

Integrated Multiple Linear Regression-One Rule Classification Model for the Prediction of Stock Price Trend

^{1,2}Lock Siew Han and ²Md Jan Nordin

¹Malaysian Institute of Information Technology, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia

²Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Article history

Received: 21-12-2016

Revised: 31-03-2017

Accepted: 17-05-2017

Corresponding Author:

Lock Siew Han

Malaysian Institute of Information Technology, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia and

Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

Email: lockshan@ymail.com

Abstract: One of the main problems of predicting stock price with regression approach is overfitting a model. An overfit model becomes tailored to fit the random noise in the dataset rather than reflecting the overall population. For this it is necessary to construct an integrated regression-classification model to approximate the true model for the entire population in the dataset. The proposed model integrates Multiple Linear Regression algorithm and One Rule (OneR) classification algorithm. Initially the prediction was treated with regression approach where the outputs were in numerical values. After that a classification model was used to interpret the regression outputs and then classified the outcomes into Profit and Loss class labels. The test results were compared to those obtained with standard classification algorithms which included OneR, Zero Rule (ZeroR), Decision Tree and REP Tree. The results showed that the regression-classification model were significantly more successful than the standard classification algorithms.

Keywords: Stock Price Prediction, Regression, Classification

Introduction

Stock price is an important vigilant factor in financial market since it changes with time in the financial domain of the market economy. When people invest, their investment values might rise or fall because of market conditions. Risk lies at the core of security investments. Low levels of risk are associated with low potential returns and vice versa (Wang and Wang, 2016a). The balance between the desire for the lowest possible risk and the highest possible return is the tradeoff of risk and return (Liu and Ren, 2015). Due to the complexity of the stock market, traditional statistical models for time series forecasting is restricted (Cai *et al.*, 2012). Fluctuation in time series leads to a lot of noises due to the nonlinear, non-stationary, selective and dynamic nature of the approach (Wang and Wang, 2016b). Investors make decisions based on available data by considering the data which have impacts on the financial instruments (Shynkevich *et al.*, 2015). The issues related to up and down price fluctuation in the stock market is defined as a binary classification problem (Sirohi *et al.*, 2014). In econometric analysis, forecasting of stock price has been one of the most active and important research topics. ARIMA model is one of the emerging fields of research in stock price prediction (Ariyo *et al.*, 2014). The defined model needs to overcome the difficulties in financial forecasting due to complexity in the stock

market. A model can maximize the benefits of many machine learning algorithms by incorporating different machine learning techniques to form a hybrid model (Liang *et al.*, 2015; Wang *et al.*, 2012; Wei, 2013). There are many ways to maximize returns and minimize risk so that the risk is worthwhile to the investors. The need of investors for various forecasting methods leads researchers to develop new predictive models (Ma, 2013). Effective prediction of stock price uses data mining and machine learning techniques to solve critical issues in stock value prediction.

To overcome difficulties in prediction of stock value, financial analysts and researchers have explored various data mining forecasting models. Some researchers developed time series model using financial theories whereas some developed time series based on characteristics of the collected data. The traditional approaches developed for predictions of stock price are moving average, exponential smoothing, statistical methods and linear regression method (Luo *et al.*, 2012). Markov model and ARMA model were developed for forecasting of stock prices (Shi *et al.*, 2012). The time-varying and nonlinear characteristics of the market enable existing traditional approaches to reveal internal pattern of the stock market (Xiao *et al.*, 2013). To overcome drawback in traditional approaches, Artificial Neural Network (ANN) and Artificial Intelligent (AI) techniques

were enhanced with the use of Radial Basis Function (RBF) (Shi and Liu, 2014) and Support Vector Machine (SVM) (Gong *et al.*, 2016) for stock price prediction and improved forecasting performance (Zhang *et al.*, 2012). Over the years, machine learning algorithms such as Neural Network (NN) and Genetic Algorithms were framed for financial assets prediction (Yang *et al.*, 2012). Researchers have exhibited their interest on selecting appropriate Kernel for various data sets which is also known as Multiple Kernel Learning (MKL) (Wang *et al.*, 2014). ANN is usually preferred as the stock price prediction tool compared with other methods (Mingyue *et al.*, 2016). Those stated approaches do not always work effectively for predictions since they are subjected to external impacts and uncertainties due to unknown factors. The existing approaches still face difficulties on forecasting movements in the stock market trends (Li *et al.*, 2015), hence it is challenging to develop a model for forecasting stock market values of firms.

Machine learning techniques are commonly used as predictive analytics in many financial applications. Regression and classification are both related to prediction, where regression predicts a continuous value and classification predicts a discrete value. The proposed model incorporates both Multiple Linear Regression and OneR algorithms. The Multiple Linear Regression approach is for modeling the linear relationship between the stock prices and the financial ratios which include debt to equity, asset turnover, cash flow and return on equity. Stock price trend is a dependent variable and financial ratios are the explanatory variables or independent variables. The Multiple Linear Regression approach produces the predicted stock prices in large range of number. The content of the numbers is then used as the content for the new dataset in the classification model. For the classification dataset, the independent variable's records consist of the predicted stock prices from the Multiple Linear Regression model, whereas the dependent variable's records consist of records from the original dataset's dependent variable. The learned OneR classification algorithm classifies the stock prices as Profit class or Loss class by performing a split of the values within the independent variable records relative to the values of the dependent variable records in the transformed dataset. The opposing sides of the split points determine the predicted values of either Profit class or Loss class. The combined approach of regression and classification produces the optimal results.

Methodology

The process flow of the proposed approach is illustrated as flow chart in Fig. 1. The model starts with the collection of the stock market data for this research and corresponding classification features are described. Dataset 1 is in classification form where the class values

are in nominal format. Initially the regression-classification model considers the prediction as a regression problem by using the regression algorithm. In order to apply the regression algorithm, the data type of the class variable must be converted into numeric format since the predicted stock values for regression are numerical numbers. This conversion results in dataset 2 which is in regression form. The records in dataset 2 are partitioned as training and testing data. The regression classifier learns from the training data to formulate the regression rules. Then the regression rules use the testing data for stock value prediction.

After the regression process, the model considers the prediction as a classification problem by using the OneR algorithm. The role of OneR is to classify the stock values into either Profit class or Loss class. In order to apply the OneR algorithm, the data type of the class variable must be converted back into nominal format since the predicted output for OneR classification is in nominal values. One of the nominal values denotes Profit class and the other nominal value denotes Loss class. This conversion results in dataset 3 which is in classification form. Dataset 3 contains only one independent variable with one dependent class variable. The independent variable contains the predicted stock values from the regression rules. The class variable contains the values derived from the class variable of dataset 1. Data are partitioned as training and testing data in dataset 3. The OneR classifier learns from the training data to formulate the OneR rules. Then the OneR rules use the testing data for stock value classification. The problem of overfitting still exists when OneR categorizes too many ranges of data as Profit class or Loss class. The value of the Minimum Bucket Size parameter for OneR is configured to produce only two ranges of data where one range is classified as Profit class and the other range is classified as Loss class.

Linear Regression Classifier

Linear regression technique is used to evaluate the uncertain data point and determines the best straight line passing through that point (Köeppen *et al.*, 2014). Regression technique offers a relationship between quantitative and explanatory variable for analysis method (Zhou *et al.*, 2016). Due to the simplicity of the process, regression technique is commonly used for analyzing larger amount of data (Ding *et al.*, 2015). The equation for Multiple Linear Regression, given n observations, is written as (Shin *et al.*, 2015):

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \epsilon \quad (1)$$

where, b_0, b_1, \dots, b_n , which are called the regression coefficients, are unknown constants to be determined from the data (Xie *et al.*, 2016).

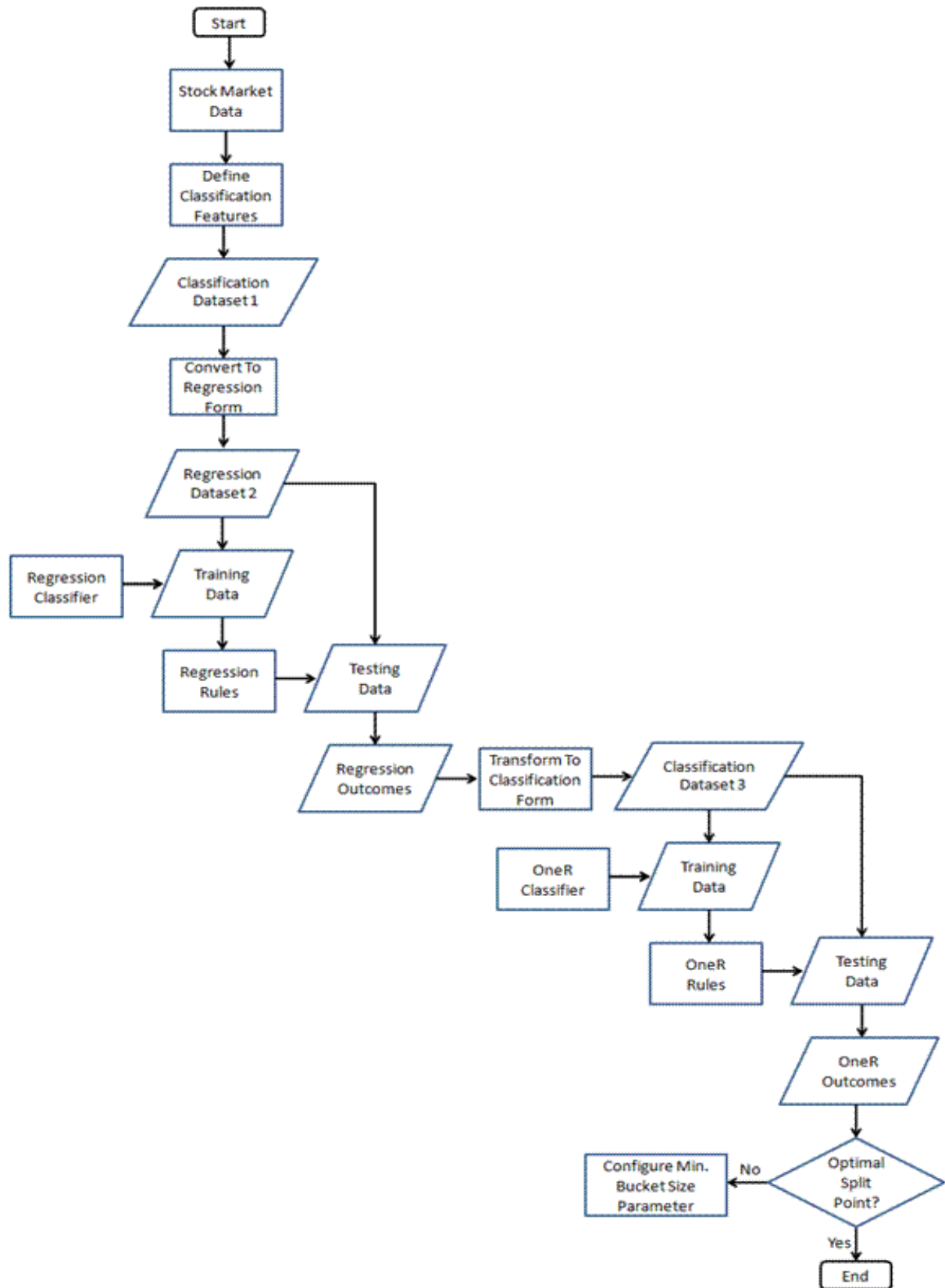


Fig. 1. Regression-classification model flowchart

In vectorial notation, the Multiple Linear Regression can be expressed as (Shin *et al.*, 2015):

$$Y = X\beta + \epsilon \quad (2)$$

In this equation, 'X' and 'Y' are two variables which are related to the parameters 'β' and 'ε', 'β' is the unknown parameter and 'ε' is the error term. A common term for any parameter estimate used in an equation for predicting Y from X is coefficient.

Linear regression focuses on the conditional probability distribution of Y given X. The structural model implies that for each value of X the population mean of Y can be calculated using the simple linear expression $X\beta + \epsilon$. The precise calculation cannot be made due to the fact that the two parameters are unknown factors. In practice, we make estimates of the parameters and substitute the estimates into the equation.

OneR Classifier

OneR which stands for One Rule is a classification algorithm used to generate a one-level decision tree. OneR is developed by Robert C. Holte. OneR produces a one-rule rule based on the value of a single attribute. The classification algorithm generates one rule for each attribute in the data and then selects the rule with the smallest error rate as its rule. To generate a rule for the obtained data, a frequency table is created for the obtained values based on the input values. The pseudo code of OneR algorithm (Buddhinath and Derry, 2006) is illustrated in Table 1.

OneR classifier rule is generated using a frequency table where it contains three columns. The first column of the table contains attributes related to the research and corresponding tally values will be included in column two. The third column contains frequency value assigned for individual attributes in the research. For individual input attribute in the initial step, the number of occurrence of each attribute needs to be evaluated as stated in Table 1. In the next step, frequency attribute occurrence and corresponding rule are assigned for each individual attribute in the research. In the final step, error rate for the individual attribute will be calculated and finally, attribute with minimum error rate is considered as a rule for the particular attributes.

Table 1. OneR pseudo code

For each attribute,
For each value of that attribute, make a rule as follows:
Count how often each class appears
Find the most frequent class
Make the rule assign that class to this attribute value.
Calculate the error rate of the rules
Choose the rules with the smallest error rate

Results

The proposed research is to predict stock price trend using integrated regression-classification techniques. The dataset input attributes considered in this research are made up of financial ratios and a class variable. All the attributes are numeric data type except for the price trend class variable which is of nominal data type. Dataset 1 is structured in classification form as shown in Table. 2.

The financial ratios in dataset 1 are based on a fundamental analysis where the first parameter considered in the prediction is debt equity. Debt equity ratio is equal to total liabilities divided by shareholders' equity. It measures how much financial leverage a company is using to finance its assets. Typically high exposure to debt is often associated with high level of risk. The second parameter is asset turnover which is a ration of total sales during a year to its average total assets for the same year. It measures the total sales for every dollar of assets a company owns since it indicates how efficiently a company is using its assets to promote sales. Cash flow ratio is the third attribute parameter which illustrates the net amount of cash and cash-equivalents moving into and out of business. It is a measure of a company's financial health in terms of liquidity which are crucial for a business' continued operation. The fourth parameter is return on equity ratio which is calculated by dividing net income of business for the period by the stockholders' equity during the same period. It indicates how much profit a company earned relative to the amount of shareholder equity. The price trend is the prediction based on the relationship between the dependable variable and the independent variables. Dataset 1 contains the original pre-processed data in classification form as illustrated in Fig. 1.

Dataset Conversion Using Unsupervised Attribute Filters

The regression model predicts the values of stock prices based on the financial ratios. The predicted output from the regression model contains decimal numbers which are in real data type. In order to execute the regression model, the dependent class variable named Price_Trend must be converted into numeric type from the nominal data type. The conversion is done through data preprocessing tool's unsupervised attribute filters. The steps adopted for conversion of nominal data into numeric data format is illustrated as follows:

Steps:

- Preprocess: Unsupervised attribute filters
- Convert attribute from nominal to binary (numeric)
- Apply to Price_Trend attribute
- Set Price_Trend attribute to contain no class value

Table 2. Dataset 1

Attribute	Data type
Debt_Equity	Numeric
Asset_Turnover	Numeric
Cash_Flow	Numeric
Return_On_Equity	Numeric
Price_Trend	Nominal {Profit, Loss}

Table 3. Dataset 2

Attribute	Data type
Debt_Equity	Numeric
Asset_Turnover	Numeric
Cash_Flow	Numeric
Return_On_Equity	Numeric
Price_Trend	Numeric

Table 4. Dataset 3

Attribute	Data type
Classification	Numeric
Price_Trend	Nominal {Profit, Loss}

Table 3 shows the converted data type from dataset 1 to dataset 2.

Dataset Transformation Using Unsupervised Attribute Filters

Using dataset 2, the regression algorithm predicts the stock prices in large range of value. In order to classify the predicted stock prices into Profit class or Loss class, the dataset is transformed into dataset 3 which has a single independent variable that contains the predicted stock prices. The independent variable is named 'Classification' in dataset 3. Dataset 3's dependent variable which is named 'Price_Trend' contains the original class values from dataset 1. The transformed dataset is in classification form that can be validated with dataset 1 which is also in classification form. Figure 1 illustrates the transformed dataset as dataset 3. Table 4 illustrates the structure of dataset 3.

OneR classifier is used as the classification model on dataset 3. The transformation into dataset 3 is done through preprocessing tool's supervised attribute filters. The steps are illustrated as follows:

Steps:

- Preprocess: Supervised attribute filters
- Add a classification attribute
- Choose classifier as Linear Regression
- Set output classification from Linear Regression
- Preprocess: Unsupervised attribute filters
- Convert attribute from numeric to nominal type
- Apply to Price_Trend attribute
- Set Price_Trend attribute to be a class attribute
- Remove and substitute Debt_Equity, Asset_Turnover, Cash_Flow and Return_On_Equity attributes with Classification attribute

- Select classifier OneR to predict class values
- Evaluate OneR model for overfitting
- Change Minimum Bucket Size parameter's value to overcome the overfitting problem
- Continue changing Minimum Bucket Size parameter's value until a single split point is achieved
- Evaluate and confirm the performance of OneR model

Multiple Linear Regression Model

A Multiple Linear Regression rule is formulated by computing the linear relationship between the dependent variable and the independent variables. The Multiple Linear Regression model is expressed as follows:

$$\begin{aligned}
 \text{class} = & \\
 & 0.2526 * \text{Debt_Equity} + \\
 & 0.3593 * \text{Asset_Turnover} + \\
 & 0.3362 * \text{Cash_Flow} + \\
 & 0.2585 * \text{Return_On_Equity} + \\
 & 0.013
 \end{aligned} \tag{3}$$

The equation of '(3)' takes the Multiple Linear Regression form from '(1)' as shown below:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \epsilon$$

The 'class' denotes y . The value of b_0 is zero and the value of ϵ is 0.013. The four independent variables have positive correlation relationships with the class variable; their correlations range from +0.2526 to +0.3593. Among the independent variables, Debt_Equity has the lowest correlation and Asset_Turnover has the highest correlation with the class variable.

OneR Model

The OneR rule is described in Fig. 2. The rule was configured to contain only one split point by setting the Minimum Bucket Size parameter's value to 20.

The split point's value lies at 0.4482274968036301. The '0' denotes predicted label of Loss class for all the continuous values that fall below 0.4482274968036301. The '1' denotes predicted label of Profit class for all the continuous values that are equal or greater than 0.4482274968036301.

Regression Predictor Output

The regression classifier used for the proposed regression-classification model is evaluated for performance by comparing the regression algorithms. Table 5 shows the results obtained using only the regression type classifiers which include Multiple Linear Regression, SMO Regression, RBF Network, Simple Linear Regression and Isotonic Regression.

Table 5. Prediction results for regression classifiers

Regression classifier	Correlation coefficient	Mean absolute error	Root mean square error
Multiple Linear Regression	0.7842	0.2340	0.2980
SMO Regression	0.6209	0.1944	0.4310
RBF Network	0.5732	0.3204	0.3937
SimpleLinearRegression	0.5679	0.2902	0.3964
Isotonic Regression	0.5679	0.2902	0.3964

Table 6. Prediction results for classification classifiers

Classification classifier	Accuracy	Mean absolute error	Root mean square error
Multiple Linear Regression-OneR	85.0746	0.1493	0.3863
OneR	71.6418	0.2836	0.5325
ZeroR	64.1791	0.4619	0.4805
J48 Decision Tree	82.0896	0.2121	0.3796
REP Tree	76.1194	0.2764	0.4207

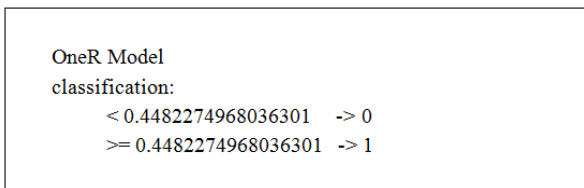


Fig. 2. OneR model

The Multiple Linear Regression classifier has the highest correlation coefficient of 0.7842, mean absolute error rate of 0.2340 and root mean square error rate of 0.2980. The results show that overall the Multiple Linear Regression classifier outperforms the other regression classifiers.

The proposed algorithm with classification outcome is evaluated for performance by comparing with other classification algorithms. Table 6 shows that the Multiple Linear Regression-OneR model has the highest accuracy rate of 85.0746, low mean absolute error rate of 0.1493 and root mean square error rate of 0.3863. The results demonstrate that the integrated model outperforms the other classification classifiers.

Discussion

The study shows that better result can be achieved by combining the strengths of different types of algorithms. The integrated approach taps into the benefits of Multiple Linear Regression technique and One Rule classification technique. In the regression model, the dependent variable contains the predicted stock prices and the outcomes depend on the values of the independent variables which are the financial ratios that include debt equity, asset turnover, cash flow and return on equity. The formula to calculate the relationship between the stock price and each financial ratio is called covariance calculation and corresponding equation are stated below:

$$Cov(x, y) = \frac{\sum (x_n - x_u)(y_n - y_u)}{N}$$

This calculation shows the direction of the relationship as well as its relative strength between the stock price and the financial ratio. If stock price variable increases and the other financial variable tend to also increase, the covariance would be positive. If stock price variable goes up and the other financial variable tends to go down, then the covariance would be negative.

The correlation calculation is used to standardize the covariance in order to better interpret and use it in forecasting (Nguyen, 2016). The equation of correlation coefficient can be expressed as:

$$Correlation = \rho_{xy} = \frac{Cov_{xy}}{s_x s_y}$$

The correlation calculation takes the covariance and divides it by the product of the standard deviation of the stock price variable and each financial ratio variable. The correlation has a value between -1 and +1. A correlation of +1 can be interpreted as both stock price variable and financial ratio variable move perfectly positively with each other and a -1 implies that they are perfectly negatively correlated. The regression model shows that the stock price variable and the financial ratio variables are all positively correlated with values in the range of 0.2526 to 0.3593 as shown in the equation of '(3)'.

The data of the predicted outcome of the regression model are made up of continuous values. The process of the regression model is followed by a classification model. The task of the classification model is to classify the continuous values into given set of categories which are Profit or Loss. The classification model uses OneR classifier to perform a split on the continuous values to determine the Profit class and Loss class categories. As there is only one transformed independent attribute in

dataset 3, OneR's 'one rule' fits the classification model on dataset 3. There were too many split points when the values of Minimum Bucket Size's parameter were set below 20. There is only one split point at 0.4482274968036301 when the Minimum Bucket Size parameter's value is set to 20 and this produces the required result in a 2-class scenario on Profit or Loss. A higher value of bucket size has obtained rules that are not overfitting.

Table 5 presents the stage 1 results obtained using the regression classifiers which predict the class values based on the independent variables' values. The class values are the forecasted stock prices. The Multiple Linear Regression classifier obtained the highest correlation coefficient with a low mean absolute error and the lowest root mean square error.

The values of the predicted stock prices from the Multiple Linear Regression classifier were used in a new dataset in stage 2. The new dataset was transformed into classification form to suit the classification classifiers. Table 6 shows the results for classification classifiers which predict discrete variable in nominal values.

Conclusion

The Multiple Linear Regression-OneR algorithm provides better results with lower prediction errors, when compared to the other standard classification algorithms. The integrated algorithm that capitalizes on the strength of both regression and classification techniques leads to significantly better performance.

The end result of the proposed model is based on binary classification where instances are classified into one of the two Profit and Loss classes. Further research can be done to transform the model into multiclass classification. The multiclass classification extends the two-class scenario to more classes. A potential model with five classes could include more class labels such as 'Strong_Profit', 'Weak_Profit', 'Neutral', 'Weak_Loss' and 'Strong_Loss'. More comprehensive information can be obtained by using a multiclass classification.

Acknowledgement

This work is supported by the Malaysian Institute of Information Technology, Universiti Kuala Lumpur. This research is sponsored by a research grant DIP-2016-018 from Universiti Kebangsaan Malaysia.

Author's Contributions

Lock Siew Han: Conception, design, data collection, analysis of data and writing the manuscript.

Md Jan Nordin: Analysis and interpretation of data, reviewing the manuscript, research supervision and monitoring of the research.

Ethics

The authors confirm that this manuscript has not been published elsewhere and that no ethical issues are involved.

References

- Ariyo, A.A., A.O. Adewumi and C.K. Ayo, 2014. Stock price prediction using the ARIMA model. Proceedings of the UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Mar. 26-28, IEEE Xplore Press, pp: 106-112. DOI: 10.1109/UKSim.2014.67
- Buddhinath, G. and D. Derry, 2006. A simple enhancement to one rule classification.
- Cai, X., S. Hu and X. Lin, 2012. Feature extraction using restricted Boltzmann machine for stock price prediction. Proceedings of the IEEE International Conference on Computer Science and Automation Engineering, IEEE Xplore Press, pp: 80-83. DOI: 10.1109/CSAE.2012.6272913
- Ding, Y., L. Cheng, W. Pedrycz and K. Hao, 2015. Global nonlinear kernel prediction for large data set with a particle swarm-optimized interval support vector regression. IEEE Trans. Neural Netw. Learn. Syst., 26: 2521-2534. DOI: 10.1109/TNNLS.2015.2426182
- Gong, X., Y.W. Si, S. Fong and R.P. Biuk-Aghai, 2016. Financial time series pattern matching with extended UCR suite and support vector machine. Expert Syst. Applic., 55: 284-296. DOI: 10.1016/j.eswa.2016.02.017
- Köeppen, M., K. Yoshida and K. Ohnishi, 2014. Evolving fair linear regression for the representation of human-drawn regression lines. Proceedings of the International Conference on Intelligent Networking and Collaborative Systems, Sept. 10-12, IEEE Xplore Press, pp: 296-303. DOI: 10.1109/INCoS.2014.89
- Li, X., H. Xie, Y. Song, S. Zhu, Q. Li and F.L. Wang, 2015. Does summarization help stock prediction? A news impact analysis. IEEE Intell. Syst., 30: 26-34. DOI: 10.1109/MIS.2015.1
- Liang, X., X. Liang, W. Xu and X. Wang, 2015. A hybrid model for stock price based on wavelet transform and support vector machines. Proceedings of the 12th International Conference on Service Systems and Service Management, IEEE Xplore Press, pp: 1-7. DOI: 10.1109/ICSSSM.2015.7170268
- Luo, J., Y.W. Si and S. Fong, 2012. Trend following with float-encoding genetic algorithm. Proceedings of the 7th International Conference on Digital Information Management, Aug. 22-24, IEEE Xplore Press, pp: 173-176. DOI: 10.1109/ICDIM.2012.6360100

- Ma, Y., 2013. The research of stock predictive model based on the combination of CART and DBSCAN. Proceedings of the 9th International Conference on Computational Intelligence and Security, Dec. 14-15, IEEE Xplore Press, pp: 159-164. DOI: 10.1109/CIS.2013.40
- Mingyue, Q., L. Cheng and S. Yu, 2016. Application of the artificial neural network in predicting the direction of stock market index. Proceedings of the 10th International Conference on Complex, Intelligent and Software Intensive Systems, Jul. 6-8, IEEE Xplore Press, pp: 219-223. DOI: 10.1109/CISIS.2016.115
- Nguyen, J., 2016. Regression basics for business analysis. Investopedia.
- Liu, P. and Y. Ren, 2015. Bp neural network model for prediction of listing corporation stock price of Qinghai province. Proceedings of the International Conference on Logistics, Informatics and Service Sciences, Jul. 27-29, IEEE Xplore Press, pp: 1-6. DOI: 10.1109/LISS.2015.7369738
- Shi, H. and X. Liu, 2014. Application on stock price prediction of Elman neural networks based on principal component analysis method. Proceedings of the 11th International Computer Conference on Wavelet Active Media Technology and Information Processing, IEEE Xplore Press, pp: 411-414. DOI: 10.1109/ICCWAMTIP.2014.7073438
- Shi, S., W. Liu and M. Jin, 2012. Stock price forecasting using a hybrid ARMA and BP neural network and Markov model. Proceedings of the IEEE 14th International Conference on Communication Technology, Nov. 9-11, IEEE Xplore Press, pp: 981-985. DOI: 10.1109/ICCT.2012.6511341
- Shin, Y., C.B. Chae and S. Kim, 2015. Multiple regression models for a big data empowered SON framework. Proceedings of the 7th International Conference on Ubiquitous and Future Networks, Jul. 7-10, IEEE Xplore Press, pp: 982-984. DOI: 10.1109/ICUFN.2015.7182693
- Shynkevich, Y., T.M. McGinnity, S. Coleman and A. Belatreche, 2015. Stock price prediction based on stock-specific and sub-industry-specific news articles. Proceedings of the International Joint Conference on Neural Networks, Jul. 12-17, IEEE Xplore Press, pp: 1-8. DOI: 10.1109/IJCNN.2015.7280517
- Sirohi, A.K., P.K. Mahato and V. Attar, 2014. Multiple Kernel learning for stock price direction prediction. International Conference on Advances in Engineering and Technology Research, Aug. 1-2, IEEE Xplore Press, pp: 1-4. DOI: 10.1109/ICAETR.2014.7012901
- Wang, F., Z. Zhao, X. Li, F. Yu and H. Zhang, 2014. Stock volatility prediction using multi-kernel learning based extreme learning machine. Proceedings of the International Joint Conference on Neural Networks, IEEE Xplore Press, pp: 3078-3085. DOI: 10.1109/IJCNN.2014.6889651
- Wang, J.J., J.Z. Wang, Z.G. Zhang and S.P. Guo, 2012. Stock index forecasting based on a hybrid model. Omega, 40: 758-766. DOI: 10.1016/j.omega.2011.07.008
- Wang, Y. and Y. Wang, 2016a. Using social media mining technology to assist in price prediction of stock market. Proceedings of the IEEE International Conference on Big Data Analysis, Mar. 12-14, IEEE Xplore Press, pp: 1-4. DOI: 10.1109/ICBDA.2016.7509794
- Wang, Y. and Y. Wang, 2016b. Using social media mining technology to assist in price prediction of stock market. Proceedings of the IEEE International Conference on Big Data Analysis, Mar. 12-14, IEEE Xplore Press, pp: 1-4. DOI: 10.1109/ICBDA.2016.7509794
- Wei, L.Y., 2013. A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX. Economic Modell., 33: 893-899. DOI: 10.1016/j.econmod.2013.06.009
- Xiao, Y., W.G. Che, Z. Wang and C.C. Yang, 2013. The research of morphological characteristics in time series of stock prices based on CBR. 3rd International Conference on Intelligent System Design and Engineering Applications, IEEE Xplore Press, pp: 1518-1521. DOI: 10.1109/ISDEA.2012.364
- Xie, G., Z. Wang, X. Hei, S. Takahashi and H. Nakamura, 2016. Data-based axle temperature prediction of high speed train by multiple regression analysis. Proceedings of the 12th International Conference on Computational Intelligence and Security, IEEE Xplore Press, pp: 349-353. DOI: 10.1109/CIS.2016.0085
- Yang, K., M. Wu and J. Lin, 2012. The application of fuzzy neural networks in stock price forecasting based On Genetic Algorithm discovering fuzzy rules. Proceedings of the 8th International Conference on Natural Computation, May 29-31, IEEE Xplore Press, pp: 470-474. DOI: 10.1109/ICNC.2012.6234684
- Zhang, Y., R. Shan, H. Wang and F. Jin, 2012. A new wavelet-neural network-ARIMA shares index combination forecast model. Proceedings of the International Conference on Automatic Control and Artificial Intelligence, Mar. 3-5, IEEE Xplore Press, pp: 199-201. DOI: 10.1049/cp.2012.0953
- Zhou, F., M. Ghalwash and Z. Obradovic, 2016. A fast structured regression for large networks. IEEE International Conference on Big Data, Dec. 5-8, IEEE Xplore Press, pp: 106-115. DOI: 10.1109/BigData.2016.7840594