Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

Comparative Analysis of LSTM, FB Prophet, and Moving Average Methods for Fuel Sales Prediction: A Time Series Forecasting Approach

Ahmad Rizky Fadhilah[1,*], Arbi Haza Nasution[1], Winda Monika[2]

[1] Department of Informatics Engineering, Universitas Islam Riau, Pekanbaru 28284, Indonesia
[2] Department of Library Information, Universitas Lancang Kuning, Riau 28266, Indonesia
Email: [1*]ahmadrizkyfadhilah@student.uir.ac.id, [2]arbi@eng.uir.ac.id, [3]windamonika@unilak.ac.id
Correspondence Author Email: ahmadrizkyfadhilah@student.uir.ac.id
Submitted: **02/09/2024**; Accepted: **01/12/2024**; Published**: 03/12/2024**

Abstract−Fuel is an important part of vehicles and machinery where sales demand is very high and has various fluctuations. The uncertainty in these fuel sales patterns poses serious problems in inventory management and fuel distribution planning in Indonesia, which can result in excess stock or fuel scarcity in various regions. Additionally, changing trends in vehicle usage and the impact of the COVID-19 pandemic have made accurate sales predictions increasingly difficult. Therefore, this research aims to understand the current and future sales patterns and trends of fuel sales in Indonesia. Careful analysis of prices and other factors such as data processing and other variables is required. This study uses time series analysis methods and compares four models, namely Long Short-Term Memory (LSTM), FB Prophet, Simple Moving Average (SMA), and Exponential Moving Average (EMA). By comparing the results using statistics such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) over various prediction time frames, we assessed the patterns of each model. The results of the analysis show that the LSTM model outperformed all other methods with the lowest MAPE for the prediction of gasoline in the next 31 days, which is 17.11%, while the FB Prophet outperformed all other methods with the lowest MAPE for the prediction of diesel in the next 31 days, which is 18.32%. Although the LSTM model generally outperformed all other algorithms, the FB Prophet model can be used to predict future trends, such as increased use of diesel and decreased use of gasoline which are expected to last within one year. This analysis also provides insights for choosing the right model for a time series problem, including the characteristics of the data to be predicted and analyzed, as well as the assumptions of stationarity and normality of the data. The results of this study indicate that machine learning algorithms can improve the accuracy of time series predictions significantly compared to traditional statistical methods.

Keywords: Fuels Analysis; Sales Prediction; Forecasting; LSTM; FB Prophet; Moving Average

1. INTRODUCTION

The use of fuels plays an important role in the world of vehicles and industrial machinery. In Indonesia, fuel consumption is particularly significant due to the country's reliance on fossil fuels for a wide range of activities. Whether for personal vehicles or large-scale industrial machinery, fuel is an indispensable resource. This high dependence on fuel creates a substantial demand that fluctuates regularly due to factors such as seasonal variations, economic conditions, and external disruptions like the COVID-19 pandemic. Such fluctuations have serious implications for inventory management and distribution planning, often leading to situations where fuel stocks are either over-supplied or insufficiently available. Consequently, an accurate forecast of fuel sales is crucial to optimize supply chains, maintain economic stability, and aid in strategic decision-making for stakeholders such as government bodies, industries, and consumers.

One of the most prominent challenges in fuel management is predicting the future demand, given that the market is subject to a wide range of uncontrollable factors. For instance, price fluctuations not only affect individual consumption but also ripple through various sectors, creating uncertainty that can disrupt national economic stability. Moreover, with the rise of alternative energy sources, changing vehicle usage patterns, and economic slowdowns, particularly during the pandemic, it becomes increasingly difficult to develop accurate sales forecasts using conventional methods. These challenges underscore the need for sophisticated predictive models that can account for the complexity and volatility of fuel demand [1]. To address this issue, time series analysis emerges as a powerful tool for forecasting trends based on historical data. Time series analysis allows organizations to predict future events by analyzing sequences of data points collected at consistent intervals. Through the application of time series forecasting, companies and governments alike can make informed decisions regarding inventory control, price setting, and long-term strategic planning. Unlike simple data collection methods, time series analysis reveals patterns, highlights shifts in trends, and offers valuable insights into the underlying factors driving these changes.

Previous research has explored several forecasting methods to predict fuel prices and demand, with machine learning techniques like Long Short-Term Memory (LSTM) proving especially effective in handling complex data patterns. Studies have demonstrated that LSTM models are highly accurate compared to traditional statistical models like regression analysis, especially in the context of time series predictions[2], [3]. However, despite their strengths, neural network models such as LSTM are not without limitations. Cumulative errors in neural networks and models like FB Prophet can significantly reduce the reliability of long-term forecasts. These cumulative errors, often caused by minor inaccuracies in individual data points, can compound over time, creating larger discrepancies in predictions[4].

Despite these advancements, each model has its limitations. For example, LSTM and FB Prophet are both susceptible to cumulative errors, where minor inaccuracies at different data points compound over time, affecting the reliability of the forecast. On the other hand, traditional models such as the Moving Average (MA), which includes

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

Simple Moving Average (SMA) and Exponential Moving Average (EMA), provide more stable predictions by smoothing out short-term fluctuations. However, Moving Average models are restricted by their inherent simplicity, as they are only influenced by errors within a specific time window. These methods are less effective in volatile markets where data is subject to rapid changes [5]. Nevertheless, Moving Averages remain widely used in market analysis, especially when dealing with stable trends over short periods [6].

The use of hybrid models that combine traditional statistical methods like MA with machine learning algorithms such as LSTM has been shown to enhance prediction accuracy, especially in volatile markets. These models are better at capturing complex data patterns compared to traditional approaches like SMA or ARIMA[7], [8]. Hybrid models help mitigate some of the weaknesses of individual models by combining the robustness of statistical models with the learning capabilities of neural networks. For instance, Random Forest has proven to be effective at capturing subtle and diverse patterns in historical data compared to traditional models like ARIMA and MA [9]. In recent studies, LSTM models have consistently outperformed other models in predicting time-series data, including fuel demand. LSTM can capture long-term dependencies that simpler methods, like SMA, cannot, making it a superior choice for complex forecasting tasks [10], [11]. While FB Prophet is excellent for handling seasonal trends, LSTM excels in capturing both short-term fluctuations and long-term trends, making it highly applicable to fuel demand forecasting, where both periodicity and long-term consumption changes are essential. Recent advancements in hybrid models which combine traditional statistical approaches with machine learning techniques have shown promise in improving forecast accuracy in volatile markets. Hybrid models, such as combining LSTM with ARIMA or FB Prophet, excel at capturing complex patterns in fuel sales data and outperforming single models in terms of accuracy and robustness. Research by [4] in 2021, for instance, demonstrated that combining machine learning models with statistical approaches yielded superior results in predicting commodity prices in markets experiencing high volatility. Similarly, [12] explored the use of Random Forest in conjunction with time series models, finding that the machine learning-based approach was able to capture subtle patterns that traditional models often overlook.

Price prediction and fuel sales forecasting are among the most pressing challenges for businesses and policymakers alike. In the context of fuel, the dynamic and future oriented nature of the market means that stakeholders must rely on forecasts to manage supply chains, regulate prices, and anticipate future demand [3], [6]. While statistical models have traditionally been used to simplify the complexities of market phenomena, artificial intelligence (AI) and machine learning models are now emerging as essential tools for analyzing non-linear data structures and producing more reliable forecasts. Studies such as demonstrate that AI-based models, particularly LSTM, are well-suited for addressing the long-term dependencies often found in time series data, making them ideal for predicting fuel demand [5].

While several studies have focused on the application of LSTM and other models in forecasting fuel prices in international markets, there remains a significant gap in the literature concerning the use of these models for predicting fuel sales in Indonesia. Given the unique characteristics of the Indonesian market including geographical diversity, regional consumption patterns, and economic disparities a tailored approach to fuel sales prediction is required. Moreover, Indonesia's reliance on imported fuel adds an extra layer of complexity to forecasting, as fluctuations in global oil prices directly affect local markets. Therefore, this research aims to address these gaps by conducting a comparative analysis of four key forecasting models: Long Short-Term Memory (LSTM), FB Prophet, Simple Moving Average (SMA), and Exponential Moving Average (EMA). By evaluating the performance of these models using metrics such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), this study seeks to determine which model offers the most accurate predictions for fuel sales in Indonesia. Furthermore, this research takes into account external factors such as price trends, economic conditions, and the impact of global market fluctuations, aiming to provide a robust predictive framework that can aid decision-making in fuel inventory management and distribution planning.

A lot of research was done for time series forecasting using Fb Prophet for various problems. Predicts product selection based on future requests, describing time series curve approach for a particular product within the product portfolio [13]. In another research adaptive filter is used in conjunction with FB Prophet to maximize prediction demand with better prediction, even though the FB Prophet model is relatively new [14]. Works on forecasting trend that occur in crude oil market volatility by using the Simple Moving Average (SMA), Exponential Moving Average (EMA), and Kaufman Adaptive Moving Average (KAMA) taking into account the non-linear nature and varying times over time [7]. Therefore, this research aims to address gaps in fuel sales forecasting by comparing the performance of LSTM, FB Prophet, SMA, and EMA models. By evaluating their accuracy using metrics such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), this study seeks to determine the most effective model for forecasting fuel sales in Indonesia. The unique characteristics of the Indonesian market, including geographical diversity, regional consumption patterns, and reliance on imported fuel, make it crucial to adopt accurate predictive models that can accommodate market fluctuations and external factors[15].

Machine learning algorithms are also widely applied in various other fields, such as peat fires prediction. It was concluded that the LSTM could also be the basis for predicting peat fires which are very vulnerable to occur in South and Southeast Asian countries [16]. Another study compared several machine learning algorithms that produce predictions for tourists who come by relying on hybrid models such as conventional models and AI-based models, but these models are not sufficient in making predictions if only using these two models [17]. Performs forecasting of water use consumption in high accuracy and short term [18].

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

2. RESEARCH METHODOLOGY

2.1 Dataset

This raw data is not open to the public and obtained directly from four refueling stations located in Riau, Indonesia. We get fuel data with octane 90 (pertalite) and diesel fuel for the period of early 2016 to ended 2021. The data format is numeric and consists of the number of sellers and the amount of fuel in the scale of tons. After getting the data in excel form, we immediately convert the data into csv form and preprocess the data by looking at the required variables such as date, amount and selling price which are useful in the analysis process.

Table 1. Gasoline dataset.

Date	Qty	Sales
2016-01-01	8.694	63.467.067
2016-01-02	10.044	73.320.032
2016-01-03	13.928	101.671.188
2016-01-04	12.518	91.382.972
2016-01-05	11.112	77.424.375
...	...	...
2021-10-25	12.239	93.628.350
2021-10-26	11.319	86.590.350
2021-10-27	12.502	95.640.300
2021-10-28	12.456	95.288.400
2021-10-29	11.228	85.894.200

From Table 1 contains daily sales data for gasoline. Each day, the number of liters sold and the total revenue is recorded. For Example, on 01 January 2016, 8.694 liters of gasoline were sold. This dataset focuses exclusively on gasoline, meaning the sales pattern may differ from other fuels like diesel, which is covered in Table 2. During the COVID-19 pandemic, gasoline sales experienced a significant decline. The pandemic caused widespread restrictions on movement, resulting in reduced travel for both personal and commercial vehicles. This led to a substantial drop in fuel demand, as people stayed home and industries scaled down operations.

Table 2. Diesel dataset.

Date	Qty	Sales
2016-01-01	836	5.600.932
2016-01-02	964	6.460.341
2016-01-03	1.900	12.727.990
2016-01-04	1.944	13.023.929
2016-01-05	7.653	43.238.490
...	...	...
2021-12-27	3.892	20.043.800
2021-12-28	3.848	19.817.200
2021-12-29	3.915	20.162.250
2021-12-30	2.715	13.982.250
2021-12-31	3.177	16.361.550

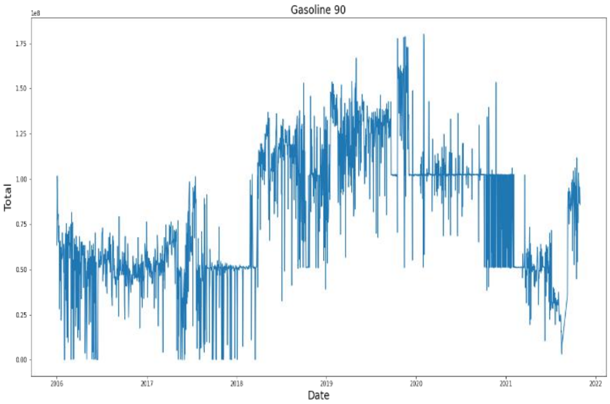
Table 2 provides daily sales data for diesel, much like Table 1 but focused on diesel. On 01 January 2016, 836 liters of diesel were sold. Diesel is typically used in large vehicles or industrial machinery. Compared to gasoline, diesel follows a different usage pattern, which reflects in its sales volume and period. Diesel consumption is more influenced by industrial activity than personal vehicle usage.From several conditions that the data may not be able to determine, such as using the LSTM algorithm to analyze future predictions, it will be more accurate if there is a lot of training data available. With the data that has been collected, we will then carry out the process of cleaning the dataset.

Table 3. Describe of dataset.

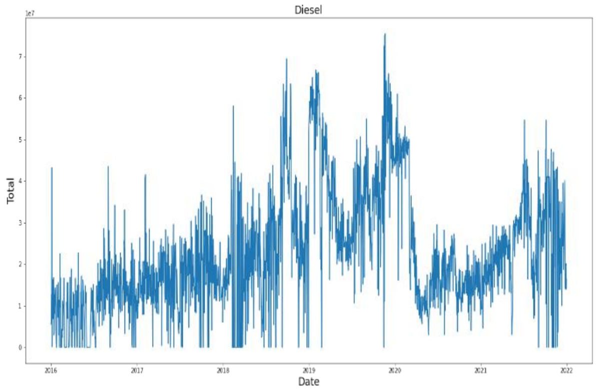
Type	Gasoline	Diesel
Rows	1.998	2.088
Train Set	1.598	1.670
Test Set	400	418
Min Qty	1	20
Max Qty	27.922	14.659
Mean	12.431,88	4.625,58

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

Table 3 provides an overview of the dataset size. The dataset size and distribution differ between gasoline and diesel. The average daily gasoline sales 12,431 liters are significantly higher than diesel 4,625 liters, indicating different consumption patterns between the two fuels. There are 1,998 rows for gasoline and 2,088 rows for diesel. The data is split into a training set 80% and a test set 20% for building the prediction models. Next, we compare the results from LSTM, MA and FB Prophet by looking at the error rate by using the Root-Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).



(a) Gasoline.



(b) Diesel.

Figure 1. (a) Sales plot of gasoline fuels. **(b)** Sales plot of diesel fuels.

From Figure 1a graph illustrates the fluctuations in gasoline sales from 2016 to 2021. There is a significant drop during the COVID-19 pandemic, showing how external events impact sales. The fluctuation in this graph is more extreme than in diesel sales, as gasoline is more influenced by personal vehicle activity, which was heavily affected by the pandemic. Figure 1b shows diesel sales, which are generally more stable compared to gasoline, though there are dips during the pandemic and diesel sales increased after the economic recovery. Diesel demand grows more steadily, influenced by the industrial sector, which doesn't halt as easily as personal vehicle use during crises. From Figure 1 it can also be seen that there are intensity value of zero (0) due to a delay in the distribution of these materials which resulted in sales being delayed that day. Bildirici et al. [19] also experienced poor performance in analyzing price impact during COVID 19 pandemic. As the most important source of energy for economic activity is petroleum, fuel is also the most demanded commodity on the globe [20].

2.2 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is an alternative architecture developed to complement deficiencies in Recurrent Neural Networks (RNN). Long Short-Term Memory (LSTM) was introduced in 1997 by Hochreiter and Schmidhuber. LSTM replaces the hidden layer nodes in RNN with LSTM cells that which serves to store previous information. In RNN, a repeating layer or hidden layer consists of repeating cells whose state is affected by the past state and the current input with a feedback connection. Repeat layers can be arranged in various architectures to form different RNNs. The LSTM architecture is shown in Figure 3. RNNs are a type of neural network that is constructed to handle sequential input [21]. This structure consists of neurons that function as cells c_t, and three types of gates, called input gates i_t, forget gates f_t, and output gates o_t. These three different gate types work to direct the input from the (new) incoming data point x_t, to the neuron cell, which on exposure directly affects the computational output h_t, of the entire LSTM unit.

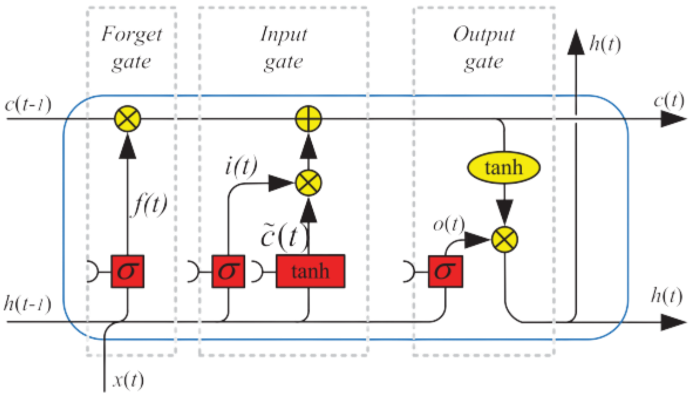


Figure 2. Architecture of LSTM.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

Based on Figure 2 LSTM architecture can be described mathematically as:

$$\begin{aligned} f_t &= \sigma(W_f h_{t-1} + W_f x_t + b_f) \\ i_t &= \sigma(W_i h_{t-1} + W_i x_t + b_i) \\ \tilde{c}_t &= \sigma(W_c h_{t-1} + W_c x_t + b_c) \\ c_t &= f_t . c_{t-1} + i_t . \tilde{c}_t \\ o_t &= \sigma(W_o h_{t-1} + W_o x_t + b_o) \\ h_t &= o_t . tanh(c_t) \end{aligned} \quad (1)$$

The forget gate (f_t) in the LSTM controls how much information is deleted or retained from the previous cell. When f_t has a value of 1, it means all information will be retained, whereas when f_t has a value of 0, it means all information will be deleted from the previous cell. Meanwhile, input cell (i_t) and output cell (o_t) in LSTM are described by current input (x_t), previous output (h_{t-1}), and current output (y_t) of each LSTM equation.

2.3 FB Prophet

FB Prophet Model is a time series algorithm that is used to make predictions from data based on additive models where non-linear trends will be matched in time series on an annual, weekly and daily basis. This model is great for missing data and trending movements and it can also produce accurate and reliable forecasts.

$$y(t) = g_{(t)} + h_{(t)} + s_{(t)} + e_{(t)} \quad (2)$$

From Equation 2 it is shown that $g_{(t)}$ represents non-periodic changes in the period of time, $s_{(t)}$ represents periodic changes in the period of weeks or even years, $h_{(t)}$ represents the impact of holidays that occurs on a potentially irregular schedule of one or more days and $e_{(t)}$ represents a change error not facilitated by this model. [22] explains that the trend factor at $g_{(t)}$ is modeled in two ways:

a. Logistic Growth Model

$$f_{(x)} = \frac{L}{1+e^{-k(x-x0)}} \quad (3)$$

As shown in Equation 3, this model represents the growth stages which are observed to be about exponential once the saturation stage is reached from there growing linearly.

b. Piece-Wise Linear Model

$$y = \beta_0 + \beta_1 x + \beta_2 (x-c)^+ + \epsilon \quad (4)$$

Formula on Equation~\ref{eq4} is a modified version of the linear model in which different ranges of x have different linear relationships.

2.4 Moving Average

Moving Average (MA) is an algorithm that is often used in stock analysis. By calculating a moving average, the impact of short-term random fluctuations on the price over a certain period of time is reduced. Moving Average (MA) is divided into several types, in this research, we only use 2 of several types of Moving Average, namely the Simple Moving Average (SMA) which uses a simple arithmetic average of prices within a certain period and the Exponential Moving Average (EMA) which gives more weight to more recent prices than older prices during that time period. The reason for calculating stock moving averages is to help smooth out price data by creating a constantly updated average price. Karasu et al. [7] also explains that SMA is a simple average of the price of a security over a certain period of time, while EMA is a commonly used technical indicator that gives greater weight to exponentially newer prices.

$$SMA = \frac{A_1 + A_2 + A_3 + \cdots + A_n}{n} \quad (5)$$

The calculation of this SMA formula is by taking the arithmetic average of a set of data over a certain period and making it in the form of a set of numbers represented by A_1 and then adding up to A_n as a whole and dividing it by the number of prices in that set.

$$EMA_{today} = EMA_{yesterday} + \alpha \times (Price_{today} - EMA_{yesterday}) \quad (6)$$

EMA is a type of MA that uses weighting and filters data infinitely. In other words, the EMA will not throw away old data like the SMA, but the weight of the data will decrease exponentially. So to calculate the EMA, we first compute the SMA over a certain period and then calculates a multiplier to give the weight to the EMA. The smoothing factor is combined with the previous EMA to get the current value. EMA gives a higher weight to the current price, while SMA gives the same weight to all values. So the longer the data from the calculated EMA, the less weight the data will have in the calculation.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

2.5 Evaluation Metrics

In predictive analysis evaluations are definitely needed to predict future values or events based on historical data. To determine how accurate the predictions from the model are, we need the right evaluation metrics. One of the commonly used evaluation metrics for time-series analysis is the RMSE. RMSE calculates the root of the average squared difference between the predicted value and the actual value as shown in Equation 7. The lower the RMSE value, the more accurate the model prediction, because it indicates a lower error rate.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y_t})^2} \tag{7}$$

In addition, another evaluation metric that can be used is the MAPE. MAPE measures the percentage of prediction error relative to the true value as shown in Equation 8. MAPE provides a more intuitive picture of prediction accuracy, because it measures error in percentage terms. If the MAPE value is low, it indicates that the model prediction is quite accurate.

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y_t}}{y_t}\right| \times 100\% \tag{8}$$

3. RESULT AND DISCUSSION

After getting all the required variables and clean data, we conducted series of experiments comparing those four algorithms which are LSTM, Fb Prophet, SMA and EMA. This comparison is generated by visualizing the predictions and at this stage we do a comparison of each layer and the proportions of existing RMSE and MAPE calculations to get fairly good its mean has lower error rate results in this comparative analysis. The models are trained using AMD Ryzen 9 5900X with a 3.7 GHz base clock computer with 32 Gb of RAM.

For the first algorithm, i.e., LSTM, initially, we would like to identify the best epoch for both Gasoline and Diesel datasets. In Table 4 it can be seen that starting from 50 epoch, there is a trend that the error rate is decreasing. However, starting from 750 epoch, the error rate is increasing, thus we stop the experiment on 1000 epoch. The best result is with 500 epoch for both Gasoline and Diesel datasets. Therefore, in the following experiments, we will use 500 epoch.

Table 4. Experiments with LSTM algorithm on several epochs.

Type of Fuels	LSTM Model	MAPE
Gasoline	10 epoch	24.21 %
	50 epoch	24.34 %
	100 epoch	23.96 %
	250 epoch	23.11 %
	500 epoch	**22.47 %**
	750 epoch	22.89 %
	1000 epoch	31.00 %
Diesel	10 epoch	24.01 %
	50 epoch	23.91 %
	100 epoch	23.76 %
	250 epoch	23.71 %
	500 epoch	**23.67 %**
	750 epoch	25.76 %
	1000 epoch	41.87 %

Table 4 shows the results of using Long Short-Term Memory (LSTM) model with varying numbers of epochs (iterations of training). The best result with the lowest error rate (MAPE) is achieved at 500 epochs. LSTM requires

(a) 31 Days.

(b) 365 Days.

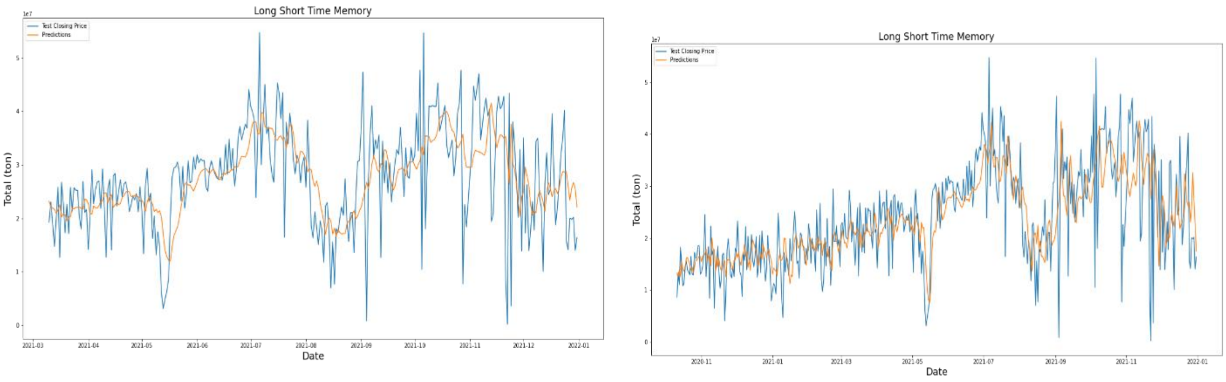
Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

longer and more complex training compared to simpler models like Moving Average, due to its ability to understand intricate data patterns.

The experiment for gasoline dataset is presented in Figure 3 and the visualization for diesel is presented in Figure 4. From the visualization it can also be seen that the best graph flow can be seen from the smaller margin of errors than the others. As shown in Figure 3a, LSTM model that predict 31 days ahead for gasoline has a smallest percentage of errors due to the small prediction time span. The same goes for diesel dataset where LSTM model that predict 31 days ahead has a smallest percentage of errors due to the small prediction time span. Long-term predictions 365 days have higher error rates compared to short-term predictions 31 days.

Figure 4a shows the LSTM prediction for diesel sales over 31 days, with the model demonstrating strong performance for short-term predictions. LSTM model performs well for short-term diesel sales. Figure 4b illustrates the diesel sales prediction for 365 days, where the error rate increases and the model's accuracy declines over time. However, Figure 4b shows that long-term diesel forecasts are less reliable, much like with gasoline, though diesel has a more stable demand curve.

The second algorithm, i.e., FB Prophet is a predictive tool that utilizes time series models to predict trends and patterns in real-time data. FB Prophet can also accurately predict future fuel usage by taking into account factors such as seasonal patterns, long-term trends, as well as special events such as Covid-19 and the holiday season that affect fuel use. In addition, by looking at Figures 5 and Figure 6, it can be seen that there is a different trend in the prediction result between the gasoline and the diesel. In this regard, it is important to consider that the farther the prediction time is, the more difficult it is to accurately predict future fuel usage. As shown in Figure 5, it can be seen that the observed variable value trend tends to decrease in the analysis pattern for gasoline fuel, while increase steadily for the type of diesel fuel as shown in Figure 6. However, there are fluctuations in the resulting predicted value, which can be caused by external factors that cannot be predicted by the model. This model must also be evaluated periodically to ensure the accuracy of the predictions in this model. In this condition it can also be assumed that the development of technology such as electric vehicles has increased its use which has resulted in a decrease for gasoline fuel, whereas

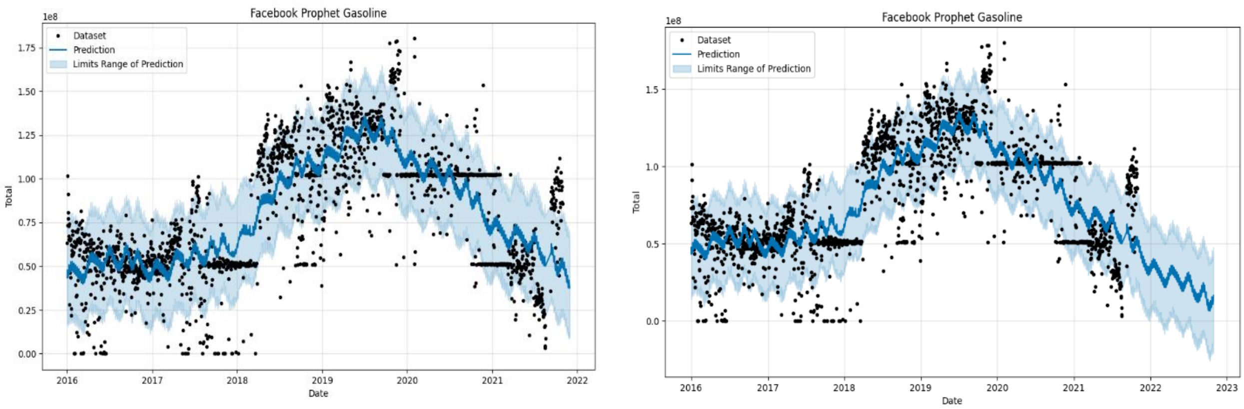


(a) 31 Days.

(b) 365 Days.

Figure 3. Output LSTM model for diesel

diesel fuel such as trucks, production machines and other means of transportation are not easy to replace.



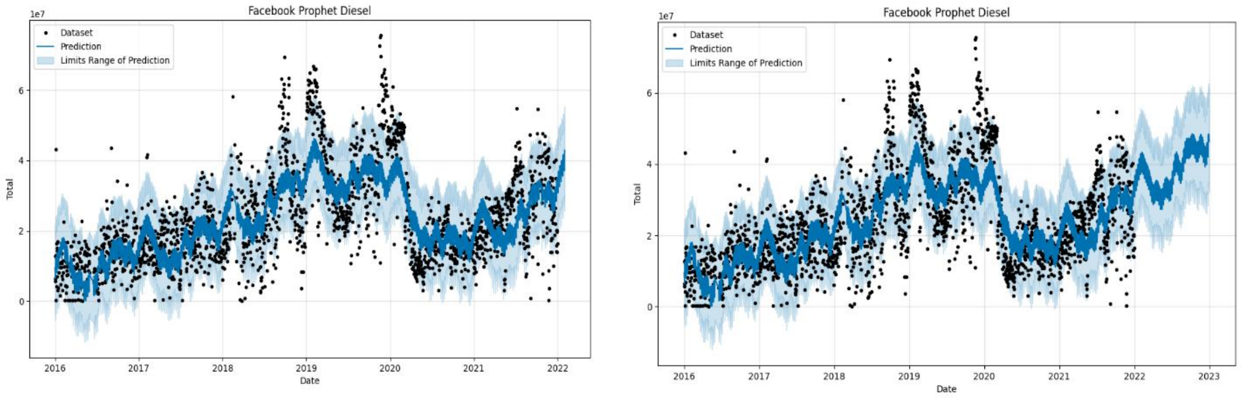
(a) 31 Days.

(b) 365 Days.

Figure 4. Output FB Prophet model of gasoline.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

Figure 5a presents the prediction for gasoline sales using FB Prophet over a 31 days period, showing a relatively good fit with actual data and fluctuations in gasoline demand effectively. However, Figure 5b displays the 365 days forecast, where the model captures long-term trends but struggles with daily fluctuations. while Figure 5b focuses more on long-term patterns, such as seasonal trends. However, accuracy decreases over the longer prediction horizon.



(b) 31 Days. **(a)** 365 Days.

Figure 5. Output FB Prophet model of diesel.

Since in this case FB Prophet produces a graph that does not change significantly between the horizons, the difference in the MAPE and RMSE percentage plays an important part in determining the best model. In addition, the prediction results on the 31, 62 ,123, and 365 days horizons provide insight into how the model can predict future data over various time frames. Although the time series graphs do not show significant changes, the differences in the MAPE and RMSE figures highlight the importance of understanding evaluation metrics to measure the performance and accuracy of time series analysis models with FB Prophet. In addition, the prediction results on these horizons provide an overview of the data quality and prediction precision of this model in different time frames. Figure 6a shows a 31 days diesel forecast using FB Prophet, which captures short-term trends reasonably well. In Figure 6b, The 365 days solar forecast focuses on long-term trends with improved daily prediction accuracy.

Lastly, for the third algorithm, i.e., Moving Average, we run predictions for a range of 31 and 62 periods for SMA and EMA respectively, due to the limitation of the MA prediction for a long period of time especially if there are significant changes in historical data. As shown in Figure 7 and Figure 8, we can see the extreme fluctuation of sales, which are extreme increase of sales for gasoline dataset and extreme drop of sales for diesel dataset. This method is suitable for stable data trends without significant changes over time. Predictive analysis patterns in this method are carried out by calculating the average of sales depending on the desired period, in this case, a range of 31 and 62 periods.

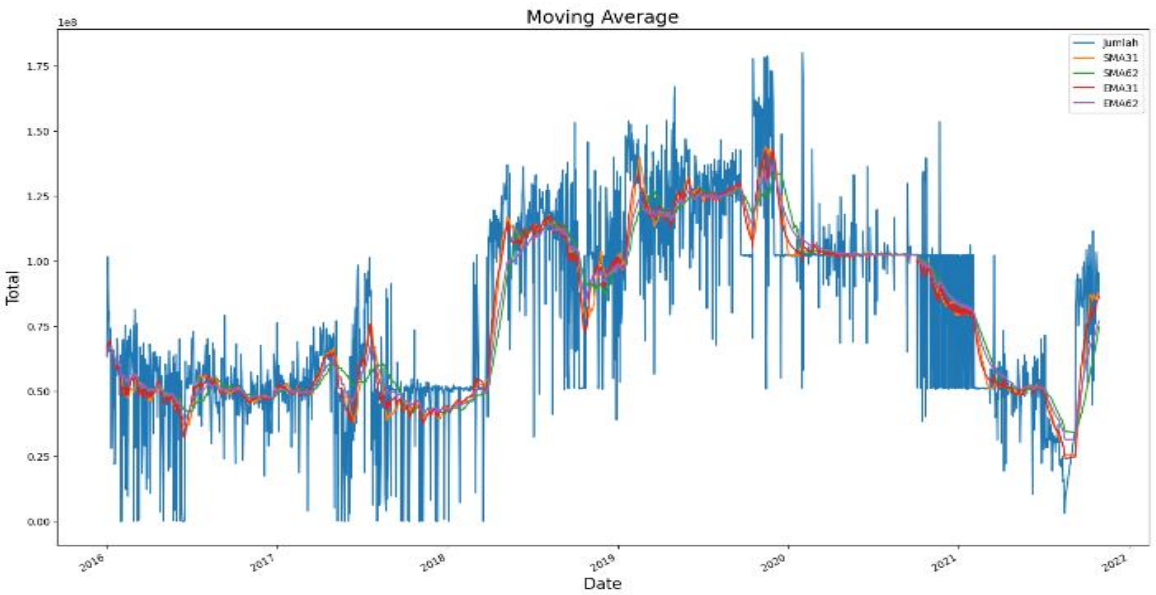


Figure 6. Output MA model for gasoline

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

As shown Figure 7, Moving Averages with 31 periods are more focused on short-term trends, offering faster adjustments and responsiveness to recent gasoline sales shifts. In contrast, 62 periods Moving Averages provide a more smoothed and stable view of the data, focusing on medium-term trends but with reduced sensitivity to immediate changes in gasoline demand.

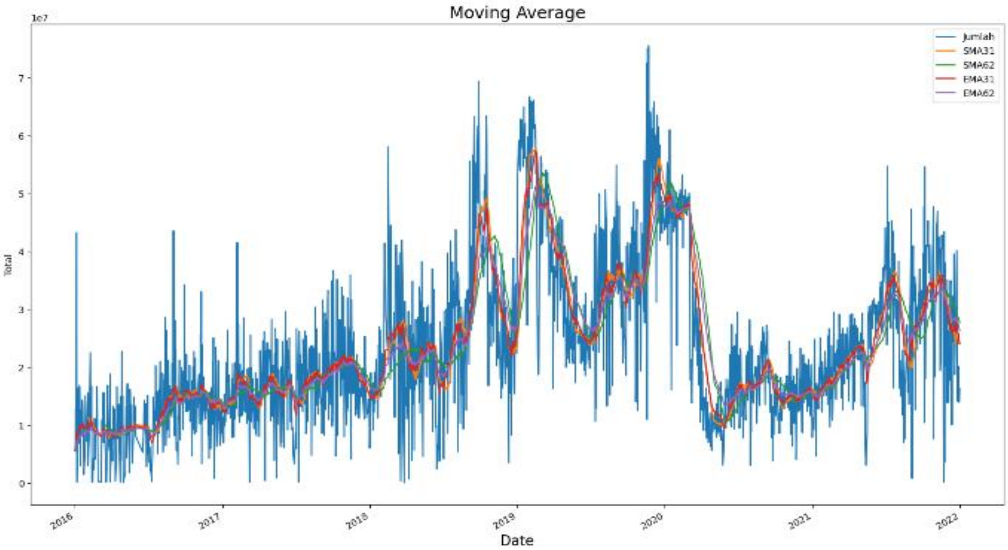


Figure 7. Output MA model for diesel.

Figure 8 presents the predictions for diesel sales using SMA and EMA over 31 and 62 periods. The 31 period SMA provides a stable view by equally weighing all data points, ideal for minimizing short-term fluctuations in diesel demand. However, the 62 period SMA further smooths the trend, making it even less reactive to sudden changes but better suited for identifying long-term patterns. The 31 period EMA is more responsive, giving greater weight to recent sales data and thus better for capturing short-term trends in the diesel market. In contrast, the 62 period EMA balances the responsiveness with longer-term data, making it a more flexible tool for medium-term predictions in diesel sales.

Table 5. Error rate in comparison of all models used.

Type of Fuels	Model	RMSE	MAPE %
Gasoline	LSTM 31 days	**51.3687**	**17.11**
	LSTM 62 days	78.0678	19.87
	LSTM 128 days	82.2219	20.76
	LSTM 365 days	98.7611	22.47
	FB Prophet 31 days	**62.1770**	**17.87**
	FB Prophet 62 days	65.7811	18.48
	FB Prophet 128 days	69.2734	19.43
	FB Prophet 365 days	109.3882	23.23
	SMA 31 periods	59.2701	17.93
	SMA 62 periods	68.8170	18.73
	EMA 31 periods	**58.3998**	**17.81**
	EMA 62 periods	69.8551	18.79
Diesel	LSTM 31 days	**52.8477**	**18.44**
	LSTM 62 days	81.3338	19.12
	LSTM 128 days	86.3677	20.81
	LSTM 365 days	107.2149	23.67
	FB Prophet 31 days	**65.9871**	**18.32**
	FB Prophet 62 days	70.6817	19.49
	FB Prophet 128 days	68.7771	19.01
	FB Prophet 365 days	107.4652	24.13
	SMA 31 periods	**59.6372**	**18.42**
	SMA 62 periods	70.8516	19.34
	EMA 31 periods	61.6436	18.78
	EMA 62 periods	69.8502	18.91

Table 5 shows a summary of evaluation results for both gasoline and diesel datasets using LSTM, FB Prophet, SMA, and EMA with RMSE and MAPE as the evaluation metrics. The LSTM model with 31 days prediction outperformed all other algorithms for gasoline dataset while the FB Prophet with 31 days prediction outperformed all other

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

algorithms for diesel. However, for 365 days prediction, LSTM outperformed FB Prophet on both gasoline and diesel datasets.

This paper presents a prediction analysis of gasoline and diesel fuel, by carrying out the stages of data processing and modeling using the LSTM, FB Prophet, Simple Moving Average and Exponential Moving Average methods for predictive analysis of future fuel use by looking at analysis indicators based on RMSE and MAPE from the comparison of these algorithms. From the results, it can be seen that fuel sales tend to have high fluctuations in a certain time depending on the problems that occur at that time. This experiment was carried out with different settings, where the LSTM uses a range of analyzes on predictions of 31, 62, 123, and 365 days horizon, while the FB Prophet uses a range of analyzes on predictions of 31, 62, 123, and 365 days horizon, and the Moving Average uses only 31 and 62 periods.

The experiment results show that the machine learning approach, i.e., LSTM and FB Prophet are more suitable in the time series prediction of the fuel price in Indonesia compared to the traditional statistical approach, i.e. the Moving Average. The two machine learning methods have their respective advantages and can predict with long-term resistance. In contrast to the Moving Average, which only requires low computation because it uses ordinary mathematical calculations, this means that MA is only able to predict with a short time window. In this sense, the MA method is not effective for long-term predictions. LSTM and FB Prophet, on the other hand, are able to model complex relationships between variables in the data and learn patterns to make more accurate predictions over a longer period of time. LSTM as a form of recurrent neural network, has the ability to remember long-term information and overcome the vanishing gradient problem that often occurs in traditional neural networks. Meanwhile, Facebook Prophet is a method developed by Facebook to predict trends and seasonality of time-related data. This method is specifically designed to be able to overcome some of the challenges commonly encountered in time series data analysis, such as the presence of missing data or outliers. By combining these two methods, analysis can produce more accurate and reliable predictions in the long term.

From the analysis that has been done, the LSTM model generally outperformed the rest of algorithms and can produce a smaller error rate in the predictions compared to the FB Prophet and the Moving Average, especially for longer prediction periods, since the LSTM model is a deep learning model that can be used to predict time series data by taking into account the relationship between data current and past data. Even though this LSTM model only has a slightly smaller error rate than other models, it does not rule out the possibility that this model will work optimally with a larger amount of data and a higher distance.

The predictions fuel of diesel using FB Prophet algorithm with 31, 62, 123, 365 days provide an overview of the direction of future trends and predict the level of sales that will occur in the market. Since the demand for diesel production is also very high, it is not surprising that the scarcity of this fuel affects sales. The results of this predictive analysis can also help companies that want to make decisions about taking a high amount of fuel and can also be a reference for planning a more effective business.

4. CONCLUSION

This research aims to understand the current and future sales patterns and trends of fuel sales in Indonesia by utilizing four time series analysis models, i.e., LSTM, FB Prophet, SMA, and EMA. By comparing the results using statistics such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), it is shown that the machine learning approach, i.e., LSTM and FB Prophet are more suitable in the time series prediction of the fuel price in Indonesia compared to the traditional statistical approach, i.e. the Moving Average. The two machine learning methods have their respective advantages and can predict with long-term resistance. The LSTM model is useful for predicting data that has long-term dependencies, while the Moving Average and FB Prophet can also be an alternative for making accurate predictions for the short term. The LSTM model achieved the lowest MAPE for gasoline at 17.11% with a 31-day prediction, categorized as fairly accurate. For diesel, FB Prophet had the smallest MAPE at 18.32% for the same period. These findings suggest that LSTM and FB Prophet can aid in decision-making by reducing uncertainty in fuel stock management. However, these models demand robust computational systems, especially LSTM. In contrast, the MA model requires minimal computational power due to its simplicity. Each method has its strengths: LSTM is ideal for long-term data dependencies, while Moving Average and FB Prophet excel in short-term predictions.

REFERENCES

[1] S. Ou *et al.*, “Machine learning model to project the impact of COVID-19 on US motor gasoline demand,” *Nat Energy*, vol. 5, no. 8, pp. 666–673, 2020, doi: 10.1038/s41560-020-0662-1.
[2] C. Y. Lam, J. N. Yam, and H. W. Lau, “Time series forecasting with machine learning for fuel demand prediction,” *Energy Policy*, vol. 144, p. 111980, 2022.
[3] B. Zhang, Y. Wang, and X. Liu, “Hybrid LSTM-ARIMA model for oil price prediction,” *Energy Econ*, vol. 88, p. 104726, 2021.
[4] P. S. Rao and M. H. Hariri, “Improving forecasting accuracy using hybrid deep learning models for fuel sales prediction,” *Journal of Energy Markets*, vol. 12, pp. 72–83, 2020.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 3, Desember 2024 Page: 1321−1331
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i3.5877

[5] A. Gupta and P. Kumar, "Fuel demand forecasting using machine learning algorithms: A case study of India," *Energy Reports*, vol. 6, pp. 1127–1136, 2020.
[6] J. Wang and Z. Li, "Forecasting future fuel sales trends: A comparison of LSTM and Prophet," *Journal of Applied Energy*, vol. 283, p. 116287, 2021.
[7] S. Karasu, A. Altan, S. Bekiros, and W. Ahmad, "A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series," *Energy*, vol. 212, p. 118750, 2020.
[8] J. Li, W. Liu, and H. Zhang, "Uncertainty in fuel demand forecasting and its implications for energy policy," *Energy Policy*, vol. 135, p. 111939, 2020.
[9] X. Liu, Y. Wang, and H. Zhang, "A hybrid ARIMA and deep learning model for fuel sales prediction in volatile markets," *Journal of Energy Markets*, vol. 13, no. 2, pp. 45–62, 2020.
[10] A. Gunawan and P. Hartono, "Machine Learning Approaches for Fuel Demand Prediction in the Transportation Sector," *Journal of Energy Markets*, vol. 13, pp. 92–108, 2021.
[11] J. Wang, X. Zhang, and Y. Li, "Attention-based LSTM for time series forecasting in energy markets," *Energy Econ*, vol. 86, p. 104645, 2020.
[12] S. Adhikary and K. Roy, "Predictive modeling of fuel demand in volatile markets using random forests," *Appl Energy*, vol. 290, p. 116529, 2021.
[13] E. Zunic, K. Korjenic, K. Hodzic, and D. Donko, "Application of facebook's prophet algorithm for successful sales forecasting based on real-world data," *arXiv preprint arXiv:2005.07575*, 2020.
[14] C. Guo, Q. Ge, H. Jiang, G. Yao, and Q. Hua, "Maximum power demand prediction using fbprophet with adaptive Kalman filtering," *IEEE Access*, vol. 8, pp. 19236–19247, 2020.
[15] B. Lakshmanan, P. S. N. Vivek Raja, and V. Kalathiappan, "Sales Demand Forecasting Using LSTM Network," in *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, Springer, 2020, pp. 125–132. doi: 10.1007/978-981-15-0199-9_11.
[16] E. A. Kadir *et al.*, "Fires Hotspot Forecasting in Indonesia Using Long Short-Term Memory Algorithm and MODIS Datasets," in *Vegetation Fires and Pollution in Asia*, Springer, 2023, pp. 589–602.
[17] E. houssin Ouassou and H. Taya, "Forecasting regional tourism demand in morocco from traditional and AI-based methods to ensemble modeling," *Forecasting*, vol. 4, no. 2, pp. 420–437, 2022.
[18] A. Boudhaouia and P. Wira, "A real-time data analysis platform for short-term water consumption forecasting with machine learning," *Forecasting*, vol. 3, no. 4, pp. 682–694, 2021.
[19] M. Bildirici, N. Guler Bayazit, and Y. Ucan, "Analyzing crude oil prices under the impact of COVID-19 by using LSTARGARCHLSTM," *Energies (Basel)*, vol. 13, no. 11, p. 2980, 2020.
[20] J. Peng, Z. Li, and B. M. Drakeford, "Dynamic characteristics of crude oil price fluctuation—from the perspective of crude oil price influence mechanism," *Energies (Basel)*, vol. 13, no. 17, p. 4465, 2020.
[21] D. Swain, P. K. Pattnaik, and P. K. Gupta, *Machine Learning and Information Processing*. Springer, 2021.
[22] B. K. Jha and S. Pande, "Time series forecasting model for supermarket sales using FB-prophet," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 2021, pp. 547–554.