risks

MDPI

Article

Estimating the BIS Capital Adequacy Ratio for Korean Banks Using Machine Learning: Predicting by Variable Selection Using Random Forest Algorithms

Jaewon Park [1], Minsoo Shin [1] and Wookjae Heo [2,*]

[1] Division of Management Information System, College of Business Administration, Hanyang University, Seoul 04763, Korea; jwhanyang@hanyang.ac.kr (J.P.); minsooshin@hanyang.ac.kr (M.S.)
[2] Department of Consumer Sciences, College of Education and Human Sciences, South Dakota State University, Brookings, SD 57007, USA
* Correspondence: wookjae.heo@sdstate.edu; Tel.: +1-605-688-5835

Abstract: The purpose of this study is to find the most important variables that represent the future projections of the Bank of International Settlements' (BIS) capital adequacy ratio, which is the index of financial soundness in a bank as a comprehensive and important measure of capital adequacy. This study analyzed the past 12 years of data from all domestic banks in South Korea. The research data include all financial information, such as key operating indicators, major business activities, and general information of the financial supervisory service of South Korea from 2008 to 2019. In this study, machine learning techniques, Random Forest Boruta algorithms, Random Forest Recursive Feature Elimination, and Bayesian Regularization Neural Networks (BRNN) were utilized. Among 1929 variables, this study found 38 most important variables for representing the BIS capital adequacy ratio. An additional comparison was executed to confirm the statistical validity of future prediction performance between BRNN and ordinary least squares (OLS) models. BRNN predicted the BIS capital adequacy ratio more robustly and accurately than the OLS models. We believe our findings would appeal to the readership of your journal such as the policymakers, managers and practitioners in the bank-related fields because this study highlights the key findings from the data-driven approaches using machine learning techniques.

Keywords: bank; Bayesian regulatory neural network; random forest algorithms; BIS capital adequacy ratio; capital adequacy; machine learning

check for updates

Citation: Park, Jaewon, Minsoo Shin, and Wookjae Heo. 2021. Estimating the BIS Capital Adequacy Ratio for Korean Banks Using Machine Learning: Predicting by Variable Selection Using Random Forest Algorithms. *Risks* 9: 32. https://doi.org/10.3390/risks9020032

Academic Editor: Mogens Steffensen
Received: 7 December 2020
Accepted: 14 January 2021
Published: 1 February 2021

1. Background

The Bank for International Settlements' (BIS) capital adequacy ratio is an internationally accepted key indicator for measuring a bank's capital adequacy, which is defined and suggested by the BIS regardless of local financial, regulatory systems, policies, and laws (Bank for International Settlements 2020). BIS is an international financial institution that is owned by 63 central banks over countries (Bank for International Settlements 2020). BIS has been established in 1930 and have these missions (Bank for International Settlements 2020): (a) collaborating with participating central banks; (b) promoting financial stability; (c) supporting research about financial stability; (d) being prime counterparty for financial transactions; and (e) connecting international financial operations. As the missions imply, BIS is made for stabilizing the international finances as international efforts by participating central banks.

The BIS capital adequacy ratio represents prudent finance regulation indexes for monitoring domestic banks' capital levels and raised capital (National Law Information Center 2020a). It is also an important macro index against systemic risks because banks have the most important position in the financial system, and it control the highest percentage of payment systems.

The BIS capital adequacy ratio in South Korea is calculated with the formula of the equity capital ratio, which is the percentage of equity capital divided by the risk-weighted assets (RWA). Under Korean financial supervisory institutions' regulations for calculating RWA, domestic banks can select one of the standard methods or an internal ratings-based (IRB) approach for credit RWA and market RWA. The calculation of operation RWA can be selected among the basic indicator approach, standardized approach, and advanced measurement approach (National Law Information Center 2020b). Credit risk is the largest and most significant risk that banks face as a current or a potential risk. Credit risk directly represents a loss to a bank's net income and capital when the borrower does not fulfill the terms of the contract or does not pay the debt to the bank. Regarding of credit risk, most of the domestic banks also use the IRB approach, such as Shinhan Bank, KB Kookmin Bank, Nonghyup Bank, KEB Hana Bank, and Woori Bank (Fsc.go.kr 2020).

Total RWA designates credit risk, market risk, and operation risk as the three main categorized risks. Credit risk means the risk of a potential economic loss that is likely to be incurred by a debtor's default or an obligator's failure to fulfill a contract or an obligation. Operational risk means the risk of loss that is likely to be incurred as a result of inappropriate internal processes, human resources, or an external event, which could include legal risk (National Law Information Center 2020e). Market risk means the risk of losses incurred in a bank's on and off balance-sheet positions from movements of market prices. Market risk is further divided into general market risk and specific risk-weighted assets (National Law Information Center 2020c).

However, the current methodologies need some improvements in relation to the denominator of the BIS ratio. There have been many questions and recommendations about credibility and comparability because the standard approach did not reflect well on high-risk discrimination of RWA, and the IRB has shown considerable gaps for each bank (Fss.or.kr 2020). The IRB approach with high autonomy has a problem when calculating risk assets that are relatively small compared to the standard approach. This can cause the banks to be more likely to comply with regulatory ratios by reducing the size of RWA with the IRB approach if capital costs rise while banks' profitability deteriorates (Korea Institute of Finance (Korea Institute of Financia 2016)). For this reason, regulatory reform has required improvements of both the standard approach and IRB. This reform focuses on sensitivity, simplicity, and comparability (Korea Institute of Financia 2016). Correspondingly, Korea's FSC has been reviewing regulatory improvements according to Basel III Finalizing Post-Crisis Reforms, issued in December 2017, as proposed by The Basel Committee on Banking Supervision ([BCBS]; Fss.or.kr 2020).

Under these circumstances, this study investigated data from the past 12 years, including all the information of domestic banks in South Korea. As a final point, this study shows the most important indicators representing the BIS capital adequacy ratio as a supplement index. Considering the large number of variables (1929) and strong prediction performance, machine learning techniques were used. Specifically, the dynamic ecological systemic framework suggested by Heo (2020) was used as the basic assumption of utilizing all 1929 variables as predictors for the BIS capital adequacy ratio. Heo explained that large numbers of predictors from multiple systems around an organization are interacting and influencing the managerial performance of the organization. Machine learning techniques, such as artificial neural networks (ANN), are suggested as the main methodology for the dynamic ecological systemic framework due to the large number of predictors used as dynamic ecological factors (Heo 2020).

Regarding these points, the main purpose of this study was to predict the BIS capital adequacy ratio more accurately. Accordingly, this study has two sub-purposes. First, this study determined which variables are the most important predictors for the BIS capital adequacy ratio. Second, all the variables were analyzed for better prediction by machine learning since using traditional statistical approaches for such a large number of variables was not possible. In short, the study confirmed the prediction performance of selected variables to determine whether the machine learning technique has a better prediction performance than the classical linear prediction method: Bayesian Generalized Linear Regression Model (BGLM).

2. Literature Review

2.1. Financial Stability and BIS Capital Adequacy Ratios of Korean Banks

The BIS capital adequacy ratio regulation was implemented by BCBS under the BIS in July 1988. To secure the banks' capital adequacy, the banks must hold at least 8% of RWA as equity capital. Korean banks introduced BIS I in July 1992. In January 2008, BIS II was introduced to clarify the characteristics of the risk-oriented equity capital regulation. On 1 December 2013, BIS III was introduced to enhance the quality and quantity of the capital and introduced the capital buffer concept and liquidity expansion. In the beginning, the BIS capital adequacy ratio was introduced for the concept of risk and for the regulation of equity capital, starting with credit risk. It continues to expand to include market risk, operation risk and elaborate requirements for accreditation of the IRB approach. Equity capital regulation is playing a role not only as a supervisory tool to improve the financial soundness of companies, but also as an administrative regulation tool (Fss.or.kr 2020).

2.2. Korean Laws and Regulations about BIS Capital Adequacy Ratio

Korean laws and regulations stipulate that all banks must follow the Bank Act and enforce its decrees, including the Regulations on Supervision of Banking Business and the Detailed Regulations on Supervision of Banking Business. The laws and regulations support the BIS capital adequacy ratio with management guidance standards and with the principle of bank soundness. All banks must comply with the management guidelines set by the Financial Services Commission (FSC) as prescribed by presidential decrees in relation to capital adequacy to maintain sound management. While the FSC establishes management guidelines, the principles of bank soundness supervision recommended by the BIS must be fully included (National Law Information Center 2020a).

Enforcement of the Banking Act stipulates that the standards for equity capital ratio corresponding to a bank's credit risk should be included in the standards for management guidelines. The FSC can ask for necessary measures for management improvement when there is a risk of seriously harming the soundness of management, such as failing to meet the management guidance standards according to the BIS capital adequacy ratio or if it is deemed inevitable for maintaining the soundness of management (National Law Information Center 2020g).

Regarding the BIS management guidance ratio, it is stipulated in the regulations on the supervision of the banking business. All banks must maintain a management guidance ratio corresponding to the minimum for the capital ratio; a common equity capital ratio of 4.5/100; a core capital ratio of 6/100; and a total capital ratio of 8/100. Each calculation methodology is determined by the governor of the Financial Supervisory Services (FSS), and it needs to follow the standard guidelines recommended by the BIS (National Law Information Center 2020h). Accordingly, when the equity capital ratio falls under the BIS management guidance ratio as a result of evaluation and analysis of management status, timely corrective actions are triggered as one of the three stages: management improvement recommendations, management improvement requirements, and management improvement orders. Timely corrective measures are mandatory regulations. These must be taken by the governor of the FSS or the FSC when the requirements for initiation are met (National Law Information Center 2020i).

The first stage of timely corrective measures is management improvement recommendations by the FSC. This action is required that the bank has a total capital ratio of less than 8/100 or a core capital ratio of less than 6.0/100 or a common equity capital ratio of less than 4.5/100 (National Law Information Center 2020j). The second stage of timely corrective measure is management improvement requirements by FSC. This action requires that a bank has a total capital ratio of less than 6/100 or a core capital ratio of less than 4.5/100 or a common equity capital ratio of less than 3.5/100 (National Law Information

Center 2020k). The final stage of timely corrective measures is Management Improvement Order by the FSC. These actions require that a bank has a total capital ratio of less than 2/100 or a core capital ratio of less than 1.5/100 or a common equity capital ratio of less than 1.2/100. The FSC can order the bank to take necessary actions and the governor of the FSS also has to monitor execution after receiving necessary actions within two months as determined by the FSC (National Law Information Center 2020l).

2.3. Predicting Financial Ratio with Machine Learning Techniques

As such, the credit risk of bank such as BIS is better to be expected by some tools such as prediction modeling. Recently, machine learning techniques were utilized in predicting credit risk and BIS (e.g., Gambacorta et al. 2019; Petropoulos et al. 2019). Gambacorta et al. (2019) utilized Chinese fintech company data and found the effectiveness of using machine learning technique (i.e., fintech credit scoring models that was combined with big data) for predicting the loss in banks. Petropoulos et al. (2019) used data from Greek banking system and forecasted the credit quality. Specifically, considering that credit risk is the foundational concept of BIS, this literature implies that machine learning techniques are useful for prediction in credit issues in banks. Beyond the credit issues of banks industry, it is valuable to utilize the machine learning technique in larger realm such as assessing the various risks (i.e., credit risk, market risk, and operation risk) of the banks industry (i.e., BIS).

In addition, Bazarbash (2019) introduced the strength and the weakness of using machine learning in predicting the credit risks. Even though there are many strengths of using machine learning (e.g., better in classification, good for forecasting, and predicting the risks), the machine learning technique is still having weakness so called as black box. It meant that, by using machine learning, it is not possible to know which factors are precisely predicting the outcome. Therefore, in this study, three steps of machine learning were utilized to check which factors performed to predict BIS. Specifically, the first two steps are expected to sort out the important variables from the total predictors. As a result, the three steps of sorting the variable lists make the important variables to predict the BIS of banks.

To sum up, the uniqueness of this study that is different with the previous literature are three: (a) assessing the risks in the macro-size level with banks' data, (b) utilizing multiple steps of machine learning to find the important orders of predictors (i.e., overcoming black box issue), and (c) using South Korea banks data that were not predicted yet.

3. Theoretical Background: BIS Capital Adequacy Ratio

Standards for the Calculation of Management Guidance Ratios are stipulated in the Detailed Regulations on the Supervision of Banking *Business*, which follows the BIS Rule of the Bank for International Settlement (National Law Information Center 2020b). According to The Principles for the Calculation of Equity Capital Ratio Based on BIS I, the ratio of capital to risk-weighted assets must be calculated based on the capital and gross assets on the consolidated balance sheet. When calculating this ratio, a bank must classify the capital ratio based on the credit risk and the combined risks according to the gravity of the risks. The capital ratio based on the market risk must be calculated, where a bank whose ratio of the sum of the trading book to the total assets on the consolidated balance sheet is at least a 5% maximum per day, or where a bank whose sum of trading book is 100 billion Won maximum per day. As a result, the formula for the calculation of equity capital ratio based on credit risk, BIS I is calculated by Function (1) (National Law Information Center 2020e).

$$Equity\ capitalratio\ based\ on\ credit\ RWA\ = \frac{Equity\ capital}{Credit\ risk - weighted\ assets} \times 100 \geq 8\% \quad (1)$$

The principles for the calculation of the equity capital ratio based on BIS II are as follows: banks calculate the capital ratio based on credit and operational risks, and separately calculate the capital ratio based on credit, operation and market risks. The equity capital ratio based on market risk is calculated according to the same criteria as the market risk based on BIS I. The capital ratio based on credit and operation risks must be calculated by dividing the equity capital by the sum of the risk-weighted assets aggregated credit RWA, operational RWA, and the risk evaluation adjustment. Equity capital must be calculated by aggregating core capital and supplementary capital then subtracting deductions as 50% from both of the core capital and supplementary capital, except as otherwise prescribed expressly. As a result, the BIS II is calculated by Function (2) (National Law Information Center 2020f).

$$Equity\ capital\ ratio\ based\ on\ credit\ and\ operation\ risks = \frac{Equity\ capital}{Risk{-}weighted\ assets + Risk\ evaluation\ adjustment} \times 100 = \frac{Core\ capital + Supplementary\ capital - Deductions}{Credit\ risk{-}weighted\ assets + Operational\ risk{-}weighted\ assets + Risk\ evaluation\ adjustment} \times 100 \quad (2)$$

Principles for the calculation of the equity capital ratio based on BIS III are as follows. The capital ratio based on credit and operational risks should be calculated by dividing capital calculated by the formula. Total capital should be calculated by aggregating Tier 1 capital (the aggregate of Common Equity Tier 1 and Additional Tier 1 capital) and Tier 2 capital then dividing by the sum of the risk-weighted assets aggregated credit risk-weighted assets, operational risk-weighted assets, and the risk evaluation adjustment. BIS III is calculated by the ratio of Common Equity Tier 1, the ratio of Tier 1 capital, and the ratio of total capital, depending upon the scope of components of each capital. As a result, BIS III is calculated by Functions (3)–(5) (National Law Information Center 2020d).

$$Ratio\ of\ Common\ Equity\ Tier\ 1\ based\ on\ credit\ and\ operational\ risks = \frac{Common\ Equity\ Tier\ 1}{Credit\ risk{-}weighted\ assets + Operational\ risk{-}weighted\ assets + Adjustment\ of\ risk\ assessment} \times 100 \quad (3)$$

$$Ratio\ of\ Tier\ 1\ capital\ based\ on\ credit\ and\ operational\ risks = \frac{Tier\ 1\ capital}{Risk{-}weighted\ assets + Adjustment\ of\ risk\ assessment} \times 100 = \frac{Common\ Equity\ Tier\ 1 + Additional\ Tier\ 1\ capital}{Credit\ risk{-}weighted\ assets + Operational\ risk{-}weighted\ assets + Adjustment\ of\ risk\ assessment} \times 100 \quad (4)$$

$$Ratio\ of\ total\ capital\ based\ on\ credit\ and\ operational\ risks = \frac{Total\ capital}{Risk{-}weighted\ assets + Adjustment\ of\ risk\ assessment} \times 100 = \frac{Common\ Equity\ Tier\ 1 + Additional\ Tier\ 2\ capital}{Credit\ risk{-}weighted\ assets + Operational\ risk{-}weighted\ assets + Adjustment\ of\ risk\ assessment} \times 100 \quad (5)$$

4. Statistical Background: Machine Learning Algorithms

This study sought to find the key variables from 1929 variables that represent the banks' BIS forecast. Machine learning algorithms in this study consisted of three steps: Random Forest Boruta, Random Forest Feature Elimination, and the Bayesian Regularized Neural Network (BRNN), See Figure 1. Regarding feature selection, it is known that many machine learning algorithms reduce prediction accuracy when using large numbers of variables compared to the optimal one (Kohavi and John 1997). It is also advisable to choose a small feature set that yields the best possible classification result according to minimal-optimal problem (Nilsson et al. 2007). In some situations, it is also important to find all the relevant variables to the classification by all-relevant problem (Nilsson et al. 2007).

Considering this practical point of view, Random Forest was used in this study for precise predictions in cases of large numbers of variables (Genuer et al. 2010). The Random Forest technique represents an early type of machine learning data classification modeling process (Breiman 2001). This technique is based on ensemble learning, which is a procedure that attempts to merge two or more algorithms into one algorithm. In the case of a random forest, the decision tree, average prediction, and bagging sample are all merged into a single algorithm (Ho 1998; Liaw and Wiener 2002).

Kursa and Rudnicki (2010) improved the variable selection performance of the Random Forest algorithm, which is now called the Boruta algorithm. In addition, it is known for the importance of ranking the predictor variables can be found by utilizing a combination of the same family of algorithms (e.g., Random Forest), if the predictor variables are numerous (1939 predictors) (Zhou et al. 2014). Thus, Random Forest RFE and a random forest-based recursive variable removal method, is used as an algorithm that predicts weights to rank predictors in ascending order (Guyon et al. 2002).

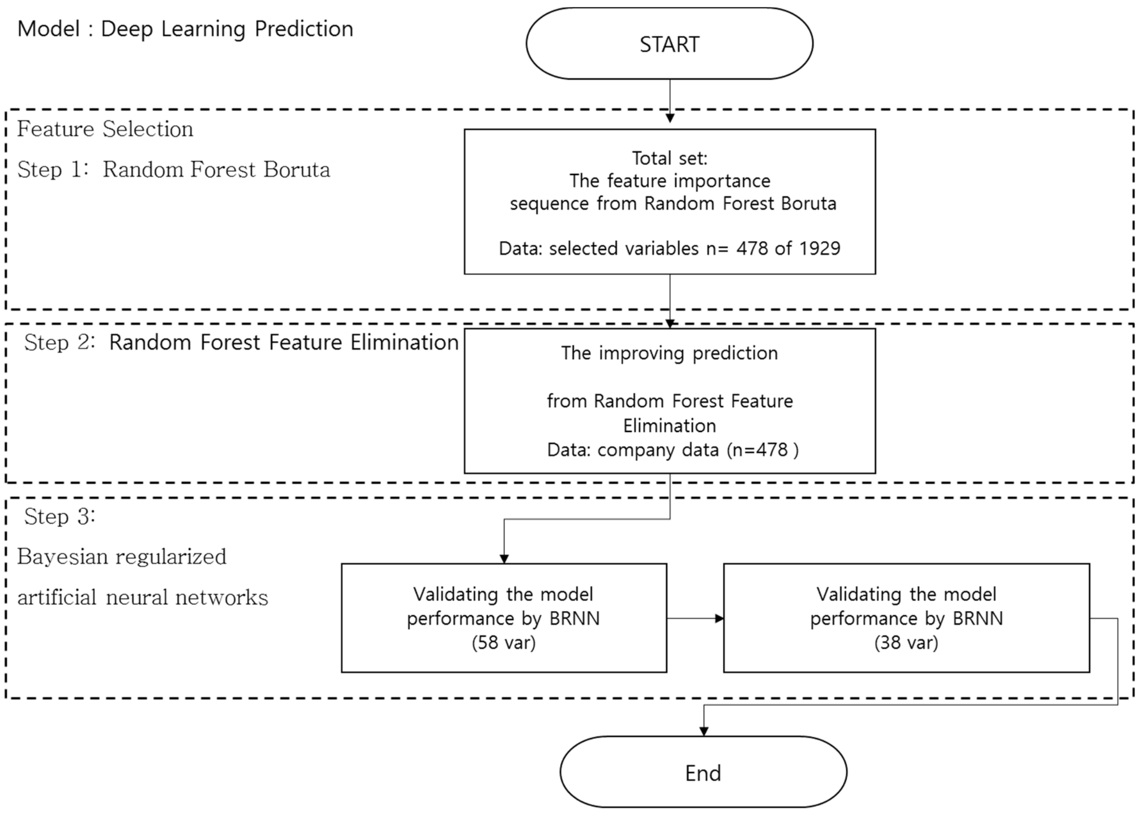

Figure 1. Analytic Algorithm Methodology: Deep Learning Prediction and Regression Prediction Algorithms.

In the first stage, the Random Boruta algorithm analyzes 1929 variables from the financial and non-financial information of domestic banks. As a result of the analysis, 489 variables were identified. In the second stage, the root mean of squared error (RMSE) and mean of absolute error (MAE) were analyzed by Random Forest RFE. As a result, 58 and 38 variable sets were defined by importance rank. In the final stage, BRNN algorithms confirmed the statistical validity of the predictive power of the variables found in Random Forest RFE.

BRNN is a type of artificial neural network (ANN) (Sariev and Germano 2020), which estimates the efficient output of nonlinear data without overfitting and overtraining (Burden and Winkler 2008). BRNN is an advanced algorithm based on ANN, a type of feed forward neural network (Sariev and Germano 2020). The feed forward neural network assumes that the gradient process is expected only through an iterative process. The BRNN is also based on the Gauss-Newton approximation using Bayesian estimation and normalization approaches (MacKay 1992). Considering this normalization process, BRNN is superior to existing prediction models for predicting dependent variables composed of continuous variables (Sariev and Germano 2020).

5. Methods

The data for this study are the statistical information from the FSS, which is being updated quarterly by the Korean FSS. This is 12-year time series data from 2008 to 2019 covering all domestic banks. The data include all financial and non-financial variables of the bank, including general information, financial information, key management indices, major business activities, and metadata. Specifically, all information was downloaded from FSS website that were transformed to numeric variables. Total number of downloadable information were 1933. Among those 1,933 variables, company name was used as identification of analytic unit; year and quarter were used as time indicator; and BIS was used as outcome variable. Therefore, the total number of predictors at the initial procedure was 1929.

Domestic banks are classified into general banks established under the Banking Act and special banks established under the individual special Bank Act. The total number of banks in the FSS dataset were eighteen banks. Among those eighteen banks, sixteen banks were utilized in this study and two banks were excluded like following explanation.

First, twelve banks established in accordance with the Banking Act include commercial banks, local banks, foreign banks, domestic branches, and Internet banks (i.e., KEB Hana Bank, Shinhan Bank, Kookmin Bank, Woori Bank, Korean Bity Bank, SC Bank, Busan Bank, Kyungnam Bank, Daegu Bank, Gwangju Bank, Citi bank, and Jeonbuk Bank). Second, five banks as local banks have their own businesses of deposits, loans, and payment settlements (i.e., IBK, Korea Development Bank, Korea Export-Import Bank, Nonghyup, and Suhyup). These five banks were established in accordance with individual special laws to supply funds to certain sectors where it is difficult to supply sufficient funds due to financial constraints and difficulties in securing profitability. Therefore, there were seventeen banks were analyzed in this study.

However, this study excluded Internet banks such as Kakao Korea Bank and K-Bank that started operating in 2016 and 2017 after the approval. As the first step of data analysis, the Boruta algorithm analyzed 1929 predictors representing the financial and non-financial indicators of banks, and 489 variables were identified as significant variables, excluding the output value of avg BIS and the BIS capital ratio, which is the default value. As the second step, the Random Forest Feature Elimination Method was used to analyze these variables and find the optimal selection by order of importance for the volatility of root mean of squared error (RMSE) and mean absolute error (MAE) after inputting the selected 478 variables through the Boruta algorithm. As the third step, the Bayesian Regularized Neural Network (BRNN) showed the comparison results against traditional prediction models to examine the statistical validity. Specially, the results of this study compared the deep learning algorithm to the classical linear prediction by the Bayesian Generalized Linear Regression Model (BGLM) by RMSE and MAE results, since it cannot be directly compared due to the difference of model results. RMSE and MAE are commonly used by researchers as criteria for evaluating predictive models (Hyndman and Koehler 2006; Wooldridge 2016). In general, it means that the higher the RMSE, the lower the model's accuracy, and the higher the MAE, the lower the accuracy.

6. Results

6.1. Stage 1. Feature Selection Using Random Forest Boruta

A total of 1929 variables were analyzed using the Random Forest Boruta technique after being classified as important, tentative, or unimportant in the primary classification criteria. As shown in Table 1 and Figure 2, the results of the analysis classified 299 as important, 465 as tentative, and 1165 as unimportant. The results by the second classification criteria showed 478 as important and 1451 as unimportant, excluding the 299 important variables. Finally, 478 variables were selected as significant predictors of the BIS capital adequacy ratio.

Table 1. Variables Predicting Bank for International SettlementValues Using the Random Forest Boruta Method.

- The 1299 variables were selected as 299 important, 1165 unimportant, and 465 tentative. Boruta performed 99 iterations in 23.7025 min.
  (1) 299 attributes confirmed as important: avg BIS, Tier 1 Capital Ratio, Household Money-In-Trust (amount), Household Loans Total, Households and 294 more.
  (2) 1165 attributes confirmed as unimportant: Retained earnings; loan loss reserve; expected transfer amount—Reserve for credit loss; Unaccumulated carryover amount; Additional transfer amount, Automated Teller Machine, Certificate of Deposit, Household Loans, Household Loans Depreciation and 1160 more variables.
  (3) 465 tentative attributes left: Accumulated Other Comprehensive Income\\Allowance, Certificate of Depsit, Automated Teller Machine, Household Money Trust Percentage, Household Loans Allowance, Household Loans for Repayment, and 460 more variables.
- Machine learning analysis was performed by reclassifying the secondary variable selection criteria into important and unimportant. Using 1165 unimportant and 465 tentative, Boruta performed 99 iterations in 23.7025 min. Tentative roughness fixed over the last 99 iterations.
  (1) 478 attributes confirmed as important: avgBIS, Tier 1 Capital Ratio, Household Money-In-Trust (amount), Household Loans Allowance, Household Loans Total, and 486 more.
  (2) 1451 attributes confirmed unimportant: Accumulated Other Comprehensive Income Allowance, Retained Earnings, Loan Loss Reserve, Expected Transfer Amount - Reserve for Credit Loss, Unaccumulated Carryover Amount, Additional Transfer Amount, ATM, CD, CDATM, and 1445 more variables.

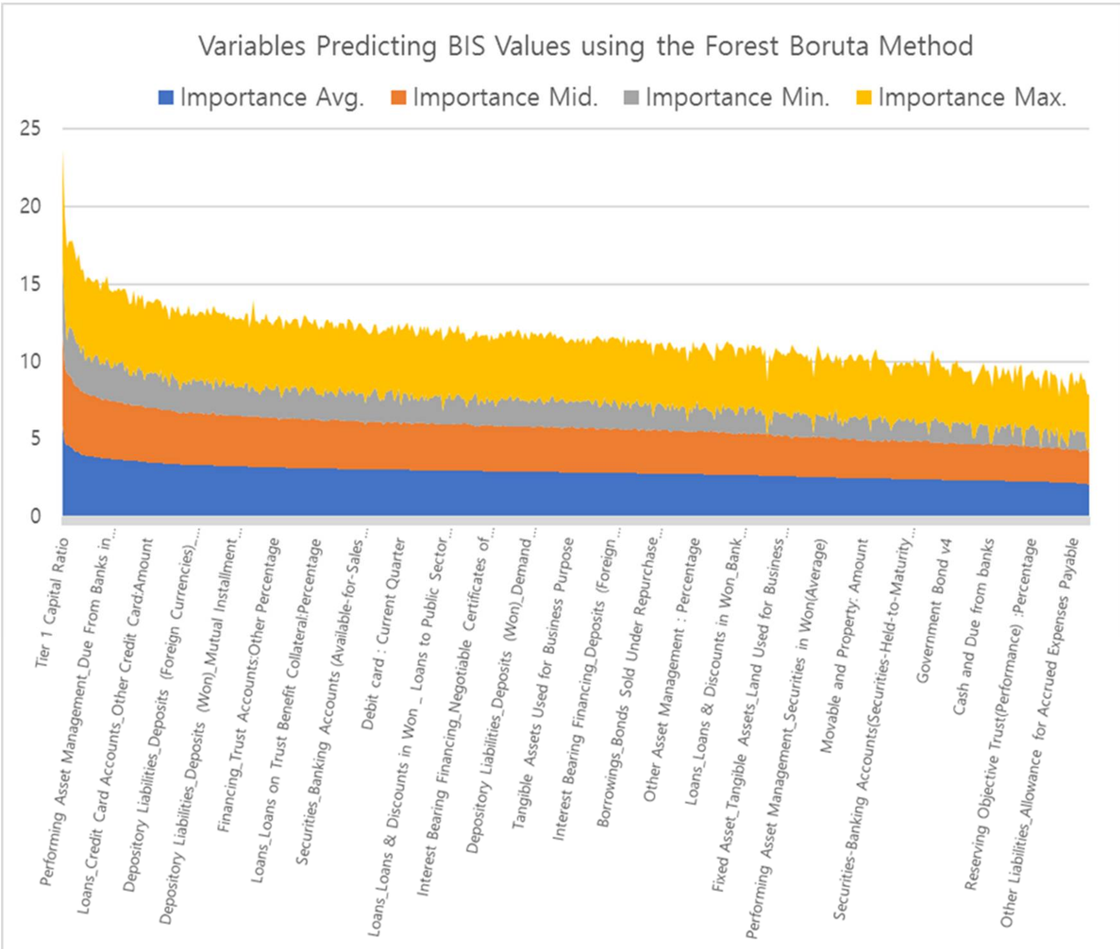

Figure 2. Significant Variables Using the Random Forest Boruta Method.

6.2. Stage 2. Feature Selection by Importance Rank Using Random Forest Feature Elimination

The Random Forest Feature Elimination Algorithm was used for marginal predictions of 478 significant variables selected through Random Forest Boruta. The variables were sequentially introduced in Random Forest Feature Elimination Algorithm to analyze the volatility of RMSE and MAE. As a result, 58 important variables and 38 most important variables were selected as determinants representing the BIS prediction as shown in Figure 3. The list of 38 variables and 58 variables were discussed in the discussion.

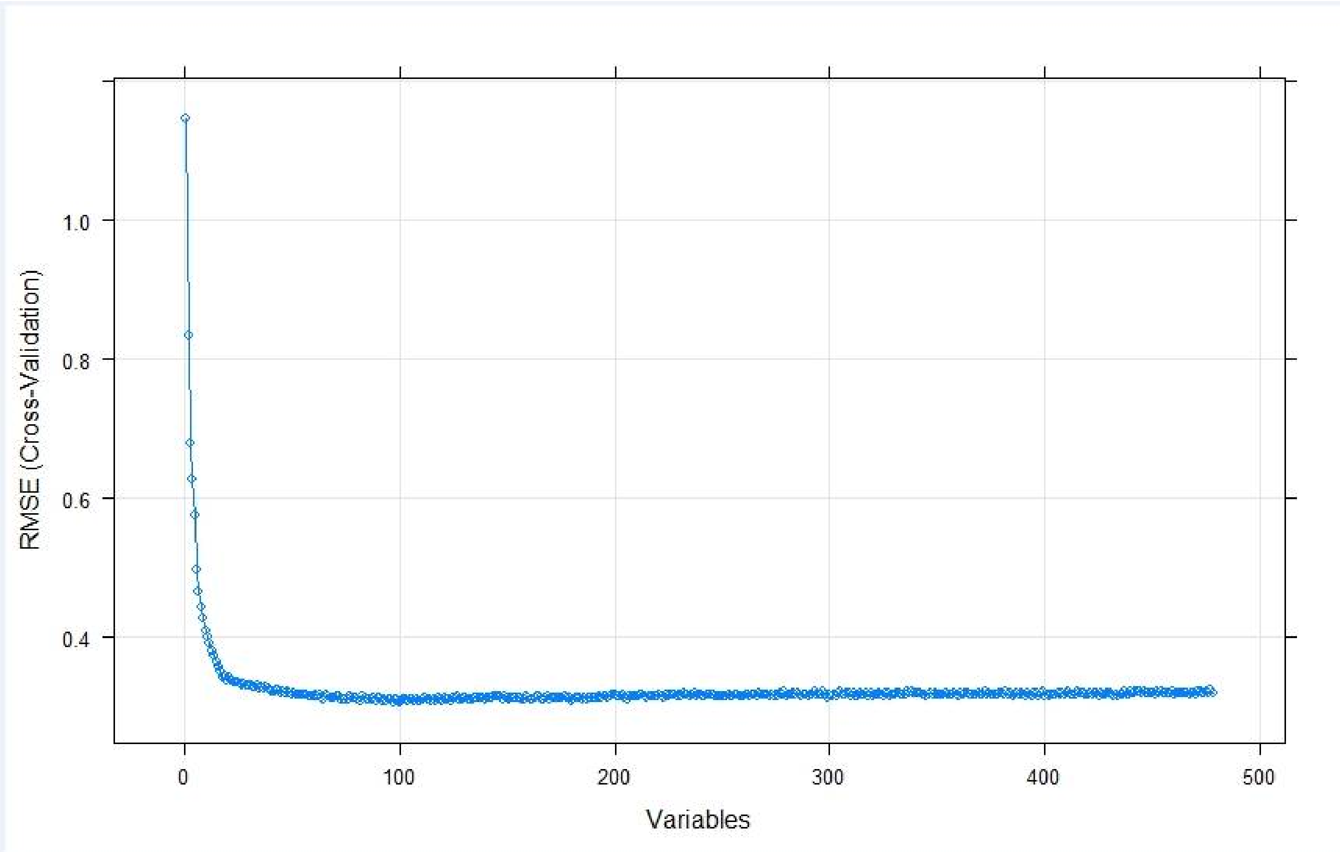

Figure 3. Importance of Variables Using Random Forest Recursive Feature Elimination Method.

6.3. Stage 3. BIS Prediction Using Bayesian Regularized Neural Network Model

To examine the statistical feasibility with the accuracy, this research compared the performance of the deep learning modeling technique and the Bayesian Regularized Neural Network model to the comparison model, which was BGLM, based on the results error of RMSE and MAE.

All models were resampled 50 times and iterated nine times. Specifically, 50 folds resampling procedure per each iteration were utilized as cross validation in the study. The multiple times of folds for resampling are expected to eliminate the overfitting issue of the combination between Random Forest Feature Elimination and BRNN. Specifically, the comparison of indicators between training model and testing model is expected to indicate whether the overfitting issue occurred in this study. As shown in Table 2, the differences of indicators (i.e., MAE and RMSE) between training model and testing model were not big across nine iterations. Therefore, the combination of Random Forest and BRNN did not have overfitting issue.

Table 2. Difference of Indicators between Training Model and Testing Model.

Iteration	Indicator	BRNN With 38 vars		BRNN With 58 vars	
		Training Model	**Prediction Model**	**Training Model**	**Prediction Model**
1	Neuron #	20	20	16	16
	MAE	0.4008	0.4059	0.3807	0.4072
	RMSE	0.5265	0.5737	0.5288	0.5775
	R2	0.8901		0.8937	
2	Neuron #	4	4	3	3
	MAE	0.4475	0.3858	0.4461	0.3849
	RMSE	0.6214	0.5507	0.5801	0.5497
	R2	0.8577		0.8727	
3	Neuron #	5	5	3	3
	MAE	0.4742	0.4429	0.4815	0.4430
	RMSE	0.6551	0.8090	0.6360	0.8091
	R2	0.8270		0.8505	
4	Neuron #	4	4	10	10
	MAE	0.4330	0.4040	0.3878	0.4065
	RMSE	0.5617	0.6254	0.5445	0.6274
	R2	0.8808		0.8996	
5	Neuron #	18	18	3	3
	MAE	0.4177	0.3851	0.3961	0.3854
	RMSE	0.5788	0.5469	0.5383	0.5471
	R2	0.8731		0.8690	
6	Neuron #	16	16	20	30
	MAE	0.4437	0.4455	0.4082	0.3493
	RMSE	0.5975	0.6718	0.5575	0.5102
	R2	0.8374		0.8768	
7	Neuron #	3	3	10	10
	MAE	0.4394	0.3954	0.3688	0.4604
	RMSE	0.5701	0.5843	0.8042	0.6297
	R2	0.8694		0.8950	
8	Neuron #	4	4	20	20
	MAE	0.4114	0.3585	0.3752	0.3667
	RMSE	0.5413	0.5586	0.4988	0.5977
	R2	0.8893		0.9086	
9	Neuron #	3	3	10	10
	MAE	0.4539	0.3971	0.3502	0.4984
	RMSE	0.5771	0.5757	0.4764	0.7602
	R2	0.8752		0.9176	

Note. MAE denotes mean absolute error; RMSE is root mean of squared error; BRNN means Bayesian Regularization Neural Networks. Neuron # indicates the optimal number of neurons that was selected by machine learning. 50-folds resampling per each iteration was used as cross validation to calculate MAE, RMSE, and R^2.

The performance of nine times iteration was shown in Appendix A. As shown in Table 3, BRNN with 38 variables showed slightly better MAE and RMSE, but statistically, they were not significantly different from the t-test of MAE (See Table 4). BRNN with 38 variables showed significantly better prediction compared to the other two kinds of BGLMs from t-test of MAE. BRNN with 58 variables showed significantly better prediction compared to the other two kinds of BGLMs from t-test of MAE. BGLM with 58 variables showed slightly better MAE and RMSE, but statistically, they are not significantly different from the 70% of RMSE (See Table 5). Both BRNNs with 38 variables and 58 variables showed lower than 70% of BGLM with 38 variables from 70% of RMSE. Both BRNNs with 38 variables and 58 variables showed very close to 70% of BGLM with 58 variables.

Table 3. MAE and RMSE from 9 iterations of 340 companies.

	BRNN With 38 vars	BRNN With 58 vars	Bayesian GLM With 38 vars	Bayesian GLM With 58 vars
MAE (S.D.)	0.4022 (0.4663)	0.4113 (0.4776)	0.5477 (1.0433)	0.5119 (0.7689)
RMSE (S.D.)	0.6107 (0.0796)	0.6232 (0.0943)	1.000 (0.6229)	0.8583 (0.3412)

Note. GLM means generalized linear modeling.

Table 4. MAE Comparison: Mean Comparison (*t*-test, n = 340 companies * 9 iterations).

	BRNN (38 var)	BRNN (58 var)	BGLM (38 var)	BGLM (58 var)
BRNN (38 var)				
BRNN (58 var)	−0.7542			
BGLM (38 var)	−7.0432 ***	−6.5759 ***		
BGLM (58 var)	−6.7482 ***	−5.6656 ***	1.5280	

* $p < 0.05$; *** $p < 0.001$.

Table 5. RMSE comparison: 70% of compared model.

	BRNN (38 var)	BRNN (58 var)	BGLM (38 var)	BGLM (58 var)
RMSE	0.6107	0.6232	1.000	0.8583
70% of RMSE			0.700	0.6008

In addition, seeing the standard deviation of MAE and RMSE, the prediction from BGLMs is more fluctuated than the prediction from BRNNs regardless of the number of variables. With the standard deviation of MAE and RMSE, the prediction from 38 variables with BRNN showed lower fluctuation than the prediction from 58 variables with BRNN.

7. Discussion

This study analyzed 1929 variables out of all the financial and non-financial information that consisted of general information, financial information, key management indices, and major business activities of all domestic banks in South Korea. First, this study confirmed that 478 variables out of 1929 variables are the significant indicators affecting the total BIS ratio when using the Random Forest Boruta method. Second, 58 out of 478 variables were selected as important variables for representing BIS ratio analyzed by the Random Forest recursive feature elimination technique. The 38 variables in Table 6 were confirmed as the most important determinants for the prediction of the total BIS ratio. The 58 variables were introduced in Appendix B.

The following determinants were selected as the Top 10 key indicators of the 38 variables through importance ranking: Tier 1 Capital Ratio; Borrowings_Bonds Payable_(Discount Present Value):Percentage; Borrowings:Percentage; Acceptances and Guarantees Others; Acceptances and Guarantees; Borrowings_Bonds Payable: Percentage; Borrowings_Borrowings: Percentage; Receivable Charge-Offs; Other Liabilities_(Transfer from National Pension):Amount; and Fixed Asset_Tangible Assets Used for Business Purpose_(Accumulated Depreciation):Amount.

The construction of Loans Receivable (Industries) asset can be selected as a Top 10 variable by excluding a Tier 1 capital ratio representing default values. Loans Receivable (Industries) asset is composed of Agricultural, Forestry, Fishery; Manufacturing; Construction; Wholesale and Retail; Transportation; Hotels and Restaurants; Telecommunications; Real Estate, Leasing and Others for a total of eight items.

Table 6. 38 Important Predictors According to Random Forest RFE by Machine Learning Techniques.

Report (Financial Statics Information System)	**(Importance Rank) by Variance (Item) of Domestic Bank**
Capital Adequacy	(1) Tier 1 Capital Ratio
Consolidated Balance Sheet (Liabilities & Shareholders' Equity-Banking Account)	(2) Borrowings_Bonds Payable_(Discount Present Value):Percentage
	(3) Borrowings:Percentage
	(6) Borrowings_Bonds Payable:Percentage
	(7) Borrowings_Borrowings:Percentage
Loans Receivable (Industries)	(11) Construction
Off-balance Accounts (Bank Accounts)	(4) Acceptances and Guarantees Others
	(5) Acceptances and Guarantees
	(8) Receivable Charge-Offs
	(25) Derivative Contracts
Principal Sources of Cash Flows in Bank Accounts	(12) Financing Without Cost_Other Non-cost Bearing Financing:(Average)
	(13) Financing Without Cost_Provision for Other Allowances:Percentage
	(14) Performing Asset Management_Due From Banks in Won:(Average)
	(16) Financing Without Cost_Demand Deposits:Percentage
	(18) Financing With Cost_Borrowings in Won:Percentage
	(19) Non-Performing Asset Management_Others:Average
	(20) Performing Asset Management_Due From Banks in Won:Percentage
	(22) Financing Without Cost_Other Non-cost Bearing Financing:Percentage
	(24) Non-Performing Asset Management_Cash & Checks and Foreign Currency:Percentage
	(29) Asset Management for Benefit_Other Won-Denomiated Currency Asset Management:Average
	(38) Non-Performing Asset Management_Fixed Assets Used for Business Purposes:Percentage
Principal Sources of Cash Flows in Trust Accounts	(23) Operation_Loans & Discounts:Percentage
Profitability	(17) Loans in won _ Average Interest rate
	(33) Deposits in Won _ Average Interest rate
Summarized Balance Statement (Assets-Banking Account)	(10) Fixed Asset_Tangible Assets Used for Business Purpose_(Accumulated Depreciation):Amount
	(27) Loans_Loans & Discounts in Won Loans to Enterprise:Percentage
	(28) Fixed Asset_Tangible Assets_Buildings Used for Business Purpose:Amount
	(30) Securities_Banking Accounts (Available-for-Sales Securities) Available-for-Sales Securities in Won_Others: Amount
	(35) Loans_Credit Card Accounts_Cash Service:Percentage
	(36) Securities-Banking Accounts(Subsidiaries)_Equity Investment (Won) Consolidated Subsidiary Stock: Amount
Summarized Balance Statement (Assets-Trust Account)	(31) Bond Accounts:Amount
	(37) Loans & Discounts_Loans on Real Estate Collateral:Amount
Summarized Balance Statement (Liabilities & Trust Account)	(32) Personal Pension Trust:Percentage
Summarized Balance Statement (Liabilities &Shareholders' Equity-Banking Account)	(9) Other Liabilities_(Transfer from National Pension):Amount
	(15) Other Liabilities_Account for Agency Business_Grio Account:Amount
	(26) Other Liabilities_Allowance Accounts_Allowance for Severance and Retirement Benefits_(Plan Assets)_(Due from Pension Plan):Percentage
Summarized Income Statement (Banking Account)	(21) General and Administrative Expenses_Amortization of Intangible assets:Current Quarter
	(34) Interest_Interest and Dividends on Securities_Interest on Trading Securities:Current Quarter

Additionally, Top 10 variables were analyzed by classifying them into Assets, Stockholders' Equity, and Liabilities, which are factors related to direct measurement of the financial position based on the Balance Sheet. First, the Total Assets element consists of Cash and Due from Banks, Securities, Loans, Tangible Assets, and Other Assets on the Balance Sheet. As a result of the analysis, only Tangible Assets used for Business (Accumulated Depreciation (-)) was included, as it is an account item of the Tangible Assets category.

Second, the Total Liabilities element consists of Depository Liabilities, Borrowings & Bonds Payable, Other Liabilities, Share Capital Repayable on Demand category of the Balance Sheet. As a result of the analysis, four variables are confirmed for each, Borrowings (Discount Present Value) (-); Borrowings & Bonds Payable (Percentage); Bonds Payable (Percentage); and Borrowings (Percentage) in the Borrowings & Bonds Payable element. Allowance Accounts (Transfer from National Pension (-)) was selected in the Other Liabilities element. Accordingly, a total of five account items were confirmed in the Liabilities element.

Third, the Total Stockholders' Equity element consist of Capital Stock, Other Equity Instruments, Capital Surplus, Retained Earnings, Capital Adjustment, and Accumulated Other Inclusive Gain(loss) category on the Balance Sheet. The results of the analysis confirmed the Tier 1 capital ratio, which is the key indicator corresponding those total stockholders' equity elements.

Finally, Footnotes and Off-balance Accounts consist of Acceptances and Guarantees Outstanding, Acceptances and Guarantees, Commitments, Bills Endorsed, Derivative Contracts, Contracts on Credit Derivatives Purchased, Checks & Bills on Clearing, Receivable Charge-Offs, OTC Bonds Sold and Loans Sold under Repurchase Agreements category on the Balance Sheet. Among them, three variables, Acceptances and Guarantees On L/C and On Others, Acceptances and Guarantees, and Receivable Charge-Offs, were confirmed as account items representing the Footnotes.

As shown in the above results, five account items of Total Liabilities, three account items of Footnotes and one account item of Total Assets were selected with the Tier 1 capital ratio in top 10 variables. The Tier 1 capital ratio was an expected result because it showed banks' core equity capital against their total risk weighted assets. In the asset item, only Tangible Assets Used for Business (Accumulated Depreciation (-)) was selected. Accumulated depreciation means all recorded depreciation to calculate the present value of Fixed Assets on the asset line in the balance sheet. In the Footnotes, Acceptances and Guarantees On L/C and On Others consist of Payment guarantees confirmed by a bank, but the main obligation has not yet been determined. The Receivable Charge-Offs Accounts was selected for the account for the process of the amortized receivable balance.

In the Liabilities, five of the top 10 variables were covered for classifying into Assets, Stockholders' Equity, and Liabilities based on the Balance Sheet. Especially, Borrowings & Bonds Payable, which are borrowed after the contract of the return of the principal and interest, showed most of the account items with four variables, except for the Transfer from the National Pension amount for Other Liabilities. Interestingly, Deposits in Won (Mutual Installment Deposits: Percentage) was ranked at 106th as account item of Depository Liabilities, thus it was excluded from the final result even though it represents the structure of sources of funds for bank with Equity Capital, Borrowings, Bonds Payable, and Other Liabilities. In addition, Liabilities cannot be found by BIS formula since BIS capital adequacy ratio is not considering the liability in the formula.

As such, the results by the machine learning technique confirmed that the total liabilities elements are important predictors, which represent the BIS capital adequacy ratio.

This result shows various factors affected both directly and indirectly, which needs to be considered for BIS prediction since various factors are influenced by several factors with either direct or indirect correlation.

In this study, there was found a direct or indirect correlation between Assets, Stockholders' Equity and Liabilities. According to Heo (2020), the various factors from multiple systems are interacting and influencing the outcome on the dynamics of ecological and systemic framework. Therefore, large numbers of predictors from multiple systems around an organization need to be considered carefully for granular control of the financial soundness managerial performance. In addition, the total BIS ratio is calculated by the standard approach or the IRB approach, based on the exposure with risk. Therefore, it is not possible to simply predict total risk for BIS Capital Adequacy Ratio based on a preset of the determinants from previous literature. However, since this study highlighted the key findings (i.e., various determinants from diverse systems) by using machine learning approach, it has contributed to the policymakers as well as the managers and practitioners in the bank-related fields.

To support the discussion above, the statistical validity for the future predictions was compared and showed that the machine learning BRNN method was significantly better at predicting the total BIS ratio, which showed superior accuracy and performance compared to the BGLM. This result supports the previous literature (Bosarge 1993; Heo 2020; Heo et al. 2020; Linoff and Berry 2011; Thompson 2014; Ye 2013).

Author Contributions: J.P.—Conceptualization, data collection, original draft, investigation; M.S.—Project administration, writing revision, supervision; W.H.—Methodology, writing and review, data management. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Resampling Performance Over Subset Size

Variables	RMSE	Rsquared	MAE	RMSESD	Rsquaredsd	MAESD
1	1.14645624	0.52161392	0.91597048	0.36944915	0.25013672	0.30473594
2	0.83488915	0.69543209	0.65871893	0.28540853	0.24394932	0.23308458
3	0.67893878	0.79362339	0.54014738	0.24017445	0.20280734	0.20077234
4	0.62692811	0.83468561	0.50001168	0.21732637	0.17058545	0.17891839
5	0.57497985	0.86323263	0.45396974	0.2005359	0.12218131	0.16157981
6	0.49643268	0.89350869	0.38584788	0.19108713	0.10980485	0.14230643
7	0.46477758	0.91060933	0.36186155	0.1849704	0.09282203	0.14001393
8	0.44432276	0.91851811	0.34425833	0.16859113	0.08582431	0.12287298
9	0.42694648	0.92633685	0.33015263	0.16704546	0.07508313	0.11987712
10	0.41042964	0.93148302	0.31652693	0.1591851	0.0778689	0.1117455
11	0.40045495	0.93270929	0.30961061	0.1577901	0.07636589	0.10924144
12	0.39109651	0.93500684	0.30124706	0.15665837	0.07570652	0.11049778
13	0.37949307	0.93999163	0.29305774	0.15246007	0.07213541	0.10936046
14	0.3747662	0.94274916	0.28856782	0.14681898	0.06354868	0.10608978
15	0.36460401	0.94557023	0.28159099	0.14678479	0.05899202	0.10561506
16	0.35801908	0.94814348	0.27776522	0.14022683	0.0590924	0.10467481
17	0.35161432	0.9503722	0.27264381	0.14029614	0.05575943	0.10401329
18	0.34306366	0.95269971	0.26457882	0.13749041	0.05533791	0.09966472
19	0.34346663	0.95415872	0.26487792	0.13132092	0.04871966	0.09553778
20	0.33867216	0.95426285	0.26067439	0.13142173	0.05040102	0.09473011
21	0.34125451	0.95226657	0.26075325	0.13809573	0.0571921	0.0977487
22	0.33740905	0.95361833	0.25860115	0.13957096	0.06041624	0.09890045
23	0.33451822	0.95309857	0.25596187	0.14202812	0.06127314	0.09942867

Variables	RMSE	Rsquared	MAE	RMSESD	Rsquaredsd	MAESD
24	0.33585402	0.95350444	0.25592416	0.14270643	0.06258873	0.09930184
25	0.33440436	0.95371627	0.25683429	0.14407921	0.0643146	0.10102568
26	0.33309569	0.95497805	0.25530976	0.14001755	0.05874175	0.09814218
27	0.33059904	0.95465301	0.25294532	0.13792976	0.05935399	0.09573769
28	0.3326236	0.95406408	0.25504138	0.14039886	0.06459986	0.09684133
29	0.33070512	0.95423625	0.25471799	0.13988459	0.06143998	0.09745162
30	0.33026133	0.95437832	0.25394703	0.14270839	0.06413126	0.09832754
31	0.3314382	0.95483292	0.25452592	0.13697178	0.05975762	0.09509734
32	0.32950484	0.95535379	0.25411515	0.13907456	0.05805611	0.09703359
33	0.32951393	0.95553131	0.25386125	0.13939743	0.05872402	0.09830013
34	0.33021771	0.95548004	0.25430648	0.1402402	0.05673706	0.09915603
35	0.32709189	0.95600184	0.25101174	0.13678294	0.05795215	0.09605582
36	0.3287237	0.95555893	0.25298414	0.13940936	0.05596266	0.09943597
37	0.32690524	0.95650936	0.25045244	0.13706089	0.05292801	0.09645311
38	0.32903835	0.95533334	0.25032301	0.13927293	0.05454318	0.09675861
39	0.32595764	0.95616563	0.24930408	0.13698386	0.05553099	0.09553185
40	0.32334027	0.95752941	0.24737394	0.13580949	0.0549632	0.09521656
41	0.32298461	0.95756363	0.24815907	0.13296681	0.05177316	0.09404707
42	0.32119002	0.95789169	0.24544974	0.13752653	0.05510958	0.09449879
43	0.32329842	0.95752497	0.24776468	0.13657025	0.05486717	0.09473953
44	0.32446705	0.95739704	0.24895191	0.13672597	0.05607182	0.09412781
45	0.32005437	0.9590659	0.24430884	0.13759181	0.05003251	0.09617369
46	0.32301006	0.95843716	0.24760451	0.1372733	0.05021254	0.09559924
47	0.32080698	0.95862485	0.24543791	0.13473029	0.04856581	0.09411708
48	0.3209867	0.95894209	0.24464988	0.13712268	0.05111523	0.09657376
49	0.31926599	0.95900948	0.24396994	0.13760497	0.05116915	0.09720571
50	0.31770839	0.95986892	0.24167917	0.13685467	0.04782006	0.09547324
51	0.31975591	0.95916842	0.24299139	0.13772705	0.05335469	0.09567098
52	0.31759363	0.96003622	0.2421658	0.13634134	0.04886002	0.09466409
53	0.31824483	0.95970273	0.24286984	0.13192711	0.0497733	0.0929538
54	0.31859278	0.95967006	0.24238003	0.13649142	0.05269658	0.09526267
55	0.31716627	0.96080082	0.24218971	0.13396621	0.04801456	0.09443763
56	0.3164921	0.96084916	0.24044904	0.13153527	0.04782532	0.09130198
57	0.31723576	0.96012941	0.24097047	0.13678016	0.05194551	0.09499573
58	0.31443086	0.96133116	0.23965306	0.13413816	0.04642494	0.09368547

Appendix B. Resampling Performance Over Subset Size: Important Predictors by RFE

Num	Overall	Variance (English)
1	22.151384	Tier 1 Capital Ratio
2	11.111107	Borrowings_Bonds Payable_(Discount Present Value):Percentage
3	9.3877522	Borrowings:Percentage
4	8.9370711	Acceptances and guarantees others
5	8.6561452	Acceptances and Guarantees
6	8.2861521	Borrowings_Bonds Payable:Percentage
7	7.9801142	Borrowings_Borrowings:Percentage
8	7.8963313	Receivable Charge-Offs
9	7.7832859	Other Liabilities_(Transfer from National Pension):Amount
10	7.7440828	Fixed Asset_Tangible Assets Used for Business Purpose_((Accumulated Depreciation)):Amount
11	7.6228497	Construction
12	7.5957596	Financing Without Cost_Other Non-cost Bearing Financing:(Average)
13	7.2821885	Financing Without Cost_Provision for Other Allowances:Percentage
14	7.24095	Performing Asset Management_Due From Banks in Won:(Average)
15	6.6978429	Other Liabilities_Account for Agency Business_Grio Account:Amount
16	6.5836249	Financing Without Cost_Demand Deposits:Percentage
17	6.4640151	Loans in won _ Average Interest rate
18	6.4579122	Financing With Cost_Borrowings in Won:Percentage
19	6.4524425	Non-Performing Asset Management_Others:Average
20	6.4328985	Performing Asset Management_Due From Banks in Won:Percentage
21	6.4170021	General and Administrative Expenses_Amortization of Intangible assets:Current Quarter
22	6.3741191	Financing Without Cost_Other Non-cost Bearing Financing:Percentage
23	6.3035924	Operation_Loans & Discounts:Percentage
24	6.2622486	Non-Performing Asset Management_Cash & Checks and Foreign Currency:Percentage
25	6.1880531	Derivative Contracts
26	6.1864739	Other Liabilities_Allowance Accounts_Allowance for Severance and Retirement Benefits_(Plan Assets)_(Due from Pension Plan):Percentage
27	6.0419083	Loans_Loans & Discounts in Won Loans to Enterprise:Percentage
28	6.0393983	Fixed Asset_Tangible Assets_Buildings Used for Business Purpose:Amount
29	6.0326002	Asset Management for Benefit_Other Won-Denomiated Currency Asset Management:Average
30	6.0161628	Securities_Banking Accounts (Available-for-Sales Securities)_ Available-for-Sales Securities in Won_Others: Amount
31	5.9820987	Bond Accounts:Amount
32	5.9736885	Personal Pension Trust:Percentage
33	5.9424913	Deposits in Won _ Average Interest rate
34	5.9243753	Interest_Interest and Dividends on Securities_Interest on Trading Securities:Current Quarter

Num	Overall	Variance (English)
35	5.8140847	Loans_Credit Card Accounts_Cash Service:Percentage
36	5.7222446	Securities-Banking Accounts (Subsidiaries)_Equity Investment (Won)_Consolidated Subsidiary Stock: Amount
37	5.7078509	Loans & Discounts_Loans on Real Estate Collateral:Amount
38	5.7019038	Non-Performing Asset Management_Fixed Assets Used for Business Purposes:Percentage
39	5.6784716	Other Liabilities_Accrued Expenses Payable:Percentage
40	5.6422311	Interest_Available-for-Sales Securities Interest:Current Quarter
41	5.6110259	Deposits in Won_Mutual Installment Deposits
42	5.6005534	Loans_Allowance for Credit Losses on Other Loans_Credit Card Accounts:Percentage
43	5.5961849	(Allowance for Credit Losses) Amount
44	5.5531814	Collateral_Others
45	5.4702801	Nonoperating Income_Rental income:Current Quarter
46	5.4684209	Securities-Banking Accounts (Subsidiaries)_Equity Investment (Won)_Consolidated Subsidiary Stock:Percentage
47	5.4506053	Allowance for Credit Losses on Other Loans:Amount
48	5.4476118	Loansoff-Shore Loans in Foreign Currency:Percentage
49	5.4371039	Loans_Loans & Discounts in Won Interbank Loans:Amount
50	5.4077926	Total Financing & Operation:Average
51	5.4046407	Securities-Banking Accounts (Securities-Held-to-Maturity Securities):Securities (Foreign Currencies):Percentage
52	5.380378	Financing_Trust Accounts:Average
53	5.3780784	Depository Liabilities_Deposits (Won)_Trust Account:Percentage
54	5.3771052	Performing Asset Management_Securities (Foreign Currencies):Percentage
55	5.3696979	Securities-Banking Accounts (Securities-Held-to-Maturity Securities):Securities (Foreign Currencies)_Debentures:Amount
56	5.3564302	Performing Asset Management_Loans in Foreign Currency:Percentage
57	5.3204636	Loans_Loans on Trust Benefit Collateral:Percentage
58	5.3196173	Consolidated Capital Surplus:Amount

References

Bazarbash, Majid. 2019. *FinTech in Financial Inclusion Machine Learning Applications in Assessing Credit Risk*. IMP Working Paper No. WP/19/109. Basel: International Monetary Fund.

Bank for International Settlements. 2020. About BIS—Overview. Bank for International Settlement. Available online: https://www.bis.org/about/index.htm?m=1%7C1 (accessed on 25 December 2020).

Bosarge, W. E. 1993. Adaptive processes to exploit the nonlinear structure of financial markets. In *Neural Networks in Finance and Investing*. Chicago: Probus Publishing, pp. 371–402.

Breiman, Leo. 2001. Random forests. *Machine Learning* 45: 5–32. [CrossRef]

Burden, Frank, and Dave Winkler. 2008. Bayesian regularization of neural networks. In *Artificial Neural Networks*. Tolowa: Humana Press, pp. 23–42.

Fsc.go.kr. 2020. Financial Services Commission, Early Implementation of Basel III Credit Risk Framework. Available online: https://www.fsc.go.kr/eng/pr010101/22439 (accessed on 20 December 2020).

Gambacorta, Leonardo, Yiping Huang, Han Qiu, and Jingyi Wang. 2019. How do Machine Learning and Non-Traditional Data Affect Credit Scoring? New Evidence from a Chinese Fintech Firm (BIS Working Paper No. 834). Bank of International Settlements. Available online: https://www.bis.org/publ/work834.htm (accessed on 20 December 2020).

Genuer, Robin, Jean-Michel Poggi, and Christine Tuleau-Malot. 2010. Variable selection using random forests. *Pattern Recognition Letters* 31: 2225–36. [CrossRef]

Guyon, Isabelle, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine Learning* 46: 389–422. [CrossRef]

Heo, Wookjae. 2020. Theoretical background: A new theoretical framework for financial planning with the case of life insurance demand—Dynamic ecological systemic framework. In *The Demand for Life Insurance*. Cham: Palgrave Pivot, pp. 19–46.

Heo, Wookjae, Jae Min Lee, Narang Park, and John E. Grable. 2020. Using artificial neural network techniques to improve the description and prediction of household financial ratios. *Journal of Behavioral and Experimental Finance* 25: 100273. [CrossRef]

Ho, Tin Kam. 1998. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20: 832–44.

Hyndman, Rob J., and Anne B. Koehler. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22: 679–88. [CrossRef]

Korea Institute of Financia. 2016. The improvement and implications for risk-weighted asset calculation method of Basel Bank Supervisory Commission. *International Financial Issues* 25: 14–15.

Kohavi, Ron, and George H. John. 1997. Wrappers for feature subset selection. *Artificial Intelligence* 97: 273–324. [CrossRef]

Kursa, Miron B., and Witold R. Rudnicki. 2010. Feature selection with the Boruta package. *Journal of Statistical Software* 36: 1–13. [CrossRef]

Liaw, Andy, and Matthew Wiener. 2002. Classification and regression by random forest. *R News* 2: 18–22.

Linoff, Gordon S., and Michael J. A. Berry. 2011. *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. Hoboken: John Wiley & Sons.

MacKay, David J. C. 1992. Bayesian Interpolation. *Neural Computation* 4: 415–47. [CrossRef]

Nilsson, Roland, José M. Peña, Johan Björkegren, and Jesper Tegnér. 2007. Consistent feature selection for pattern recognition in polynomial time. *Journal of Machine Learning Research* 8: 589–612.

Petropoulos, Anastasios, Vasilis Siakoulis, Evaggelos Stavroulakis, and Aristotelis Klamargias. 2019. A robust machine learning approach for credit risk analysis of large loan level datasets using deep learning and extreme gradient boosting. In *Bank for International Settlements*. IFC Bulletins. Basel: Bank for International Settlements, vol. 49.

Sariev, Eduard, and Guido Germano. 2020. Bayesian regularized artificial neural networks for the estimation of the probability of default. *Quantitative Finance* 20: 311–28. [CrossRef]

Thompson, Wayne. 2014. Data mining methods and the rise of big data. In *Big Data, Mining, and Analytics*. New York: Auerbach, p. 71.

Wooldridge, Jeffrey M. 2016. *Introductory Econometrics: A Modern Approach*. Mason: Nelson Education.

Ye, Nong. 2013. *Data Mining: Theories, Algorithms, and Examples*. Boca Raton: CRC Press.

Zhou, Qifeng, Hao Zhou, Qingqing Zhou, Fan Yang, and Linkai Luo. 2014. Structure damage detection based on random forest recursive feature elimination. *Mechanical Systems and Signal Processing* 46: 82–90. [CrossRef]

Fss.or.kr. 2020. Financial Supervisory Service, Basel Regulations Related Report. Available online: http://www.fss.or.kr/fss/kr/bbs/list.jsp?bbsid=1207396739395&url=/fss/kr/1207396739395 (accessed on 25 August 2020).

National Law Information Center. 2020a. Banking Act §34. Available online: http://www.law.go.kr/%eb%b2%95%eb%a0%b9/%ec%9d%80%ed%96%89%eb%b2%95 (accessed on 24 August 2020).

National Law Information Center. 2020b. Detailed Regulations on Supervision of Banking Business. Appendix §17. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ec%97%85%eb%ac%b4%ec%8b%9c%ed%96%89%ec%84%b8%ec%b9%99 (accessed on 25 August 2020).

National Law Information Center. 2020c. Detailed Regulations on Supervision of Banking Business. Appendix §3-2-2. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ec%97%85%eb%ac%b4%ec%8b%9c%ed%96%89%ec%84%b8%ec%b9%99 (accessed on 25 August 2020).

National Law Information Center. 2020d. Detailed Regulations on Supervision of Banking Business. Appendix §3. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ec%97%85%eb%ac%b4%ec%8b%9c%ed%96%89%ec%84%b8%ec%b9%99 (accessed on 25 August 2020).

National Law Information Center. 2020e. Detailed Regulations on Supervision of Banking Business. Appendix §3-3. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ec%97%85%eb%ac%b4%ec%8b%9c%ed%96%89%ec%84%b8%ec%b9%99 (accessed on 25 August 2020).

National Law Information Center. 2020f. Detailed Regulations on Supervision of Banking Business. Appendix §3-4. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ec%97%85%eb%ac%b4%ec%8b%9c%ed%96%89%ec%84%b8%ec%b9%99\T1\textgreater{} (accessed on 25 August 2020).

National Law Information Center. 2020g. Enforcement Decree of The Banking Act §20. Available online: http://www.law.go.kr/%eb%b2%95%eb%a0%b9/%ec%9d%80%ed%96%89%eb%b2%95%ec%8b%9c%ed%96%89%eb%a0%b9 (accessed on 25 August 2020).

National Law Information Center. 2020h. Regulations on Supervision of Banking Business §26. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ea%b7%9c%ec%a0%95 (accessed on 25 August 2020).

National Law Information Center. 2020i. Regulations on Supervision of Banking Business §33. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ea%b7%9c%ec%a0%95 (accessed on 25 August 2020).

National Law Information Center. 2020j. Regulations on Supervision of Banking Business §34. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ea%b7%9c%ec%a0%95 (accessed on 25 August 2020).

National Law Information Center. 2020k. Regulations on Supervision of Banking Business §35. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ea%b7%9c%ec%a0%95 (accessed on 25 August 2020).

National Law Information Center. 2020l. Regulations on Supervision of Banking Business §36. Available online: http://www.law.go.kr/%ed%96%89%ec%a0%95%ea%b7%9c%ec%b9%99/%ec%9d%80%ed%96%89%ec%97%85%ea%b0%90%eb%8f%85%ea%b7%9c%ec%a0%95 (accessed on 25 August 2020).