Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia
Vol. 9, 2022, page 259-269

Forecasting the Bank Loan Approval for Residential Property Purchased using Time Series Model

Muhammad Shafie Yusop, Nur Arina Bazilah Kamisan
Department of Mathematical Sciences, Faculty of Science
Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia.

Abstract
Forecasting time series data is a common occurrence that needs to be addressed by researchers in order to help our bank industries. This type of study are very limited in Malaysia. Plus, the banking industry data are mostly complex which is quite difficult to make the prediction for decision making. Therefore, this research's output may help to enhance the prediction performance for bank industry in Malaysia. In this research, the data used is the Loan Approved for Residential Property that are purchased from January 2006 and up to August 2020. This data was obtained from Bank Negara Malaysia (BNM) website under the Monthly Highlights & Statistics publications. The objectives of this research are to examine the pattern of the forecasted data using the best model chosen and to produce the most suitable forecasting model in predicting the bank loan approval data of Bank Negara Malaysia. Seasonal ARIMA, Holt-Winters Exponential Smoothing and Multiple Regression are utilized to make future prediction of the Loan Approval for Residential Property Purchased by comparing these three forecasting models. The accuracy and performance of the forecasted data will be measured using Mean Absolute Error (MAE), Root-Mean-Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). From this study, Multiple Regression were identified as the most suitable forecasting model for this bank loan approval data which perform with the least error value for MAE, RMSE and MAPE. Followed by Holt-Winters Exponential Smoothing and SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ with just small difference of error estimation.
Keywords: forecasting; banking; loans; SARIMA; Holt-Winters; multiple regression; time series.

1. Introduction
Forecasting is a prediction of future based on the historical data as indicators to cope with the future uncertainty [7]. Forecasting has been used for a long time as a decision makers with the forecasted data. It can help businesses in avoiding any possible obstacles and meticulous planning can be made in maximizing their profit. Therefore, people make predictions all the time in order to achieve their target. Some are assuming based on their experience and some are based on the scientific calculation [5]. Unfortunately, there is also people that are only guessing, which are quite risky for someone who run businesses.

Loan approval was a significant problem to the banks since there are numerous applications received daily, which are difficult for staff to handle and where the risks of making mistakes are great [5]. Almost every bank in the world's primary business is loan distribution. However, [9] found that housing provident fund loans have lower interest rates than commercial bank loans, and provident fund payments are now considered standard employee benefits by most enterprises and government bodies. Therefore, providing an efficient and non-biased system that can lead to more sound and profitable loan decisions, resulting in significant cost savings for banks and businesses, as well as a more stable financial banking system is crucial for financial stability and economic recovery [4].

Therefore, the purpose of this study is to come up with a suitable time series model. In order to do that, Seasonal AutoRegressive Integrated Moving Average (SARIMA), Holt-Winters Exponential Smoothing and Multiple Regression are chosen to forecast the data from Bank Negara Malaysia. These models will be compared by using the error checking tools which is Mean Absolute Error (MAE), Root-Mean-Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Model with the minimum value of these error checking tools will be chosen as the best forecasting model for this series since it performs the most accurate values in forecasting. Besides, this study can predict the pattern of forecasted data so that the financial industry can plan their strategy in maximizing their profit.

2. Literature Review

The aim of this section is to give an overview of some important related works.

2.1. The Conventional Model for Finance

In finance and insurance, forecasting financial time series is a critical subject [3]. [2] perform the series of stationary analysis, the correlogram test shows that the data series from Beta Bank Inc. for monthly personal loan sales in 2004 to 2010 was not stationary. Then, unit root test was applied by taking the first and second difference. The result shows that the series was stationary at the second difference of root test. Since the stationary analysis series shows the unstable result, this study used Box-Jenkins Method. The AIC and SIC shows that the ARIMA model gives smaller value which means that the application of time series creates a very rational results for businesses.

In [3] study, they are using the internal rating based (IRB) approach is a well-known credit risk measurement technique. It enables banks to simulate the risk of default. For converting the original observations into a time-scale domain they are using two type of wavelet transform models, which are the discrete wavelet transform (DWT) and the continuous wavelet transform (CWT). In order to forecast the credit risk type of data, this study compared the four main wavelet transform function, which are the Haar Wavelet Transform (Haar), Daubechies wavelet transform (d4), Coiflets Wavelet Transform (C6), and Least Asymetric (La8).

The author, [14] has demonstrate the accuracy of multiple regression-based models as investigative tools. The authors concentrated particularly on a real-world case study of a private Tunisian bank. In fact, they created a multiple regression model that allows the bank system to forecast the overall turnover of an organization—i.e., a customer—within a given time frame, taking into account potential changes in financial, macroeconomic, and microeconomic data/variables. Extensive simulated experiments has proved the accuracy of our approximation.

In the banking business, the ARIMA model is also used to develop a regression model and estimate expected trends for bank credit to public and private sectors. [7] analyses three types of ARIMA models for both public and private sector bank loan forecasting in order to find the best model. ARIMA model (1,1,0) is adequate for predicting bank credit to the public sector, whereas ARIMA model (3,2,3) is acceptable for forecasting bank credit to the private sector, according to a comparison study of fitted models for bank credit to the public and private sectors. Using the Sharwarz Information Criteria (SIC), Durbin Watson Test (DW), Standard Error of Regression (SE), and Akaike Information Criteria with the lowest value (AIC).

2.2. The Advanced Model of Time Series Analysis

With the advancement in technology, the number of application for loan is increasing every day in banking sector. As a result, banks are beginning to examine artificial intelligence (AI) as a way for data analysis [5]. The author employed a machine learning strategy, which is a sort of AI for loan bank prediction, to reduce time wastes in checking every applicant on a priority basis to deliver a loan to deserving clients. The link between the dependent variable and the other independent variables was defined using logistic regression [5]. This study is able to produce a model that can come up with the status of loan applicant which are either safe or risky applicant in HTML, CSS or Django at the local server. This result was reached using a mix of random forest (RF) and parameter correlation. Heat map was used for correlation, which shows the magnitude of a phenomenon as colour in two dimensions. RF was used by combining the predictions of estimators to produce more accurate predictions, while heat map was used for correlation by combining the predictions of estimators to produce more accurate predictions.

[14] investigate the use of Neuro-Fuzzy Systems (NFS) to solve global business difficulties. They combined research on NFS application approaches using a mixture of Artificial Neural Network (ANN) ability for many business domains in this study. Finance, marketing, distribution, business planning, information systems, production, and operations are just a few examples. According to many studies, the implementation of NFS is more accurate than other AI and statistical methods in handling business challenges. However, the author prove that the combination of NFS approach and ANN generalization ability to come up with the adaptive neuro-fuzzy inference system (ANFIS) have been favoured by most researchers. This model shows the ability to perform successfully for various of business domain.

Apart from that, [15] was also proposed on using machine learning approach in forecasting the cash flow prediction. This research compared LSTM with the ARIMA model, Facebook's ^{TM} Prophet and multi-layered perceptron (MLP) with the aim of saving money. To come up with the most accurate model, this study propose a new performance measure, interest opportunity cost (IOC) along with MAE and MAE. As a result, this study shows that LSTM and MLP perform better in forecasting the cash flow for this type of data.

Similarly, the study on the secondary data from the Credit Union of the Minescho Credit Union, Tarkwa. [9] used a combination of Regression Analysis and the Box-Jenkins method of Time Series to come up with a good predictive model for the trend of loan default. In regression analysis, they used the simple linear regression since this only one independent variable is involved and hypothesis testing to test the statistical supposition. The result shows a positive relationship and upward trend for the monthly record of Loan Default at Minescho Credit Union. On the other hand, the Box-Jenkins Analysis come up with the suitable ARIMA model with the lowest AIC, RMSE and MAPE.

2.3. The Application of Time Series in Malaysia

The study of fuzzy time series has gotten a lot of attention lately because of its unique ability to deal with ambiguous and partial data. Several forecasting models have been developed in order to enhance the predicting accuracy. The new fuzzy time series model that was introduced by [17] which combine the concept of the Fibonacci sequence, the framework of Song and Chissom's model and the weighted method of Yu's model where the length of intervals has not been invertigated. Therefore, [18] have studied the model despite [19] advocated that length of intervals affects forecasting results by testing it into Kuala Lumpur Composite Index (KLCI). Experimental data sets were created from KLCI stock index data during a two-year weekly period. The results suggest that the [17] model's frequency density-based partitioning functioned well with the datasets tested. The importance of interval lengths in forecasting performance is reaffirmed by this outcome.

Furthermore, [11] study have developed a hybrid model that combines DES and the Additive Holt-Winters approach. In their study, the accuracy of three forecasting models were tested: Additive Holt-Winters, Seasonal Auto Regressive Integrated Moving Average (SARIMA), and a hybrid model. In comparison to SARIMA and the Additive Holt-Winters model, the hybrid methodology produces the lowest RMSE and MAE values, as well as the second lowest MAPE value. Therefore, hybrid model can be viewed as an effective forecasting model specifically for forecasting of financial for banking system.

In conclusion, pursuing study in financial forecasting are a proper decision especially in loan approval. Especially in Malaysia, there are limited source of reference. This study aims to propose a forecasting model based on time series analysis. Based on the literature above, Seasonal ARIMA model, exponential smoothing and multiple regression was the most suitable approach on loan approval in Malaysia. AIC, MAPE and RMSE also can be used as the term for goodness-of-fit to check the performance of these three models' prediction.

3. Methodology

3.1. Description of Data

In this study, the data that will be used was the Loan Approval for Residential Property that are purchased from January 2006 and up to August 2020. This data was obtained from Bank Negara Malaysia (BNM's) website under the Monthly Highlights & Statistics publications.

3.2. Seasonal ARIMA Model

Seasonal ARIMA models with non-negative parameters p, d, q, P, D, and Q are referred to as SARIMA (p, d, q)(P, D, Q). The parameters p, d, and q are for the seasonal portion, while the capital P, D, and Q are for the lag order, degree of regular differencing, and moving average order, respectively.

$$\phi_p(B)\Phi_p(B^S)(1-B)^d(1-B^S)^D y_t = \theta_q(B)\phi_Q(B^S)u_t \tag{1}$$

where,

B	: non-seasonal backward operators,
B^S	: seasonal backward operators,
ϕ_p	: non-seasonal AR component coefficients with order p,
Φ_p	: seasonal AR component coefficients with order P,
θ_q	: non-seasonal MA component coefficients with order q,
ϕ_Q	: seasonal MA component coefficients with order Q,
d	: non-seasonal differencing order,
D	: seasonal differencing order,
y_t	: the time series,
u_t	: the white noise residual

3.3. Holt-Winters Exponential Smoothing

The Holt Winter's Exponential Smoothing method is based on three smoothing elements, namely the stationary data element (mean), the trend element, and the seasonal element for each period, and provides three weightings in the prediction, namely α, β and γ. The magnitude of the coefficient of α, β and γ were determined by minimize the error value of the estimate.

$$F_{t+m} = (L_t + mb_t)S_{t+m-s} \tag{2}$$

where

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \qquad \text{(level series)} \tag{3}$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \qquad \text{(trend estimate)} \tag{4}$$

$$S_t = \gamma \frac{y_t}{L_t} + (1-\gamma)S_{t-s} \qquad \text{(seasonality factor)} \tag{5}$$

F_{t+1}	:	Forecast for the next period
F_t	:	Forecast for the current period
α	:	Weight constant for exponential smoothing
β	:	Weight constant for trend element
γ	:	Weight constant for seasonal elements
y_t	:	Observed value of series on period t
n	:	Amount of actual data

3.4. Multiple Regression

Multiple regression analysis is one of the most widely used for all statistical methods. It is oftenly used to determine the relationship of one dependent variable, y with more than one independent variables, X_n. The estimated multiple linear regression can be formulated as follow :

$$y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + + \beta_p X_p \tag{6}$$

The parameters of the model can be calculated using the least square estimates as :

$$\widehat{\beta} = (X^T X)^{-1}(X^T Y) \tag{7}$$

where

$$\hat{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

and β_n for n = 0, 1, 2, …, p are parameters.

3.5. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is a metric for comparing errors between paired observations that describe the same phenomenon by measuring the average magnitude of the error in a set of forecast. The mean absolute error formulated as the sum of absolute error divided by the sample size, n as follow :

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n} \tag{8}$$

where

e_i	:	Error between actual data and forecasted data
$\hat{y}_i$	:	Prediction data
y_i	:	Actual data
n	:	Amount of actual data

3.6. Root-Mean-Square Error (RMSE)

RMSE is a measure of accuracy that can be used to evaluate predicting errors of different models for a single dataset rather than across datasets because it is scale-dependent. The RMSE can be formulated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{9}$$

where

y_t	:	The actual value of time t
$\hat{y}_t$	:	Forecast value of time t
n	:	The total number of observation

3.7. Mean Absolute Percentage Error (MAPE)

MAPE is the most useful measure to compare the accuracy of forecast between differenct products since it measures relative performance. The accuracy is usually used as a ratio defined by the formula:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{10}$$

where

y_t	:	The actual value of time t
$\hat{y}_t$	:	Forecasted Value of time t
n	;	the total number of observation

4. Results and discussion

4.1. Data Stationarity

In this study, the data obtained is Loan Approved for Residential Property that are purchased from January 2006 and up to August 2020.Figure 1 display the trend analysis plot of the actual loan approval data from January 2006 to August 2020. Based on the graph, we can observe that there is a trend and seasonality which shows the data are not stationary.

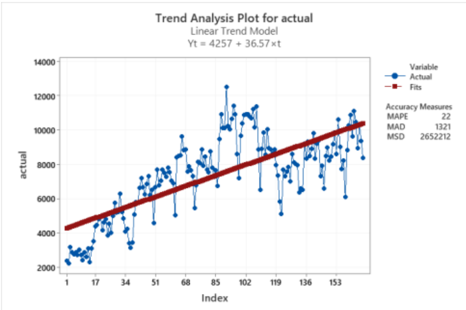

Figure 1 Trend Analysis Plot of the Actual Loan Approval Data

The autocorrelation function (ACF) graph is another approach in checking for the stationarity of time series data. We can see that the delays have a diminishing trend in Figure 4.2, indicating that the series are non-stationary. Plus, we can also see an obvious seasonal component from the ACF in Figure 4.2.

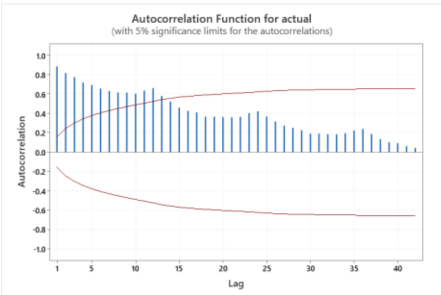

Figure 2 ACF Plot for Loan Approval Data

4.2. Seasonal ARIMA model

4.2.1. Model Identification

SARIMA model is suitable model for forecasting the non-stationary data. Figure 4.1 and Figure 4.2 shows that the series are non-stationary, differencing are needed before we can proceed to the next stage. After regular differencing, the ACF graph from Figure 4.1 shows that there exists a seasonality for each 12 lags sequence. Therefore, seasonal differencing are needed to be done.

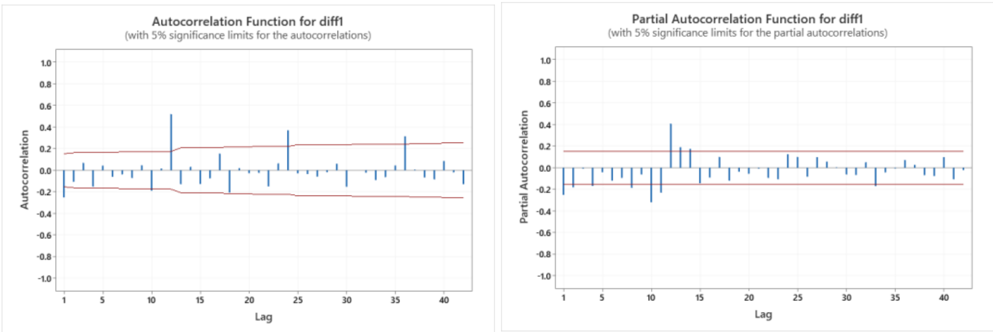

Figure 3 ACF and PACF plot of regural differencing

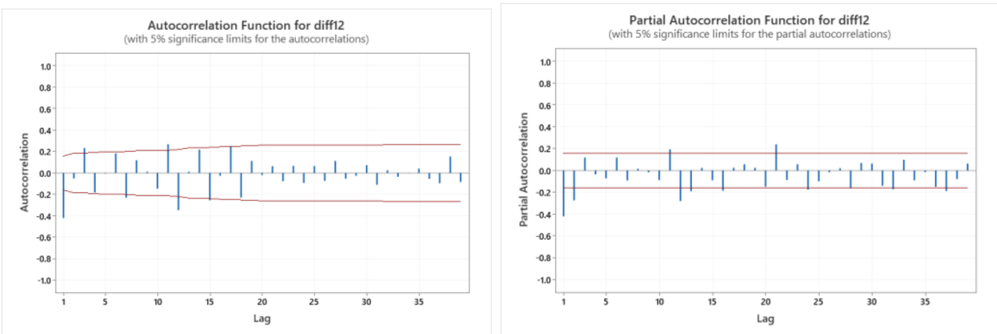

Figure 4 ACF and PACF plot of seasonality differencing

4.2.2. Parameter Estimation

From the regular differencing and seasonal differencing, we can fill in the d and D with 1 since we conduct only one time for each differencing. For P and p, we can obtain from the PACF plot while Q and q from ACF plot. From Figure 4.4, we can see that the PACF plot has two spikes on the first and second lag and there is spikes on lag 12 and 24 where we can fill our p and P with 2. For q and Q, we can fill with 1 since there is only one spike on the first lag and lag 12. From the ACF and PACF plot, we got a model SARIMA $(2, 1, 1)(2, 1, 1)^{12}$. There are also another possible SARIMA model can be used with difference parameter such that SARIMA $(0, 1, 1)(0, 1, 1)^{12}$ and SARIMA $(2, 1, 0)(2, 1, 0)^{12}$.

Lag	12	24	36	48
Chi-Square	31.04	57.16	71.27	82.80
DF	5	17	29	41
P-Value	0.000	0.000	0.000	0.000

Lag	12	24	36	48
Chi-Square	25.59	55.63	68.12	83.67
DF	9	21	33	45
P-Value	0.002	0.000	0.000	0.000

Lag	12	24	36	48
Chi-Square	13.64	36.81	55.10	73.06
DF	7	19	31	43
P-Value	0.058	0.008	0.005	0.003

Table 1 Ljung-Box Test for SARIMA $(2, 1, 1)(2, 1, 1)^{12}$, SARIMA $(0, 1, 1)(0, 1, 1)^{12}$ and SARIMA $(2, 1, 0)(2, 1, 0)^{12}$

However, based on Ljung-Box test shows that these three SARIMA model are not an adequate forecasting model since the *p*-value for SARIMA $(2, 1, 1)(2, 1, 1)^{12}$, SARIMA $(0, 1, 1)(0, 1, 1)^{12}$ and SARIMA $(2, 1, 0)(2, 1, 0)^{12}$ are all less than α=0.05. In which we can say that there is no series of autocorrelation of residuals if the p-values of Ljung-Box test statistics are greater than the significance level of α = 0.05.

Therefore, we will use the ACF and PACF plot from the first regular differencing to estimate the parameter for the SARIMA model. From Figure 3, we can see that the PACF plot has two spikes on the first and second lag and on lag 12 where we can fill our p with 2 and P with 1. In the ACF plot, q can equal 1 since there is only one spike on the first lag and Q will be equal to 0 since there is no spike in lag 12. From the ACF and PACF plot, we got a model SARIMA $(2, 1, 1)(1, 0, 0)^{12}$ where our D is equal to 0 since we are only consider the regular differencing. Then we can consider the model SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ as another possible model.

4.2.3. Model Selection

Both model shows the *p*-value that are greater than α=0.05 which are significant for Ljung-Box Test. Then T-test will be consider the best SARIMA model where if the p-value of T-test is lower than critical value at the significance level α = 0.05, we can conclude the residuals of the model is not autocorrelated.

Lag	12	24	36	48
Chi-Square	9.37	27.65	34.93	58.95
DF	7	19	31	43
P-Value	0.227	0.090	0.287	0.053

Lag	12	24	36	48
Chi-Square	13.27	31.27	39.00	61.66
DF	8	20	32	44
P-Value	0.103	0.052	0.184	0.040

Table 2 Ljung-Box Test for SARIMA $(2, 1, 1)(1, 0, 0)^{12}$ and SARIMA $(2, 1, 0)(1, 0, 0)^{12}$

From Table 3, we can see that the *p*-value of MA of SARIMA $(2, 1, 1)(1, 0, 0)^{12}$ is 0.194 which are greater than α = 0.05. Therefore we can say that SARIMA $(2, 1, 1)(1, 0, 0)^{12}$ are not an adequate model for this series. However, Table 4.10 shows that all the *p*-values of SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ are less than 0.05. Therefore, we can observed that SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ are the most suitable model for this series.

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.805	0.242	-3.33	0.001
AR 2	-0.378	0.102	-3.73	0.000
SAR 12	0.6891	0.0646	10.67	0.000
MA 1	-0.337	0.258	-1.30	0.194
Constant	26.1	85.3	0.31	0.760

Type	Coef	SE Coef	T-Value	P-Value
AR 1	-0.4897	0.0766	-6.39	0.000
AR 2	-0.2497	0.0759	-3.29	0.001
SAR 12	0.6859	0.0649	10.57	0.000
Constant	20.8	64.0	0.33	0.745

Table 3 T-test for SARIMA $(2, 1, 1)(1, 0, 0)^{12}$ and SARIMA $(2, 1, 0)(1, 0, 0)^{12}$

4.3 Holt-Winters Exponential Smoothing

Holt-Winters exponential smoothing is suitable to use for data that exhibit both trend and seasonality. It has three parameters model which are α, β, and γ. The value of these parameters is used to minimize the sum of square error for the smoothing constant, which $\alpha = 0.6, \beta = 0.1$ and $\gamma = 0.5$. These parameters were obtained automatically from Solver function in Microsoft Excel.

4.4. Multiple Regression

Multiple regression is used to determine the relationship between one dependent and more than one independent variables. It is a suitable model in forecasting for data series that exhibit both trend and seasonality. Based on Figure 1, it shows that the data have both the trend and seasonality. The seasonal dummy model can also be obtained which is:

actual = 6946 - 817 MONTH_1 - 1716 MONTH_2 + 336 MONTH_3
+ 798 MONTH_4 + 774 MONTH_5 + 871 MONTH_6

+ 1244 MONTH_7 + 1102 MONTH_8 + 473 MONTH_9
+ 1017 MONTH_10 + 743 MONTH_11

Based on the equation, Month_12 which is December are automatically removed by Minitab software since the term can not be estimated.

4.5. Comparison Between Forecasting Models

In this study, we already observed that SARIMA $(2, 1, 0)(1, 0, 0)^{12}$, Holt-Winters method and Multiple regression are all significant for this series. In order to come up with the best forecasting model for this loan approval data. Therefore, these three models will be compared by the error checking which include MAE, RMSE and MAPE. The model with the lowest value of the error will be considered as the best fitted model for the accuracy and performance.

Forecasted model	**MAE**	**RMSE**	**MAPE**
SARIMA $(2, 1, 0)(1, 0, 0)^{12}$	3168.70	4154.62	83.86
Holt-Winter ES	3152.17	4148.71	83.50
Multiple Regression	**2203.11**	**2696.85**	**55.51**

Table 4 Error Measurement for the out-sample

Based on Table 4, Multiple Regression model managed to attain the lowest value for all MAE, RMSE and MAPE compared to SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ and Holt-Winter exponential smoothing. Therefore, we can safely conclude that the Multiple Regression model is the best model in forecasting monthly loans issued by the Bank Negara Malaysia. Then, Holt-Winter exponential smoothing is the second best forecasting model followed by the SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ since it show the highest error for each three error checking tools. The accuracy and performance of all the three forecasting models can also be seen through the line plot as shown in Figure 5.

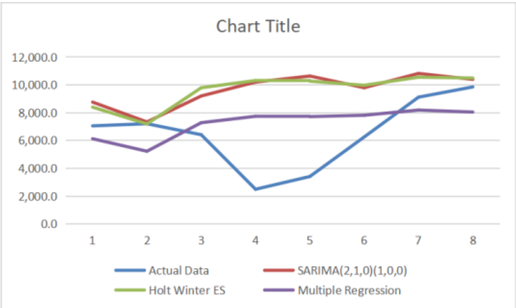

Figure 5 Comparison of loan data for actual and forecasted data

Conclusion

Our findings has succeed to observed that the best forecasting model which Multiple regression as it yields the lowest for all error checking tools that include MAE, RMSE and MAPE. Holt-Winters exponential smoothing was the second best forecasting model and followed by SARIMA $(2, 1, 0)(1, 0, 0)^{12}$. This is because SARIMA $(2, 1, 0)(1, 0, 0)^{12}$ shows the highest value for all three error checking tools even with only small difference of error with Holt-Winters exponential smoothing. To sum up, Multiple regression can be considered as the best forecasting model.

To improve this research, further study should use more out-sample data since all three forecasted model show more accuracy and better performance with more data. Besides that, this study can be improved by comparing between traditional forecasting model and advanced forecasting model would shows a better study since we can observe the best forecasting model for this loan data series. Therefore, this type of study could provide more forecasting model so that it can be used in forecasting with the best accuracy and better performance. Besides, future researcher should also consider on hybrid models since there are a lot of study that shows the hybrid model could perform better prediction on loan data series.

Acknowledgement

The researcher wish to express my sincere appreciation to my main thesis supervisor, Dr. Nur Arina Bazilah Kamisan, for encouragement, guidance, patience, critics and friendship. Hence, towards the parents and to all my siblings for their unwavering love, unflinching support and always be the rock that I have leaned on throughout the journey of finishing thid dissertation.

References

[1] Jeong, D. B. (2017). Forecasting for a Credit Loan from Households in South Korea. The International Journal of Industrial Distribution & Business, 8(4), 15-21. http://dx.doi.org/10.13106/ijidb.2017.vol8.no4.15

[2] Ali, İ. Ö. (2021). Personal Loan Sales Forecasting Through Time Series Analysis. Prizren Social Science Journa, 5(1).

[3] Jaber, J. J., Ismail, N., Ramli, S. N., Wadi, S. A., & Boughaci, D. (2020). Assessment of Credit Losses Based on ARIMA-Wavelet Method. Journal of Theorotical and Applied Information Technology, 98(9).

[4] Aliaj, T., Anagnostopoulos, A., & Piersanti, S. (2020). Firms Default Prediction with Machine Learning. Mining Data for Financial Applications. 10.1007/978-3-030-37720-5_4.

[5] Gupta, A., Pant, V., Kumar, S., & Bansal, P. K. (2020). Bank Loan Prediction System using Machine Learning. 9th International Conference on System Modelling & Advancement in Research Trend, 4-5.

[6] Liang, L., & Cai, X. (2020). Forecasting Peer-to-Peer Platform Default Rate with LSTM Neural Network. Electronic Commerce Research and Applications, 43.

[7] Nooreen, A., Asif, R., Nisar, S., & Qayyum, N. (2017). Model Building and Forecasting of Bank Credit to Public and Private Sector. Universal Journal of Accounting and Finance, 5(4), 73-77. 10.13189/ujaf.2017.050401

[8] Arutjothi, G., & Senthamarai, Dr. C. (2017). Prediction of Loan Status in Commercial Bank using Machine Learning Classifier. Proceedings of the International Conference on Intelligent Sustainable Systems.

[9] Otoo, H., Takyi Appiah, S., & Wiah, E. N. (2015). Regression and Time Series Analysis of Loan Default at Minescho Cooperative Credit Union, Tarkwa. International Journal of Scientific & Technology Research, 4(8).

[10] Ke, L., Li, C., Zhong, T., Cai, Z., Wen, J., Wang, R., Qian, F., & Tang, H. (2021). Loan Repayment Behaviour Prediction of Provident Fund Users Using a Stacking-Based Model. IEEE 6th International Conference on Cloud Computing and Big Data Analytics. 10.1109/ICCCBDA51879.2021.9442613

[11] Jafry, N. A., & Kamisan, N. A. B. (2021). Forecasting of Loans Approved by Purpose for Banking System in Malaysia. Proceedings of Science and Mathematics, 2, 193-202.

[12] Dosalwar, S., Kinkar, K., Sannat, R., & Pise, Dr. N. (2021). Analysis of Loan Avaibility using Machine Learning Techniques. International Journal of Advanced Research in Science, Communication and Technology, 9(1).

[13] Byanjankar, A. (2017). Predicting Credit Risk in Peer-to-Peer Lending with Survival Analysis. 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 1-8. 10.1109/SSCI.2017.8280927

[14] Hasni, M., & Layeb, S. (2017). Multiple Regression-Based Models for Accurate Credit Risk Management. International Journal of Economics & Strategic Management of Business Process, 10, 53-57.

[15] Weytjens, H., Lohmann, E., & Kleinsteuber, M. (2019). Cash Flow Prediction: MLP and LSTM Compared to ARIMA and Prophet. Electronic Commerce Research (2021), 21, 371-391. https://doi.org/10.1007/s10660-019-09362-7

[16] Zhang, T., & Du, Y. (2021). Research on User Credit Score Model Based on Fusion Neural Network. 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). 1391-1395. 10.1109/IAEAC50856.2021.9390672

[17] Chen, T. L., Cheng, C. H., & Teoh, H. J. (2007). Fuzzy time-series based on Fibonacci sequence for stock price forecasting, Physica A, 380, 377-390.

[18] Abdullah, L., & Ling, C. Y. (2011). A fuzzy time series model for Kuala Lumpur Composite Index forecasting, 2011 Fourth International Conference on Modeling, Simulation and Applied Optimization, 1-5, doi: 10.1109/ICMSAO.2011.5775515.

[19] Huarng, K. H., & Yu, H. K. (2005). A type-2 fuzzy time-series model for stock index forecasting, Physica A, 353, 445-462.