EARLY WARNING MODEL WITH MACHINE LEARNING FOR TURKISH INSURANCE SECTOR

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÜNAY BURAK KOÇER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

SEPTEMBER 2019

Approval of the thesis:

EARLY WARNING MODEL WITH MACHINE LEARNING FOR TURKISH INSURANCE SECTOR

submitted by **GÜNAY BURAK KOÇER** in partial fulfillment of the requirements for the degree of **Master of Science in Financial Mathematics Department, Middle East Technical University** by,

Prof. Dr. Ömür Uğur
Director, Graduate School of **Applied Mathematics** __________

Prof. Dr. A. Sevtap Kestel
Head of Department, **Financial Mathematics** __________

Prof. Dr. A. Sevtap Kestel
Supervisor, **Financial Mathematics, METU** __________

Examining Committee Members:

Prof. Dr. Fatih Tank
Actuarial Sciences, ANKARA UNIVERSITY __________

Prof. Dr. A. Sevtap Kestel
Financial Mathematics, METU __________

Assoc. Prof. Dr. Ceylan Yozgatlıgil
Statistics, METU __________

Date: __________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: GÜNAY BURAK KOÇER

Signature :

ABSTRACT

EARLY WARNING MODEL WITH MACHINE LEARNING FOR TURKISH INSURANCE SECTOR

Koçer, Günay Burak
M.S., Department of Financial Mathematics
Supervisor : Prof. Dr. A. Sevtap Kestel

September 2019, 103 pages

Early warning models are needed to ensure that relevant stakeholders would not cause worse outcomes by ignoring the risks. Specific to the insurance sector, this risk is about meeting the obligations and its sustainability. In this study, an early warning model is formed using the ratios obtained from the financial statements of insurance companies. The goal of the model is to identify the risk areas and to support the strengthening of the financial structure and taking timely measures in companies. Since the data consisted of annual periods, only non-life insurance companies are included in the analysis. The annual balance sheet and statements of income declared by the Insurance Association of Turkey (IAT) and the annual reports about insurance and private pension activities published by The Republic of Turkey Ministry of Treasury and Finance (TRMTF) are examined based on companies, and ratios to be used in the model are calculated. The data set is composed of 70 financial ratios obtained from the financial statements of all non-life insurance companies operating between 2011-2018. Classifications are determined as credit, liquidity, market, reinsurance, underwriting, technical provisions, reputational, operational, profitability, and capital risks. In the developed model, the values realized in 2018 are estimated with machine learning methods by using the data of 2011-2017. Random Forest, Neural Networks, Gradient Boosting Machine and Extreme Gradient Boosting are used as analysis methods, and Boruta is used as a feature selection method. Capital requirement ratio is chosen as the dependent variable. The other 69 ratios are the independent variables and this set is reduced to 22 independent variables by the Boruta method. Analyzes are conducted on two datasets with 69 and 22 independent variables and the results are compared. Furthermore, after predicting the 2018 values for above mentioned 38 non-life insurance companies and doing stage classifications with those values, whether the actual stages match with the predicted stages is evaluated. The best estimate accuracy belongs to Random Forest and Gradient Boosting Machine methods with 87%. The predictive power decreases with 22 independent variables, but the results are still close to each other. Then, performances of machine learning models are compared over capital adequacy classification. In this comparison, the best predictive method is Neural Networks with 95% accuracy.

Keywords: Early Warning Model, Financial Ratio Analysis, Capital Requirement Ratio, Random Forest, Neural Networks, Gradient Boosting Machine, eXtreme Gradient Boosting, Boruta, etc.

ÖZ

TÜRK SİGORTA SEKTÖRÜ İÇİN MAKİNE ÖĞRENİMİ İLE ERKEN UYARI MODELİ

Koçer, Günay Burak
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Prof. Dr. A. Sevtap Kestel

Eylül 2019, 103 sayfa

Erken Uyarı modellerine ilgili paydaşların riskleri görmezden gelerek daha kötü sonuçlara sebep olmamaları adına gerek duyulmaktadır. Sigorta sektörü özelinde bu risk, yükümlülükleri karşılama ve bunun sürdürülebilirliği ile ilgilidir. Bu çalışmada, sigorta şirketlerinin finansal tablolarından elde edilen oranlar ile bir erken uyarı modeli oluşturulmuştur. Bu modelin amacı şirketlerdeki risk alanlarını tespit etmek ve mali bünyenin güçlendirilmesini ve zamanında önlem alınmasını desteklemektir. Verilerin yıllık dönemlerden oluşması nedeniyle sadece hayat dışı sigorta şirketleri analize dahil edilmektedir. Türkiye Sigorta Birliği tarafından ilan edilen yıllık bilanço ve gelir tabloları ile T.C. Hazine ve Maliye Bakanlığı'nın yayınladığı sigortacılık ve bireysel emeklilik faaliyetleri hakkındaki yıllık raporlar, şirketler bazında incelenmiş ve modelde kullanılacak oranlar hesaplanmıştır. Veri seti 2011-2018 yılları arasında faaliyet gösteren tüm hayat dışı sigorta şirketlerinin mali tablolarından elde edilen 70 finansal orandan oluşmaktadır. Sınıflandırmalar; alacak, likidite, piyasa, reasürans, üretim, teknik karşılıklar, itibar, operasyon, karlılık ve sermaye riskleri olarak belirlenmiştir. Oluşturulan modelde, 2011-2017 yılları verisi kullanılarak 2018 yılında gerçekleşen değerler makine öğrenimi yöntemleri ile tahmin edilmektedir. Analiz yöntemleri olarak Rastgele Orman, Sinir Ağları, Gradyan Güçlendirme Makinesi ve Aşırı Gradyan Güçlendirme ve özellik seçimi yöntemi olarak Boruta kullanılmıştır. Bağımlı değişken olarak sermaye yeterlilik oranı seçilmiştir. Diğer 69 oran bağımsız değişkenlerdir ve bu set Boruta yöntemiyle 22 bağımsız değişkene düşürülmüştür. 69 ve 22 bağımsız değişkenli iki veri setine de analiz yapılarak sonuçlar karşılaştırılmıştır. Ayrıca, bahsi geçen 38 hayat dışı sigorta şirketine ait 2018 yılı verileri farklı makine öğrenimi yöntemleriyle tahmin edildikten ve ilgili aşama sınıfına yerleştirildikten sonra, 2018 gerçekleşmeleriyle uyumluluğu değerlendirilmiştir. En iyi tahmin doğruluğu %87 ile Rastgele Orman ve Gradyan Güçlendirme Makinesi yöntemlerine aittir. 22 bağımsız değişkeni ile tahmin gücü azalmaktadır ancak yine de sonuçlar birbirine yakındır. Ardından sermaye yeterliliği sınıflandırması üzerinden makine öğrenmesi modellerinin performansları karşılaştırılmıştır. Bu karşılaştırmada en iyi tahmin gücüne sahip yöntem %95 doğruluk ile Sinir Ağlarıdır.

Anahtar Kelimeler: Erken Uyarı Modeli, Finansal Oran Analizi, Sermaye Yeterlilik Oranı, Rastgele Orman, Sinir Ağları, Gradyan Güçlendirme Makineleri, Aşırı Gradyan Güçlendirme, Boruta

To My Family

ACKNOWLEDGMENTS

I would like to express my very great appreciation to my thesis supervisor Prof. Dr. A. Sevtap Kestel for her patient guidance, enthusiastic encouragement and valuable advices during the development and preparation of this thesis. Her willingness to give her time and to share her experiences has brightened my path.

I would like to thank the examining committee members, Prof. Dr. A. Sevtap Kestel, Prof.Dr. Fatih Tank and Assoc. Prof. Dr. Ceylan Yozgatlıgil for their suggestions, corrections and comments.

I would like to thank my dear professors Prof. Dr. Ayşen Apaydın, Prof. Dr. Fatih Tank and Assoc. Prof. Furkan Başer at Actuarial Science Department, Ankara University for their endless support and understanding.

I would like to sincerely thank Yiğit Yeldan for his support and friendship. Also, I would like to thank my friends Kemal Öztürk and Oğuz Koç for their encouragement and contribution.

Above all, I owe my deepest gratitude to my beloved father and mother, Fahri Koçer and Hilal Koçer for their endless support and love. I always feel lucky for being their son.

I would like to thank my dear sister Burcu Koçer Oruç, who makes me feel like she is beside me even if she is miles away. I also want to express my appreciation to Sercan Oruç for being a brother to me.

A very special gratitude goes out to my love Dilan Teber for her limitless love, rescue, and patience. She and I will accomplish and experience many more milestones together.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

TRMTF	Republic of Turkey Ministry of Treasury and Finance
GDI	General Directorate of Insurance
SSI	Social Security Institution
IAT	Insurance Association of Turkey
IRIS	Insurance Regulatory Information System
NAIC	National Association of Insurance Commissioners
SCR	Solvency Capital Requirements
MCR	Minimum Capital Requirements
FX	Foreign Exchange
IBNR	Incurred But Not Reported
LA/TA	Liquid Assets / Total Assets
NPR/TA	Net Premium Receivables / Total Assets
PCR	Premium Collection Ratio
GLR	Gross Loss Ratio
P/PC	Profit / Paid Capital
PP/C	Premium Production / Coverage
PRO/E	Payables on Reinsurance Operation / Equity
TR/NP	Total Reserve / Net Premium
TR/LA	Total Reserve / Liquid Asset
TP/NWP	Technical Profit / Net Written Premium
TI/TA	Total Income / Total Assets
RS/GP	Reinsurance Share / Gross Premium
NLR	Net Loss Ratio
NLR/GLR	Net Loss Ratio / Gross Loss Ratio
GPWE	Gross Premiums Written to Equity
NPWE	Net Premiums Written to Equity
CNWP	Change in Net Written Premiums
CGWP	Change in Gross Written Premiums

DR/TA	Doubtful Receivables / Total Assets
SRP/E	Share of Reinsurance of Provisions / Equity
CashR	Cash Ratio
ACTP/TP	Assets Covering Technical Provisions / Technical Provisions
IP/TA	Investment Properties / Total Assets
ROI	Return on Investment
ROA	Return of Assets
TP/GWP	Technical Profit / Gross Written Premium
CRR	Capital Requirements Ratio
E/TP	Equity / Total Payables
CE	Changes in Equity
AFC/LFC	Assets in Foreign Currency / Liabilities in Foreign Currency
NER	Net Expense Ratio
MPS	Motor Portfolio Share
NPL/GPL	Net Paid Losses / Gross Paid Losses
CiRR	Changes in Retention Ratio
RRR	Reinsurance Rate of Return
RR	Reinsurance Risk Ratio
RPOL/APOL	Change in Required Provision for Outstanding Losses / Available Provision for Outstanding Losses
GPOL/E	Gross Provision for Outstanding Losses / Equity
NPOL/E	Net Provision for Outstanding Losses / Equity
CPOL	Change in Provision for Outstanding Losses
PPPP	Premium Production Per Personnel
GCPR	Gross Compensation Payment Ratio
NCPR	Net Compensation Payment Ratio
GPL/GWP	Gross Paid Losses / Gross Written Premiums
CGPOL/GWP	Change in Gross Provision for Outstanding Losses / Gross Written Premiums
SRPL/GPL	Share of Reinsurance of Paid Losses / Gross Paid Losses
SR/GPOL	Share of Reinsurance / Gross Provision for Outstanding Losses
SRPTOL/PTOL	Share of Reinsurance Provision for Transferred Outstanding Losses / Provision for Transferred Outstanding Losses
NPL/NWP	Net Paid Losses / Net Written Premiums

NOL/NWP	Net Outstanding Losses / Net Written Premiums
OE/GWP	Operating Expenses / Gross Written Premiums
NCR	Net Commission Ratio
OE/NEP	Operating Expenses / Net Earned Premiums
COMR	Combined Ratio
MShr	Market Share
TI/TE	Non-Life Technical Income / Non-Life Technical Expense
FLR	Financial Leverage Ratios
CURR	Current Ratio
SFR	Self-Financing Ratio
TA/E	Tangible Assets / Equity
TA/TA	Tangible Assets / Total Assets
NCA/LE	Non-Current Assets / Long Term Liabilities and Equity
RCR	Receivables Cycle Ratio
FP	Financial Profitability
EP	Economic Profitability
ROE	Return on Equity
CR	Collection Ratio
CA/TA	Current Assets / Total Assets
CLA	Change in Liquid Assets
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
RF	Random Forest
GBM	Gradient Boosting Machine
NN	Neural Networks
XGBoost	Extreme Boosting Machine

CHAPTER 1

INTRODUCTION

Recently unsteady economic conditions have led to controversy about the robustness of the economic system and the measurement of it. Some actors are looking for ports to act as insurance for themselves, while others may want to reduce costs by avoiding being insured. Even though these behaviors of all natural and legal persons vary according to their attitudes towards risk, this may lead to further deterioration in the economy. While taking all these risks into consideration, in all financial environments, all the stakeholders may need to measure the risks they took due to their work in the economy. The situation of banks and insurance companies, two of the main and most important actors of the economy, is wondered by supervisory and regulatory authorities and all related stakeholders. For the sake of the stability of the economy, it is desirable to predict and take prevention of the failure of companies.

The global economic crisis has resulted in analysts and policymakers around the globe to make significant efforts to understand and predict systemic financial crises. Therefore, they are concentrated on creating early warning mechanisms to anticipate potential crises in the empirical literature on financial crisis prediction [1]. Based on past crisis experiences, it is very important for all components of the economy to detect overheating of economic and financial activities and to take preventive measures against possible crises [37].

The insurance industry is also an important part of the economy and a strong insurance sector is a guarantee for the strength of the financial system. Early warning systems are established for the purpose of achieving the goals of the companies or institutions, minimizing the uncertainties that may arise in the realization of their missions. Early warning systems are used to produce pre-warnings about possible problems that may arise or develop in the future due to the prior notification of the current problems and the current risk profile of the sectors. Early warning model can assist: [57]

- Systematic evaluation of institutions during both on-site and off-site inspections;
- Identification of the problematic sections or possible risks;
- Prioritization of institution reviews for pre-review planning with optimal allocation of audit resources; and

- Initiation of guaranteed and timely transaction by regulator and supervisory authority.

Statistical and mathematical models are intended to be a real "Early Warning Models". Essentially, they are data-driven and use sophisticated quantitative methods that try to convert multiple economic power and efficiency indices into risk assessments. These estimates show that models are designed to separate future companies with a high risk of failure from the companies with a low risk of failure. Different estimation methods for statistical models may be used. Most model designs are focused on the qualitative reaction method, which examines the correlation between dependent and certain independent variables. In such designs, the dependent variables may be failure, survival or ordered results [57]. Ratio analysis enables us to obtain information about the economic and financial structure, profitability and operational status of the companies by determining the dependent and independent variables by meaningful relationships between the items in the financial statements [26].

The motivation and purpose of the study can be listed as follows;

- decide which Machine Learning method can be used for better development of the early warning system established to prevent insurers from insolvent,
- determine which ratios better detect early warning,
- determine which early warning model works better in determining the financial situation of companies,
- emphasize the importance of ratio selection in early warning.

Moreover, the ratios used in both national and international standards are used in this study. The most important feature of this study is being the first study that uses such detailed ratios and machine learning techniques.

1.1 The Solvency Directives

Solvency is an application that provides information about possible changes in operating accounts and whether companies will meet their obligations to the insured. With this application, it is measured whether the companies are sufficient to meet their debts at a certain time [32]. Solvency for an insurance firm, where the assets exceed liabilities, is the minimum economic safety standard [56]. Different methods are used for life and non-life companies in the calculation of capital adequacy in European Union insurance companies. The first regulation for life insurance companies was made with the Directive 79/267/EEC [20] issued in 1979 and the Directive 73/239/EEC [19] for non-life in 1973. The steps and regulations taken to regulate capital adequacy in the insurance market in the European Union are called Solvency I. In the following periods, Solvency II [24] system was introduced to help make comprehensive arrangements for insurance companies by considering the weaknesses of Solvency I, including non-risk-sensitive, market, credit and operational risks. The new structure is seen as a harmonious, robust and strong framework for insurance companies in the European Union in 2009. The main reasons for the need for Solvency II are the fact that the current system is based on using historical data and does not take into consideration the risks that may occur in the future, is inadequate in preventing company failures, inability to ensure effective supervision of the groups within the insurance companies and not being in compliance with the regulations in the international and sectors. Solvency II consists of 3 Pillars that includes Pillar 1 that is quantitative requirements, Pillar 2 that is requirements for risk management and Pillar 3 that is transparency. Capital Requirement, calculated by quantitative requirements in Pillar 1, provides a structure that can be adjusted to the risk of each insurer. There is also a two-tier structure, including Solvency Capital Requirements (SCR) and Minimum Capital Requirements (MCR) under Solvency II to increase control power [23]. As far as the SCR is concerned, an insurer can calculate its requirements using a predefined standard formula, a self-developed complete internal model, partial internal model as a mixture of the first two options. Pillar Two includes a number of qualitative components to guarantee adequate handling of quantitative requirements. Specifically, Pillar Two includes supervisory assessments of governance and risk management systems. Solvency II's third pillar addresses European insurance companies' legislative reporting requirements. Their goal is to provide both supervisory agencies and the public with market transparency. As a consequence, report respondents will be able to acquire data on the solvency situation and economic circumstances of a company [28].

Since January 2016, the insurers and reinsurers of the European Union have been governed by the Solvency II regulatory regime. With the 2 major revisions made in the regulation in 2018, the possibility for any improvement in the system is provided. In the first revision made in late 2018, the main focus is on simplification and resolution of technical problems through capital calculations. With the second revision planned for the end of 2020, as Solvency II introduces unnecessary limitations for long term guarantees and long term investors, major changes addressing wider issues covering concerns are planned to be made. With these revisions, the European Commission considers that Solvency II will be improved and the barriers to long-term investments will be removed. It is considered that the risk margin design should be taken into consideration as a priority in the changes envisaged to be made in 2020. Furthermore, this arrangement emphasizes the need to review the interest rate risk [35].

1.2 Literature Survey

Financial ratios were first used in the literature by Beaver in 1966 to estimate the failures of companies. In this study, 79 failed and 79 non-failed companies were examined. Data were obtained from Moody's Industrial Manual and cover the period between the years 1954 and 1964. In this study, he used 30 financial ratios which were collected under six groups. These groups are cash flow ratios, net income ratios, debt to total asset ratios, liquid asset to total asset ratios, liquid asset to current debt ratios and turnover ratios. He divided the data into three sections while analyzing. He described the comparison of mean values, which was the first section, as profile analysis. Beaver, who indicates that profile analysis shows the difference between failed and non-failed companies, stated that the lack of this analysis not being able to respond to the magnitude of the difference. The profile analysis showed that the average asset size of non-failed companies is greater than that of failed companies. In the second part, the dichotomous test, a predictive test unlike the profile analysis, is mentioned. This test estimates the failure status of companies. As a result of the study, he classified the bankrupted companies with 78% accuracy, 5 years before their bankruptcy [5].

Another study to predict the failure of companies belongs to Altman. The aim of this study which was conducted in 1968 is to measure the quality of ratio analysis as an analytical technique. In this study, multiple discriminant analysis in addition to financial and economic ratios were used to investigate business collapse probability estimations using 22 ratios. The reason for using these two analyzes at the same time is the compatibility of multiple discriminant analyses with ratio analysis in estimating the probability of companies to bankrupt. In this study, which examined 33 bankrupted and 33 non-bankrupted firms, the data covers the years 1946-1965. As a result of this study, the discriminant-ratio model predicted group 1 (bankrupt companies) with 94% accuracy and proved that the process of assigning companies to groups is done 95% correctly. The model that emerged with this study is Z-score which constitutes a general index of study [2]. Previous studies by Beaver and Altman were compared, and the predictive power of a large proportion group was measured by using both univariate and multivariate statistics by Deakin (1972). As a result of the analysis, it was seen that discriminant analysis can predict business failures up to three years before the failure with a very high accuracy [21]. Libby devised a study in 1975 to find out whether accounting ratios provide loan officers with information that can be used to predict business failures. Deakin's (1972) randomly selected data group consisting of 30 failed and 30 non-failed firms were evaluated by the participants. Principal component analysis and varimax rotation were applied to the analysis to reduce the number of ratios used in the analysis phase and to eliminate the excess. As a result of the analysis, it was seen that the small experimental set obtained from accounting ratios enables bankers to predict the failures of companies with high accuracy and reliability [44]. Edmister (1972) investigated whether these rates can be used to predict the failure of companies with small asset size. As a result of the study, it was found that most but not all of examined methods and analyzes were predictors of failures [22].

The purpose of Meyer and Pifer's study conducted in 1970 is to predict future failures rather than to explain previous failures. Data were obtained from solvent and insolvent banks operating between 1948-1965. In the research, the stepwise regression is used. According to the results of the research, first of all, it is concluded that financial measures can evaluate the relative strength of companies. It was observed that 80% of the observations were correctly classified with $R^2 = 0.70$ [46].

Blum (1974) developed the Failing Company Model to predict failure in his study. In the analysis, 115 failed and 115 non-failed firm data were examined. As a result of the study, the model distinguishes failed companies from non-failed companies with 94% accuracy [6].

In the study conducted by Brockett, Cooper, Pitaktong, and Golden (1994), the artificial intelligence neural network model is used as an early warning model to estimate the insolvency status of insurance companies. In order to measure the susceptibility of an insurance company to bankruptcy, the data from two years ago of the companies that had failed in 1991 and 1992 and the existing data of the existing firms were taken into consideration. Early Warning Model is established with regulatory annual statements. The model is first established with 24 variables and then reduced to 8 variables by stepwise Logistic Regression. The results of artificial neural networks are compared with discriminant analysis, the National Association of Insurance Commissioners' Insurance Regulatory Information System ratings and A. M. Best ratings. The findings of the neural network show elevated predictability and generalizability, indicating that this technique is useful in anticipating potential insurance insolvency. The total percentage correctly classified of the Neural Network is 89.3%.[10].

Isseveroglu and Gucenme (2009) compared Logit Analysis with Multiple Discriminator Analysis and Multiple Regression Analysis to measure the power of models to predict the failure of companies. In the research, an early warning model was developed to analyze the situation of 45 non-life insurance companies operating between 1992 and 2006, whose financial situation is deteriorating. In the data set containing 17 financial ratios of 45 companies. According to the results of the generated model, it was seen that the logit model has better predictive power than the other two models. In addition, it is concluded in the study that the Logit Model is strong enough to define the current situation of companies [36].

In a study conducted in Australia on insurance companies whose financial situation is deteriorating, an early warning model has been developed to analyze to the situation of companies by Australian Prudential Regulation Authority (2006). The data is between 1999 and 2001. The model is tried to be estimated by logit regression. As a result, it is seen that insurers living financially distressed are small in size, have low profitability and have low cession rates. Moreover, these companies have reinsurance assets and properties rather than liquid assets [59].

The aim of the study by Tornoa and Tiub (2014) is to examine the factors affecting the survival and failures of non-life insurance companies in the Philippines and to classify them as strong, moderate, weak and insolvent. The data of the study covering the period between 2002 and 2011 are obtained from the Insurance Commission, which includes annual reports. 10 years panel data of 79 non-life insurance companies are used. While ANOVA is used for the strong, moderate, weak and insolvent classifications of the companies, the Stochastic Frontier Regression technique is used to estimate the probability of survival and failure. As a result of the analyzes, variables such as firm size, degree of diversification, leverage, paid-up capital were shown among the important factors affecting the survival of non-life insurance companies in the Philippines, and the study is found to be effective in the decision-making processes of stakeholders in non-life insurance companies [62].

The early warning model in the study conducted by Genç (2002) consists of 3 parts. The first part of the early warning model for Turkish non-life insurance companies is the leading indicator research. These are indicators that need to be constantly observed. In the second part, the multiple regression model, capital analysis and management evaluation study were conducted. In the multiple regression model, 14 financial ratios were used. Ratios of non-life insurance companies operating in Turkey in 2001 were calculated utilizing the financial statements. The analysis uses 3 companies that stopped their operations and 19 companies that continued their activities in 2001. The data set consists of 22 rows and 15 columns. Multiple regression model was found to be significant with the Stepwise method and R^2 of the best model was 62.4%. Then, it is investigated whether the capital of the company is sufficient. In the management evaluation, it was examined in terms of the backgrounds of the managers, their reputations, successes and failures, the stability in management, the frequency of change of managers and the institutionalization of the company. In the third part, monthly claims payments and outstanding claims development should be examined, and monthly premium production and collection should be followed. In the Model, multiple regression and qualitative analysis were used together to score. As a result of the model, 10 out of 22 companies are said to have insufficient capital and 6 out of financially insufficient[26]. In the study conducted by Ocak (2015), in addition to the 14 financial ratios in the study of Genç (2002), there are data on premium production and the year in which the companies operate. The financial information of 41 companies between 1998-2012 was analyzed. In addition to the 14 financial ratios, a model was created with the variable of premium production and year of operation of the company. The data set is 41 rows and 16 columns. The existence of the firm was chosen as the dependent variable and the classification of the bankrupted firms was made according to the high loss premium ratio, account deficit amount, negative ratio value and high debt ratios. According to these variables, the worst year of the firm is selected and the rates of that year are included in the analysis. Linear, Logistic and Bayesian Regression and Multivariate Discriminant Analysis are used. The bayesian Regression model is selected as the best model and a warning index is created. R^2 of this method is 90.3%. The warning index is created to obtain an idea about the companies by comparing them with the average value for each year [48].

In the study conducted by Oktar and Yüksel (2015), it was aimed to estimate the bank crises with early warning signals. In this study, the 1994 and 2000 crisis occurred in Turkey was estimated with the data obtained from 1988 to 2014. The data used are quarterly data and there are 108 observations. In the analysis conducted with 19 independent 1 dependent variables, the dependent variable has a value of 1 for crisis years and 0 for other years. According to the results of MARS model, using derivative products for speculative purposes rather than risk management purpose, inflation rate and net profit / total assets ratio are the most important determinants of Turkish banking crisis [50].

1.3 The Aim of the Study and Scope

In this study, it is aimed to estimate the possible failure events on behalf of insurance companies in the light of historical data. Thanks to the early warning model, the financial situation of both companies and the entire sector can improve with timely and accurate interventions. The purpose is to predict whether the companies are insolvent by means of the machine learning methods and to make a situation analysis about the companies. Different machine learning methods have been tried to create the best model. The ratios used in national and international standards are included in the analyzes. This thesis is the first study using a detailed ratio set having 70 financial ratios and using machine learning methods.

Because of the fact that conditions of insurance companies or sector change from country to country, the data of insurance companies only in Turkey are included in our analyses. There are 2 reinsurance companies, 4 life, 38 non-life, and 18 pension and life insurance companies that operating in Turkey in 2019. Only non-life insurance companies are used in our analysis. The reason for this is that data is obtained from annual balance sheets and income statements and non-life insurance companies generally produce annual premiums and provide coverages. Due to the differences in the balance sheet and income statement in consequence of that the changes in the regulations, it is deemed more appropriate to collect the data of the last 10-years period. In this part of the thesis, brief information about the Solvency directives is given and the recent and planned improvements are mentioned. Then, a literature search is made about the studies that predict failures with early warning models. Afterward, information about the aim of the study and the scope is given.

Financial indicators are mentioned in Chapter 2 that includes all ratios to be used in the analysis. The definitions of financial ratios, how they are calculated, what extent they should be, and how the Turkish non-life insurance sector has been followed over the years have been shared. In addition, the risk classification of all ratios has been made. There are 4 main headings which are financial, non-life, operational and profit and solvency risks. There are a total of 10 risk categories, including credit, liquidity, market, reinsurance, underwriting, technical provision, reputational, operational, profitability, and capital risk. When making this ratio classification, any ratio is considered below the risk group to which it belongs most. Although a ratio is associated with multiple risk groups, it is followed under the main heading.

In Chapter 3, some machine learning algorithms are discussed. These are Random Forest, Neuron Networks, Gradient Boosting Machine, and Extreme Gradient Boosting methods. In addition, information is given about Boruta that is a feature selection method. in the third chapter, shows how these machine learning algorithms work and what their mathematical background.

Data descriptions and analyses are explained in Chapter 4. In this chapter, how data is acquired, and the selection of dependent and independent variables are mentioned. Capital requirement ratio that is one of the financial ratios in Chapter 2 is selected as the dependent variable and the other 69 ratios are selected as the independent variables. The grid method is used to create the best model and the analyzes are applied to both two data sets and the results are printed. These two data sets are full data set with 69 ratios, and data set with 22 ratios selected by the Boruta feature selection method. R program is utilized in our analyzes.

Finally, Chapter 5 addressed conclusions and comments. After briefly mentioning the whole thesis, how the analyses are performed and their successes are mentioned.

CHAPTER 2

FINANCIAL INDICATORS USED IN TURKISH INSURANCE SECTOR ANALYSES

The financial statements of companies or sectors and their analysis are of great importance for all actors in the economy. Lenders, creditors, investors, employees and employers, and regulatory agencies; each everyone has a unique desire to analyze the companies they are interested in. Lenders may be interested to determine if the money they lent will be repaid. Creditors may be keen to determine if the organization can be lent. The possible investment in one organization may be compared with that of another by investors. Employees may wish to compare their employer's present efficiency or economic status with previous phases. Regulatory agencies often need to evaluate economic health and efficiency of organizations or industries. A financial assessment is always supported by a number of issues and the issues which require responses rely on who the person is and why the assessment is conducted [11].

In economic analyzes, quantities or ratios can be contrasted with sector standards, the same measurement in a preceding era, the same measurement in the organization of a competitor, or with earlier defined quantities. It is necessary to decide which method should be chosen in order to make the best use of accounting information. Even if many ideas can be obtained with ratio analysis, it can be misleading to look at each of the financial ratios alone. A ratio is not always an indication of health, status, or performance of the organization even though it is the best [11]. For these reasons, all ratios used in the analysis should be examined as an entire.

In this chapter, the first 12 ratios and 21st and 28th ratios are the ratios used by Genç(2002) [26] in his analyses. The eleven ratios are those calculated by the IAT [33] in the balance sheet detail. To attain international standards, 4 other ratios used by NAIC [61] are taken into account.

Sector average ratios are calculated on the basis of total balance sheet values obtained by adding the relevant values of all related non-life insurance companies in Turkey.
The "(-)" sign in the equations indicates that the account is a negative balance sheet item. These accounts are recorded as negative.

The financial statements used by insurance companies are categorized below:

- Statement of financial position (Balance Sheet),
- Income statement,
- Statement of changes in shareholders' equity,
- Cash flow statement,
- Footnotes and other explanatory notes summarizing significant accounting policies,
- Profit distribution table.

The ratios are defined and their development in the Turkish insurance sector are illustrated.

2.1 Liquid Assets / Total Assets

This ratio shows the weight of liquid assets in total assets. Liquid assets are calculated by subtracting financial investments whose risk is assumed by life insurance policyholders from the aggregation of cash and cash equivalents and financial assets and financial investments whose risk is assumed by policyholders. Since only non-life companies and their activities are taken into consideration, financial investments whose risk is assumed by life insurance policyholders account has been excluded from the calculation. Total assets consist of current and non-current assets.

$$\text{LA/TA} = \frac{\left(\begin{array}{c}\text{Cash and cash equivalents}\\ \text{+ Financial assets and financial investments whose risk is assumed by policy}\\ \text{holders}\\ \text{- Financial investments whose risk is assumed by life insurance policy holders}\end{array}\right)}{(\text{ Current assets + Non-current assets })} \tag{2.1}$$

Figure 2.1 shows the development of the ratios over the years. It is expected to be higher than the sector average.

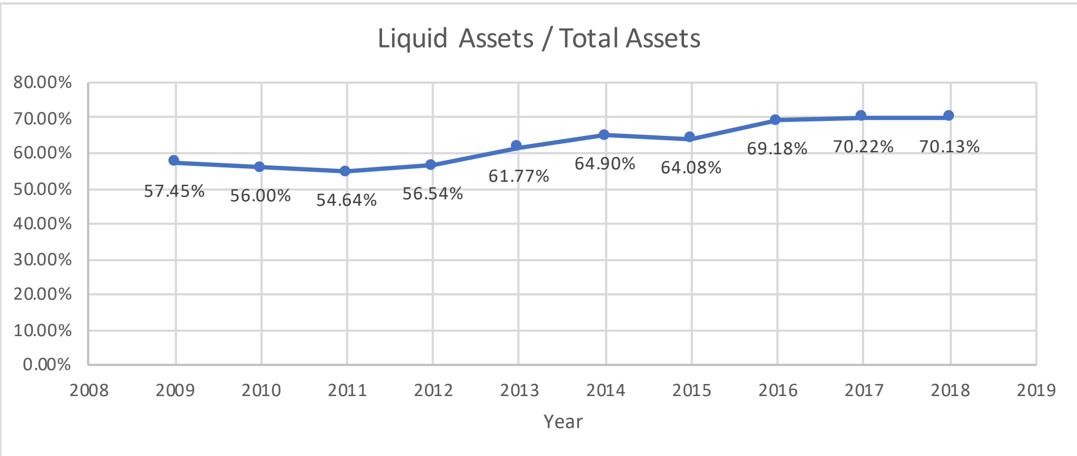


Figure 2.1: Liquid Assets / Total Assets, Non-Life Insurance Sector, 2009-2018

2.2 Net Premium Receivables / Total Assets

Receivables from operating activities contain receivables from insurance and reinsurance activities, Doubtful receivables from operating activities, Credits (loans) to policyholders and their provisions. This ratio shows the level of receivables from the main activities of the company. The size of premium receivables is tolerable according to the size of the total assets, however, it is more desirable to collect the receivables. It may be misleading to examine only the receivables of companies because large-scale companies are more likely to have more receivables. Therefore, it will allow us to see the actual position of the company by observing its share in total assets by making comparisons with it.

$$\text{NPR/TA} = \frac{(\text{Receivables from operating activities})}{(\text{Current assets} + \text{Non-current assets})} \tag{2.2}$$

This ratio is expected to be lower than the sector average. Based on the data set, this ratio has a pattern shown in Figure 2.2.

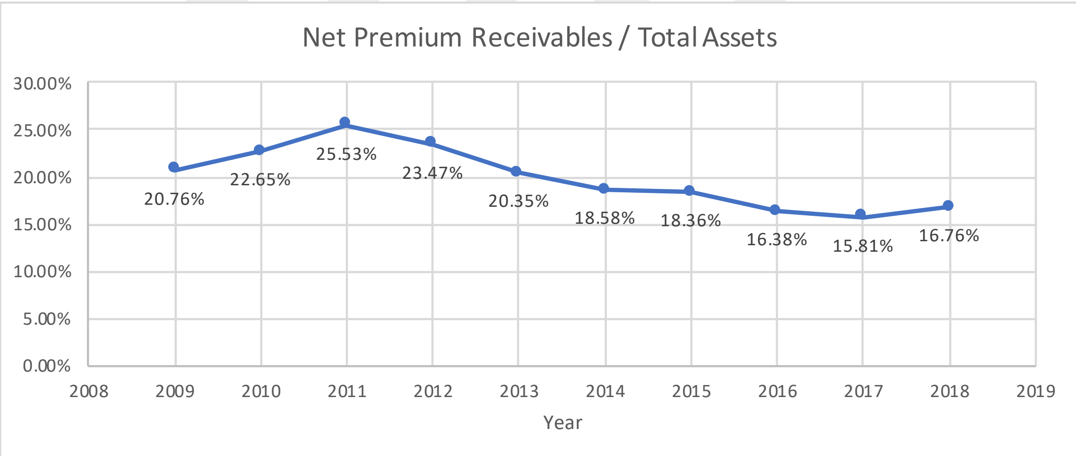


Figure 2.2: Net Premium Receivables / Total Assets, Non-Life Insurance Sector, 2009-2018

2.3 Premium Collection Ratio

Although collecting premium receivables is easier than ever with developing technology and payment systems, this risk remains essential for insurance companies. For instance, credit card receivables with a bank guarantee and shorter term than three months account is tracked in cash and cash equivalents. This ratio is used in the past year and current year receivables from the insured and intermediary.

$$\text{PCR} = \frac{\left(\begin{array}{c}\text{Receivables from Insureds (Previous Year)}\\ +\text{ Receivables from Intermediaries (Previous Year)}\\ +\text{ Written premiums (Net) (Current Year)}\\ +\text{ Acquisition Commission (-) (Current Year)}\\ -\text{ Receivables from Insureds (Current Year)}\\ -\text{ Receivables from Intermediaries (Current Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Receivables from Insureds (Previous Year)}\\ +\text{ Receivables from Intermediaries (Previous Year)}\\ +\text{ Written premiums (Net) (Current Year)}\\ +\text{ Acquisition Commission (-) (Current Year)}\end{array}\right)} \tag{2.3}$$

PCR is expected to be higher than the sector average. The higher this ratio is, the better we can say about the company can collect the receivables from the insured and intermediaries. Based on the data set, the pattern shown in Figure 2.3 shows this ratio.

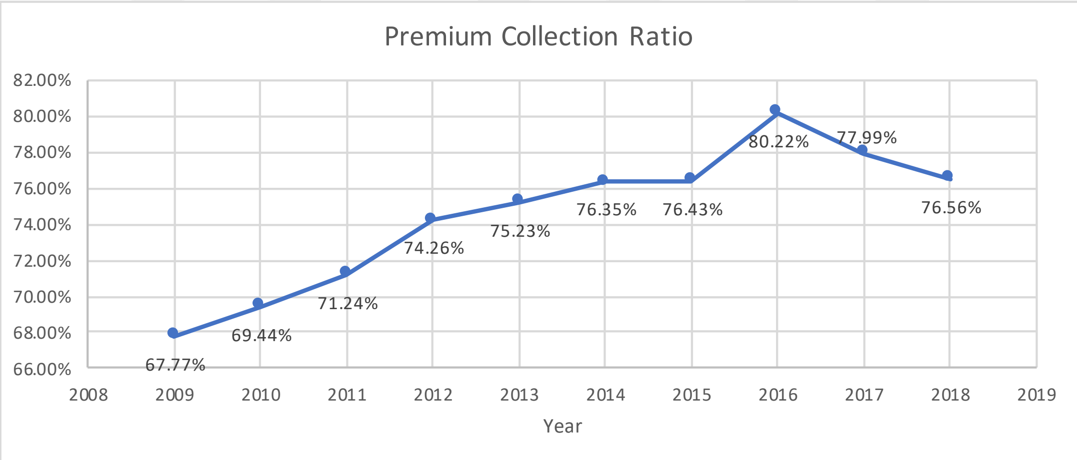


Figure 2.3: Premium Collection Ratio, Non-Life Insurance Sector, 2009-2018

2.4 Gross Loss Ratio

Gross paid losses, gross provision for outstanding losses and gross provision for transferred outstanding losses equal to gross claims(losses) incurred. The sum of gross written premiums, premiums transferred to the SSI, gross provision for unearned premiums, gross provision for transferred unearned premiums, SSI share of provision for unearned premiums, SSI share of provision for transferred unearned premiums, gross provision for ongoing risks and gross provision for transferred ongoing risks is equal to gross earned premiums. Although our calculation is gross, the reason for issuing SSI share is transfer of all premiums written on behalf of SSI. Loss ratio, also called Loss / Premium ratio, indicates how much damage has been paid against 100 units of earned premium. With the gross loss ratio, we can see success and failure in both insurance and reinsurance operations of an insurance company.

$$\mathrm{GLR} = \frac{\begin{pmatrix}\text{- Paid losses (Gross) (-)} \\ \text{- Provision for outstanding losses (Gross) (-)} \\ \text{- Provision for transferred outstanding losses (Gross)}\end{pmatrix}}{\begin{pmatrix}\text{Written premiums (Gross)} \\ \text{+ Premiums transferred to the SSI (-)} \\ \text{+ Provision for unearned premiums (Gross) (-)} \\ \text{+ Provision for transferred unearned premiums (Gross)} \\ \text{+ Provision for unearned premiums (SSI share)} \\ \text{+ Provision for transferred unearned premiums (SSI share)} \\ \text{+ Provision for ongoing risks (Gross)} \\ \text{+ Provision for transferred ongoing risks (Gross)}\end{pmatrix}} \tag{2.4}$$

This ratio is expected to be lower than the sector average. GLR values are shown in Figure 2.4. If this ratio greater than the sector average, the company needs to review pricing, portfolio selection policies, and losses.

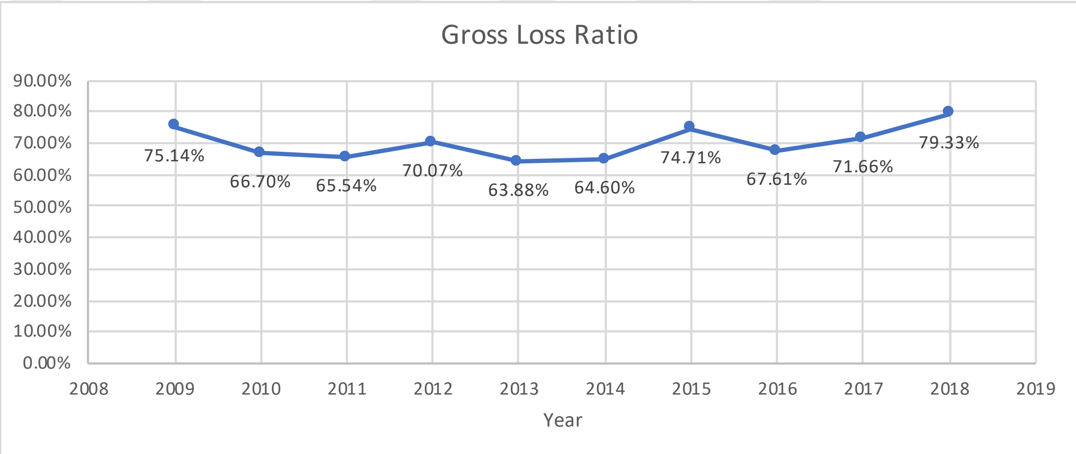


Figure 2.4: Gross Loss Ratio, Non-Life Insurance Sector, 2009-2018

2.5 Profit / Paid Capital

Paid capital or paid-in capital represents the capital investments made by shareholders at the balance sheet date. [58]

It consists of the sum of nominal capital, unpaid capital, positive and negative differences of capital adjustment and capital with pending registration.

$$\mathrm{P/PC} = \frac{(\text{Net profit or loss of the period})}{(\text{Paid capital})} \tag{2.5}$$

In 2010, 2011, 2012 and 2015, the Turkish non-life insurance sector suffered losses. In the last three years, it has seen a rising trend. It is expected to be higher than sector average that is shown in Figure 2.5.

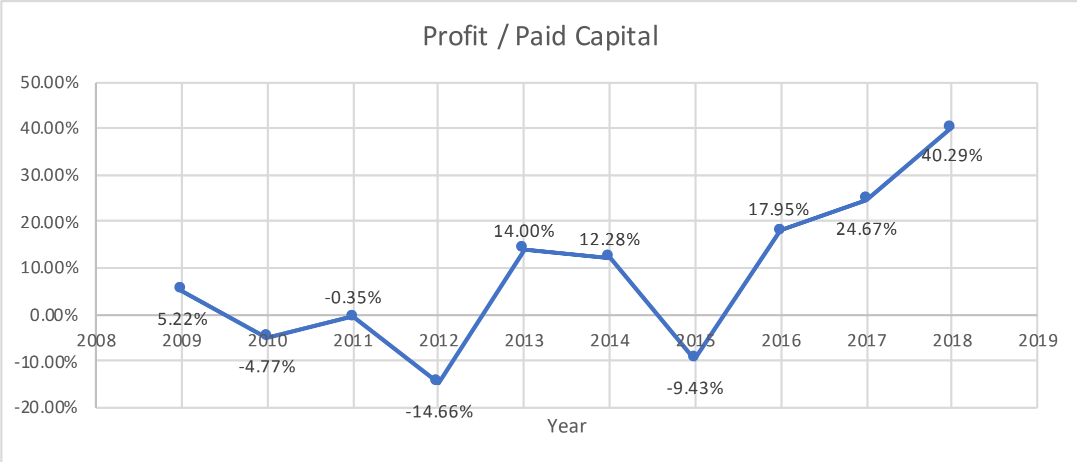


Figure 2.5: Profit / Paid Capital, Non-Life Insurance Sector, 2009-2018

2.6 Premium Production / Coverage

Premium production and coverage values have been collected from Insurance and Individual Pension Annual Report Tables published by TRMTF.

In respect of any risk, the amount paid in cash by the insured or the insurer against the guarantee provided by the insurer is the premium. The coverage is the guarantee of the insurer, the insured or it's beneficiary, in case of damage to the insured in whole or in part, that the damage will be compensated in accordance with the general principles and policy conditions of the insured.

The gross written premium has three sub-accounts. They are direct, indirect, and to be transferred to SSI premiums. The production of insurance policies is recorded to the direct premium account if it is occurred by the insurer itself or the intermediary; otherwise, production by the other is written indirect premium. Direct premium production is realized with central, regional and branch directorates, bank and non-bank agencies and brokers. Insurance pools and hidden co-insurance agreements are monitored in the indirect production account. In addition, premiums written to be transferred to SSI are also examined in a separate account. The premiums written to be transferred to SSI are written by the company but all of them are transferred to SSI. Compulsory earthquake, subsidized agriculture, SSI, Green Card, and high-risk insurance pool are excluded from the direct premium valuation.

$$\text{PP/C} = \frac{(\text{Direct Premium})}{(\text{Coverage})} \tag{2.6}$$

PP/C ratio is expected to be higher than the sector average. Turkish insurance sector averages are shown in Figure 2.6.

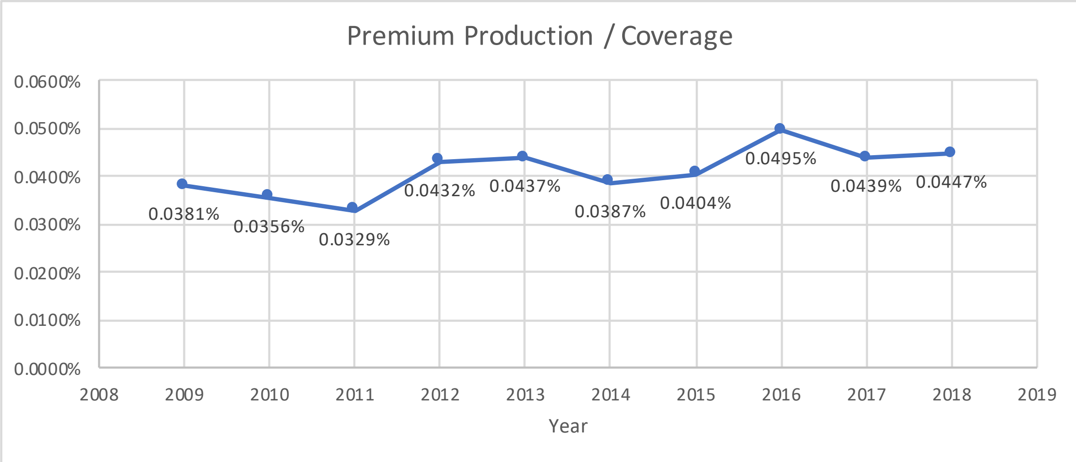


Figure 2.6: Premium Production / Coverage, Non-Life Insurance Sector, 2009-2018

2.7 Payables on Reinsurance Operation / Equity

Reinsurance is insurance of insurance. It is the again insurance of some or all of the insured risk [31]. Reinsurance debts are essential for the continuity of their agreements and acceptance of new business. The reinsurance debts of insurance companies may vary according to the reinsurance policies of the firm, but it may be more accurate to compare them with equity.

$$\text{PRO/E} = \frac{\left(\begin{array}{c}\text{Debts to reinsurance companies}\\ +\text{ Bill of debt rediscount for debts to reinsurance companies}\\ +\text{ Debts due to reinsurance activities}\end{array}\right)}{(\text{ Equity })} \tag{2.7}$$

It is expected to be lower than the sector average. Turkish non-life insurance industry averages are shown in Figure 2.7.

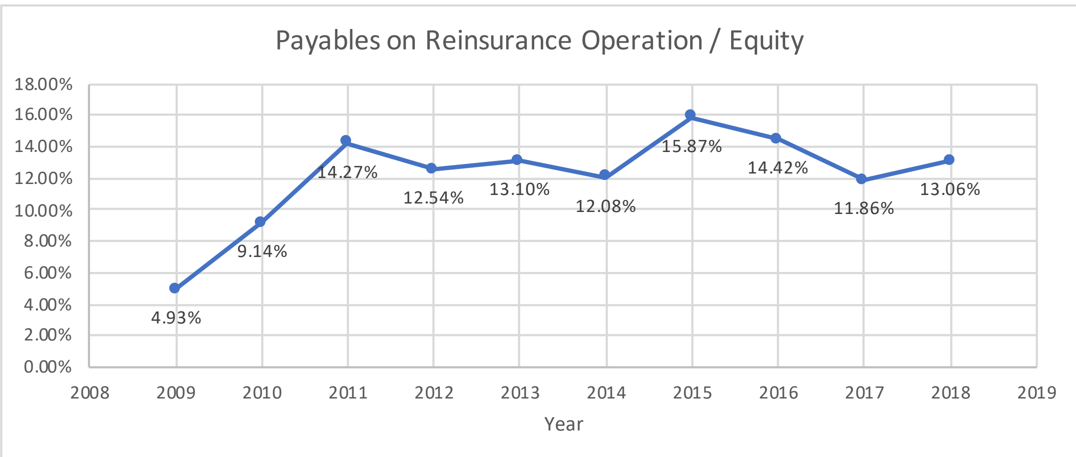


Figure 2.7: Payables on Reinsurance Operation / Equity, Non-Life Insurance Sector, 2009-2018

2.8 Total Reserve / Net Premium

Total reserve is found by the totality of capital reserve and profit reserve. The former has share certificate issuance premiums, share certificate cancellation profits, sale profits to be included in the capital, foreign currency translation differences and other capital reserves. The latter has legal, status, extraordinary, special funds and other profit reserves and financial asset valuation.

$$\text{TR/NP} = \frac{(\text{Capital reserves} + \text{Profit reserves})}{(\text{Written Premiums (Net)})} \tag{2.8}$$

The purpose of the comparison with the net written premium is to examine whether it allows the appropriate reserves. It is expected to be higher than the sector average that is indicated in Figure 2.8.

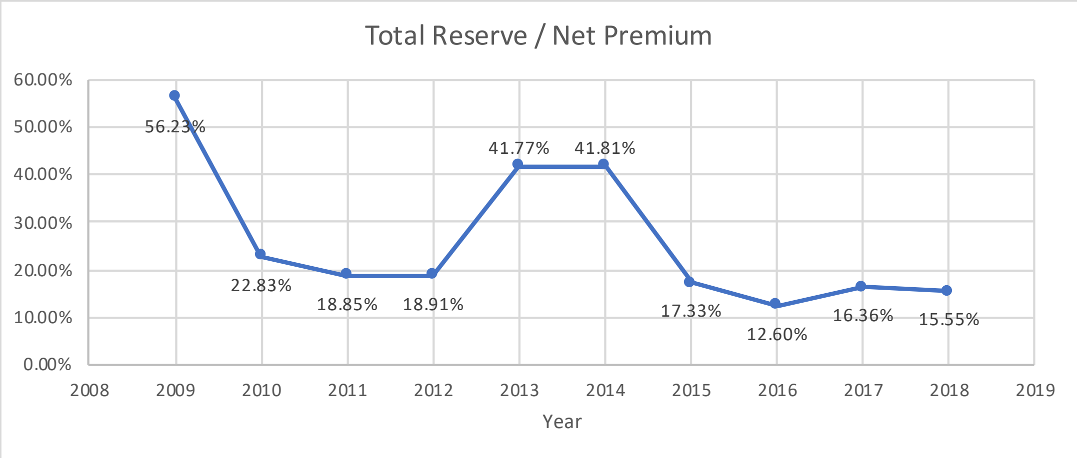


Figure 2.8: Total Reserve / Net Premium, Non-Life Insurance Sector, 2009-2018

2.9 Total Reserve / Liquid Asset

It is aimed to examine the relations of companies' reserves and liquid assets.

$$\text{TR/LA} = \frac{(\text{Capital reserves} + \text{Profit reserves})}{\left(\begin{array}{c}\text{Cash and cash equivalents} \\ + \text{ Financial assets and financial investments whose risk is} \\ \text{assumed by policy holders} \\ - \text{ Financial investments whose risk is assumed by life insurance} \\ \text{policy holders}\end{array}\right)} \tag{2.9}$$

It is expected to be higher than the sector average. We can see in Figure 2.9, how over the years the insurance sector has been following.

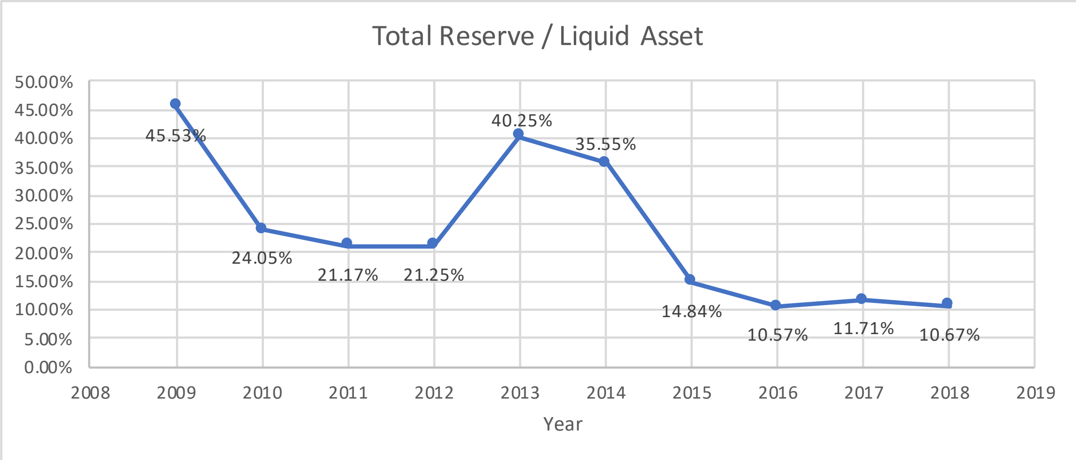


Figure 2.9: Total Reserve / Liquid Asset, Non-Life Insurance Sector, 2009-2018

2.10 Technical Profit / Net Written Premium

This ratio measures the ability of premiums to generate technical profit.

$$\text{TP/NWP} = \frac{\begin{pmatrix}\text{Technical income of Non-Life branches}\\ + \text{Technical expense of Non-Life branches}\end{pmatrix}}{\begin{pmatrix}\text{Written Premiums}\\ \text{(Net)}\end{pmatrix}} \tag{2.10}$$

The positive difference in technical income and expense of Non-Life insurance companies gives profit technically. As can be seen Figure 2.10, the Turkish Non-Life Insurance Sector in 2012 and 2015 declared technical loss. It is expected to be higher than the sector average.

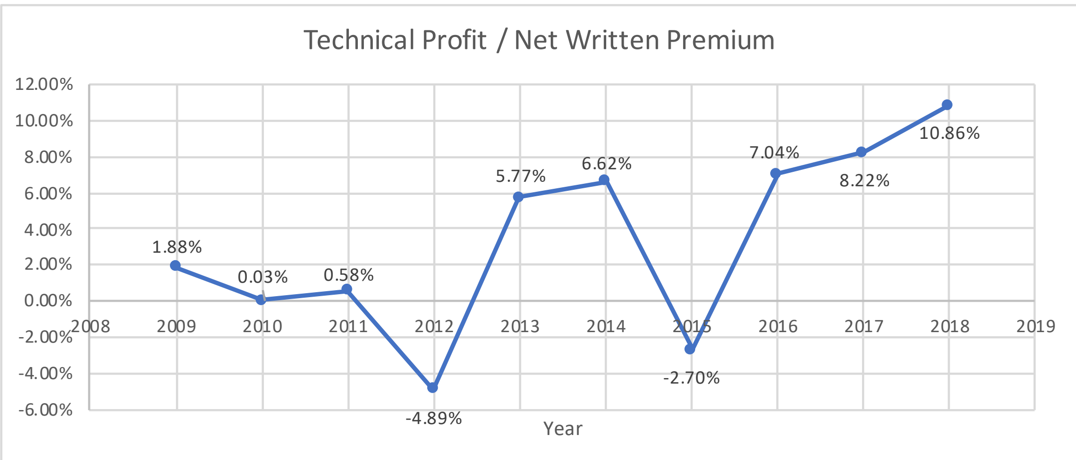


Figure 2.10: Technical Profit / Net Written Premium, Non-Life Insurance Sector, 2009-2018

2.11 Total Income / Total Assets

With this ratio, we can learn to what extent the company can generate revenue with its total assets. Total asset is the sum of current and non-current assets, just like other ratios. Total income means income from non-life insurance. It covers net written premiums, net provision for unearned premiums, net provision for transferred unearned premiums, net provision for ongoing risks, net provision for transferred ongoing risk, investment incomes transferred from the non-technical section, other technical incomes and accrued recourse and recovery incomes.

$$\text{TI/TA} = \frac{(\text{Technical income of Non-Life branches})}{(\text{Current assets} + \text{Non-current assets})} \tag{2.11}$$

TI/TA ratio is expected to be higher than the sector average. Turkish non-life insurance industry averages of this ratio are indicated in Figure 2.11.

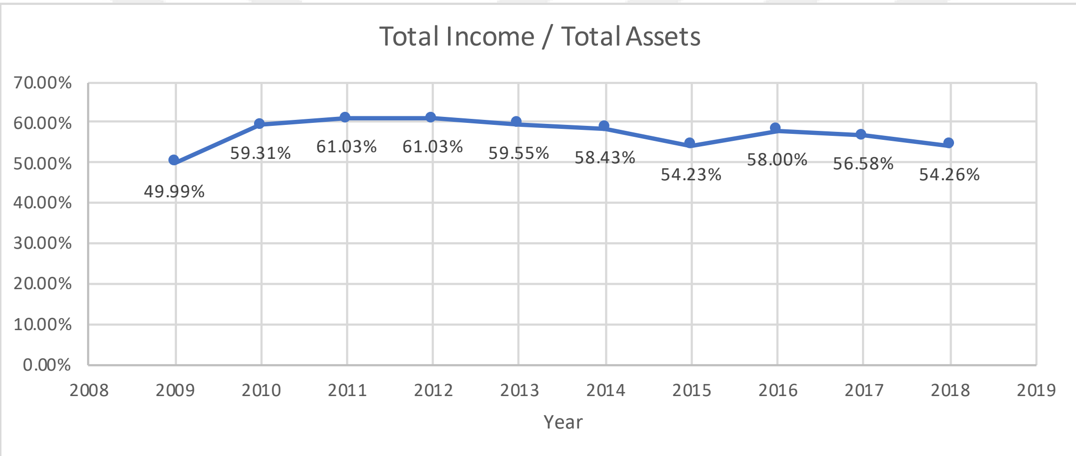


Figure 2.11: Total Income / Total Assets, Non-Life Insurance Sector, 2009-2018

2.12 Reinsurance Share / Gross Premium

Gross written premiums encompass net written premiums and transferred to reinsurer premiums. This ratio points out the transferred premium with reinsurance contract within the gross written premiums. The sum of this ratio and retention ratio is equal to 100%.

$$\text{RS/GP} = -\frac{(\text{Premiums transferred to reinsurer (-)})}{(\text{Written premiums (Gross)})} \tag{2.12}$$

The data of the Turkish Non-Life insurance sector is in Figure 2.12.

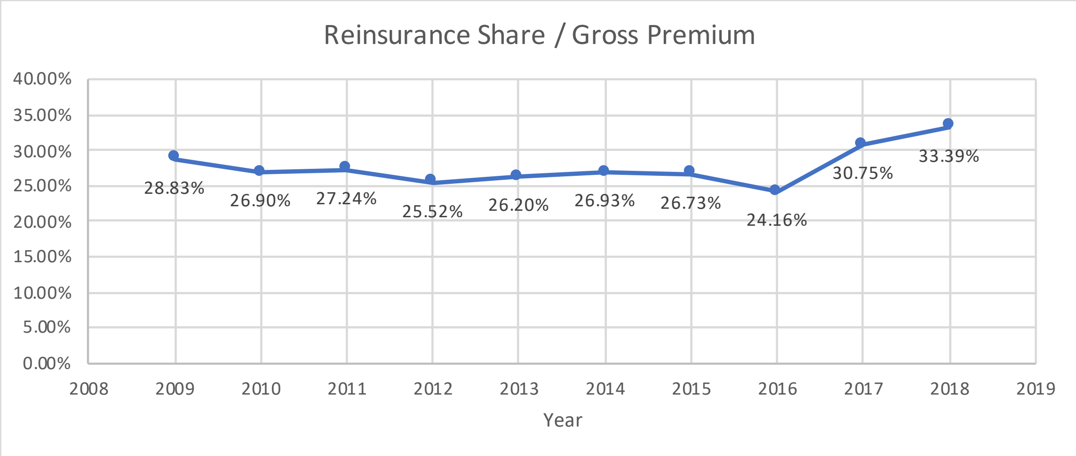


Figure 2.12: Reinsurance Share / Gross Premium, Non-Life Insurance Sector, 2009-2018

2.13 Net Loss Ratio

Net Loss Ratio is the relation of net incurred losses to net earned premiums.

Net paid losses are equal to the sum of gross paid losses and share of reinsurance of paid losses.

Net provision for outstanding losses is equal to the sum of gross provision for outstanding losses and share of reinsurance of provision for outstanding. Net provision for transferred outstanding losses is equal to the sum of gross provision for transferred outstanding losses and share of reinsurance provision for transferred outstanding losses. Change in net provision for outstanding losses is equal to the sum of net provision for outstanding losses and net provision for transferred outstanding losses.

Net paid losses and change in net provision for outstanding losses is equal to net claims(losses) incurred.

The sum of gross written premiums, premiums transferred to reinsurer and premiums transferred to the SSI is net written premiums.

Net provision for unearned premiums is equal to the sum of gross provision for unearned premiums, share of reinsurance of Provision for unearned premiums and SSI share of Provision for unearned premiums. Net provision for transferred unearned premiums is equal to the sum of gross provision for transferred unearned premiums, share of reinsurance of provision for transferred unearned premiums and SSI share of provision for transferred unearned premiums. Change in net provision for unearned premiums is equal to the sum of net provision for unearned premiums and net provision for transferred unearned premiums.

Net provision for ongoing risks is equal to the sum of gross provision for ongoing risks and share of reinsurance of provision for ongoing risks. Net provision for transferred ongoing risks is equal to the sum of gross provision for transferred ongoing risks and share of reinsurance of provision for transferred ongoing risks. Change in net provision for ongoing risks is equal to the sum of net provision for ongoing risks and net provision for transferred ongoing risks.

Net written premiums, change in net provision for unearned premiums and change in net provision for ongoing risks are equal to net earned premiums.

$$\text{NLR} = -\frac{\left(\begin{array}{c}\text{Paid losses (Net) (-)}\\ \text{+ Change in provision for outstanding losses (Net)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums (Net)}\\ \text{+ Change in provision for unearned premiums (Net)}\\ \text{+ Change in provision for onging risks (Net)}\end{array}\right)} \tag{2.13}$$

With the net loss ratio, thanks to the transferred premium and losses transferred to the reinsurance, the actual loss payment and the premium earned are considered. 80% net loss ratio represents the break-even point. It is expected to be lower than the 80% for Turkish non-life insurance companies. Turkish insurance industry average can be looked at in Figure 2.13.

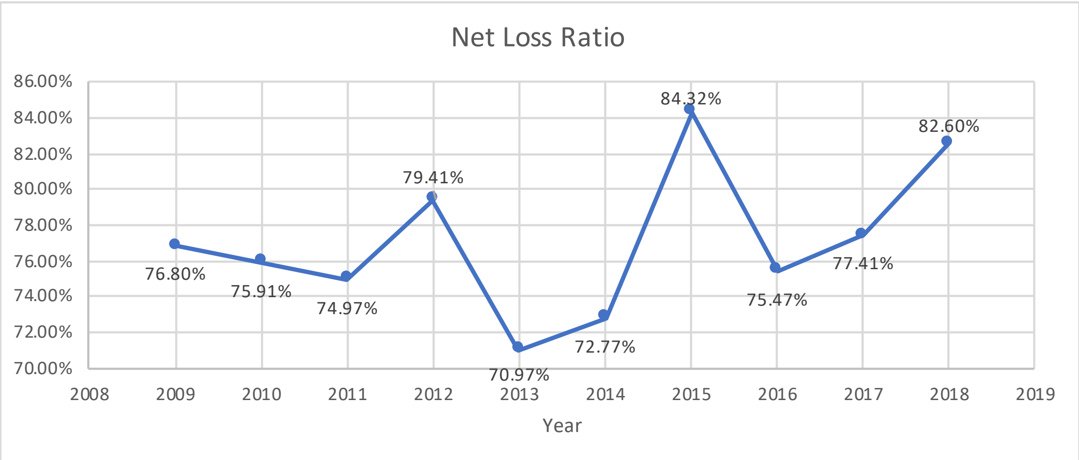


Figure 2.13: Net Loss Ratio, Non-Life Insurance Sector, 2009-2018

2.14 Net Loss Ratio / Gross Loss Ratio

This ratio is calculated by dividing net to gross loss ratios.

$$\text{NLR/GLR} = \frac{\dfrac{\begin{pmatrix}\text{-Paid losses (Net) (-)}\\ \text{- Change in provision for outstanding losses (Net)}\end{pmatrix}}{\begin{pmatrix}\text{Written premiums (Net)}\\ \text{+ Change in provision for unearned premiums (Net)}\\ \text{+ Change in provision for onging risks (Net)}\end{pmatrix}}}{\dfrac{\begin{pmatrix}\text{- Paid losses (Gross) (-)}\\ \text{- Provision for outstanding losses (Gross) (-)}\\ \text{- Provision for transferred outstanding losses (Gross)}\end{pmatrix}}{\begin{pmatrix}\text{Written premiums (Gross)}\\ \text{+ Premiums transferred to the SSI (-)}\\ \text{+ Provision for unearned premiums (Gross) (-)}\\ \text{+ Provision for transferred unearned premiums (Gross)}\\ \text{+ Provision for unearned premiums (SSI share)}\\ \text{+ Provision for transferred unearned premiums (SSI share)}\\ \text{+ Provision for ongoing risks (Gross)}\\ \text{+ Provision for transferred ongoing risks (Gross)}\end{pmatrix}}} \tag{2.14}$$

The net loss ratio should be lower than the gross loss ratio. Otherwise, there are two possible cases. One of these is wrong reinsurance contract; second is SSI share. The former is in case of significant damages by making wrong reinsurance agreement, most of the damage remains on the company. The latter it can be seen that the damages arising from the premiums transferred to SSI are not transferred to what extent.

It is expected to be lower than the sector average. Data of the Turkish non-life insurance sector are indicated in Figure 2.14.

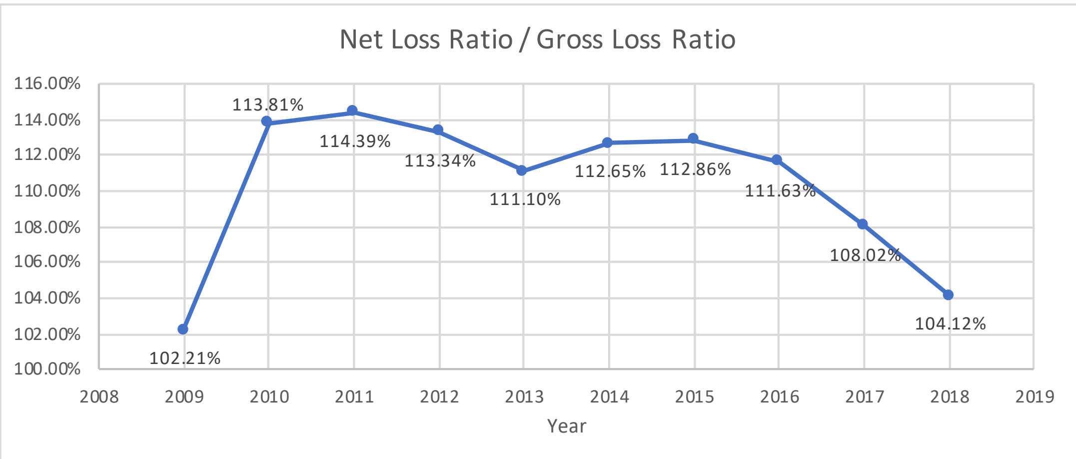


Figure 2.14: Net Loss Ratio / Gross Loss Ratio, Non-Life Insurance Sector, 2009-2018

2.15 Gross Premiums Written to Equity

Equity provides protection for loss payments. The risk for equity and the possible loss burden of the insurer increase with the rising in premiums, which the reinsurance agreements are ignored. This ratio examines the relationship between equity and premiums without the effect of premiums ceded to reinsurance. The problem can arise as a result of a high written gross premium compared to equity. Since reinsurance at this ratio is not taken into consideration, it is vital to interpret with net written premiums to equity. If the difference is enormous, the insurer may rely heavily on the reinsurer. This may not be a problem if the financial strength of reinsurance companies is sound, and there is no disruption in payments. However, even if there is a reinsurance agreement, the insurer's liability to the policyholder remains. In general, the proportion of companies that have a sturdy reinsurance agreement and receive regular payments, which are financially strong, profitable, and stable, are likely to be higher. [61].

$$\text{GPWE} = \frac{(\text{ Written premiums (Gross) })}{(\text{ Equity })} \tag{2.15}$$

With this ratio, it is tested whether the shareholders' equity meets the obligation undertaken due to gross written premiums. According to IRIS, unusual range for the ratio is equal to or over 900%, but in the Turkish Insurance Sector, this ratio is observed between 140% and 374% for ten years. For this reason, we can say that It should be smaller than 400%. For the Turkish insurance industry, the path of this ratio by years can be seen in Figure 2.15.

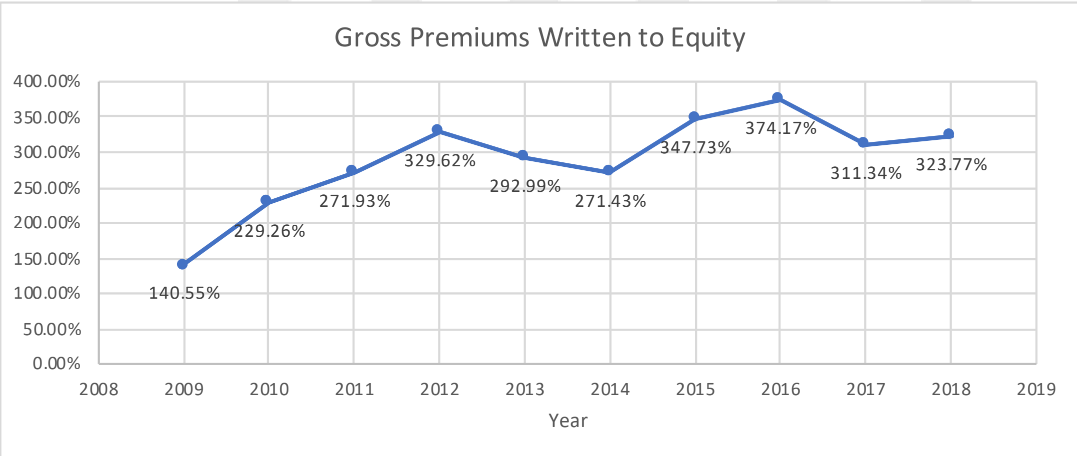


Figure 2.15: Gross Premiums Written to Equity, Non-Life Insurance Sector, 2009-2018

2.16 Net Premiums Written to Equity

With this ratio, it tests whether the equity meets the obligation undertaken due to the production of premiums ceded to reinsurers. The higher the net premium to equity ratio, the more risk the insurer bears in relation to equity. The result from high net premiums written in relation to equity could be a problem. Strong and large companies with stable profits can better tolerate higher Net Premiums Written to Equity ratio than financially non-clear companies.

The main branches of firms should also be taken into consideration. The variability of losses increases in branches that are difficult to estimate potential losses. Therefore, it is better to have a lower ratio in these branches. At the same time, the level of adequacy of reinsurance assurance against major losses is important. Reinsurance contracts and levels of retention should be reviewed on behalf of this. Financial stability and financial statements of reinsurers should be examined. The ability of reinsurance companies to pay high losses is critical for an insurer. Unusual range for the ratio is equal to or over 400% for IRIS; however, in the Turkish insurance sector, this ratio is observed between 97$\%$ and 270$\%$ for ten years in Figure 2.16. Because of this reason, we can say that It should not be bigger than 200%.

$$\text{NPWE} = \frac{(\text{Written premiums (Net)})}{(\text{Equity})} \tag{2.16}$$

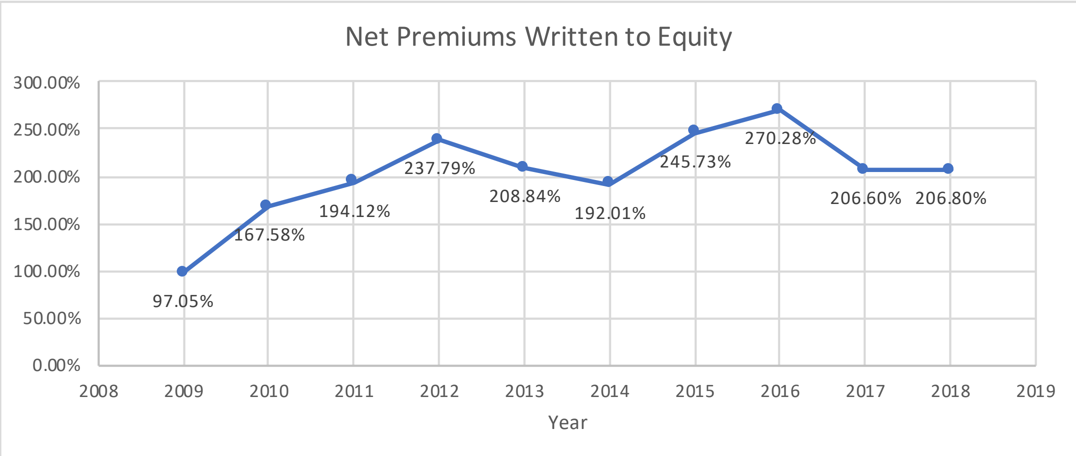


Figure 2.16: Net Premiums Written to Equity, Non-Life Insurance Sector, 2009-2018

2.17 Change in Net Written Premiums

Important changes in net written premiums could demonstrate a lack of stability in the insurer's operations or management. A major increase in premiums might hint entry into new lines of business, branch or geographic locations. Moreover, such an increase in written premiums may be a sign that the insurer is trying to increase cash flow in order to meet current or expected loss payments. Companies entering the insurance sector may have outlier values at this rate. On the other hand, a major decrease in written premiums might indicate the stopping of certain lines of business or branch, scaled back writings due to huge losses in specific lines, loss of market share because of competition, or increased use of reinsurance choice. Based on the insurer's activities and history, it is useful to evaluate the ratio results. These outcomes generally may contain dramatic changing in product mix, marketing areas, or underwriting policy. Moreover, whether the insurer's assets are properly valued and adequate liquidity is available to meet cash demands should determine. If the company reserves are insufficient, they may be more willing to increase the premium. An increment in writings, especially in the liability lines, to pay current claims provides a very short-term cure to underlying problems and quickly increases the risk of bankruptcy. However, an increase in premiums does not indicate difficulties that would threaten an insurer's solvency if they are accompanied by a reasonably low Gross Premiums Written to Equity ratio, enough reserving, operations that can profit, and a relatively steady product mixture. A decrease in net premiums written with stable gross premiums written can report that the insurance company is attempting to increase cash flow related to ceding commissions from reinsurance.

$$\text{CNWP} = \frac{\begin{pmatrix} \text{Written premiums (Net) (Current Year)} \\ \text{- Written premiums (Net) (Previous Year)} \end{pmatrix}}{(\text{Written premiums (Net) (Previous Year)})} \tag{2.17}$$

Values between -33% and 33% are acceptable for IRIS, and we can also accept it for the Turkish Insurance Sector because the last ten years this ratio is formed between 2.56% and 33.93%. According to the data set, the Turkish insurance sector can be shown in Figure 2.17.

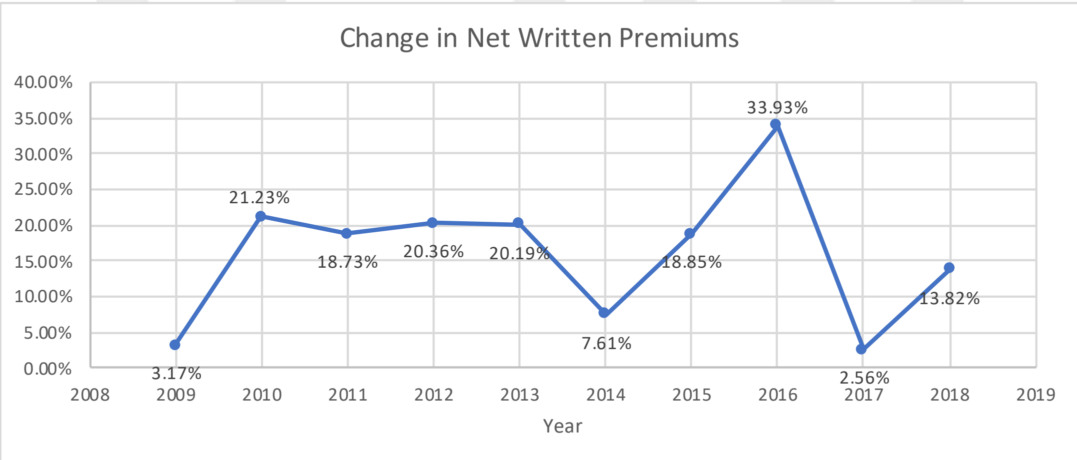


Figure 2.17: Change in Net Written Premiums, Non-Life Insurance Sector, 2009-2018

2.18 Change in Gross Written Premiums

Change in gross premiums is a ratio that needs to be closely examined, just like the change in net premiums. Transferred premium to reinsurers is not taken into account in the calculation of change in gross premiums ratio. The increase and decrease of the premiums written throughout the year vary according to the operational preferences of the companies or the competition conditions. This ratio, adjusted from the reinsurance agreements of the company, shows the gross premium production of the company. High premium growth may have been due to the fact that the company has started a new branch or because it wants to create cash flow as a company policy. The reasons for high premium decreases may be the decrease in production as a result of competition or the decrease in income as a result of the increase in insurance policy prices. Other strong sides of companies, such as low loss premiums, high profits or strong liquid assets, can compensate for huge increases and decreases in net written premiums.

$$\text{CGWP} = \frac{\begin{pmatrix} \text{Written premiums (Gross) (Current Year)} \\ \text{- Written premiums (Gross) (Previous Year)} \end{pmatrix}}{(\text{Written premiums (Gross) (Previous Year)})} \tag{2.18}$$

-10% and 50% are acceptable for insurance firms at the Turkish Insurance Sector. For the last ten years, this ratio is formed between 2.52% and 33.93%. It is shown in Figure 2.18.

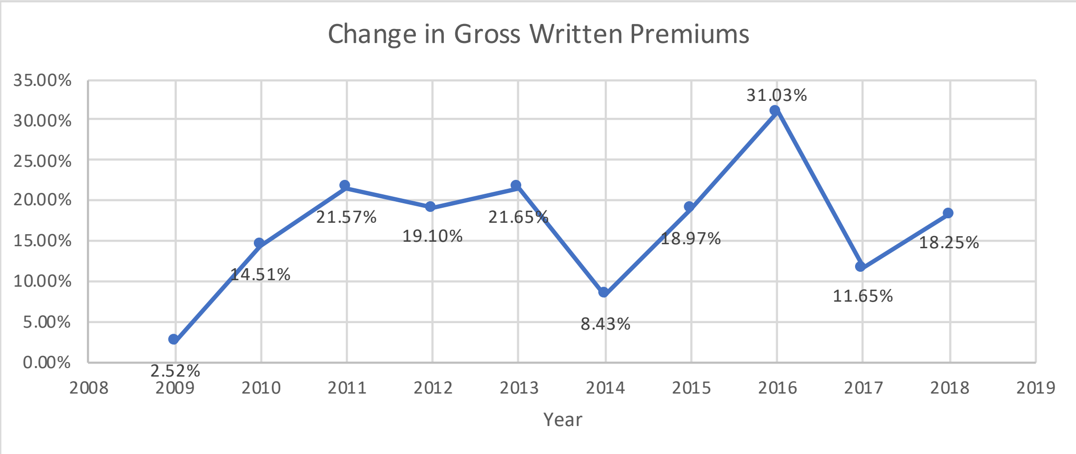


Figure 2.18: Change in Gross Written Premiums, Non-Life Insurance Sector, 2009-2018

2.19 Doubtful Receivables / Total Assets

This ratio measures the weight of doubtful receivables between current and non-current assets. Receivables that cannot be collected during the payment period are recorded in the doubtful receivables account. The maturity of these receivables may be extended or protested.

$$\text{DR/TA} = \frac{(\text{Doubtful receivables from operating activities})}{(\text{Current assets + Non-current Assets})} \tag{2.19}$$

It is expected to be lower than the sector average. The higher this ratio may jeopardize the company's situation. The proportions followed by the insurance industry average over the years are shown in Figure 2.19.

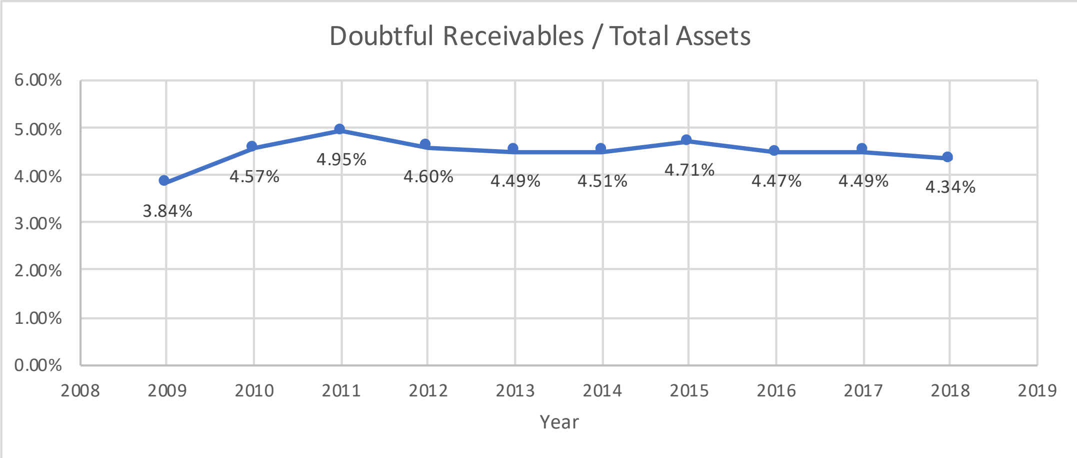


Figure 2.19: Doubtful Receivables / Total Assets, Non-Life Insurance Sector, 2009-2018

2.20 Share of Reinsurance of Provisions / Equity

The definitions of all provisions are made in Article 16 named Technical reserves in Chapter 4 of the No. 5684 Insurance Law [52]. Insurance and reinsurance companies are obliged to allocate sufficient provisions for their liabilities arising from insurance agreements in accordance with the principles specified in the law. The share of reinsurer in technical provisions should be commensurate with the transferred risk and premium. These technical provisions are unearned premiums, unexpired risks, equalization, mathematical, outstanding claims, bonus and rebate, and other technical reserves.

$$\text{SRP/E} = \frac{\begin{pmatrix} \text{Provision for unearned premiums (Share of Reinsurance) (-)} \\ + \text{Provision for ongoing risks (Share of Reinsurance) (-)} \\ + \text{Mathematic provision (Share of Reinsurance) (-)} \\ + \text{Provision for outstanding losses (Share of Reinsurance) (-)} \\ + \text{Provisions for bonuses and discounts (Share of Reinsurance) (-)} \\ + \text{Other technical provisions (Share of Reinsurance) (-)} \end{pmatrix}}{(\text{Equity})} \tag{2.20}$$

This ratio is the result of comparing the reinsurance share of all technical provisions with the equity of the insurance company. The non-life insurance sector average is shown in Figure 2.20 and the values of the companies are expected to be above this average.

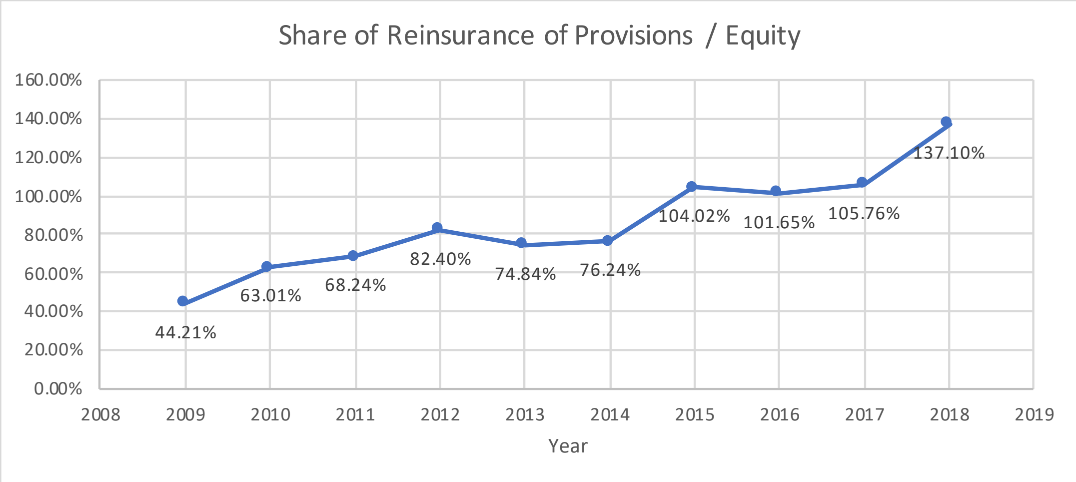


Figure 2.20: Share of Reinsurance of Provisions / Equity, Non-Life Insurance Sector, 2009-2018

2.21 Cash Ratio

Cash ratio measures the company's ability to meet its short term liabilities with the most liquid assets in total assets. Short-term liabilities express liabilities that are due within one year or less. Cash ratio which can be used as the quick ratio in some sources [8], has importance because of the possibility to respect its responsibility in due time.

$$\text{CashR} = \frac{\left(\begin{array}{c}\text{Cash and cash equivalents}\\ +\text{ Financial assets and financial investments whose risk is}\\ \text{assumed by policy holders}\\ -\text{ Financial investments whose risk is assumed by life insurance}\\ \text{policy holders}\end{array}\right)}{(\text{Short term liabilities})} \tag{2.21}$$

The cash ratio can be more helpful when it is compared with sector and competitor averages. This ratio lower than 0.95 does sometimes indicate that insurance firm is at risk of having financial difficulty. This ratio has a trend shown in Figure 2.21 based on the information collection.

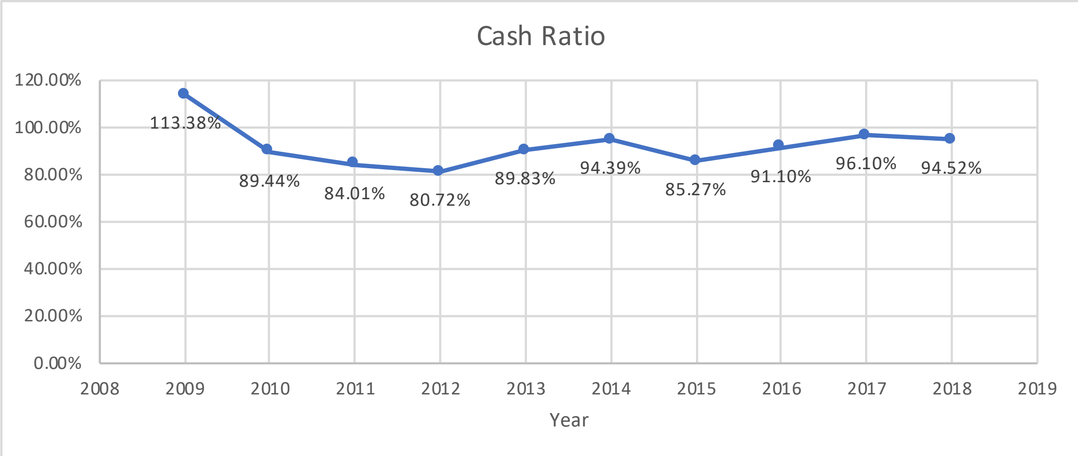


Figure 2.21: Cash Ratio, Non-Life Insurance Sector, 2009-2018

2.22 Assets Covering Technical provisions / Technical Provisions

Assets covering technical provision value calculation conditions are stated in named Regulation on Technical Provisions of Insurance, Reinsurance and Pension Companies, and Assets on which Such Provisions are to be Invested published by Undersecretariat of Treasury, Prime Ministry of Republic of Turkey. The objective of this Regulation is to ensure that insurance and reinsurance companies set aside technical provisions to meet their existing and potential liabilities, and to regulate the methods and principles regarding the assets on which those provisions shall be invested.

In this regulation, technical provisions are defined, and then assets to be invested in technical provisions are specified. These assets are Turkish Lira, foreign currencies bought and sold by the Central Bank of Republic of Turkey, current and time-deposit bank accounts in Turkish Lira, current accounts and participation accounts in participation banks, blocked accounts for credit cards, foreign exchange deposit accounts, government bonds, treasury bonds and other financial assets issued by the state, bonds and other fixed-income financial assets issued by the private sector, stocks and other variable-income financial assets, investment fund participation certificates, repo transactions, receivables from technical operations and reinsurer shares in technical provisions, loans offered with or without respect to insurance contracts, real estate, fixed assets other than real estate, taxes and funds paid in advance, and deferred tax assets [53].

$$\text{ACTP/TP} = \frac{(\text{Assets covering technical provisions})}{(\text{Technical provisions})} \tag{2.22}$$

We can find the adequacy of assets in terms of technical provisions with the ratio of assets that can be invested in technical provisions to technical provisions. It can be said to be more powerful for non-life insurance companies with a higher ratio. It must be at least 100%. This ratio shows a pattern in Figure 2.22 based on the data set.

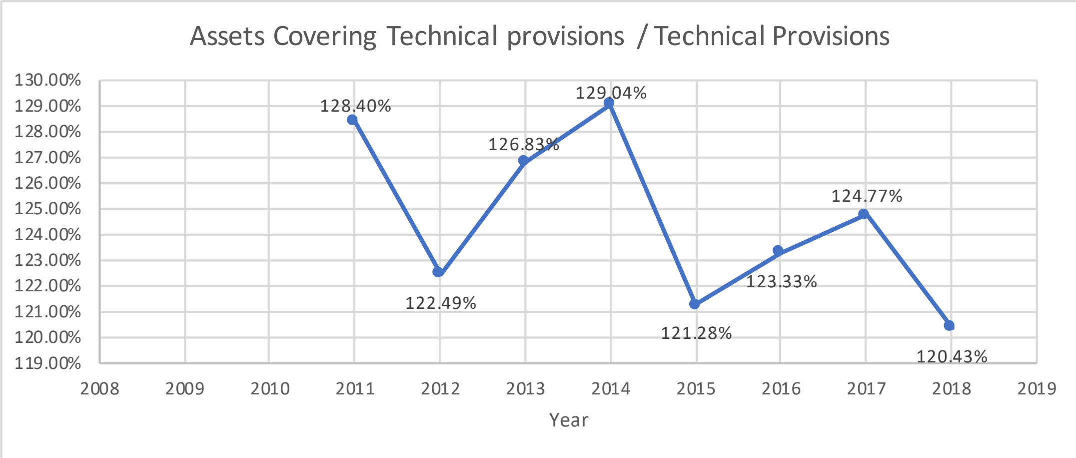


Figure 2.22: Assets Covering Technical provisions / Technical Provisions, Non-Life Insurance Sector, 2011-2018

2.23 Investment Properties / Total Assets

Investment property account under tangible assets is the assets that the company does not use as fixtures and directs to investment. Companies may wish to obtain more stable returns by diversifying their assets by investing in land, buildings and other real estates. This ratio is divisional fixed assets held for investment to total assets.

Accounting the real estate investment in the insurance sector in Turkey is performed according to a standard called Turkish Accounting Standard 40 - Investment Property. In order to be considered as investment property in accordance with this standard, it is necessary that it is possible to enter future economic benefits related to real estate and the cost of investment property can be measured reliably [55].

$$\text{IP/TA} = \frac{(\,\text{Investment properties}\,)}{(\,\text{Current assets} + \text{Non-current Assets}\,)} \tag{2.23}$$

Figure 2.23 shows the sector average of non-life insurance sector values between 2009 and 2018. Among these years, a generally decreasing graph was observed. The company's ratio should be greater than the sector average.

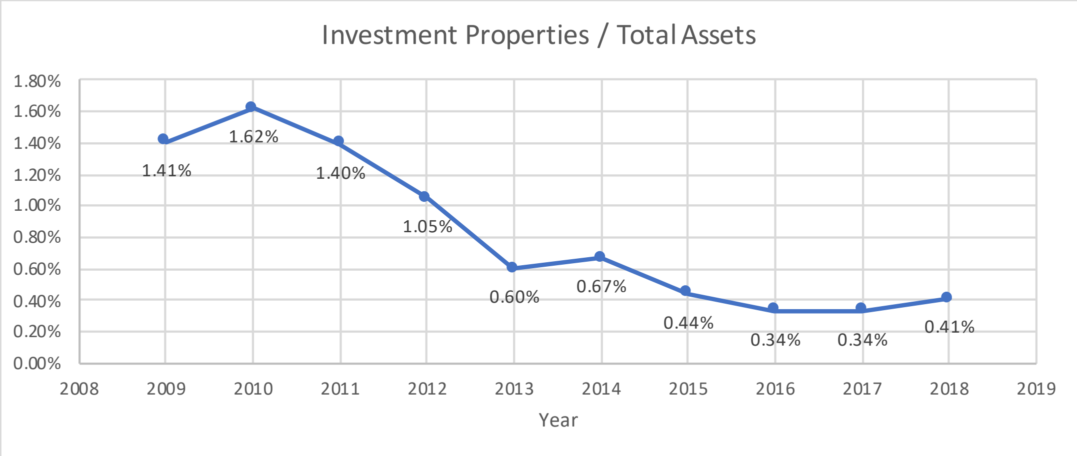


Figure 2.23: Investment Properties / Total Assets, Non-Life Insurance Sector, 2009-2018

2.24 Return on Investment

ROI ratio is calculated by dividing the net investment profit and all invested assets. The sum of investment income and expenses, and the total assets invested are considered. It is used to measure the investment success of the company and shows the quality of a company investment portfolio [49]. It is expected to be higher than the sector average.

$$\mathrm{ROI} = \frac{\left(\begin{array}{c}\text{Investment incomes transferred from the Non-technical section} \\ + \text{Investment incomes} \\ + \text{Investment expenses}\end{array}\right)}{\left(\begin{array}{c}\text{Cash and cash equivalents} \\ + \text{Financial assets and financial investments whose risk is} \\ \text{assumed by policy holders} \\ + \text{Financial assets} \\ + \text{Investment properties} \\ + \text{Provision for impairment of investment properties (-)}\end{array}\right)} \tag{2.24}$$

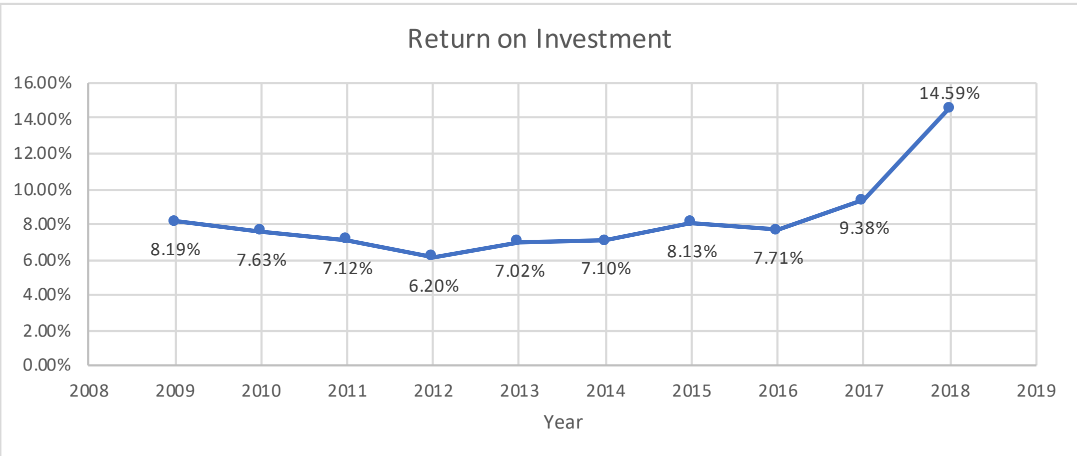


Figure 2.24: Return on Investment, Non-Life Insurance Sector, 2009-2018

2.25 Return of Assets

The ratio of net profit or loss of the period to total assets measures the return on total assets after interest and taxes. This ratio is expected to be high. The low return from sector average can due to the firm's low earning strength and too high-interest costs resulting from its above average debt having [9].

$$\text{ROA} = \frac{(\text{ Profit or loss of the period })}{(\text{ Current assets} + \text{Non-current Assets })} \tag{2.25}$$

As seen in Figure 2.25, Roa is negative due to the loss of the insurance sector in 2012 and 2015. In the last three years, revenues of insurers have increased.

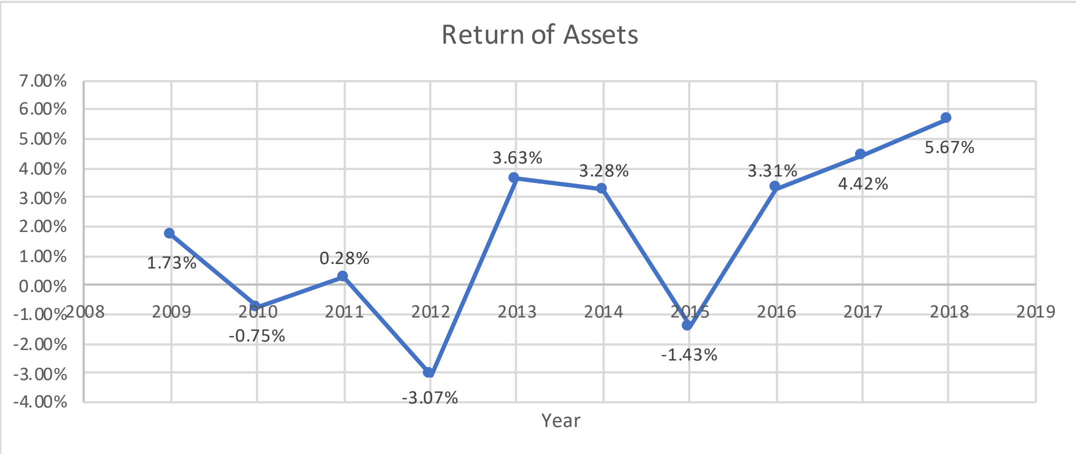


Figure 2.25: Return of Assets, Non-Life Insurance Sector, 2009-2018

2.26 Technical Profit / Gross Written Premium

TP/GWP ratio measures the power of premiums to generate technical profit. Unlike technical profit to net premiums written ratio, calculation of this gross ratio has reinsurance transferred premiums.

$$\text{TP/GWP} = \frac{\left(\begin{array}{c}\text{Technical income of Non-Life branches} \\ + \text{Technical expense of Non-Life branches}\end{array}\right)}{(\text{Written premiums (Gross)})} \tag{2.26}$$

The values shown in Figure 2.26 are the non-life sector averages of TP/GWP ratio. According to our data sets, a company has to higher ratio than the sector average. It was negative in 2012 and 2015.

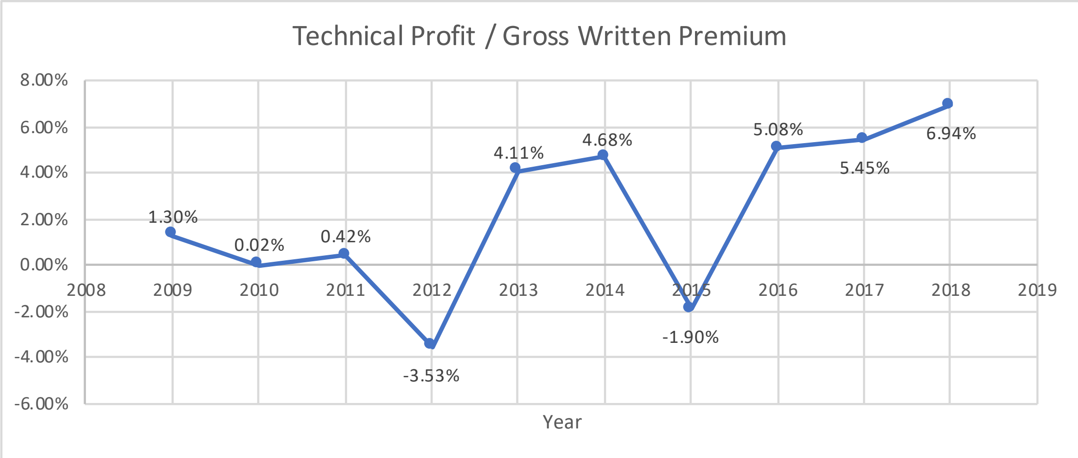


Figure 2.26: Technical Profit / Gross Written Premium, Non-Life Insurance Sector, 2009-2018

2.27 Capital Requirements Ratio

Capital requirement ratio shows whether the equity of the insurance company is sufficient for the risks to which it is exposed due to insurance business. This ratio is calculated in accordance with the Regulation on Measurement and Assessment of Capital Requirements of Insurance and Reinsurance and Pension Companies published in the Official Gazette No. 26761 by the Republic of Turkey Undersecretariat of Treasury (Prime Ministry) on 19.01.2008 [54]. The goal of this regulation is to secure that insurance, reinsurance and pension companies have adequate capital to cover claims which might appear due to their present liabilities and probable risks.

The equity capital of the company is calculated by making deductions items specified in the fifth article of regulation from the sum of the main capital, additive capital and other capital items. Main capital is the sum of paid capital, capital reserves, profit reserves, the gains and losses reflected in equity, net profit after deducting dividends and retained profits. Additive capital is the sum of equalization reserve, all subordinated loans for companies established in the form of cooperatives and up to 30% of the required equity for other companies. Items to be deducted are loss for the period, previous years losses, company's own shares, the amount calculated by multiplying the paid-up capital of subsidiaries, affiliates, affiliate securities and joint ventures with the share of the parent company in such and other items to be determined by the Undersecretariat.

The required capital is calculated by two methods and the larger calculation shows the required capital value of the insurer. After the details of the ratio calculation, administrative measures are mentioned.

Equity should not be lower than the required equity. If capital requirements ratio of insurance company is between 100% and 115% is the self-evaluation stage, between 70% and 99.99% is take precaution stage, between 33% and 69.99% is emergency take precaution stage and less than 33% is intervention stage.

$$\text{CRR} = \frac{(\text{ Equity based to capital requirement })}{(\text{Required equity capital})} \tag{2.27}$$

The values of the capital requirements ratio calculated on the basis of Turkish non-life insurance companies' total sector data are shown in Figure 2.27. In view of these information, in 2012 and 2015, the Turkish insurance sector is in the self-evaluation phase. The reason for this may be the loss of the industry in those years. Other years' ratios are better than these.

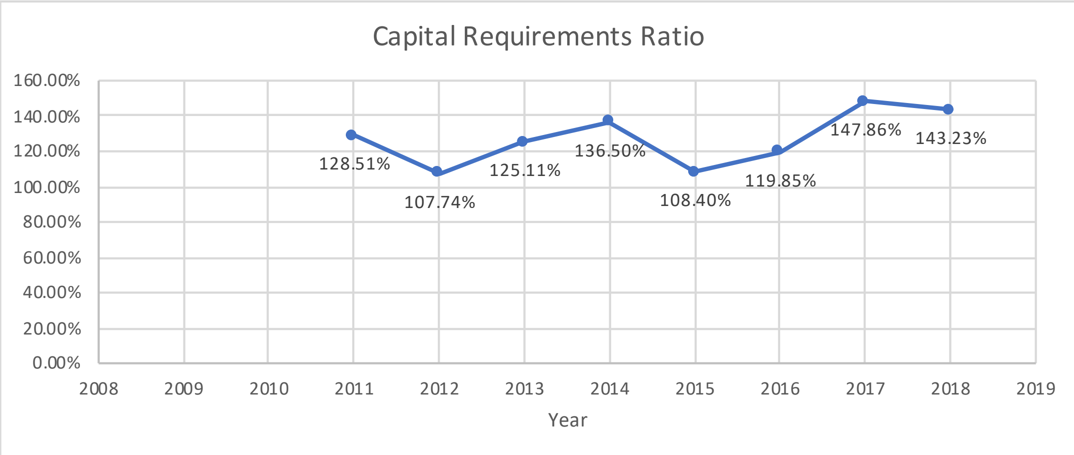


Figure 2.27: Capital Requirements Ratio, Non-Life Insurance Sector, 2011-2018

2.28 Equity / Total Payables

It is calculated by dividing equity by the total payables. Long and short term liabilities are total payables. The reason for examining the ratio is to measure how much of the insurance company's total liabilities can be met with its equity.

$$\text{E/TP} = \frac{(\text{ Equity })}{(\text{Short term liabilities + Long term liabilities})} \tag{2.28}$$

The reflection of the Turkish non-life insurance sector average is in Figure 2.28. According to our data, E/TP ratio of the insurance firms should be higher than 25%.

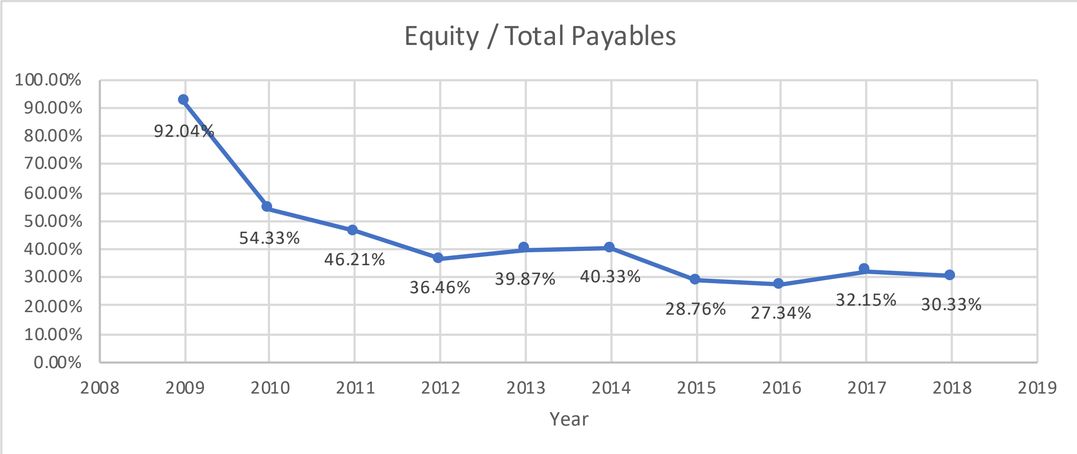


Figure 2.28: Equity / Total Payables, Non-Life Insurance Sector, 2009-2018

2.29 Changes in Equity

The rate of change in equity provides information about the insurer's financial position during the year. This is a sign of recovery or deterioration. As a result of the decrease in equity, our ratio has a negative value. Huge increases may also be undesirable. We can see like this increase in the history of many bankrupt insurance companies. These significant increases may also be indicative of instability, or as a result of significant mergers or acquisitions. Companies with very low or very high rates should be examined closely. In particular, the realization of these extreme values in consecutive years requires detailed examination and consultation. This ratio is expected to be higher than -10% and lower than 50% for both IRIS[61] and the Turkish insurance industry.

$$\mathrm{CE} = \frac{\left(\begin{array}{c}\text{Equity (Current Year) - Equity (Previous} \\ \text{Year)}\end{array}\right)}{(\text{Equity (Previous Year)})} \tag{2.29}$$

In Figure 2.29, we can see how the insurance sector has been following over the years. In 2009, 2010 and 2015, this ratio is negative. In other words, there has been a decrease in equity values in these years.

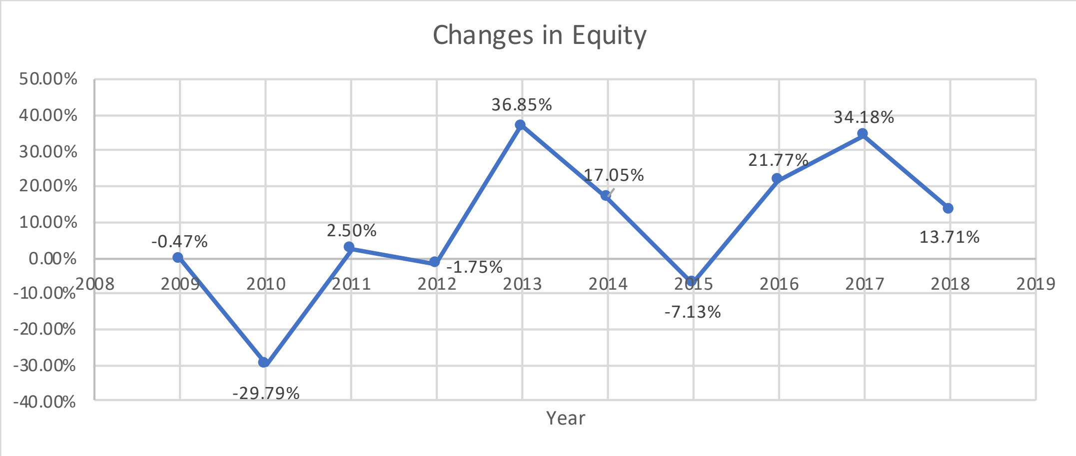


Figure 2.29: Changes in Equity, Non-Life Insurance Sector, 2009-2018

2.30 Assets in Foreign Currency / Liabilities in Foreign Currency

The goal is measuring foreign exchange assets and liabilities of insurance companies in Turkey arising from their transactions about non-life branches to provide indicators for the FX risk they bear. This ratio calculated as dividing foreign currency assets to foreign currency liabilities. The foreign exchange risk that the Company is exposed to is related to the value of this rate. Companies should take positions and make decisions according to market expectations.

$$\text{AFC/LFC} = \frac{(\text{Total assets in foreign currency})}{(\text{Total liabilities in foreign currency})} \tag{2.30}$$

While assets are expected to exceed liabilities in general, foreign currency assets and liabilities should be in the same way. Otherwise, the company is exposed to both asset risk and foreign currency risk. As shown in Figure 2.30, the average rates of Turkish insurance companies were at least 138.5% in 2012.

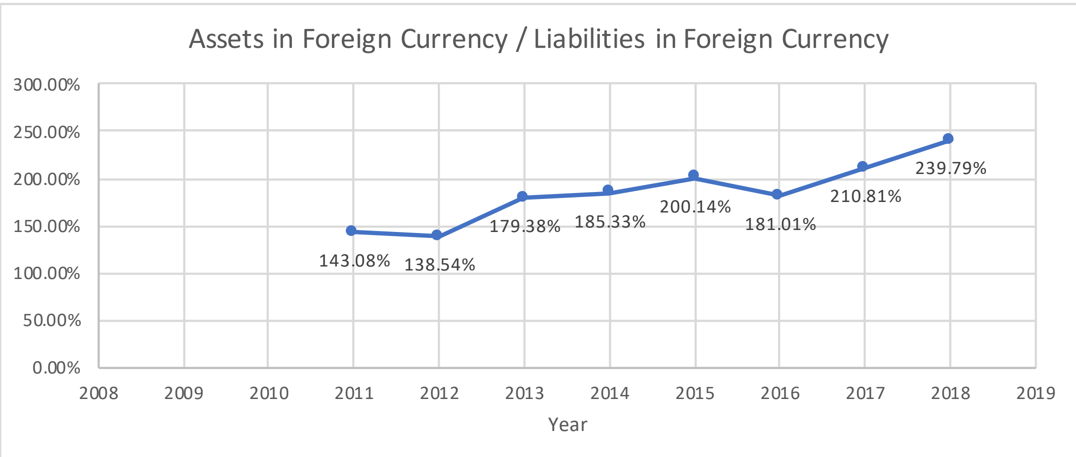


Figure 2.30: Assets in Foreign Currency / Liabilities in Foreign Currency, Non-Life Insurance Sector, 2011-2018

2.31 Net Expense Ratio

The expense ratio measures the level of non-life operating expenses in the technical part. This ratio measures operating expenses to net earned premiums. Operating expenses contain production, personnel, management, research and development, outside benefit and service, and other operation expenses and reinsurance commissions. This ratio is a measure of profitability. Moreover, the expense ratio states the non-life insurance company's efficiency before factoring in claims on its policies and investment profits or losses.

$$\text{NER} = -\frac{(\text{Operating expenses (-)})}{\begin{pmatrix}\text{Written premiums (Net)} \\ \text{+ Change in provision for unearned premiums (Net)}\end{pmatrix}} \tag{2.31}$$

NER ratio was observed between 22% and 32.7% according to the data set in Figure 2.31. The smaller the ratio, the better for the company.

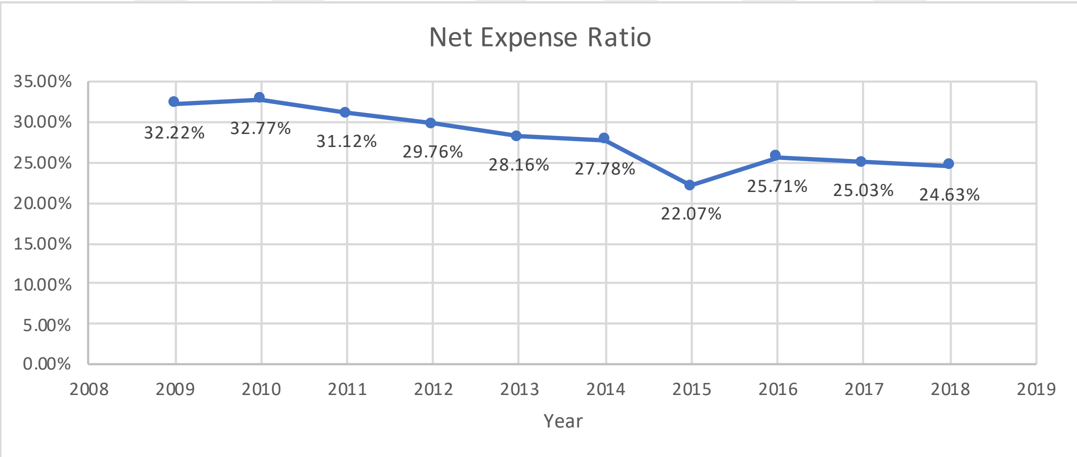


Figure 2.31: Net Expense Ratio, Non-Life Insurance Sector, 2009-2018

2.32 Motor Portfolio Share

Motor portfolio share ratio measures the share of motor vehicles and motor vehicles liability insurance in gross premium generated in all branch types. Motor vehicles liability insurance is motor vehicles compulsory third party liability and contains compulsory road transportation financial responsibility, highway motor vehicles compulsory liability insurance and facultative financial liability branches.

The Turkish insurance sector has a stable structure. Premium production of the majority of a non-life insurer is obtained by the first ten companies. In 2017, 73% of the premium in non-life insurance was produced by the first ten firms. Therefore, the activities of small scale companies in the industry have existed barrier [34].

$$\text{MPS} = \frac{\begin{pmatrix}\text{Premium production of Motor Vehicles} \\ +\text{ Premium production of Motor Vehicles Liability}\end{pmatrix}}{(\text{ Total gross premium production })} \tag{2.32}$$

Turkish non-life companies' average ratios are shown in Figure 2.32.

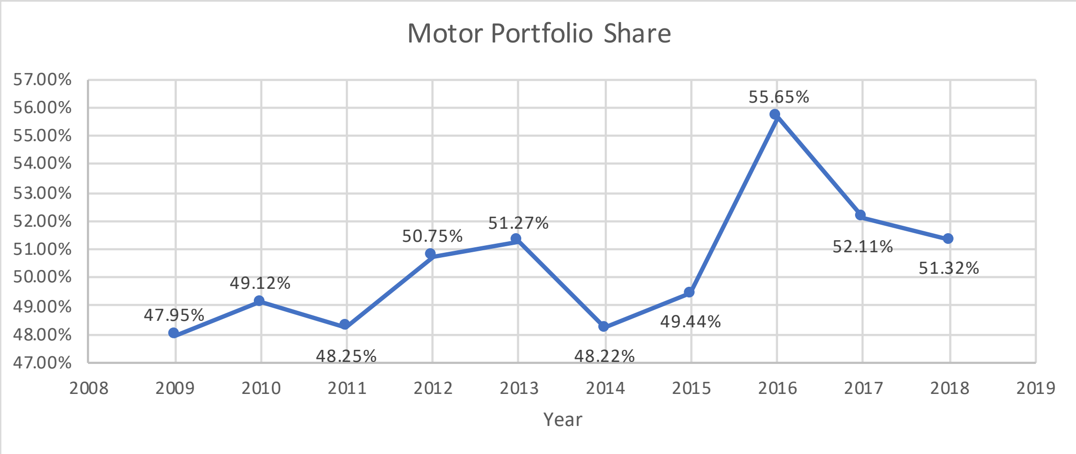


Figure 2.32: Motor Portfolio Share, Non-Life Insurance Sector, 2009-2018

2.33 Net Paid Losses / Gross Paid Losses

The difference between gross and net losses is the amount of loss transferred to the reinsurer. The indicator of the reinsurance success of insurance companies is the paid premiums for reinsurance and the ceded payment of claims in the events of loss in return for premiums. The less the share of reinsurance of premiums and the higher transferred loss to reinsurance company has, the more successful it gets for operations. Claims incurred are calculated by summing paid losses and changes in outstanding claims reserve. The net/gross incurred loss ratio is not considered as the sole indicator of the success or failure of reinsurance. However, reducing the burden of damage payments can be seen as successful.

$$\text{NPL/GPL} = \frac{\begin{pmatrix}\text{Paid Losses (Net) (-)} \\ +\text{ Change in provision for outstanding losses (Net)}\end{pmatrix}}{\begin{pmatrix}\text{Paid losses (Gross) (-)} \\ +\text{ Change in provision for outstanding losses (Gross)}\end{pmatrix}} \tag{2.33}$$

Turkish non-life companies in Figure 2.33 are represent average values and values below these rates are considered good for insurance companies.

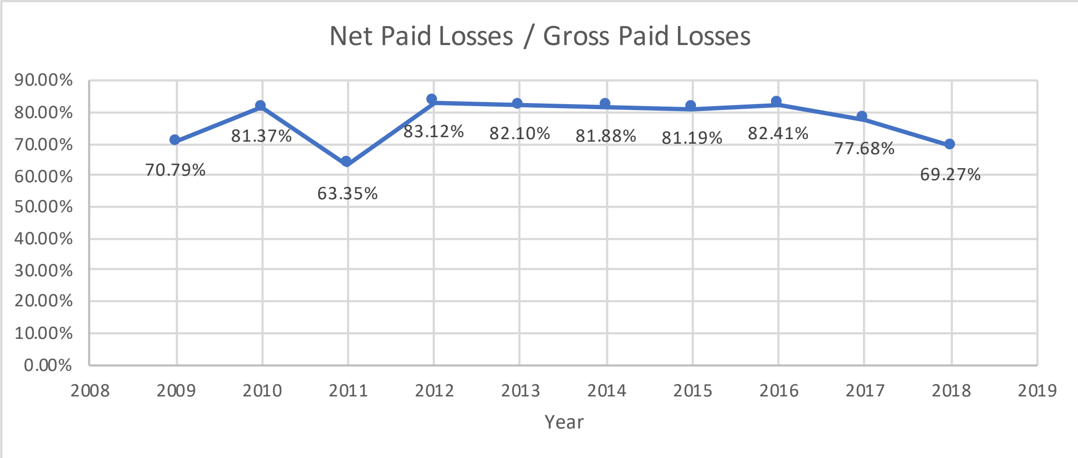


Figure 2.33: Net Paid Losses / Gross Paid Losses, Non-Life Insurance Sector, 2009-2018

2.34 Changes in Retention Ratio

The conservation shows how much the company holds on written premiums. Companies transfer some of their premium production to reinsurance companies through a reinsurance agreement. The remaining portion after the reinsurer agreements shows the net premium production of the company. This ratio is a measure of the insurer's retention for own account. Insurers having a low retention ratio and high solvency requirement ratio would appear to be acting like an agent and relying on earning a commission of reinsurance. Nevertheless, the branch of insurance is important to decide about the retention ratio. If the majority of the written premiums are of a specialized nature, then retention tends to be low because of the extra high-risk exposure of this kind of business. The retention ratio is the net written premium to the gross written premiums [49]. The sum of the retention ratio and the reinsurance share to gross premiums ratio equals to 100%.

$$\text{Retentio Ratio} = \frac{\text{Written premiums (Net)}}{\text{Written premiums (Gross)}}$$

This ratio provides information about the change in the retention ratio described above.

$$\text{CiRR} = \frac{\dfrac{\left(\begin{array}{c}\text{Written premiums (Net)}\\ \text{(Current Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ \text{(Current Year)}\end{array}\right)} - \dfrac{\left(\begin{array}{c}\text{Written premiums (Net)}\\ \text{(Previous Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ \text{(Previous Year)}\end{array}\right)}}{\dfrac{\left(\begin{array}{c}\text{Written premiums (Net)}\\ \text{(Previous Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ \text{(Previous Year)}\end{array}\right)}} \tag{2.34}$$

Changes in retention ratio are shown Figure 2.34 based on the Turkish non-life insurance sector averages data sets. The ratio of companies should be in -33% and 33%.

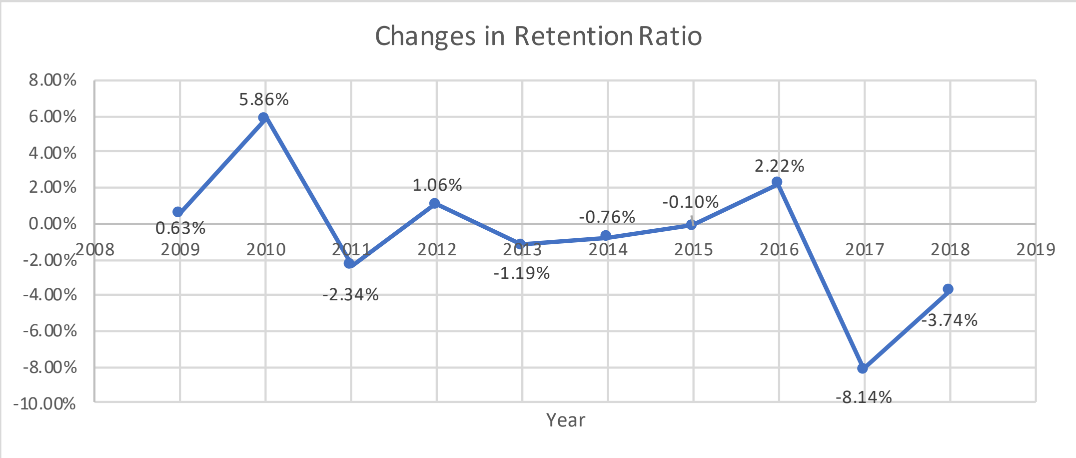


Figure 2.34: Changes in Retention Ratio, Non-Life Insurance Sector, 2009-2018

2.35 Reinsurance Rate of Return

The reinsurance rate of return is calculated by the ratio of the reinsurance received commissions to the profit or loss of the contract with reinsurers. The profit or loss from the reinsurance is calculated by subtracting the transferred realized loss from the premium of reinsurer share. The transferred premium consists of premiums transferred to reinsurer and change in provisions for gross unearned premium of reinsurance share. The transferred incurred loss is obtained by subtracting the net incurred loss from the gross incurred loss. The final formula after simplification is as follows.

$$\text{RRR} = -\frac{(\text{ Reinsurance Commissions })}{\left(\begin{array}{c}\text{Premiums transferred to reinsurer (-)}\\ + \text{ Change in provisions for gross unearned premium}\\ \text{(Share of Reinsurance)}\\ + \text{ Paid losses (Share of Reinsurance)}\\ + \text{ Provision for outstanding losses (Share of Reinsurance)}\\ + \text{ Provision for transferred outstanding losses (Share of}\\ \text{Reinsurance)(-)}\end{array}\right)} \tag{2.35}$$

If this ratio is negative, it can be said that this business has the highest possible rate. If the company's value is positive, the greater the ratio is, the better the outcome will be. The path can be seen in Figure 2.35, this ratio over the years for the non-life insurance sector in Turkey.

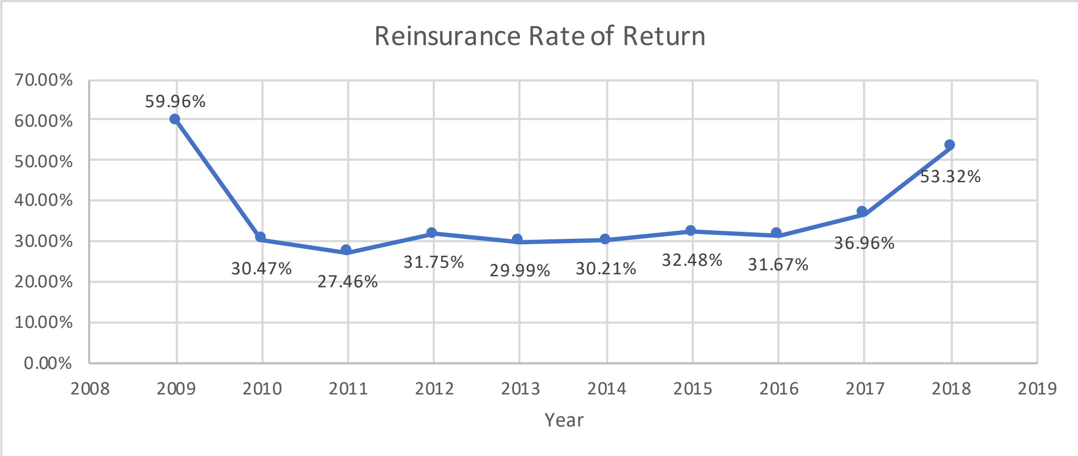


Figure 2.35: Reinsurance Rate of Return, Non-Life Insurance Sector, 2009-2018

2.36 Reinsurance Risk Ratio

Reinsurance risk ratio represents dividing of the earned premium on ceded reinsurance to the gross earned premiums. It represents the premium earned by reinsurance in the 100 units earned premiums.

$$\mathrm{RR} = -\frac{\left(\begin{array}{c}\text{Premiums transferred to reinsurer (-)}\\ \text{+ Change in Provision for unearned premiums (share of reinsurance)}\\ \text{(Groos)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums(Gross)}\\ \text{+ Change in Provision for unearned premiums (Gross)}\end{array}\right)} \tag{2.36}$$

It is expected to be higher than the sector average. In Figure 2.36, non-life insurance sector averages are shown data as per reinsurance risk ratio.

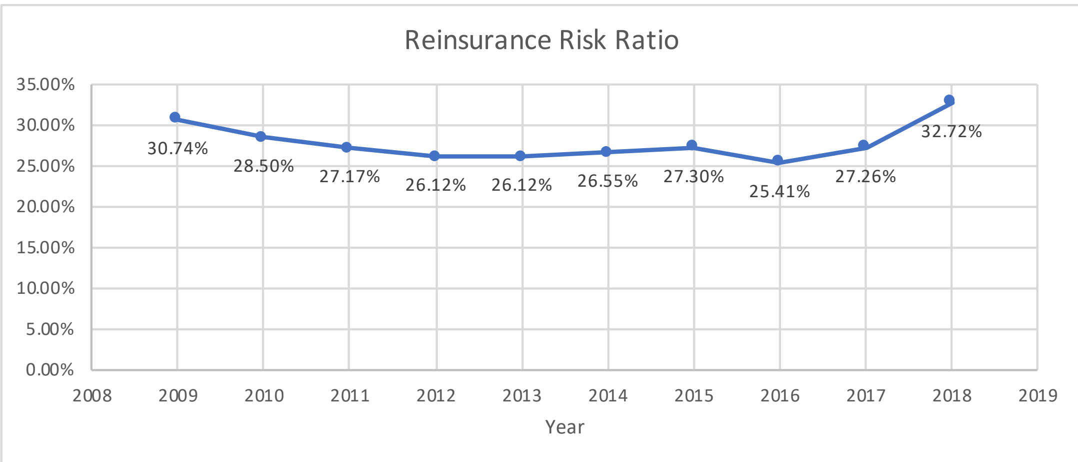


Figure 2.36: Reinsurance Risk Ratio, Non-Life Insurance Sector, 2009-2018

2.37 Change in Required Provision for Outstanding Losses / Available Provision for Outstanding Losses

This ratio also is the change in outstanding claims reserve adequacy ratio. The actuary shall calculate the provision for IBNR claims and the difference in outstanding claims reserve adequacy in the calculations relating to the freshly established divisions during the first five years from the beginning of activities. In addition, the adequacy of the provision for outstanding claims shall be calculated for the eliminated large claims in branches where large claims are eliminated by the actuary due to insufficient number of claim files. These are mentioned at Article 7 of the regulation on Technical provisions of insurance, reinsurance and pension companies, and assets on which such provisions are to be invested [53].

$$\text{RPOL/APOL} = \frac{\left(\begin{array}{c}\text{Required Provision for}\\ \text{Outstanding Losses}\end{array}\right)}{\left(\begin{array}{c}\text{Available Provision for}\\ \text{Outstanding Losses}\end{array}\right)} \tag{2.37}$$

Turkish non-life companies' average ratios are shown in Figure 2.37.

Figure 2.37: Required Provision for Outstanding Losses / Available Provision for Outstanding Losses, Non-Life Insurance Sector, 2011-2018

2.38 Gross Provision for Outstanding Losses / Equity

This percentage is a gross provision for outstanding losses to equity calculation.

$$\text{GPOL/E} = \frac{(\text{Provision for outstanding losses (Gross)})}{(\text{ Equity })} \tag{2.38}$$

This percentage of all non-life insurance companies has acquired the scores in Figure 2.38. According to our information collection, this proportion has risen in 2010, 2012, 2015 and 2018.

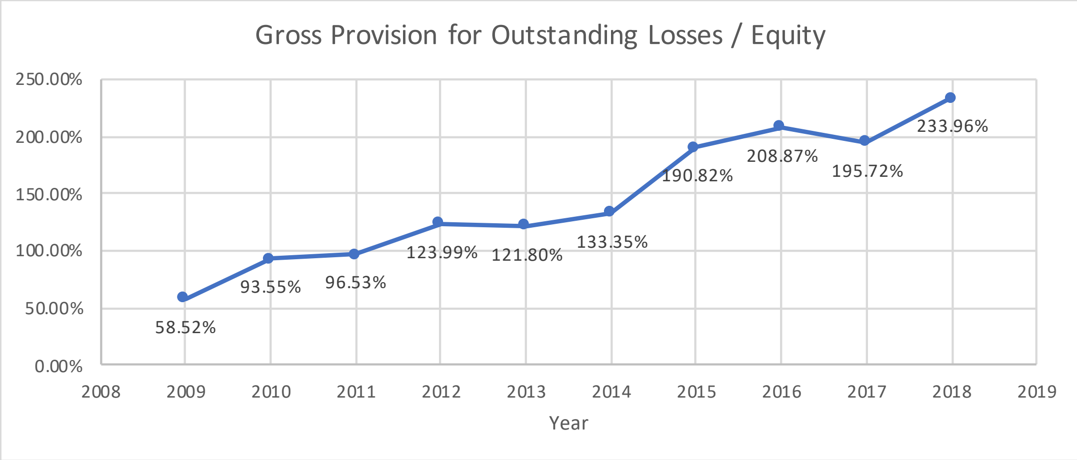


Figure 2.38: Gross Provision for Outstanding Losses / Equity, Non-Life Insurance Sector, 2009-2018

2.39 Net Provision for Outstanding Losses / Equity

This proportion is the calculation of the net reserves for outstanding losses to the equity.

$$\text{NPOL/E} = \frac{(\text{Provision for outstanding losses (Net)})}{(\text{ Equity })} \tag{2.39}$$

This proportion of all non-life insurance companies sets the outcomes in Figure 2.39. This percentage increased in 2010, 2012, 2014 and 2018, according to our data compilation.

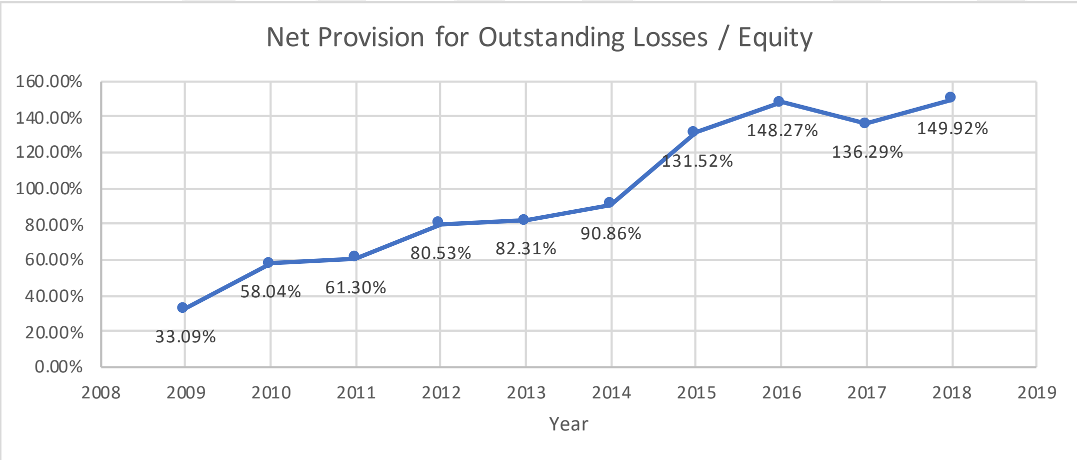


Figure 2.39: Net Provision for Outstanding Losses / Equity, Non-Life Insurance Sector, 2009-2018

2.40 Change in Provision for Outstanding Losses

Outstanding claims reserve consists of accrued but unpaid losses and estimated incurred but not reported claims provisions under this account. In addition, there are provisions for expenses and additional provisions for qualifications in accordance with the rules set by the Undersecretariat [52]. These accounts are calculated based on the statistical knowledge and data of the insurance companies and recorded in the balance sheet. Thanks to this proportion, we can see how many times, according to the prior year's information, the businesses improved or reduced the provision for outstanding losses.

$$\text{CPOL} = \frac{\left(\begin{array}{c}\text{Accrued provision for outstanding losses (Current Year)}\\ + \text{Not reported provision for outstanding losses(Current Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Accrued provision for outstanding losses(Previous year)}\\ + \text{Not reported provision for outstanding losses(Previous year)}\end{array}\right)} \tag{2.40}$$

Figure 2.40 demonstrates this ratio's average industry values over the years. This ratio has approximately doubled in 2012 and 2018, according to Turkish non-life insurance information.

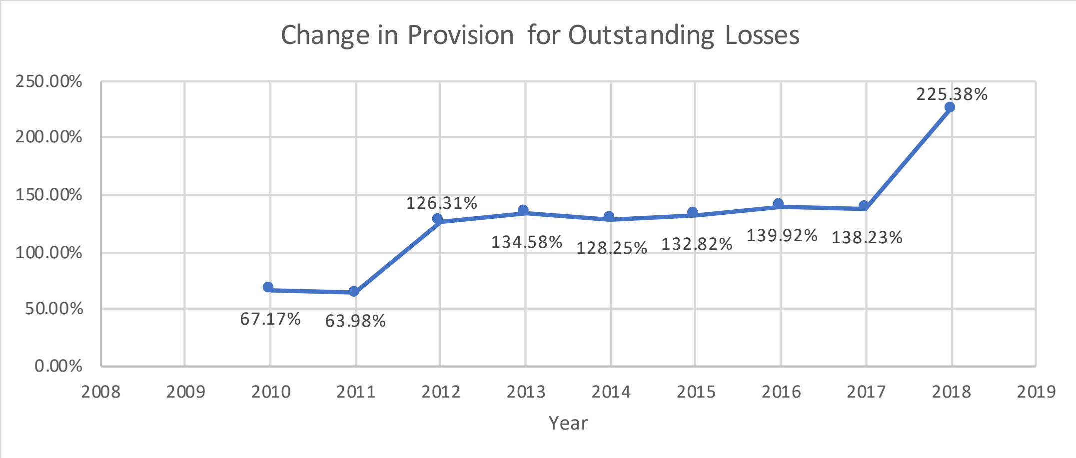


Figure 2.40: Change in Provision for Outstanding Losses, Non-Life Insurance Sector, 2010-2018

2.41 Premium Production Per Personnel

Gross earned premium to number of staffs employed measures the premium production per person for company. Share of reinsurance of written premiums and reinsurance share of all provision's values are ignored because of that reinsurance success or failure is not been taken into consideration. The number of personnel is covered both the central employees of the insurance companies and the marketing staff. A high rate indicates an increase in the effectiveness of staff work.

$$\text{PPPP} = \frac{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ + \text{Provision for unearned premiums (Gross) (-)}\\ + \text{Provision for transferred unearned premiums (Gross)}\\ + \text{Provision for ongoing risks (Gross) (-)}\\ + \text{Provision for transferred ongoing risks (Gross)}\end{array}\right)}{(\text{Number of staffs employed})} \tag{2.41}$$

The values in Figure 2.41 have been obtained by the ratio of gross earned premiums to the number of total personnel of all non-life insurance companies. From 2009 to 2018, this value has always increased for our data sets.

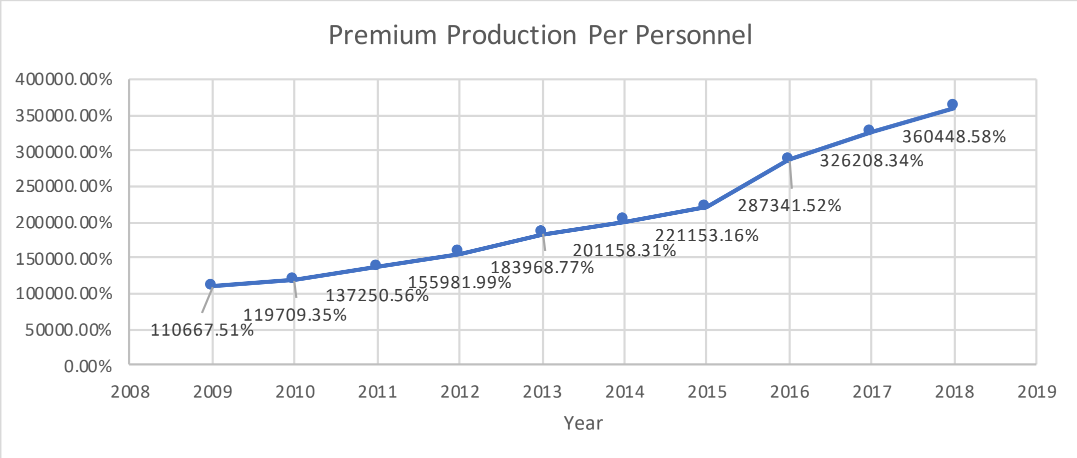


Figure 2.41: Premium Production Per Personnel, Non-Life Insurance Sector, 2009-2018

2.42 Gross Compensation Payment Ratio

Gross paid losses and Change in outstanding losses reserves represents the amount of damage incurred. GCPR calculated dividing of paid losses to incurred claims. This ratio should not be lower than the sector average. If the company has higher ratio, it means that the insurers make payment acceptably to insured and beneficiary of policy. This case may increase confidence in the company and can be seen as a result of operational success.

$$\text{GCPR} = \frac{(\text{Paid losses (Gross) (-)})}{\left(\begin{array}{c}\text{Paid losses (Gross) (-)}\\ \text{+ Change in provision for outstanding losses (Gross)}\end{array}\right)} \tag{2.42}$$

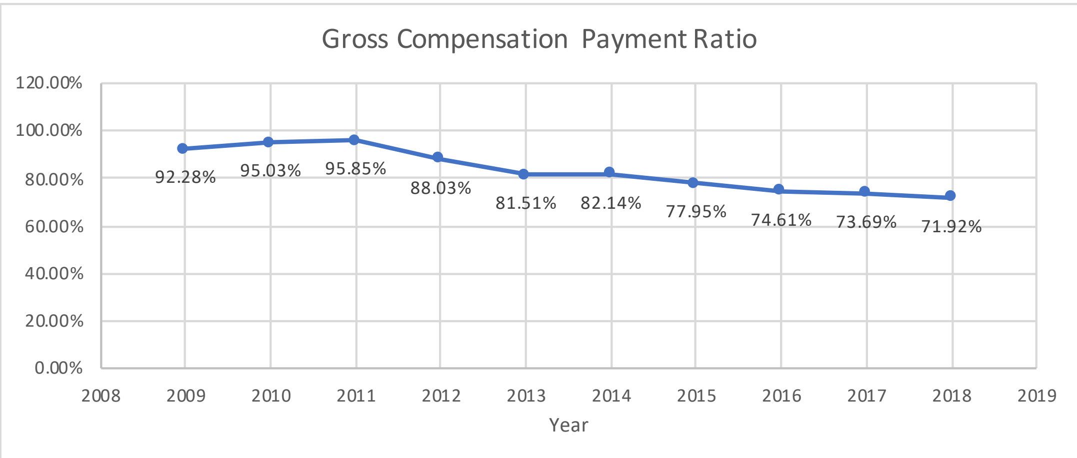


Figure 2.42: Gross Compensation Payment Ratio, Non-Life Insurance Sector, 2009-2018

2.43 Net Compensation Payment Ratio

Net compensation payment ratio is the calculation of the compensation payment ratio with net calculations. In this ratio, reinsurer loss payments are not considered. Only the insurer's loss defrayment is reasoned and it does not matter whether reinsurance payments are made or not.

$$\text{NCPR} = \frac{\big(\text{Paid losses (Net) (-)}\big)}{\left(\begin{array}{c}\text{Paid losses (Net) (-)}\\ + \text{Change in provision for outstanding losses (Net)}\end{array}\right)} \tag{2.43}$$

If this ratio is higher than the sector average, positive opinions about the company's situation may increase. As can be seen in Figure 2.43, the average of Turkish non-life insurance companies has decreased slightly in general.

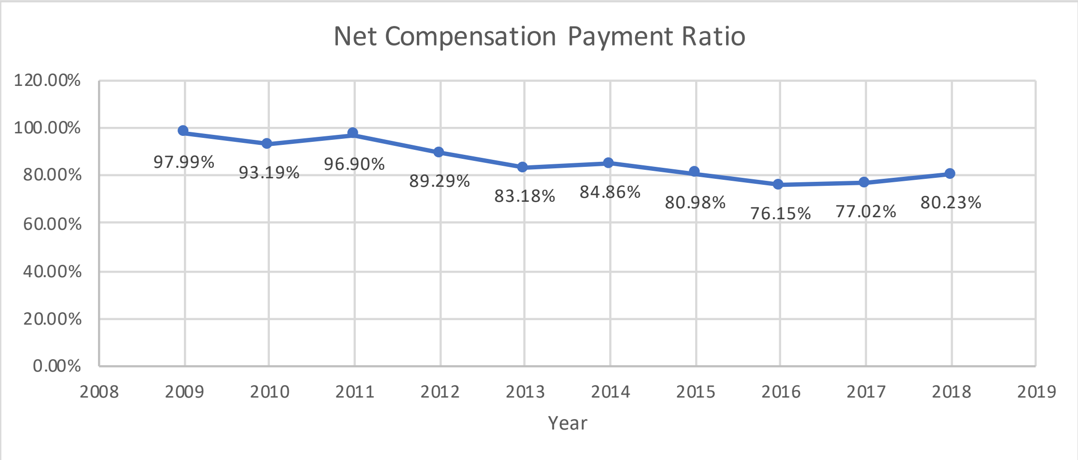


Figure 2.43: Net Compensation Payment Ratio, Non-Life Insurance Sector, 2009-2018

2.44 Gross Paid Losses / Gross Written Premiums

Thanks to this ratio, relations with gross loss and gross premium amount are wanted to be examined. Reinsurance is not taken into consideration, but SSI premiums are subtracted from the written premium. The lower ratio is the better company situation.

$$\text{GPL/GWP} = -\frac{\text{(Paid losses (Gross) (-))}}{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ + \text{Premiums transferred to the SSI (-)}\end{array}\right)} \tag{2.44}$$

Figure 2.44 indicates how many units of damage paid against 100 units of premium for our data sets.

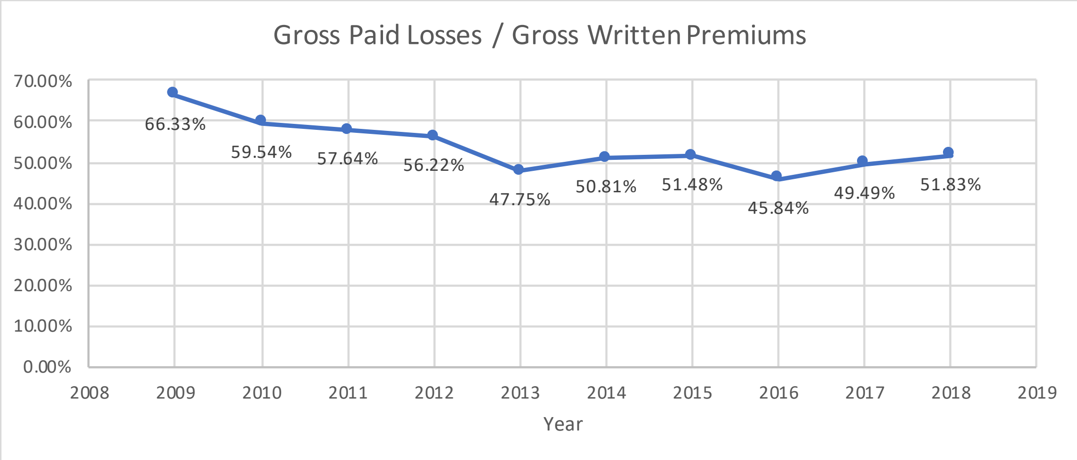


Figure 2.44: Gross Paid Losses / Gross Written Premiums, Non-Life Insurance Sector, 2009-2018

2.45 Change in Gross Provision for Outstanding Losses / Gross Written Premiums

This ratio means that a proportion of gross provision for outstanding losses to gross written premiums. The amount of gross provision for outstanding losses and gross provision for outstanding losses transmitted is the change in gross provision for outstanding losses.

$$\text{CGPOL/GWP} = -\frac{\left(\begin{array}{c}\text{Provision for outstanding losses (Gross) (-)}\\ \text{+ Provision for transferred outstanding losses (Gross)}\end{array}\right)}{\left(\begin{array}{c}\text{Written premiums (Gross)}\\ \text{+ Premiums transferred to the SSI (-)}\end{array}\right)} \tag{2.45}$$

This ratio should be more than the average sector. Non-life business Turkish information is in Figure 2.45.

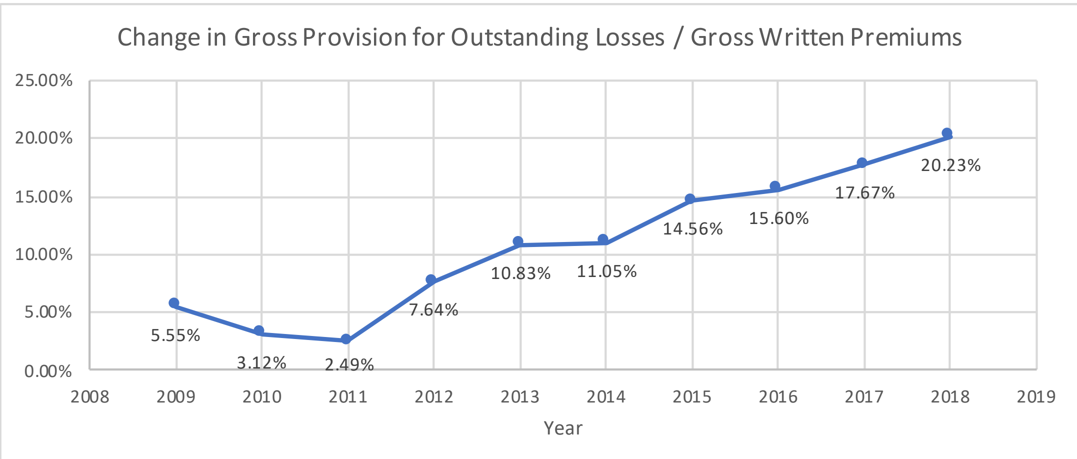


Figure 2.45: Change in Gross Provision for Outstanding Losses / Gross Written Premiums, Non-Life Insurance Sector, 2009-2018

2.46 Share of Reinsurance of Paid Losses / Gross Paid Losses

This ratio indicates how many percentages of the paid loss is transferred to the reinsurer.

$$\text{SRPL/GPL} = -\frac{(\text{Paid losses (Share of reinsurance)})}{(\text{Paid losses (Gross) (-)})} \tag{2.46}$$

It is better to be higher than the sector average. Figure 2.46 shows it.

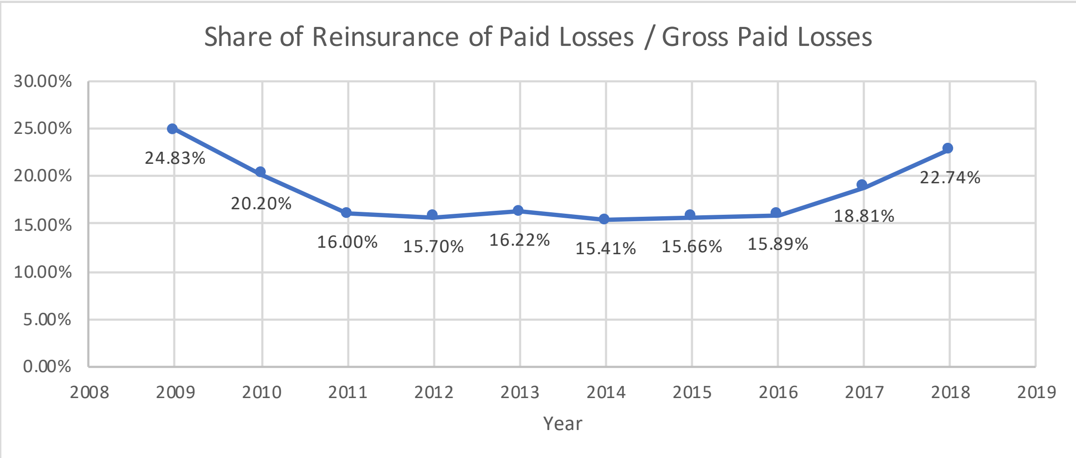


Figure 2.46: Share of Reinsurance of Paid Losses / Gross Paid Losses, Non-Life Insurance Sector, 2009-2018

2.47 Share of Reinsurance / Gross Provision for Outstanding Losses

Explicitly says how much is shared with the reinsurer of the gross outstanding loss provision.

$$\text{SR/GPOL} = -\frac{(\text{Provision for outstanding losses (Share of reinsurance)})}{(\text{Provision for outstanding losses (Gross) (-)})} \tag{2.47}$$

Figure 2.47 shows, according to our data, the average national Turkish non-life insurance industry values. This ratio should be greater than the average industry with respect to our assessment.

2.48 Share of Reinsurance Provision for Transferred Outstanding Losses / Provision for Transferred Outstanding Losses

Specifies how much of the transferred outstanding losses reserve is shared with the reinsurer.

$$\text{SRPTOL/PTOL} = -\frac{\left(\begin{array}{c}\text{Provision for transferred outstanding losses}\\ \text{(Share of Reinsurance) (-)}\end{array}\right)}{\left(\begin{array}{c}\text{Provision for transferred outstanding losses}\\ \text{(Gross)}\end{array}\right)} \tag{2.48}$$

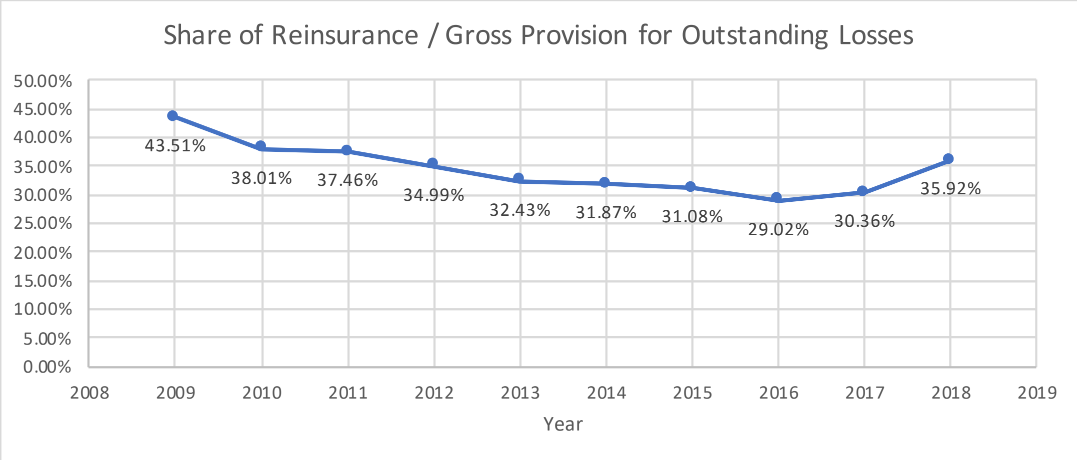


Figure 2.47: Share of Reinsurance / Gross Provision for Outstanding Losses, Non-Life Insurance Sector, 2009-2018

Figure 2.48 demonstrates the Turkish non-life insurance industry's average values according to our information. With regard to our evaluation, this ratio should be higher than the sector average.

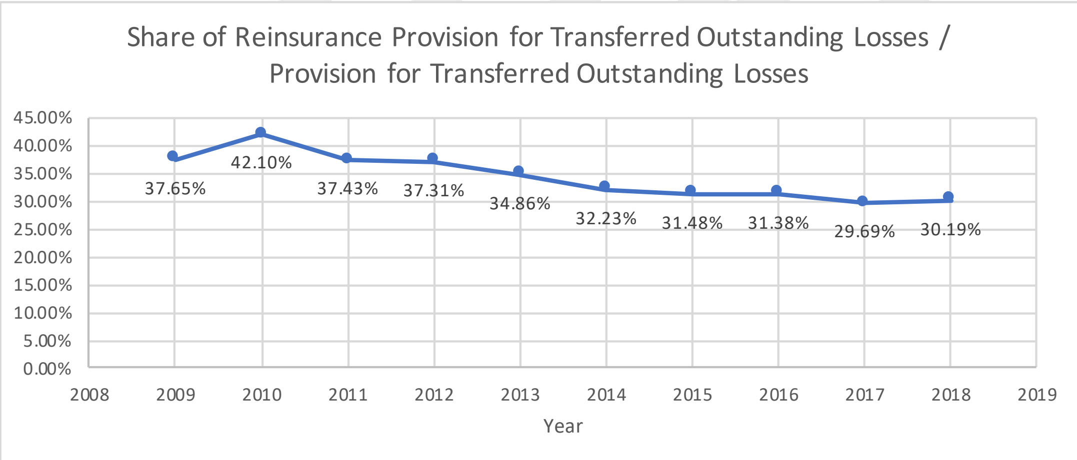


Figure 2.48: Share of Reinsurance Provision for Transferred Outstanding Losses / Provision for Transferred Outstanding Losses, Non-Life Insurance Sector, 2009-2018

2.49 Net Paid Losses / Net Written Premiums

The connection with net paid loss and net premium amount is wanted to be examined with this ratio. Ceded reinsurance and transferred SSI premium are subtracted from gross premium and we obtain the net premium written. The reduced percentage is the better condition of the business. Furthermore, it is essential for decision-makers to analyze the connection between gross and net.

$$\text{NPL/NWP} = -\frac{(\text{Paid losses (Net) (-)})}{(\text{Written premiums (Net)})} \tag{2.49}$$

Turkish non-life companies' average ratios are shown in Figure 2.49.

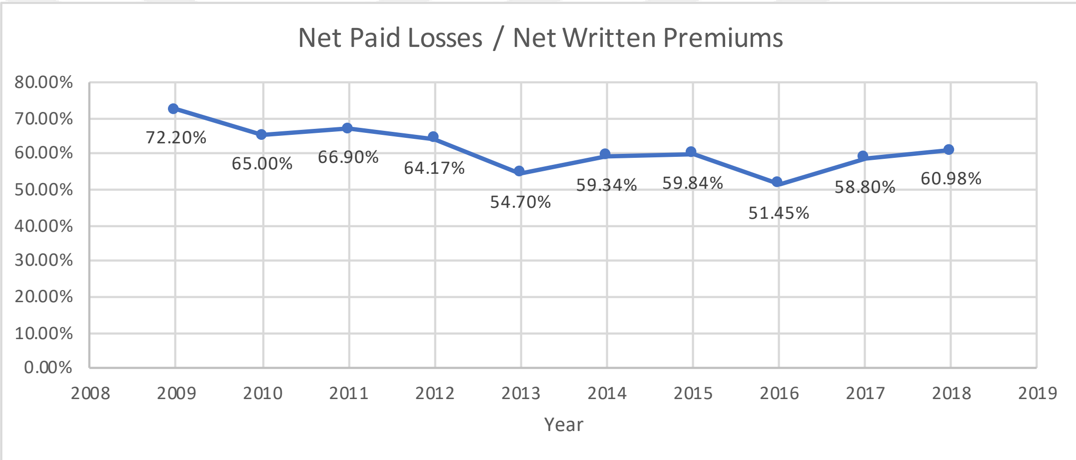


Figure 2.49: Net Paid Losses / Net Written Premiums, Non-Life Insurance Sector, 2009-2018

2.50 Net Outstanding Losses / Net Written Premiums

This ratio calculated the division of net losses reserves to net written premiums.

$$\text{NOL/NWP} = -\frac{(\text{Change in provision for outstanding losses (Net)})}{(\text{Written premiums (Net)})} \tag{2.50}$$

Turkish non-life company industry values are created in Figure 2.50 of this ratio. For all insurers, this ratio should be smaller than the industry scores.

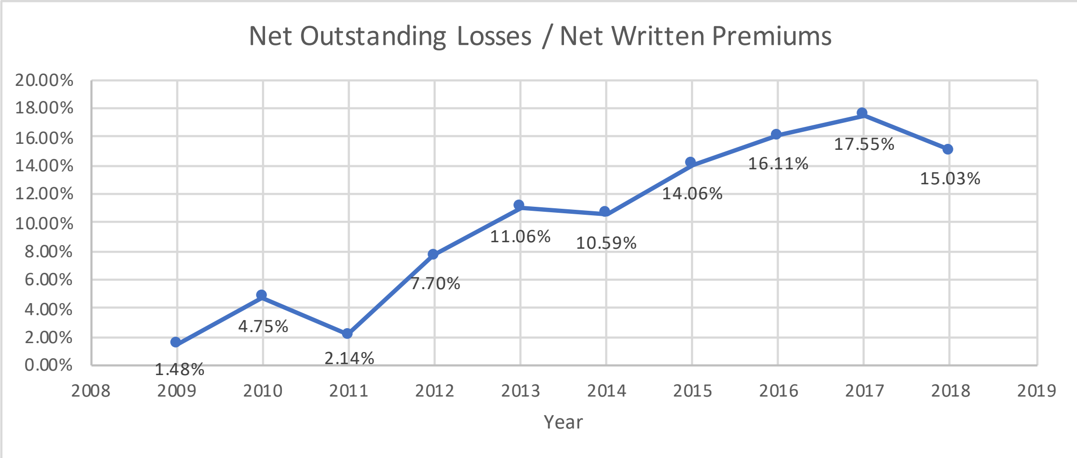


Figure 2.50: Net Outstanding Losses / Net Written Premiums, Non-Life Insurance Sector, 2009-2018

2.51 Operating Expenses / Gross Written Premiums

It is calculated by the ratio of the expenses incurred by insurance companies to their activities and gross written premiums without transferred to the SSI.

$$\text{OE/GWP} = -\frac{(\text{Operating expenses (-)})}{\left(\begin{array}{c}\text{Written premiums (Gross)} \\ + \text{ Premiums transferred to the SSI (-)}\end{array}\right)} \tag{2.51}$$

Company ratio values are required to be lower than the Turkish insurance industry average values in Figure 2.51.

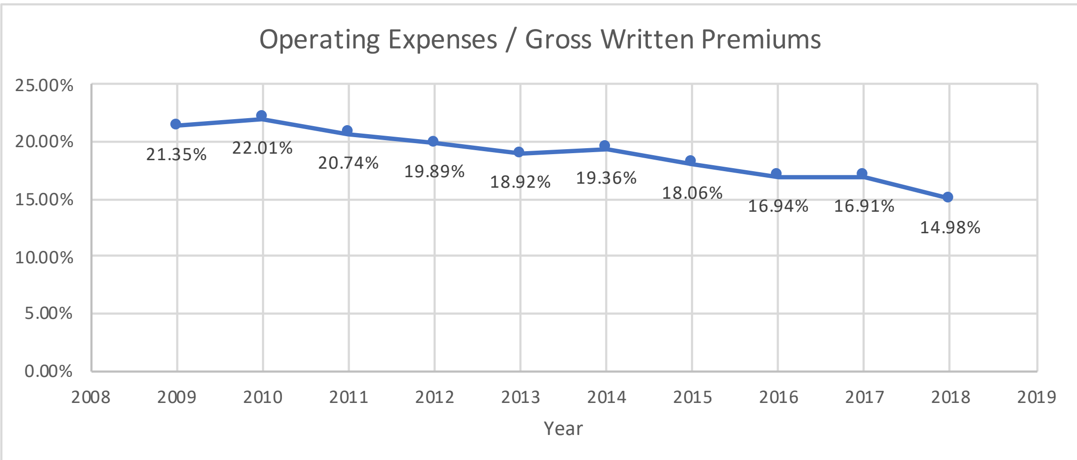


Figure 2.51: Operating Expenses / Gross Written Premiums, Non-Life Insurance Sector, 2009-2018

2.52 Net Commission Ratio

Net commission ratio surveys the obtaining costs of insurance operations. Also, It provides ease for comparison of this commission ratio over against other insurance company and insurance sector. If the ratio is higher, the cost of acquisition may be high or the earned premium may be inadequate [49].

$$\text{NCR} = -\frac{\begin{pmatrix}\text{Production costs (-)}\\ +\text{ Reinsurance commissions}\end{pmatrix}}{\begin{pmatrix}\text{Written Premiums (Net)}\\ +\text{ Change in provision for unearned premiums (Net)}\end{pmatrix}} \tag{2.52}$$

The sector values of the Turkish non-life companies of the net commission ratio in Figure 2.52 are formed. This ratio should be lower than sector averages for all insurance firms.

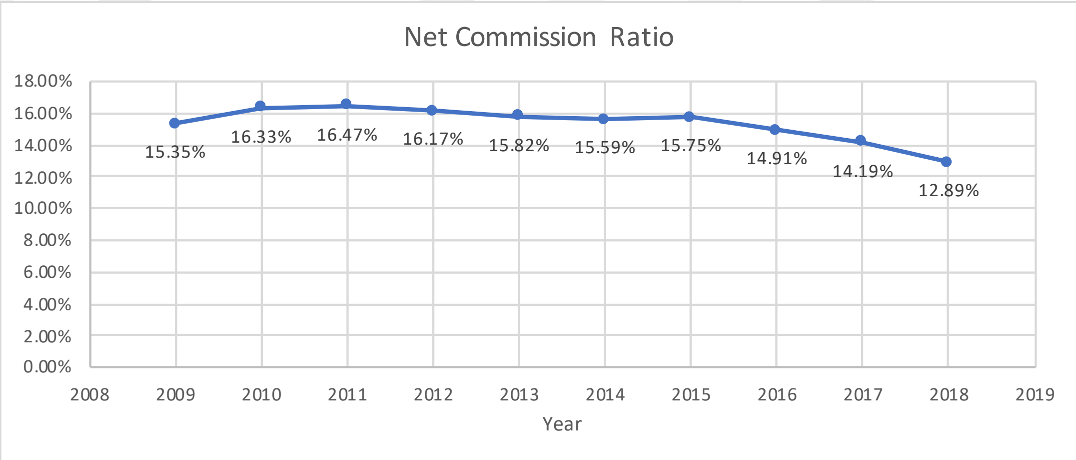


Figure 2.52: Net Commission Ratio, Non-Life Insurance Sector, 2009-2018

2.53 Operating Expenses / Net Earned Premiums

OE/NEP is the ratio of expenses related to the main activity of the enterprise to earned premiums.

$$\text{OE/NEP} = -\frac{\begin{pmatrix}\text{Operating expenses (-)}\\ +\text{ Other technical expenses (-)}\end{pmatrix}}{\begin{pmatrix}\text{Written Premiums (Net)}\\ +\text{ Change in provision for unearned premiums (Net)}\end{pmatrix}} \tag{2.53}$$

The sector values of the Turkish non-life companies of the operating expenses to net earned premiums ratio in Figure 2.53 are formed. This ratio should be lower than the sector averages.

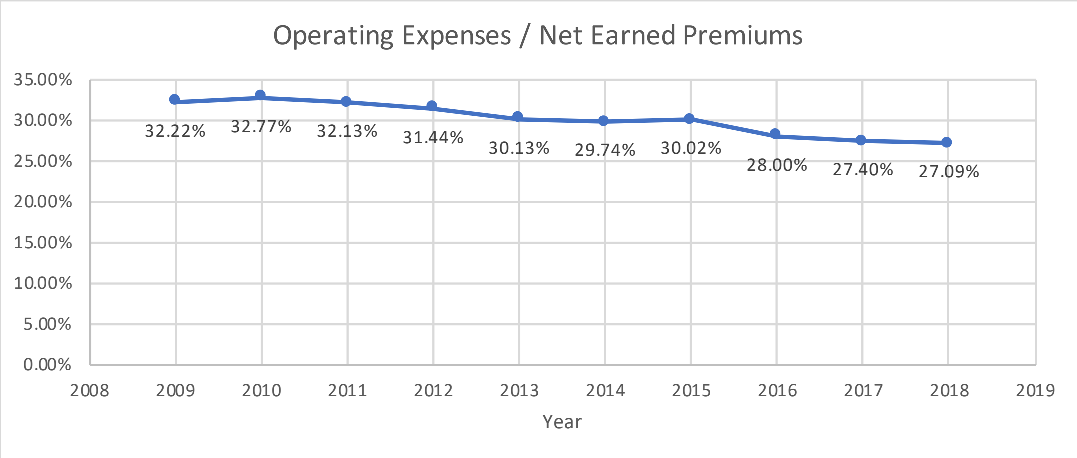


Figure 2.53: Operating Expenses / Net Earned Premiums, Non-Life Insurance Sector, 2009-2018

2.54 Combined Ratio

Combined ratio measures the company's technical profitability based on the ratio of the net incurred losses plus net operating expenses to net earned premiums [38].

$$\text{COMR} = -\frac{(\text{Non-Life technical expense (-)})}{\left(\begin{array}{c}\text{Written premiums (Net)}\\ + \text{Change in provision for unearned premiums (Net)}\\ + \text{Change in provision for ongoing risks (Net)}\end{array}\right)} \tag{2.54}$$

Combined ratio is the sum of loss ratio and expense ratio. This ratio should be lower than 110%. A ratio of over 110% symbolizes a technical loss. Figure 2.54 represents of the non-life insurance of the Turkish sector.

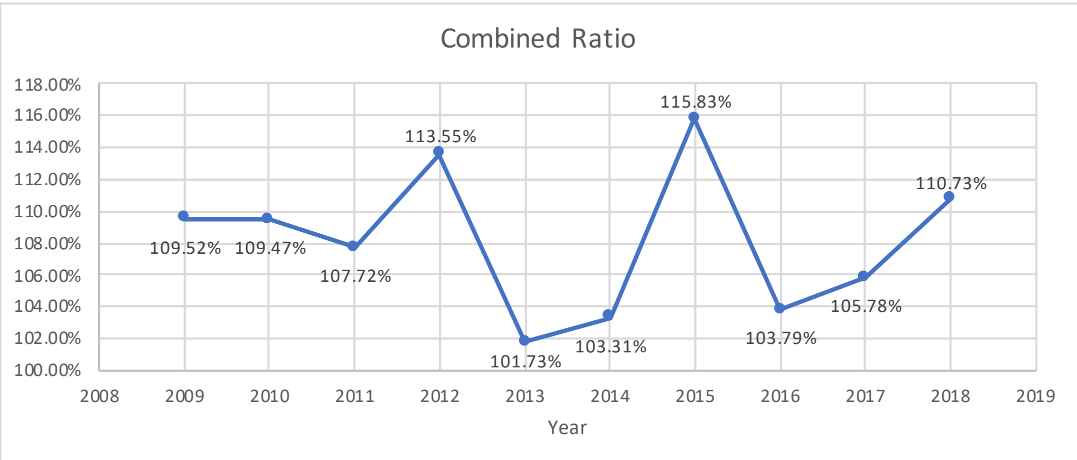


Figure 2.54: Combined Ratio, Non-Life Insurance Sector, 2009-2018

2.55 Market Share

The ratio has been calculated by how much of the total gross premiums generated in the insurance sector are produced by a company. It is the market share of this insurance company. The rate of premium production can also provide information about the size of the company. However, attention should be paid to companies that start operations in the new business or increase premium production for cash inflows. If premium production is consistently high, we can make a better decision about the size of the company.

$$\text{MShr} = \frac{(\text{Insurance company's written premium(Gross)})}{(\text{Insurance sector's written premium(Gross)})} \tag{2.55}$$

Figure 2.55 shows the weight of non-life insurances in the premium production of the Turkish insurance sector. All insurance sector includes non-life, life and pension and reinsurance companies. For example, in 2018, 81.83% of total premium production was generated in the non-life insurance branch. These rates indicate that the industry is predominantly having non-life branches. In our analysis, If this ratio is greater than 2%, the company may belong to the large company category. Because approximately the first fifteen companies have at least 2%, others have below 2% market share.

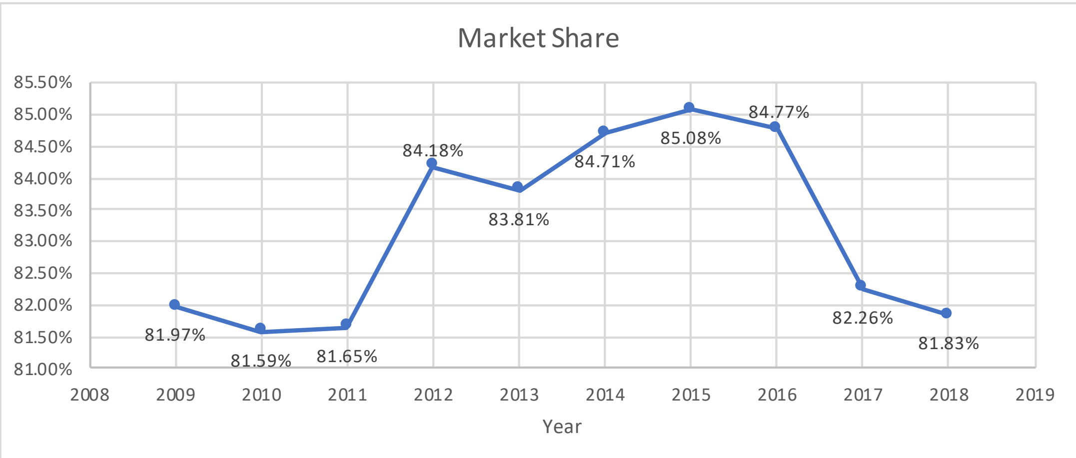


Figure 2.55: Market Share, Non-Life Insurance Sector, 2009-2018

2.56 Non-Life Technical Income / Non-Life Technical Expense

It is investigated to test the relationship between technical income and technical expenses. This ratio is calculated by dividing the income and expenses of the insurance business.

$$\text{TI/TE} = -\frac{(\text{ Technical income of Non-Life branches })}{(\text{Technical expense of Non-Life branches})} \tag{2.56}$$

The graph that is shown in Figure 2.56 is the average value of insurance companies working in the non-life branch. When this ratio is less than 100%, the sector has not revenue. This ratio is requested to be greater than 100%.

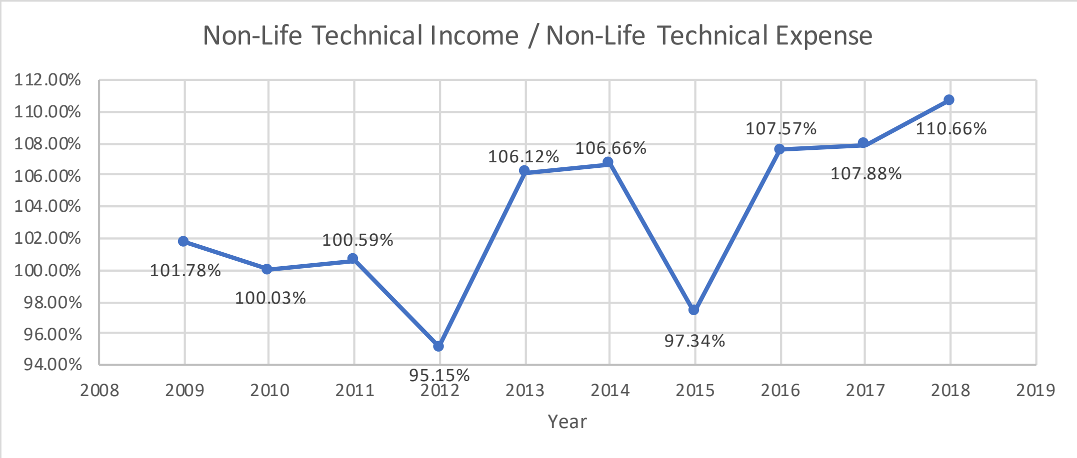


Figure 2.56: Non-Life Technical Income / Non-Life Technical Expense, Non-Life Insurance Sector, 2009-2018

2.57 Financial Leverage Ratios

The ratio of short and long term liabilities to current and non-current assets measures the percentage of funds provided by creditors. It is also known that debt ratio or total debt ratio at some sources [9]. Economy actors prefer low financial leverage ratios. Creditors may think that there may be a disruption in the repayment of loans with companies owned high leverage rates. For these reasons, they may be reluctant to give credit. If this ratio is not above the sector average, we can assume that the probability of bankruptcy increases.

$$\text{FLR} = \frac{(\text{Short term liabilities + Long term liabilities})}{(\text{Current assets + Non-current assets})} \tag{2.57}$$

We can see in Figure 2.57 that average of all Turkish non-life insurance companies have been followed between 52% and 78.5% in the last 10 years. This ratio has been more stable in recent years.

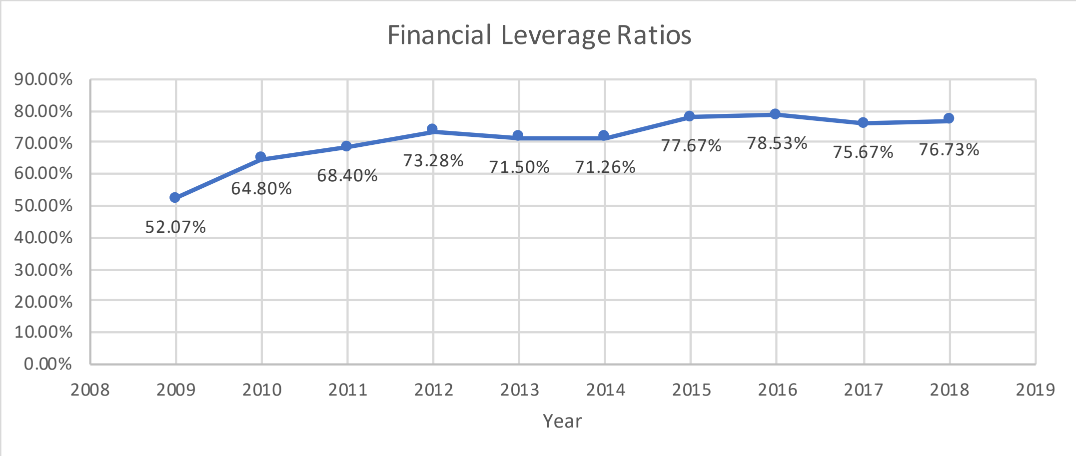


Figure 2.57: Financial Leverage Ratios, Non-Life Insurance Sector, 2009-2018

2.58 Current Ratio

Current ratio is found by dividing current assets by short term liabilities. Indicates the extent to which current liabilities which expected due in one year or less, are covered by those assets expected to be converted to cash in the near future [9]. Current assets are the assets that the entity can use in the short term. Short term liabilities are liabilities with maturities less than one year. Current ratio includes less liquid assets than cash ratio.

$$\text{CURR} = \frac{(\text{ Current assets })}{(\text{ Short term liabilities })} \tag{2.58}$$

According to our data set, as seen in Figure 2.58, current ratio was observed in the range of 119.20% and 167.5%. It can be said that companies' ratio values should be bigger than the sector average in order to interpret the situation.

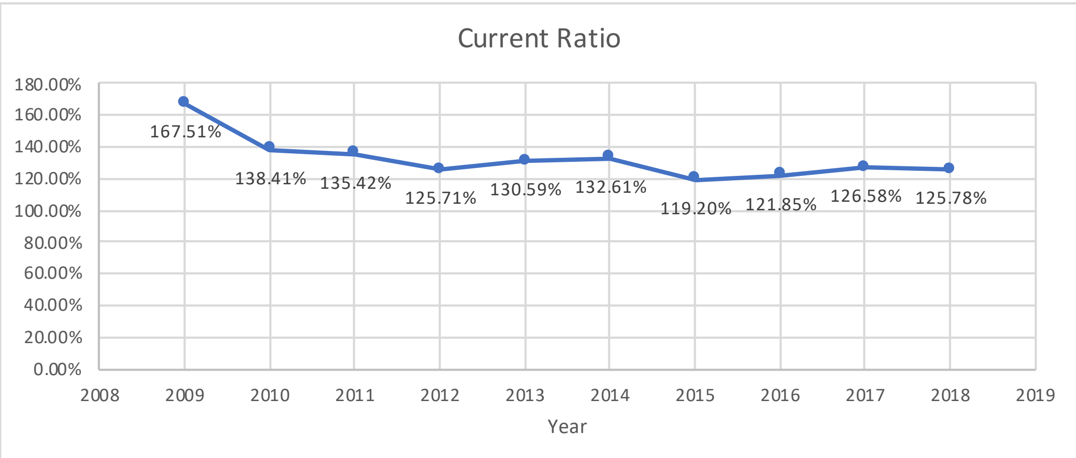


Figure 2.58: Current Ratio, Non-Life Insurance Sector, 2009-2018

2.59 Self-Financing Ratio

Self-financing means that a portion of the profit of the insurance company in the balance sheet year is not subject to distribution and left to the company. In this way, the insurance company can obtain equity. This ratio is the proportional representation of the resources obtained by the insurer through self-financing. Profit reserves have legal, status, extraordinary and other profit reserves, special funds (reserves) and financial asset valuation.

$$\text{SFR} = \frac{(\text{ Profit reserves + Profit/Loss difference of the previous year })}{(\text{ Paid capital })} \tag{2.59}$$

This ratio should be higher than the Turkish non-life insurance industry values that are shown in Figure 2.59.

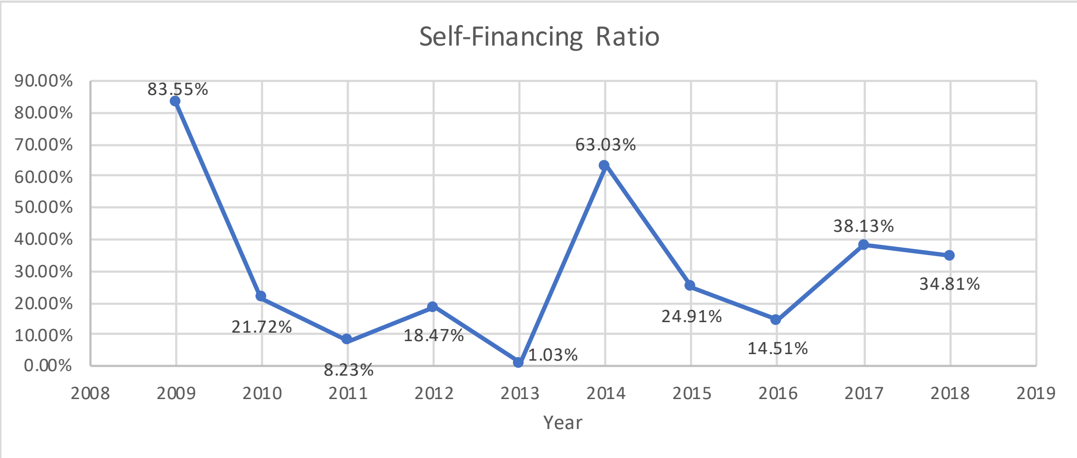


Figure 2.59: Self-Financing Ratio, Non-Life Insurance Sector, 2009-2018

2.60 Tangible Assets / Equity

A tangible asset is an account that is used to monitor the physical assets acquired for use in the activities of insurance companies with an estimated life of more than one year. Tangible assets enclose investment properties, provision for impairment of investment properties, usage properties, machines and equipment, fixtures and installations, motor vehicles, other tangible assets (including special cost prices), tangible assets acquired by leasing, accumulated depreciations and advances are given for tangible assets (including ongoing investments). This ratio calculated as tangible assets to equity.

$$\text{TA/E} = \frac{(\text{Tangible assets})}{(\text{Equity})} \tag{2.60}$$

This ratio is required for companies to be less than the sector average data shown in Figure 2.60.

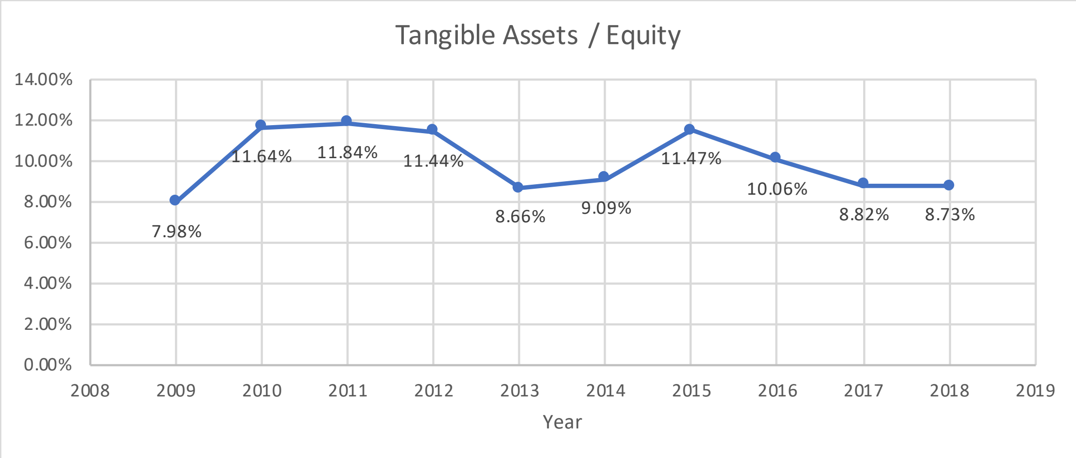


Figure 2.60: Tangible Assets / Equity, Non-Life Insurance Sector, 2009-2018

2.61 Tangible Assets / Total Assets

This ratio is calculated dividing tangible assets to current and noncurrent assets.

$$\text{TA/TA} = \frac{(\text{Tangible assets})}{(\text{Current assets} + \text{Non-current Assets})} \tag{2.61}$$

This ratio ought to be lower than the sector average. A high ratio may cause the company to become in a difficult situation during liquidity needs. Sector averages for the industry are showing in Figure 2.61.

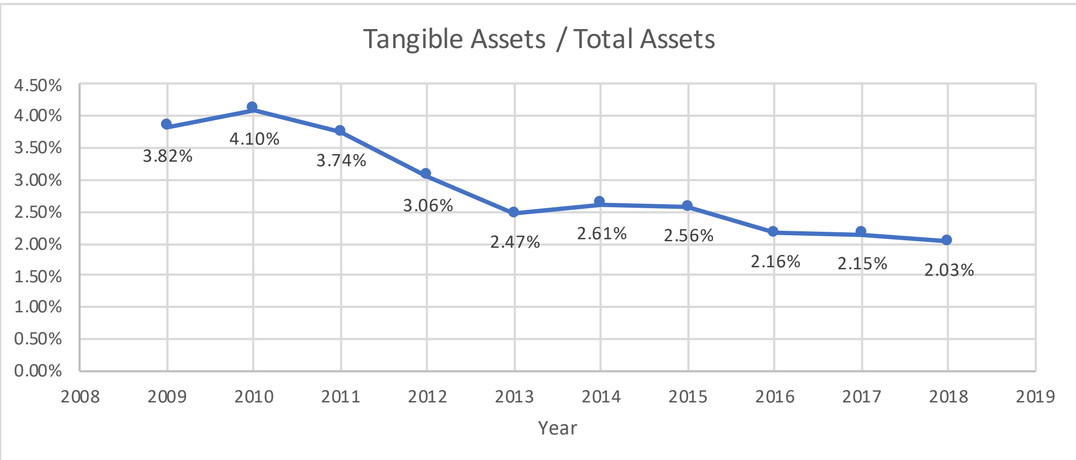


Figure 2.61: Non-Life Insurance Sector, 2009-2018

2.62 Non-Current Assets / Long Term Liabilities and Equity

It shows how much of the non-current assets are financed by the total of long term liabilities and equity.

$$\text{NCA/LE} = \frac{(\text{ Non-current assets })}{(\text{ Long term liabilities + Equity })} \tag{2.62}$$

The shapes that stand out in Figure 2.62, are those of the non-life Turkish insurance sector. The industry average should be smaller than this ratio of companies.

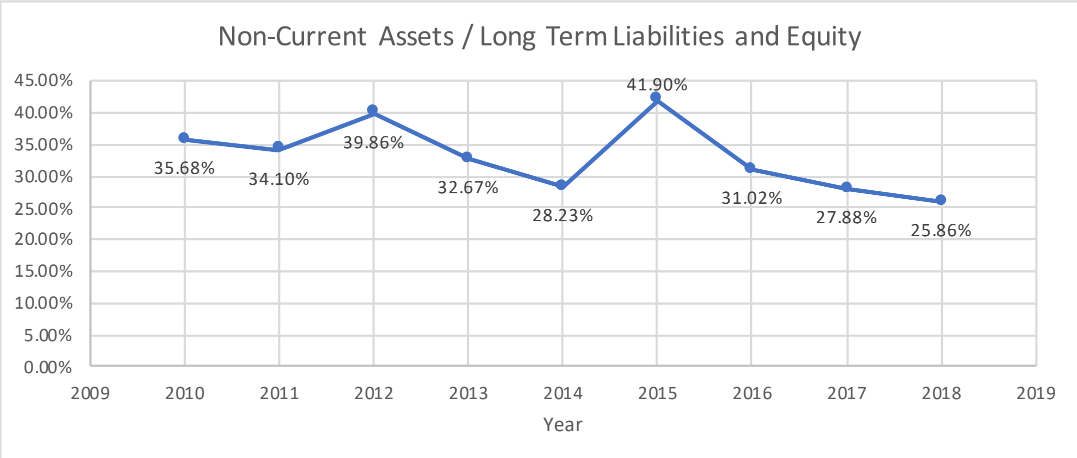


Figure 2.62: Non-Current Assets / Long Term Liabilities and Equity, Non-Life Insurance Sector, 2009-2018

2.63 Receivables Cycle Ratio

This ratio shows the number of times insurance companies collect their premium receivables with the premiums written.

$$\text{RCR} = \frac{(\text{ Written premiums (Gross) })}{(\text{Receivables from operating activities})} \tag{2.63}$$

It is better for the company's situation to have the ratio upstairs the sector average. The insurance average for Turkish industry data is printed on Figure 2.63.

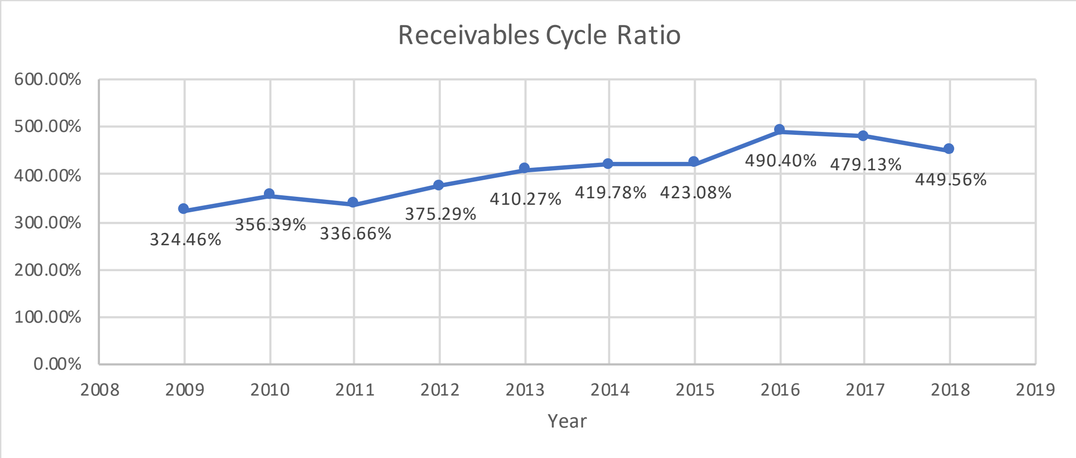


Figure 2.63: Receivables Cycle Ratio, Non-Life Insurance Sector, 2009-2018

2.64 Financial Profitability

It shows the relationship between the insurance company's average equity in the last two years and its post-tax profit. A high level of business profitability indicates a good level.

$$\text{FP} = \frac{\left(\begin{array}{c}\text{Net profit or loss of the period}\\ \text{(Current Year)}\end{array}\right)}{\text{((Equity (Current Year) + Equity (Previous Year)) / 2)}} \tag{2.64}$$

The average values of the Turkish insurance companies are shown in Figure 2.64. In 2010, 2012 and 2015, a negative rate is observed.

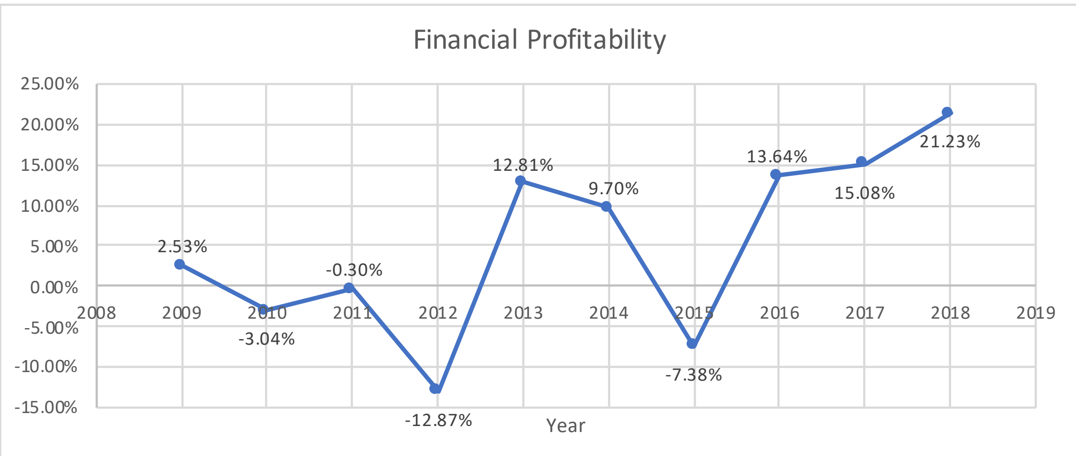


Figure 2.64: Financial Profitability, Non-Life Insurance Sector, 2009-2018

2.65 Economic Profitability

EP ratio is calculated by dividing gross profit of the period to short and long term liabilities. Indicates whether resources are used efficiently. Short and long term liabilities are the total liabilities of the company. The fact that it is higher than the sector average can be a good indicator of the company's situation.

$$\text{EP} = \frac{(\text{Profit or loss of the period})}{(\text{Short term liabilities} + \text{Long term liabilities})} \tag{2.65}$$

Based on the data of the Turkish insurance sector average, this ratio has a pattern shown in Figure 2.65.

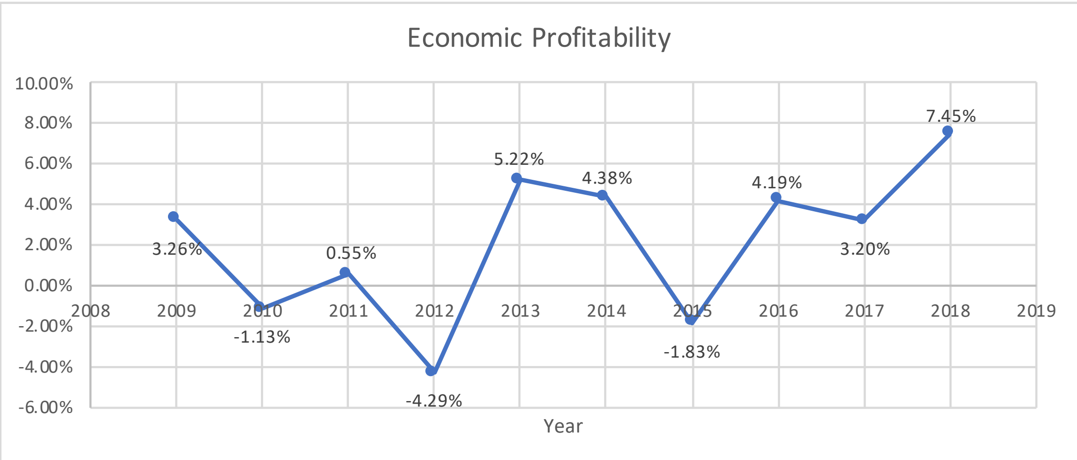


Figure 2.65: Economic Profitability, Non-Life Insurance Sector, 2009-2018

2.66 Return on Equity

The ratio of net income to common equity measures the return on equity. Stockholders invest money and put forward reputation to company and they want to gain a return on their investments. It can be misleading to measure it only by looking at ROE. Comparing with ROA allows us to compare total assets and equity [9]. This ratio should be greater than the sector mean.

$$\text{ROE} = \frac{(\text{ Profit or loss of the period })}{(\text{Equity})} \tag{2.66}$$

Figure 2.66 shows the sector average. In accordance with it, in 2010, 2012 and 2015, the values take negative ratios because, in these years, the Turkish insurance industry records losses.

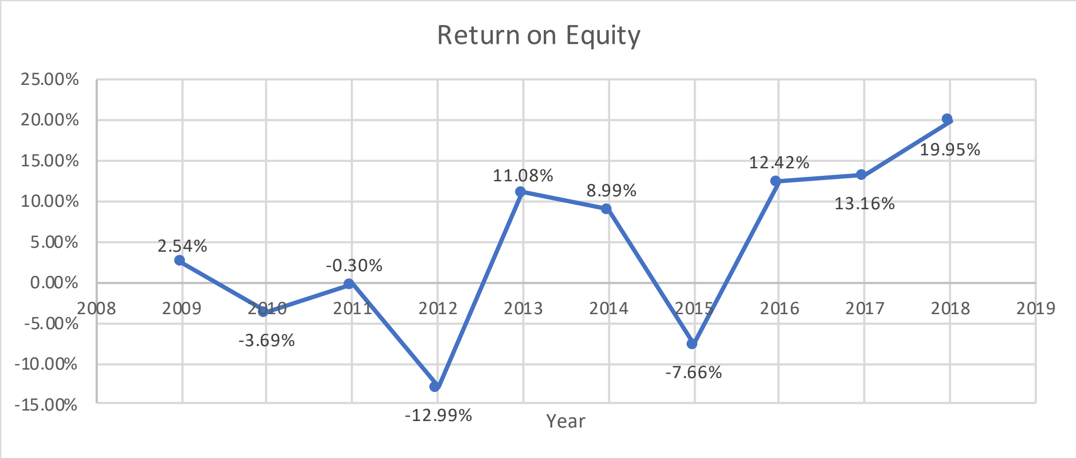


Figure 2.66: Return on Equity, Non-Life Insurance Sector, 2009-2018

2.67 Collection Ratio

The collection ratio rate does not take into account production commission costs, unlike the premium collection ratio. This ratio is the rate at which premium receivables are collected. We can say that companies with higher rates than the sector average make a better premium collection. The sector average of non-life insurance is indicated in Figure 2.67.

$$CR = \frac{\begin{pmatrix} \text{Receivables from insurance activities (Previous Year)} \\ \text{+Written premiums (Gross) (Current Year)} \\ \text{-Receivables from insurance activities (Current Year)} \end{pmatrix}}{(\text{Written premiums (Gross) (Current Year)})} \tag{2.67}$$

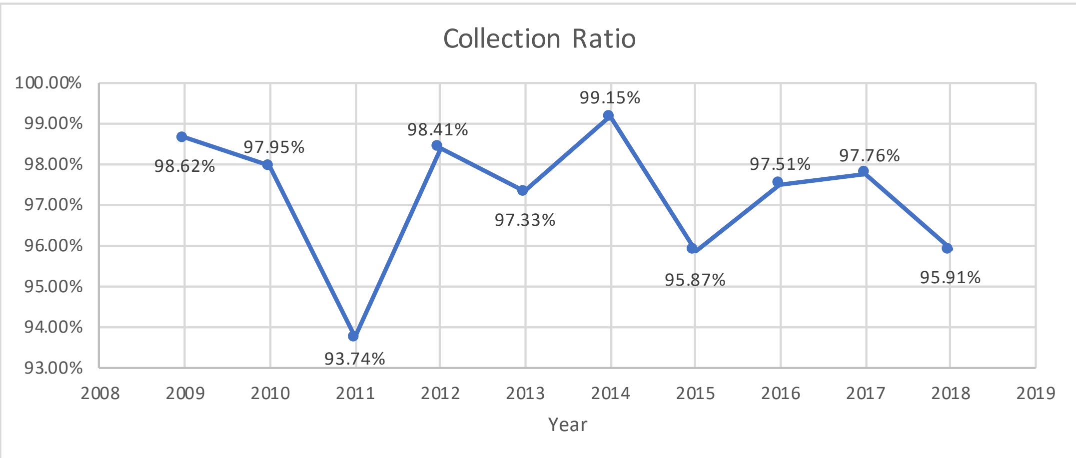


Figure 2.67: Collection Ratio, Non-Life Insurance Sector, 2009-2018

2.68 Current Assets / Total Assets

Current assets consist of the money held by the insurance company in cash, the demand deposits and term deposits up to one year and the assets that are expected to be converted into money within one year. Figure 2.68 shows how this ratio is realized by years. CA/TA ratio of Turkish elementary companies' average was at least 85% and it has generally increased over the years. It is expected to be above the sector average.

$$\text{CA/TA} = \frac{(\text{Current assets})}{(\text{Current assets + Non-current assets})} \tag{2.68}$$

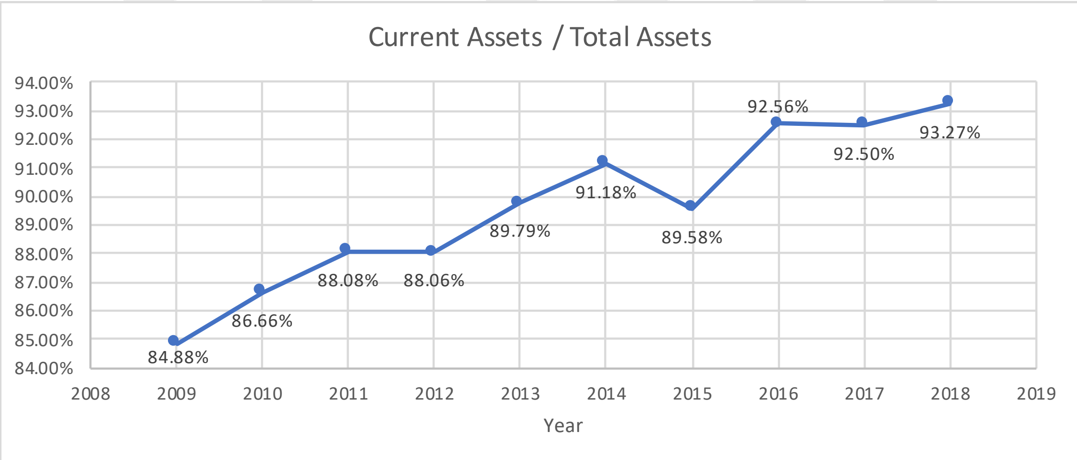


Figure 2.68: Current Assets / Total Assets, Non-Life Insurance Sector, 2009-2018

2.69 Change in Liquid Assets

Most liquid assets of companies are cash and cash equivalents and financial assets and financial investments whose risk is assumed by policyholders. Because of the fact that life insurance business of companies is not considered in our analysis, financial investments whose risk is assumed by life insurance policyholders have been removed. A strong liquid asset serves as an important cushion to meet its obligations in catastrophic damage. The reason for the decrease in liquid assets should be investigated and measures should be taken to prevent the company to become a cash consuming company.

$$\text{CLA} = \frac{\left(\begin{array}{c}\text{Cash and cash equivalents (Current Year)}\\ \text{- Cash and cash equivalents (Previous Year)}\\ \text{+ Financial assets and financial investments whose risk is}\\ \text{assumed by policy holders (Current Year)}\\ \text{- Financial assets and financial investments whose risk is}\\ \text{assumed by policy holders (Previous Year)}\\ \text{- Financial investments whose risk is assumed by life insurance}\\ \text{policy holders (Current Year)}\\ \text{- Financial investments whose risk is assumed by life insurance}\\ \text{policy holders (Previous Year)}\end{array}\right)}{\left(\begin{array}{c}\text{Cash and cash equivalents (Previous year)}\\ \text{+ Financial assets and financial investments whose risk is}\\ \text{assumed by policy holders (Previous year)}\\ \text{- Financial investments whose risk is assumed by life insurance}\\ \text{policy holders (Previous year)}\end{array}\right)} \tag{2.69}$$

Figure 2.69 shows the change of liquid assets for Turkish insurance industry data. A decrease was observed in 2010 and an increase was observed in other years.

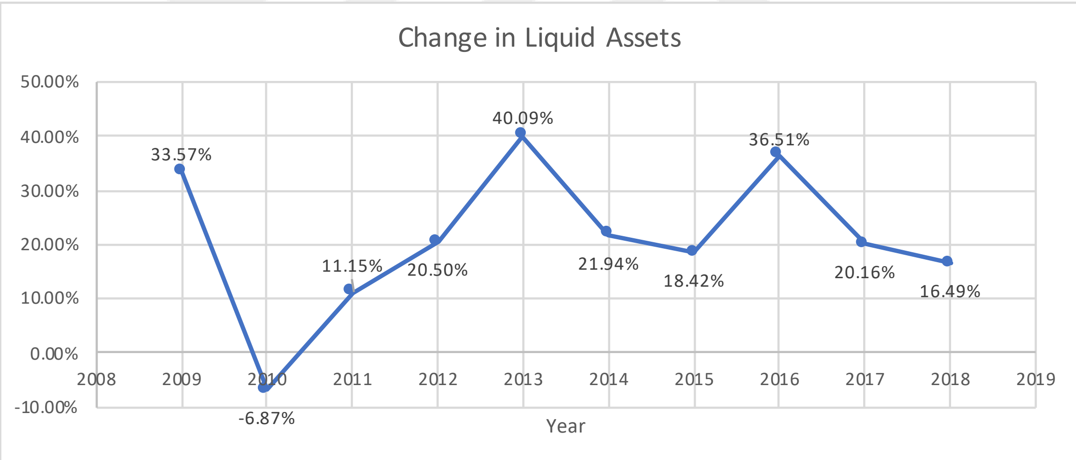


Figure 2.69: Change in Liquid Assets, Non-Life Insurance Sector, 2009-2018

2.70 Market Risk

The risk to a financial portfolio from movements in market worth such as FX rates, interest, and inflation rates and capital and stock prices is defined as market risk [17]. Another definition of it, is the possibility of investors experiencing losses due to factors that affect the overall performance of the financial markets. Another definition of it is the possibility of investors experiencing losses due to factors that influence the entire performance of the financial markets. Market risk is used to measure the amounts of assets that the insurer directs to invest in cash and cash equivalents and investable assets.

$$\text{MR} = \frac{\left(\begin{array}{c}\text{Sellable financial assets}\\ +\text{ Financial assets to be held by the due date}\\ +\text{ Financial assets for purchase and sale}\\ +\text{ Credits}\\ +\text{ Provision for credits (-)}\\ +\text{ Corporate shares}\\ +\text{ Provision for impairment of financial assets (-)}\end{array}\right)}{\left(\begin{array}{c}\text{Cash and cash equivalents}\\ +\text{ Financial assets and financial investments whose}\\ \text{risk is assumed by policy holders}\end{array}\right)} \tag{2.70}$$

Figure 2.70 shows the average values of the Turkish non-life insurance industry according to our data. This ratio should be greater than the industry average with respect to our analysis. The decrease in 2010 is due to approximately 60% reduce in the numerator part of the rate.

Figure 2.70: Market Risk, Non-Life Insurance Sector, 2009-2018

2.71 Classification of Risks

Classification of risks is a grouping of various risks based on projected price or probable effect, probability of occurrence, necessary countermeasures, etc [30]. We created the risk grouping by adding the calculated ratio to the closest class. In other words, a ratio may be associated with more than one risk group but we assume that it belongs to one class. The risk classification of all ratios is shown in 2.1.

(i) Financial Risks

Financial risk can be considered as uncertainty about the valuation of the company's property and the company's future profits or receivables [4]. Seventeen ratios are calculated under the name of financial risks.

(a) Credit Risk

Credit risk is a matter of concern when an organization owes cash or is required to make a payment to it or on its behalf by another organization [29]. Six ratios are available.

(b) Liquidity Risk

Liquidity is merely the convenience of securities trading [3]. The easier an asset can be converted to cash, the more liquid it is. Illiquidity can be a problem when occurred catastrophic risks or emergency cash needs. There are seven ratios about liquidity risk of the insurance company.

(c) Market Risk

Market risk is defined as the change in asset quality and liability intensities resulting from changes in market dynamics. In our classification, there are four ratios under market risk and generally related to investment income.

(ii) Non-life Risks

This main class includes the reinsurance, underwriting, technical provisions, and operational risks to non-life processing of non-life insurance companies related. Non-life Risks contains twenty-six ratios.

(a) Reinsurance Risk

In this risk group, considering the relations of the company with reinsurance, we can see the ratios related to the risk it carries or transfers. In total, ten ratios are under reinsurance risk.

(b) Underwriting Risk

The risk of loss carried by the insurers is an underwriting risk. In insurance, the danger of underwriting may be the result of incorrect risk evaluation, or of uncontrollable variables, in the form of an insurance policy [39]. Underwriting risk has twelve ratios and they are related premiums production, claims, and its consequences.

(c) Technical Provision Risk

Below this risk class, there are four ratios concerning technical provision. All ratios in there are related to outstanding claims reserves.

(iii) Operational Risks

Human mistakes and fraud, processes and procedures, and structures are the results of operational risks. In addition to these, reputational risks are covered in our classification. Besides, there are eight ratios in this main title.

(a) Reputational Risk

The ratios here are essential for the company's reputation, but also measure the size and success of its operations. There are four ratios.

(b) Operational Risk

Expenses, outgoings, and commissions arising from non-life insurance transactions are considered under this risk heading. There are four ratios related to this.

(iv) Profit and Solvency Risks

In this main heading, the company's profitability and capital adequacy have been taken into consideration. There are nineteen ratios related to profit and solvency risk topics.

(a) Profitability Risk

The profitability ratios are intended to assess the company's profit margins. Based on the income generated in the form of dividends, an analysis of profit is essential for stockholders. Creditors also have a keen interest in earnings because earnings are a source of debt coverage resources. Management also utilizes profit as a metric of efficiency [12]. There are ten ratios about profitability risks.

(b) Capital Risk

Ratios related to capital evaluate the capacity of a company to fulfill its debts and obligations and the level of debt financing of the firm's equity. They disclose the equity accessible to absorb any losses [12]. It has nine ratios in the matter of capital risk.

Table 2.1: Risk Classification of the Ratios

Financial Risks			Non-Life Risks			Operational Risks		Profit and Solvency Risks	
Credit Risk	**Liquidity Risk**	**Market Risk**	**Reinsurance Risk**	**Underwriting Risk**	**Technical Provision Risk**	**Reputational Risk**	**Operational Risk**	**Profitability Risk**	**Capital Risk**
CR	ACTP/TP	AFC/LFC	CiRR	CGPOL/GWP	CPOL	PPPP	NCR	EP	CE
DR/TA	CA/TA	IP/TA	NPL/GPL	CGWP	GPOL/E	GCPR	NER	FP	CRR
FLR	CashR	NCA/LE	PRO/E	CNWP	NPOL/E	NCPR	OE/GWP	P/PC	E/TP
NPR/TA	CLA	MR	RR	COMR	RPOL/APOL	MShr	OE/NEP	ROA	GPWE
PCR	CURR		RRR	GLR				ROE	NPWE
RCR	LA/TA		RS/GP	GPL/GWP				ROI	SFR
	TA/TA		SR/GPOL	MPS				TI/TA	TA/E
			SRP/E	NLR				TI/TE	TR/LA
			SRPL/GPL	NLR/GLR				TP/GWP	TR/NP
			SRPTOL/PTOL	NOL/NWP				TP/NWP	
				NPL/NWP					
				PP/C					

CHAPTER 3

MACHINE LEARNING ALGORITHMS

Machine learning is a computer science discipline that uses standard programming methods to study algorithms and techniques and to automate alternatives too complicated issues that are difficult to program. The machine learns from data set. The bigger the dataset, the more correct it becomes. The purpose of a machine learning algorithm is learning a model or set of rules from a labeled dataset. In this way, this algorithm can predict data points which are not in the dataset correctly [51]. In this chapter, some of the machine learning techniques, their features, how they work, and their background describes. Boruta which is a feature selection method, and machine learning techniques which are Random Forest, Neural Networks, Gradient Boosting Machine, and eXtreme Gradient Boosting are mentioned respectively.

3.1 Boruta

Boruta is a feature selection method. With this method, we can decide which of the variables in our data set are important and which are insignificant [41]. Making this decision is an important step for machine learning. Too many variables should be defined to create an appropriate model from data sets, but some of these variables may be irrelevant to the classification or regression. Working with too many variables can lead to disadvantages such as reduced speed algorithms and excessive resource usage. If the number of variables is greater than necessary, the accuracy may be reduced. In order to obtain more practical and accurate results, dimension reduction methods are intensively studied. The algorithm of the Boruta is intended as a wrapper around an algorithm for the classification of Random Forest. In Slavic mythology, Boruta is known as the Forest God. It iteratively removes characteristics that are proven to be less important than random samples by a statistical test. The algorithm uses a wrapper approach built around a random forest classifier [42]. Algorithm is an extension of the idea put forward by Stoppiglia, Dreyfus, Dubois, and Oussar (2003)[60], and determines the relevance of actual properties by comparing the relevance of random probes.

The algorithm of Boruta feature selection follows the following steps:

Create copies of all independent variables. All copy variables are blended the original data, but their values are mixed to remove their relationship to the target variable. This is called permuted copies or shadow features. Random forest classifier is run with the combined data and performs a variable importance measure to evaluate the importance of each variable where the higher means more important. After that, the mean of accuracy loss divided by the standard deviation of accuracy loss (Z_{score}) is calculated. The maximum Z_{score} is found among shadow attributes (MZSA). Variables are labeled as unimportant when they are too low than MZSA. And this variable is permanently removed from the process. Variables are labeled as important when they are higher than MZSA. The random forest is refreshed to a predefined number so that all variables are labeled as important or unimportant [42].

3.2 Random Forest

Random Forest is a decision tree algorithm with relatively few parameters that are effective in a wide range of data sets. Decision trees are one of the most easily understood machine learning approaches. There is a flow chart of decision trees separated from the black box of artificial neural networks and mathematical equations of linear models. The most popular class is chosen after a great number of decision trees are produced. These procedures are referred to as random forests[18].

The algorithm for Random Forest means that more than one model is produced and their outcomes are used together. The objective is to better deal with unexpected circumstances such as avoiding overfitting [18]. They do not overfit due to the law of large numbers [7].

For both classification and regression, the random forest algorithm is as follows:

Each tree in a random forest will learn from a random sample of the information points during training. Draw samples of n_{tree} bootstrap from the initial data. Bootstrapping implies that some samples are used numerous times in a single tree. For each bootstrap sample, develop an uncut classification or regression tree. Instead of choosing the best division between all the unit predictors, sample the predictors by random and select the best division between them. At the moment of the testing, estimates are produced by averaging the results of each choice tree for regression and majority votes for classification. This method of training each individual learner on distinct bootstrapped subsets of the data set and then averaging the estimates is regarded as bagging, short for aggregating bootstrap [40, 43].

Definition 3.1. A classification of random forest is consist of a collection of decision tree-structured classifiers $h(x, \Theta_k), k = 1, ...$ where the Θ_k are random vectors which is independent identically distributed and each tree a vote for the most popular class at input x [7].

The regression method of random forest is created by trees evolving according to a random vector such that numerical values are adopted by the predictor tree $h(x, \Theta)$ rather than class marks. The values for output are quantitative, and the training sets are supposed to be separate of the random vector distribution Y, X. The mean-squared generalization error for any quantitative predictor is $h(x)$.

$$h(x) = E_{X,Y}(Y - h(X))^2 \tag{3.1}$$

The random forest forecaster consists of getting the average over k of all trees $h(x, \Theta_k)$.

When the amount of trees in the forest reaches endless we can say that is the generalization error of the forest.

$$E_{X,Y}(Y - av_k h(X, \Theta_k))^2 \rightarrow PE^*(forest) = E_{X,Y}(Y - E_\Theta h(X, \Theta))^2 \tag{3.2}$$

If we define the average generalization error of only one tree as $PE^*(tree)$

$$PE^*(tree) = E_\Theta E_{X,Y}(Y - h(X, \Theta))^2 \tag{3.3}$$

Assume that $\forall\ \Theta$, $EY = E_X h(X, \Theta)$. Then,

$$PE^*(forest) \leq \bar{\rho} PE^*(tree) \tag{3.4}$$

where $\bar{\rho}$ is the weighted correlation of residues between $Y - h(X, \Theta)$ and $Y - h(X, \Theta')$ that Θ and Θ' are independent[7]. $PE^*(forest)$ equals to

$$E_{X,Y}\left[E_\Theta(Y - h(X, \Theta))\right]^2 = E_\Theta E_{\Theta'} E_{X,Y}(Y - h(X, \Theta))(Y - h(X, \Theta')) \tag{3.5}$$

and, covariance is $E_\Theta E_{\Theta'}(\rho(\Theta, \Theta'))sd(\Theta)sd(\Theta')$, $sd(\Theta) = \sqrt{E_{X,Y}(Y - h(X, \Theta))^2}$ and weighted correlation can be written as:

$$\bar{\rho} = \frac{E_\Theta E_{\Theta'} \rho(\Theta, \Theta') sd(\Theta) sd(\Theta')}{(E_\Theta sd(\Theta))^2} \tag{3.6}$$

3.2.1 Grid Search for Random Forest

The parameters of all learning algorithms often influence the efficiency of a model. However, it can be complicated to interact with the parameters. The hard work route is to attempt and assess a model, then play with any of the parameters and repeat it. This can be the most effective route if you are lucky or have patience. A more systematic approach would be to create recursive circuits for all the attributes you believe might be essential for each parameter. The method of hyperparameter tuning is to attempt to enhance the default configurations.

Grids provide the answer to this challenge, which is presently in the form of H2O implementations in two types. These are the Cartesian and the RandomDiscrete. The former is comprehensive, and the latter is random mode. Cartesian method attempts all combinations that you specify in the code. It provides the most optimal result by comparing the results of the models. RandomDiscrete mode is used when you need to specify too many hyperparameters to try out all combinations exhaustively. This creates a random combination and controls when it should stop with some additional parameters[18].

3.3 Neural Networks (Deep Learning)

The fresh, trendy word for neural networks is deep learning. It is trendy because it is currently the source of some of the most impressive developments in machine learning. Neural Networks algorithms are often the best performers for problems that people experience easily. However, there are some disadvantages, such as being slow, getting black boxes, not being able to their thoughts and deal with the categorical data inputs.

A neuron is a function that takes multiple numerical inputs and delivers one numerical output. The neurons have been organized into layers, and the outputs of every neuron in one layer become inputs in the next layer for each neuron [18].

The first layer is a training sample or test sample out data. The last layer is our outputs. If it is a regression with learning a single value, then the output layer will have one neuron. If the classification is, then for each possible answer the output layer will have one neuron, and the highest probability is chosen for our input set. There are the hidden layers between the input layer and the output layer.

Each neuron in each hidden layer carries weight for every input and how the network learns to change these weights. Every neuron also has a "bias" input, which can be understood as a weight linked to a permanent input and also adjusted during training. This process is started to the first training sample with random weights then is calculated the error. After that, go back and tweak each of the weights in order to obtain a bit less error. Then you receive the other training sample and recalculate. The processing of each part of the training set is called epoch, and you can specify the number during the algorithm code arranging. The shape of the network is set by using the number of layers, the number of neurons in each of those layers and the number of epochs [18]. More particular, the goal is to minimize the loss function for every example of training j [13].

3.3.1 Loss Functions

The system principle is focused on whether regression or classification is being performed.

$$\text{Loss function} = L(W, B|j) \tag{3.7}$$

where $i \in [1, N-1]$, $W_i \subset W$ denotes the weight matrix connecting layers i and $i+1$ for a network of the number of N layers. Likewise, $i \in [1, N-1]$, $b_i \subset B$ indicates the column vector of biases for layer $i+1$ [13].

3.3.1.1 Mean Squared Error

The Mean Squared Error evaluates a predictor's performance for regression. y_j is actual observed output data of j_{th}, and $\hat{y}_j$ represent the predicted output value.

$$L(W, B|j)_{MSE} = \frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2 \tag{3.8}$$

3.3.1.2 Root Mean Squared Error

The fundamental hypothesis in displaying the Root Mean Squared Error is that the mistakes are unbiased and adopt a normal distribution[14].

$$L(W, B|j)_{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}{n}} \tag{3.9}$$

3.3.1.3 Mean Absolute Error

The Mean Absolute Error is appropriate for describing uniformly distributed errors[14].

$$L(W, B|j)_{MAE} = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \tag{3.10}$$

3.3.1.4 Cross Entropy

A typical use of the cross entropy is classification[13].

$$L(W, B|j)_{CE} = -\sum_{o \in O}\left(ln(y_o^j)(\hat{y}_o^j) + ln(1 - y_o^j)(1 - \hat{y}_o^j)\right) \tag{3.11}$$

where o represents the output data in the output layer of O.

3.3.2 Activation Functions

There are some choices for the nonlinear activation function of g. x_i represents the neuron's input values and w_i represents their weights. α denotes the weighted combination and the bias b represents the activation threshold of the neuron.

$$\alpha = \sum_i wixi + b \tag{3.12}$$

3.3.2.1 Rectifier

Rectifier function is default and the most common activation function. It produces the amount of its weighted inputs, but it draws all adverse scores to zero.

$$g(\alpha) = max(0, \alpha) \tag{3.13}$$

where $g(\alpha) \in \mathbb{R}_+$.

3.3.2.2 Tanh

It is short for hyperbolic tangent. According to tanh function, the input range is $-\infty$ to $+\infty$ and transforms that to the output range of -1 to $+1$.

$$g(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} - e^{-\alpha}} \tag{3.14}$$

where $g(\alpha) \in [-1, 1]$.

3.3.2.3 Maxout

Maxout function merely produces the largest inputs, meaning weighted inputs are used immediately and not summed up [13].

$$g(\alpha) = max(x_1, x_2) \tag{3.15}$$

where $g(\alpha) \in \mathbb{R}$.

H2O supports six possible values for the activation parameter. The three are above, others are RectifierWithDropout, MaxoutWithDropout, and TanhWithDropout. They enable the use of hidden dropout ratios to control the frequency at which outputs are randomly put to zero [18].

Finally, $g\Big(\sum_i wixi + b\Big)$ gives output data.

3.3.3 Grid Search for Deep Learning

H2O promotes model tuning in grid search by enabling customers to define value sets for parameter statements and monitor adjustments in model behavior[13].

3.4 Gradient Boosting Machines

Another decision tree algorithm, like random forest, is the gradient boosting machines. It is an ensemble technique of regression and classification tree models: we are trying to make more than one tree, then combine their outputs. There are forward learning techniques to achieve predictive outcomes using gradually enhanced estimates. In addition, the main idea is boosting. Boosting, which is a flexible nonlinear regression procedure, is used because of the training data that is hard to learn and helps enhance the accuracy of trees [18, 45].

Gradient boosting machines is a method of machine learning that incorporates two strong instruments that is gradient-based optimization and boosting. Optimization based on gradient utilizes gradient computations to minimize the loss function of a model in terms of training data. Additively boosting gathers an ensemble of weak systems to generate a solid learning scheme for predictive assignments.

It has more parameters and needs a little more effort to tune compared to random forest, but it may achieve somewhat stronger outcomes. It has been observed that it gives better results especially in studies on regression. The primary risk is that if it is continued to give it more and more trees, it can easily overfit [18].

GBM for K-class classification algorithm starts the observation weights $w_i = 1/N$, $i = 1, 2, \ldots, N$. The sample weights are separately altered for each consecutive iteration $m = 2, 3, \ldots, M$ and the classification algorithm is refilled to the weighted results. Use weights w_i to fit a classifier $G_m(x)$ to the training data.

$$G(x) = sign\Big(\sum_{m=1}^{M} \alpha_m G_m(x)\Big) \tag{3.16}$$

where $G(x)$ is final classifier. To obtain this calculate with

$$err_m = \frac{\sum_{i=1}^{N} w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^{N} w_i} \tag{3.17}$$

and compute

$$\alpha_m = \log\Big(\frac{1 - err_m}{err_m}\Big) \tag{3.18}$$

set $w_i.exp[\alpha_m.I(y_i \neq G_m(x))] \rightarrow w_i$ where $i = 1, 2, \ldots, N$ [25].

GBM setting for the regression, the connection between $L(y, f(x)) = (y - f(x))^2$ squared-error loss and $L(y, f(x)) = |y - f(x)|$ absolute loss is similar to the connection between exponential loss and binomial log-likelihood where y is actual data and x is predicted values. The population solutions are for both $f(x) = E(Y|x)$ for squared-error loss, and $f(x) = median(Y|x)$ for absolute loss. Thus, these are the same for symmetric error distributions. One of the loss functions is Huber loss for regression and calculation is[25]

$$L(y, f(x)) = \begin{cases} (y - f(x))^2 & \text{for } |y - f(x)| \leq \delta, \\ 2\delta|y - f(x)| - \delta^2 & \text{otherwise.} \end{cases} \tag{3.19}$$

3.5 Extreme Gradient Boosting (XGBoost)

XGBoost is the abbreviation of eXtreme Gradient Boosting. It involves an effective linear model solver and an algorithm for decision tree learning and supports different objective functions, such as regression, ranking, and classification [16].

Extreme gradient boosting brings parallel tree boosting that quickly and accurately resolves many data scientific issues. It is one of today's finest gradient boosting frames for many issues [27].

XGBoost's scalability in all scenarios is the most significant consideration behind achievement. The scalability is due to various major technologies and algorithmic optimizations. These developments include a new algorithm for tree learning and parallel and distributed computing [15].

In the XGboost algorithm, an objective function is defined with two components instead of optimizing a simple squared error loss; a loss function is defined over the training data and a regularization term is defined that penalizes the model's complexity [47]:

$$\mathcal{L}(\phi) = \sum_i L(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{3.20}$$

$L(y_i, \hat{y}_i)$ may be any differentiating convex loss function that estimates the predictive distinction with the real label for a particular training data.

$$\Omega(f_k) = \gamma T + 0.5\lambda w^2 \tag{3.21}$$

where $\Omega(f_k)$ dexcribes the f_k tree complexity. The number of f_k tree leaves is T and the leaf weight is w. If $\Omega(f_k)$ is included in the objective function, we must optimize a smaller tree, which minimizes $L(y_i, \hat{y}_i)$ at the same time. This contributes to reducing overfitting. γT normally provide for each extra leaf of tree a steady charge and λw^2 will penalize extreme weights. Parameters for user configuration are γ and λ.

Boosting proceeds is iterated and newest objective function is:

$$\mathcal{L}^t = \sum_i L(y_i, \hat{y}_i{}^{(t-1)} + f_k(x_i)) + \sum_k \Omega(f_k) \tag{3.22}$$

and minimal objective function will be found [47].

CHAPTER 4

DATA DESCRIPTIONS AND ANALYSES

In the ratio analysis, the ratios are calculated using the company based financial and technical statements published by the IAT and the tables section of the Insurance and individual pension system financials activity report published by TRMTF. ACTP/TP, CRR, AFC/LFC, and RPOL/APOL ratios of 2011-2018 are obtained from TRMTF. All the other ratios are calculated based on non-life insurance companies for the period between 2009 and 2018. However, in order to have data integrity, our analysis covers only the data of Turkish non-life insurance companies that are active between 2011 and 2018.

In our analysis, 70 ratios, which are stated in Chapter 2, are used. Among these ratios, the capital requirements ratio in Section 2.27 is selected as the dependent variable because of that TRMTF, General Directorate of Insurance (GDI) considers this ratio to take action on insurance companies. The remaining 69 ratios are included in the analysis as independent variables. The reason for considering only non-life insurance companies in our analysis is that they generally provide annual premiums and collaterals and have shorter policy periods compared to life insurances. Our entire data set consists of 301 observations and 70 ratios. Companies that do not voluntarily produce premiums, whose premium production is stopped by the insurance authorities, and most of ratios which could not be calculated are not used in this analysis. Also, all analyses in this section are performed using the 'R' program.

The high correlation between some of the ratios used has led us to machine learning methods. Because of that multivariate or linear regression models are adversely affected by this dependency and machine learning methods such as Neural Networks and Random Forest can cope with this high positive or negative relationship. Autocorrelation control is not performed because the time effect at the ratios is not considered. Random Forest, Neural Network (Deep Learning), Gradient Boosting Machine, and eXtreme Gradient Boosting described in Chapter 3 are used as analysis methods. In the first stage, the most appropriate model is established by using grid methods of all these methods. Afterward, the analyses are run with the parameters of the best model. At the same time, 22 out of 69 ratios in data set used in this study are confirmed with the Boruta feature selection method and the new data set is obtained as 301 observations and 22 ratios. Another feature selection method such as principal component analysis or factor analysis is not implemented as Boruta stems from a decision tree approach which copes with the methods selected in this study. The 2011-2017 is the train data set and

2018 is the test data set. Train data set consists of 263 observations and test data set consists of 38 observations. Our aim in all analyses is to estimate the 2018 values by training the values from 2011 to 2017.

4.1 Boruta with Data Sets

The results and the table of the Boruta feature selection made with the train data set (263x70) are as follows. All "X" ratios are set according to the section number of the ratios in Chapter 2. In other words, X1 refers to the ratio in Section 2.1.

Table 4.1: The R Output of Boruta

Boruta performed 99 iterations in 18.77481 secs.
Attributes confirmed as important: X1, X2, X5, X7, X15, X16, X20, X21, X25, X26, X28, X29, X38, X39, X44, X51, X57, X58, X60, X62, X64, X66
Attributes confirmed as unimportant: X6, X8, X9, X14, X17, X18, X19, X23, X24, X30, X31, X32, X33, X34, X35, X37, X40, X42, X43, X45, X47, X48, X50, X52, X54, X59, X61, X67, X68, X69, X70
Tentative attributes left: X3, X4, X10, X11, X12, X13, X22, X36, X41, X46, X49, X53, X55, X56, X63, X65

According to the Boruta method, 22 out of 69 independent variables are confirmed as important and 31 as unimportant while the remaining 16 variables have tentative attributes.

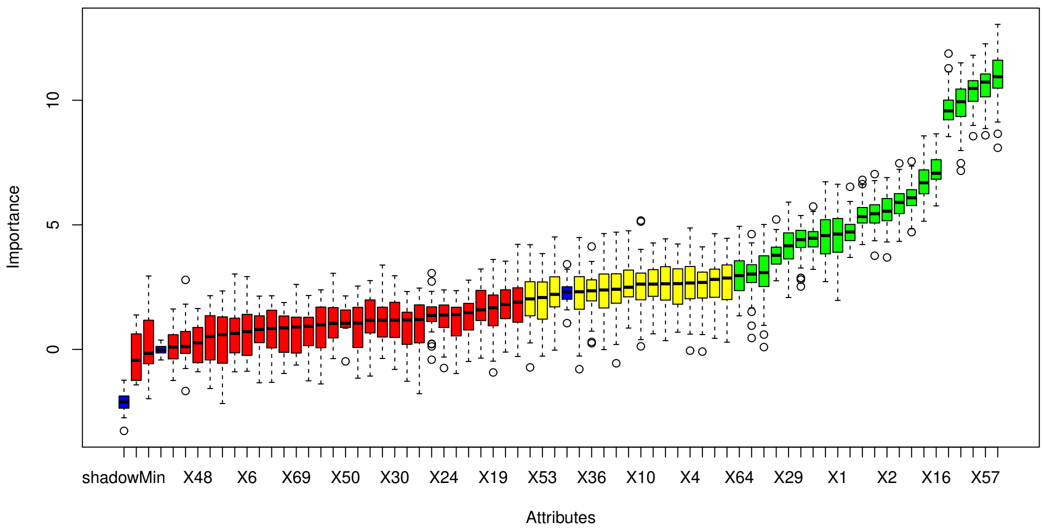


Figure 4.1: Plotting of Boruta Feature Selection

The ratios confirmed as important are as follows and constitute the new data set. In Figure 4.1, important variables are shown in green and unimportant variables are shown in red.

All the analysis methods used in this study are run with both the data set (301x70) and the data set (301x23) which is obtained from the Boruta. Therefore, there are two data sets which are used in the following analyses. The 2018 data is separated from all data. The 2018 data is intended to be the test data sets and is approximately 13% of all our data. In short; Train, Test, Boruta Train, and Boruta Test data sets have data points of (263x70), (38x70), (273x23) and (38x23) respectively, given as (Observation x Feature).

4.2 Random Forest with Data Sets

Firstly, the grid search for Random Forest is applied to the training data and the following output is obtained.

Table 4.2: Grid for Random Forest

Used Hyper-Parameters			max_depth
			mtries
			ntrees
Number of Models			200
Number of Failed Models			0
Hyper-Parameter Search Summary: ordered by increasing rmse			
max_depth	mtries	ntrees	RMSE
15	7	1000	0.513536526
15	7	300	0.517067667
10	7	500	0.526636139
15	6	300	0.527177147
10	6	1000	0.531556650
⋮	⋮	⋮	⋮
15	2	2	0.987258192
2	3	2	0.991902714
4	3	5	1.011027667
5	7	5	1.034234971
3	4	2	1.097505538
2	6	2	1.161763593

In order to obtain the best results, 200 different models have been created and these models have been ordered from small to large according to their RMSE values. According to the result, the parameters of the best model are max_depth=15, mtries=7, ntrees=1000.

Then, the Random Forest with the best model parameters determined is run. In the first Random Forest model, the train data set is the data with 70 ratios between 2011 and 2017, and the test data set is the data set with 70 ratios in 2018. Train and test performances of Random Forest are as follows. Moreover, the graph of RMSE values of training according to the number of trees can be seen in Figure 4.2.

Table 4.3: Random Forest Details

Model Details	
number_of_trees	1000
number_of_nternal_trees	1000
min_depth	15
max_depth	15
mean_depth	15
min_leaves	86
max_leaves	180
mean_leaves	143.747

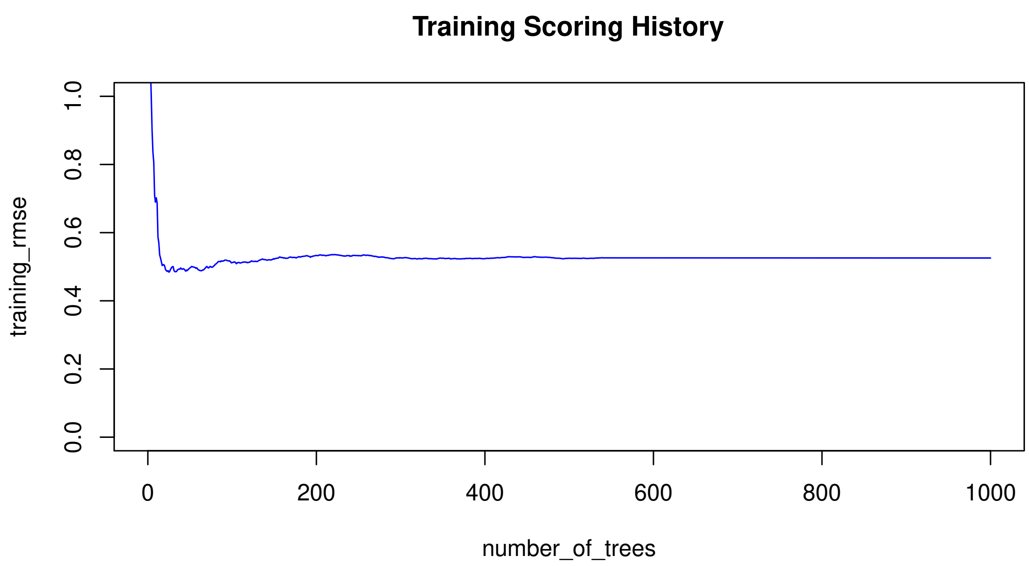


Figure 4.2: Plotting of Random Forest

Performances of train and test are shown in Table 4.4. Here, NaN stands for not a number. As a result of the first Random Forest analysis, the RMSE value of the training data is 0.1953, while the RMSE value of the test data is 0.4820.

Table 4.4: Train and Test Performance of Random Forest

	RF Train	RF Test
MSE	0.0381597	0.2323729
RMSE	0.1953452	0.4820508
MAE	0.1045557	0.2524219
RMSLE	NaN	0.1527293
Mean Residual Deviance	0.0381597	0.2323729

Sets of data with 22 ratios obtained with the Boruta in the same random forest model are set as the 2011-2017 train and 2018 test data. Then, the model is run. Random Forest with Boruta data set details are given in Table 4.5. The performance of the test and train data obtained with Boruta data are compared in Table 4.6. Again, the graph of RMSE values of training according to the number of trees set can be seen in Figure 4.3.

Table 4.5: Random Forest with Boruta Data Set Details

Model Details	
number_of_trees	1000
number_of_nternal_trees	1000
min_depth	14
max_depth	15
mean_depth	14.992
min_leaves	118
max_leaves	182
mean_leaves	154.516

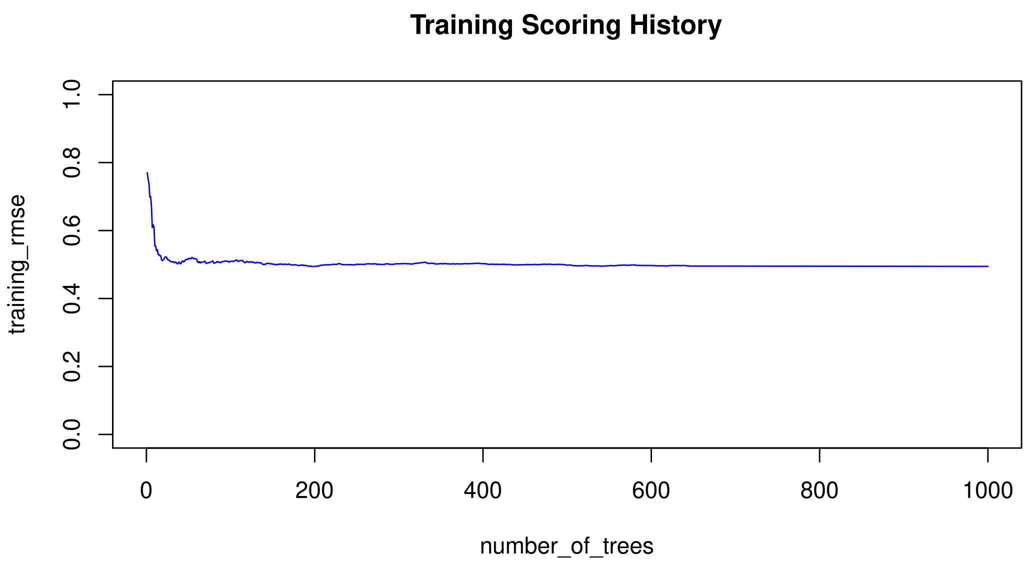


Figure 4.3: Plotting of Random Forest with Boruta Data

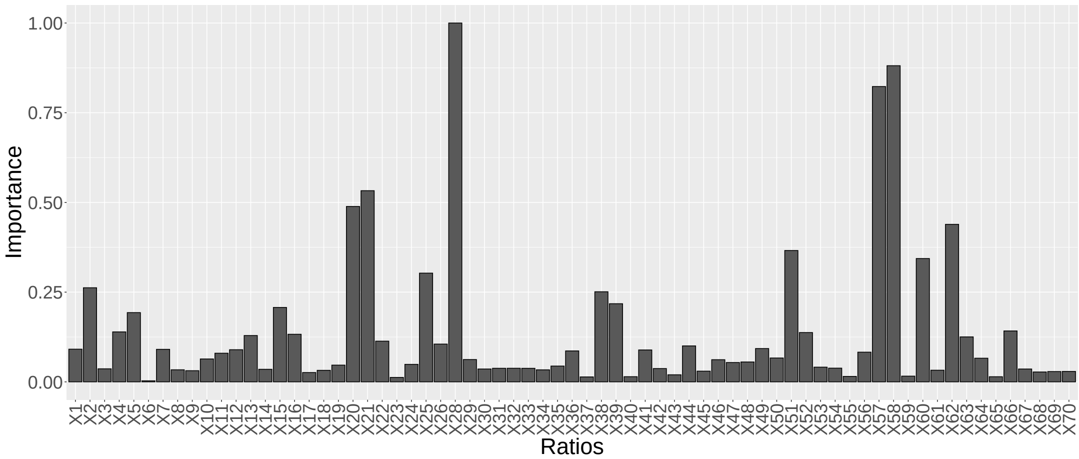
Table 4.6: Train and Test Performance of Random Forest with Boruta Data Set

	RF Train Boruta	RF Test Boruta
MSE	0.03365679	0.2149195
RMSE	0.18345790	0.4635942
MAE	0.09570488	0.2331598
RMSLE	NaN	0.1345234
Mean Residual Deviance	0.03365679	0.2149195

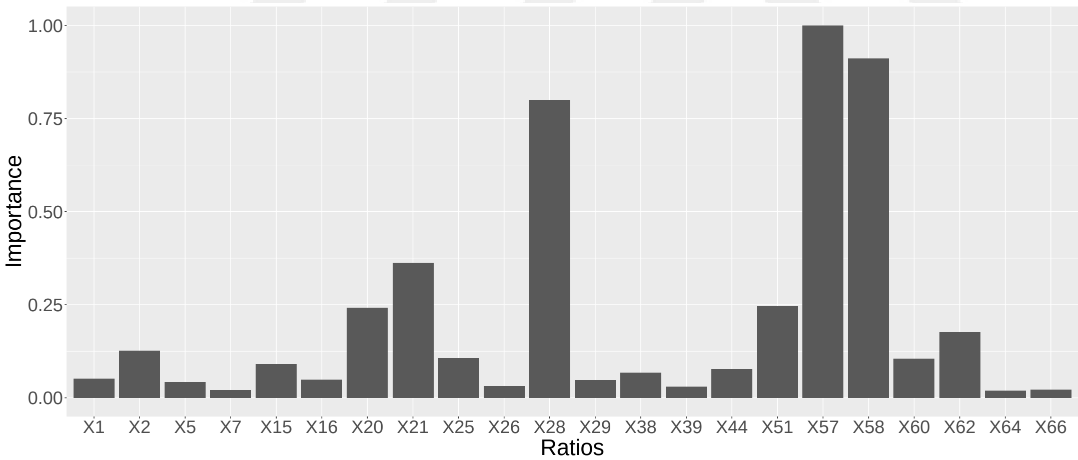
As a result, when the number of independent variables is reduced from 69 ratios to 22 ratios in the Random Forest method, RMSE of the test and train predictions made by the model get lower. In short,

$$\text{RF\_Train RMSE} > \text{RF\_Train\_Boruta RMSE},$$
$$\text{RF\_Test RMSE} > \text{RF\_Test\_Boruta RMSE}.$$

The variable importance in Random Forest method for two data sets shows that the most significant variables with order are X28, X58, X57 and so on as can be seen from Figure 4.4. When we look at the Boruta data set, the variable importance of similar ratios is high. For other methods, the variable importance is shown in Appendix A.



(a) Full Data



(b) Boruta Data

Figure 4.4: Variable Importance for Random Forest with (a) Full Data and (b) Boruta Data

4.3 Neural Networks with Data Sets

Secondly, an analysis is performed using the deep learning method. First of all, parameter selection with the Gird search is made for the Neural Networks. The aim is to decide the most suitable parameters for the data set. The grid details on the train data set can be seen in Table 4.7.

Table 4.7: Grid for Neural Networks

Used Hyper-Parameters	activation
	epochs
	hidden
Number of Models	200
Number of Failed Models	0
Hyper-Parameter Search Summary: ordered by increasing rmse	

activation	epochs	hidden	RMSE
Tanh	1000	[100, 100, 100, 100]	0.0221396
Rectifier	1000	[100, 100, 100, 100]	0.0242539
Tanh	1000	[50, 50, 50, 50]	0.0258954
Rectifier	1000	[10, 10, 10, 10]	0.0282181
Tanh	500	[50, 50, 50, 50]	0.0302755
⋮	⋮	⋮	⋮
MaxoutWithDropout	1000	[5, 5, 5, 5]	1.1196544
RectifierWithDropout	500	[5, 5, 5, 5]	1.1225386
RectifierWithDropout	200	[5, 5, 5, 5]	1.1301783
RectifierWithDropout	1000	[5, 5, 5, 5]	1.1301783
RectifierWithDropout	500	[5, 5, 5, 5]	1.1301783
RectifierWithDropout	1000	[5, 5, 5, 5]	1.2518754

In total 200 models are created and the 87th model with the lowest RMSE value is chosen as the best model. In this model, parameters of the Neural Networks are that activation function is Tanh, the number of epochs is 1000 and hidden layers are [100,100,100,100]. Then, the Deep Learning is computed. There are four hidden layers and Tanh function is used in the best model.

These are run for the 2011-2017 train and 2018 test data, and then run on the Boruta data set in the same way. Details of Neural Networks are shown in Table 4.8.

Table 4.8: Neural Networks Details

Model Details						
Layer	1	2	3	4	5	6
Units	69	100	100	100	100	1
Type	Input	Tanh	Tanh	Tanh	Tanh	Linear
Mean_rate	NA	0.021173	0.022002	0.031763	0.070175	0.001931
Rate_rms	NA	0.030987	0.009629	0.016052	0.165822	0.000360
Mean_weight	NA	0.000482	-0.001923	-0.000578	0.001869	-0.005482
Weight_rms	NA	0.108848	0.102471	0.101055	0.101426	0.129181
Mean_bias	NA	0.001957	-0.000764	-0.000136	0.000333	-0.012445
Bias_rms	NA	0.232300	0.031455	0.033008	0.016403	0.000000

The RMSE value of Neural Networks, which completed machine learning through training data, is 0.1604. Test and train perfomance details are shown in Table 4.9.

Table 4.9: Train and Test Performance of Neural Networks

	NN Train	NN Test
MSE	0.0257287	0.3838322
RMSE	0.1604017	0.6195419
MAE	0.1224909	0.3165279
RMSLE	NaN	0.2084888
Mean Residual Deviance	0.0257287	0.3838322

On the data set obtained with Boruta, Neural Networks method is repeated with the parameters in the first Neural Networks analysis. Details of this analysis are given in Table 4.10. and test and train performances are compared in Table 4.11.

Table 4.10: Neural Networks with Boruta Data Set Details

Model Details						
Layer	1	2	3	4	5	6
Units	22	100	100	100	100	1
Type	Input	Tanh	Tanh	Tanh	Tanh	Linear
Mean_rate	NA	0.017824	0.026041	0.037163	0.125428	0.001906
Rate_rms	NA	0.017799	0.013471	0.020876	0.233592	0.001331
Mean_weight	NA	0.003158	-0.002035	-0.000678	0.001673	-0.004221
Weight_rms	NA	0.129332	0.103666	0.103382	0.103323	0.126631
Mean_bias	NA	0.005579	-0.006194	-0.005570	0.006488	-0.004000
Bias_rms	NA	0.053002	0.080232	0.073778	0.045576	0.000000

Table 4.11: Train and Test Performance of Neural Networks with Boruta Data Set

	NN Train Boruta	NN Test Boruta
MSE	0.06065461	0.3861052
RMSE	0.24628160	0.6213736
MAE	0.18071440	0.3078977
RMSLE	NaN	0.2925063
Mean Residual Deviance	0.06065461	0.3861052

According to the results of the Deep Learning study, RMSE value is 0.1604 in training and 0.61954 in testing. In the study conducted with the Boruta data, RMSE value is 0.2462 in training while the testing RMSE is calculated as 0.6213. When the size of the data set used in Neural Networks analysis reduced, decreases in test and train performances can be observed.

4.4 Gradient Boosting Machine with Data Sets

Thirdly, the Grid search is added to the training data for the Gradient Boosting Machine and the following output is reached (Table 4.12). With this analysis, it is aimed to select the best model.

Table 4.12: Grid for Gradient Boosting Machine

Used Hyper-Parameters	col_sample_rate
	learn_rate
	max_depth
	ntrees
	sample_rate
Number of Models	360
Number of Failed Models	0
Hyper-Parameter Search Summary: ordered by increasing rmse	

col_sample_rate	learn_rate	max_depth	ntrees	sample_rate	RMSE
1.0	0.1	9	20	0.8	0.45901759
1.0	0.1	9	30	0.8	0.45912757
1.0	0.1	9	40	0.8	0.46246190
1.0	0.1	9	50	0.8	0.46753371
1.0	0.1	9	20	1.0	0.47064592
⋮	⋮	⋮	⋮	⋮	⋮
0.2	0.01	9	10	1.0	0.76189831
0.2	0.01	5	10	1.0	0.76496634
0.2	0.01	3	10	0.8	0.76691161
0.5	0.01	3	10	1.0	0.76790154
0.5	0.01	3	10	0.8	0.76809016
0.2	0.01	3	10	1.0	0.77193972

In total, 360 models are tried and the lowest RMSE value is found for the model with col_sample_rate = 1, learn_rate 0.1, max_depth = 9, ntrees = 20.

With these parameters, 2011-2017 data are trained, and 2018 data are tested. The model details can be seen in Table 4.13. Figure 4.5 is the plotting of the RMSE training score history of the Gradient Boosting Machine. Then the regression performances of test and train data are shown in Table 4.14.

Table 4.13: Gradient Boosting Machine Details

Model Details	
number_of_trees	20
number_of_nternal_trees	20
min_depth	8
max_depth	9
mean_depth	8.75
min_leaves	10
max_leaves	18
mean_leaves	16

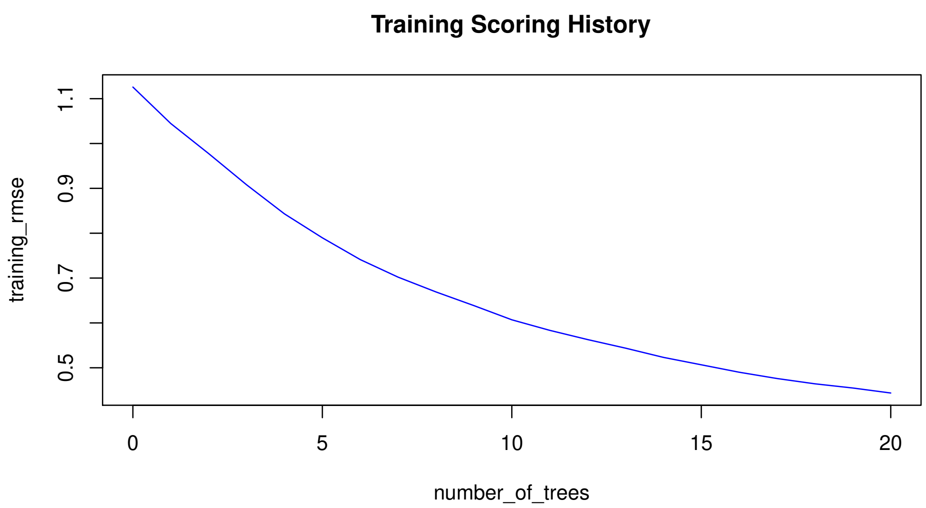


Figure 4.5: Plotting of Gradient Boosting Machine

Table 4.14: Train and Test Performance of Gradient Boosting Machine

	GBM Train	GBM Test
MSE	0.1968865	0.2221473
RMSE	0.4437190	0.4713250
MAE	0.1892934	0.2652365
RMSLE	NaN	0.1475609
Mean Residual Deviance	0.1968865	0.2221473

The analysis is repeated with the same parameters and splits for the data set chosen by Boruta. Model results are as follows in Table 4.15. Also, train and test performance of Gradient Boosting Machines with Boruta data set is shown below in Table 4.16. Training scoring history of RMSE as per number of trees is shown in Figure 4.6. As seen in this Figure, while the number of trees increases from 0 to 20, the RMSE value decreases from 1.1 to 0.4419. When the results of this method on two data sets are compared, the RMSE value of the test performance of the Boruta data set is slightly higher, while the train performance is almost the same.

Table 4.15: Gradient Boosting Machine with Boruta Data set Details

Model Details	
number_of_trees	20
number_of_nternal_trees	20
min_depth	7
max_depth	9
mean_depth	8.60
min_leaves	13
max_leaves	18
mean_leaves	16.05

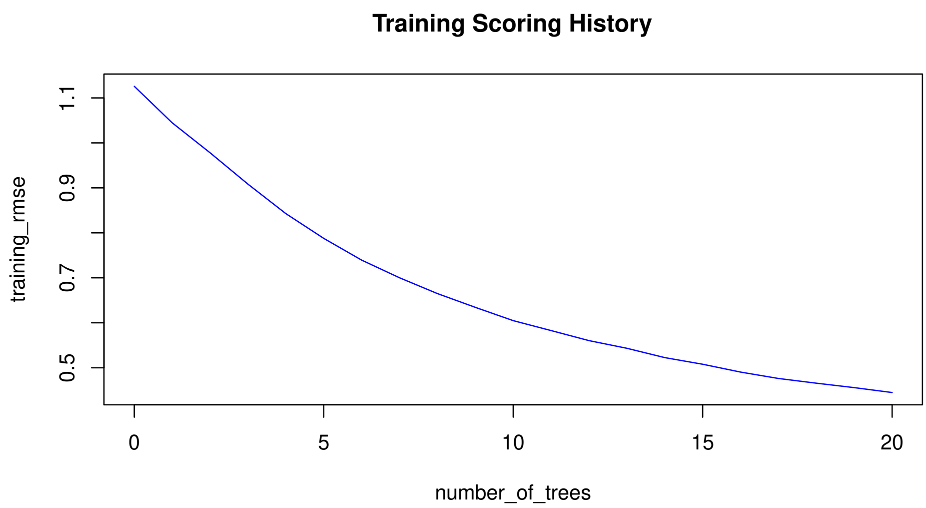


Figure 4.6: Plotting of Gradient Boosting Machine with Boruta Data

Table 4.16: Train and Test Performance of GBM with Boruta Data Set

	GBM Train Boruta	GBM Test Boruta
MSE	0.1953357	0.2342573
RMSE	0.4419680	0.4840013
MAE	0.2011118	0.2812078
RMSLE	NaN	0.1580579
Mean Residual Deviance	0.1953357	0.2342573

4.5 Extreme Gradient Boosting with Data Sets

Fourthly and lastly, the grid method is used to determine the parameters of the XGboost method. The output of the method designating the most suitable parameters for the train data set is as in Table 4.17

Table 4.17: Grid for XGBoost

Used Hyper-Parameters					col_sample_rate
					learn_rate
					max_depth
					ntrees
					sample_rate
Number of Models					50
Number of Failed Models					0
Hyper-Parameter Search Summary: ordered by increasing rmse					
col_sample_rate	learn_rate	max_depth	ntrees	sample_rate	RMSE
0.79	0.186	4	990	0.94	0.4637661
0.87	0.181	14	740	0.85	0.4794839
0.41	0.066	18	120	0.9	0.4806477
0.84	0.141	12	850	0.87	0.4888496
0.84	0.076	8	880	0.88	0.4928266
⋮	⋮	⋮	⋮	⋮	⋮
0.48	0.181	12	700	0.62	0.5491531
0.42	0.131	7	240	0.64	0.5560640
0.26	0.171	6	510	0.68	0.5856609
0.21	0.011	1	520	0.51	0.5982795
0.75	0.006	14	90	0.89	0.9336443
0.35	0.001	15	570	0.74	0.9451144

Grid analysis for XGBoost including 50 models is performed and RMSE values are sorted from small to large. The model with the smallest RMSE is the best model and has parameters of col_sample_rate=0.79, learn_rate=0.186, max_depth=4, ntrees=990, sample_rate=0.94. Using the parameters of the model with the best results, 2018 data are tested by training of the 2011-2017 data set.The output of XGBoost model is given in Table 4.18.

Table 4.18: XGBoost Details

Model Details	
col_sample_rate	0.79
learn_rate	0.186
max_depth	4
ntrees	990
sample_rate	0.94

The performance results of the test and train data are in Table 4.19. There are 990 trees in XGBoost method. The RMSE value of the train data is 0.0004 while the RMSE value of the test data is 0.4539.

Table 4.19: Train and Test Performance of XGBoost

	XGBoost Train	XGBoost Test
MSE	2.06211E-07	0.2060784
RMSE	0.000454105	0.4539586
MAE	0.000344849	0.2266698
RMSLE	NaN	0.1369161
Mean Residual Deviance	2.06211E-07	0.2060784

XGBoost Machine Learning algorithm, which is applied with the first data set, is also applied to Boruta dataset with the same parameters. Model details of XGBoost model with Boruta data sets is shown in Table 4.20.

Table 4.20: XGBoost with Boruta Details

Model Details	
col_sample_rate	0.79
learn_rate	0.186
max_depth	4
ntrees	990
sample_rate	0.94

In Table 4.21, test and train performance of the Boruta data set are shown and the train RMSE is 0.0005 and the test RMSE is 0.4207. Compared to the performance of our first dataset, the train performance is similar, however, the test performance decreased.

Table 4.21: Train and Test Performance of XGBoost with Boruta

	XGBoost Train Boruta	XGBoost Test Boruta
MSE	2.90264E-07	0.1770168
RMSE	0.000538761	0.4207336
MAE	0.00040923	0.2470194
RMSLE	NaN	0.1363722
Mean Residual Deviance	2.90264E-07	0.1770168

To sum up, in Table 4.22, the outputs of the MSE values of all analyses can be seen for both train and test. Table 4.23 and Table 4.24 represent RMSE and MAE values respectively.

Table 4.22: Comparative MSE Values for the ML Methods Used

	RF	**NN**	**GBM**	**XGB**	**RF_b**	**NN_b**	**GBM_b**	**XGB_b**
train_mse	0.03816	0.025729	0.196887	2.06E-07	0.033657	0.060655	0.195336	2.90E-07
test_mse	0.232373	0.383832	0.222147	2.06E-01	0.21492	0.386105	0.234257	1.77E-01

Table 4.23: Comparative RMSE Values for the ML Methods Used

	RF	**NN**	**GBM**	**XGB**	**RF_b**	**NN_b**	**GBM_b**	**XGB_b**
train_rmse	0.195345	0.160402	0.443719	0.000454	0.183458	0.246282	0.441968	0.000539
test_rmse	0.482051	0.619542	0.471325	0.453959	0.463594	0.621374	0.484001	0.420734

Table 4.24: Comparative MAE Values for the ML Methods Used

	RF	**NN**	**GBM**	**XGB**	**RF_b**	**NN_b**	**GBM_b**	**XGB_b**
train_mae	0.1045557	0.1224909	0.1892934	0.000345	0.09570488	0.1807144	0.2011118	0.000409
test_mae	0.2524219	0.3165279	0.2652365	0.22667	0.23315982	0.3078977	0.2812078	0.247019

Estimates of all methods for the y independent values are obtained as outputs and are displayed as $\hat{y}$. The MSE values are calculated according to the Equation (3.8) using the observed values and the estimated values. The values for RMSE and MAE are calculated by the Equations (3.9) and (3.10) respectively.

Considering the tables above, the XGBoost method is the best machine learning model according to MSE, RMSE, MAE values. However, very small RMSE values could be an alert of overfitting. If a model receives such a big amount of information, it begins to learn from the noise and incorrect information in the data set. Then, because of too much detail and noise, the model does not classify the information properly. Very high training performance can lead to very low test performance. When the test values are compared, the XGBoost method has also the smallest MSE, RMSE and MAE values.

4.6 Analysis on Administrative Measures Stages

According to the Regulation on Measurement and Assessment of Capital Requirements of Insurance, Reinsurance and Pension Companies, TRMTF takes administrative measures with respect to capital requirement ratio. The capital requirement ratio of insurance companies covers 4 stages and takes different names according to the percentage range. $100\% - 115\%$ is the self-evaluation stage, $70\% - 99.99\%$ is take precaution stage, $33\% - 69.99\%$ is emergency take precaution stage, and less than 33% is intervention stage.

In this case, we aimed to measure whether the predicted values fall within the correct range. The observed values of capital requirements ratio for 2018 and estimated by all methods are numbered according to the above values. According to this; intervention stage takes '-2', emergency take precaution stage takes '-1', take precaution stage takes '0', self-evaluation stage takes '1' and greater than 115% takes '2'.

In the analysis, where the years of 2011-2017 are used as training data, the values obtained from the 2018 forecasts are shown in Table 4.25. If the predicted value and the actual value correspond to the same range, it is considered 'TRUE', otherwise 'FALSE'. The accuracy rate is calculated by dividing the number of observations correctly estimated to the total number of observations.

Table 4.25: Predictions by Ranges of Capital Requirements Ratio

Company	y	**Random Forest** $\hat{y}$	**Random Forest with Boruta** $\hat{y}$	**Neural Networks** $\hat{y}$	**Neural Networks with Boruta** $\hat{y}$	**Gradient Boosting Machine** $\hat{y}$	**Gradient Boosting Machine with Boruta** $\hat{y}$	**eXtreme Gradient Boosting** $\hat{y}$	**eXtreme Gradient Boosting with Boruta** $\hat{y}$
1	0	2	2	0	0	2	2	1	2
2	2	2	1	2	2	2	1	1	0
3	2	2	2	2	2	2	2	2	2
4	2	2	2	2	1	2	2	2	2
5	2	2	2	2	2	2	2	2	2
6	0	1	2	0	1	2	2	2	2
7	2	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2	2
9	2	2	2	1	2	2	1	2	1
10	-1	-1	-1	2	1	-1	0	-1	-1
11	1	1	0	1	0	1	1	1	0
12	2	2	2	1	2	2	2	2	2
13	2	2	2	2	2	2	2	2	2
14	2	2	2	2	2	2	2	2	2
15	2	2	2	2	2	2	2	2	2
16	2	2	2	2	2	2	2	2	2
17	2	2	2	2	2	2	2	2	2
18	1	0	1	2	1	1	1	0	1
19	2	2	2	2	2	2	2	2	2
20	0	0	0	-1	-1	0	0	-1	-1
21	2	2	2	1	0	2	2	2	2
22	1	1	0	2	0	1	0	0	0
23	2	2	2	2	2	2	2	2	2
24	0	0	0	-1	0	0	0	0	0
25	2	2	2	2	1	2	2	2	2
26	2	2	2	0	2	2	2	2	2
27	2	2	2	2	2	2	2	2	2
28	2	0	0	2	0	0	0	0	0
29	-1	-1	-1	-1	-1	-1	-1	-1	-1
30	2	2	2	2	2	2	2	2	2
31	1	1	1	1	0	1	1	1	1
32	2	2	2	1	2	1	2	2	1
33	2	2	2	2	2	2	2	2	2
34	2	2	2	2	2	2	2	2	2
35	1	1	1	2	0	2	2	2	2
36	-1	0	-1	-1	0	-1	-1	0	-1
37	-2	-2	-2	-2	-2	-2	-2	-2	-2
38	2	2	2	2	2	2	2	2	2
FALSE		5	6	11	12	5	8	9	10
TRUE		33	32	27	26	33	30	29	28
PERCENTAGE		0.87	0.84	0.71	0.68	0.87	0.79	0.76	0.74

There are 38 observations in 2018. If the actual and forecast values of these observations fall within a different range, we can conclude that our method predicts fail for that company. As shown in Table 4.25, if the actual and predicted values are different, the predicted value is painted in red. In this way, the methods with the least mistakes are Random Forest and GBM. Both has an accuracy of 87%. Neural Networks method estimates with 71% accuracy. The most striking feature of the Neural Networks method is that it accurately predicts some observations that all other methods cannot foresee. Only Neural Networks can estimate correctly the 1st, 6th, and 28th observations. While GBM made 5 incorrect estimates with the full data set, the number of errors increased to 8 in the analysis made with the Boruta data set. Although the estimated accuracy of the XGBoost method with two data sets is close to one another, it produces fewer errors with the full data set. The error number of the XGBoost process that is run with the Boruta data is higher than XGBoost with complete data.

In short, the following order can be made:

$$\text{Random Forest} = \text{Gradient Boosting Machine} > \text{Random Forest with Boruta} >$$
$$\text{Gradient Boosting Machine with Boruta} > \text{XGBoost} >$$
$$\text{XGBoost with Boruta} > \text{Neural Networks} > \text{Neural Networks with Boruta}.$$

4.7 Analysis on Capital Requirements

An examination is made on the estimation of whether the capital of the company is sufficient or not. According to the Regulation on Measurement and Assessment of Capital Requirements of Insurance, Reinsurance and Pension Companies, the equity should not be lower than the required equity. This means that the capital requirement ratio should be greater than 100%. Even though the initial range estimates investigation provides more sensitive accuracy; it can be understood what extent it can be accurately predicted whether the capital of the companies is sufficient or not thanks to the second examination. If the capital requirement ratio of the company is greater than 100%, this company has 1; otherwise, it has 0. Thus, we can see that the capital adequacy status is sufficient or insufficient. The outputs of the actual and the estimated values are also shown in Table 4.26. If these two values do not match, the estimate values are painted in red. The false numbers, true numbers, and accuracy percentages of all methods are given at the bottom of the table. The best predictive analysis method is found as Neural Networks. With this method, we can predict whether the capital of the companies is adequate or not with 95% accuracy. While 3 incorrect estimates are made by the GBM method, the number of errors in Random Forest is 4. Accuracy success is 92% and 89% respectively. The predictive power of Random Forest is the same as that of GBM with Boruta. In all methods, the percentage of success with the Boruta data set reduces. At the same time, the Neural Networks method which is run with the Boruta data set has the worst prediction of dependent variable, and it is 79%.

Table 4.26: Predictions Capital Requirements Ratio 1 - 0

Company	y	Random Forest $\hat{y}$	Random Forest with Boruta $\hat{y}$	Neural Networks $\hat{y}$	Neural Networks with Boruta $\hat{y}$	Gradient Boosting Machine $\hat{y}$	Gradient Boosting Machine with Boruta $\hat{y}$	eXtreme Gradient Boosting $\hat{y}$	eXtreme Gradient Boosting with Boruta $\hat{y}$
1	0	1	1	0	0	1	1	1	1
2	1	1	1	1	1	1	1	1	0
3	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1
6	0	1	1	0	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1
10	0	0	0	1	1	0	0	0	0
11	1	1	0	1	0	1	1	1	0
12	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1
18	1	0	1	1	1	1	1	0	1
19	1	1	1	1	1	1	1	1	1
20	0	0	0	0	0	0	0	0	0
21	1	1	1	1	0	1	1	1	1
22	1	1	0	1	0	1	0	0	0
23	1	1	1	1	1	1	1	1	1
24	0	0	0	0	0	0	0	0	0
25	1	1	1	1	1	1	1	1	1
26	1	1	1	0	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1
28	1	0	0	1	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0
30	1	1	1	1	1	1	1	1	1
31	1	1	1	1	0	1	1	1	1
32	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1
35	1	1	1	1	0	1	1	1	1
36	0	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0	0
38	1	1	1	1	1	1	1	1	1
FALSE		4	5	2	8	3	4	5	6
TRUE		34	33	36	30	35	34	33	32
PERCENTAGE		0.89	0.87	0.95	0.79	0.92	0.89	0.87	0.84

In short, the following order can be made:

$$\text{Neural Networks} > \text{Gradient Boosting Machine} > \text{Random Forest} = \text{Gradient Boosting Machine with Boruta} > \text{Random Forest with Boruta} = \text{XGBoost} > \text{XGBoost with Boruta} > \text{Neural Networks with Boruta.}$$

CHAPTER 5

CONCLUSION

Generally, the main concern of the insurance sector is whether companies can pay their debts. This problem is very important for the continuity and robustness of the company, the sector, and even the whole economy. We can say that the most indicator that is whether the companies can pay their liabilities is sufficient for their capital. The biggest supporter of this is the use of capital requirement ratio by regulator to take action on insurance companies. As a result of this ratio, the situation of the company can be seen well, the company can be warned and even intervened. Establishing an early warning model with machine learning methods can be seen as the starting point of this thesis. With the light of the previous studies, we aim to introduce an early warning model using Machine Learning algorithms using the financial data from the Turkish non-life insurance companies. For all these reasons, CRR is chosen as the dependent variable. With sixty-nine independent variables, while the 2011-2017 data set is defined as training data, dependent variables in 2018 are wanted to be estimated. Likewise, the same method is repeated in the data set consisting of 22 ratios obtained by the Boruta feature selection method. Random Forest, Neural Network, Gradient Boosting Machine, and eXtreme Gradient Boosting methods are applied to both two data sets. The Grid method is applied for the selection of method parameters and the best model is selected. Firstly, it is analyzed whether the predicted values and actual values coincide with the same range. The highest two accuracy percentages are Random Forest and Gradient Boosting Machine with 87%. Then, the estimation and actual values are compared according to whether capital is sufficient or not, and the best model is the Neural Networks with 95% accuracy.

The number of variables is reduced by Boruta feature selection method and all analyzes are applied to this data set. The predictive power is increased as the information we teach to the machine increases. However, the estimates from Boruta Data set show good performance.

In the light of the analyses, this model can be used to monitor the current situation of insurance companies and can be used as a guide in the measures to be taken. For supervisory and regulatory authorities, it will support their ability to conduct risk-focused supervision.

REFERENCES

[1] L. Alessi, A. Antunes, J. Babeckỳ, S. Baltussen, M. Behn, D. Bonfim, O. Bush, C. Detken, J. Frost, R. Guimaraes, et al., Comparing Different Early Warning Systems: Results from a horse race competition among members of the macro-prudential research network, SSRN, 2015.

[2] E. I. Altman, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, The Journal of Finance, 23(4), pp. 589–609, 1968.

[3] Y. Amihud, H. Mendelson, L. H. Pedersen, et al., Liquidity and asset prices, Foundations and Trends in Finance, 1(4), pp. 269–364, 2006.

[4] S. M. Bartram, G. W. Brown, and W. Waller, How important is financial risk?, Journal of Financial and Quantitative Analysis, 50(4), pp. 801–824, 2015.

[5] W. Beaver, Financial ratios as predictors of failure journal of accounting research, Empirical Research in Accounting: Selected Studies, 4(1), pp. 71–111, 1966.

[6] M. Blum, Failing company discriminant analysis, Journal of Accounting Research, pp. 1–25, 1974.

[7] L. Breiman, Random forests, Machine Learning, 45(1), pp. 5–32, 2001.

[8] A. Breuer, M. L. Frumusanu, B. L. Breuer, and A. Manciu, Cash and liquidity/liquidity and liquidity ratio, Annals - Economy Series, 4I, pp. 78–82, 2012.

[9] E. F. Brigham and M. C. Ehrhardt, *Financial Management: Theory and Practice*, Thomson ONE Publishers, 2010.

[10] P. L. Brockett, W. W. Cooper, L. L. Golden, and U. Pitaktong, A neural network method for obtaining an early warning of insurer insolvency, Journal of Risk and Insurance, 61, pp. 402–402, 1994.

[11] W. Bruns, *Introduction to Financial Ratios and Financial Statement Analysis*, President and Fellows of Harvard College, 1992.

[12] BusinessDictionary.com. WebFinance, Inc., Risk classification, `http://www.businessdictionary.com/definition/risk-classification.html`, 2019.

[13] A. Candel, E. LeDell, and A. Bartz, Deep learning with h2o, H2O.ai, Inc., 2019.

[14] T. Chai and R. R. Draxler, Root mean square error (rmse) or mean absolute error (mae)?– arguments against avoiding rmse in the literature, Geoscientific Model Development, 7(3), pp. 1247–1250, 2014.

[15] T. Chen and C. Guestrin, Xgboost: A scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 08 2016.

[16] T. Chen, T. He, M. Benesty, V. Khotilovich, and Y. Tang, Xgboost: eXtreme Gradient Boosting, R Package Version 0.4-2, pp. 1–4, 2015.

[17] P. Christoffersen, *Elements of Financial Risk Management, Second Edition*, Academic Press, 2 edition, 2011, ISBN 0123744482,9780123744487.

[18] D. Cook, *Practical machine learning with H2O: powerful, scalable techniques for deep learning and AI*, " O'Reilly Media, Inc.", 2016.

[19] Council of the European Union, First council directive 73/239/eec of 24 july 1973 on the coordination of laws, regulations and administrative provisions relating to the taking-up and pursuit of the business of direct insurance other than life assurance, Official Journal of the European Communities, 1973-07-27, `https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A31973L0239`.

[20] Council of the European Union, First council directive 79/267/eec of 5 march 1979 on the coordination of laws, regulations and administrative provisions relating to the taking up and pursuit of the business of direct life assurance, Official Journal of the European Communities, 1979-03-5, `https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:31979L0267`.

[21] E. B. Deakin, A discriminant analysis of predictors of business failure, Journal of Accounting Research, pp. 167–179, 1972.

[22] R. O. Edmister, An empirical test of financial ratio analysis for small business failure prediction, Journal of Financial and Quantitative Analysis, 7(2), pp. 1477–1493, 1972.

[23] European Commission, Solvency II Overview - Frequently Asked Questions, 2015, `https://europa.eu/rapid/press-release_MEMO-15-3120_en.htm`.

[24] European Parliament, Council of the European Union, Directive 2009/138/ec of the european parliament and of the council of 25 november 2009 on the taking-up and pursuit of the business of insurance and reinsurance (solvency ii), Official Journal of the European Communities, 2009-11-25, `https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32009L0138`.

[25] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*, volume 1, Springer Series in Statistics New York, 2001.

[26] A. Genç, *Measuring Financial Adequacy for Non-life Insurance Companies and Early Warning Model Recommendation for Turkey*, Ph.D. thesis, Ankara University, 2002.

[27] H2O.ai, Xgboost algorithm, last update Jul 15, 2019, http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/xgboost.html.

[28] M. Heep-Altiner, M. Mullins, and T. Rohlfs, *Solvency II in the Insurance Industry*, Springer, 2018.

[29] K. A. Horcher, *Essentials of financial risk management*, volume 32, John Wiley & Sons, 2011.

[30] S. ICAP GROUP, Financial ratios explanation, 2006, https://docplayer.net/7887774-Icap-group-s-a-financial-ratios-explanation.html.

[31] Insurance Association of Turkey, Insurance definitions, https://www.tsb.org.tr/sigorta-tanimlari.aspx?pageID=648.

[32] Insurance Association of Turkey, Solvency, https://www.tsb.org.tr/sermaye-yeterliligi-solvency.aspx?pageID=821.

[33] Insurance Association of Turkey, Statistics, https://www.tsb.org.tr/gelir-tablosu.aspx?pageID=911.

[34] Insurance Association of Turkey, Annual Report of Insurance Association of Turkey, 2017, https://www.tsb.org.tr/Document/Publications/2017Annual%20Report.pdf.

[35] Insurance Europe, Insight Briefing Solvency II Review, Insurance Europe AISBL, 2018, https://www.insuranceeurope.eu/insight-briefing-2018-solvency-ii-review.

[36] G. Isseveroglu and U. Gucenme, Early warning model with statistical analysis procedures in turkish insurance companies, African Journal of Business Management, 4(5), pp. 623–630, 2010.

[37] Y. Ito, T. Kitamura, K. Nakamura, T. Nakazawa, et al., New financial activity indexes: Early warning system for financial imbalances in japan, Technical report, Bank of Japan, 2014, https://www.boj.or.jp/en/research/wps_rev/wps_2014/data/wp14e07.pdf.

[38] J. James, *Lloyd's and the London Insurance Market: An Overview*, Lloyd's of London, 2007.

[39] J. KAGAN, Underwriting risk, https://www.investopedia.com/terms/u/underwriting-risk.asp, updated Feb 20, 2018.

[40] W. Koehrsen, An Implementation and Explanation of the Random Forest in Python, 2018, www.towardsdatascience.com.

[41] M. B. Kursa, Boruta for those in a hurry, 2018, https://rdrr.io/cran/Boruta/f/inst/doc/inahurry.pdf.

[42] M. B. Kursa, W. R. Rudnicki, et al., Feature selection with the boruta package, J Stat Softw, 36(11), pp. 1–13, 2010.

[43] A. Liaw, M. Wiener, et al., Classification and regression by random forest, R News, 2(3), pp. 18–22, 2002.

[44] R. Libby, Accounting ratios and the prediction of failure: Some behavioral evidence, Journal of Accounting Research, 13, pp. 150–161, 1975.

[45] M. Malohlava, A. Candel, and A. Bartz, Gradient boosting machine with h2o, H2O.ai, Inc., 2019.

[46] P. A. Meyer and H. W. Pifer, Prediction of bank failures, The Journal of Finance, 25(4), pp. 853–868, 1970.

[47] R. Mitchell and E. Frank, Accelerating the xgboost algorithm using gpu computing, PeerJ Computer Science, 3, p. 127, 2017.

[48] G. Ocak, An Early Warning Model for Turkish Insurance Companies, 2015, MSc Thesis.

[49] OECD, *Insurance regulation and supervision in Asia*, OECD, 2000, ISBN 9789264171589,9264171584.

[50] S. Oktar and S. Yüksel, Bankacılık krizlerinin erken uyarı sinyalleri: Türkiye üzerine bir uygulama, İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, 2015.

[51] G. Rebala, A. Ravi, and S. Churiwala, *An Introduction to Machine Learning*, Springer, 2019.

[52] Republic of Turkey, No. 5684 Insurance Law, 14.6.2007, `http://www.resmigazete.gov.tr/eskiler/2007/06/20070614-2.htm`.

[53] Republic of Turkey Prime Ministry (Undersecretariat of Treasury), Regulation on technical provisions of insurance, reinsurance and pension companies, and assets on which such provisions are to be invested, 07.08.2007, `https://www.tsb.org.tr/Default.aspx?pageID=987`.

[54] Republic of Turkey Prime Ministry (Undersecretariat of Treasury), Regulation on measurement and assessment of capital requirements of insurance and reinsurance and pension companies, 19.01.2008, `https://www.tsb.org.tr/Document/Yonetmelikler/Sermaye_yeterliligi-son.pdf`.

[55] Republic of Turkey, Public Oversight Accounting and Auditing Standards Authority, Turkish Accounting Standards 40 - Investment Property, 17.03.2006, `https://kgk.gov.tr`.

[56] H. W. Rubin, *Dictionary of Insurance Terms*, Barron's Business Guides, sixth edition, 2000.

[57] R. Sahajwala and P. Van den Bergh, *Supervisory risk assessment and early warning systems*, Citeseer, 2000.

[58] M. Sarıaslan, *Insurance Accounting*, İstanbul: On İki Levha Yayıncılık AŞ, 2015.

[59] I. G. Sharpe and A. Stadnik, A statistical early warning model of financial distress in australian general insurers, Australian Prudential Regulation Authority, 2006.

[60] H. Stoppiglia, G. Dreyfus, R. Dubois, and Y. Oussar, Ranking a random feature for variable and feature selection, Journal of Machine Learning Research, 3(Mar), pp. 1399–1414, 2003.

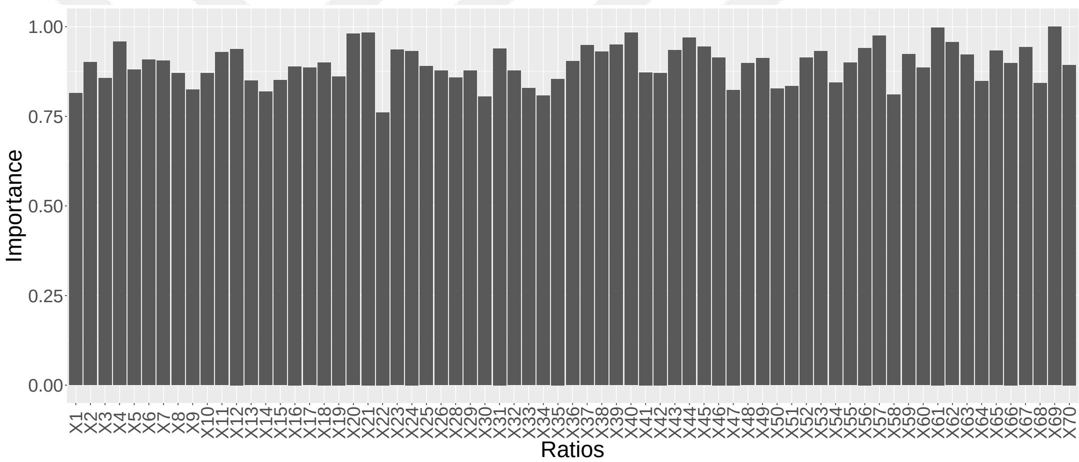
[61] The National Association of Insurance Commissioners, Insurance regulatory information system ratios manual, 2018, `https://www.naic.org/prod_serv/UIR-ZB-18.pdf`.

[62] E. T. Tornoa and T. S. Tiub, An early warning system on the propensity of survival and failure of non-life insurance firms in the philippines, Journal of Business and Finance, 2(01), p. 2014, 2014.

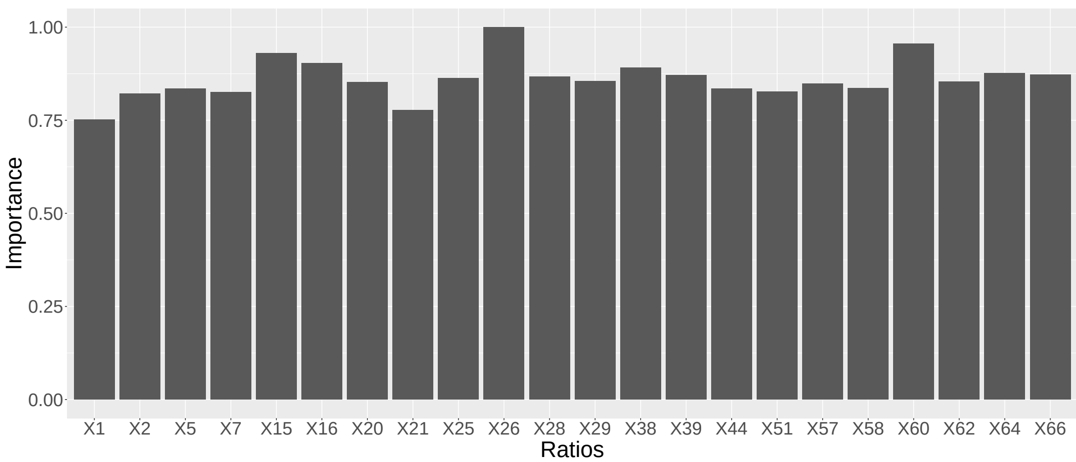
APPENDIX A

VARIABLE IMPORTANCE

In this section, the importance ranking is plotted and presented for Neural Networks, Gradient Boosting Machine, and XGBoost methods implemented. We can explain the variables of the importance of all methods used, and their analyses with the Boruta data set in order to increase the explanatory power of used machine learning methods. The significance level of the variables of Random Forest is shown in Chapter 4.



(a) Full Data

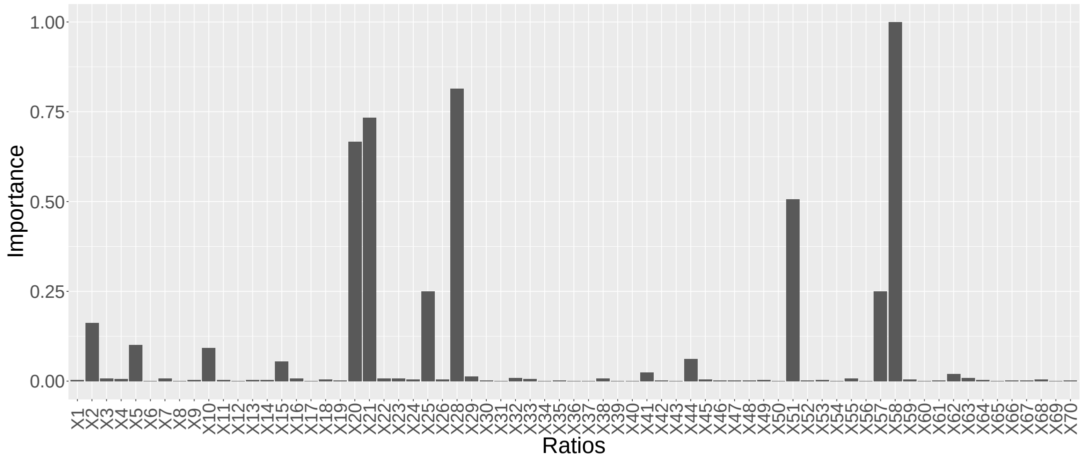


(b) Boruta Data

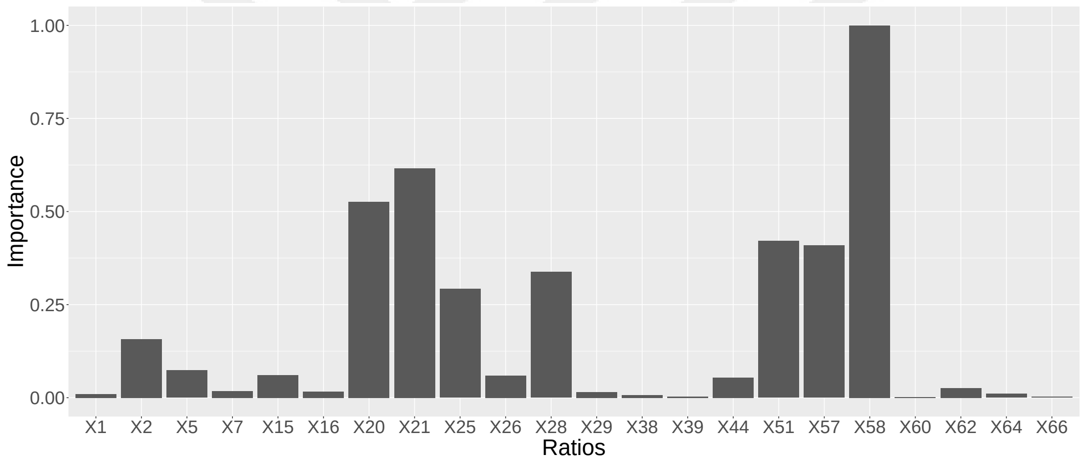
Figure A.1: Variable Importance for Neural Networks with (a) Full Data and (b) Boruta Data

Neural Networks procedure is weaker than others about the explanation of feature importance. As can be seen in Figure A.1a and Figure A.1b, the significance of all variables entering the model is not very different from each other.

Variable importance for Gradient Boosting Machine with two data sets is shown in Figure A.2.



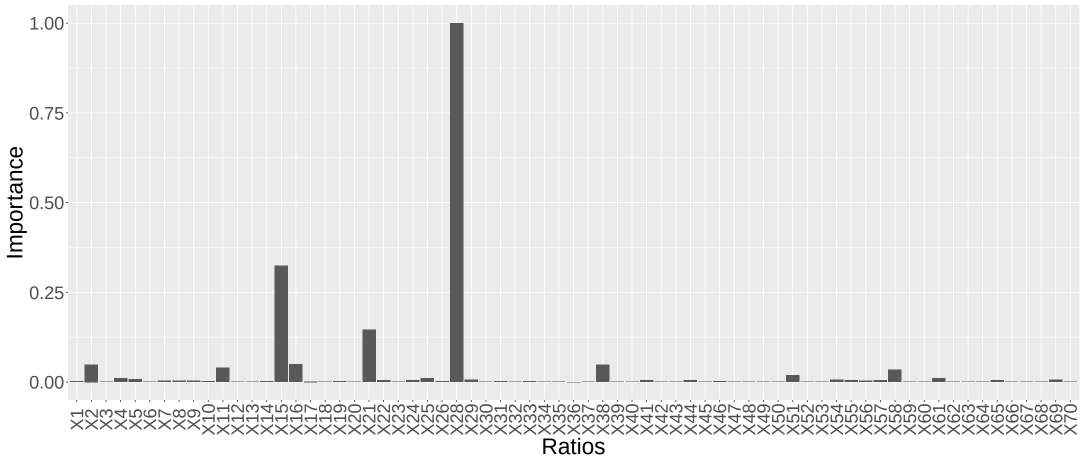
(a) Full Data



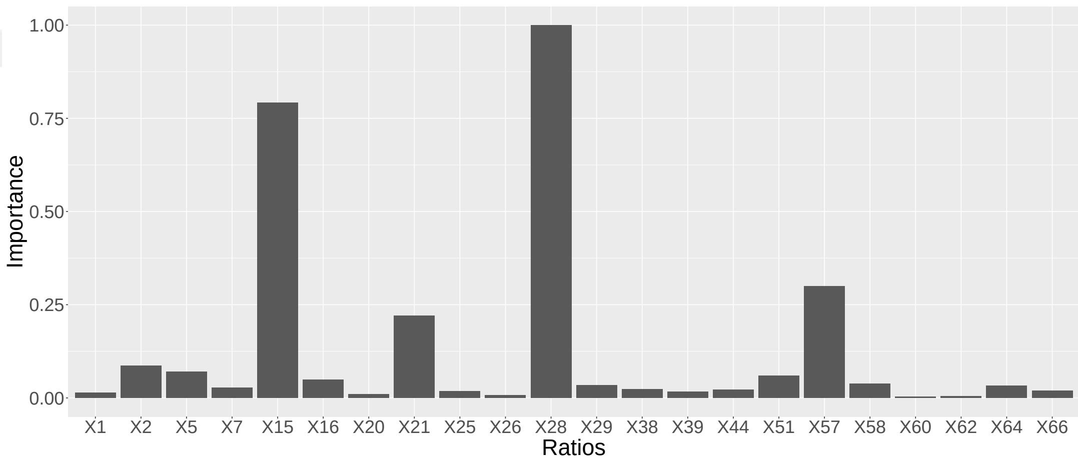
(b) Boruta Data

Figure A.2: Variable Importance for Gradient Boosting Machine with (a) Full Data and (b) Boruta Data

Variable importance for XGBoost with two data sets is shown in Figure A.3.



(a) Full Data



(b) Boruta Data

Figure A.3: Variable Importance for XGBoost with (a) Full Data and (b) Boruta Data