Information Management and Business Review (ISSN 2220-3796)
Vol. 16, No. 2, pp. 104-114, SI(2) 2024

Forecasting Short-Term FTSE Bursa Malaysia Using WEKA

*Nurhuda Nizar[1], Ahmad Danial Zainudin[2], Ali Albada[3], Chua Mei Shan[1],
[1]Faculty of Business and Management, UiTM Cawangan Selangor, Puncak Alam, Malaysia
[2]School of Accounting and Finance, Asia Pacific University of Technology & Innovation, Technology Park Malaysia, Kuala Lumpur, Malaysia
[3]Faculty of Business, Sohar University, Sohar, Oman
*hudanizar@uitm.edu.my

Abstract: This study investigates the use of machine learning methods, specifically utilizing the WEKA software, to predict stock prices of the FTSE Bursa Malaysia Kuala Lumpur Composite Index (KLCI). Two algorithms, Sequential Minimal Optimization Regression (SMOreg) and Multilayer perceptron (MLP), were employed for data analysis. Historical data from January 3, 2023, to December 29, 2023, was used to forecast open, high low, and close prices for ten days. Results from both algorithms were compared, with SMOreg proving to be more accurate than MLP for the dataset. However, it's important to note that further exploration of different forecasting algorithms may lead to even more precise results in the future. The findings of this analysis hold significant implications for investors, as they can use the insights gained to inform their investment strategies. By leveraging machine learning techniques like SMOreg within the WEKA framework, investors can potentially make more informed decisions regarding their stock market investments, leading to improved portfolio performance and risk management.

Keywords: *WEKA, Data Mining, stock market indices prediction, FTSE Bursa Malaysia KLCI.*

1. Introduction and Background

In today's ever-changing financial markets, the ability to forecast stock prices is crucial for investors, financial institutions, and market analysts (Gandhmal & Kumar, 2019; Kumar et al., 2022). The accuracy of these predictions is essential for achieving financial gains and minimizing losses (Emami, 2018; Farias Nazário et al., 2017). However, predicting stock market behavior is a difficult task due to its volatile and dynamic nature (Rouf et al., 2021). The stock market is affected by many interconnected factors, such as economic conditions, company performance, government policies and regulations, interest rates, and market sentiment. This complexity makes it challenging to forecast stock price movements, especially in the field of time series research (Dong et al., 2021). As a result, incorporating Machine Learning (ML) techniques has emerged as a promising approach to address these challenges in predicting stock market behavior (Zou et al., 2023). ML techniques, such as the use of WEKA (Waikato Environment for Knowledge Analysis), demonstrate the application of cutting-edge technologies in finance for stock market forecasting. These techniques have the potential to capture complex patterns in financial data and improve forecasting accuracy compared to traditional methods.

In the past, investors primarily relied on traditional tools to predict stock prices. These methods included fundamentals analysis (Al-Radaideh et al., 2013; Khadjeh Nassirtoussi et al., 2014), technical analysis (Farias Nazário et al., 2017; Kumbure et al., 2022; Navarro et al., 2023), market sentiment (Peng et al., 2023) and statistical techniques (Shah et al., 2022). However, with the widespread adoption of machine learning, investors now have an additional tool to enhance their predictions and make more informed decisions. Machine learning allows for the analysis of large amounts of data, identification of complex patterns, and adaptation to changing market conditions (Mintarya et al., 2022). Although traditional methods are still commonly used, the integration of machine learning has provided investors with valuable capabilities.

The popularity of machine learning has increased because it can identify non-linear trends and adapt to market changes that may be missed by traditional statistical methods. WEKA, a leading open-source project written in Java, is one of the most comprehensive tools for data mining. It includes state-of-the-art machine-learning algorithms for data pre-processing, classification, regression, clustering, association rules, and visualization (Frank et al., 2017). With this extensive range of tools, researchers and practitioners can explore different modeling techniques and choose the most suitable approach for their specific forecasting needs. WEKA's user-friendly interface makes it accessible to users with varying levels of expertise in machine learning including researchers and practitioners (Rehman & Soomro, 2019). This accessibility reduces the barriers to utilizing advanced modeling techniques, enabling more individuals to leverage the power of data-driven insights for stock price prediction.

This paper examines predicting future stock market performance, particularly stock prices, by analyzing historical data of open, high, low, and close prices using the WEKA software. The process involves collecting and preprocessing relevant data. Two algorithms, SMOreg and MLP, are employed for stock price prediction due to their ability to effectively manage the intricacies of financial data and identify non-linear patterns in stock market activity. Our contributions are twofold. Firstly, this paper is the first to present the accuracy of the WEKA approach in forecasting Malaysian equities. Extant literature such as Abd Samad, Mutalib & Abdul-Rahman (2019) reports the applicability of using machine learning algorithms for stock price prediction however the study is limited to CIMB, Sime Darby, Axiata, Maybank, and Petronas stocks using the Random Forest algorithm. We, on the other hand, analyze the Malaysian equity index which is the bellwether for the country's economic sentiment. Secondly, unlike the conventional econometric methods used by Choudhry, Hasan & Zhang (2019), Dong, Guo, Reichgelt & Hu (2020), and Rubio & Alba (2022), we examine the data as-is, which means there is no data medication procedures to observe and stringent statistical requirements. Henceforth, with the new effort on using machine learning for stock market forecasting, we intend to go beyond the existing literature to enhance the body of knowledge and provide another significant contribution to academic and finance professionals. Leveraging machine learning techniques within the WEKA framework can enable investors to navigate the complexities of the stock market with greater confidence, leading to improved portfolio performance and risk management.

The remainder of this paper is organized as follows: Section 2 discusses the literature review, Section 3 explains the data description and methodology, Section 4 presents the results of empirical tests, and Section 5 provides the conclusion.

2. Literature Review

Stock price prediction plays a crucial role in financial markets as it enables investors, traders, and financial institutions to make informed decisions about buying, selling, or holding stocks. Accurate predictions help stakeholders mitigate risks, optimize portfolio returns, and capitalize on investment opportunities. In this context, machine learning techniques have emerged as powerful tools for enhancing forecasting accuracy and have gained significant attention. Among the various tools available for this purpose, WEKA (Waikato Environment for Knowledge Learning) stands out as a versatile and widely used open-source software developed by the University of Waikato, New Zealand. Equipped with a plethora of machine learning algorithms, WEKA offers comprehensive tools for data pre-processing, classification, clustering, regression, association, and visualization (Frank et al., 2017).

Kumar and Ravi (2016) conducted a thorough review of text-mining applications in finance, shedding light on how textual data analysis through text-mining methods contributes to predicting trends and outcomes in the financial domain. Their study underscored the significance of text mining in forecasting financial events. In a similar vein, Kulkarni and More (2016) applied machine learning techniques, specifically utilizing WEKA, to predict stock prices. Their research explored different stock market forecasting processes within WEKA's forecasting plugin and conducted experiments on various stocks to analyze the tool's predictive capabilities. Further, Žmuk and Jošic (2020) aimed to forecast stock market indices using machine learning algorithms, including those available in the WEKA tool. Their study evaluated the efficiency of machine learning algorithms for predicting stock market indices, emphasizing WEKA's role in tasks such as data pre-processing, classification, regression, clustering, association rules, visualization, and forecasting. Rehman and Soomro (2019) leveraged the WEKA tool to analyze OGDCL stock prices and predict open, high, low, and close prices for ten days using SMOreg and MLP algorithms.

Machine learning methods, including those implemented in WEKA, have shown superior performance compared to traditional forecasting methods like time series analysis, econometric models, and fundamental analysis. Studies have demonstrated that machine learning algorithms outperform traditional methods for medium to long-term predictions (Iaousse et al., 2023). Additionally, incorporating machine learning approaches, significantly outperforms conventional methods, showcasing the potential of machine learning in handling complex datasets with consistent and reliable performance (Patil et al., 2024). Furthermore, research utilizing WEKA for share price prediction found that Sequential minimal optimization provided more accurate results compared to other methods, highlighting the effectiveness of data mining techniques in financial forecasting Devi, et al. (2023).

Overall, these studies highlight the diverse applications of WEKA in predicting stock market trends and prices. From text mining applications to forecasting individual stock prices and overall market performance, WEKA emerges as a valuable tool in the arsenal of financial analysts and researchers.

3. Methodology

Data Description

In our study, we collected data from Refinitiv over a 1-year sample period, spanning from January 3, 2023, and ending on December 29, 2023. We excluded weekends and public holidays from our data collection. The amount of historical data required depends on the prediction horizon. For short-term predictions such as the next 10 days, 1-2 years of daily price data can be sufficient. Previous studies have used different lengths of historical data for stock prediction using WEKA. For instance, Rehman and Soomro (2019) utilized one year and three months of daily data to forecast OGDCL stock prices, while Jošić, (2020) used historical daily close prices from 1, 5, and 10-year periods to predict major stock indices 5, 10, 15, and 20 days ahead.

The daily data collected was used to evaluate the effectiveness of Weka 3.8.6 in forecasting the FTSE Bursa Malaysia KLCI index. Our objective was to assess Weka's performance as a forecasting tool following the methodology of the previous study by Rehman and Soomro (2019). We used the Weka to predict the open, high, low, and close prices for the next ten days, comparing them to the actual prices. Two forecasting algorithms, namely SMOreg and MLP, were utilized. Data preprocessing was conducted to address missing values, and outliers, and normalize variables for consistent input into the machine learning models.

Multilayer Perceptron (MLP)

Multilayer Perceptron (MLP) in a Weka is a neural network classifier known for its effectiveness in solving classification and prediction tasks (Md Salim Chowdhury et al., 2024). The process begins with the input layer, where input variables from the dataset are received. The information then passes through hidden layers, updating weights and utilizing activation functions to interpret the data's nonlinearity. The output layer ultimately produces the final result, which may be a class label or a prediction. This tool is widely utilized in different applications and can be trained using the backpropagation algorithm. MLPs have the ability to learn complex, nonlinear relationships between inputs and outputs. They have been successfully applied to diverse problems such as image classification, natural language processing, and financial forecasting.

Sequential Minimal Optimization (SMOreg)

SMOreg, also known as Sequential Minimal Optimization, is a technique utilized to predict numerical values using Support Vector Machine (SVM) for regression. It addresses the intricate mathematical problem of quadratic programming optimization, which is crucial in SVM training. SMOreg simplifies this process by dividing the larger problem into smaller segments and solving them individually.

Both SMOreg and MLP algorithms are utilized for regression tasks and can achieve high accuracy in predicting results. The primary distinction between the two is that SMOreg utilizes the Sequential Minimal Optimization algorithm to solve the quadratic programming optimization problem associated with SVM for regression. In contrast, MLP typically relies on backpropagation to adjust connections between points based on prediction errors (Rashedi et al., 2024).

Mean Absolute Error (MAE)

When evaluating the accuracy of a prediction model, the Mean Absolute Error (MAE) metric is commonly used, especially in regression analysis. It measures the average absolute difference between the actual values and the predicted values of a dataset. Lower MAE values signify a higher level of predictive performance. Previous studies such as Nirob and Hasan (2023) and Villavicencio et al. (2021) have employed the MAE to assess predictive accuracy.

The formula for Mean Absolute Error is:

$$MAE = \frac{1}{n}\sum_{i}^{n} = 1\ |Y_i - \widehat{Y_i}|$$

Where:
In the dataset, n represents the total number of observations, yi refers to the observed value for the i^{th} observations while y$\hat{}i$ represents the predicted value for the i^{th} observation.

Root Mean Square Error (RMSE)
The error in a regression model can be measured using RMSE, which involves squaring the difference between the predicted and actual values and taking the average. In WEKA software, RMSE is used to evaluate the accuracy of a regression model in predicting continuous numerical outcomes. A lower RMSE value indicates better prediction accuracy. The calculation of RMSE follows this formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n} = 1\ |Y_i - \widehat{Y_i}|}$$

Where n represents the number of observations in the dataset, yi is the observed value for the i^{th} observations and $y\hat{}i$ is the predicted value for the i^{th} observations.

Mean Squared Error (MSE)
MSE is a widely used measure for assessing the precision of regression models. It calculates the average of the squared differences between the predicted values and the actual values in a dataset. A lower MSE signifies superior model performance, with a value of 0 representing a perfect fit (perfect predictions). The formula for calculating MSE is as follows:

$$MSE = \frac{1}{n}\sum i = 1^n\ (Yi - \widehat{yi})^2$$

Where n is the number of observations in the dataset yi is the actual (observed) value for the i^{th} observations and $y\hat{}i$ represents the predicted value for the i^{th} observations.

Mean Absolute Percentage Error (MAPE)
MAPE is a metric used to evaluate the accuracy of a forecasting or predictive model, especially in the context of time series data. MAPE calculates the average percentage difference between the predicted values and the actual values in a dataset. It is particularly useful for assessing the relative accuracy of a model across different scales and periods. A lower MAPE value indicates a more accurate predictive model. A MAPE of 0% would indicate a perfect fit where the predicted values exactly match the actual values.
The Mean Absolute Percentage Error (MAPE) formula is as follows:

$$MAPE = \frac{100\%}{n}\sum_{i}^{n} = 1\ \left\lfloor\frac{yi - \widehat{yi}}{yi}\right\rfloor$$

Where:
n is the number of observations in the dataset, yi is the actual (observed) value for the i^{th} observations and $y\hat{}$ i represents the predicted value for the i^{th} observations.

The distinction among these three methods lies in how they measure the difference. MAE calculates the average of the absolute differences between predicted and actual values. MSE calculates the average of the squared differences between predicted and actual values. As for RMSE, it's the square root of MSE.

4. Findings

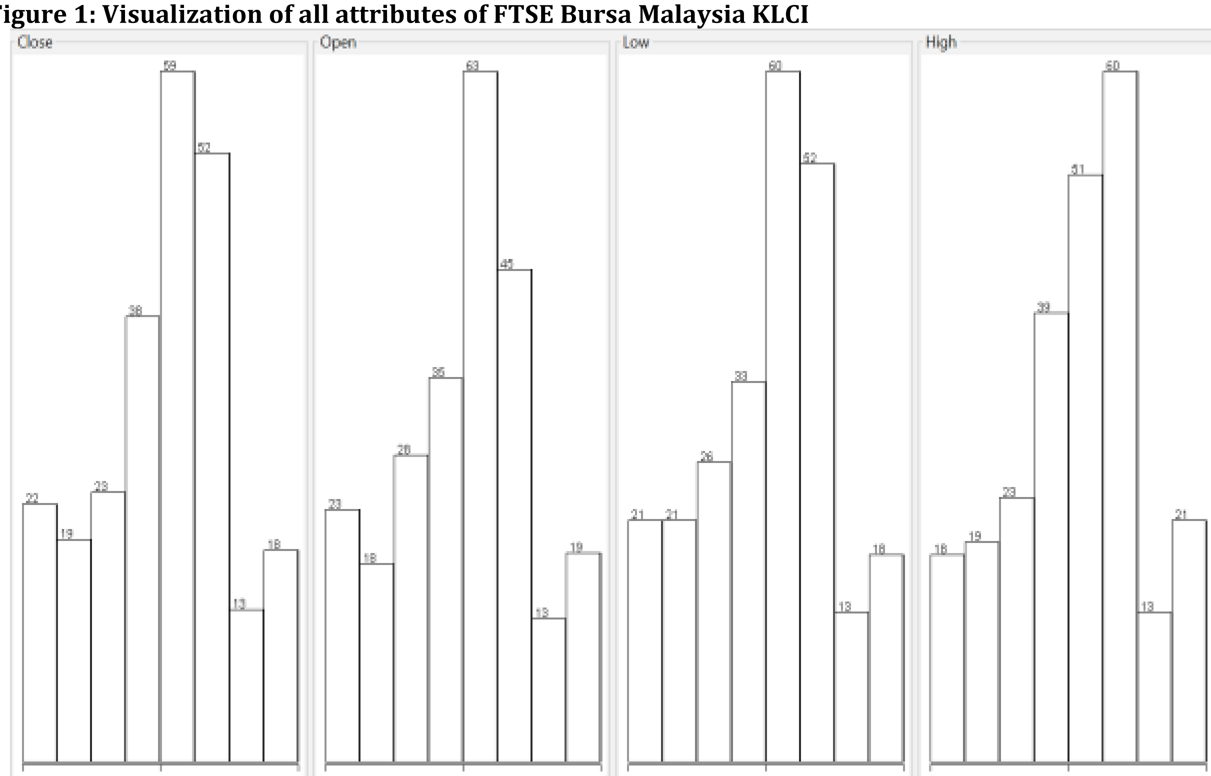
Descriptive Analyses: In this analysis, we utilized the WEKA tool to explore key descriptive statistics, aiming to gain insights into the historical trends and patterns that can help us forecast the stock market. The dataset used includes historical stock market data covering the period from January 3, 2023, to December 29, 2023, including daily open, high, low, and closing prices. The data spans a specified period, allowing us to forecast the 10-day future prices.

Table 1: Minimum, Maximum, Mean & Standard Deviation of Attributes

	Minimum	Maximum	Mean	Standard Deviation
Open	1374.64	1500.33	1438.891	29.978
High	1376.62	1501.86	1442.673	29.454
Low	1369.41	1495.8	1434.309	30.023
Close	1374.64	1500.33	1438.891	29.978

The descriptive statistics, presented in Table 1, provide a comprehensive overview of the data distribution. Notably, the minimum values for the open, high, low, and closing prices are recorded at 1369.41, while the maximum price peaks at 1501.86. This range between the maximum and minimum values reflects the extent of variability within the dataset, offering valuable insights into its distribution. The mean values for both the open and closing prices stand at 1438.891, representing central tendencies within the dataset. Additionally, examining the standard deviation reveals the dispersion of data points around the mean. In this context, the highest standard deviation of 30.023, observed in the low prices, underscores the extent of variability and potential fluctuations within this particular aspect of the dataset over the given period. Figure 1 shows the visualization for each attribute to visualize their distributions.

Figure 1: Visualization of all attributes of FTSE Bursa Malaysia KLCI



Results: In this study, we utilized Weka 3.8.6, which incorporates time series forecasting features. Four attributes, namely the open, high, low, and close prices of the KLSE, were included in the sample dataset that was put into the Weka experimental environment in CSV format. For the dataset, two algorithms SMOreg and MLP were applied for forecasting stock market predictions in Weka. Out of many other regression metrics, we choose MAE, MAPE, RMSE, and MSE as the basis of evaluation for the two forecasting models as shown in Table 2 below. These metrics are commonly used to measure the model's performance on stock price prediction (Jiang, 2021).

As depicted in Table 2, we compared the closing price of both algorithms and discovered that the values for metrics MAE, MAPE, RMSE, and MSE are lower for SMOreg over the next ten days. This trend persisted across the open, low, and high prices as well. This aligns with the findings of previous studies done by Ahmed and Hussain (2022); Žmuk and Jošic (2020); and Rehman and Soomro, (2019). Lower metric values indicate that the SMOreg algorithm may serve as a more accurate predictor in this context, as it produces forecasts that are closer to actual close prices.

$$MAE, MAPE, RMSE, MSE_{SMOreg} < MAE, MAPE, RMSE, MSE_{MLP}$$

To summarize, our results suggest that the SMOreg algorithm exhibits superior predictive capabilities compared to MLP, offering valuable insights for stock market forecasting applications.

Connection to the previous works

Our findings are in accordance with Abd Samad et al. (2019) and Ismail, Noorani, Ismail, Razak & Alias (2020) in the sense that incorporating the hybrid algorithm for financial forecasting tends to yield better results. Logically, the financial market dynamics consist of outliers hence we need the combination of machine learning algorithms to capture the actual market turmoil. As the financial market keeps evolving, there is no one-method-fits-all. Furthermore, the emerging stock markets have proven more efficient recently due to the increased sensitivity to global sentiment (Patra & Hiremath, 2022). Our paper presents a way of forecasting the Malaysian stock index with better computational power as a way to highlight the trajectory of the national financial sentiment.

5. Conclusion

In summary, the study aimed to use machine learning techniques in the WEKA framework to forecast stock prices. The SMOreg and MLP algorithms were employed to analyze historical data and make predictions. While SMOreg showed better accuracy in the dataset, further research is needed to explore forecasting algorithms. Diversifying methodologies is crucial for improving the reliability of stock price predictions in different market conditions.

These findings have practical implications for investors looking to enhance their investment strategies. By incorporating machine learning-based forecasts, investors can make better decisions on portfolio allocation, risk management, and timing of investments. The potential of SMOreg as a forecasting tool suggests its use in real-world investment practices, leading to improved performance and profitability.

In conclusion, the study provides valuable insights into the effectiveness of machine learning in stock price predictions, emphasizing the importance of ongoing exploration and innovation in this field. Advancing our understanding of forecasting algorithms can empower investors to navigate financial markets confidently and accurately.

Table 2: MAE, MAPE, RMSE & MSE for Open, High, Low & Close attributes obtained from SMOreg

		SMOreg															
		Close				Open				Low				High			
Target	N	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
I step ahead	237	1.547	0.108	2 .300	5.294	4.059	0.283	5.456	29.777	2.690	0.188	3.672	13.488	2.402	0.167	3.424	11.720
2 step ahead	236	4.679	0.326	6.192	38.337	6.358	0.443	8.182	66.940	5.361	0.375	7.154	51.178	4.931	0.342	6.665	44.426
3 step ahead	235	6.854	0.477	8.993	80.866	8.452	0.589	10.666	113.758	7.440	0.520	9.781	95.676	7.047	0.489	9.278	86.075
4 step ahead	234	8.838	0.616	11.352	128.866	9.887	0.689	12.431	154.533	9.345	0.654	11.965	143.152	8.815	0.612	11.342	128.639
5 step ahead	233	10.316	0.719	13.062	170.604	11.054	0.771	13.984	195.557	10.598	0.741	13.521	182.819	10.160	0.706	13.039	170.009
6 step ahead	232	11.477	0.800	14.553	211.785	11.820	0.824	15.093	227.790	11.663	0.816	14.823	219.717	11.330	0.788	14.361	206.245
7 step ahead	231	12.246	0.854	15.556	241.988	12.878	0.892	15.978	255.284	12.449	0.871	15.767	248.605	12.021	0.836	15.318	234.652
8 step ahead	230	13.234	0.923	16.473	271.358	13.515	0.942	16.760	280.906	13.383	0.937	16.641	276.923	12.884	0.896	16.048	257.528
9 step ahead	229	13.940	0.972	17.243	297.335	14.262	0.995	17.529	307.271	14.099	0.987	17.422	303.518	13.628	0.947	16.804	282.361
I0 step ahead	228	14.700	1.025	17.984	323.431	14.940	1.042	18.216	331.820	14.800	1.036	18.118	328.259	14.274	0.992	17.466	305.047

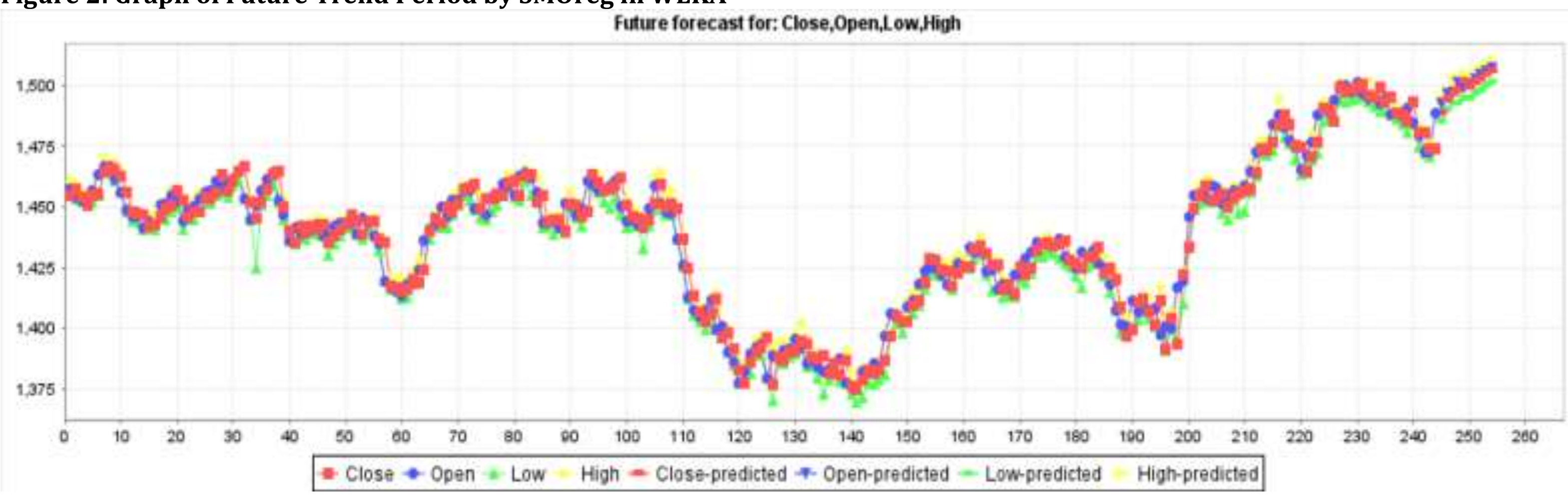
Table 3: MAE, MAPE, RMSE & MSE for Open, High, Low & Close attributes obtained from Multilayer Perceptron (MLP)

MLP															
Close				Open				Low				High			
MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE	MAE	MAPE	RMSE	MSE
2.197	0.153	2.886	8.332	7.284	0.506	8.538	72.891	7.067	0.493	7.766	60.315	3.335	0.231	4.316	18.626
8.237	0.574	9.683	93.765	14.870	1.032	16.617	276.135	14.728	1.027	16.107	259.435	10.209	0.708	11.753	138.132
16.319	1.134	18.385	337.995	24.018	1.667	26.075	679.926	23.319	1.625	25.055	627.772	18.613	1.289	20.514	420.811
25.723	1.787	28.098	789.473	33.140	2.299	35.423	1254.780	32.771	2.282	34.878	1216.502	27.541	1.908	29.695	881.792
34.869	2.421	37.432	1401.166	42.528	2.949	45.023	2027.069	41.786	2.908	44.072	1942.338	36.580	2.533	39.030	1523.302
44.684	3.100	47.389	2245.750	51.374	3.561	54.112	2928.070	50.091	3.485	52.571	2763.675	45.501	3.149	48.180	2321.284
54.566	3.783	57.425	3297.635	60.226	4.173	63.441	4024.802	58.588	4.073	61.383	3767.898	54.398	3.763	57.349	3288.866
64.071	4.439	67.369	4538.569	68.806	4.764	72.882	5311.739	67.113	4.664	70.537	4975.516	62.911	4.348	66.446	4415.087
72.240	5.002	76.357	5830.344	76.210	5.274	81.384	6623.344	74.631	5.184	78.921	6228.513	70.641	4.785	75.189	5653.356
78.999	5.466	84.118	7075.816	81.659	5.649	87.898	7726.081	80.738	5.606	85.855	7371.096	76.692	5.293	82.407	6790.901

Table 4: Comparing the Predicted and Actual Prices of FTSE Bursa Malaysia KLCI

SMOreg				MLP				Actual Rates			
Close	Open	Low	High	Close	Open	Low	High	Close	Open	Low	High
1488.874	1492.632	1485.534	1495.186	1393.52	1377.99	1375.44	1395.42	1453.100	1452.200	1446.360	1453.560
1494.086	1496.829	1490.454	1497.782	1496.022	1486.265	1484.062	1493.028	1462.370	1452.540	1450.170	1465.690
1496.795	1497.489	1493.068	1502.383	1492.412	1483.185	1480.849	1489.514	1477.260	1462.850	1461.030	1477.560
1498.231	1500.983	1492.948	1502.640	1485.604	1476.874	1476.651	1485.361	1487.610	1476.890	1476.850	1487.610
1502.065	1499.388	1495.027	1504.857	1480.366	1471.422	1469.627	1478.579	1495.700	1488.600	1488.600	1498.520
1499.686	1500.659	1494.804	1504.209	1474.889	1466.198	1464.044	1473.474	1498.830	1497.170	1496.840	1503.930
1501.260	1502.238	1496.877	1504.620	1468.882	1461.362	1460.237	1468.202	1486.860	1497.800	1486.720	1497.800
1502.936	1504.258	1497.945	1506.695	1463.708	1456.488	1454.345	1463.740	1483.000	1487.150	1483.000	1492.440
1505.172	1505.724	1500.119	1508.514	1457.953	1451.132	1448.690	1461.532	1487.340	1483.240	1482.020	1487.340
1506.124	1507.546	1501.701	1510.414	1453.038	1445.673	1444.384	1456.984	1501.110	1488.280	1487.810	1501.860

Figure 2: Graph of Future Trend Period by SMOreg in WEKA



References

Abd Samad, P. H. D., Mutalib, S., & Abdul-Rahman, S. (2019). Analytics of stock market prices based on machine learning algorithms. *Indonesian Journal of Electrical Engineering and Computer Science, 16*(2), 1050. http://doi.org/10.11591/ijeecs.v16.i2.pp1050-1058.

Ahmed, M. U., & Hussain, I. (2022). Prediction of Wheat Production Using Machine Learning Algorithms in northern areas of Pakistan. *Telecommunications Policy, 46*(6), 102370. https://doi.org/10.1016/j.telpol.2022.102370.

Al-Radaideh, Q., Assaf, A., & Alnagi, E. (2013). Predicting stock prices using data mining techniques. *The International Arab Conference on Information Technology,* December 2013, 8.

Choudhry, T., Hasan, M., & Zhang, Y. (2019). Forecasting the daily dynamic hedge ratios in emerging European stock futures markets: evidence from GARCH models. *International Journal of Banking, Accounting and Finance, 10*(1), 67-100. https://doi.org/10.1504/IJBAAF.2019.099316.

Devi, A. U., Vishal Kumar, R., Priya, P. S., Amzad Basha, M. S., Sucharitha, M. M. and Oveis, P. M. (2023). A Data Mining approach on the Performance of Machine Learning Methods for Share Price Forecasting using the Weka Environment, *Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT), Erode, India*, 2023, 01-06, doi: 10.1109/ICECCT56650.2023.10179772.

Dong, S., Wang, J., Luo, H., Wang, H., & Wu, F. X. (2021). A dynamic predictor selection algorithm for predicting stock market movement. *Expert Systems with Applications, 186*(August), 115836. https://doi.org/10.1016/j.eswa.2021.115836.

Dong, H., Guo, X., Reichgelt, H., & Hu, R. (2020). The predictive power of ARIMA models in forecasting equity returns: a sliding window method. *Journal of Asset Management, 21*, 549-566. https://doi.org/1057/s41260-020-00184-z.

Emami, S. S. (2018). Predicting Trends of Stock Prices by Developing Data Mining Techniques with the Aim of Gaining Profit. *Journal of Accounting & Marketing, 07*(04). https://doi.org/10.4172/2168-9601.1000304

Farias Nazário, R. T., e Silva, J. L., Sobreiro, V. A., & Kimura, H. (2017). A literature review of technical analysis on stock markets. *Quarterly Review of Economics and Finance, 66*, 115–126. https://doi.org/10.1016/j.qref.2017.01.014

Frank, E., Hall, M. A., & Witten, I. H. (2017). The WEKA workbench. Data Mining, 553–571. https://doi.org/10.1016/b978-0-12-804291-5.00024-6

Gandhmal, D. P., & Kumar, K. (2019). Systematic analysis and review of stock market prediction techniques. *Computer Science Review, 34*, 100190. https://doi.org/10.1016/j.cosrev.2019.08.001

Ismail, M. S., Noorani, M. S. M., Ismail, M., Razak, F. A., & Alias, M. A. (2020). Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange. *Applied Soft Computing, 93*, 106422. https://doi.org/10.1016/j.asoc.2020.106422.

Iaousse, M., Jouilil, Y., Bouincha, M., & Mentagui, D. (2023). A Comparative Simulation Study of Classical and Machine Learning Techniques for Forecasting Time Series Data. *International Journal of Online and Biomedical Engineering, 19*(8), 56–65. https://doi.org/10.3991/ijoe.v19i08.39853.

Jiang, W. (2021). Applications of deep learning in stock market prediction: *Recent progress. Expert Systems with Applications, 184*(June), 115537. https://doi.org/10.1016/j.eswa.2021.115537.

Jošić, B. Ž. and H. (2020). Forecasting Stock Market Indices Using Machine Learning Algorithms. *Interdisciplinary Description of Complex Systems, 18*(4), 471–489. https://doi.org/10.7906/indecs.18.4.7.

Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications, 41*(16), 7653–7670. https://doi.org/10.1016/j.eswa.2014.06.009.

Kulkarni, A., & More, A. (2016). Formulation of a Prediction Index with the Help of WEKA Tool for Guiding the Stock Market Investors. *Oriental Journal of Computer Science and Technology, 9*(3), 212–225. https://doi.org/10.13005/ojcst/09.03.07.

Kumar, B. S., & Ravi, V. (2016). A survey of the applications of text mining in the financial domain. *Knowledge-Based Systems, 114*, 128–147. https://doi.org/10.1016/j.knosys.2016.10.003.

Kumar, D., Sarangi, P. K., & Verma, R. (2020). A systematic review of stock market prediction using machine learning and statistical techniques. *Materials Today: Proceedings, 49*, 3187–3191. https://doi.org/10.1016/j.matpr.2020.11.399.

Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications, 197*(February), 116659. https://doi.org/10.1016/j.eswa.2022.116659.

Mintarya, L. N., Halim, J. N. M., Angie, C., Achmad, S., & Kurniawan, A. (2022). Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science, 216*, 96–102. https://doi.org/10.1016/j.procs.2022.12.115.

Md Salim Chowdhury, Norun Nabi, Md Nasir Uddin Rana, Mujiba Shaima, Hammed Esa, Anik Mitra, Md Abu Sufian Mozumder, Irin Akter Liza, Md Murshid Reja Sweet, & Refat Naznin. (2024). Deep Learning Models for Stock Market Forecasting: A Comprehensive Comparative Analysis. *Journal of Business and Management Studies, 6*(2), 95–99. https://doi.org/10.32996/jbms.2024.6.2.9.

Navarro, M. M., Young, M. N., Prasetyo, Y. T., & Taylar, J. V. (2023). Stock market optimization amidst the COVID-19 pandemic: Technical analysis, K-means algorithm, and mean-variance model (TAKMV) approach. *Heliyon, 9*(7), e17577. https://doi.org/10.1016/j.heliyon.2023.e17577.

Nirob, F. A., & Hasan, M. M. (2023). Predicting Stock Price from Historical Data using LSTM Technique. *Journal of Artificial Intelligence and Data Science (JAIDA), 3*(1), 36–49. https://dergipark.org.tr/pub/jaida.

Patil, H., Bolla, B. K., Sabeesh, E., & Bhumireddy, D. R. (2024). Comparative Study of Predicting Stock Index Using Deep Learning Models. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, 536*, 45–57. https://doi.org/10.1007/978-3-031-48888-7_4.

Patra, S., & Hiremath, G. S. (2022). An Entropy Approach to Measure the Dynamic Stock Market Efficiency. *Journal of Quantitative Economics, 20*(2), 337-377. https://doi.org/10.1007/s40953-022-00295-x.

Peng, K. L., Wu, C. H., Lin, P. M. C., & Kou, I. T. E. (2023). Investor sentiment in the tourism stock market. *Journal of Behavioral and Experimental Finance, 37*, 100732. https://doi.org/10.1016/j.jbef.2022.100732.

Rashedi, K. A., Ismail, M. T., Al Wadi, S., Serroukh, A., Alshammari, T. S., & Jaber, J. J. (2024). Multi-Layer Perceptron-Based Classification with Application to Outlier Detection in Saudi Arabia Stock Returns. *Journal of Risk and Financial Management, 17*(2). https://doi.org/10.3390/jrfm17020069.

Rehman, I., & Soomro, T. R. (2019). Data Mining for Forecasting OGDCL Share Prices Using WEKA. MACS 2019 - 13th International Conference on Mathematics, Actuarial Science, *Computer Science and Statistics, Proceedings*, 1–7. https://doi.org/10.1109/MACS48846.2019.9024780.

Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. *Electronics (Switzerland), 10*(21). https://doi.org/10.3390/electronics10212717.

Rubio, L., & Alba, K. (2022). Forecasting selected Colombian shares using a hybrid ARIMA-SVR model. *Mathematics, 10*(13), 2181. https://doi.org/10.3390/math10132181.

Shah, A., Gor, M., Sagar, M., & Shah, M. (2022). A stock market trading framework based on deep learning architectures. *Multimedia Tools and Applications, 81*(10), 14153–14171. https://doi.org/10.1007/s11042-022-12328-x.

Villavicencio, C. N., Macrohon, J. J. E., Inbaraj, X. A., Jeng, J. H., & Hsieh, J. G. (2021). COVID-19 prediction applying supervised machine learning algorithms with comparative analysis using Weka. *Algorithms, 14*(7). https://doi.org/10.3390/a14070201.

Žmuk, B., & Jošic, H. (2020). Forecasting Stock Market Indices Using Machine Learning Algorithms. *Interdisciplinary Description of Complex Systems, 18*(4), 471–489. https://doi.org/10.7906/indecs.18.4.7.

Zou, J., Zhao, Q., Jiao, Y., Cao, H., Liu, Y., Yan, Q., Abbasnejad, E., Liu, L., & Shi, J. Q. (2023). Stock Market Prediction via Deep Learning Techniques: *A Survey.* 1(1). https://doi.org/10.48550/arXiv.2212.12717.