Predicting and Analyzing the Intangible Assets Value with Machine Learning Techniques

Yu-Hsin Lu

Department of Accounting, Feng Chia University, Taichung, Taiwan

Yu-Feng Hsu*

Department of Accounting and Information Technology, National Chung Cheng University, Chia-yi, Taiwan

Abstract

With the shift to a knowledge-based economy, the importance of intangible assets to management has increased. Therefore, evaluating the real value of intangible assets has become a critical issue. This study develops and compares two models based on machine learning techniques: single classifier models and ensemble learning methods. The results show that the ensemble learning methods do not perform better than the single best classifiers. On the other hand, a well-performing single classifier model produces the optimal intangible asset value prediction model. Otherwise, this study shows and explains essential features in different periods. Finally, moving window-based and fixed window-based evaluation methods are adopted to verify the prediction algorithms, since the dynamic economy and business environment changes over time may affect the data. The results show that the 1Y Random Tree method provides the best results. These prediction models can provide valuable information to management and shareholders to evaluate the status of companies and make more effective decisions.

Keywords: Intangible Assets Value, Artificial Neural Networks, Linear Regression, Support Vector Regression, Ensemble Learning.

JEL Classifications: M41, M48

*Corresponding author

E-mail: benyfhsu@ccu.edu.tw

Address: No.168, Sec. 1, University Rd., Minhsiung, Chiayi 621301, Taiwan.

1. Introduction

The value of a firm is a critical issue for the management and stockholders in business operations. A company with high value usually implies that the management provides good performance in operating a company and implies more profit for the stockholders. Therefore, in order to enhance the firm value, it is crucial to figure out what would affect the firm value (Ahmed Haji & Mohd Ghazali, 2018; Gamayuni, 2015; Lev, Radhakrishnan, & Zhang, 2009; Osinski, Selig, Matos, & Roman, 2017; Penman, 2009). In the past, firm value has been estimated based on the tangible assets shown on the balance sheet. However, with the shift to a knowledge-based economy, the importance of intangible assets has increased, and a considerable change has occurred over the past twenty years (Zeb & Rashid, 2016). As a result, the source of firm value is changed from physical manufacturing factors (e.g., machinery and factory) to intangible factors such as knowledge and resources.

A company's market price reflects a firm's value, and some studies prove that a firm's market value is usually greater than its book value (Gamayuni, 2015). Morris & Alam (2012) studied the relationship between the stock market value and the value obtained from traditional accounting methods in the 1990s. They suggested that earning quality could explain part of the relationship. The gap between the market value and the book value of a firm has prompted researchers to investigate the importance of intangible assets for identifying firm value. Moreover, it is necessary to find an effective method to evaluate intangible assets' value and provides helpful information to the management.

Several studies applying machine learning techniques to construct prediction models for intangible assets found that the prediction models outperform the traditional statistical methods on performance (Huang, Chen, Hsu, Chen, & Wu, 2004; Min & Lee, 2005). However, most studies have applied a single classifier or classifier ensembles. To the best of our knowledge, no study has applied machine learning techniques to evaluate the actual value of intangible assets. In order to provide a more precise estimation result and suggestion to the management, this study aims to examine the performance of single classifier models and ensemble learning methods for predicting the actual value of intangible assets.

In the past studies, six well-known machine learning algorithms, including Linear Regression (LR), Multi-layer Perception (MLP), Support Vector Machine (SVM), Classification and Regression Tree (CART), Random Tree (RT), and K-Nearest Neighbors (KNN), and two ensemble learning methods (i.e., Boosting and Bagging) have been employed to construct the prediction model (Cover & Hart, 1967; Haykin, 1994; Quinlan, 1987; Vapnik, 2013). To fair examine the prediction performance for these methods, we consider all the methods above in this study.

The contributions of this study are two-fold. First, apply machine learning algorithms, including single classifier models and ensemble learning methods. This study is the first to apply machine learning to estimate the value of intangible assets. The performance of these algorithms has not yet been fully assessed for this purpose. The experiments are implemented using datasets for companies listed on NASDAQ, NYSE, and AMEX to determine the best prediction model. Second, due to fewer regulations and limitations for disclosing intangible assets, the value cannot be precisely reflected in financial reports. Therefore, this study can provide more information on intangible assets' value to the management when assessing the company's performance.

The remainder of this study organizes as follows. Section 2 reviews related literature about the value of intangible assets. The experimental methodology is described in Section 3, and the experimental results are presented in Section 4. Finally, in Section 5, some conclusions and their implication are offered.

2. Literature Review

2.1 Value of Intangible Assets

Knowledge and information technology are the fundamental driving forces that trigger dramatic changes to the structure of companies. These forces drive the companies to shift their focal point from tangible to intangible assets. The intangible assets are increasingly important, influencing organizational competitiveness (Edvinsson, 2013) and stimulating future profitability growth. Therefore, they have prevailed as a measure of core competency and competitive advantage of a company and explain the gap between the market value and book value of an organization; significantly when the financial reports are decreasing in usefulness (Ahmed Haji & Mohd Ghazali, 2018; Han & Han, 2004; Lev et al., 2009; Penman, 2009).

The evidence shows that the company's market value is usually higher than the book value, especially the hidden values are over 80 percent in some cases (Ahmed Haji & Mohd Ghazali, 2018; Haji, 2016; Lev et al., 2009). Therefore, many researchers studied intangible assets and tried to figure out the main factors that affect the market value. Stewart & Ruckdeschel (1998) define intangible assets like knowledge, information, intellectual property, and experience that can be used to create wealth. Sveiby (1997) determined that the intangible assets of a firm consist of internal (e.g., patents, administrative system, organizational structure), external (e.g., brands, trademarks, relations with customers and suppliers), and organizational structures as well as of the competence of their employee. Adriessen & Tissen (2000) provided a considerably broader definition of intangible assets. They distinguished five assets groups that may be referred to as intangible assets: 1. assets and endowments, 2. skills and tacit knowledge, 3. collective values and norms, 4. technology and explicit knowledge, 5. primary and management processes. However, there is no uniform definition, even though a general understanding of intangible assets composition exists.

Intangible and tangible assets must be combined to ascertain the actual firm market value. However, the value of intangible assets is more challenging to determine than that of tangible assets. Although many companies voluntarily provide information regarding intangible assets (Burgman & Roos, 2007; Eleftherios, Dimitrios, & Panagiotis, 2021; Tsai, Lu, Hung, & Yen, 2016; Vandemaele, Vergauwen, & Smits, 2005; Yunhong, 2009); estimating intangible assets' value is still a difficult task, and lack of an accurate and objective estimating method. Therefore, it is crucial to build an intangible assets prediction model and provide accurate estimating results to the management for decision-making.

The target variable used in this paper is the same as those in previous studies (Asghar, Sajjad, Shahzad, & Matemilola, 2020; Buallay, Hamdan, & Zureigat, 2017; Mishra & Kapil, 2017; Silva, Silva, & Chan, 2019; Tsai et al., 2016); Tobin's Q is applied as a proxy for intangible assets. It is based on the difference between a firm's market value and the replacement cost of the tangible assets. Tobin's Q can be used to represent the value of the intangible assets; while Tobin's Q is more significant than one, the firm's market value is greater than the book value.

2.2 Determinants of Intangible Assets Value

A survey of the literature shows many factors affecting the value of intangible assets (Asghar et al., 2020; Dang, Pham, & Vu, 2018; Horn, De Klerk, & De Villiers, 2018; Kong, Famba, Chituku-Dzimiro, Sun, & Kurauone, 2020; Li, 2016; Nekhili & Cherif, 2011; Yoong, Alfan, & Devi, 2015); the literature described as follows:

Many studies (Black, Jang, & Kim, 2006; Fukui & Ushijima, 2007; Horn et al., 2018; Rao, Agarwal, & Dahlhoff, 2004; Yoong et al., 2015) use the market value as a forward-looking performance measure, it represents the expectation of the investors for the companies' future potential profit. Gleason and Klock (2006) find that innovation and brand loyalty are viewed as investments which with forecastable positive effects on the future cash flow and market-based value. Klock & Megna (2000) argued that advertising and R&D stocks are determinants of Tobin's Q in the wireless communications industry. Gleason & Klock (2006) suggested that advertising and R&D intensity are statistically significant determinants of Tobin's Q and explained 20% of the variation in their sample. Nekhili & Cherif (2011) and Rao et al. (2004) indicate that advertising and R&D expenditure are positive and significant in statistics to be the determinants of the market value of a corporation.

There are several firm-specific factors affecting the intangible assets value directly or indirectly. Some factors are indicators of historical business operation, and others reflect future growth opportunities (Rao et al., 2004). Sales growth is a proxy for future growth opportunities which increase firm value. Firm size is likely inversely related to expected growth opportunities (Asghar et al., 2020; Fukui & Ushijima, 2007; Gleason & Klock, 2006; Kong et al., 2020; Li, 2016; Mishra & Kapil, 2017; Silva et al., 2019; Yoong et al., 2015). Asghar et al. (2020) and Dang et al. (2018) found that high financial leverage will cause financial distress and reduce the value of a business, even leading to bankruptcy. Li (2016) indicated that higher leverage is related to higher risk and affects firm value. However, some studies (Black et al., 2006; McConnell & Servaes, 1990; Yoong et al., 2015) have found that firms with higher leverage can obtain tax benefits because they can deduct the interest costs; this may result in greater cash flow and thus have a positive relationship with firm value. Moreover, capital intensity and capital expenditure also affect firm value because they can be viewed as a proxy for investment opportunities (Horn et al., 2018; Jentsch, 2019; Khediri & Folus, 2010). Although many studies suggest a positive relationship between capital intensity and firm value, most provide insignificant or negative results. Allayannis & Weston (2001) found that if the management cannot obtain the necessary funding, they will be forced to forgo the investment projects; however, Tobin’s Q may remain high because they will only be able to undertake positive net present value (NPV) projects. A firm paying higher dividends may result in fewer profitable investment opportunities and thus obtain a lower Tobin’s Q(Jentsch, 2019; Khediri & Folus, 2010).

A profitable firm generally triggers investors' expectations of high cash flow, which drives the intangible value. Furthermore, higher intangible values are significantly associated with higher profitability (Horn et al., 2018; Kong et al., 2020; Li, 2016). On the other hand, some studies spend their efforts on one and multiple earnings measures and their impact on firm value through the cost of capital effect. For example, Barth and his colleagues (Barth, Konchitchki, & Landsman, 2013) indicated an insignificant linkage between earnings quality and the cost of capital. However, while the earnings quality of financial reports increased, the cost of capital decreased since the information asymmetry was reduced, which will increase the firm value. Moreover, financial accounting improves the quality of the information available to stakeholders and thus improves firm value (Asghar et al., 2020). In addition to the firm's characteristics, the characteristics of the different industries will affect the value of their intangible assets (Hong, 2017; Jentsch, 2019; Ni, Huang, Cheng, & Huang, 2020).

Land, capital, and labor are critical factors in assessing a firm's value in the traditional manufacturing industry. However, in knowledge-intensive industries (e.g., high-tech companies), knowledge and innovation are the dominant resources far more critical than physical assets (Tseng & James Goo, 2005). In these cases, intangible assets determine a large part of a firm's value. For example, Klock & Megna (2000) showed that in the communication industry, a firm's market value is about ten times higher than the book value; but in traditional industries, most firms' Tobin's Q is usually equal to or less than one. In other words, there is a considerable variation in the value of intangible assets in different industries. According to the above, the factors affecting intangible assets value are listed in Table 1.

Table 1. Item-Total Statistics for Each Variable

Variables	Measurement	References
R&D INTENSITY	Research and development expenditure in the current year.	(Asghar et al., 2020; Black et al., 2006; Dang et al., 2018; Fukui & Ushijima, 2007; Gleason & Klock, 2006; Li, 2016; Mishra & Kapil, 2017; Silva et al., 2019; Yoong et al., 2015) Allayannis & Weston, 2001; (Asghar et al., 2020; Barth et al., 2013); (Black et al., 2006; Fukui & Ushijima, 2007; Gleason & Klock, 2006; Khediri & Folus, 2010; Klock & Megna, 2000; Nekhili & Cherif, 2011; Rao et al., 2004; Yoong et al., 2015).
ADVERTISING INTENSITY	Advertising expenditure in the current year.	
SALE GROWTH	Percentage change in sales.	
SIZE	The log of total assets.	
LEVERAGE	The ratio of total debt to total assets.	
CAPITAL INTENSITY	The ratio of capital expenditures to sales.	
CAPITAL EXPENDITURE	Capital expenditure in the current year.	
DIVIDEND	Dummy variable; which is equal to 1 if the firm paid dividends in the current year.	
PROFITABILITY	The ratio of net income to total assets.	
EARNING MANAGEMENT	Total net income in the current year - total cash flows from operating activities in the current year.	
INDUSTRY	Indicator variables; industry category of the stock exchanges.	

According to the literature discussion, the factors are critical to determining the value of intangible assets. However, they are usually applied in the regression models in the literature and assume that dependent and independent variables are linearly correlated. This assumption can be easily challenged since the results provided in prior studies are insignificant or mixed. In order to overcome this problem, this study employs machine learning techniques to explore the relationship between intangible assets value and the critical factors; and construct the intangible assets prediction model.

3. Research Design

3.1 The research samples

This study collects the data from the Capital IQ database[1] for companies listed on NASDAQ, NYSE, and AMEX. The period covered by the dataset ranges from 2000 to 2019. After eliminating the missing values and unidentifiable data, 74,427 observations remain to be used for the experiments.

3.2 Prediction methods

This study developed and compared two types of intangible assets value prediction models:

1 The data are available on the Capital IQ database. S&P Capital IQ Pro | S&P Global Market Intelligence (spglobal.com)

single classifier models and ensemble learning methods (i.e., Bagging and Boosting). The techniques used to construct these prediction models include LR, MLP, SVM, CART, RT, and KNN algorithms.

3.2.1 Single Classifiers

3.2.1.1 Linear Regression

Linear regression is a widely used statistical method (Yan & Su, 2009), which predicts possible outcomes of the dependent variable through a study of the independent variables. The relationship between the independent variables and the dependent variable is estimated through a linear prediction function. The parameters of the linear function are obtained by analyzing the data; this type of model is called a linear model (Seal, 1968).

3.2.1.2 Multi-Layer Perception Neural Network

The MLP neural network method has been commonly used in many disciplines. The architecture of an MLP neural network can be roughly divided into three parts: parameters of learning epochs, numbers of network layers, and numbers of hidden nodes. These parts are critical factors in the MLP neural network construction process. The network layers usually consist of one input layer, one or more hidden layers, and one output layer; the input layer includes several input nodes and several neural nodes in the hidden layers and the output layer, respectively. Moreover, the connection between the layers is implemented by passing the weighted values; the values will be dynamically updated during the training process of the neural network (Haykin, 1994).

3.2.1.3 Support Vector Regression

The SVM classification algorithms can be used to construct prediction models by applying a training and testing process. This process involves several steps: first, the training data are projected onto a higher-dimensional space; then, the kernel function creates a hyperplane on the higher-dimensional feature space (Vapnik & Vapnik, 1998). This hyperplane separates the data into two parts; the data in the different parts are assigned different labels. Finally, the data points closest to the hyperplane are recognized as the so-called support vectors used to determine the boundaries of the hyperplane in the classification process. Support Vector Regression (SVR) is an application of the SVM for regression. It operates similarly to the SVM, but rather than label prediction, it predicts discrete values.

3.2.1.4 Classification and Regression Trees

The Decision Tree (DT) algorithm is a prevalent method for prediction and classification problems that establishes a prediction model by constructing many nodes and branches. The model can produce several decision rules. Several different decision tree algorithms have been developed. For example, Breiman (1984) developed the Classification and Regression Tree (CART) method as a non-parametric statistical method to solve classification and regression problems. CART constructs classification trees or regression trees depending on the variable type, which can be categorical or numerical (Breiman, 1984). CART can be broken down into three stages. In the first step, a maximal tree is grown using a recursive partitioning technique to select variables and split points using a splitting criterion. The overgrown tree, which shows overfitting, is pruned in the next step. Then, cross-validation or a testing sample can be used to provide estimates of future classification errors for each subtree. The optimal tree, which corresponds to the tree yielding the lowest cross-validation or testing set error rate, is selected in the last stage.

3.2.1.5 Random Tree

The Random Tree (RT) is another tree-based prediction method built on the classification and regression tree methodology. This algorithm uses recursive partitioning to split the training data into segments with similar label values. The tree construction process starts from the root node by examining the input variables to find the best split, measured by the reduction in an impurity index resulting from the split. The split defines two subgroups, each of which is then split into two more subgroups until one of the stopping criteria is triggered. All splits are binary (only two subgroups).

The RT node uses bootstrap sampling with replacements to generate sample data. The sample data are used to grow a tree model. During tree growth, RT will not sample the data again. Instead, it randomly selects part of the predictors and uses the best one to split a tree node. This process is repeated when splitting each tree node. RT is the basic idea for growing a tree in a random forest.

3.2.1.6 K-Nearest Neighbors Algorithm

The K-Nearest Neighbor (KNN) algorithm is a conventional non-parametric classifier for pattern classification (Bishop, 1995). KNN classifies an unknown instance represented by feature vectors as a point in feature space; the KNN algorithm calculates distances between one point and other points in a training dataset. It then assigns the point to a class among its KNNs (k is an integer).

Compared to an inductive learning approach, KNN can also be called instance-based learning (Mitchell, 1997) because, without off-line training (i.e., model generation), the KNN algorithm only needs to search all examples of a given training dataset to classify a new instance. Therefore, the direct computation of the KNN algorithm is the online scoring of training examples to find the KNNs of a new instance. According to Jain, Duin, & Mao (2000), KNN can be conveniently used as a benchmark for all the other methods since it can perform well in most applications.

3.2.2 Ensemble Learning Method

The ensemble learning methods are based on combinations of multiple classifiers/regressors. There are two families of multiple regressor combinations, serial and parallel. In the former, the order of the arrangement is crucial for the system's performance, while the individual performance of each regressor does not have as much influence as the system performance. In the latter, system performance depends on combining different techniques (Kim, Kim, & Lee, 2003). The method used in this paper is based on the parallel combination since this study aims to compare the performance of six regression algorithms. Therefore, to build a prediction model based on the ensemble learning method, we combine different numbers of the same regressor techniques until the prediction performance does not change. In addition, both bagging and boosting combination methods are considered (Breiman, 1996; Freund & Schapire, 1996). The six machine learning algorithms are used to construct both bagging and boosting ensemble methods. In addition, the Random Forest (RF) method is a very flexible and powerful ensemble regressor/classifier decision tree-based technique first developed by Breiman (2001). RF is a popular ensemble regressor/classifier since it can avoid the problem of overfitting because each tree is grown using a different subset of data and features. Therefore, RF is also considered in the experiments.

We conduct a preliminary experiment for parameter setting for the intangible assets prediction models to find the best parameters for the algorithms. For the MLP neural network, a comparison of the prediction performance using different learning epochs and numbers of hidden nodes is implemented. For support vector regression, a few well-known kernel functions are used in the preliminary experiment to determine the most suitable kernel function and the values of cost and gamma. In the preliminary KNN experiment, K-values from 1, 3, 5, …, to 51 are considered in the testing. The number of trees in the Random Forest experiment ranges from 1, 2, 3, …, 1001. The default settings are used for the CART, LR, and Random Tree algorithms, with Tobin's Q used as a proxy variable for intangible assets value.

3.3 Evaluation Methods

In this study, the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are employed to measure the prediction performance of the algorithms. The model offering the lowest MAE and RMSE values is selected as the most accurate since the MAE and RMSE values are inversely proportional to the accuracy rate. The MAE and RMSE are given by equations 1 and 2 below.

$$MAE = \sum_{i=1}^{N} \frac{|p_i - x_i|}{N} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_i - x_i)^2} \tag{2}$$

Here, p represents the predicted Tobin's Q value (i.e., predicted intangible assets value), x represents the actual Tobin's Q value (i.e., actual intangible assets value), and N represents the number of instances.

4. Experimental Results

The predicted Tobin's Q values are used to determine the most suitable model for the user (i.e., the shareholders, investors, or others). This study conducts a series of comparisons and presents the results below. The results are divided into three parts (i.e., Before the crisis 2000-2006; During the crisis 2007-2008, and After the crisis 2009-2019). First, 2007-2008, also known as the global financial crisis (GFC), started from a subprime mortgage event and caused a severe worldwide economic crisis. The results sent stock markets spiraling and economies severely shaken. Before the COVID-19 recession in 2020, GFC was recognized by many economists as the most severe financial crisis in history. After the three-part division, the results for individual industries are provided and analyzed. Finally, the moving window-based and fixed-based evaluation techniques are applied to explore the value of Tobin's Q.

Table 2 shows descriptive statistics for Tobin's Q value for the different datasets. There are significant differences between the three datasets. For example, the mean of the During crisis period is higher than either the After crisis or Before crisis periods because of the immense Tobin's Q value of 3,360,000. In addition, the During crisis period (2007-2008) standard deviation is huge compared to the other two datasets, indicating the enormous influence of the financial crisis on companies.

Table 2. Descriptive statistics for the Tobin's Q value of the datasets

Descriptive Statistics				
Period	Full period 2000-2019	Before crisis 2000-2006	During crisis 2007-2008	After crisis 2009-2019
Mean	82.74	13.08	503.63	52.58
Median	0.96	0.96	0.85	0.98
Stdev	13,196.30	1,046.48	41,006.18	5,870.87
Skewness	230.72	103.59	81.94	157.62
Kurtosis	57,134.94	11,012.08	6,713.93	26,520.14
Min	-0.15	0.00	0.00	-0.15
Max	3,360,000.00	120,503.9195	3,360,000.00	1,072,634.64

4.1 Results for the Three Periods

First, the study applies the algorithms, including the single classifiers and two ensemble learning methods (Bagging based and Boosting based). Figure 1 shows the experimental results obtained on the Before crisis dataset. The MAEs obtained with the SVR, SVR Boosting, and SVR Bagging algorithms are relatively higher than those obtained with the other methods. On the other hand, the KNN and K-NN Boosting methods outperform the others and obtain the lowest MAE value of 17.63. The results may imply that the KNN and K-NN Boosting-based methods are more suitable for predicting Tobin's Q value for the 2000-2006 (Before crisis) dataset.

Comparing the single and ensemble learning methods helps understand whether the ensemble learning methods can outperform the single ones. The results indicate that the ensemble learning methods do not necessarily perform better than the best single classifier, as the differences in performance between the single and ensemble learning methods are insignificant based on the Analysis of Variance (ANOVA). A well-performing regression technique will produce the optimal model for intangible assets value prediction.

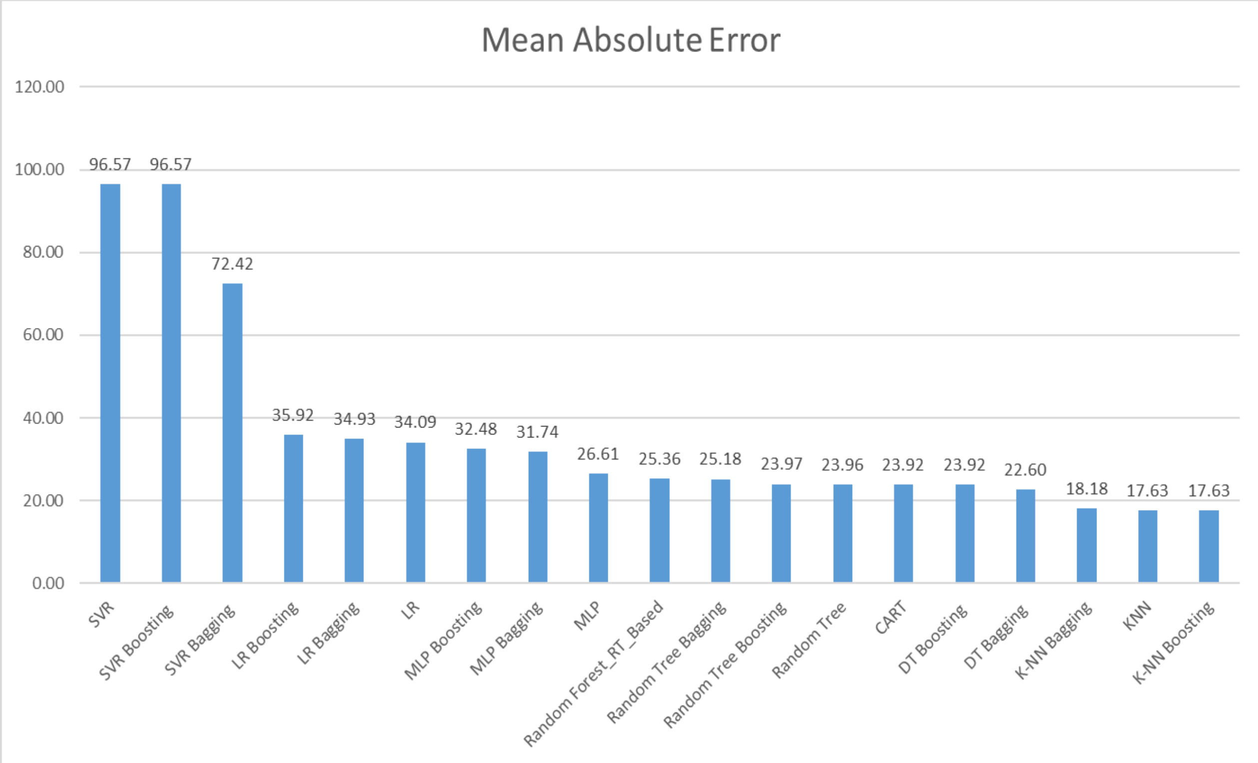

Figure 1. MAE for the prediction models over the 2000-2006 (Before crisis) dataset

In addition to the 2000-2006 dataset results, the experimental results for During crisis dataset are shown in Figure 2. The MAE values are all high, regardless of the method; the MAE results are significantly higher than those obtained with the Before crisis dataset. These results are probably caused by the subprime mortgage event and financial crisis, which resulted in massive damage to various industries in the US and made it challenging to predict Tobin's Q value with a lower MAE. The Random Tree and Random Tree Boosting-based methods outperform the other methods with MAEs of 503.94 and 504.76, respectively. On the other hand, the LR Boosting-based method performs the worst, with an MAE value of 2721.78. In contrast to the experiment using the Before crisis dataset, there is a change in the companies' financial status and Tobin's Q value, which has become volatile.

The results prove that the ensemble learning methods do not necessarily perform better than the best single classifier. This result is consistent with the results obtained with the Before crisis dataset.

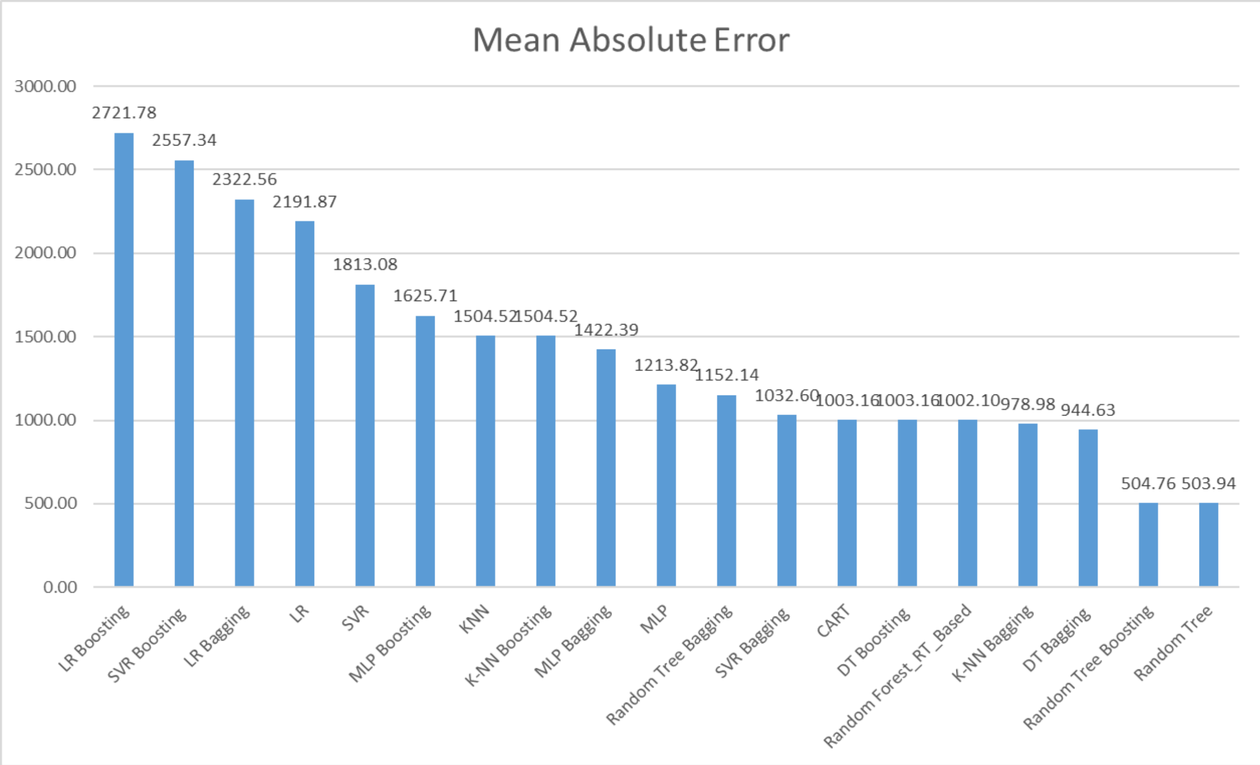

Figure 2. MAE for the prediction models over the 2007-2008 (During crisis) dataset

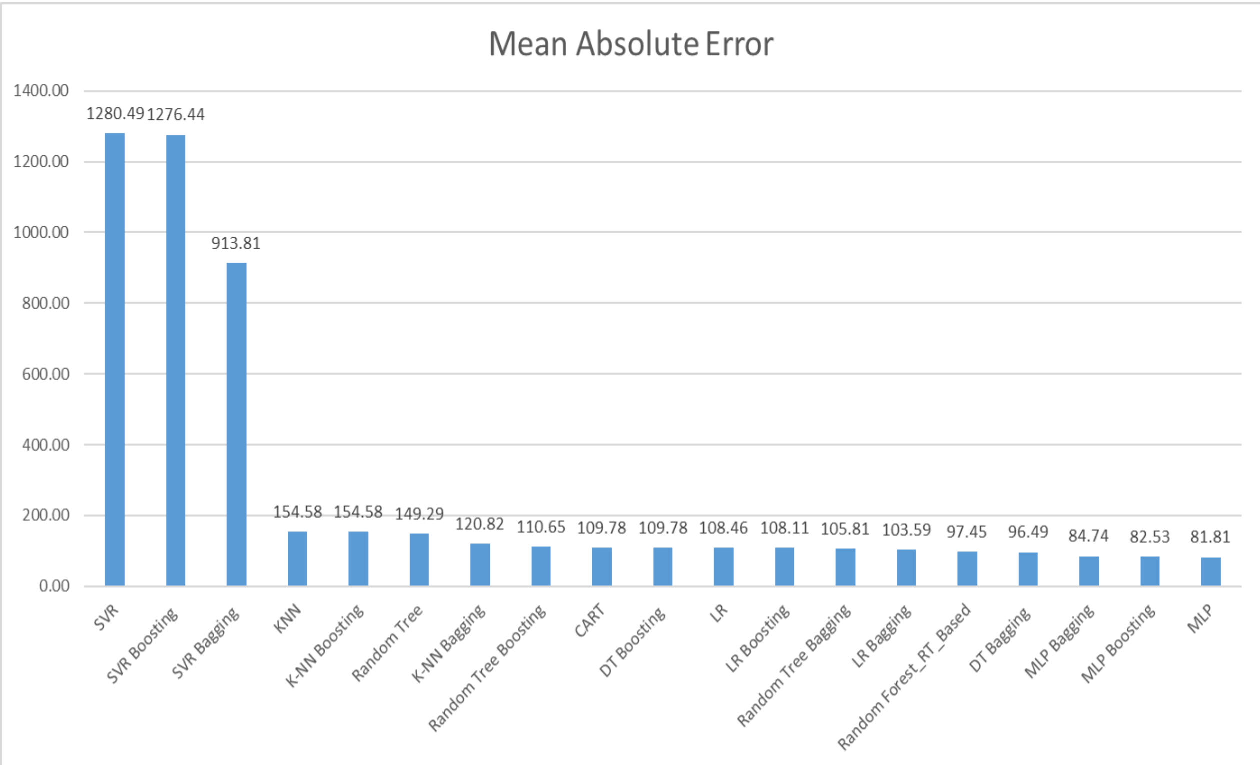

Figure 3. MAE for the prediction models over the 2009-2019 (After crisis) dataset

Figure 3 presents the experimental results for the 2009-2019 dataset. Most algorithms obtain relatively low MAEs for this dataset, with the SVR-related methods, including SVR, SVR Boosting, and SVR Bagging methods producing higher results. Otherwise, the results prove that the ensemble learning methods do not necessarily perform better than the best single classifier. On the other hand, to correctly forecast Tobin's Q value for the industry, the MLP algorithm is the best, with the lowest predicted MAE value.

After reviewing the experimental results, this study finds varying prediction performance for the algorithms and determines that the difficulty of correctly forecasting Tobin's Q value is different for the three datasets. Most algorithms obtained similar MAE values for the Before crisis dataset to the SVR and SVR-based methods. The MAE values descend with a low slope pattern, implying that most algorithms have a similar ability to forecast Tobin's Q value correctly. Furthermore, the results consistently show that the ensemble learning methods, even the commonly used Random Forest methods, perform no better than the best single classifier. A well-performing single classification technique produces the optimal intangible assets value prediction model.

As shown in Figure 2, the prediction performance of the algorithms for During crisis dataset is quite different from that for the Before crisis dataset. The MAE values descend with a higher slope, which results in a considerable distinction between the two datasets. As the same algorithms applied over both datasets, the prediction results reflect the significant increase in the complexity of the structure of the During crisis dataset in comparison to the previous one. The considerable increase in MAE values supports this viewpoint. The During crisis dataset is also helpful for testing the predictive ability of the algorithms, as the steep slope represents the score for each algorithm. The subprime mortgage event and financial crisis should be the reasons for this.

In 2007, the subprime mortgage crisis undermined the US financial market, resulting in global credit and liquidity issues and altering the structure of the world financial market. In addition, the crisis became the cause of much more rigorous regulations in the US and led to changes in the financial structuring of industries. Before this, the adoption of International Financial Reporting Standards (IFRS) in the US had been debated by regulators, practitioners, and academicians since 2002. Although the Securities and Exchange Commission (SEC) did not switch from the US generally accepted accounting principles (GAAP) to IFRS, it has continued reviewing a proposal to allow IFRS information to supplement US financial filings since 2009.

The pattern of prediction performance for the After crisis dataset is similar to the Before crisis dataset; the SVR and SVR-related methods produce the worst MAE values 985compared to the other algorithms. However, the MAE values for the After crisis period are significantly higher than those for the Before crisis period, which means that, in most cases, it has become more difficult to precisely forecast Tobin's Q. The results illustrate how the financial crisis and IFRS have affected the financial structure of companies in the US. However, no single method can also outperform all other algorithms over all three different periods. This situation may partially be due to changes in the data structure.

Let us now consider the consistency of the prediction outcomes by reviewing the RMSE values for all algorithms over the three datasets. Table 3 indicates that the lowest RMSE values are obtained using the Random Tree and MLP algorithms on the During and After-crisis datasets. This result means that the prediction outcomes of these two algorithms are less varied than other methods. The KNN algorithm was chosen to conduct further analysis based on the MAE results, even though it did not produce the lowest RMSE result for the Before crisis dataset.

Table 3. RMSE results for the algorithms

Root Mean Squared Error			
Model	**2000-2006**	**2007-2008**	**2009-2019**
CART	1051.2959	41009.5646	6365.5220
DT Bagging	1053.6295	41777.4630	5962.4185
DT Boosting	1051.2959	41009.5646	6365.5220
KNN	**1054.3329**	71023.3676	10567.1401
K-NN Bagging	1056.5349	47951.8105	8116.6431
K-NN Boosting	1054.3329	71023.3676	10567.1401
LR	1046.8121	52182.7437	5871.3117
LR Bagging	1046.8055	80343.3759	5872.6837
LR Boosting	1046.9480	78830.2892	5871.642
MLP	1046.6830	41054.1537	**5871.2197**
MLP Bagging	1047.0310	41128.2676	5871.2887
MLP Boosting	1046.8985	41094.8292	5871.2292
Random Forest	1161.0700	49162.4610	6321.8584
Random Tree	1357.5999	**41006.0976**	10857.4368
Random Tree Bagging	1128.1647	49680.2737	6327.9301
Random Tree Boosting	1357.4313	41006.1500	8754.2807
SVR	1050.1062	41041.0827	5997.6841
SVR Bagging	1047.9682	41011.5624	5930.5463
SVR Boosting	1050.1062	41077.2547	5999.3961

4.2 Results for the Individual Industries

As discussed above, there is some discrepancy in prediction performance between the three datasets even though the applied algorithms are precisely the same. Moreover, the datasets include data for various industries, which raises the question: "Is the same algorithm also the best for each industry?" In order to find the answer, this study separates the three datasets into several sub-datasets according to industry category from the three stock exchanges (i.e., NASDAQ, AMEX, and NYSE), resulting in 155, 131, and 157 sub-datasets for the three periods, respectively, as shown in the Appendix.

4.2.1 Before Crisis

For industries before the crisis, the algorithm with the lowest MAE (i.e., KNN) for the Before crisis dataset was used to conduct the experiments. Figure 4 shows the results for the top 10 industries in descending order; the MAE values in the figure indicate the good prediction performance of the KNN algorithm, except for the Health Care Technology industry. The KNN algorithm performs well for predicting Tobin's Q value for most industries before the crisis and is a reliable tool for the user. Health Care Technology is the only industry with a worse MAE value, but this might have been because of the highly complicated structure of the financial data and Tobin's Q value. The data are composed of financial indicators and corresponding Tobin's Q values, from which the algorithm learns the instances. Similar financial indicators combined with volatile Tobin's Q values may increase the difficulty of learning and forecasting.

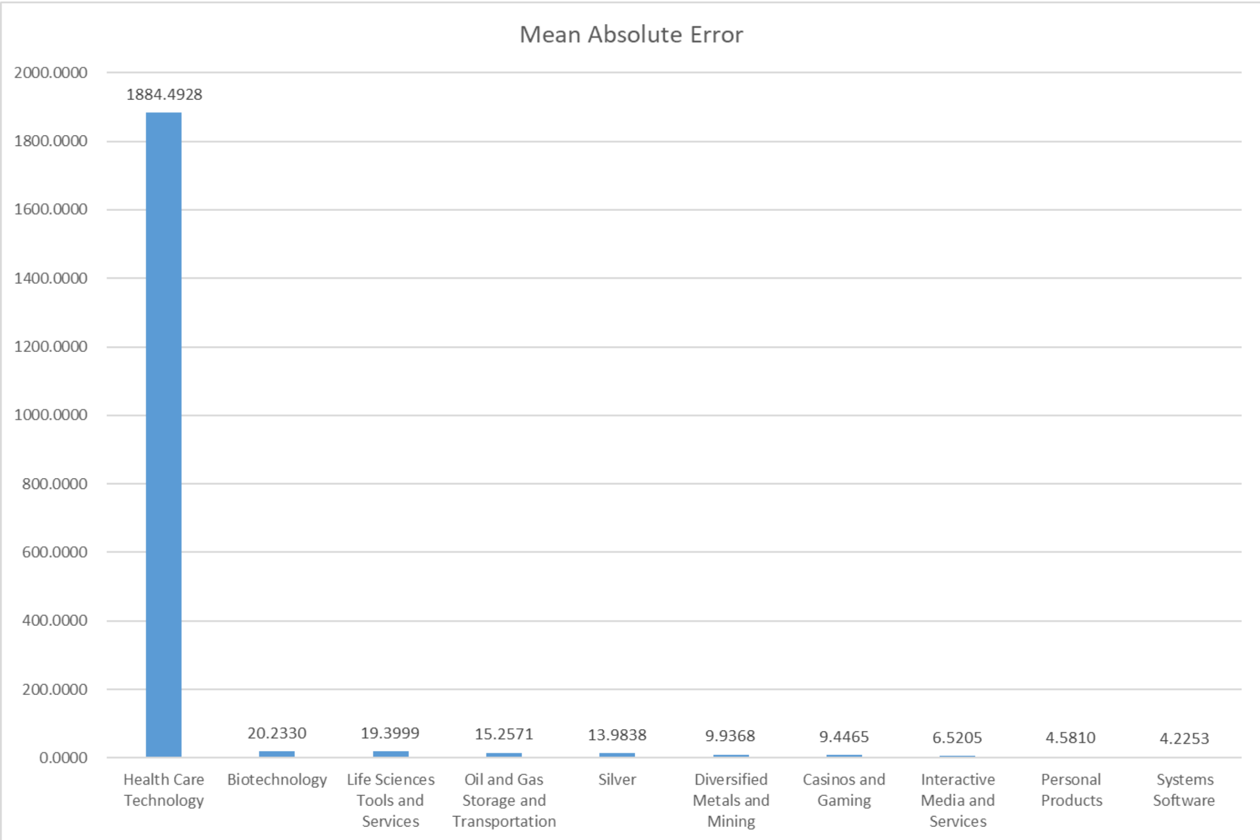

Figure 4. MAE values for the various industries obtained using the Before crisis dataset; the value for the Health Care Technology industry is clearly higher than the others

4.2.2 During Crisis

Figure 5 presents the results for the During crisis dataset. The best algorithm (i.e., Random Tree) maintained good performance for most industries except Biotechnology. Moreover, the low MAE values in the figure make it easier to precisely forecast Tobin's Q, meaning high consistency between the financial indicators and the corresponding Tobin's Q values in most industries. Such a result is hard to obtain, especially during a huge event.

In contrast, the Random Tree algorithm failed to predict Tobin's Q value for the Biotechnology industry. Part of the reason may be the consistency issue in the Before crisis dataset, and part may be due to the unusually high Tobin's Q values for a few companies.

Figure 5. MAEs for the various industries for the During crisis dataset; the Biotechnology industry value is significantly higher than the others

4.2.3 After Crisis

The After crisis dataset presented in Figure 6 reveal a significant increase in the difficulty of precisely forecasting Tobin's Q. There is a sharp increase in the MAE values of the top 10 industries, especially in comparison to the previous two periods. Such a result may reflect doubts about the quality of financial earnings reports after adopting the IFRS. In addition, some external users perceive the US GAAP to be higher than the IFRS (Coopers, 2009). Finally, it has been noted in some studies that momentum towards the adoption of IFRS in the US has slowed (Atwood, Drake, Myers, & Myers, 2011) and some have recently recommended reversing the tentative SEC decision to adopt the IFRS (Selling, 2013).

Despite the considerations above, worldwide adoption of the IFRS would potentially enhance cross-border comparability and the usefulness of accounting information to external users by enhancing disclosure and transparency and reducing information costs and asymmetry (Ball, 2006; Barth, Landsman, & Lang, 2008). In addition, the widespread application of the IFRS by non-US firms has improved financial reporting comparability with US firms (Barth, Landsman, Lang, &

Williams, 2012). This result may make it easier to apply the intangible assets value prediction model in samples from various countries in future research.

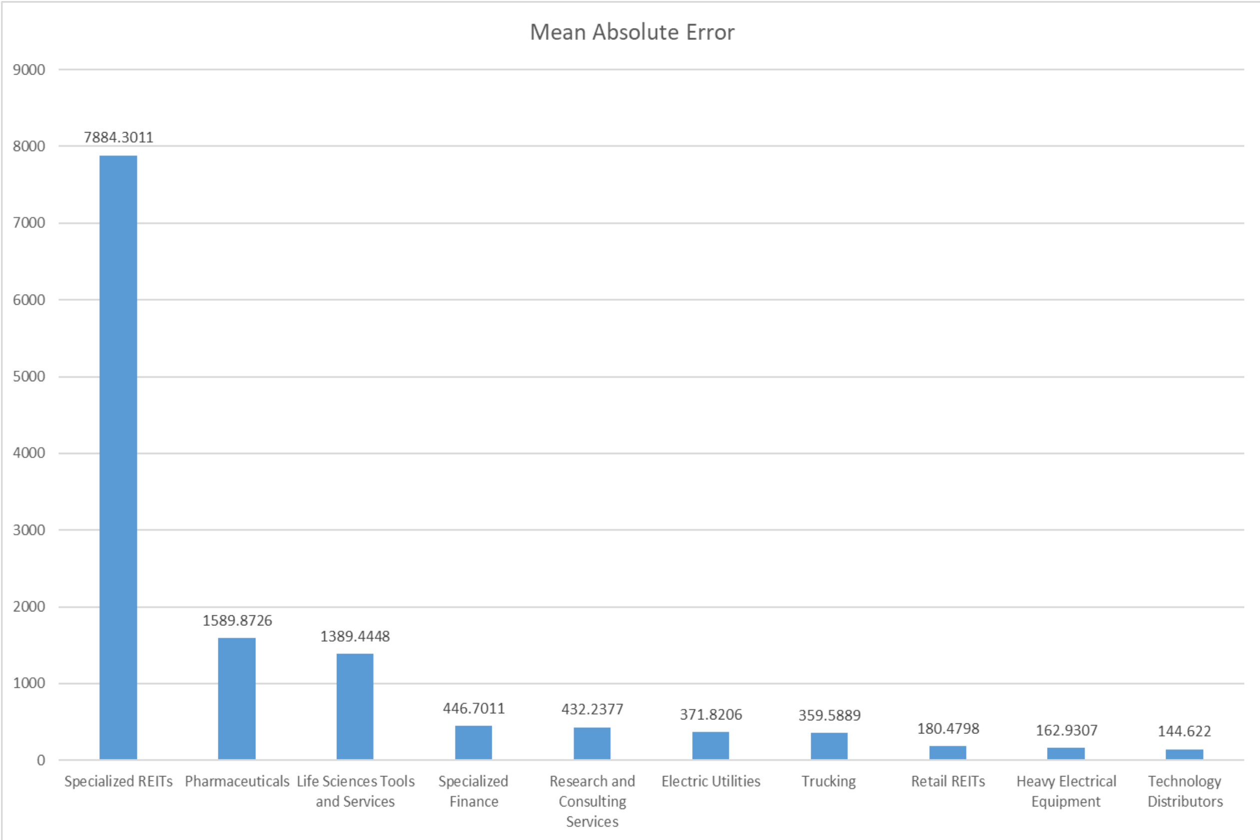

Figure 6. MAE values obtained for various industries using the After crisis dataset

4.3 Feature Importance for Different Periods

In addition to providing the results of prediction performance, this study also shows the importance of variables in different periods. These results can help the user to understand which feature is the most important in different periods. The rank of average merit of features in different periods is shown in Table 4-7. In Table 4(i.e., Before Crisis), firm size affected Tobin's Q significantly, and SIZE with 0.6044 average merit is the most important feature when we predicted intangible assets value. Much prior literature indicated that firm size is likely to be inversely related to expected growth opportunities and had a significantly negatively relationship with Tobin's Q (Asghar et al., 2020; Fukui & Ushijima, 2007; Gleason & Klock, 2006; Kong et al., 2020; Li, 2016; Mishra & Kapil, 2017; Silva et al., 2019; Yoong et al., 2015). PROFITABILITY and INDUSTRY are the second and third features in the rank; however, their average merits are significantly lower than SIZE.

However, during Crisis (i.e., Table 5), the rank of features is changed. In 2007, the subprime mortgage crisis injured the US financial market; in this situation, if the companies can make earnings or survive will become a critical factor when investors assess an investment opportunity. PROFITABILITY with 0.6164 average merit is the most important feature when we predicted intangible assets value in this period. LEVERAGE and INDUSTRY are the second and third features in the rank; however, the average merit is significantly lower than PROFITABILITY. In addition, during Crisis, SIZE becomes an unimportant feature different from Before Crisis.

The results of After Crisis shows in Table 6. After the crisis, the issue of company survival continues to attract public attention. The long-term solvency measured by LEVERAGE becomes the most important feature, followed by PROFITABILITY. SIZE with 0.1129 average merit is ranked and becomes the third feature in the rank. Furthermore, the average merit differences of the first three features are reduced.

Finally, as all periods are shown in Table 7, the results indicated that the first three features are PROFITABILITY, INDUSTRY, and ADVERTISING INTENSITY. The average merit of PROFITABILITY is the highest, and INDUSTRY is also a critical feature. According to the above results, we can confirm that the study's different periods and industry experiments are reasonable.

Table 4. Rank of Average merit of feature in Before Crisis period (2000~2006)

Features	Average merit
SIZE	***0.6044***
PROFITABILITY	***0.1367***
INDUSTRY	***0.1225***
CAPITAL INTENSITY	0.0873
DIVIDEND	0.0367
LEVERAGE	0.0046
EARNING MANAGEMENT	0.0041
SALE GROWTH	0.0031
R&D INTENSITY	0.0003
CAPITAL EXPENDITURE	0.0003
ADVERTISING INTENSITY	0.0001

Table 5. Rank of Average merit of feature in During Crisis period (2007~2008)

Features	Average merit
PROFITABILITY	***0.6164***
LEVERAGE	***0.1099***
INDUSTRY	***0.1026***
R&D INTENSITY	0.0909
SIZE	0.0794
EARNING MANAGEMENT	0.0006
SALE GROWTH	0.0001
CAPITAL INTENSITY	0.0001
ADVERTISING INTENSITY	0.0000
CAPITAL EXPENDITURE	0.0000
DIVIDEND	0.0000

Table 6. Rank of Average merit of feature in After Crisis period (2009~2019)

Features	Average merit
LEVERAGE	***0.3040***
PROFITABILITY	***0.1930***
SIZE	***0.1129***
CAPITAL EXPENDITURE	0.0658
R&D INTENSITY	0.0591
ADVERTISING INTENSITY	0.0587
DIVIDEND	0.0568
EARNING MANAGEMENT	0.0530
SALE GROWTH	0.0444
INDUSTRY	0.0426
CAPITAL INTENSITY	0.0098

Table 7. Rank of Average merit of feature in After Crisis period (2009~2019)

Features	Average merit
PROFITABILITY	**0.8222**
INDUSTRY	**0.1120**
ADVERTISING INTENSITY	**0.0335**
LEVERAGE	0.0218
EARNING MANAGEMENT	0.0048
R&D INTENSITY	0.0027
DIVIDEND	0.0014
CAPITAL INTENSITY	0.0007
SIZE	0.0006
CAPITAL EXPENDITURE	0.0001
SALE GROWTH	0.0000

4.4 Results for Moving Window-Based and Fixed-Based Evaluations

Business operations can be viewed as a time series, where the company’s financial status and the corresponding Tobin’s Q value will continuously be affected by the economy and the business environment. Therefore, this study adopts a moving window-based evaluation method to verify the prediction algorithms in detail.

The moving window-based method starts with the first year's data and continues until the last year's data. For example, the prediction model is trained using each year's data and verified by the following year's data. No study in the literature uses a moving window-based method to evaluate a model for predicting the value of Tobin's Q. To identify the most suitable periods to train the prediction model, we applied training sets with one year, two years, and three years of data and verified the model using the data from the following year. Moreover, a fixed-based method is also applied in this study.

In contrast to the moving window-based method, the fixed-based method assembles the data yearly. For example, data for the year 2000 are initially used as a training set, and the dataset for the following year (2001) is used to verify the model. After the verification process, the fixed-based method assembles the 2000s and 2001 to produce a new training set for building the prediction model. Then the dataset for the following year (2002) is used as a test set, and so on.

4.4.1 Moving Windows Based on One Year

Figure 7 shows the results for moving windows based on one year. The MAE values of the three algorithms (i.e., KNN, MLP, and Random Tree) are similar from 2001 to 2007 but sharply increase in 2007. This situation reflects the influence of the financial crisis. However, the performance of the three algorithms in the next year (2008) varies. There is a decrease in the MAE value for the Random Tree algorithm, which is similar to those for the period from 2001 to 2006; the KNN result is similar to the 2007 result, and the MAE obtained with the MLP method is higher than the 2007 value. The diversified results reflect the similarity of the data for 2007 and 2008. Therefore, it is not difficult to use the Random Tree approach to build a prediction model through learning the data for 2007 (i.e., subprime mortgage event) and then predict a precise result for 2008 (i.e., financial crisis).

After the financial crisis, the prediction performances again diverged in 2010. However, the Random Tree method still outperforms the other algorithms with the lowest MAE value, and there is no significant divergence in the performance over the remaining data.

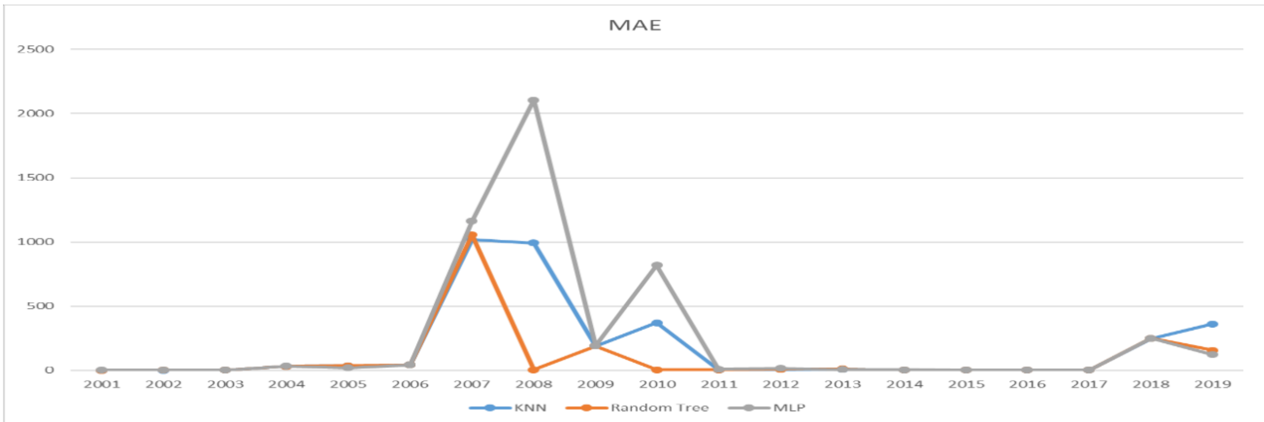

Figure 7. MAE results for the One Year-based Moving Window method; Random Trees produces the lowest value

Figure 8 shows the results for the RMSE. The results are similar to the MAE results. However, the Random Tree algorithm outperforms the other algorithms, especially for data near and after the crisis. It may be because of the structure of the data. The event and crisis occurred in 2007 and 2008, respectively, and the movement of the window is based on one year. Therefore, the MAE and RMSE results show the Random Tree algorithm to be the most suitable for predicting Tobin's Q for moving windows based on one year.

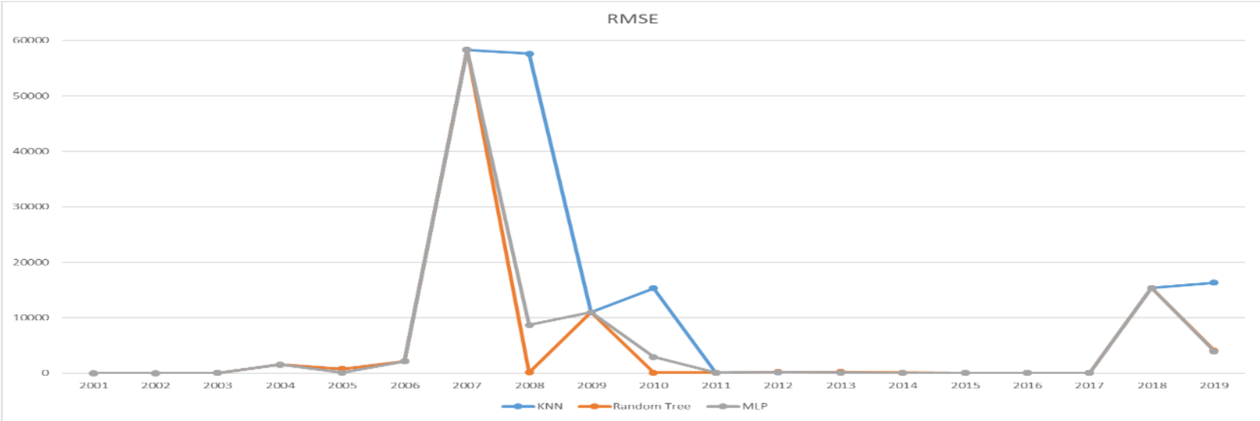

Figure 8. RMSE results for the One Year-based Moving Window method: Random Trees produces the lowest RMSE for most of the periods

4.4.2 Moving Windows Based on Two Years

The MAE and RMSE results for two-year-based moving windows are shown in Figures 9 and 10. In Figure 9, it can be seen that the Random Tree and the KNN methods obtained lower MAE values than the MLP on the data for 2007, 2008, 2009, and 2011. The results suggest that these two algorithms maintained similar learning and verification ability during these periods. In addition, there was a decrease in the Random Tree MAE in 2010 but a sharp increase for the KNN MAE. Since the training data are comprised of a combination of the 2008 and 2009 datasets, the structure may not be favorable for the KNN algorithm. However, the Random Tree outperforms the other two algorithms in the MAE.

Considering variance in the prediction results of the algorithms, the MLP outperforms the others in terms of the RMSE. Although the performance of the Random Tree algorithm is similar for most periods, it fell behind the MLP in 2008. For seeking the most suitable algorithm for predicting the value of Tobin's Q, the MAE and RMSE performance should be considered simultaneously. Since the Random Tree method produces the lowest MAE for the entire period and only produces a high RMSE in 2008, it should be recognized as the most suitable algorithm for two-year-based moving windows. Figure 10 shows the results below.

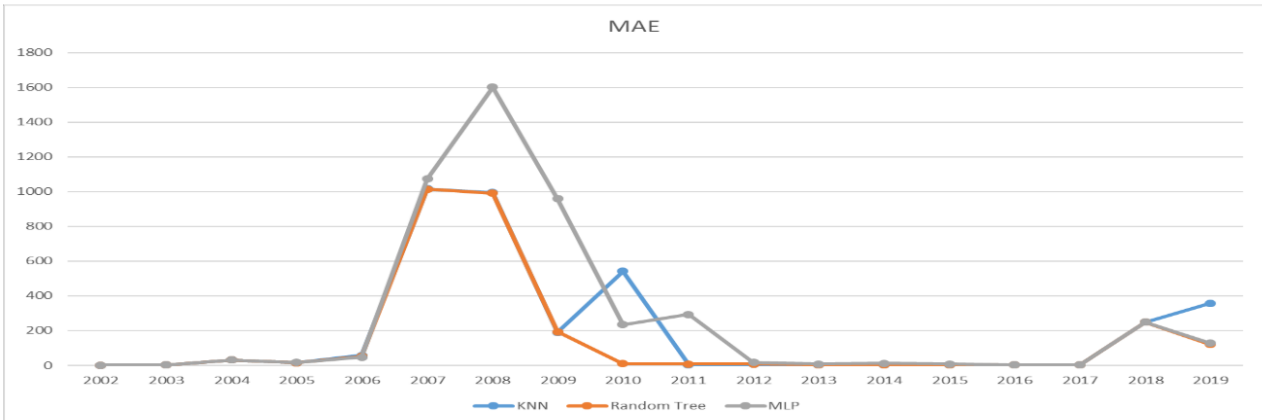

Figure 9. MAE results for the Two Year-based Moving Window method: Random Trees outperforms the other algorithms

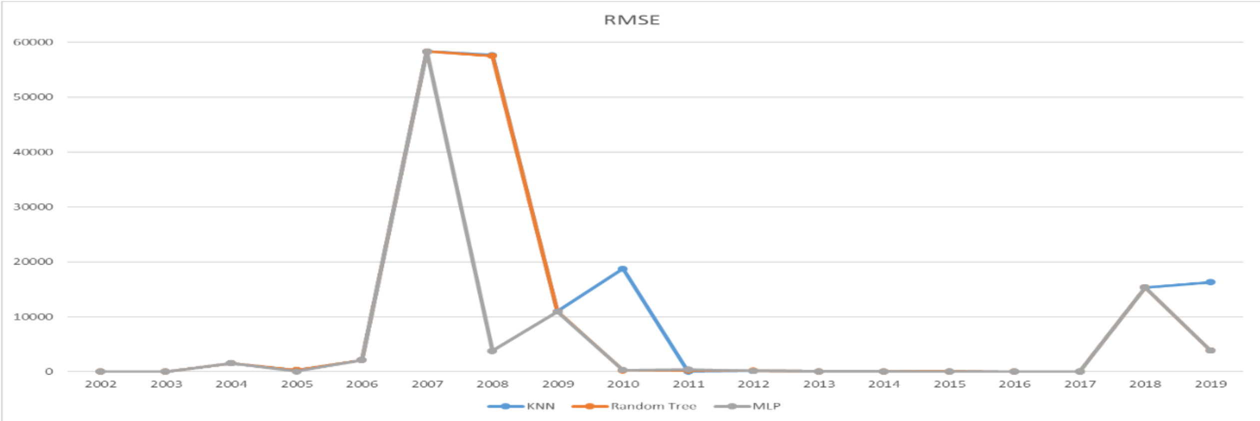

Figure 10. RMSE results for the Two Year-based Moving Window method: MLP outperforms the other algorithms

4.4.3 Moving Windows Based on Three Years

The moving windows based on three years are shown in Figures 11 and 12. The algorithms produced volatile results in contrast to the previous sections for the data from 2007 to 2012. Generally, the three-year-based moving window data are comprised of three years of data instances. More data for learning should improve the verification ability of the prediction model; however, the volatile results for the measure of MAE are not consistent with this supposition. The reason may be that there might be quite a difference between the structure of the training and the testing data, in which case it would be difficult for the prediction model to precisely identify the corresponding Tobin's Q value for the testing data.

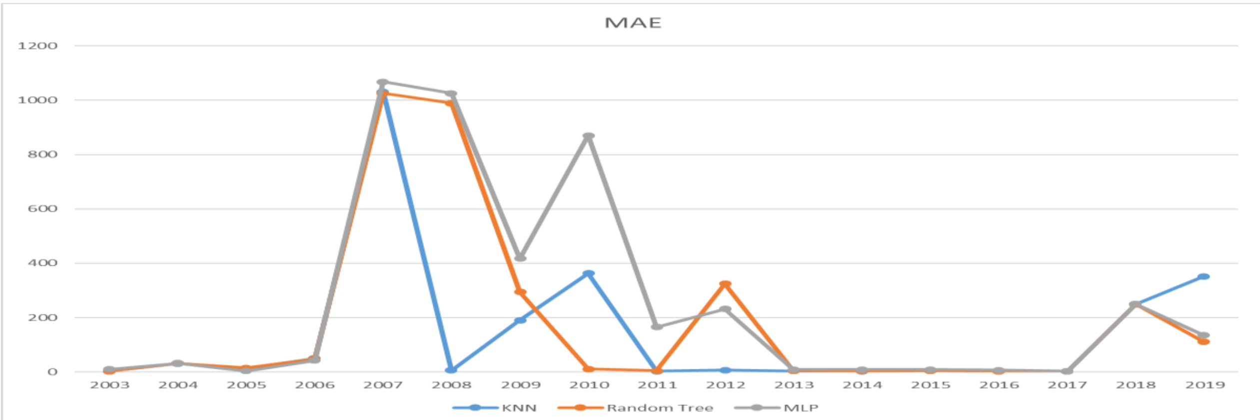

Figure 11. MAE results for the Three Year-based Moving Window method: KNN produces the lowest MAE

As shown in figure 12, although the performance of the algorithms becomes volatile when the movement is based on three years, one algorithm still stands out as more suitable. The KNN algorithm outperforms the other algorithms in most years for MAE and RMSE, except in 2010.

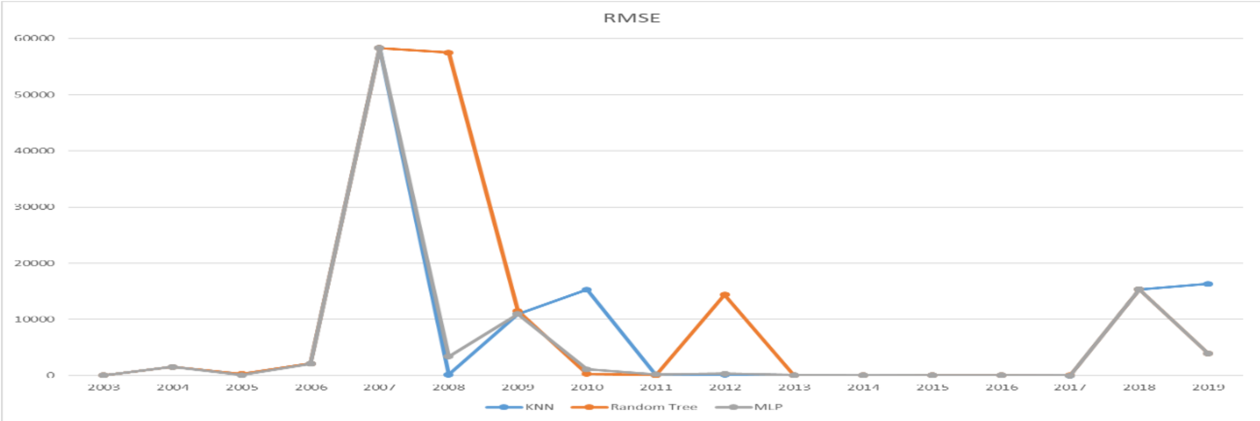

Figure 12. RMSE results for the Three Year-based Moving Window method: KNN produces the lowest RMSE for most of the periods

4.4.4 Fixed Based Results

In addition to the moving window-based method, a fixed-based learning mechanism was adopted to determine the best learning mechanism. In this method, the data for years n and n + 1 are used as the training and testing data for the learning and verification process. After completion of the process, then the year n and n + 1 become the training dataset, and the year n + 2 becomes the testing dataset. This process continues until the last year becomes the testing data, and the others become training data.

The results for the fixed-based method are shown in figures 13 and 14 below. It is not difficult to see that the fixed-based method results in a higher MAE, especially for the MLP algorithm. On the other hand, given that the purpose is to find the most suitable algorithm for predicting Tobin's Q, the KNN method must be the best candidate due to its less volatile MAE than the Random Tree method. Furthermore, in terms of the measure of RMSE, KNN also outperforms the other methods. According to the MAE and RMSE results, KNN seems the most suitable algorithm for the fixed-based learning mechanism.

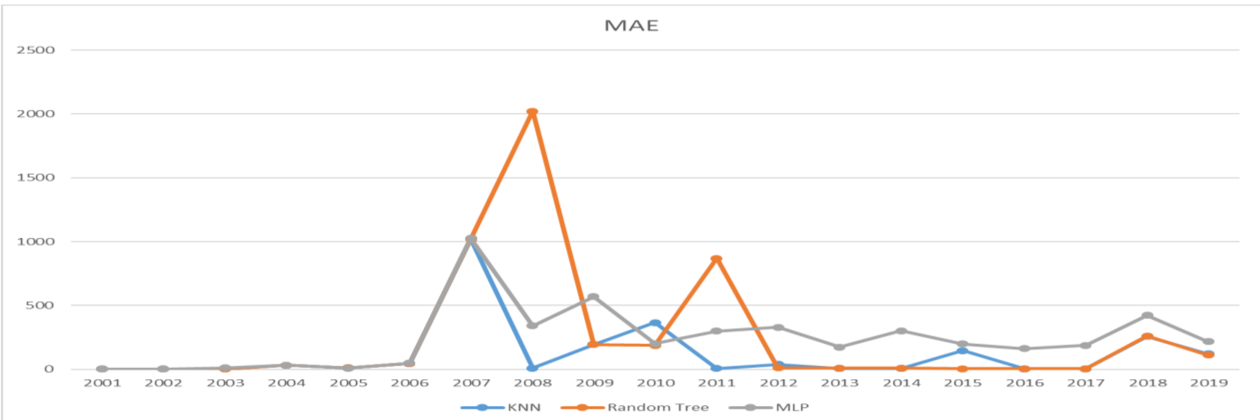

Figure 13. MAE results for the fixed-base method: KNN gives the lowest MAE

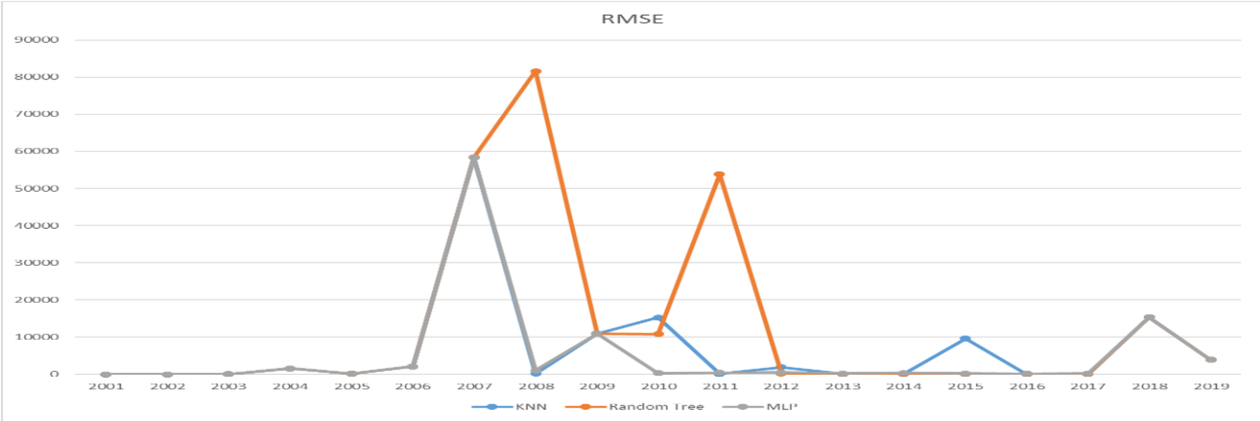

Figure 14. RMSE results for the fixed-base method: MLP gives the lowest RMSE for most of the periods

4.4.5 Comparison of Moving Window-Based and Fixed-Based Learning Mechanisms

After reviewing the results for the moving-window-based and fixed-based learning mechanisms, this study provides the four most optimal algorithms for predicting the value of Tobin’s Q under each set of circumstances under different learning mechanisms. We, therefore, try to identify the best learning mechanism and most suitable algorithm. The results show in figures 15 and 16. The algorithm performance is similar for most data, except for 2007, 2008, 2009, and 2010. Therefore, these periods can further be used to test the ability of these methods to precisely forecast Tobin’s Q value. According to the experimental results, the lowest MAE is obtained with the 1Y Random Tree algorithm (i.e., use one-year data to train the algorithm) when used on 2008, 2009, and 2010 datasets; the MAE is similar to that of the others in the year 2007. Moreover, the 1Y Random Tree algorithm produces the best RMSE results on 2008 and 2010 datasets; and results similar to the others for 2007 and 2009.

Considering the Random Tree algorithm's prediction performances and learning mechanism, it would be the most suitable method for predicting Tobin's Q value. Furthermore, since the learning mechanism is moved on a one-year basis, the Random Tree needs only one year of data as the training set to train the algorithm and produce superior performance most of the time, even during the financial crisis.

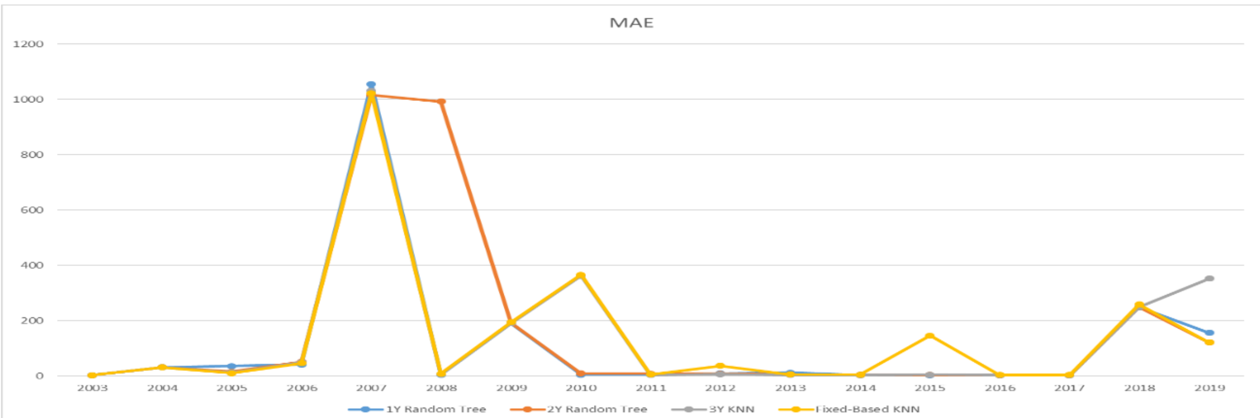

Figure 15. MAE results: 1Y Random Tree outperforms the other methods

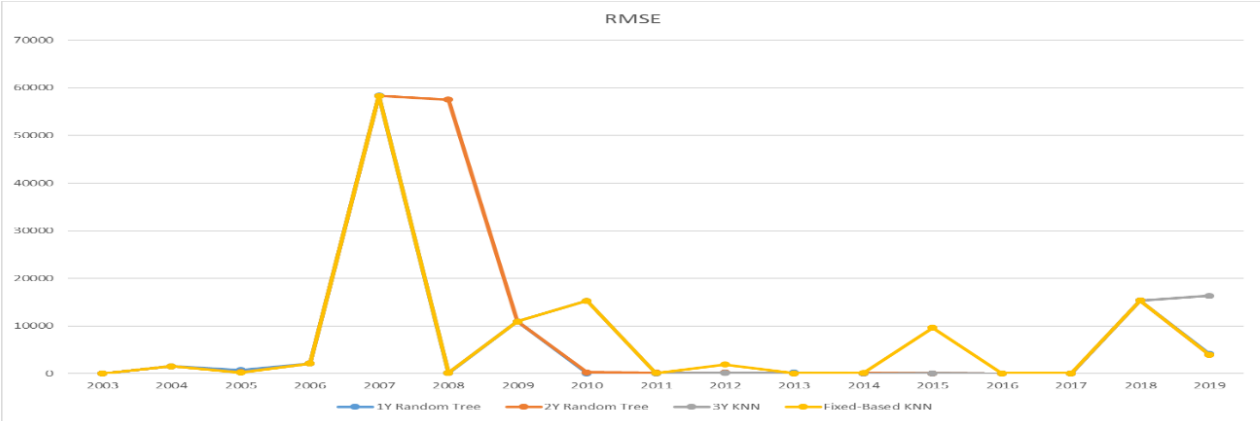

Figure 16. RMSE results: 1Y Random Tree outperforms the other methods

The results indicate that the ensemble learning methods do not necessarily perform better than the best single classifier. A well-performing single classification technique produces the optimal model for intangible assets value prediction in three periods (i.e., before, during, and After crisis). Therefore, three optimal algorithms, KNN, Random trees, and MLP are used to construct intangible assets value models for various industries in these three periods. Finally, moving window-based and fixed-based evaluation methods are adopted to verify the prediction algorithms since the data may be affected by changes in the dynamic economy and business environment over time. The results indicate that the 1Y Random Tree method produces the best results in predicting intangible assets values. These prediction models can provide valuable information to the management and stockholders to assess companies' status and make more effective decisions.

5. Conclusion

Along with the evolution of the knowledge-based economy in recent decades, the primary method for creating firm value has moved from traditional physical assets to intangible knowledge. As a result, intangible assets become an essential proportion of the valuation of a company; therefore, it is crucial to carefully evaluate intangible assets in the enterprise value evaluation (Chan, Lakonishok, & Sougiannis, 2001; Eckstein, 2004). However, it is more difficult to measure the future potential of intangibles than to evaluate the benefits of investment in fixed assets. Therefore, a prediction model with precise results to evaluate intangible assets value is critical. This study focuses on comparing the ability of various machine learning algorithms to make a more accurate evaluation and prediction model for assessing intangible assets value. To be specific, six single classification algorithms are compared, Linear regression (LR), Multi-layer perception (MLP), Support Vector Machine (SVM), Classification and Regression Tree (CART), Random Tree (RT), and K-Nearest Neighbors (KNN); two ensemble learning methods (i.e., Boosting and Bagging) are also examined. The methods are tested for different periods and various industries.

The practical contribution of this study includes as follows. For the management and stockholders, the prediction results of the model can provide a precise estimating value for intangible assets of the companies when they are evaluating the operating performance. This model assists the management and stockholders in identifying performance in the decision-making process for business operation. On the other hand, the auditors should audit the financial reports released from the companies, in which the predicted intangible assets value from the prediction model can be used as a reference for the auditors when auditing intangible assets of the companies.

It should be noted that although several widely used techniques are considered to develop the baseline prediction models, other algorithms for modifying current single classifier techniques could also provide improved performance. Furthermore, for future work, other clustering and classification methods could be applied to compare the prediction models discussed here to reach a more reliable conclusion, for example, K-means, Self-organizing map, Convolutional neural network, etcetera. However, it is difficult to comprehensively study all existing methods available in this area from a practical standpoint. In addition, different input variables could also be considered for comparison purposes since different studies identify them as necessary for affecting intangible assets values. Furthermore, different techniques and data sources could also be considered, such as Text mining techniques and companies news. Finally, as per the earlier discussion, this study mainly evaluates the intangible assets value. However, this procedure could be extended to other domain problems in future work.

References

Adriessen, D. and R. Tissen, (2000), *Weightless wealth: find your real value in a future of intangibles assets*, London: Financial Times Prentice Hall.

Ahmed Haji, A. and N. A. Mohd Ghazali, (2018), "The Role of Intangible Assets and Liabilities in Firm Performance: Empirical Evidence," *Journal of Applied Accounting Research*, 19(1), 42-59.

Allayannis, G. and J. P. Weston, (2001), "The Use of Foreign Currency Derivatives and Firm Market Value," *Review of Financial Studies*, 14(1), 243-276.

Asghar, A., S. Sajjad, A. Shahzad, and B. T. Matemilola, (2020), "Role of Discretionary Earning Management in Corporate Governance-Value and Corporate Governance-Risk Relationships," *Corporate Governance: The International Journal of Business in Societ*, 20(4), 561-581.

Atwood, T., M. S. Drake, J. N. Myers and L. A. Myers, (2011), "Do Earnings Reported under IFRS Tell Us More About Future Earnings and Cash Flows?" *Journal of Accounting Public Policy*, 30(2), 103-121.

Ball, R., (2006), "International Financial Reporting Standards (IFRS): pros and cons for investors,"*Accounting Business Research*, 36, 5-27.

Barth, M. E., Y. Konchitchki, and W. R. Landsman, (2013), "Cost of Capital and Earnings Transparency," *Journal of Accounting Economics*, 55(2-3), 206-224.

Barth, M. E., W. R. Landsman, M. Lang, and C. Williams, (2012), "Are IFRS-Based and US GAAP-Based Accounting Amounts Comparable?" *Journal of Accounting Economics*, 54(1), 68-93.

Barth, M. E., W. R. Landsman, and M. H. Lang, (2008), "International Accounting Standards and Accounting Quality," *Journal of Accounting Research*, 46(3), 467-498.

Bishop, C. M., (1995), *Neural networks for pattern recognition*, Oxford university press.

Black, B. S., H. Jang, and W. Kim, (2006), "Does Corporate Governance Predict Firms' Market Values? Evidence from Korea," *The Journal of Law, Economics, and Organization*, 22(2), 366-413.

Breiman, L., (1996), "Bagging Predictors," *Machine Learning*, 24(2), 123-140.

Breiman, L., (2001), "Random Forests," *Machine Learning*, 45(1), 5-32.

Breiman L, F. J., R. Olshen, C. Stone, (1984), *Classification and regression trees*, New York: Chapman & Hall/CRC.

Buallay, A., A. Hamdan, and Q. Zureigat, (2017), "Corporate Governance and Firm Performance: Evidence from Saudi Arabia," *Australasian Accounting, Business Finance Journal*, 11(1), 78-98.

Burgman, R. and G. Roos, (2007), "The Importance of Intellectual Capital Reporting: Evidence and Implications," *Journal of Intellectual Capital*, 8(1), 7-51.

Chan, L. K., J. Lakonishok, and T. Sougiannis, (2001), "The Stock Market Valuation of Research and Development Expenditures," *The Journal of Finance*, 56(6), 2431-2456.

Coopers, P. W., (2009), "Views of the Capital Market," A Panel Discussion on IFRS Video Learning Center.

Cover, T. and P. Hart, (1967), "Nearest Neighbor Pattern Classification," *IEEE transactions on information theory*, 13(1), 21-27.

Dang, N., D. C. Pham, and T. Vu, (2018), "Effects of Financial Statements Information on Firms' Value: Evidence from Vietnamese Listed Firms," *Investment Management Financial Innovations*, 15(4), 210-218.

Eckstein, C., (2004), "The Measurement and Recognition of Intangible Assets: Then and Now," *Paper presented at the Accounting Forum*, 28(2), 139-158.

Edvinsson, L., (2013), "IC 21: Reflections from 21 Years of IC Practice and Theory," *Journal of Intellectual Capital*, 14(1), 163-172.

Eleftherios, P., G. Dimitrios, and B. Panagiotis, (2021), "Can Intangible Assets Predict Future Performance? A Deep Learning Approach," *International Journal of Accounting & Information Management*, 30(1), 61-72.

Freund, Y. and R. E. Schapire, (1996), "Experiments with a New Boosting Algorithm," *Paper presented at the icml.*

Fukui, Y. and T. Ushijima, (2007), "Corporate Diversification, Performance, and Restructuring in the Largest Japanese Manufacturers," *Journal of the Japanese and International Economies*, 21(3), 303-323.

Gamayuni, R. R., (2015), "The Effect of Intangible Asset, Financial Performance and Financial Policies on the Firm Value," *International Journal of Scientific Technology Research*, 4(1), 202-212.

Gleason, K. I. and M. Klock, (2006), "Intangible Capital in the Pharmaceutical and Chemical Industry," *The Quarterly Review of Economics and Finance*, 46(2), 300-314.

Haji, A. A., (2016), "Trend of Hidden Values and Use of Intellectual Capital Information: Evidence from Malaysia," *Accounting Research Journal*, 29(1), 81-105.

Han, D. and I. Han, (2004), "Prioritization and Selection of Intellectual Capital Measurement Indicators Using Analytic Hierarchy Process for the Mobile Telecommunications Industry," *Expert Systems with Applications*, 26(4), 519-527.

Haykin, S., (1994), *Neural networks: a comprehensive foundation*, Prentice Hall PTR.

Hong, S., (2017), "The Effect of Debt Choice on Firm Value," *Journal of Applied Business Research*, 33(1), 135-140.

Horn, R., M. De Klerk, and C. De Villiers, (2018), "The Association between Corporate Social Responsibility Reporting and Firm Value for South African Firms," *South African Journal of Economic Management Sciences*, 21(1), 1-10.

Huang, Z., H. Chen, C. J. Hsu, W. H. Chen, and S.Wu, (2004), "Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study," *Decision Support Systems*, 37(4), 543-558.

Jain, A. K., R. P. W. Duin, and J. Mao, (2000), "Statistical Pattern Recognition: A review," *IEEE Transactions on pattern analysis machine intelligence*, 22(1), 4-37.

Jentsch, V., (2019), "Board Composition, Ownership Structure and Firm Value: Empirical Evidence from Switzerland," *European business organization law review*, 20(2), 203-254.

Khediri, K. B. and D. Folus, (2010), "Does Hedging Increase Firm Value? Evidence from French Firms," *Applied Economics Letters*, 17(10), 995-998.

Kim, E., W. Kim, and Y. Lee, (2003), "Combination of Multiple Classifiers for the Customer's Purchase Behavior Prediction," *Decision Support Systems*, 34(2), 167-175.

Klock, M. and P. Megna, (2000), "Measuring and Valuing Intangible Capital in the Wireless Communications Industry," *The Quarterly Review of Economics Finance*, 40(4), 519-532.

Kong, Y., T. Famba, G. Chituku-Dzimiro, H. Sun, and O. Kurauone, (2020), "Corporate Governance Mechanisms, Ownership and Firm Value: Evidence from Listed Chinese Firms," *International Journal of Financial Studies*, 8(2), 20.

Lev, B., S. Radhakrishnan, and W. Zhang, (2009), "Organization Capital," *Abacus*, 45(3), 275-298.

Li, B., (2016), "Impact of Information Disclosure Violation on Firm Value in Chinese Listed Firms," *Journal of Applied Business Research*, 32(3), 903-916.

McConnell, J. J. and H. Servaes, (1990), "Additional Evidence on Equity Ownership and Corporate Value," *Journal of Financial Economics*, 27(2), 595-612.

Min, J. H. and Y.-C. Lee, (2005), "Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters," *Expert Systems with Applications*, 28(4), 603-614.

Mishra, R. and S. Kapil, (2017), "Effect of Ownership Structure and Board Structure on Firm Value: Evidence from India," *Corporate Governance: The International Journal of Business in Society*, 28(4), 603-614.

Mitchell, T. M., (1997), *Machine learning*, Burr Ridge, IL: McGraw Hill.

Morris, J. J. and P. Alam, (2012), “Value Relevance and the Dot-Com Bubble of the 1990s,” *The Quarterly Review of Economics Finance*, 52(2), 243-255.

Nekhili, M. and M. Cherif, (2011), “Related Parties Transactions and Firm's Market Value: the French Case,” *Review of Accounting and Finance*, 10(3), 291-315.

Ni, Y., P. Huang, Y. Cheng, and W. C. Huang, (2020), “Does the Role of Relatives on Ownership Structure Affect Firm Value?” *Journal Review of Business*, 40(1), 16-34.

Osinski, M., P. M. Selig, F. Matos, and D. J. Roman, (2017), “Methods of Evaluation of Intangible Assets and Intellectual Capital,” *Journal of Intellectual capital*, 18(3), 470-485.

Penman, S. H., (2009), “Accounting for Intangible Assets: There Is Also An Income Statement,” *Abacus*, 45(3), 358-371.

Quinlan, J. R., (1987), “Simplifying Decision Trees,” *International Journal of Man-machine Studies*, 27(3), 221-234.

Rao, V. R., M. K. Agarwal, and D. Dahlhoff, (2004), “How Is Manifest Branding Strategy Related to the Intangible Value of A Corporation?” *Journal of Marketing*, 68(4), 126-141.

Seal, H. L., (1968), “The Historical Development of the Gauss Linear Model,” Yale University.

Selling, T. I., (2013), “Bumps in the Road to IFRS Adoption: Is A U-turn Possible?” *Accounting Horizons*, 27(1), 155-167.

Silva, J. R., A. F. d. Silva, and B. L. Chan, (2019), “Enterprise Risk Management and Firm Value: Evidence from Brazil,” *Emerging Markets Finance Trade*, 55(3), 687-703.

Stewart, T. and C. Ruckdeschel, (1998), *Intellectual Capital: The New Wealth of Organizations*, Wiley Online Library.

Sveiby, K. E., (1997), *The New Organizational Wealth: Managing & Measuring Knowledge-Based Assets*, Berrett-Koehler Publishers.

Tsai, C. F., Y. H. Lu , Y. C. Hung, and D. C. Yen, (2016), “Intangible Assets Evaluation: The Machine Learning Perspective,” *Neurocomputing*, 175, 110-120.

Tseng, C. Y. and Y. J. James Goo, (2005), “Intellectual Capital and Corporate Value in An Emerging Economy: Empirical Study of Taiwanese Manufacturers,” *RD Management*, 35(2), 187-201.

Vandemaele, S., P. Vergauwen, and A. Smits, (2005), “Intellectual Capital Disclosure in The Netherlands, Sweden and the UK: A Longitudinal and Comparative Study,” *Journal of Intellectual Capital*, 6(3), 417-426.

Vapnik, V., (2013), *The Nature of Statistical Learning Theory*, Springer Science & Business Media.

Vapnik, V. N. and V. Vapnik, (1998), *Statistical Learning Theory* (Vol. 1), Wiley New York.

Yan, X. and X. G. Su, (2009), *Linear Regression Analysis: Theory and Computing*, World Scientific.

Yoong, L. C., E. Alfan, and S. S.Devi, (2015), “Family Firms, Expropriation and Firm Value: Evidence from Related Party Transactions in Malaysia,” *The Journal of Developing Areas*, 49(5), 139-152.

Yunhong, C., (2009), “The Research on the Recognition and Measurement of Intangible Assets for High-Tech Enterprise,” *Management Science and Engineering*, 3(2), 55-60.

Zeb, S. and A. Rashid, (2016), “Impact of Financial Health and Capital Structure on Firm's Value, with Moderating Role of Intangible Assets,” *Global Management Journal for Academic Corporate Studies*, 6(1), 37.