Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Corn sales analysis using linear regression and svm methods to improve production planning

Ahmad Hakiki Saputra[1], Dadang Priyanto[2], Rifqi Hammad[3]
[1,2,3]Department of Information Technology, Faculty of Engineering, Universitas Bumigora, Indonesia

Article Info

Article history:

Received June 28, 2025
Revised August 3, 2025
Accepted August 3, 2025

Keywords:

Corn sales prediction
Data preprocessing
Forecasting
Linear regression
Support vector machine

ABSTRACT

This research aimed to analyze and predict corn sales at UD Muara Kasih to improve production planning accuracy. The study used historical corn sales data collected over a specific period, covering 42 data entries from January 2021 to December 2024. The dataset included variables such as sales date, quantity sold, selling price per ton, total sales value, weather conditions, market demand (in tons), and the number of laborers. Prior to model training, the data underwent comprehensive preprocessing involving data cleaning, feature extraction, and normalization to ensure its quality and readiness for analysis. Two predictive models were applied: Linear Regression and Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel. Simulation data for 2024 and 2025 were generated based on the monthly averages derived from the historical dataset. The results showed that the Linear Regression model produced more stable predictions with a lower Root Mean Squared Error (RMSE) of 255.84 compared to the SVM model's RMSE of 256.42. While the SVM model showed greater responsiveness to seasonal variations, the Linear Regression model was identified as the most suitable for the given dataset. The predictive models developed in this study are expected to support UD Muara Kasih in making more accurate and informed production decisions in the future.

Corresponding Author:

Ahmad Hakiki Saputra
Department of Information Technology,
Faculty of Engineering, Universitas Bumigora, Indonesia
Email: grasmahakiki24@gmail.com
https://doi.org/10.52465/joscex.v6i3.591

1. INTRODUCTION

The availability of increasingly abundant historical data on agricultural sales and production has not been fully utilized optimally by business actors and farmers to support data-based decision making [1]. Many business actors and farmers still rely on intuition or subjective experience in determining production volume and sales strategies. This is a serious problem, especially when market demand is fluctuating and influenced by various external factors such as weather, price fluctuations, and government policies that can pose serious risks to business actors and farmers [2]. One of them is the increasingly abundant corn production data that has not been utilized properly in times of fluctuating market demand, especially in Indonesia.

Referring to a press release issued by the Coordinating Ministry for Economic Affairs of the Republic of Indonesia in August 2022, world corn prices improved in January–June 2022, increasing by 21.53% compared to the same period in 2021 [3]. This is an opportunity for Indonesia to export corn. In addition to meeting export needs, the government also continues to emphasize increasing corn development for national needs. Data on corn growth has continued to increase in recent years in various corn-producing regions in Indonesia. One of the areas that participated in the euphoria of increasing corn harvests is East Lombok Regency. Based on data released by the East Lombok Central Statistics Agency, the average corn harvest in East Lombok Regency is 2,053 tons. However, the abundance of the harvest often causes the selling price to decrease [4].

Corn is one of the strategic commodities in Indonesia because it plays a major role in meeting national food needs and the animal feed industry. As demand for corn increases, effective production planning becomes increasingly important so that production can adjust to fluctuating market needs [5]. However, this planning is often faced with the challenge of uncertainty, such as weather changes, price fluctuations, and government policies that have a direct impact on the availability and price of corn. One real example occurs at UD Muara Kasih, a corn processing and sales business unit located in Lenek Lauq Village, Aikmel District, East Lombok Regency, West Nusa Tenggara. Although this business unit has historical sales data, the data has not been processed into predictive information to assist production planning. As a result, an imbalance between supply and demand often occurs, which results in falling prices when the harvest is abundant [4]. Therefore, accurate sales analysis and corn production predictions are needed to anticipate these changes and support the sustainability of adequate supply [6].

Several previous studies have tried to apply data analysis-based prediction methods to agricultural commodities. Maulana and Danar Dana in 2024 used Linear Regression to predict corn prices in West Java and achieved an accuracy level of 99.10% based on MAE, MSE, and RMSE values. However, this approach has limitations in capturing complex data patterns [7]. In contrast, Saadah et al in 2021 applied Support Vector Machine (SVM) to predict agricultural commodity prices and successfully overcame the problem of non-linear data, although it required a more complex parameter tuning process [8]. Meanwhile, Rozaini and Silaban in 2023 developed Multiple Linear Regression to analyze the effect of weather factors and input prices on red chili production, but this model was less accurate when faced with non-linear patterned data [9].

Based on this background, this research aims to implement two analytical methods, namely Linear Regression and Support Vector Machine (SVM), in analyzing and predicting corn sales at UD Muara Kasih, East Lombok. This research will compare the performance of the two methods based on the level of prediction accuracy, with the aim of providing recommendations for the best approach to support a more effective corn production planning strategy. The outputs of this research include a corn sales prediction model, comparative analysis of model performance, and a data mining-based production planning strategy that can be applied to the local agricultural sector.

2. METHOD

In this research, the analysis process is carried out through several stages that have been systematically designed to achieve the research objectives. The series of stages can be seen more clearly in Figure 1 below, which presents the overall research workflow.

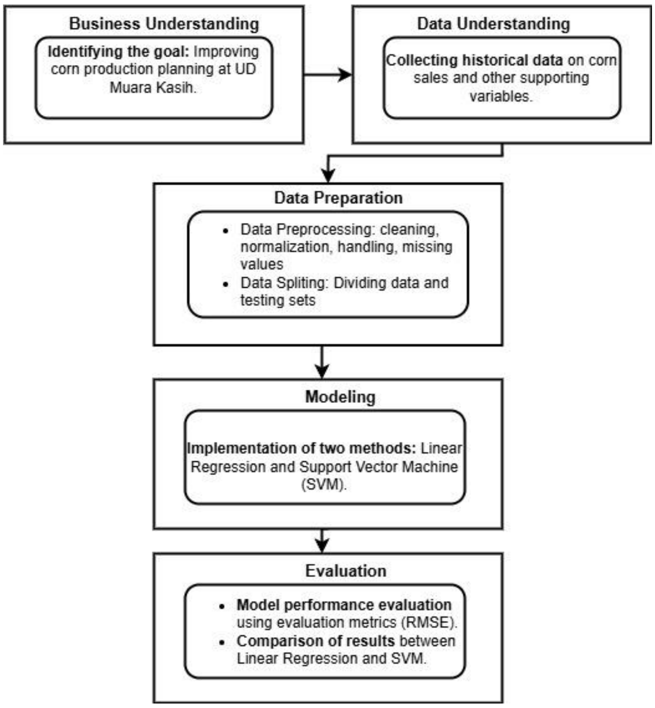

Figure 1. Research workflow

General Overview of the Research

This research aims to compare the Linear Regression and Support Vector Machine (SVM) methods for analyzing and predicting corn sales to improve production planning. The research was conducted using historical corn sales data from UD Muara Kasih, which included variables such as price, sales volume, planting season, and weather. The analysis process was carried out through systematic stages to build an accurate and reliable prediction model for production decision-making. This approach helps to better understand historical corn sales patterns, enabling the identification of factors that influence sales performance, as well as more accurately estimating future demand.

Business Understanding

The primary goal of this stage is to thoroughly understand the business needs of UD Muara Kasih, which has been experiencing difficulties in planning corn production due to the uncertainty caused by fluctuating demand and sales. These unpredictable changes pose the risk of producing either too much or too little, both of which can lead to inefficiencies and reduced profitability. To address this issue, a corn sales forecasting model is being developed to support the company in making production decisions that are more timely, accurate, and aligned with market conditions. By leveraging reliable predictions, UD Muara Kasih can better adjust its production levels to match demand trends, minimize potential losses, and improve overall business performance [10].

Data Understanding

The data used in this research comes from historical corn sales data at UD Muara Kasih, located in Lenek Lauq Village, Aikmel District, East Lombok Regency. Data were obtained through the company's internal records, covering the period from April to November 2023. Data was collected in tabular format and reflects operational activities as well as variables considered to influence corn sales [11]. The description of the data is explained in the following Table 1.

Table 1. Data description of corn sales at UD muara kasih

No	Variable Name	Data Type	Unit / Format	Description
1	sale_date	Date	yyyy-mm-dd	The date of the corn sales transaction
2	sales_volume	Numeric	Rupiah (IDR)	Total income from corn sales on the corresponding date
3	corn_sold	Numeric	Ton	Total amount of corn sold (in metric tons)
4	weather	Categorical	Sunny, Rainy, Cloudy	Weather condition during or before the transaction
5	market_demand	Numeric	Rupiah (IDR)	Market demand value for corn on that date (assumed to be equal to sales)
6	laborers	Numeric	Persons	Number of workers involved in sales or production process
7	sales_class	Categorical	Low, High	Classification of sales performance based on volume or threshold criteria

This data serves as the input for the predictive analysis and model training process. The goal is to assist UD Muara Kasih in creating more accurate production plans based on previous sales patterns and external factors like weather. samples of the dataset used in this research is presented in Table 2.

Table 2. Samples of the dataset

Date	Sales Volume (IDR)	Corn Sold (Ton)	Weather	Market Demand (Ton)	Laborers (People)	Sales Class
2021-01-01	378,078	4.26	Sunny	378,078	16	Low
2021-02-01	380,976	5.09	Rainy	380,976	12	Low
2021-03-01	380,711	4.08	Sunny	380,711	18	High
2021-04-01	376,204	5.29	Rainy	376,204	6	Low
2021-05-01	380,563	4.09	Sunny	380,563	15	High
2021-06-01	376,087	5.49	Sunny	376,087	16	High
2021-07-01	377,202	5.42	Cloudy	377,202	5	Low
2021-08-01	380,234	4.07	Sunny	380,234	10	High
2021-09-01	377,999	5.05	Rainy	377,999	9	High
2021-10-01	380,178	5.79	Cloudy	380,178	16	High
...	...	...	...	...	...	...
2024-12-28	378,658	5.50	Rainy	378,658	10	High

Data Preparation

This stage includes the data preprocessing process to ensure that the data is in the appropriate format for modeling purposes [12]. Several important steps are taken, including cleaning duplicate data and handling missing values, transforming numerical and categorical data, normalizing data if necessary, dividing the dataset into training and test data, and selecting features using the Forward Feature Selection method. In addition, testing is also carried out on various historical periods, ranging from one to four months prior, to analyze their impact on model performance.

Data preprocessing is crucial to ensure that the dataset used is clean, consistent, and ready for modeling [13]. Some key steps in this process include data cleaning, which involves removing duplicate entries and handling missing data using methods such as mean imputation, interpolation, or row deletion. Additionally, date formats and numerical units such as kilograms or currency are standardized. Data transformation is performed by converting categorical variables, such as weather conditions, into numerical values using label encoding or one-hot encoding. To improve the performance of the SVM model, numerical features are normalized using Min-Max Scaling. In the time feature engineering stage, additional features are derived from the sales date, such as day of the week, month, and harvest season. Finally, daily weather data from external sources is combined with sales data to form a more comprehensive and informative dataset.

Modeling

At this stage, two modeling approaches were developed to predict corn sales, namely Linear Regression (LR) and Support Vector Machine (SVM) [14]. These two methods were chosen to provide a comparison between simple linear models and nonlinear machine learning-based models, in order to produce accurate predictions that can be applied to the production planning of UD Muara Kasih.

Linear Regression

In Linear Regression, the process begins with manual calculations to understand the basic concepts of prediction and validate the modeling process. In this approach, the independent variable used is the selling price per kilogram (X), while the dependent variable is the amount of corn sold (Y). The regression parameters are calculated manually, namely the slope coefficient (b) and the intercept (a), using the formula:

$$b = \frac{n\sum XY - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2} \tag{1}$$

$$a = \frac{n\sum Y - b(\sum X)}{n} \tag{2}$$

.

Then, a simple linear regression equation was formed in the form of $Y = a + bX$ (3) [15]. Sales predictions were made based on this equation, and the results were compared with actual data. As an evaluation measure, Root Mean Absolute Error (RMAE) was used to assess the performance of manual predictions. The results of this manual approach were also used as an initial benchmark to evaluate the linear regression model built automatically using Python, to ensure consistency of logic and accuracy of results [16].

Support Vector Machine (SVM)

The Support Vector Machine (SVM) model was developed using the Python programming language and the Scikit-learn library [17]. The modeling process was carried out in several stages to ensure optimal prediction accuracy. The first stage was feature selection, in which relevant variables such as selling price, production cost, day of sale, weather conditions, and lag features were selected using the Forward Feature Selection method [18]. Feature selection is important so that the model only uses variables that truly contribute to the output, thereby reducing the risk of overfitting and improving interpretability.

Next, the data is divided into two parts, namely 80% training data and 20% test data. Additionally, various historical period lengths, such as one to four months prior, were tested to evaluate their impact on model performance [11]. To handle nonlinear relationships between variables, the Radial Basis Function (RBF) was used as the kernel function. The RBF kernel function is mathematically defined as [19]:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \tag{3}$$

where γ is the kernel parameter that controls the extent to which each training data point influences the prediction. In this stage, several model parameters such as C (regularization parameter), epsilon (error tolerance margin), and gamma are optimized using the grid search method to obtain the best parameter combination [18]. The model is then trained using the processed training data. The prediction function in Support Vector Regression (SVR) can be written as [20]:

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b \tag{4}$$

where α_i and α_i^* are Lagrange multipliers obtained during the training process, $K(x_i, x_j)$ is the kernel function (in this case RBF), and b is the bias or intercept. This function allows the model to predict the output with a maximum deviation of ∈ from the actual value, while keeping the function as simple as possible.

The model was evaluated using the Root Mean Square Error (RMSE) metric, which calculates the average prediction error in its original units. RMSE was chosen for its ability to detect large prediction errors, thereby providing a more realistic picture of the model's performance [21]. The evaluation process was carried out by comparing the prediction results with the actual data and observing the performance of various kernel configurations and historical data lengths used. The model with the lowest RMSE value was selected as the best model, also considering aspects of interpretability and ease of implementation in the UD Muara Kasih production planning system [22]. This approach is expected to produce more accurate, data-driven decision support that is responsive to changing market conditions.

Evaluation

Model evaluation was conducted to assess the accuracy and reliability of the prediction results produced by each method, namely Linear Regression and Support Vector Machine (SVM). At this stage, two main metrics were used, namely Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). RMSE measures the magnitude of prediction errors by giving greater weight to extreme errors, making it sensitive to outliers. The RMSE formula is written as follows [23]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{5}$$

Meanwhile, MAE measures the absolute average of the difference between the actual value and the predicted value without taking into account the direction of the error, using the following formula, where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of test data points [23]:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{6}$$

The evaluation was performed on both models in various training scenarios, based on variations in the length of the historical period. The test results show that the Linear Regression model provides stable and easy-to-interpret results, but is less capable of capturing nonlinear patterns that may appear in sales data. Conversely, the SVM model with a Radial Basis Function (RBF) kernel is better able to adapt to fluctuating data trends, although it requires longer computation time and tends to be more sensitive to parameter settings. The best model was selected based on the lowest RMSE and MAE values, as well as its performance stability across various data configurations. This evaluation stage is crucial to ensure that the developed model is not only statistically superior but also consistently and reliably applicable in supporting operational decision-making at UD Muara Kasih.

3. RESULTS AND DISCUSSIONS

The data used in this study consists of historical corn sales records from UD Muara Kasih over a specified period. The dataset includes variables such as the sales date, quantity of corn sold, selling price per ton, total sales revenue, weather conditions, market demand (in tons), and the number of laborers involved. A total of 42 data entries were collected, covering the period from January 2021 to December 2024. Before being used in the model training process, the data will undergo a thorough preprocessing phase to ensure its quality and readiness for analysis.

Data Preprocessing Results

Preprocessing is a crucial step to ensure that the dataset is clean, consistent, and suitable for model development. The dataset used in this study consists of historical corn sales records ranging from January 2021 to December 2024. The main attributes in this dataset include the sales date, quantity of corn sold (in tons), sale price per kilogram (in IDR), weather conditions, market demand (in tons), and the number of laborers involved. The preprocessing process focused on cleaning, transforming, and restructuring the data to improve its quality and make it ready for modeling. The preprocessing procedure was carried out in four main steps. First, the date format was converted into the standardized YYYY-MM-DD format to unify the temporal structure of the data. Second, time-based features such as month, year, and month name were extracted in order to identify seasonal patterns that may influence the volume of corn sales. Table 4.2 presents the results of the date conversion and time feature extraction that were successfully performed.

Date Formatting and Time Feature Extraction

After converting the sales date into a standardized format, several temporal features were extracted, including the combined Month-Year format (e.g., Jan 2021) and Month Name (both in English and Indonesian). These derived features are valuable for exploring potential seasonal patterns in corn sales, identifying temporal trends, and enhancing feature selection in time-sensitive models. Table 3 presents the dataset after date conversion and time feature extraction have been performed.

Table 3. Results of date conversion and time feature extraction

Sales Date	Corn Sold (Ton)	Sale Price (IDR)	Weather	Market Demand (Ton)	Laborers	Month-Year	Month Name (EN)	Month Name (ID)
2021-01-01	4.80	379,000	Sunny	379,000	13	Jan 2021	January	Januari
2021-02-01	4.90	379,500	Rainy	379,500	12	Feb 2021	February	Februari
2021-03-01	4.50	380,000	Sunny	380,000	15	Mar 2021	March	Maret
2021-04-01	5.00	377,000	Rainy	377,000	10	Apr 2021	April	April
...	...	...	...	...	...	...	...	...
2024-12-28	5.10	378,500	Rainy	378,500	12	Dec 2024	December	Desember

This step provided meaningful temporal context that can be leveraged to detect monthly or annual fluctuations in sales volume. Such patterns are particularly relevant in agricultural domains where climate and harvest cycles impact production and sales trends.

.

Normalization of Numerical Features

The next phase of preprocessing involved scaling the numerical data to ensure uniformity in range and improve model efficiency. Raw values in variables such as corn sold, sale price, and market demand varied significantly in scale, potentially biasing machine learning algorithms that rely on distance-based calculations. To address this, the Min-Max Scaling method was applied, which rescales the features into the [0, 1] interval based on the formula:

$$X_{scaled} = \frac{X - X_{min}}{X_{min} - X_{max}} \quad (7)$$

Table 4 shows the normalized values for selected numerical features over various months:

Table 4. Normalized numerical data using min-max scaling

Month-Year	Corn Sold	Sale Price	Market Demand
Jan 2021	0.6667	0.3333	0.6667
Feb 2021	0.8333	0.4444	0.8333
Mar 2021	1.0000	0.0000	1.0000
Apr 2021	0.0000	0.5556	0.0000
...	...	...	...
Dec 2024	0.5000	0.6667	0.5000

This scaling ensures that all numerical attributes contribute equally to the training process. By eliminating differences in scale, it prevents dominant variables from disproportionately influencing the learning algorithm, which is particularly vital for Support Vector Machine (SVM) models.

Final Preprocessed Dataset

After completing normalization, cleaning, and feature engineering, the final preprocessed dataset was constructed. It contains both transformed and original features that are ready to be used for model training. Table 5 provides a snapshot of the final dataset after preprocessing.

Table 5. Final dataset after preprocessing

Sales Date	Corn Sold	Sale Price	Market Demand	Weather
2021-01-01	0.425	0.298	0.425	Sunny
2021-02-01	0.862	0.703	0.862	Rainy
2021-03-01	0.823	0.261	0.823	Sunny
2021-04-01	0.000	0.778	0.000	Rainy
...	...	...	...	...
2024-12-28	0.494	0.872	0.494	Rainy

This structured dataset includes normalized numerical values, encoded categorical attributes, and extracted temporal features, enabling comprehensive model training and predictive analytics. Additionally, the sales class variable used as the target in classification tasks has also been prepared in numerical format to support algorithm compatibility. The preprocessing phase, as described, ensures that the dataset adheres to the necessary standards for machine learning applications. By standardizing the input data and enriching it with meaningful features, the modeling process is expected to achieve higher accuracy, better generalization, and more reliable forecasting results for corn sales in subsequent years (2025 and 2026).

Prediction Results Using Linear Regression

The Linear Regression model was developed using the normalized dataset and trained on historical sales data. It was then used to forecast monthly corn sales for the period of 2025–2026. The prediction results, as shown in Figure 2, indicate a relatively stable sales pattern throughout the two-year period, with subtle fluctuations across the months.

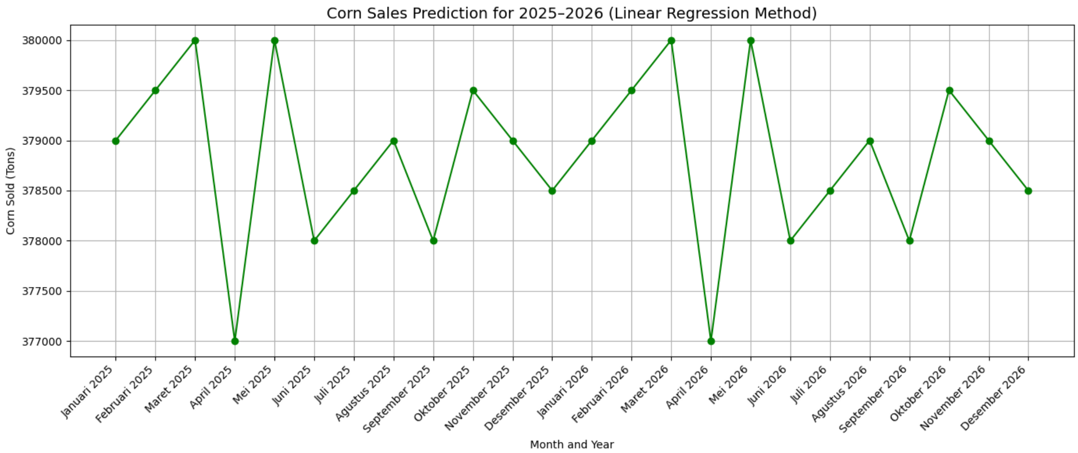

Figure 2. Linear regression sales predictions

Figure 2 illustrates the visualization of the predicted corn sales trends over the next two years. The forecasted pattern reveals recurring sales peaks in March, May, and October, while the lowest sales consistently occur in April. This indicates that the linear regression model assumes a regular trend, which may reflect a consistent production and distribution cycle but lacks sensitivity to more complex seasonal variations.

Prediction Results Using Support Vector Machine (SVM)

A Support Vector Machine (SVM) model with a Radial Basis Function (RBF) kernel was also applied to the same dataset to capture more complex, non-linear patterns in the sales data. As shown in Figure 3, this model yields more varied predictions compared to the linear approach.

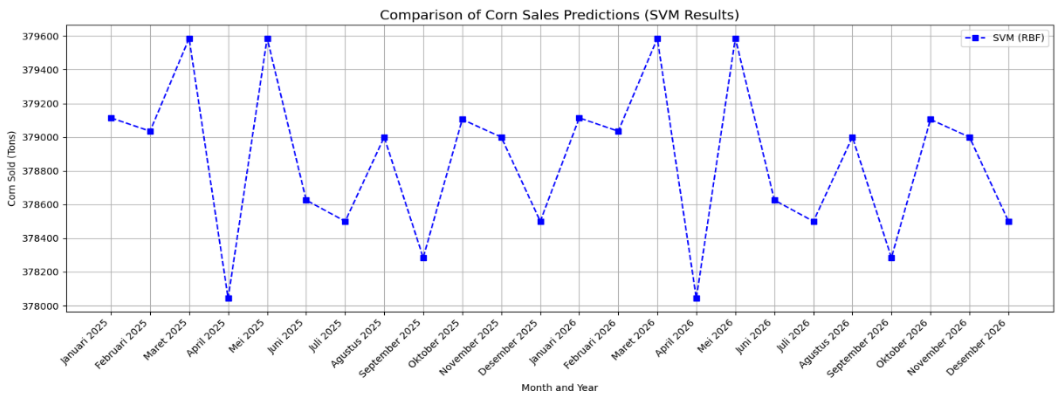

Figure 3. SVM sales predictions

The SVM model highlights sharper month-to-month changes. Peak sales are observed in March and May, while the lowest point is projected in April for both 2025 and 2026. This fluctuation pattern may better represent real-world dynamics such as seasonal harvest variations or sudden market demand shifts.

Model Comparison and Evaluation

Model performance was evaluated using the Root Mean Squared Error (RMSE) metric. The comparison results are presented in Table 6**.**

Tabel 6. Model comparison and evaluation

Model	RMSE
Linear Regression	255.84
SVM (RBF Kernel)	256.42

.

The evaluation shows that the Linear Regression model achieved a slightly lower RMSE compared to the SVM model, indicating better predictive accuracy in this case. However, the difference was marginal (0.58), suggesting that both models are capable of providing reliable predictions. The simplicity and interpretability of Linear Regression make it a practical choice, but SVM remains valuable when dealing with more complex data patterns.

The findings of this research indicate that both Linear Regression and Support Vector Machine (SVM) can be used effectively to predict corn sales at UD Muara Kasih. The Linear Regression model provides consistent and easy-to-understand results, while SVM with Radial Basis Function (RBF) kernel shows a more adaptive response to seasonal variations and market dynamics. The seasonal patterns captured by SVM align with agricultural production cycles, where sales are typically low at the beginning of the year and increase toward the end of the year, likely due to harvest periods or high market demand.

Evaluation results using the Root Mean Square Error (RMSE) metric show that the Linear Regression model achieved an RMSE value of 255.84, slightly lower than the SVM model, which recorded an RMSE of 256.42. This difference indicates that while Linear Regression is slightly more accurate in the context of the available data, SVM remains a promising alternative, especially if the data used becomes more complex or large in size. As a follow-up, future research could consider applying more complex time series-based predictive models, such as ARIMA or Long Short-Term Memory (LSTM), to enhance accuracy and capture seasonal variations and market trend changes more sharply.

4. CONCLUSION

This research successfully developed and implemented corn sales prediction models for UD Muara Kasih using two approaches: Linear Regression and Support Vector Machine (SVM). The results demonstrate that both models can effectively predict monthly sales trends based on historical data and influencing variables such as selling price, production costs, market demand, labor, and weather conditions.

The Linear Regression model provided stable and easy-to-interpret predictions with a lower Root Mean Squared Error (RMSE) value of 255.84, indicating slightly higher accuracy compared to the SVM model. On the other hand, the SVM model using the Radial Basis Function (RBF) kernel produced more dynamic predictions, better capturing seasonal fluctuations and market variations, even though its RMSE value was slightly higher at 256.42.

The results align with the objectives stated in the introduction, where accurate prediction models are expected to support production planning and minimize risks due to supply-demand imbalances. The developed models are highly applicable for supporting decision-making in the production and inventory management processes at UD Muara Kasih.

For future research, the application of more advanced time series models such as ARIMA or Long Short-Term Memory (LSTM) is highly recommended to further improve prediction accuracy, especially in handling complex seasonal patterns and long-term sales trends. Additionally, integrating external factors such as market competition, commodity prices, and broader economic indicators could enhance the robustness of the prediction system.

The prediction system developed in this research has practical potential to be implemented as a decision support tool for agricultural product sales management, not only for UD Muara Kasih but also for similar agribusinesses.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to UD Muara Kasih for providing the necessary data and support for this research. Special thanks are also extended to our academic supervisors for their valuable guidance and feedback throughout the research.

REFERENCES

[1] N. Chergui and M. T. Kechadi, “Data analytics for crop management: a big data view,” *J. Big Data*, vol. 9, no. 1, p. 123, Dec. 2022, doi: 10.1186/s40537-022-00668-2.

[2] K. R. Mukhamedova, N. P. Cherepkova, A. V Korotkov, Z. B. Dagasheva, and M. Tvaronavičienė, “Digitalisation of Agricultural Production for Precision Farming: A Case Study,” *Sustainability*, vol. 14, no. 22, p. 14802, Nov. 2022, doi: 10.3390/su142214802.

[3] E. B. Santoso, E. W. Budiman, and P. Puspitorini, “Analisis Pendapatan Usahatani Jagung Kelompok Tani Rahayu II Mojoroto Kota Kediri,” *Grafting J. Ilm. Ilmu Pertan.*, vol. 12, no. 2, pp. 52–60, Sep. 2022, doi: 10.35457/grafting.v12i2.2568.

[4] A. Bayhaqi, “Implementasi Metode Single Exponential Smoothing dalam Sistem Peramalan Hasil Panen Jagung,” Universitas Nahdlatul Ulama Sunan Diri, Galang Tanjung, 2023.

[5] F. Melia, F. M. Aldian, M. S. F. Pahlevi, R. N. I. Risqullah, and S. Oktaffiani, “Peran Pemerintah dalam Meningkatkan Volume Ekspor Jagung,” *J. Econ.*, vol. 2, no. 1, pp. 1305–1320, Jan. 2023, doi: 10.55681/economina.v2i1.287.

[6] A. P. W. Suci, “Prediksi Nilai Ekspor Komoditas Perkebunan di Pasar Internasional dengan Algoritma Linear Regression,” Universitas Nusa Putra, Sukabumi, 2024.

[7] G. Maulana and R. D. Dana, “Prediksi Hasil Produksi Jagung di Jawa Barat dengan Metode Algoritma Regresi Linear Menggunakan Google Colab,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 827–837, Mar. 2024, doi: 10.36040/jati.v8i1.8816.

[8] S. Saadah, F. Z. Z, and H. H. Z, “Support Vector Regression (SVR) Dalam Memprediksi Harga Minyak Kelapa Sawit di Indonesia dan Nilai Tukar Mata Uang EUR/USD,” *J. Comput. Sci. Informatics Eng.*, vol. 5, no. 1, pp. 85–92, Jun. 2021, doi: 10.29303/jcosine.v5i1.403.

[9] N. Rozaini and S. J. Silaban, “Pengaruh Biaya Produksi Dan Harga Jual Terhadap Pendapatan Petani Cabai Merah Di Kecamatan Doloksanggul Kabupaten Humbang Hasundutan,” *J. Publ. Sist. Inf. dan Manaj. Bisnis*, vol. 2, no. 2, pp. 128–141, Apr. 2023, doi: 10.55606/jupsim.v2i2.1314.

[10] Y. A. Singgalen, “Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT),” *J. Media Inform. Budidarma*, vol. 7, no. 3, p. 1551, Jul. 2023, doi: 10.30865/mib.v7i3.6461.

[11] A. Angdresey, L. Sitanayah, and I. L. H. Tangka, “Sentiment Analysis for Political Debates on YouTube Comments using BERT Labeling, Random Oversampling, and Multinomial Naïve Bayes,” *J. Comput. Theor. Appl.*, vol. 2, no. 3, pp. 342–354, Jan. 2025, doi: 10.62411/jcta.11668.

[12] S. Alden and B. N. Sari, “Implementasi Algoritma CNN Untuk Pemilahan Jenis Sampah Berbasis Android Dengan Metode CRISP-DM,” *J. Inform.*, vol. 10, no. 1, pp. 62–71, Mar. 2023, doi: 10.31294/inf.v10i1.14985.

[13] I. Pii, N. Suarna, and N. Rahaningsih, “Penerapan Data Mining pada Penjualan Produk Pakaian Dameyra Fashion Menggunakan Metode K-Means Clustering,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 423–430, Mar. 2023, doi: 10.36040/jati.v7i1.6336.

[14] Y. Anggraini, D. Pasha, Damayanti, and A. Setiawan, “Sistem Informasi Penjualan Sepeda Berbasis Web Menggunakan Framework CodeIgniter (Studi Kasus: Orbit Station),” *J. Teknol. dan Sist. Inf.*, vol. 1, no. 2, pp. 64–70, 2020.

[15] A. Nazir and others, “Estimation and Forecasting of Rice Yield Using Phenology-Based Algorithm and Linear Regression Model on Sentinel-II Satellite Data,” *Agriculture*, vol. 11, no. 10, p. 1026, Oct. 2021, doi: 10.3390/agriculture11101026.

[16] C. Catal, K. Ece, B. Arslan, and A. Akbulut, “Benchmarking of Regression Algorithms and Time Series Analysis Techniques for Sales Forecasting,” *Balk. J. Electr. Comput. Eng.*, vol. 7, no. 1, pp. 20–26, Jan. 2019, doi: 10.17694/bajece.494920.

[17] T. S. Yange, C. O. Egbunu, O. Onyekwere, and K. A. Foga, “Prediction of Agro Products Sales Using Regression Algorithm,” *Am. J. Data Min. Knowl. Discov.*, vol. 5, no. 1, p. 11, 2020, doi: 10.11648/j.ajdmkd.20200501.12.

[18] I. C. R. Drajana and B. Betrisandi, “Prediksi Harga Jagung Menggunakan Support Vector Machine dengan Fitur Seleksi Forward Selection di Kabupaten Pohuwato,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 7, no. 5, pp. 1248–1255, Sep. 2024, doi: 10.32672/jnkti.v7i5.8059.

[19] I. S. Al-Mejibli, J. K. Alwan, and D. H. Abd, “The effect of gamma value on support vector machine performance with different kernels,” *Int. J. Electr. Comput. Eng.*, vol. 10, no. 5, pp. 5497–5506, Oct. 2020, doi: 10.11591/ijece.v10i5.pp5497-5506.

[20] H.-C. Zhang, Q. Wu, F.-Y. Li, and H. Li, “Multitask Learning Based on Least Squares Support Vector Regression for Stock Forecast,” *Axioms*, vol. 11, no. 6, p. 292, Jun. 2022, doi: 10.3390/axioms11060292.

[21] B. Setiadi, E. Purwanto, and H. Permatasari, “Optimisasi Klasifikasi Sentimen Pada Review Hotel Bahasa Inggris Dengan Model Roberta Twitter,” *SINTECH (Science Inf. Technol. J.*, vol. 7, no. 2, pp. 70–79, Aug. 2024, doi: 10.31598/sintechjournal.v7i2.1547.

[22] R. Faurina and E. Sitanggang, “Implementasi Metode Content-Based Filtering dan Collaborative Filtering pada Sistem Rekomendasi Wisata di Bali,” *Techno.Com*, vol. 22, no. 4, pp. 870–881, Nov. 2023, doi: 10.33633/tc.v22i4.8556.

[23] Tatachar and V. Abhishek, “Comparative Assessment of Regression Models Based On Model Evaluation Metrics,” *Int. Res. J. Eng. Technol.*, vol. 8, no. 9, 2021.

.