deep learning methods are used for segmentation. Segmentation of cloud images is provided by two different deep learning methods. U-NET is found to be 87% as Dice coefficient and 45% as loss during the training phase. FCN network, stochastic slope drop, Adam's momentum and Nesterov momentum techniques have been applied to different training methods and the highest success is achieved with Adam's momentum technique is 63.12%.

Sorkun (2018), focused on statistical time series estimation methods. The aim of this study is to investigate the suitability and competitiveness of time series estimation methods with deep learning on solar radiation data. In this context, time series estimation on solar irradiation is made by using the regenerative neural network variation Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models. By optimizing the parameters, values are found to form the model that best represents the solar radiation data. The seasonal performances of established RNN models and other machine learning methods are compared and a hybrid model is proposed according to experimental results. Finally, the effect of additional meteorological parameters on solar radiation estimation is investigated. The results show that LSTM and GRU models are suitable and competitive for estimating time series over 1 hour horizon on solar radiation data. Experiments have shown that the hybrid approach and additional meteorological parameters improve the performance of the model.

Anwer (2017) defined deep learning which is used in medical filed, to propose deep learning based clinical support system for diagnosis breast cancer. Deep learning is a subfield of machine learning to solve problems of artificial intelligence. This study aims to investigate whether the deep learning approaches are also capable of producing successful results in the medical area. Many of deep learning algorithms have been used in the diagnosis of breast cancer, such as fully connected neural networks and recurrent neural network, the new order is the first time that the new architecture of convolution neural networks is used to diagnosis breast cancer. Wisconsin Breast Cancer datasets from the UCI Machine Learning repository is used to test the ability of deep learning different techniques. The experiments with deep learning approaches show a promising direction towards supporting medical decision making.

Eriş (2018) aimed to use the deep learning methods which have been successful in many use cases in judicial evidence examinations. For this purpose, a method of extracting evidence to assist in forensic examinations is proposed using deep learning object detection algorithms. The results show that the classification of images obtained from digital evidences according to their contents can be performed 93% faster using deep learning. It has also been shown that human errors will be reduced as a result of classification. It is aimed to integrate the proposed method into the forensic evidence analysis software tools by applying it to audio, text and video data as well as image data in further studies.

Kutlu (2019) delineated in recent years, artificial intelligence algorithms are one of the most researched and developed applications. Machine learning, artificial neural networks, classification, clustering algorithms are used in artificial intelligence applications. One of these methods is deep learning. Deep learning is an advanced machine learning class. Using the Deep Learning method, video analysis, image classification, speech recognition and natural language processing are very successful. The data and experiences to be provided by the projects that will cover the areas of Deep Learning and Unmanned Aerial Vehicles will contribute to the development of high value-added products in these technologies by increasing the number of qualified studies on these issues. In this study, a control software that evaluates image data from unmanned aerial vehicles and makes various inferences (classification, positioning, marking) is created. By using the method of retraining the last layers of the pre-trained artificial neural network models with our data set, it is tried to reduce the training time and increase the success. In these studies, 2 pre-educated models are used and as a result of training of these models, as a result of 190 thousand step training, 25.39 and 27.87 mAP values are reached.

Kantepe (2019) producted recommendation system is designed using the deep learning method AutoEncoder. This designed recommendation system is implemented using the Python language on the TensorFlow platform. MovieLens 1M dataset is used which consists of the movie ratings ranging from 1 to 5 and collected by the GroupLens researchers from the users of MovieLens website. The study is written in four sections. Furthermore, it is tried to find the best optimization algorithm to improve the success of the system and the effect of increasing the amount of data that is analyzed. The results of the algorithms of Gradient Descent, Gradient Descent with Momentum, RmsProp and Adam are depicted in both tables and graphs. In the third section, findings are evaluated and as a result, the Adam is shown to be the best algorithm with 1,363 points of test error. It is observed that the more data a training set has, the more successful the recommendation system is. In the fourth section, all the results from the processes mentioned above are summarized and the studies which can be conducted in order to implement a better recommendation system as future plans are stated.

Koyun and Afşin (2017), used 2160000 characters to train the convolutional neural network. The authors state that the method developed by themselves is better when compared to the OCR tool embedded in the Matlab environment.

Uçar and Bingöl (2018), have briefly introduced the layers of DKSA (Deep Convolutional Neural Networks), which is a kind of deep learning and is widely used in image processing applications, and gave information about their architecture. Using the Caffe program to implement DKSAs, the authors built the applications on two computers with embedded Nvidia Jetson TK1 / TX1 development cards and Nvidia GTX550 / GTX960 graphics cards. On the cards and the computer, they use the LeNet network to recognize handwritten numbers, and in their assessment of performance, speed and accuracy, they stated that GPUs are faster than CPUs, although they are close to accuracy.

In the study conducted by Büber and Şahingöz in 2017, it is aimed to perform image processing by using deep learning approach and to recognize numerical characters. In the application, 50000 data from MNIST data set consisting of handwriting pictures taken from 250 different people are used for training and 10000 data are used for testing. In addition, the effects of hyperparameters on the performance of the algorithm should be determined before using the deep learning method. The authors stated that the success rate of the numerical hyper parameters determined to be used in the test step for the best case is 94.75%.

In 2016, Kaynar et al. presented a deep learning-based approach to automatically detect spam mail. Authors using 2 different data sets used 75% of the data in Turkish and English data sets for educational purposes and 25% for testing purposes. The proposed model is a classifier with 2 deep auto encoders connected back to back and a softmax layer at the outputs. This classifier has a 98% success rate.

In the studies conducted by Işık and Artuner in 2016, it is aimed to identify the radio signals received through Software Based Radio with deep learning neural networks. The authors made a convolutional neural network study with the images obtained from the Signal Identification Wiki and used 82 of 137 images for training, 22 for verification and the remaining 33 for testing. It is stated that 82% performance is achieved at the end of the operation of the Caffe platform-trained network.

Light and Artuner (2016) worked on recognition of radio signals with deep learning Neural Networks. In their work, the authors state that they aim to correctly identify the properties of fabrics with a distinctive texture rather than detecting fabric defects. They used 1000 sample data, 153 of which are incorrect during the training phase, and 100 of which 60 are error-free during the testing phase. The authors concluded that the proposed method yielded 88% success.

In 2017, the study conducted by Razavi and Yalçın aimed to identify the plant type from images collected from smart agricultural stations. The authors proposed a deep learning method that can automatically extract attributes from two-dimensional plant images. In the construction of CNN architecture, 4800 images of 16 plant classes in TARBIL data set are used. The authors state that 3 different deep learning methods are used, the CNN-based approach works on 16 kinds of plants with an accuracy of approximately 97.47% and the classification accuracy is better than other methods. Yalçın examined to automatically identify human activities in RGB-D video images. Yalçın wanted to classify the activities of people from skeletal movements and deep learning through 3D skeletal joint data. The raw data from human images are pre-processed with operations such as translation, rotation and scaling. Using the Human3.6M activity data set for the training of the proposed method, the author used a model with 5 hidden layers as a deep understanding network, 47 neurons per layer, a non-instructional learning rate of 0.001 and an instructional learning rate of 0.12.

Arık et al. 2017, delineated a system that classifies skin lesions for the detection of melanoma by deep learning methods is proposed. In this study, approximately 65% of 1279 healthy and diseased images from ISIC archive are used for education and 35% for testing. The authors stated that because of the data limitation, the system they recommended is 70% successful and expert dermatologists predicted the disease at an average rate of 75%.

Yetkin and Hamamcı (2016), worked on a deep learning-based method for estimating the relative exposure between the patient's planar image and the previously taken head MR images is presented. In practice, the MR volume obtained from a healthy subject using the 3D gradient-echo sequence is used. After this volume is rendered in 3D, 324 images are obtained with 18 different light angles, and 260 images are used for training and 64 for testing in the convolutional neural network.

(Shahbazi et al. 2014) studied the elastic buckling of smart lightweight column structures integrated with a pair of surface piezoelectric layers using artificial intelligence. The finite element modeling of Smart lightweight columns is found using ANSYS® software. Then, the first buckling load of the structure is calculated using eigenvalue buckling analysis. To determine the accuracy of the present finite element analysis, a compression study is carried out with literature. Later, parametric studies for length variations, width, and thickness of the elastic core and of the piezoelectric outer layers are performed and the associated buckling load data sets for artificial intelligence are gathered. Finally, the application of soft computing-based methods including artificial neural network (ANN), fuzzy inference system (FIS), and adaptive neuro fuzzy inference system (ANFIS) are carried out. The comparison of the results reveal that, the ANFIS model with Gaussian membership function provides high accuracy on the prediction of the buckling load in smart lightweight columns, providing better predictions compared to other methods. However, the results obtained from the ANN model using the feed-forward algorithm are also accurate and reliable.

Yuan and Bao (2018) assessed a guided wave (GW)-CNN based fatigue crack diagnosis method is proposed. GW features extracted from GW signals of different excitation-acquiring paths with different excitation frequencies are employed as the input of a trained CNN for fatigue crack diagnosis. An experiment verification is performed on a kind of attachment lug specimen, which is an important connecting component on aircraft structures. The results show that this proposed method is promising.

Shon et al. developed a deep conditional generative model for structured output prediction using Gaussian latent variables. The model is trained efficiently in the framework of stochastic gradient variational Bayes, and allows for fast prediction using stochastic feedforward inference. In experiments, we demonstrate the effectiveness of our proposed algorithm in comparison to the deterministic deep neural network counterparts in generating diverse but realistic structured output predictions using stochastic inference. Furthermore, the proposed training methods are complimentary, which leads to strong pixel-level object segmentation and semantic labeling performance on Caltech-UCSD Birds 200 and the subset of Labeled Faces in the Wild dataset.

Merwe, analyzed the most effective column design method is to be found by comparing the different column design criteria and methods of design which are in use for different materials. The materials under consideration are carbon and low-alloy steels (structural steels), stainless steels and an aluminum alloy that is regarded as suitable for structural design.

Durmuş et. al (2006) Effects of ageing conditions at various temperatures, load, sliding speed, abrasive grit diameter in 6351 aluminum alloy have been investigated by using artificial neural networks. The experimental results are trained in an ANNs program and the results are compared with experimental values. It is observed that the experimental results coincided with ANNs results.

Hassan et. al (2009) studied the potential of using feed forward backpropagation neural network in prediction of some physical properties and hardness of aluminum–copper/silicon carbide composites synthesized by compocasting method. Two input vectors are used in the construction of proposed network; namely weight percentage of the copper and volume fraction of the reinforced particles. Density, porosity and hardness are the three outputs developed fromthe proposed network. Effects of addition of copper as alloying element and silicon carbide as reinforcement particles to Al–4 wt.% Mg metal matrix have been investigated by using artificial neural networks. The maximum absolute relative error for predicted values does not exceed 5.99%. Therefore, by using ANN outputs, satisfactory results can be estimated rather than measured and hence reduce testing time and cost.

Kechagias et. al assesed Al7075 alloy surface quality, achieved in slot milling, constitutes the subject of the current research study. Twenty seven slots are cut using different cutting conditions by a KC633M drill-slot end mill cutter. The independent variables considered are the depth of cut (ap, mm), cutting speed (Vc, m/min), and feed rate (f,mm/rev). Process performance is estimated using the statistical surface texture parameters Ra, RSm, and Rt; all measured in microns. To predict the surface roughness within the limits of the parameters involved, an artificial feed forward back propagation neural network model is designed for the data obtained.

CHAPTER 3 MATERIALS & METHODS

3.1 Aluminum Alloy Columns

Aluminum tubular elements are used in curtain walls, space structures and other structural applications. Aluminum tubular elements are normally made of aluminum for heat treatment because heat treatment alloys have higher yield stress than important properties of non-heat treated alloys. Aluminum structural elements are specially based on ultimate strength of compact sections. Because of this, it is necessary to investigate the behavior and design of aluminum columns, plates, beams and beam-columns of slender sections. Although aluminum alloys are well suited for a variety of applications in structures, their unique material properties make structural reactions different from steel. In the literature, important experimental researches about Aluminum column test have been made. Although aluminum alloys are well suited for some applications in marine structures, their unique material characteristics make the structural response different than steel. When structural elements are subjected to compressive loads, the buckling and collapse capacity is one of the most crucial factors governing the design. Ultimate strength of longitudinally stiffened panels is very important because it governs the structural capacity. Such panels are subjected to longitudinal compression, transverse compression, shear, and local bending. As the mechanical properties of aluminum alloy typically vary more significantly between the parent metal, weld metal and HAZ (Heat Affected Zone), as compared to those of steel, it would be anticipated that the existing formulations for steel structures may not be accurate when applied to aluminum alloy panels. The scope of research on the extensive experimental data and numerical data made in the European Construction Steels Convention (ECCS) in the 1960s and 1970s. Based on the results of these tests, column curves of welds referred to as ECCS a-, b- and c-curves (ECCS, 1978) are recommended for aluminum alloys, where a- and b-curves are accepted by ECCS. Heat treated and non-heat treated alloys, respectively. The major difference in the sequence of different curves for all heat-treated and non-heat-treated alloys is that the greater difference in softening of non-heat-treated alloys is greater than for heat-treated alloys. ECCS column curves cannot be used in accordance with the design as they are in tabular form.

Liu et. al. 2015 presented experimental and quantitative research on the buckling behavior of shape alloyed aluminum columns under axial compression. The structural component having the cross section examined is generally used as columns in an aluminum alloy framed structure as shown in Figure 3.1 (a). The Fiber Reinforced Plastic (FRP) wall can be easily fixed to the channel of the section as shown in Figure 3.1 (b).

Figure 3.1. Application of aluminum alloy column with section.

Advantages of using aluminum alloys as a structural material are high strength / weight ratio, formability, electrical-thermal conductivity, recycle, light weight, corrosion resistance and production purpose. There are the disadvantages of using aluminum alloys for structural applications, such as the aluminum modulus, which is roughly one third of the steel and causes the aluminum element to fail easily by buckling. Previous research of aluminum structural elements focused mainly on the ultimate strength of compact (non-thin) sections. However, the use of aluminum thin-walled sections has increased in recent years. In aluminum structures, welds are divided into two types: (1) transverse welds, (2) longitudinal welds, to share their effects on element strength. While transverse welds are often used in connections, longitudinal welds are used for the production of built-up elements. However, there is little research on the behavior of aluminum columns containing transverse welds.

Therefore, it is necessary to investigate the behavior and design of aluminum columns, beams and beam columns of thin sections.

3.2 Artificial Intelligence (AI)

From the beginning of time to the end, it went hand in hand with technological advances to improve the lives of people. Since the first tools and the discovery of the wheel, technology, the concrete offspring of human intelligence, has made humanity longevity, comfort, and provided means for understanding the cosmos that surrounds it (Makaritou, 2019). The disrupting processes are named Revolutions denoting that those events are abrupt and radical in nature. The progress of a human kind and civilizations has been driven by the progress of technology. Though, perception of a word technology of a modern person might be associated more with a digital technology, such as computers, internet and smartphones, for the ancient people, their technology are invention of the wheel and utilization of simple tools (Rahimov, 2019). In fact, since human beings survive in a digital age, it is very important to examine and investigate related issues further and make logical predictions to come to a conclusion and eventually prepare for developments in future differences in industries. Artificial intelligence can be broadly defined as the intelligence formed by software programs in machines, which can be seen as similar to the natural intelligence shown in humans and other animals. In practice, intelligence is attributed to artificial entities when those entities simulate human-like functions such as learning and problem solving features (Özmen, 2019). Artificial intelligence applications are used in many areas to make everyday life more convenient and efficient. They are used to clean houses, take care of old or sick people, to work at dangerous jobs replacing human beings, to give medical advisory and healthcare services, to prevent fraudulent situations in finance, etc. AI is just one of the areas of computer science that aims to create intelligent machines that work and respond, such as human preparation, solutions to problems, speaking, recognizing, introducing, teaching and learning. Also, Artificial intelligence is a branch of computer science that aims to create intelligent machines. It has become an necessary part of the technology production. Research related with artificial intelligence is highly technical and specialized. The core problems of artificial intelligence contain programming computers for certain traits such as: Information, sensivity, becomes skilled, program solving, ability to manipulate and move objects (Okyay, 2018).

Also it can be defined as, ANN is a programming technique that uses a large number of cells that mimic the functions of the human brain and functional connections between these cells. The functions on the created network (Figure 3.2) are optimized by changing the variables on the neurons that symbolize each cell. Using the input and output data obtained, ANN is trained and it is tried to reduce the error between the outputs of the ANN and the outputs that should have. When this error level goes down to the minimum level, ANN training is accepted as complete (Kara, 2019). Artificial neural networks (ANNs) or connective systems can be used in conjunction with the so-called unlimited inspiration and intelligent computer systems using biological neural networks. Such systems or programs "learn" to successfully consider tasks, taking into account examples, without programming a particular system with specific rules. In recognizing images in the system, they can define the desired contents by analyzing the sample images entered in the system in a healthy way. For example, they can learn to identify images that contain cats. Using "cat" or "no cat" in the entered data according to the desired result and using the results to identify cats in other images. They do this without prior knowledge of cats, such as feathers, sharp ears, tail, whiskers and cat-like faces. Instead, they automatically generate descriptive properties from the samples they process. An ANN is based on a collection of connected units or nodes called artificial neurons that loosely model neurons in a brain in a biological system. In this collection, each neuron has connection tasks. Each link, ie neurons, can transmit a signal to other neurons, ie other links, such as synapses in a biological brain. An artificial neuron that receives a signal can then process it, understand it, and point to the neurons connected to it. In this way, they can work in a systematic way depending on each other. This enables communication between neurons. In embodiments of the ANN system, the signal in a connection is a real number, and the output of information from each neuron is calculated by a nonlinear function of the sum of its inputs. These functions are calculated differently. These connections are called edges. Knowledgeable neurons and so-called connecting edges typically have a weight adjusted as learning progresses. Weight increases or decreases the strength of the signal in a link.

Neurons may have a threshold value at which a signal is sent only when the sum of the signals delivered exceeds this threshold. That is to say, the situation depends on the sum of the inputs. Typically, neurons exist in layers and are processed into the system in a healthy manner by means of edges. Different layers can perform different transformations in their input. Sometimes these conversions can be achieved. The signals probably cross the layers more

than once, and once this transition is made, the signals go from the first layer (the input layer) to the last layer (the output layer) as expected, and the system is completed. The major objective of the ANN approach is to solve the problems systematically, that is, not artificially, but as if they are natural within a human brain. However, over time, attention led to the fulfillment of specific tasks and deviations from biology. Over time, these problems led to different functions. ANNs have been used in a variety of tasks, including computer vision, communication, speech, processing, machine translation, social networks filtering for people, video games, medical diagnostics and even traditionally considered activities reserved for people. Over time, these tasks have made people's lives easier.

Figure 3.2. General view of an ANN.

Artificial neural networks are another data mining technique that should be used for estimation and classification. Neural networks are signal processing systems that attempt to mimic the biological nervous system by presenting a numerical data model of multiple neuronal compositions connected to a network (Haykin & Lippmann, 1994). Neurons founded in the input and output layers and hidden layers or layers, if any. When a neuron significantly entered the system, the weights associated with that neuron can change. This
means that the neuron will be more effective than other neurons of the same level as the neuron itself. Artificial neural networks learn by adjusting the weights between neurons. (Mumbuçoğlu, 2019).

3.3 Deep Learning

Deep Learning history can be traced back to 1943 when a computer model is developed by Walter Pitts and Warren McCulloch based on neural networks, with imitation of human brain. They used a combination of algorithms and mathematics they called "threshold logic" to mimic the thought process. Furthermore, in the 90s Yann LeCun & Yoshua Bengio (1995) take a further step from natural Multilayer Perceptron (MLP) with the sole purpose to reduce both high computational tasks and availability of high number of dimensions. Subsequent to that, a paper is published that presented a better pattern recognition system by canceling unrelated variables (Kamachy, 2019).

With the importance and popularity of analysis task, lots of algorithms have been tried for having better scores for this. Various natural language processing and machine learning techniques are applied for accuracy classification. Recently, deep learning algorithms like CNN and LSTM networks have shown state of the good scores in accuracy analysis. These scores are high in the literature but they are not satisfactory for the implementation needs in real life. Therefore, it is an important interest area for researchers. People have been trying different combinations and ensembles of these deep learning methods for increasing accuracy (Kamış, 2019).

Structurally, deep learning can be thought as a more complex form of artificial neural networks (ANN). There is a feed-forward structure in traditional neural networks. Each input neuron has a hidden layer to which each neuron is attached, and each neuron in the hidden layer is connected to each neuron in the output layer. For classification problems, the number of units in the output layer is equal 3 to the number of classes. Typically, a single linear output neuron is used to predict a continuous output. All connections in the network are lead from the input layer to the output layer, and it is possible to create deep networks by adding more hidden layers to which each neuron binds to each neuron in the following layer. Deep learning provides a very strong framework for the supervised learning process. By adding more units to a layer, a deep network can represent increased complexity functions. It is easy

to match an input vector to an output vector, with the help of human easily. But, that process can be accomplished through deep learning model with modeled large data sets and exemplary training samples. To put it more simply, classical (simple) neural networks have only one hidden layer, whereas deep neural networks have more than one hidden layer (Figure 3.3) (Paker, 2019).

Figure 3.3. Difference between a simple neural network and a deep learning neural network.

The main definition of DL is that it is a Neural Network with many hidden layers, so "deep" here refers to the depth of layers consisting more than 2 hidden layers. It provides automatic feature extraction by determining the properties of the input data which can be used as a pointer to label the input data accurately. Each layer extracts features from the output of previous layer. This is a revolution in computer vision tasks. In contrast, shallow networks required manual designs for feature extraction and a great amount of experience in the image processing field. On the other hand, transformations from input images to vectors led to losses of much interesting information.

There are 7 main applications that are mainly practiced in DL (Hordri et. al 2016):

- \Box Automatic Speech Recognition (ASR)
- \Box Image Recognition
- \Box Natural Language Processing
- \Box Drug Discovery and Toxicology
- \Box Customer Relationship Management

 \Box Recommendation Systems

- \Box Bioinformatics
- \Box Deep Learning Frameworks
- \Box Tensor Flow
- \Box Theano
- \Box Keras
- \Box Torch
- \Box Caff

On the other hand, Deep learning is part of a family of artificial neural networks based on a broader and more useful machine and learning methods. In this way, healthy learning systems are established by working with artificial neural networks together with deep learning. Learning may not always be controlled. Learning in different forms can be controlled, semicontrolled or not. Deep learning architectures such as deep neural networks, deep belief networks, repetitive neural networks and convoluted neural networks have been applied to fields such as computer vision, speech, communication, introduction, recognition, natural language processing, voice recognition, machine translation, bioinformatics, drug design. Deep learning has made human life easier as can be predicted after entering human life. Medical image analysis, material examination, and board game programs are comparable to human experts here, and in some cases have produced superior results. Usage area is increasing day by day.

Deep Learning is a machine learning method. It allows us to train the data set and the artificial intelligence to predict the output after a given or desired information is input. In this way, it provides convenient classifications and to reach the result after the required details are entered in the required information as soon as possible. Both supervised and unsupervised learning can be used to train artificial intelligence. The more data input, the better artificial intelligence features will be revealed. Because as classifications enter, the desired data will be reached more easily, but on the other hand, as the data increases, things will become more complex, and as they become complex, undesirable shifts from artificial intelligence to machine learning will occur. As it becomes more complex, the shifts from machine learning to deep learning will begin and user will see how useful learning is. In the machine learning, the

experiences of human beings to date are introduced to the machine by means of parameters. Because how much data, parameters, experience is entered in technology, results will increase in direct proportion. After entering the information such as pear and apple shape, color, size, stalk, the result will be easily reached. However, deep learning can learn this different on its own. Only by showing the apple and pear to the deep learning system itself creates its own rules, to reveal the differences in color and shape are the main distinguishing features itself. Thus, it can perform its operations by creating its own discriminatory abilities without the need of basic human abilities.

3.3.1 Deep Learning Library

The so-called deep learning library is actually a library of software systems. There are multiple systems, soft wares. These are ; TensorFlow(Python), Caffe(Python), Theano(Python), Keras(Python), Torch(C++), Deeplearning4j(Java), Covnetjs(Java), Mxnet(Python), PyLearn2(Python), Deep Learn Toolbox-Matlab, Accord.NET(C#), Sci-Kit Learn(Python), Accord.MachineLearning. These software systems work in different ways from each other. The common points are software, many of which are used for deep learning.

3.4 Keras

Keras is a python library for deep learning. It also works with Theano or Tensorflow libraries, which are also based on symbolic operations and used for deep learning. GPU or CPU work on these basic libraries and helps to reach a conclusion. Since it is a higher level library, user can develop applications more easily than Theano or Tensorflow. It is quite common use. It is a library that can be easily used by many developers around the world. Also many examples can even find about the problem in the Kaggle competitions. There are two main API structures in Keras. The models to be installed can be designed by installing one of these two different main structures. Sequential models have to be designed in layers. But it is more simple and understandable, comfortable and easy to use. Functional structure is designed as functions. It makes it possible to design more flexible and advanced models. Since the input layer data will be used in the structure call the model, the input size of the data must be specified. It does not need to be specified for, because the other layers receive data from the preceding layer. To get the results of the output layer, the output size must be determined. The output size is generally the output size of the layer before the last activation layer.

3.4.1 Keras Models

There are two ways to create Keras models: sequential and functional.The sequential API gives user opportunities to create multiple layer models for most problems.It is limited that it does not allow models that share layers or have multiple inputs or outputs.Alternatively, the functional API allows to create models with much more flexibility because they can easily identify models where layers are connected to more layers than only previous and later layers. In fact, user can connect layers to another layer (literally). As a result, siamese networks and now it becomes possible to create complex networks such as networks. Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear and actionable feedback upon user error. This makes Keras easy to learn and easy to use. As a Keras user, user are more productive, allowing user to try more ideas than userr competition, faster which in turn helps user win machine learning competitions. This ease of use does not come at the cost of reduced flexibility: because Keras integrates with lower-level deep learning languages (in particular TensorFlow), it enables to implement anything user could have built in the base language. In particular, as keras, the Keras API integrates seamlessly with userr Tensor Flow workflows (Figure 3.4).

Figure 3.4. Deep learning frameworks ranking.

With over 250,000 individual users as of mid-2018, Keras has stronger adoption in both the

industry and the research community than any other deep learning framework except TensorFlow itself (and the Keras API is the official frontend of TensorFlow, via the tf.keras module).User are already constantly interacting with features built with Keras -- it is in use at Netflix, Uber, Yelp, Instacart, Zocdoc, Square, and many others. It is especially popular among startups that place deep learning at the core of their products. Keras is also a favorite among deep learning researchers, coming in terms of mentions in scientific papers uploaded to the preprint some server. Keras has also been adopted by researchers at large scientific organizations, in particular CERN and NASA.

3.4.1.1 Functional API

The Keras functional API is the way to go for defining complex models, such as multi-output models, directed acyclic graphs, or models with shared layers. This guide already familiar with the Sequential model. A layer instance is callable (on a tensor), and it returns a tensor. Input tensor(s) and output tensor(s) can then be used to define, such a model can be trained just like Keras Sequential models. With the functional API, it is easy to reuse trained models: Any model can treat as if it are a layer, by calling it on a tensor. Note that by calling a model just reusing the architecture of the model, user are also reusing its weights.

3.4.1.2 Methods of Sequential

Choosing the right file editing system for database records or digital files affects how much data can be done and how useful and efficient user system is. The more organized the files are, the more efficient the system works. Sequential file organization means that computers store their data, information, or files in a specific file, not in a location, but in a specific order based on the data or file type. This sequence is processed regularly and makes the system useful. Sequential file organization is transparent to the user, and the methods of organizing sequential files work with a variety of data and different operating system environments. It helps the system to be preferred according to the needs and to install files to suit the preferred system, to obtain effective data and to convert it into different ways. The simplest sequential file organization is to save the files in the order in which user create them. The first file is saved first, the next is second and so on. This method is straight forward and it lets user to delete files without disturbing the sequence. While the method lets user place newly created files at the end of the sequence, there is no indication of the location of information. The low

complexity of the organizational method means finding information takes a long time since a search has to look through every file, starting with the first one. Indexed sequential files solve this problem but at a cost of organizational simplicity. The system lists files or data in a predetermined order, such as alphabetically. When user creates a file, the system either has to insert the new file in the proper sequence or re-index the whole file list. It takes longer to create new files and place them in the index, but it takes less time to find information because user know where in the indexed system to start looking. When designing a sequential file system, user have to look at how often user need to search for data and how often user plan to add or change files. Archival-type systems constantly grow as user adds new files but searches are comparatively rare. A system that files the information in sequence, as user archives it, is appropriate. A call-center database is relatively static but subject to constant searches. It makes sense to use an alphabetical or other index that lets user find specific records quickly. The two critical parameters for sequential file organization methods are speed and storage space. When user store files sequentially, there is no room to later make a file bigger. The system either has to store the additional information elsewhere or it has to make room and resave all the files. Both methods take a lot of additional time. Sequential file systems solve this problem by reserving extra space for each file when the system creates it. This avoids the delay but uses up a lot more space. When setting up the system, user has to decide whether time saved or space used is more valuable and use the corresponding method to organize the files.

The method of sequential analysis is first attributed to Abraham Wald with Jacob Wolfowitz, W. Allen Wallis, and Milton Friedman at Columbia University's Statistical Research Group. At the same time, George Barnard led a group working on optional stopping in Great Britain. Another early contribution to the method is made by K.J. Arrow with D. Blackwell and M.A. Girshick. A similar approach is independently developed from first principles at about the same time by Alan Turing, as part of the Banburismus technique used at Bletchley Park, to test hypotheses about whether different messages coded by German Enigma machines should be connected and analysed together. This work remained secret until the early 1980s.Peter Armitage introduced the use of sequential analysis in medical research, especially in the area of clinical trials. Sequential methods became increasingly popular in medicine following Stuart Pocock's work that provided clear recommendations on how to control Type 1 error rates in sequential designs.

3.5 Optimizers of Sequential

Adaptive Moment Estimation (Adam) is another method that computes adaptive learning rates for each parameter. In addition to storing an exponentially decaying average of past squared gradients Vt like Adadelta and RMSprop, Adam also keeps an exponentially decaying average of past gradients mt, similar to momentum. Norms for large pp values generally become numerically unstable, which is why ℓ 1 ℓ 1 and ℓ 2 ℓ 2 norms are most common in practice. However, $\ell \infty \ell \infty$ also generally exhibits stable behavior. For this reason, the authors propose AdaMax (Kingma and Ba, 2015) and show that vtvt with $\ell \infty \ell$ converges to the following more stable value. As we have seen before, Adam can be viewed as a combination of RMSprop and momentum: RMSprop contributes the exponentially decaying average of past squared gradients vtvt, while momentum accounts for the exponentially decaying average of past gradients mtmt. We have also seen that Nesterov accelerated gradient (NAG) is superior to vanilla momentum.

Given the ubiquity of large-scale data solutions and the availability of low-commodity clusters, distributing SGD to speed it up further is an obvious choice. SGD by itself is inherently sequential: Step-by-step, we progress further towards the minimum. Running it provides good convergence but can be slow particularly on large datasets. In contrast, running SGD asynchronously is faster, but suboptimal communication between workers can lead to poor convergence. Additionally, we can also parallelize SGD on one machine without the need for a large computing cluster. The following are algorithms and architectures that have been proposed to optimize parallelized and distributed SGD.

Whereas momentum can be seen as a ball running down a slope, Adam behaves like a heavy ball with friction, which thus prefers flat minima in the error surface Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, is introduced based on the thoughts and predictions made of adjustable simple moments. The procedures that need to be performed are easy to implement and pour into tools, are also very efficient in terms of calculations, require little memory, do not change in diagonal rescaling of gradients and are well suited for large problems in terms of data and / or parameters. This method is also suitable for non-stationary purposes and problems with very noisy and / or sparse gradients. Hyper parameters have intuitive interpretations and typically require minimal adjustments and adjustments. Empirical results show that Adam works well in practice and is compared to

other stochastic optimization methods. Finally, AdaMax, a variant of Adam based on the eternity norm (Kingma, 2014).

Reddi, Kale and Kumar is working More recently, stochastic optimization methods, which have been successfully and successfully used in the training of deep networks such as RMSProp, Adam, Adadelta, Nadam, have been imposed on using gradient updates scaled with square roots of exponential moving averages of square past gradients. In many types of applications, e.g. learning with large output spaces, empirically it has been observed that these algorithms cannot approach an optimal result and solution (or a critical point in non-convex environments). One reason for such errors is the exponential moving average used in algorithms. Here is a clear example of a simple convex optimization setting where Adam cannot merge with the optimal solution, and user can see the exact problems in the previous analysis of the Adam algorithm. Conducted analyzes indicate that convergence problems can be corrected by equipping such algorithms with long term memory of past gradients, and that the Adam algorithm not only corrects convergence problems, but can also suggest new variants that often result in improved empirical performance.

3.6 Evaluation Methods

In many sciences, as mentioned, the use of qualitative or quantitative (concrete or abstract) methods has become a matter of debate, with particular schools of thought in each discipline favoring one method and disdains another. Qualitative methods argue that quantitative methods are meant to conceal the reality of the social events studied because they underestimate, ignore, ignore or neglect the most immeasurable factors. The modern tendency (and indeed the majority tendency throughout the history of social science) is to use eclectic approaches: quantitative methods can be used with a global qualitative framework, and qualitative methods can be used to understand the meaning of numbers generated by quantitative methods.

But even this practice has raised a debate as to whether quantitative and qualitative research methods are complementary: some investigators argue that combining these two different approaches to a research method is useful and that the social world helps create a more complete situation. They believe and trust that the epistemologies that support each approach are so divergent that they cannot be reconciled even in a research project (Gadd, 2006).

3.6.1 MAD (Mean Absolute Deviation)

The mean absolute deviation (or mean absolute deviation (MAD)) about a given point of a data set (or only "mean absolute deviation") is the average of the absolute deviations of the data given or the positive difference (Usually central values). It is a summary statistics of statistical distribution or variability. In the general form, the central point may be the mean, median, mode, or any central trend measure, or the result of any random data point associated with the given data set. The absolute values of the difference between the data points and their central tendencies are summed and divided by the number of data points.

3.6.2 MAPE

Many organizations focus primarily on the MAPE when assessing forecast accuracy. Most people are comfortable thinking in percentage terms, making the MAPE easy to interpret. It can also convey information when user doesn't know the item's demand volume. For example, telling user manager, "we are off by less than 4%" is more meaningful than saying "we are off by 3,000 cases," if user manager doesn't know an item's typical demand volume.

The MAPE is scale sensitive and should not be used when working with low-volume data. Notice that because "Actual" is in the denominator of the equation, the MAPE is undefined when Actual demand is zero. Furthermore, when the Actual value is not zero, but quite small, the MAPE will often take on extreme values. This scale sensitivity renders the MAPE close to worthless as an error measure for low-volume data.

3.6.3 MEAN (Actual / Predicted)

In statistics, a forecast error is the difference between the actual or real and the predicted or forecast value of a time series or any other phenomenon of interest. Since the forecast error is derived from the same scale of data, comparisons between the forecast errors of different series can only be made when the series are on the same scale.

In simple cases, a forecast is compared with an outcome at a single time-point and a summary of forecast errors is constructed over a collection of such time-points. Here the forecast may be assessed using the difference or using a proportional error. By convention, the error is defined using the value of the outcome minus the value of the forecast.

In other cases, a forecast may consist of predicted values over a number of lead-times; in this case an assessment of forecast error may need to consider more general ways of assessing the match between the time-profiles of the forecast and the outcome. If a main application of the forecast is to predict when certain thresholds will be crossed, one possible way of assessing the forecast is to use the timing-error—the difference in time between when the outcome crosses the threshold and when the forecast does so. When there is interest in the maximum value being reached, assessment of forecasts can be done using any of, the difference of times of the peaks the difference in the peak values in the forecast and outcome, the difference between the peak value of the outcome and the value forecast for that time point. Forecast error can be a calendar forecast error or a cross-sectional forecast error, when we want to summarize the forecast error over a group of units. If we observe the average forecast error for a time-series of forecasts for the same product or phenomenon, then we call this a calendar forecast error or time-series forecast error. If we observe this for multiple products for the same period, then this is a cross- sectional performance error (Grafe et al., 2010).

3.6.4 Standart Deviation

The standard deviation is a system used for a measure of the amount of variation or distribution of a set of values. A low standard deviation indicates that the values tend to be close to the average of the set, while a high standard deviation indicates that the values are spread over a wider range. In this way, the distance and proximity to the intended or targeted results are found. The standard deviation of a random variable, statistical population, data set

or probability distribution is the square root of its variance. In practice it is algebraically simpler, although less robust than the average absolute deviation. A useful feature of standard deviation is that, in contrast to variance, it is expressed in the same units as the data. In addition to expressing the variability of a population, standard deviation is commonly used to measure confidence in statistical results. For example, the margin of error in the query data is determined by calculating the expected standard deviation in the results if the same survey is performed more than once. This derivation of the standard deviation is often referred to as the mean while the estimate is called "standard error" or "standard error of average". If an infinite number of samples are drawn and an average is calculated for each sample, it is calculated as the standard deviation of all instruments to be calculated from this population. The standard deviation of a population and the standard mean of a statistic derived from this population are quite different, but related. The reported error margin of a survey is calculated from the standard error of the mean. In science, many investigators report the standard deviation of experimental data, results, and, by convention, more than just two standard deviations from an empty expectation are considered statistically significant. Normal random error or change in measurements is thus distinguished from the possible original. Standard deviation is also important in finance where the standard deviation in the return on investment is a measure of the volatility of the investment.

3.6.5 Correlation

Correlation or dependence is any statistical relationship, whether causal or not, between two random variables or bivariate data. In the broadest sense correlation is any statistical association, though it commonly refers to the degree to which a pair of variables are linearly related. Familiar examples of dependent phenomena include the correlation between the physical statures of parents and their offspring, and the correlation between the demand for a limited supply product and its price.[citation needed]

Correlations are useful because they can indicate a predictive relationship that can be exploited in practice. For example, an electrical utility may produce less power on a mild day based on the correlation between electricity demand and weather. In this example, there is a causal relationship, because extreme weather causes people to use more electricity for heating or cooling. However, in general, the presence of a correlation is not sufficient to infer the presence of a causal relationship.

CHAPTER 4 RESULTS & DISCUSSIONS

4.1 Overall Results

In this section, the predictive results of the deep learning model using the data obtained from the study named 'Flexural buckling load prediction of aluminum alloy columns using soft computing techniques are compared with the results obtained from soft computing techniques in the same study. Sequential model of the learning method of the optimizers Adam, Adamax, RMSprops, Nadam, Adagrad, Adadelta are used and the model compiled with MAE and MSE loss is run 10 times; optimum results are used for evaluation. Dataset is divided into training set and testing set. The training set is 70% and the test set is 30%. Experimental values for 104 tests in Table 4.2 given with related material parameters.

The description of material parameters are shown in Table 4.1 taken from Cevik et al., (2009).An analytical statement was proposed (Frey and Rondal, 1978; Rondal and Maquoi, 1979) and was accepted by the ECCS. The calculation of the non-dimensional column strength is shown as:

$$
X = \frac{1}{\theta + \sqrt{\varphi^2 - \lambda^2}},
$$

$$
\varphi = \frac{1}{2}(1 + \eta + \lambda^2),
$$

The constants of the material parameters are $\alpha, \beta, \lambda_0, \lambda_1$. These are shown as:

$$
\alpha(n,e)=\frac{1.5}{(e^{0.6}+0.03)(n(\frac{0.0048}{e^{0.55}})^{1.4}+1.3)}+\frac{0.02}{e^{0.6}},
$$

$$
\beta(n, e) = \frac{0.36 \exp(-n)}{e^{0.45} + 0.007} + \tanh\left(\frac{n}{180} + \frac{6 \times 10^{-6}}{e^{1.4}} + 0.04\right),
$$

 $\lambda_0(n,e) = 0.82 \left(\frac{e}{e^{1.0.6}} \right)$ $\frac{e}{e+0.0004} - 0.01n \ge 0.2$

$$
\lambda_1(n,e)=0.8\frac{e}{e+0.0018}\left(1-\left[\left(\frac{n-55}{n+\frac{6e-0.0054}{e+0.0015}}\right)^2\right]^{0.6}\right).
$$

where $\sigma_{0.2}$, L and r are the ultimate stress, effective length and Radius of gyration respectively.

$$
\chi = \frac{\sigma_u}{\sigma_{0.2}}, \quad \lambda = \sqrt{\frac{\sigma_u}{\sigma_{E_0}}}, \quad \sigma_{E_0} = \frac{\pi^2 E_0}{(L/r)^2}
$$

The flowchart of the application and execution of deep learning model used in the thesis and is shown in Figure 4.1. According to this chart, Firstly, The data set is scaled to give more correct results of Model. After that, the parameters of the model is defined and the dataset is trained and evaluated. If the execution count is smaller than 10; the application is executed until obtain the better results than soft computing techniques. If the execution count is equal 10; the last results are taken for comparing.

Figure 4.1. The Flowchart of the Application of Model

4.2 Comparison of Nadam Optimizer's Results and Laboratuar Test Results

When Nadam is used as an optimizer and MAE is used as loss, the model's train and test accuracy are shown in Figure 4.2 Accuracy of Test dataset is 9.38 and Accuracy of Train Dataset is 82.99. Mean of Test Dataset 51.56 and mean of train dataset is 23.51. The results obtained by using both loss parameters with the Nadam optimizer of the deep learning method with GEP and NN are compared with the previous methods, GEP and NN is shown in Table 4.3. For *MSE, MAD, RMSE, MEAN and Correlation*, the best results are seen in Nadam Optimizer and MAE Loss. For *MAE, MAPE,* the best results are seen in Nadam Optimizer and MSE Loss. In this case, it can be said that MAE loss function is more successful for Nadam optimizer in deep learning model.

	MAD	MSE	RMSE	MAE	MAPE	MEAN	STD	Correlation
GEP	11.9448	300.4576	17.3337	0.0833	8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553	10.8192	0.0419	4.1852	0.9988	0.0586	0.9915
NADAM & MAE	6.4505	109.0971	10.4450	0.0384	3.8352	0.9992	0.0622	0.9920
NADAM & MSE	6.4625	123.1020	11.0951	0.0378	3.7775	1.0097	0.0611	0.9912

Table 4.3. Statistical Results of MAE and MSE loss for Nadam

In the graph in Figure 4.3, test results and model results are compared in terms of value. According to this graph, it can be said that the range with the highest deviations is between 150-250 and the R² value is 0.9841.

Nadam&MAE

Figure 4.3. Comparison of Test Results and Predicted Results of Nadam Optimizer and MAE loss

105 rows in the dataset is an example for the deep learning results for Nadam & MAE as seen in Figure 4.4. When the closeness to the correct result is examined, it is observed that in 64 of 105 samples, deep learning model gives more realistic results than NN method.

Figure 4.4. Actual datas and predicted results for Nadam Optimizer&MAE loss

The results of each line are compared with GEP and NN, which are the soft computing techniques, and the MAEs of the results obtained from the deep learning model are compared in Figure 4.5. As can be seen in the figure, the highest difference between the results belongs to GEP; it is observed that the points where the errors are minimum belong to the deep learning results.

Figure 4.5. MAE of Each Results for Nadam Optimizer and MAE loss

Train and test accuracy of the model when Nadam is used as optimizer and MSE is used as loss in Figure 4.6 seen. The comparison of the test results of each row in the dataset with the estimation results from the model is as seen in Figure 4.7. Mean of Test Dataset 172.10 and mean of train dataset is 34.35. Std of Test Dataset is 165.86 and std of train dataset is 31.67.

Figure 4.6. Comparison testing and training accuracy for Nadam Optimizer and MSE loss

Figure 4.7. Actual datas and predicted results for Nadam Optimizer&MSE loss

The comparison of the test results in the dataset with the estimation results from the model is shown in Figure 4.7. When the closeness to the correct result is examined, it is observed that the deep learning model gives more realistic results than the NN method in 67 of 105 samples.

In the Figure 4.8, the test results and the model results are compared in terms of value. According to this graph, it can be said that the range with the most deviations is 150-200 and $R²$ value is 0.9824.

Figure 4.8. Comparison of Test Results and Predicted Results of Nadam Optimizer and MSE loss.

The results of each line are compared with GEP and NN, which are Soft Computing techniques, and the MAEs of the results obtained from the deep learning model are compared in Figure 4.9. As shown in the figure, the value at the point where the differences between the results are highest belongs to GEP. Except for 6 points, deep learning results are found to be successful.

Figure 4.9. Mean Absolute Error of Each Results for Nadam Optimizer and MSE loss.

4.3 Comparison of Adamax Optimizer's Results and Laboratuar Test Results

In Table 4.4, the results of GEP and NN and Adamax optimizer of the deep learning method are compared using the two loss parameters. According to this table, the best results for *MAD, MAE, and MAPE* are found in Adamax Optimizer and MAE Loss. ; In *Correlation, MSE, RMSE,* Optimizer and MSE Loss, the last one is for standard deviation and mean for NN method. In this case, MSE and MAE Loss functions for Adamax optimizer can be interpreted as having almost equal performance.

	MAD	MSE	RMSE	MAE		MAPE MEAN	Std	Correlation
GEP	11.9448	300.4576 17.3337 0.0833			8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553 10.8192 0.0419			4.1852	0.9988	0.0586	0.9915
ADAMAX & MAE	5.8172	115.3379 10.7395 0.0337			3.3702	1.0169	0.0596	0.9922
ADAMAX & MSE	6.4842	$110.5480 \mid 10.5142 \mid 0.0463$			4.6345	1.0309	0.0657	0.9927

Table 4.4. Statistical Results of Adamax Optimizer.

Train and test accuracy of the model using Adamax as the optimizer and MAE as the loss as seen in Figure 4.10. The comparison of the test results of each row with the prediction results from the model is as presented in Figure 4.11. Accuracy of Test dataset is 9.38 and Accuracy of Train Dataset is 82.99.Mean of Test Dataset 59.002 and mean of train dataset is 23.730. Std of Test dataset is 91.266 and std of train dataset is 35.574.

Figure 4.10.Comparison testing and training accuracy for Adamax Optimizer and MAE loss.

Comparison of the test results in the dataset with the estimation results from the model as presented in Figure 4.11. When the closeness to the correct result is examined, it is observed that in 68 of 105 samples, deep learning model gives more realistic results than NN method.

In the graph in Figure 4.12, test results and model results are compared in terms of value. According to this graph, it can be said that the maximum deviations are between 160-180 and $R²$ value is 0.9845.

Figure 4.12. Comparison of Test Results and Predicted Results of Adamax Optimizer and MAE loss.

The results of each line are compared with GEP and NN, which are the soft computing techniques, and the MAEs of the results obtained from the deep learning model are compared in Figure 4.13. As can be seen in the figure, it is observed that the points where the errors are minimum belong to the deep learning results.

Figure 4.13. MAE of Each Results for Adamax Optimizer and MAE.

Train and test accuracy of the model using Adamax as the optimizer and MSE as the loss as seen in Figure 4.14. The comparison of the test results of each row with the prediction results from the model is as presented in Figure 4.15. Accuracy of Test dataset is 9.93 and Accuracy of Train Dataset is 78.40.Mean of Test Dataset 99.978 and mean of train dataset is 40.498. Std of Test dataset is 95.05 and std of train dataset is 37.953.

Figure 4.14. Comparison testing and training accuracy for Adamax Optimizer and MSE loss.

Comparison of the test results in the dataset with the estimation results from the model figure 4.15 as in. When the closeness to the correct result is examined, it is observed that in 65 of 105 samples, deep learning model gives more realistic results than NN method.

Figure 4.15. Actual datas and predicted results for Adamax Optimizer and MSE loss.

In the graph in Figure 4.16, test results and model results are compared in terms of value. According to this graph, it can be said that the maximum deviations are between 200-250 and $R²$ value is 0.9855.

Figure 4.16. Comparison of Test Results and Predicted Results of Adamax Optimizer and MSE Loss

The results of each line are compared with GEP and NN, which are the soft computing techniques, and the MAEs of the results obtained from the deep learning model are compared in Figure 4.17. As can be seen in the figure, the highest difference between the results belongs to GEP; it is observed that the points where the errors are minimum belong to the deep learning results.

Figure 4.17. MAE of Each Results for Adamax Optimizer and MSE.

4.4 Comparison of Adam Optimizer's Results and Laboratuar Test Results

Table 4.5 compared the results obtained by using both loss parameters with the Adam optimizer of the deep learning method with GEP and NN. According to this table, the best results for *MAD, MSE, RMSE, MEAN, correlation* are found in Adam Optimizer and MAE Loss; *MAE, MAPE, MEAN for* the NN method. According to this situation, MAE Loss function is better for Adamax optimizer in deep learning model.

	MAD	MSE	RMSE	MAE	MAPE	MEAN	Std	Correlation
GEP	11.9448	300.4576	17.3337	0.0833	8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553	10.8192	0.0419	4.1852	0.9988	0.0586	0.9915
ADAM & MAE	6.5120	100.0853	10.0043	0.0441	4.4052	1.0138	0.0767	0.9928
ADAM & MSE	6.9639	112.9846	10.6294	0.0461	4.6051	1.0174	0.0628	0.9920

Table 4.5. Statistical Results of Adam Optimizer.

Train and test accuracy of the model using Adam as the optimizer and MAE as the loss according to Figure 4.18. The comparison of the test results of each row with the prediction results from the model is showed in Figure 4.19. Mean of Test Dataset 54,600 and mean of train dataset is 85,735. Std of Test dataset is 5.347 and std of train dataset is 35.642 for Adam optimizer and MAE loss.

Figure 4.19. Actual datas and predicted results for Adam Optimizer and MAE loss.

In the graph in Figure.4.20, test results and model results are compared in terms of value. According to this graph, the range of 160-230 is the range with the highest deviation and \mathbb{R}^2 value is 0.9857.

.

Figure 4.20. Comparison of Test Results and Predicted Results of Adam Optimizer and MAE loss.

When the deep learning model of each row is run with GEP and NN using Adam Optimizer and MAE loss functions, the MAEs of the results are compared in Figure 4.21. As can be seen in the figure, the difference between the results at the point where the highest value belongs to GEP; the closest points to real values belong to deep learning.

Figure 4.21. MAE of Each Results for Adam Optimizer and MAE.

Train and test accuracy of the model using Adam as the optimizer and MAE as the loss Figure 4.22. The comparison of the test results of each row with the prediction results from the model is indicated in Figure4.23. Mean of Test Dataset 111.816 and mean of train dataset is 40.457. Std of Test Dataset is 106.199 and std of traindataset is 37.961 for Adam optimizer and MSE loss.

The comparison of the test results in the dataset and the estimation results from the model is as in figure 4.23. When the closeness to the correct result is examined, it is observed that in 52 of 105 samples, deep learning model gives more realistic results than NN method.

In the graph in Figure 4.24, test results and model results are compared in terms of value. According to this graph, deviations between 160-180 and 270-300 are maximum and \mathbb{R}^2 value is 0.9841

Figure 4.24. Comparison of Test Results and Predicted Results of Adam Optimizer and MSE loss.

When the deep learning model of each row is run with GEP and NN using Adam Optimizer and MSE loss functions, the MAEs of the results are compared in Figure 4.25. As can be seen in the figure, the difference between the results at the point where the highest value belongs to GEP; When the closeness to real values is considered, it is seen that NN is ahead in some points and deep learning results are closer in some points.

Figure 4.25. MAE of Each Results for Adam Optimizer and MSE

4.5 Comparison of Adadelta Optimizer's Results and Laboratuar Test Results

In Table 4.6, the results obtained by using both the loss parameters and the Adadelta optimizer of the deep learning method with the previous methods GEP and NN are compared. According to this table, the best results are found for *MAD, MSE, RMSE, MAE, MAPE, MEAN and correlation.* The standard deviation for Adadelta Optimizer and for MAE Loss is much better than NN method.

	MAD	MSE	RMSE	MAE		MAPE MEAN Std		Correlation
GEP	11.9448	300.4576 17.3337		0.0833	8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553 10.8192		0.0419	4.1852	0.9988	0.0586	0.9915
ADADELTA & MAE	6.3152	114.0623 10.6800		0.0393	3.9315	1.0178	0.0649	0.9919
ADADELTA & MSE	7.2662	120.5526 10.9796		0.0460	4.5992	1.0031	0.0666	0.9915

Table 4.6. Statistical Results of Adadelta Optimizer.

When using Adadelta as optimizer and MAE as loss, the model's train and test accuracy is shown in Figure 4.26. The comparison of the test results of each row with the prediction results from the model likes in Figure 4.27. Mean of Test Dataset 50.836 and mean of train dataset is 23.599. Std of Test Dataset is 76.909 and std of train dataset is 35.286 for Adam optimizer and MAE loss.

Figure 4.26. Comparison testing and training accuracy for Adadelta Optimizer and MAE loss.

Comparison of the test results in the dataset with the estimation results from the model illustrated in Figure 4.27. When the closeness to the correct result is examined, it is observed that in 63 of 105 samples, the deep learning model gave more realistic results than the NN method.

Figure 4.27.Actual datas and predicted results for Adadelta Optimizer and MAE loss.

In the graph in Figure.4.28, the test results and the model results are compared in terms of value. According to this graph, the range of 230-280 is the range with the highest deviations and R^2 value is 0.9838.

ADADELTA&MAE

Figure 4.28.Comparison of Test Results and Predicted Results of Adadelta Optimizer and MAE loss.

When the deep learning model of GEP and NN of each row is run using Adadelta Optimizer and MAE loss functions, the MAEs of the results are compared in Figure 4.29. As can be seen in the figure, the difference between the results at the point where the highest value belongs to GEP; the closest points to real values belong to deep learning.

Figure 4.29. MAE of Each Results for Adadelta Optimizer and MAE.

Train and test accuracy of the model when using Adadelta as optimizer and MAE as loss also seen in Figure 4.30. The comparison of the test results of each row in the dataset with the estimation results from the model is as in Figure 4.31.Mean of Test Dataset 111.792 and mean of train dataset is 46.638. Std of Test dataset is 106.252 and std of train dataset is 43.672 for Adam optimizer and MAE loss.

Figure 4.30. Comparison testing and training accuracy for Adadelta Optimizer and MSE

loss.

Comparison of the test results in the dataset with the prediction results from the model is illustrated in Figure 4.31. When the closeness to the correct result is examined, it is observed that deep learning model gives closer to the real results than the NN method in 56 of 105 samples.

Figure 4.31. Actual datas and predicted results for Adadelta Optimizer and MSE loss.

In the graph in Figure.4.32, test results and model results are compared in terms of value. According to this graph, the range of 270-290 and 180-210 are the ranges with the highest deviations and \mathbb{R}^2 is 0.983.

Figure 4.32. Comparison of Test Results and Predicted Results of Adadelta Optimizer and MSE loss.

When the deep learning model of each row with GEP and NN is run using Adadelta Optimizer and MSE loss functions, the MAEs of the results are compared in Figure 4.33.

Figure 4.33. MAE of Each Results for Adadelta Optimizer and MSE.

4.6 Comparison of Adagrad Optimizer's Results and Laboratuar Test Results

The model results that used Adagrad optimizer are shown in Table 4.7.

	MAD	MSE	RMSE	MAE	MAPE	MEAN	Std	Correlation
GEP	11.9448	300.4576	17.3337	0.0833	8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553	10.8192	0.0419	4.1852	0.9988	0.0586	0.9915
ADAGRAD & MAE	6.9978	136.4832	11.6826	0.0452	4.5225	1.0068	0.0754	0.9902
ADAGRAD & MSE	7.0842	134.2247	11.5855	0.0459	4.5881	1.0077	0.0712	0.9903

Table 4.7. Statistical Results of Adagrad Optimizer.

When using Adagrad as an optimizer and MAE as a loss, the model's train and test accuracy are also seen in Figure 4.34. The comparison of the test results for each row in the dataset with the predicted results from the model is demonstrated in Figure 4.35. Mean of Test

Dataset 71.730 and mean of train dataset is, 22.988. Std of Test Dataset is 110.752 and std of train dataset is 34.118 for Adagrad optimizer and MAE loss.

Model accuracy

Figure 4.34. Comparison testing and training accuracy for Adagrad Optimizer and MAE loss.

Comparison of the test results in the dataset with the estimation results from the model as in Figure 4.35..When the closeness to the correct result is examined, it has been observed that deep learning model gives closer results to the reality than the NN method in 64 of 105 samples.

Figure 4.35. Actual datas and predicted results for Adagrad Optimizer and MAE loss.

In the graph in Figure 4.36, test results and model results are compared in terms of value. According to this graph, the range of 220-270 is the range with the highest deviations and \mathbb{R}^2 value is 0.9805

Figure 4.36. Comparison of Test Results and Predicted Results of Adagrad Optimizer and MAE loss.

When the deep learning model of GEP and NN of each line is run using Adagrad Optimizer and MAE loss functions, the differences between the results and the actual results are compared in Figure 4.37. As shown in the figure, the value at the point where the differences between the results are highest belongs to deep learning.

Figure 4.37. MAE of Each Results for Adagrad Optimizer and MAE.

When using Adagrad as optimizer and MSE as loss, the model's train and test accuracy are indicated in Figure 4.38. The comparison of the test results for each row in the dataset and the predicted results from the model is as in Figure 4.39. Mean of Test Dataset 130.897 and mean of train dataset is 42.285. Std of Test dataset is 125.127 and std of train dataset is 39.741 for Adagrad optimizer and MAE loss.

Comparison of the test results in the dataset with the estimation results from the model seen in Figure 4.39. When the closeness to the correct result is examined, it is observed that deep learning model gave closer results to the reality than the NN method in 57 of 105 samples.

Figure 4.39. Actual datas and predicted results for Adagrad Optimizer and MSE loss. Figure 4.40 in the graph of the test results are compared in terms of the value of the model

results. According to this graph, the range of 160-180 and 240-820 is the range with the highest deviations and \mathbb{R}^2 value is 0.9806.

Figure 4.40. Comparison of Test Results and Predicted Results of Adagrad Optimizer and MSE loss.

When the deep learning model of GEP and NN of each row is run using Adagrad Optimizer and MSE loss functions, the differences between the results and the actual results are compared in Figure 4.41. As can be seen in the figure, the difference between the results at the highest point, although the value of deep learning is close to the point of 0 is also remarkable.

Figure 4.41. MAE of Each Results for Adagrad Optimizer and MSE.

4.7 Comparison of RMSProp Optimizer's Results and Laboratuar Test Results

In Table 4.8, the results obtained by using both loss parameters are compared with the previous methods GEP and NN and RMSProp optimizer of deep learning method. According to this table, the best results for *MAD, MAE, MAPE, RMSE, MSE, correlation are* seen in RMSProp Optimizer and MAE Loss. In this case, we can say that the MAE loss function is more successful in the deep learning model for RMSPROP Optimizer.

Also, according to Standart Deviation evaluation results, NN method of soft computing techniques is better than deep Learning methods that used RMSProp Optmizer.

	MAD	MSE	RMSE	MAE	MAPE MEAN STD		Correlation
GEP		11.9448 300.4576 17.3337 0.0833 8.3350 1.0322				0.1308	0.9782
NN	7.2895	\vert 117.0553 10.8192 0.0419 4.1852 0.9988 0.0586					0.9915

Table 4.8. Statistical Results of RMSPROP Optimizer.

When RMSPROP is used as an optimizer and MAE is used as a loss, the model's train and test accuracy are shown in Figure 4.42.

Figure 4.42. Comparison testing and training accuracy for RMSPROP Optimizer and MAE loss.

The comparison of the test results for each row in the dataset and the predicted results from the model is as shown in Figure.4.43. Mean of Test Dataset 91.529 and mean of train dataset is 45.341.Std of Test Dataset is 142.794 and std of train dataset is 66.679 for RMSPROP optimizer and MAE loss.

Figure 4.43. Actual datas and predicted results for RMSPROP Optimizer and MAE loss.

In the Figure 4.44, test results and model results are compared in terms of value. According to this graph, the range of 270-300 is the range with the highest deviations and \mathbb{R}^2 is 0.985.

RMSPROP&MAE

Figure 4.44. Comparison of Test Results and Predicted Results of RMSPROP Optimizer and MAE loss.

When the deep learning model of GEP and NN of each row is run using RMSPROP Optimizer and MAE loss functions, the differences between the results and the actual results are compared in Figure 4.45. As shown in the figure, the value at the point where the differences between the results are highest belongs to deep learning.

Figure 4.45. MAE of Each Results for RMSPROP Optimizer and MAE.

Train and test accuracy of the model when using RMSPROP as optimizer and MSE as loss in Figure 4.46 also seen. The comparison of the test results for each row in the dataset and the predicted results from the model is as in Figure.4.47. Mean of Test Dataset 106.457 and mean of train dataset is 49.272. Std of Test Dataset is 101.364 and std of train dataset is 46.305 for RMSPROP optimizer and MSE loss.

Figure 4.46. Comparison testing and training accuracy for RMSPROP Optimizer and MSE loss.

Comparison of the test results in the dataset with the prediction results from the model is illustrated in Figure 4.47. When the closeness to the correct result is examined, it is observed that deep learning model gave closer results to the reality than the NN method in 60 of 105 samples.

Figure 4.47. Actual datas and predicted results for RMSPROP Optimizer and MSE loss.

In the Figure 4.48 graph is compared with the test results in terms of model results. According to this graph, the range of 240-300 is the range with the highest deviations and \mathbb{R}^2 is 0.9844

RMSPROP&MSE

Figure 4.48. Comparison of Test Results and Predicted Results of RMSPROP Optimizer and MSE loss.

When the deep learning model of GEP and NN of each row is run using RMSProp Optimizer and MSE loss functions, the differences between the results and the actual results are compared in Figure 4.49. As can be seen in the figure, the difference between the results is the highest point of the GEP and 2-3 points except RMSProp optimizer and MSE loss using the deep learning model generally achieves the best results.

Figure 4.49. MAE of Each Results for RMSPROP Optimizer and MSE.

When using RMSPROP as optimizer and MSE as loss, the train and test accuracy of the model Figure 4.50. The comparison of the test results for each row in the data set and the estimation results from the model is as in Figure.4.51. Std of Test Dataset is 101.364 and std of train dataset is 46.305 for RMSPROP optimizer and MSE loss.

Figure 4.50. Comparison testing and training accuracy for RMSPROP Optimizer and MSE loss.

Comparison of the test results in the dataset with the estimated results from the model that seen in Figure 4.51. When the closeness to the correct result is examined, it is observed that in

60 of 105 samples, deep learning model gives more realistic results than NN method.

In the graph, the test results and the model results are compared in terms of value in Figure 4.52. According to this graph, the range of 240-300 is the range with the highest deviations and R^2 value is 0.9844.

Figure 4.52. Comparison of Test Results and Predicted Results of RMSPROP Optimizer and MSE loss.

When the deep learning model is run using the Adagrad Optimizer and MSE loss functions, the differences between the results and the actual results of each row are compared with GEP and NN, in Figure 4.53.As can be seen in the figure, the highest difference between the results is GEP's result and 2-3 points except the RMSPROP optimizer and MSE loss using the deep learning model generally achieves the best results.

Figure 4.53. MAE of Each Results for RMSPROP Optimizer and MSE Loss.

4.8 Comparison of All Optizimer for MAE loss in Statistical Results

When all the optimizers are compared for MAE loss according to the Table 4.9, it is seen that the best result for *RMSE, MSE* is taken from the deep learning model using Adam optimizer. For *MAD, MAPE, MAE, Standard deviation* and *correlation*, the best result appears to be the ADAMAX optimizer.

	MAD	MSE	RMSE	MAE	MAPE MEAN STD	Correl ation
GEP	11.9448	300.4576	17.3337 0.0834 8.3350 1.0322 0.1308 0.9782			
NN	7.2895	117.0553	$10.8192 0.0419 4.1852 0.9988 0.0586 0.9915$			

Table 4.9. Statistical Results of All Optizimer for MAE loss.

Also, The Comparison of All Optimizer's results are shown in Figure 4.54 for MAE Loss.

Figure 4.54. Actual values and Predicted values of all Optimizer and MAE Loss.

4.9 Comparison of All Optizimer for MSE loss in Statistical Results

When all the optimizers are compared for MSE loss according to the Table 4.10, it is seen that the best result for MAD, MAE, MAPE is taken from the deep learning model using Nadam optimizer. The best results for *RMSE, MSE and correlation* are seen in ADAMAX optimizer, and in Adadeltada for *Mean, standard deviation*. In addition, it is observed that the NN model for standard deviation is closer to zero with a slight difference from deep learning algorithms.

	MAD	MSE	RMSE	MAE	MAPE	MEAN	STD	Correlation
GEP	11.9448	300.4576	17.3337	0.0834	8.3350	1.0322	0.1308	0.9782
NN	7.2895	117.0553	10.8192	0.0419	4.1852	0.9988	0.0586	0.9915
ADAM & MSE	6.9639	112.9846	10.6294	0.0461	4.6051	1.0174	0.0628	0.9920

Table 4.10 Statistical Results of All Optizimer for MSE Loss.

 $\overline{1}$

The Comparison of All Optimizer's Results are shown in Figure 4.55 for MSE Loss.

Figure 4.55. Actual values and Predicted values of all Optimizer and MSE Loss.

4.10 Comparison of All Optimizer's Results in Statistical Evaluations

When evaluated in terms of MAD, as shown in Figure 4.56, it is seen that Deep Learning is generally more successful than soft computing. According to this graph, it can be said that SGD and MAE functions which are used as optimizer for Sequential model have the most unsuccessful results. The best results are obtained from the model using ADAMAX optimizer and MAE loss.

Figure 4.56. Mean Absolute Deviations of All Optimizers and Losses.

When evaluated in terms of MSE, as shown in Figure 4.57, it is seen that Deep Learning is generally more successful than soft computing. According to this graph, we can say that the SGD and MAE functions used as the Optimizer for Sequential model have the most unsuccessful results. However, it is seen that it is between the GEP model and the NN model when compared to soft computing techniques. In other words, when the weakest deep learning method, SGD optimizer and MAE loss functions are used, the model gives better results compared to GEP and is slightly weaker than NN.

Figure 4.57. Mean Square Errors of All Optimizers and Losses.

When evaluated in terms of Standard Deviation, as shown in (Figure 4.58), it is seen that Deep Learning technique is slightly weaker than the soft computing techniques, although GEP is more successful than NN, although approximate values are obtained. According to this graph, we can say that the SGD and MAE functions used as the Optimizer for Sequential model have the most unsuccessful results. When comparing the deep learning methods, it can be said that SGD optimizer and MAE loss functions have the weakest performance, ADAMAX optimizer and MAE Loss functions have the best performance.

Figure 4.58. Standart Deviations of All Optimizers and Losses.

When evaluated in terms of Correlation, as shown in Figure 4.59, Deep Learning is generally more successful than soft computing. According to this graph, it can be said that the SGD and MAE functions used as the Optimizer for Sequential model have the most unsuccessful results. However, it is seen that it is between the GEP model and the NN model when compared to soft computing techniques. In other words, when the weakest deep learning method, SGD optimizer and MAE loss functions are used, the model gives better results compared to GEP and is slightly weaker than NN. The best results are obtained from the model using ADAM Optimizer and MAE Loss.

Figure 4.59. Correlations(R) of All Optimizers and Losses.

When evaluated in terms of *RMSE*, as shown in Figure 4.60, the results of the Deep Learning technique and the NN method of soft computing techniques are very similar. Despite these close values, deep learning generally resulted in fewer errors. According to this graph, we have the most unsuccessful results of the SGD and MAE functions used as Loss in the Sequential model for the dataset we have here. However, when compared to soft computing techniques, it is seen that it is better than both of GEP model and the NN model. In other words, when the weakest deep learning method, SGD optimizer and MAE loss functions are used, the model gives better results compared to GEP and is slightly stronger than NN. The best results are obtained from the model using NADAM Optimizer and MAE Loss.

Figure 4.60. Root Mean Square Errors(RMSE) of All Optimizers and Losses.

When evaluated in terms of MAE, as shown in Figure 4.61, the results of the Deep Learning technique and the NN method from soft computing techniques are very similar. Despite these close values, deep learning generally resulted in fewer errors. According to this graph, the most unsuccessful results are the SGD and MAE functions used as Loss in the Sequential model for the dataset. However, it is seen that it is between the GEP model and the NN model when compared to soft computing techniques. In other words, when the weakest deep learning method, SGD optimizer and MAE loss functions are used, the model gives better results compared to GEP and is slightly better than NN. The best result is from the model using the ADAMAX Optimizer and MAE Loss.

Figure 4.61. Mean Absolute Errors(MAE) of All Optimizers and Losses.

When evaluated in terms of *MAPE*, as shown in Figure 4.62, the results of the Deep Learning technique and the NN method of soft computing techniques are very similar. In general, the comparison results of the values are the same as those of the MAE.

MAPE

Figure 4.62. Mean Absolute Percentage Errors(MAPE) of All Optimizers and Losses.

The actual values and the estimated values are evaluated in terms of MEAN obtained from the ratio shown in Figure 4.63. According to this graph, we have the most unsuccessful results of the SGD and MAE functions used as Loss in the Sequential model for the dataset we have here. In the case of the weakest deep learning method SGD optimizer and MAE loss functions according to the data available, the model has remained weak from both soft computing techniques and other deep learning methods. Of all the best results are obtained from the model using the ADAM Optimizer and MAE Loss.

Figure 4.63. Mean of All Optimizers and Losses for Average Rate of Results. The ratios of the each actual results and predicted results are shown in the Table 4.11.

Table 4. 11. The table of Mean(Actual/Predicted) of all Optimizers and Losses

CHAPTER 5 CONCLUSION

The aim of this study is to determine whether deep learning methods can be used to help the buckling load estimation of heat treated aluminum alloy columns and to analyze which method yields more successful results. Sequential is preferred as the model when applying deep learning method. Adam, Adamax, Adadelta, AdaGrad, Nadam, SGD, RMSProp optimizers are used. MAE and MSE Losses are used for each optimizer. Epoch number is kept constant as 4096 and results of all variations are considered and these results are evaluated in many respects. These evaluation methods are *RMSE, MSE, MAE, MSE, MAD, STD, MEAN, and Correlation*.

When evaluation methods *MAD, MAE, MAPE, and STD* are considered; the best results are obtained from the model using ADAMAX optimizer and MAE loss functions. Evaluation results are obtained as *MAD*: 5.8172, *MAE*: 0.0337, *MAPE*: 3.3701, STD: 0.0595 respectively.

When analyzed according to *MSE, RMSE, correlation* methods; the best results are obtained from ADAM optimizer and MAE Loss functions. Evaluation results are obtained as *MSE*: 100.0853, *RMSE*: 10.0042, *correlation*: 0.9928 respectively. When the *MEAN* results are analyzed, it is seen that the result of NN method (result: 0.9988) which is one of the soft computing techniques is closer to 1. The most successful MEAN result of the deep learning methods is ADADELTA optimizer and MSE Loss functions (result: 1.0030).

When the dataset is analyzed; it is seen that the places where the errors are most intense, are the lines where different results are obtained although the same inputs are used in the test results. In these rows, Deep learning Model makes a classification; consequently, the results of the same inputs are the same. Estimates are inaccurate because they differ in laboratory results.

In general, deep learning methods have been successful in prediction of the buckling load of aluminum alloy columns. This method is expected to be more successful if larger dataset is used to guide future studies. According to the results of the study, it is recommended to use MAE as a loss function and ADAM or ADAMAX as an optimizer in estimation processes. The SGD optimizer has the weakest results so that, it is not recommended.

REFERENCES

Abdulkader (2019), *Predicting The Maximal Oxygen Uptake Using Deep Learning,* **1**, Turkey

Anwer (2017), *Breast Cancer Dıagnosıs Wıth Deep Learnıng Methods,* **10**, Turkey

Arik, A., Gölcük, M., & Karslıgil, E. M. (2017). *Deep learning based skin cancer diagnosis. In Signal Processing and Communications Applications Conference,* Turkey

Büber, E., Sahingoz, O. K. (2017). *Image Processing and Adjustment of Optimal Parameters with the Machine-Care System* IDAP

J, Maljaar., L,Twilt., F, Soeten. (2009). *Flexural buckling of fire exposed aluminum columns,* Fire Safety Journal **44**(5), pp 711-717.

Cevik, A., Ekmekyapar, T., Atmaca, N., Guzelbey, İ. (2009). *Flexural buckling load prediction of aluminum alloy columns using soft computing techniques*, *Expert Systems with Applications*, **36**, pp 6332-6342.

Çakır (2019), *Zero-Day Attack Detectıon Wıth Deep Learnıng,* **5**, Turkey

Eriş (2018), *Evaluatıon Of Evıdence In Forensıc Informatıon Studıes Usıng Deep Learnıng Methods,* **4**, Turkey

Işık, G., Artuner, H. (2016). *Recognition of radio signals with deep learning Neural Networks. In Signal Processing and Communication Application Conference*

Kantepe (2019), *Deep Learnıng Methods Usıng Product Proposal System Desıgn,* **5,** Turkey

Kaynar, O., Görmez, Y., Işık, Y. E. (2016). *Spam Detection with Auto Coder Based Deep Learning Machines,*

Koyun, A., Afşin, E. (2017). *Two Dimensional Optical Character Recognition with Deep Learning*. Journal of Turkey Informatics Foundation of Computer Science and Engineering,

Kutlu (2019), *Obtaıned Wıth Unmanned Aırcraft Deep Of Imaged Images Analysıs Wıth Learnıng Methods,* **4**, Turkey

Munir (2019), *Handwrıtten Text Recognıtıon Usıng Deep Learnıng,* **7**, Turkey

Özgenel (2018), *Crack Detectıon Wıth Deep Learnıng: An Exemplary Study Of Data Desıgn In Archıtecture,* **5**, Turkey

Paker (2019), *Deep Learnıng For Bıologıcal Sequences,***1,** Turkey

Pala (2019), *A Comparatıve Study of Deep Learnıng Based Face Recognıtıon Algorıthms For Vıdeo Under Adverse Condıtıons*, **5**, Turkey

Rasheed (2019), *Object Detectıon From Images Usıng Deep Learnıng,* **2**, Turkey

Razavi, S., Yalcin, H. (2017). *Using convolutionalneural networks for plant classification. In Signal Processing and Communications Applications Conference (SIU)*, 2017 25th (pp. 1-4). IEEE.

Sorkun (2018), *Tıme Serıes Forecastıng On Solar Radıatıon Usıng Deep Learnıng,* **2**, Turkey Soydaş (2019), *Aırcraft Detectıon From Large Scale Remote Sensıng Images Wıth Deep Learnıng Technıques,* **23,** Turkey

Süberk (2019), "*Deep Learnıng Technıques For Buıldıng Densıty Estımatıon From Remotely Sensed Imagery",* **2**, Turkey

Şeker, A., Yüksek, A. G. (2017). "*Stacked Auto-encoder Method for Fabric Fault Detection*." Cumhuriyet University Faculty of Arts and Sciences Journal of Science,**342-354**.

Tanimu (2019), "*Gpr Raw-Data Analysıs To Detect Crack Vıa Wavelets And Deep Learnıng Methods",* **3**, Turkey

Tienin (2019), "*Cloud Coveragepredıctıonwıthdeep Learnıng",* **1**, Turkey

Uçar, A., Bingöl, M. S. (2018). "*Evaluating graphics processing cards using deep learning"*, Mühendislik Dergisi 9(1), **39-49**.

Üner (2019), Deepside: *"Predicting Drug Side Effects With Deep Learning",* **1,** Turkey

Yalçın, H. (2016). "*Human activity recognition using deep belief networks. In Signal Processing and Communication Application Conference (SIU)"*, Turkey

Yetiş (2019), "*Auto-Conversıon From 2d Drawıng to 3d Model Wıth Deep Learnıng",* **5**, Turkey

Wang et. al(2017) . "*Compressive buckling strength of extruded aluminum alloy I-section columns with fixed-pinned end conditions",* China

Yetkin, A. E., Hamamcı, A. (2016). *Data Augmentation for Head Pose Estimation From MRI* <https://keras.io/why-use-keras/#keras-prioritizes-developer-experience>

[Wald, Abraham](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQWJyYWhhbV9XYWxk) (June 1945). *"Sequential Tests of Statistical Hypotheses". The Annals of Mathematical Statistics. 16..*

[Berger, James](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvSmFtZXNfQmVyZ2VyXyhzdGF0aXN0aWNpYW4p) (2008). [Sequential Analysis.](https://www.wikizeroo.org/index.php?q=aHR0cDovL3d3dy5kaWN0aW9uYXJ5b2ZlY29ub21pY3MuY29tL2FydGljbGU_aWQ9cGRlMjAwOF9TMDAwMDk4) *The New Palgrave Dictionary of Economics, 2nd Ed. 438–439.*

[Kenneth J. Arrow,](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvS2VubmV0aF9BcnJvdw) [David Blackwell](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvRGF2aWRfQmxhY2t3ZWxs) and M.A. Girshick (1949). *"Bayes and minimax solutions of sequential decision problems". Econometrica. 17*

[Randell, Brian](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQnJpYW5fUmFuZGVsbA) *(1980), "The Colossus", A History of Computing in the Twentieth Century, 30*

Turnbull, Bruce (2000).*Group sequential methods with applications to clinical trials. Chapman & Hall.*

O'Brien, Peter C. Fleming, Thomas R. (1979). *"A Multiple Testing Procedure for Clinical Trials". Biometrics. 35*

Wang, Samuel K.; Tsiatis, Anastasios A. (1987). *"Approximately Optimal One-Parameter Boundaries for Group Sequential Trials". Biometrics. 43*

Demets, David L.; Lan, K. K. Gordon (1994). "Interim analysis: The alpha spending function approach". Statistics in Medicine. 13 (13–14): 1341–1352.

[Diederik Kingma,](https://arxiv.org/search/cs?searchtype=author&query=Kingma%2C+D) [Jimmy](https://arxiv.org/search/cs?searchtype=author&query=Ba%2C+J) Ba (2014), *On the Convergence of Adam and Beyond*

Håkan Gadd (2006). *Evaluation: definitions, methods and models, ITPS, Swedish Institute For Growth Policy Studies Studentplan 3, SE-831 40 Östersund, Sweden*

J. Andreas Graefe; *Scott Armstrong; Randall J. Jones, Jr.; Alfred G. Cuzán (2010).* ["Combining forecasts for predicting U.S. Presidential Election outcomes"](https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9kbC5kcm9wYm94LmNvbS91LzM2NjI0MDYvQXJ0aWNsZXMvR3JhZWZlX2V0X2FsX0NvbWJpbmluZy5wZGY)*Surface.* Turkey

