JINAV: Journal of Information and Visualization *ISSN: 2746-1440 (online)*
Vol. 5 No. 1 (2024)
https://doi.org/10.35877/454RI.jinav2876

Prediction of Rice Farming Yields in Padangsidimpuan City through Support Vector Machine (SVM) Algorithms

Silviana Ayu Siregar* & Yusuf Ramadhan Nasution

Department of Computer Science, Faculty of Science and Technology, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Abstract

The purpose of this study is to determine the prediction of rice farming yields in Padangsidimpuan City through SVM (Support Vector Machine) Algorithms. This type of research used quantitative methods of SVM (Support Vector Machine) with a Data-Driven development (DDD) method. This approach utilized patterns and trends in data to build accurate prediction models where the DDD method can be used when researchers have access to relevant and meaningful data to guide the development of software or prediction models.The SVM algorithm has proven to be effective in predicting rice yield trends, both in determining the direction of change (up or down) and in estimating the value of the next harvest. The implemented SVM model is able to identify patterns of change in historical data and provide relevant predictions for agricultural yields. Historical data covering a fairly long period of time provides sufficient information for models to identify trends and patterns. This model can provide better predictions with more complete and high-quality data.

Keywords: rice farming, prediction, SVM

Received: 20 March 2024 Revised: 5 June 2024 Accepted: 27 July 2024

1. Introduction

Rice is a crucial staple crop, playing a pivotal role in satisfying global food demands. This plant does not only supplies vital carbohydrates, but also forms the economic foundation for many nations, including Padangsidimpuan City. Utilizing the Support Vector Machine (SVM) algorithm to forecast rice yields can significantly enhance food security and agricultural productivity in the region (Budi et al., 2021). Prediction is a method or process for estimating sequentially and systematically about something that might happen in the future based on the available past and present information, so that the level of errors can be reduced or minimized (Huang et al., 2018; Karatzoglou et al., 2006; Nalepa & Kawulok, 2019). Predictions do not have to provide a definite answer about events that will occur in the future, but rather try to find accurate answers that might happen later (Kurniawan et al., 2023).

In this research, predictions of rice farming yields will be carried out obtained from the data obtained from the Padangsidimpuan City Agriculture Service. The data used the area of rice farming land and the number of harvests each year. The results obtained from creating this system are a prediction model that can predict the yield of rice farming in Padangsidimpuan City. From this prediction model, a picture of the predicted crop yields in Padangsidimpuan City in the coming year can be seen, so that these results can be known effectively, and can be used as a reference for controlling the price of rice agricultural crops in Padangsidimpuan City. With high accuracy, SVM (Support Vector Machine) has been used in many cases (Chen et al., 2022; Huang et al., 2018; Rodríguez-Pérez & Bajorath, 2022). Suhardjono's research, utilized Pso-based SVM to predict student graduation times, resulted in an accuracy of 85,81%. Additionally, initial studies show the best accuracy of 80,55% on average. Therefore, the algorithm is considered successful if it produces an accuracy value of 70% to 100%. In this research, the SVM (Support Vector Machine) algorithm is used to predict rice harvest yields (Al-Mejibli et al., 2020; Nugroho et al., 2022; Pepper et al., 2022; Wagle & Harikrishnan, 2021). The aim of this research is to find out how to apply the SVM (Support Vector Machine) algorithm in predicting rice farming yields in Padangsidimpuan city and to find out how

* Corresponding author.
E-mail address: ramadhannst@gmail.com

the predictions produced from application of the SVM (Support Vector Machine) algorithm (Kurniawan et al., 2023). Based on this literature, this research aims to determine the prediction of rice agricultural harvest results in the city of Padangsidimpuan using the SVM (Support Vector Machine) algorithm.

2. Research methods

This type of research used quantitative methods of SVM (Support Vector Machine) with a Data-Driven development (DDD) method (Kusumastuti, Adhi, 2020; Sugiyono, 2022; Unaradjan, 2019). This approach utilized patterns and trends in data to build accurate prediction models where the DDD method can be used when researchers have access to relevant and meaningful data to guide the development of software or prediction models. An explanation of each stage of the process in the Data-Driven Development (DDD) can be seen as follows:

a. Data Collection

The initial stage of DDD is collecting data that is relevant to the problem the researcher wants to solve or the goals the researcher wants to achieve. These data come from the Padangsidimpuan City Agriculture Service.

b. Data Processing

After the data were collected, the next step was to process the data, including the data cleaning, normalization, and data integration.

c. Data Analysis

The data analysis stage involved statistical modeling, building algorithms, or using machine learning techniques to gain a deep understanding of patterns, trends, or relationships in the data (Azizi et al., 2024; Deng et al., 2018; Shu & Ye, 2023). This analysis might include predictive modeling to make future predictions, descriptive modeling to understand the structure of the data, or inferential modeling to make conclusions about a population based on a sample of data.

d. Model or Application Development

Based on data analysis, then predictive models, applications, or other data-driven systems could be developed to solve problems or achieve set goals.

e. Testing and Evaluation

After the model or application development is complete, the next step is to carry out testing to ensure the quality and reliability of the solution as desired.

f. Implementation

After passing testing, the developed solution could be implemented as intended.

2.1. Data Collection

The data collection techniques used in this research are as follows:

a. Library Research is by looking for references from library books and also journals from the internet as a source for a theoretical basis with the aim of understanding and looking for references that can be used as a theoretical basis for completing this research.

b. Data collection from the Padangsidimpuan City Agriculture Service, this study used a dataset obtained from the Padangsidimpuan City Department year 2010 - 2023. These data have several attributes that could be processed with the Support Vector Machine (SVM) algorithm, including sub-district, land area, and harvest yield/tons.

Table 1. Database Attributes and Measurement Scales

No	Attribute	Measurement Scale
1	Sub-district	Southeast Padangsidimpuan South Padangsidimpuan Padangsidimpuan Batunadua North Padangsidimpuan

No	Attribute	Measurement Scale
		Padangsidimpuan Hutaimbaru
		Padangsidimpuan Angkola Julu
2	Land area	Hectare
3	Yields	Tons

The dataset obtained from the Padangsidimpuan City Agriculture Service can be seen in the Table 2.

Table 2. Padangsidimpuan City Agriculture Service Dataset

Year	Subdistrict	Harvested Area (Ha)	Rice Farming Yield (Tons)
2010	Southeast Padangsidimpuan	1987	10868.00
	South Padangsidimpuan	482	2631.00
	Padangsidimpuan Batunadua	2982	16431.00
	North Padangsidimpuan	943	5139.00
	Padangsidimpuan Hutaimbaru	2209	12194.00
	Padangsidimpuan Angkola Julu	2674	14760.00
2011	Southeast Padangsidimpuan	1985	10867.88
	South Padangsidimpuan	465	2538.90
	Padangsidimpuan Batunadua	2785	15428.90
	North Padangsidimpuan	935	5114.45
	Padangsidimpuan Hutaimbaru	2100	11676.00
	Padangsidimpuan Angkola Julu	2528	14030.40
2012	Southeast Padangsidimpuan	3393	19000.80
	South Padangsidimpuan	602	3371.20
	Padangsidimpuan Batunadua	2671	14690.50
	North Padangsidimpuan	781	4373.60
	Padangsidimpuan Hutaimbaru	2031	11373.60
	Padangsidimpuan Angkola Julu	3400	19040.00
2013	Southeast Padangsidimpuan	2071	11597.60
	South Padangsidimpuan	444	2486.00
	Padangsidimpuan Batunadua	2704	15142.40
	North Padangsidimpuan	850	4760.00
	Padangsidimpuan Hutaimbaru	2013	11272.80
	Padangsidimpuan Angkola Julu	3925	21980.00
2014	Southeast Padangsidimpuan	1540	8624.00
	South Padangsidimpuan	289	1618.40
	Padangsidimpuan Batunadua	2961	16581.00
	North Padangsidimpuan	762	4267.20
	Padangsidimpuan Hutaimbaru	2185	12236.00
	Padangsidimpuan Angkola Julu	1883	10544.80
2015	Southeast Padangsidimpuan	2005	11228.00
	South Padangsidimpuan	220	1232.00
	Padangsidimpuan Batunadua	2861	16021.60
	North Padangsidimpuan	883	4944.60
	Padangsidimpuan Hutaimbaru	2260	12656.00
	Padangsidimpuan Angkola Julu	2263	12672.80
2016	Southeast Padangsidimpuan	2024	11739.20
	South Padangsidimpuan	572	3317.60
	Padangsidimpuan Batunadua	2779	16118.20
	North Padangsidimpuan	1150	6672.90
	Padangsidimpuan Hutaimbaru	2995	17371.00
	Padangsidimpuan Angkola Julu	2112	12249.60
2017	Southeast Padangsidimpuan	2149.00	12922.00
	South Padangsidimpuan	396.00	2306.00
	Padangsidimpuan Batunadua	1775.00	10687.00
	North Padangsidimpuan	946.50	5614.00

Year	Subdistrict	Harvested Area (Ha)	Rice Farming Yield (Tons)
	Padangsidimpuan Hutaimbaru	2264.00	13474.00
	Padangsidimpuan Angkola Julu	1561.50	9110.00
	Southeast Padangsidimpuan	2362.00	14098.00
	South Padangsidimpuan	559.00	3332.00
2018	Padangsidimpuan Batunadua	2267.00	13250.00
	North Padangsidimpuan	1114.00	6610.00
	Padangsidimpuan Hutaimbaru	2916.00	17825.00
	Padangsidimpuan Angkola Julu	1737.00	10305.00
	Southeast Padangsidimpuan	2531.00	15361.00
	South Padangsidimpuan	537.00	3212.00
2019	Padangsidimpuan Batunadua	2018.00	11896.00
	North Padangsidimpuan	1037.00	6154.00
	Padangsidimpuan Hutaimbaru	2610.00	15955.00
	Padangsidimpuan Angkola Julu	2288.00	13689.00
2020	Southeast Padangsidimpuan	2327.00	15428.00
	South Padangsidimpuan	535.00	3323.00
	Padangsidimpuan Batunadua	2197.00	14098.00
	North Padangsidimpuan	1127.00	7353.00
	Padangsidimpuan Hutaimbaru	2760.00	17675.00
	Padangsidimpuan Angkola Julu	1848.00	11502.00
	Southeast Padangsidimpuan	2427.00	16297.00
	South Padangsidimpuan	476.00	3099.99
2021	Padangsidimpuan Batunadua	1841.00	12187.00
	North Padangsidimpuan	1075.00	7122.00
	Padangsidimpuan Hutaimbaru	2820.00	18126.00
	Padangsidimpuan Angkola Julu	1914.00	12470.00
	Southeast Padangsidimpuan	2443.00	16642.00
	South Padangsidimpuan	400.00	2614.00
2022	Padangsidimpuan Batunadua	2130.00	14186.00
	North Padangsidimpuan	768.00	5161.00
	Padangsidimpuan Hutaimbaru	2559.00	17171.00
	Padangsidimpuan Angkola Julu	2278.00	14853.00
	Southeast Padangsidimpuan	2443.00	16642.00
	South Padangsidimpuan	400.00	2614.00
2023	Padangsidimpuan Batunadua	2130.00	14186.00
	North Padangsidimpuan	768.00	5161.00
	Padangsidimpuan Hutaimbaru	2559.00	17171.00
	Padangsidimpuan Angkola Julu	2278.00	14853.00

(Source: Padangsidimpuan City Agriculture Service)

2.2. Data Pre-processing

Data pre-processing is a series of important processes in the data analysis cycle that involves transforming, processing, and structuring data to ensure its accuracy and prepare it for effective use in further analysis. The initial step in data processing involves identifying and addressing problems with the data, such as missing, duplicate, or invalid values. Next, data transformation was carried out, which includes format conversion, normalization, or simplification of variables to suit analysis needs.

Integration of data from various sources is also an important aspect in data processing, enabling the preparation of more complete and comprehensive datasets. During this process, data from multiple sources is combined and harmonized to create a more holistic picture. After that, the data is structured by designing an organized structure, such as grouping, sorting, or indexing to make searching and analysis easier. This research took data on rice agricultural harvests for 2010-2023 which contained the name of the sub-district, land area, and annual harvest. The data obtained were preprocessed, which consists of cleaning, transformation, and data normalization. The data were then classified using the SVM algorithm and evaluated with RMSE. These were divided into training data and testing data. The system testing process was carried out using agricultural production data for the 2010 - 2023 period. The statistical data consists of 84 (eighty four) data and contains 3 (three) variables, namely the name of the sub-district, land area, and annual harvest. Data preprocessing is divided into several steps, namely data cleaning, data transformation, and data normalization. Then, it was calculated using the SVM (Support Vector Machine) algorithm and then the RMSE (Root Mean Square Error) is evaluated (Anggoro, 2020; Piccialli & Sciandrone, 2022; Syahputra & Wibowo, 2023; Zulfa & Winarko, 2017).

The implementation of this system aims to predict the yield of rice farming in the city of Padangsidimpuan, based on the sub-district, land area, and harvest yields each year. The location of this research is the Padangsidimpuan City Agriculture Service.

3. Results and Discussion

The data used in this research were obtained from the Padangsidimpuan City Agriculture Service and various other related sources, which provide important information regarding rice farming. The process of predicting rice production in Padangsidimpuan City were not carried out directly but rather per sub-district. Therefore, the data model would be changed to something such the one on Table 3.

Table 3. Data Transformation

Subdistrict	Year													
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Southeast Padangsidimpuan	10868	10867	19000	11597	8624	11228	11739	12922	14098	15361	15428	16297	16642	16642
South Padangsidimpuan	2631	2538	3371	2486	1618	1232	3317	2306	3332	3212	3323	3099	2614	2614
Padangsidimpuan Batunadua	16431	15428	14690	15142	16581	16021	16118	10687	13250	11896	14098	12187	14186	14186
North Padangsidimpuan	5139	5114	4373	4760	4267	4944	6672	5614	6610	6154	7353	7122	5161	5161
Padangsidimpuan Hutaimbaru	12194	11676	11373	11272	12236	12656	17371	13474	17825	15955	17676	18126	17171	17171
Padangsidimpuan Angkola Julu	14760	14030	19040	21980	10544	12672	12249	9110	10305	13689	11502	12470	14853	14853

3.1. Data normalization and labeling

The data normalization process is required before the prediction process performed. The normalization technique used the z-score normalization technique, for example, production data in Padangsidimpuan Tenggara subdistrict.

Features and labels:

- Features (X_i)
- Label (Y_i) => $Y_i = 1$ if $X_{i+1} > X_i$ (Increase) and $Y_i = -1$ if $X_{i+1} < X_i$ (Decrease)

Table 4. Data Labeling

No	X_i	X_{i+1}	Y_i
1	10868	10867	-1
2	10867	19000	1
3	19000	11597	-1
4	11597	8624	-1
5	8624	11228	1
6	11228	11739	1
7	11739	12922	1
8	12922	14098	1
9	14098	15361	1
10	15361	15428	1
11	15428	16297	1
12	16297	16642	1
13	16642	16642	-1

After transformation, the next step was data normalization process. This research used the z-score data normalization technique. The following are the normalization results of sample data from Table 5.

Table 5. Data Normalization

No	X_i	X_{i+1}
1	-0.90245	-0.9028
2	-0.9028	1.955051
3	1.955051	-0.64629
4	-0.64629	-1.69097
5	-1.69097	-0.77595
6	-0.77595	-0.59639
7	-0.59639	-0.1807
8	-0.1807	0.232539
9	0.232539	0.676344
10	0.676344	0.699887
11	0.699887	1.005245
12	1.005245	1.126474
13	1.126474	1.126474

$$x_{norm} = \frac{x - \mu}{\delta}$$

Information:

x_{norm} = normalized data
x = initial-data
μ = column mean
δ = standard-deviation

First, find the average value of the first column:

$$\mu = \frac{10868+10867+1900+..+16297+16642}{13} = 13436$$

After the average value from the first column is obtained, calculate the standard-deviation value from the first column.

$$\delta = \sqrt{\frac{(10868-13436\,)^2+(10867-13436\,)^2+..+\,(16297-13436\,)\text{^}2+(16642-13436\,)\text{^}2}{13}}$$

$$\delta = 2845$$

Calculate the normalized value of the first data.

$$x_{norm} = \frac{10868-\ 13436}{2845} = -0.902$$

3.2. Implementation of the SVM Algorithm

The next step was to find the values of *w* and *b* that maximized margin. In this case, it focused on selected data, which became support vectors. For small datasets, the illustration used a few selected values as support vectors. Support vectors are the right data points on the margin or very close to the margin. In the actual process of implementing the SVM algorithm, it used normalized data. However, for simplicity of manual calculations, raw data were employed, selecting support vectors from the dataset.

Phase 1: Training Process

X_1 = 10867 with label Y_1 = -1
X_2 = 19000 with label Y_2 = 1

For support vectors the following equation is used.

$$y_1(w * x_1 + b) = -1 \ \ dan \ \ y_2(w * x_2 + b) = 1$$

So, the equation for each label is obtained as follows.

$label - 1 => \; -w * 10867 + b = -1$

$label\,1 \quad => \; w * 19000 + b = 1$

Complete the system of equations to find the values of *w* and *b* by substituting the equation label -1 for the equation label 1.

Equation 1:

$$-w * 10867 + b = -1$$
$$b = w.10867 - 1$$

Substitute equation 1 for equation 2.

$$w * 19000 + w.10867 - 1 = 1$$
$$w(19000 + 10867) = 2$$
$$w(29867) = 2$$
$$w = \frac{2}{29867} = 0.000067$$

After the value of *w* was obtained at 0.000067, then the value of *b* was searched by substituting the value of *w* into *b*.

$$b = w.10867 - 1$$
$$b = 0.000067 * 10867 - 1$$
$$b = 0.73 - 1$$
$$b = -0.27$$

Phase 2: Testing Process

After the *w* and *b values* were obtained, the next step was to carry out the testing process, namely the prediction process with a hyperplane using the *w* and *b* values that have been obtained from the previous training process.

For example:

Last data value = 16642

$$prediction = sign(w * 16642 + b)$$
$$prediction = sign(0.000067 * 16642 + (-0.27)$$
$$prediction = sign(1.114954 - 0.27)$$
$$prediction = sign(0.844954) = 1$$

Because the prediction result is 1, the next data value increased. For increasing/decreasing numerical values, the average change method was used. The average change method is a simple technique for predicting the next value in a time series by calculating the average change in values from historical data.

a. Calculate the average change in Δ for class 1 (increase)

$$\Delta_{increase} = \frac{(19000-10867)+(11228-8624)+ .. +(16297-15428)+(16642-16297)}{9}$$

$$\Delta_{increase} = \frac{(8133+2604+511+1183+..+1176+1263+67+869+345)}{9}$$

$$\Delta_{increase} = \frac{17151}{9} = 1905.67$$

b. Calculate the predicted value of the next production using the average value of change.

Last data: 16642

$$Prediction\ value = Last\ Data + \ \Delta_{increase}$$
$$Prediction\ value = 16642 + \ 1905$$
$$Prediction\ value = 18547$$

Based on historical data on rice production in Padangsidimpuan Tenggara district from 2010 to 2023, the predicted results for production for 2024 are around 18,547 tons. In the prediction results by the system, the values obtained are slightly different from the results of manual calculations because the computational level of the manual calculations carried out was not as high as the system computing level.

Using the same calculations as before, a prediction process was carried out for other sub-districts in Padangsidimpuan City so that predicted results for rice crop production for 2024 were obtained.

Table 6. Production Prediction Results for 2024

No	Subdistrict	Class	Amount Rice Production (Tons)
1	Southeast Padangsidimpuan	Increase	18436
2	South Padangsidimpuan	Decrease	2161
3	Padangsidimpuan Batunadua	Decrease	12615
4	North Padangsidimpuan	Increase	6158
5	Padangsidimpuan Hutaimbaru	Decrease	16079
6	Padangsidimpuan Angkola Julu	Increase	17425
Total Production			**72874**

From the SVM model implementation to the available dataset, the prediction results show that the model is able to identify patterns of change in rice yields and provide useful predictions regarding trends in increasing or decreasing production. For each dataset analyzed, the model produces the predictions that match historical patterns, as well as providing estimates of the next harvest value based on trend classification.

In order to evaluate the model that has been created, the error value was also calculated. The error value calculation technique used RMSE, where with the help of python, the RMSE obtained from the entire prediction process is 20.05%. Through this RMSE value, it can be concluded that the predictions made have an accuracy level of 79.95%.

3.3. Implementation

Prediction modeling was built in the Python programming language with *Jupyter Notebook* as a text-editor. The following is an implementation of the Python language in the modeling that has been created.

a. Import and Initialization Libraries

In predictive modeling, several libraries were used to help the process. The following is the code to import and initialize the library that will be used.

```
import numpy as np
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
```

Figure 1. Import and initialization *libraries*

b. Input Dataset

The rice production dataset is entered into code in json form. The following is a display of the code for inputting the dataset.

```
datasets = {
    'Padangsidimpuan Tenggara': [10868, 10867, 19000, 11597, 8624, 11228, 11739,
                                 12922, 14098, 15361, 15428, 16297, 16642, 16642],
    'Padangsidimpuan Selatan': [2631, 2538, 3371, 2486, 1618, 1232, 3317, 2306,
                                3332, 3212, 3323, 3099, 2614, 2614],
    'Padangsidimpuan Batunadua': [16431, 15428, 14690, 15142, 16581, 16021, 16118,
                                  10687, 13250, 11896, 14098, 12187, 14186, 14186],
    'Padangsidimpuan Utara': [5139, 5114, 4373, 4760, 4267, 4944, 6672, 5614, 6610,
                              6154, 7353, 7122, 5161, 5161],
    'Padangsidimpuan Hutaimbaru': [12194, 11676, 11373, 11272, 12236, 12656, 17371,
                                   13474, 17825, 15955, 17676, 18126, 17171, 17171],
    'Padangsidimpuan Angkola Julu': [14760, 14030, 19040, 21980, 10544, 12672, 12249,
                                     9110, 10305, 13689, 11502, 12470, 14853, 14853]
}
```

Figure 2. Input dataset

c. Labeling Data

Each data were given a label, where if the previous production amount was less than the next production amount, it will be labeled -1, which means decreasing. Meanwhile, if the previous production amount is greater than the next production amount, it will be labeled 1, which means increasing.

```
for i in range(len(data) - 1):
    features.append([data[i]])
    if data[i+1] > data[i]:
        labels.append(1)
    else:
        labels.append(0)
```

Figure 3. Labeling Data

d. Normalization Data

Before processing using an algorithm *Support Vector Machine* (SVM), data were first normalized. The normalization process was carried out using the *z-score technique*. Here is the implementation in python.

```
scaler = StandardScaler()
features = scaler.fit_transform(features)
train_features = features[:-1]
train_labels = labels[:-1]
test_features = features[-1:]
```

Figure 4. Normalization Process

e. Support Vector Machine (SVM) process

The SVM process included the model formation process and the prediction process. The following is its implementation in the Python language.

```
    model = SVC(kernel='linear')
    model.fit(train_features, train_labels)
    return model, scaler, test_features, data[-1], labels

def predict_next_value(model, scaler, test_features, current_value, labels, data):
    predicted_class = model.predict(test_features)[0]
    change_up = np.mean([data[i+1] - data[i] for i in range(len(labels)) if labels[i] == 1])
    change_down = np.mean([data[i+1] - data[i] for i in range(len(labels)) if labels[i] == 0])
    if predicted_class == 1:
        predicted_value = current_value + change_up
    else:
        predicted_value = current_value + change_down
    return predicted_class, predicted_value

for name, data in datasets.items():
    model, scaler, test_features, current_value, labels = train_svm_model(data)
    predicted_class, predicted_value = predict_next_value(model, scaler, test_features, current_value, labels, data)
```

Figure 5. SVM Process

f. RMSE Evaluation

After carrying out the prediction process, error calculations were conducted using the RMSE technique. Therefore, the error value in the process that has been carried out could be known.

```python
def calculate_rmse(actual_values, forecasted_values):
    actual_values = np.array(actual_values[1:])
    forecasted_values = np.array(forecasted_values[:-1])
    errors = np.abs((actual_values - forecasted_values) / actual_values)
    mape = np.mean(errors) * 100
    return mape
```

Figure 6. RMSE calculation

The following is a detailed visualization of the RMSE values from each dataset from each sub-district.

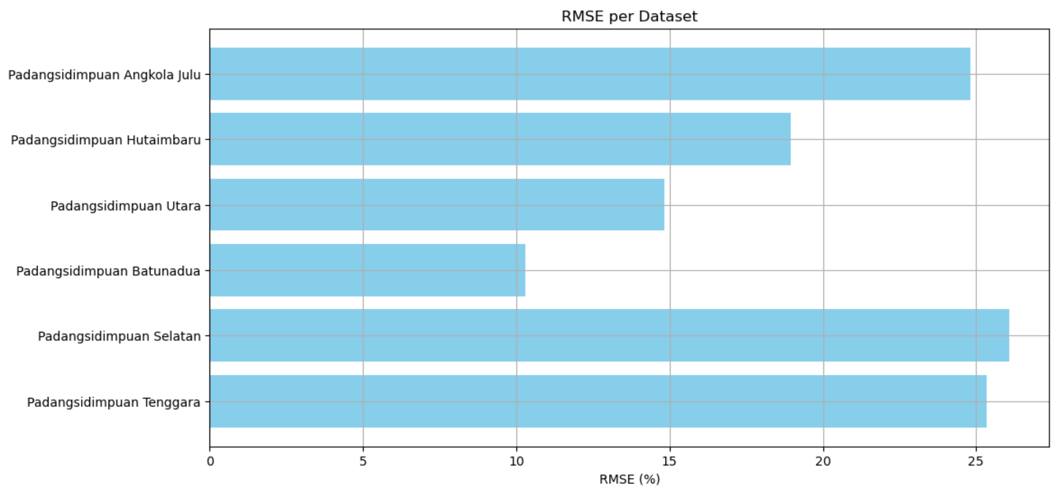

Figure 7. Details of RMSE values per sub-district

4. Conclusions and Recommendations

The SVM algorithm has proven to be effective in predicting rice yield trends, both in determining the direction of change (up or down) and in estimating the value of the next harvest. The implemented SVM model is able to identify patterns of change in historical data and provide relevant predictions for agricultural yields. Historical data covering a fairly long period of time provides sufficient information for models to identify trends and patterns. This model can provide better predictions with more complete and high-quality data. The evaluation used in the prediction process is the RMSE technique, where by using this technique the error value of the prediction process is 20,05% so that the accuracy of the prediction is 79,95%.

References

Al-Mejibli, I. S., Alwan, J. K., & Abd, D. H. (2020). The effect of gamma value on support vector machine performance with different kernels. *International Journal of Electrical and Computer Engineering*, *10*(5). https://doi.org/10.11591/IJECE.V10I5.PP5497-5506

Anggoro, D. A. (2020). Comparison of Accuracy Level of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) Algorithms in Predicting Heart Disease. *International Journal of Emerging Trends in Engineering Research*, *8*(5). https://doi.org/10.30534/ijeter/2020/32852020

Azizi, M., Baghalzadeh Shishehgarkhaneh, M., Basiri, M., Moehler, R. C., Fang, Y., & Chan, M. (2024). Wolf-Bird Optimizer (WBO): A novel metaheuristic algorithm for Building Information Modeling-based resource tradeoff. *Journal of Engineering Research (Kuwait)*. https://doi.org/10.1016/j.jer.2023.11.024

Budi, A. S., Susilo, P. H., Informatika, T., Teknik, F., Lamongan, U. I., Panen, H., & Padi, T. (2021). *583-1273-1-PB (Metode svm)*. *6*(1).

Chen, H., Yu, Y., Jia, Y., & Zhang, L. (2022). Safe transductive support vector machine. *Connection Science*, *34*(1). https://doi.org/10.1080/09540091.2021.2024511

Deng, H., Fannon, D., & Eckelman, M. J. (2018). Predictive modeling for US commercial building energy use: A comparison of existing statistical and machine learning algorithms using CBECS microdata. *Energy and Buildings*, *163*. https://doi.org/10.1016/j.enbuild.2017.12.031

Huang, S., Nianguang, C. A. I., Penzuti Pacheco, P., Narandes, S., Wang, Y., & Wayne, X. U. (2018). Applications of support vector machine (SVM) learning in cancer genomics. In *Cancer Genomics and Proteomics* (Vol. 15, Issue 1). https://doi.org/10.21873/cgp.20063

Karatzoglou, A., Meyer, D., & Hornik, K. (2006). Support vector machines in R. *Journal of Statistical Software*, *15*(9). https://doi.org/10.18637/jss.v015.i09

Kurniawan, R., Halim, A., & Melisa, H. (2023). KLIK: Kajian Ilmiah Informatika dan Komputer Prediksi Hasil Panen Pertanian Salak di Daerah Tapanuli Selatan Menggunakan Algoritma SVM (Support Vector Machine). *Media Online)*, *4*(2), 903–912. https://doi.org/10.30865/klik.v4i2.1246

Kusumastuti, Adhi, dkk. (2020). Metode Penelitian Kuantitatif. In *Google Books* (Issue April 2016). Yogyakarta: Deepublish Lickona.

Nalepa, J., & Kawulok, M. (2019). Selecting training sets for support vector machines: a review. In *Artificial Intelligence Review* (Vol. 52, Issue 2). https://doi.org/10.1007/s10462-017-9611-1

Nugroho, K. S., Sukmadewa, A. Y., Vidianto, A., & Mahmudy, W. F. (2022). Effective predictive modelling for coronary artery diseases using support vector machine. *IAES International Journal of Artificial Intelligence*, *11*(1). https://doi.org/10.11591/ijai.v11.i1.pp345-355

Pepper, N., Crespo, L., & Montomoli, F. (2022). Adaptive learning for reliability analysis using Support Vector Machines. *Reliability Engineering and System Safety*, *226*. https://doi.org/10.1016/j.ress.2022.108635

Piccialli, V., & Sciandrone, M. (2022). Nonlinear optimization and support vector machines. *Annals of Operations Research*, *314*(1), 15–47. https://doi.org/10.1007/s10479-022-04655-x

Rodríguez-Pérez, R., & Bajorath, J. (2022). Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery. *Journal of Computer-Aided Molecular Design*, *36*(5). https://doi.org/10.1007/s10822-022-00442-9

Shu, X., & Ye, Y. (2023). Knowledge Discovery: Methods from data mining and machine learning. *Social Science Research*, *110*. https://doi.org/10.1016/j.ssresearch.2022.102817

Sugiyono. (2022). *Buku Metode Penelitian Kuantitatif Kualitatif dan R&D*.

Syahputra, H., & Wibowo, A. (2023). Comparison of Support Vector Machine (SVM) and Random Forest Algorithm for Detection of Negative Content on Websites. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, *9*(1).

Unaradjan, D. (2019). *Metode penelitian kuantitatif*. Penerbit Universitas Katolik Atma Jaya, Jakarta.

Wagle, S. A., & Harikrishnan, R. (2021). Comparison of plant leaf classification using modified alexnet and support vector machine. *Traitement Du Signal*, *38*(1). https://doi.org/10.18280/TS.380108

Zulfa, I., & Winarko, E. (2017). Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *11*(2), 187. https://doi.org/10.22146/ijccs.24716