Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 10, No. 3, August, Pp. 271-282

Land price distribution prediction in Jakarta using support vector machine with feature expansion and kriging interpolation

Hadid Pilar Gautama *[1], Sri Suryani Prasetiyowati*[1], Yuliant Sibaroni*[1]
Telkom University, Indonesia[1]

Article Info

Keywords:
Land Price Prediction, Jakarta, Support Vector Machine (SVM), Time Based Feature Expansion, Kriging Interpolation

Article history:
Received: January 02, 2025
Accepted: February 13, 2025
Published: August 31, 2025

Cite:
H. Pilar Gautama, S. S. Prasetiyowati, and Y. Sibaroni, "Land Price Distribution Prediction in Jakarta Using Support Vector Machine with Feature Expansion and Kriging Interpolation", *KINETIK*, vol. 10, no. 3, Aug. 2025. https://doi.org/10.22219/kinetik.v10i3.2216

*Corresponding author.
Hadid Pilar Gautama
Sri Suryani Prasetiyowati
Yuliant Sibaroni
E-mail address:
hadidpilargautama@student.telkomuniversity.ac.id
srisuryani@telkomuniversity.ac.id
yuliant@telkomuniversity.ac.id

Abstract

Fluctuations in land prices over time are significant, especially in big cities, one of which is Jakarta. The increase in land prices is influenced by high demand, location-related needs, ease of access to various public facilities and population density. Uncontrolled prices and lack of information about the distribution of land prices cause buyers to acquire land that does not meet their needs. This study develops a land price distribution prediction system for Jakarta for 2025-2026 using Support Vector Machine (SVM) with time-based feature expansion and spatial interpolation. The SVM model with an RBF kernel demonstrated superior performance, achieving 93.14% accuracy for 2025 predictions using the t-1 model. For 2026 predictions, the t-2 model achieved 83.33% accuracy. This approach involves utilizing one to two years of historical data and systematically selected features, ensuring more accurate and relevant predictions. Ordinary kriging interpolation visualizations revealed a significant shift in land price distribution patterns, indicating a decline in affordable land availability and an increase in high-value properties across Jakarta. The integration of SVM and kriging interpolation, coupled with comprehensive evaluation metrics, provides a robust methodological framework for predicting urban land price distributions. This system offers practical implications for informed decision-making in Jakarta's dynamic land market, enabling stakeholders to make efficient, budget-based property decisions. The research contributes significantly to urban planning by providing a comprehensive tool for understanding and predicting land price trends, which can assist various stakeholders in making informed property investment decisions.

1. Introduction

Jakarta, as one of the largest cities in Southeast Asia, is experiencing rapid growth and a population density of 11.34 million people [1] as of mid-2024, covering an area of 661.23 km^2 [2]. Interestingly, despite this density, Jakarta has a low poverty rate, with a rate of 4.44 percent in 2023 [3]. The high population density and limited land in Jakarta have caused land prices to soar. This poses a challenge for the community, especially for those who want to buy land for residential or investment purposes. The land prices set by sellers are often too high, considering the various factors that affect land pricing [4]. The need for land is crucial as it provides protection, comfort, privacy, and space for safely and stably living daily life [5]. Information on the classification and price map of land in Jakarta is essential to help land seekers find a home or investment that suits their needs and budget amid the dynamics of this metropolitan city. Currently, state-of-the-art methods in land price prediction globally have evolved to incorporate artificial intelligence and machine learning techniques, with particular emphasis on spatial-temporal modeling and deep learning approaches.

Various studies have addressed the prediction and classification of land prices using various methods. The first study used Random Forest to classify land prices in Jakarta, achieving the best accuracy of 82% with the ADASYN method and Ordinary Kriging spatial analysis [6]. Random Forest was also used in house price prediction, with an accuracy of 75.10%, highlighting key factors such as land area, building, and the number of rooms[7]. Linear regression models showed good accuracy compared to Decision Tree in predicting property prices [8]. The relationship between land use and Urban Heat Island (UHI) phenomenon in Jakarta was also analyzed, showing that green non-agricultural land is more effective in reducing UHI than other land types [4]. In addition, transfer learning algorithms were proposed to overcome the limitations of small datasets and class imbalance in land price classification using SVM, Random Forest, and BP Neural Network. The PCA method with cross-validation improved feature extraction, and the rich influence factor system successfully improved the precision of residential land price assessment[9]. In time series-based feature extraction, the CART algorithm achieved the highest accuracy of 93% [10] in the last two to four years, with features such as population size and education level being the main factors[11]. Furthermore, sentiment analysis of Indihome services resulted in 87% accuracy [6], as well as prediction of students' on-time graduation [12] and stroke disease with up to 100% accuracy using a linear kernel [13]. However, SVM performance is highly dependent on parameter and kernel selection. In data regression, for example, the model accuracy only reached 71.80% [14]. The advantages of SVM in handling high-dimensional data need to be balanced with parameter optimization and outlier management for better results.

Other studies have shown the superiority of SVM in various scenarios. For example, the SVM method with PCA achieved 98.97% accuracy in the classification of tourist attractions recommendations, outperforming Decision Tree, which had 96.55% accuracy[15]. However, in the classification of diabetic patient data, the Modified Balanced Random Forest (MBRF) method produced 97.8% accuracy, higher than SVM, which reached 91.48%[16]. In network intrusion classification analysis, SVM performance showed variation based on the kernel used, where the Polynomial kernel achieved accuracies of up to 99.999%, although the ROC curve showed weaknesses in some cases [17]. In addition, SVM with a radial kernel achieved 90.20% accuracy in sentiment analysis of online transportation app reviews, which is higher than Decision Tree [18]. In this study, SVM with an RBF kernel was used for land price classification in Jakarta, resulting in the highest accuracy of 93.14%. These results support the superiority of SVM, especially compared to other algorithms in previous studies. With optimal parameter selection and relevant feature combinations, SVM can be the best method for spatial data analysis and land price classification. This confirms the relevance of SVM in handling high-dimensional data and generating accurate predictions in real estate and other domains.

The rapid urbanization and limited land availability in Jakarta have led to significant challenges in land price prediction and classification. The high demand for land and population growth in Jakarta present challenges for land seekers. There has been a continued significant increase in land prices in Jakarta[19]. Despite extensive research in this field, several critical challenges remain unaddressed: (1) existing models inadequately account for temporal changes in land prices, (2) current classification methods struggle with complex spatial relationships in urban land prices, and (3) there is a lack of comprehensive approaches combining both spatial and temporal features for land price prediction.

In this context, this research aims to provide two main solutions: price classification and mapping. The land price classification will group land in Jakarta based on similar characteristics such as location, accessibility and amenities. This will help land seekers understand the price range of land in their area of interest. In addition, land price mapping will visualize the distribution of land prices across Jakarta. Land seekers can easily and intuitively see the differences in land prices in each area.

This research also includes feature expansion, which allows the model to capture more complexity in the data, thus providing more accurate predictions. The key contributions of this research are: (1) the development of an enhanced SVM model with feature expansion capability for improved accuracy, (2) the implementation of advanced kriging interpolation for detailed spatial visualization, and (3) creation of a predictive framework that considers both spatial and temporal aspects of land prices.

In addition, this research uses kriging interpolation to map land prices in a more detailed and accurate manner. Kriging interpolation allows us to estimate land prices across Jakarta, even though land price data is only available for a few specific locations. With kriging interpolation, we can create a more refined and representative land price map, helping land buyers and sellers make better decisions based on accurate spatial information.

This research uses the SVM model, a machine learning algorithm that has proven effective in classification and prediction. The SVM model will provide accurate and easy-to-understand information to help land seekers make informed and efficient decisions. Furthermore, this research will also provide a prediction of the distribution of land prices on each street in Jakarta for the next 2 years, namely 2025 and 2026, using kriging interpolation.

The specific objectives of this study are: (1) to develop an improved SVM-based classification model for land prices in Jakarta, (2) to implement and validate the effectiveness of feature expansion in enhancing prediction accuracy, and (3) to create an accurate spatial-temporal prediction model for future land prices. The results of this research are expected to provide solutions for land seekers in Jakarta, helping them to find land that suits their needs and budget, as well as assisting sellers in setting appropriate land prices.

2. Research Method

2.1 Research Stages

The initial stage in this research began by collecting land price data in Jakarta from various relevant sources, covering the years 2022 to 2024. The data was then labeled according to predetermined attributes to achieve accurate classification results. After that, preprocessing was done to clean and prepare the data so that it can be processed properly by the machine learning model. The data was divided into training data and test data, where the training data was used to build a classification model using Support Vector Machine (SVM) with time-based feature expansion. The classification results were compared with the labels in the test data to evaluate the accuracy of the model. If any data imbalance is found, further adjustments are made. Next, the data was spatially analyzed using the kriging interpolation method. The results of the spatial analysis are visualized on a map of Jakarta in the form of a heatmap to illustrate the distribution of land prices. The flow of this experiment can be seen in Figure 1.

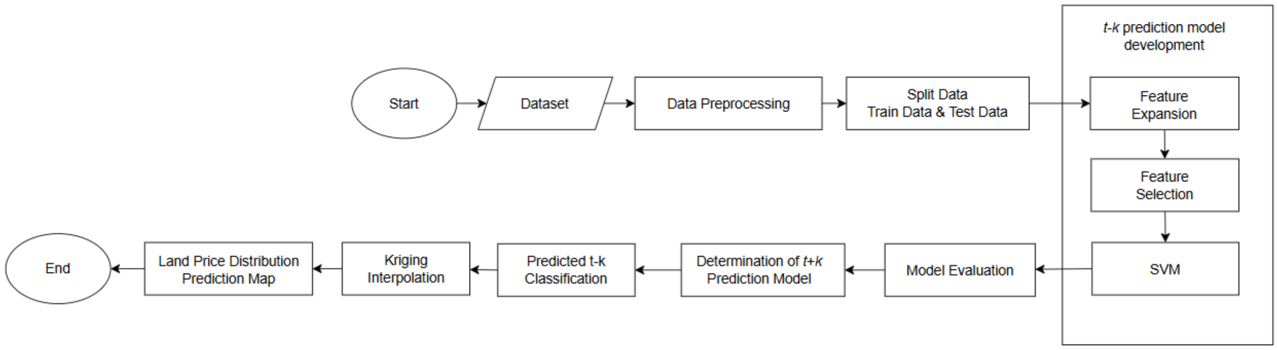


Figure 1. Experiment Flowchart

2.2 Dataset

The datasets used were obtained from rumah123.com and 99.co, containing information, such as land price, region, land area, and city. Additional attributes were obtained through observation using Google Maps based on the locations available from the web scraping data. Other attributes that were used can be seen in Table 1. In the raw dataset, there are locations and districts available for spatial mapping; but to facilitate the research model, the x and y coordinates were created first, as shown in Table 1.

Table 1. Dataset Description

Variable	Features	Description
x_1	Availability of public transport	Integer, if present '1' if not present '0'
x_2	Population crowd level	Integer, if crowded '1' if not '0'
x_3	River	Integer, if close to the river '1' otherwise '0'
x_4	Road width	Integer, road width
x_5	Main road	Integer, distance to main road
x_6	Residential area level	Integer, if luxury housing '2', if normal housing '1', if slum housing '0'
x_7	Education facilities	Integer, number of education facilities available
x_8	Health facilities	integer, number of Health facilities available
x_9	Trade in services	Integer, trade availability count
x_{10}	Recreation centre	Integer, number of recreation centres available
x_{11}	Distance to toll road	Integer, distance to toll road
x_{12}	KRL	Integer, number of KRL availability
x_{13}	Trans jakarta	Integer, number of KRL availability
x_{14}	MRT	Integer, number of KRL availability
x_{15}	District	Object, a district
x_{16}	Address	Object, containing a street address
y_1	Land price	Land price

2.3 Labeling

After scraping data from one of the land sales websites in Jakarta, the next step was to label the data, the original land price was divided by the land area to obtain the land price per square meter. Subsequently, the land price was divided into three classes, which can be seen in Table 2.

Table 2. Land Price Class

Class	Description
0	Land price < Rp 15.000.000
1	Rp 15.000.000 < Land prices < Rp 30.000.000
2	Land price > Rp 30.000.000

2.4 Preprocessing

One of the most common data mining activities is data preprocessing. Data preprocessing is an important step in improving data quality. The data preprocessing techniques performed include[20]:

1. Cleaning Data: Cleaning data involves removing errors and inconsistencies. This includes handling missing values, irrelevant values, noise, and non-uniform formats. Non-integer data type, such as "Address", "City Category" and "Land Width", are checked and deleted.
2. Normalization: Normalization, in the context of data preprocessing, is a technique to reduce data duplication and redundancy and improve data consistency [21]. For the normalization equation, see Equation 1:

$$Value\ Normalized = \frac{x_i - Mean}{\sigma} \tag{1}$$

2.5 Feature Expansion

Feature expansion is the process of adding or modifying features in a dataset to improve the model's ability to recognize complex patterns or relationships in the data [22]. The feature expansion process involves predicting classification based on time. Table 3 illustrates how data collected from three years ago can generate time predictions for the next two years.

Table 3. Scenario of Prediction Model Feature Expansion

Prediction	Model	Train Feature(x)	Class(y)
1 Years	*t*−1 1A	2023	2024
	t−1 1B	2022	2023
2 Years	*t*−2 2A	2022 & 2023	2024

To predict 2025 and 2026 based on the expanded features of data from 2022, 2023, and 2024, the scenario of feature expansion are considered. Table 3 shows examples of feature expansion combinations used in previous years.

2.6 Support Vector Machine (SVM) with time-based features expansion

SVM is a classification algorithm that separates data by finding the maximum margin between two support lines, resulting in optimal separation. For nonlinear data, SVM uses a kernel trick to map the data to higher dimensions without explicit calculations, allowing for more complex and diverse separations[23]. SVM is a popular algorithm based on statistical learning theory. This hyperplane separates the data by ensuring that data from the same class is on the same side[24], as shown in Figure 2.

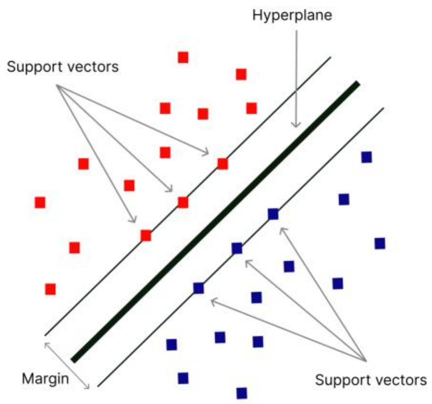


Figure 2. General Classification Hyperplane Representation of SVM Algorithm[25]

It works by finding a hyperplane, which is an optimal separating line or plane that maximizes the number of data points with the same class on the same side. To improve accuracy, SVM uses a kernel function that maps data to a higher dimensional space so that the hyperplane can better separate data [26]. Frequently used SVM formulas are in Equations 2, 3, 4, and 5.

1. Linear kernel

$$K(x,y) = x \cdot y + c \tag{2}$$

2. Polynomial Kernel

$$K(x,y) = (x \cdot y + c)^{d} \tag{3}$$

3. Radial Bias Function (RBF)

$$K(x,y) = \exp(-\gamma \, ||x - y||)^{2} \tag{4}$$

4. Sigmoid Kernel

$$K(x,y) = \tanh (\alpha \, x \cdot y + c) \tag{5}$$

Description:

(x,y)	: the data vector
·	: dot product
c	: constant
d	: degree of polynomial
$\|\|x - y\|\|^{2}$	: the squared Euclidean distance between two vectors x and y
γ	: the kernel parameter
x · y	: inner product of two vectors x and y
α dan c	: adjustable parameters

2.7 Confusion Matrix Multiclass

The SVM model used in the case study is classified as multi-class because it not only distinguishes between two categories (e.g., high and low value land prices) but also predicts property prices within a wider range of values [27]. This requires a multi-class confusion matrix to measure the prediction accuracy across different price classes.

1. Accuracy

It describes the accuracy of the model used in classification. The formula is in Equation 6.

$$Accuracy = \frac{TP_{11} + TP_{22} + \ldots + TP_{nn}}{Total\ Examples} \tag{6}$$

2. Precision

It describes the accuracy between the requested data and the prediction results provided by the model. The formula is in Equation 7.

$$\text{Precision}_i = \frac{TP_{ii}}{TP_{i1} + TP_{i2} + \ldots + TP_{in}} \tag{7}$$

3. Recall

It describes the success of the model in recovering information. The formula is in Equation 8.

$$Recall_i = \frac{TP_{ii}}{TP_{1i} + TP_{2i} + \ldots + TP_{ni}} \text{ x } 100\% \tag{8}$$

4. F1 Score

It presents a comparison of the weighted average precision and recall. The formula is in Equation 9.

$$F - 1\ score_i = \frac{Precision_i \ X \ Recall_i}{Precision_i \ X \ Recall_i} \times 2 \tag{9}$$

2.8 Ordinary Kriging Interpolation

After the SVM model predicts land prices in observed locations, Kriging interpolation is applied to estimate land prices in unobserved locations in the Jakarta area. Data visualization serves to facilitate understanding, identify patterns, and communicate information [28]. Kriging makes it possible to combine information from SVM predictions with the spatial structure of the data to produce more accurate estimates, especially in areas where observational data may not be available[29]. This study uses classification and prediction for time *t*+*k*, and ordinary kriging is used for interpolation[30]. The formula of ordinary kriging can be seen in Equation 10. Ordinary Kriging is an interpolation method that assumes that the unobserved values at a location depend on the observed values around that location, with weights assigned based on proximity and variogram[31].

$$\hat{y}_{(t+k)}(s_0) = \sum_{i=1}^{n} \lambda_i{}^{OK} \hat{y}_{(t+k)}(s_i) \tag{10}$$

Description:

$\hat{y}_{(t+k)}(s_0)$: estimate the unknown class

$\hat{y}_{(t+k)}(s_i)$: the estimation is done using class prediction $\hat{y}_{(t+k)}(s_i)$ in neighbouring locations s_i

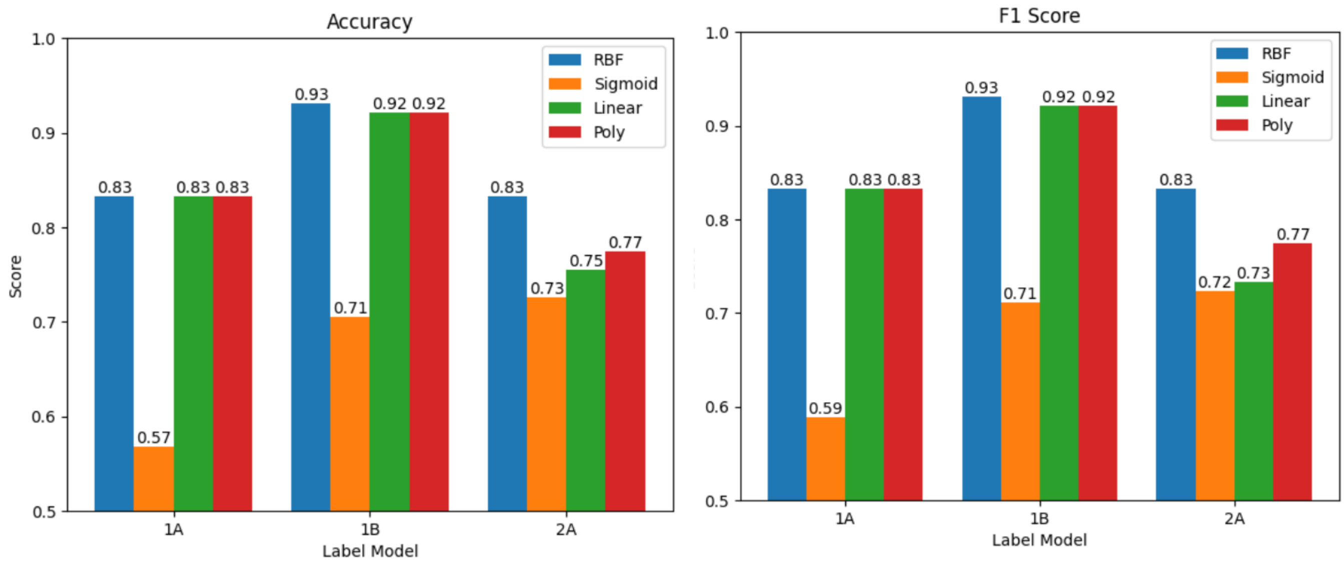
$\lambda_i{}^{OK}$: determine how much influence or contribution a location has

3. Results and Discussion

This section explains the research results obtained and provides a comprehensive discussion. The research results are presented in various forms, such as graphs, tables, and visualizations, to make it easier for readers to understand the data. The discussion is divided in several subsections to provide a structured explanation of each important aspect this research. The results presented include model performance, feature selection, spatial trends in land prices, and interpolation analysis to understand the overall pattern of land prices in Jakarta.

3.1 Time-based Feature Expansion Model Selection

Figure 3 shows the performance of various time-based feature expansion models based on the kernels used, namely RBF, Sigmoid, Linear, and Polynomial. The model was developed with the target class based on data from the previous time points (*t*−1 and *t*−2).



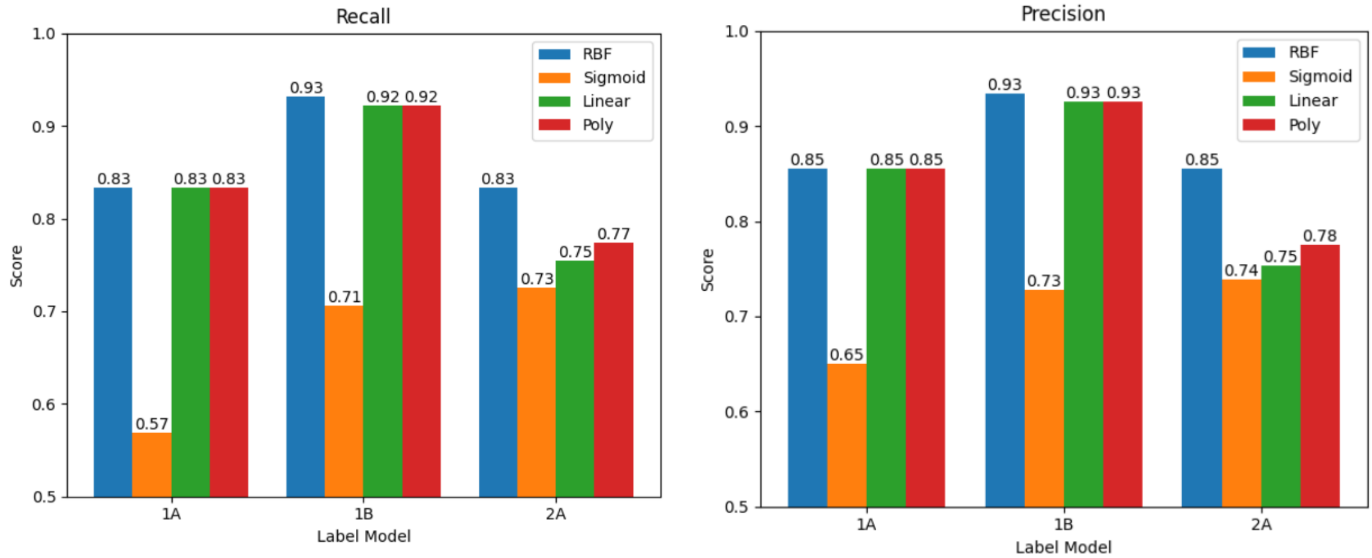


Figure 3. Comparison of SVM Predictive Model

Figure 3 shows the performance of two time-based feature expansion models, namely the *t*−1 model (1B) and the *t*−2 model (2A). These models were developed with the target class based on data from the previous time period.

Table 4. SVM Predictive Model

Kernel	Model	Label	Number of Feature Combinations	Accuracy	F1 Score	Recal	Precision
RBF	*t*−1	1B	6	0.9314	0.931	0.9314	0.9347
	t−2	2A	8	0.833333	0.833362	0.833333	0.854799

From Table 4, the model *t*−1 (1B) shows the best performance, with an accuracy of 93.14%, followed by F1 score and precision, which reach 93.1% and 93.47% respectively. The *t*−2 (2A) model has lower performance than *t*−1 model, with an accuracy of 83.33%. A time-based feature expansion process is performed to predict future land price classification (*t*+*rt* + *rt*+*r*) by utilizing a combination of models from the previous time (*t*−*rt* − *rt*−*r*). This approach aims to design a classification model that can utilize historical information to produce accurate predictions.

3.2 Selection of Influential Features

In this research, feature selection is carried out to improve accuracy and efficiency in predicting land prices in Jakarta. Feature selection aims to identify variables that have the most influence on classification results, resulting in an optimal model. Table 5 shows the features selected for each model.

Table 5. Selected Features

Model	Label	Selected features
t−1	1B	x1, x3, x7, x10, x12, x13
t−2	2A	x6, x7, x9, x16, x18, x22, x1, x2

In Table 5, the Label 2A(*t*−2) section has more than 14 features because, as shown in Table 3, *t*−2 include feature from 2022 and a year of 2023. Therefore, years x1-x14 correspond to 2022, while x15 to x28 correspond to 2023.

3.3 Land Price Classification Results in 2025 and 2026

This research also predicts the classification of land prices by specific locations for the years 2025 and 2026. In Table 10, a value of ‘0’ indicates a low-price category, a value of ‘1’ indicates a medium price category, while a value of ‘2’ indicates an expensive-price category. These predictions aim to give land seekers an idea of the potential price of a particular location.

Table 10. Year Prediction According to Location

Location	2025	2026
Jl. Raya Pondok Kelapa, Duren Sawit, East Jakarta	0	1
Jl. Raya Pondok Kopi, Duren Sawit, East Jakarta	1	2
Jl. Condet Raya, Batu Ampar, Kramat Jati, East Jakarta	1	2
Jl. Letjen T.B. Simatupang, Lingkar Selatan, Tanjung Barat, South Jakarta	1	2
...	...	...
Jl. Cipete Raya, South Jakarta	2	2

3.4 Visualization of Land Price Distribution Classification Prediction in Jakarta

The Kriging interpolation method was used to visualize the changes in land prices over time. In Figure 4, data from 2022 and 2024 is presented to highlight the differences. The results of the visualization are shown in Figure 4, which is a map showing the distribution of land prices for the years 2025 and 2026. The colors on the map represent the price categories. Based on the visualization of the distribution of land prices in Jakarta using the Kriging method in Figure 4, there is a significant change in the distribution pattern from 2023 to 2026. In 2023, areas with expensive land prices (shown in red) were concentrated in central Jakarta, while areas with cheap land prices (shown in green) dominated the periphery. Entering 2024, there is a consolidation and moderate expansion of expensive zones in the central area, followed by an increase in medium-priced areas (shown in yellow) that begin to replace some of the previous cheap zones. Predictions for 2025 show a significant expansion of expensive areas spreading out in different directions, with an increase in mid-price areas mainly on the western side of Jakarta. By 2026, the high-priced zones reach their maximum expansion and dominate the central to eastern areas of Jakarta, while the medium-priced areas further expand to replace the remaining cheap zones. This pattern of change indicates a trend of decreasing availability of affordable land and increasing dominance of high-value properties in Jakarta over the four-year period.

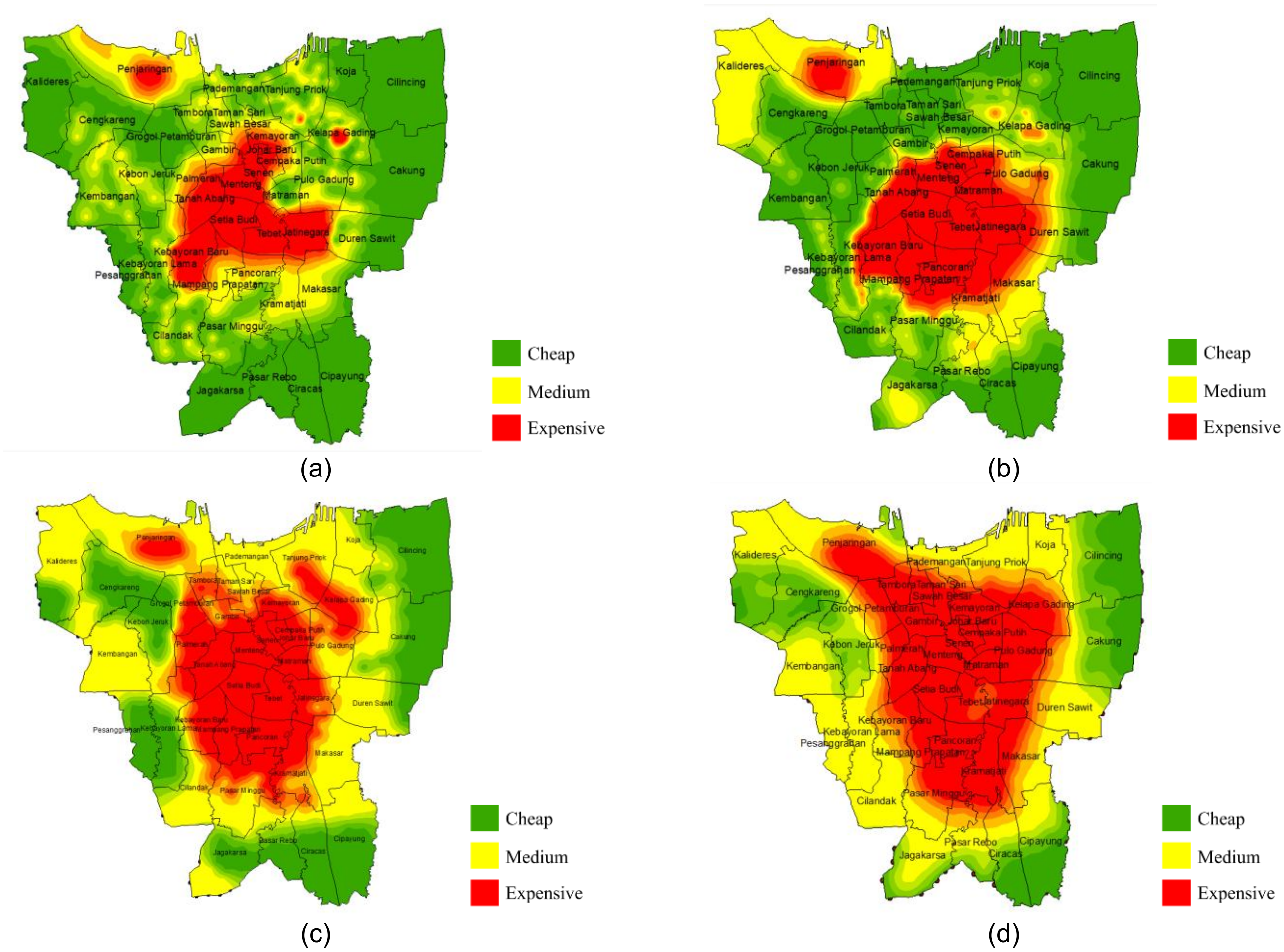


Figure 4. Predicted Land Price Distribution Classification of Jakarta, a) 2023, b) 2024, c) 2025, d) 2026

3.6 Discussion

In this study, the Support Vector Machine (SVM) model was tested using different kernels—specifically RBF, Sigmoid, Linear and Polynomial—to determine the most optimal kernel. Based on Figure 3, the RBF kernel showed the best performance, with the *t*-1 model on label 1B achieving the highest accuracy of 93.14%, an F1 score of 93.1%, and a precision of 93.47%. In contrast, model *t*-2 on label 2A produced an accuracy of 83.33%, which is lower than that of model *t*-1. Based on these results, the RBF kernel was selected for further testing, where the *t*-1 model (1B) was used for the 2025 prediction, while the *t*-2 model (2A) was used for the 2026 prediction, as it was the only model in *t*-2 with an accuracy of 83.33%.

In comparing research on land price prediction, several studies, such as [6], [7], [8], [32], [33], [34], and [35]used various methods with quite diverse accuracy results. A study in [8], for example, showed that the linear regression model had 'good enough' accuracy, as it was above 50%, but did not provide clear details regarding the accuracy value, F1 score, or other evaluation indicators. This makes it difficult to objectively assess the effectiveness of the model. In contrast, this study used a Support Vector Machine (SVM) with Kriging interpolation, which not only produced an accuracy of 93.14%, but also included additional evaluation metrics such as F1 score (0.931), precision (93.47%), and recall (93.14%).

In this research, the RBF kernel has higher accuracy. This is consistent with other studies that the RBF kernel has higher accuracy compared to other kernels. For instance, in a study[36], the RBF kernel outperformed the polynomial kernel, and in another study [37] the RBF kernel was found to be better than the linear kernel. With this approach, this research provides greater clarity in evaluating model performance and makes a significant contribution to predicting and visualizing the distribution of land prices in Jakarta in an in-depth and comprehensive manner.

Land price prediction is performed using historical data consisting of the previous two years to the next two years, as well as features selected based on their relevance and contribution to model accuracy. The feature selection process was conducted to improve computational efficiency without compromising the performance of the SVM model. By selecting relevant features, the model can become more accurate by avoiding the use of useless or redundant information[38]. Feature selection helps simplify the model by removing unimportant features, making the model easier to understand and interpret[39]. As a result of the selection, the features used for prediction on label 1B are x1, x3, x7, x10, x12 and x13, while for label 2A they are x6, x7, x9, x16, x18, x22, x1 and x2. A full description of each feature can be found in Table 1, which provides details of the variables associated with the data. To predict one year ahead (2025), the *t*-1 model uses data from one year earlier. For example, the *t*-1 model (1A) uses land price data from 2024 and selected characteristics from 2023. Meanwhile, the *t*-1 model (1B) for the prediction of the same year uses land price data from 2023 and selected characteristics from 2022. For the prediction two years ahead (2026), the *t*-2 model (2A) combines the 2024 land price data with the selected features from 2022 and 2023. This process demonstrates how historical data and feature selection play an important role in strengthening the predictive ability of SVM models that effectively integrate temporal and spatial information.

After determining the appropriate model and features, the SVM with the RBF kernel was re-run to predict land prices in Jakarta in 2025 and 2026. In this process, the selected features were used to ensure a more accurate and efficient prediction than the initial experiment using all features. This approach results in predictions that can be classified into three classes: '0' for low prices, '1' for medium prices and '2' for high prices. These classification results are presented in detail in Table 10, which shows the distribution of predictions for each class by geographical location. For the 2025 prediction, the *t*-1 model (1B) shows consistent results with the highest accuracy among the other models. These predictions provide a good insight into the distribution of land prices in that year, especially in areas where significant price increases are expected. Meanwhile, for the 2026 prediction, the *t*-2 model (2A) produced results close to high accuracy by using a combination of selected features and historical data. This method allowed the researchers to identify land price distribution patterns over a longer period, which is crucial for understanding medium-term trends. In addition, this classification makes it easier for the public and decision-makers to identify areas with cheap, medium, or expensive land prices more intuitively.

The results of the land price prediction were visualized using the ordinary kriging method to map the spatial distribution of land prices in Jakarta in different years. Using ordinary kriging, the prediction of values at unmeasured locations is more accurate because it takes into account the variability and spatial pattern of the data[40]. The visualization in Figure 4 shows land price distribution maps for the years 2023, 2024, 2025 and 2026. The map shows that the areas colored red and yellow, representing medium to expensive land prices, are expanding each year. In contrast, the green areas, which represent cheap land prices, appear to be declining. This trend shows that land prices in Jakarta will continue to rise, in line with the dynamics of high demand and limited available land. Compared to 2023 and 2024, the map shows a significant increase in the number of areas with high land prices, especially in strategic areas or business centers. The years 2025 and 2026 continue this trend, with an even more drastic reduction in areas with low land prices. This visualization not only illustrates the changes in land prices over time but also provides valuable information for policymakers and the public to understand the distribution patterns of land prices by location. In this way, the results of this research can support more targeted decision-making in investment planning, infrastructure development and land acquisition strategies. The combination of spatial visualization and data-based forecasting shows that the use of ordinary kriging is very effective in identifying and describing the dynamics of land prices geographically. In addition to helping the public to obtain transparent price information, this approach also contributes to the development of a more inclusive system, allowing all stakeholders to understand land price movements more easily and accurately.

4. Conclusion

This research successfully identified the best method for classification and prediction of land price distribution in Jakarta using a Support Vector Machine (SVM) model with an RBF kernel. Based on the test results, the RBF kernel showed the highest performance with 93.14% accuracy, 93.1% F1 score, and 93.47% precision, as achieved by model $t-1$ (1B). This model was used to predict land prices in 2025, while the $t-2$ (2A) model was used to predict 2026, with an accuracy of 83.33%. This approach involves utilizing one to two years of historical data and systematically selected features, ensuring more accurate and relevant predictions. Visualization of the prediction results using the ordinary kriging method shows changes in the distribution pattern of land prices in Jakarta. The more dispersed red and yellow colors in the visualization indicate that cheap land prices in Jakarta are decreasing, while expensive land prices are becoming more dominant. This research provides an advantage over previous studies by applying SVM and kriging interpolation, which not only produces high accuracy but also includes thorough evaluation through metrics such as F1 score, precision, and recall. With this approach, this research makes a significant contribution to predicting and visualizing the distribution of land prices in Jakarta in an in-depth and comprehensive manner.

References

[1] I. Fadhlurrahman, "29% Penduduk DKI Jakarta ada di Kota Jakarta Timur pada Pertengahan 2024," databoks.
[2] Jakarta.bps.go.id, "Luas Daerah Menurut Kabupaten/Kota," Jakarta.
[3] bps.go.id, "Garis Kemiskinan, Jumlah, dan Persentase Penduduk Miskin di Daerah Menurut Kabupaten/Kota di Provinsi DKI Jakarta, 2022-2023,".
[4] C. D. Putra, A. Ramadhani, and E. Fatimah, "Increasing Urban Heat Island area in Jakarta and it's relation to land use changes," *IOP Conference Series: Earth and Environmental Science*, vol. 737, no. 1, 2021. https://doi.org/10.1088/1755-1315/737/1/012002
[5] B. A. Bonci, E. Mccoy, and J. Cole, "A research review : the importance of families and the home environment," no. March, 2011.
[6] N. A. Chandrasa, S. S. Prasetyowati, and Y. Sibaroni, "Land Price Classification Map in Jakarta Using Random Forest and Ordinary Kriging," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 2, 2022. https://doi.org/10.47065/bits.v4i2.1896
[7] A. N. Rais, I. Alfarobi, S. W. Hadi, and W. Kurniawan, "Analisa Prediksi Harga Jual Rumah Menggunakan," vol. 6, no. 2, pp. 417–423, 2024.
[8] N. A. Az-zahra and D. B. Arianto, "Analisis Perbandingan Algoritma Regresi Linear dan Decision Tree pada Prediksi Harga Rumah," pp. 2–6.
[9] C. Shousong, G. Xiaomin, W. Xiaoguang, and C. Ying, "Notice of Removal: Research on Urban Land Price Assessment Based on Artificial Neural Network Model," *IEEE Access*, vol. 7, no. June 2018, pp. 180738–180748, 2019. https://doi.org/10.1109/ACCESS.2019.2958978
[10] A. H. Suhendar, A. A. Rohmawati, and S. S. Prasetyowati, "Performance of CART Time-Based Feature Expansion in Dengue Classification Index Rate," *Sinkron*, vol. 9, no. 1, pp. 1–9, 2024. https://doi.org/10.33395/sinkron.v9i1.13023
[11] S. Murni, D. Widiyanto, and C. N. P. Dewi, "Klasifikasi Citra Penyakit Daun Kopi Arabika Menggunakan Support Vector Machine (SVM) Dengan Seleksi Fitur Information Gain," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 700–709, 2022.
[12] Y. I. Mukti, "Sistem Prediksi Lulus Tepat Waktu Tugas Akhir Mahasiswa Menggunakan Support Vector Machine (Svm)," *JUTIM (Jurnal Teknik Informatika Musirawas)*, vol. 5, no. 2, pp. 110–115, 2020. https://doi.org/10.32767/jutim.v5i2.1050
[13] U. Amelia, J. Indra, and A. F. N. Masruriyah, "Implementasi Algoritma Support Vector Machine (Svm) Untuk Prediksi Penyakit Stroke Dengan Atribut Berpengaruh," *Scientific Student Journal for Information, Technology and Science*, vol. III, no. 2, pp. 254–259, 2022.
[14] D. Sitanggang and S. Sherly, "Model Prediksi Obesitas dengan Menggunakan Support Vector Machine," *Jurnal Sistem Informasi dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 172–175, 2022. https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2443
[15] I. Zulfahmi, H. Syahputra, S. I. Naibaho, M. A. Maulana, and E. P. Sinaga, "Perbandingan Algoritma Support Vector Machine (SVM) dan Decision Tree Untuk Deteksi Tingkat Depresi Mahasiswa," *Bina Insani Ict Journal*, vol. 10, no. 1, p. 52, 2023. https://doi.org/10.51211/biict.v10i1.2304
[16] M. D. Purbolaksono, M. Irvan Tantowi, A. Imam Hidayat, and A. Adiwijaya, "Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 2, pp. 393–399, 2021. https://doi.org/10.29207/resti.v5i2.3008
[17] M. Fluorida Fibrianda and A. Bhawiyuga, "Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 9, pp. 3112–3123, 2018.
[18] K. A. Rokhman, B. Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *Journal of Information System Management (JOISM)*, vol. 3, no. 1, pp. 1–7, 2021. https://doi.org/10.24076/joism.2021v3i1.341
[19] J. A. Nursiyono and D. M. Dewi, "Determinan Harga Tanah di Indonesia Menggunakan Big Data (Studi Kasus: www.lamudi.co.id)," *Jurnal Pertanahan*, vol. 11, no. 2, 2021. https://doi.org/10.53686/jp.v11i2.105
[20] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A Review on Data Preprocessing Techniques Toward Ef fi cient and Reliable Knowledge Discovery From Building Operational Data," vol. 9, no. March, pp. 1–17, 2021. https://doi.org/10.3389/fenrg.2021.652801
[21] P. Ferreira, D. C. Le, and N. Zincir-Heywood, "Exploring Feature Normalization and Temporal Information for Machine Learning Based Insider Threat Detection," *15th International Conference on Network and Service Management, CNSM 2019*, 2019. https://doi.org/10.23919/CNSM46954.2019.9012708
[22] S. S. Prasetiyowati and Y. Sibaroni, "Unlocking the potential of Naive Bayes for spatio temporal classification: a novel approach to feature expansion," *Journal of Big Data*, vol. 11, no. 1, 2024. https://doi.org/10.1186/s40537-024-00958-x

[23] M. Athoillah and R. K. Putri, "Handwritten Arabic Numeral Character Recognition Using Multi Kernel Learning Support Vector Machine," vol. 4, no. 2, pp. 99–106, 2019. https://doi.org/10.22219/kinetik.v4i2.724
[24] H. R. Baghaee, D. Mlakic, S. Nikolovski, and T. Dragicevic, "Support Vector Machine-Based Islanding and Grid Fault Detection in Active Distribution Networks," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 3, pp. 2385–2403, 2020. https://doi.org/10.1109/JESTPE.2019.2916621
[25] C. Li, K. Liu, and H. Wang, "The incremental learning algorithm with support vector machine based on hyperplane-distance," *Applied Intelligence*, vol. 34, no. 1, pp. 19–27, 2011. https://doi.org/10.1007/s10489-009-0176-9
[26] S. Ding, X. Hua, and J. Yu, "An overview on nonparallel hyperplane support vector machine algorithms," pp. 975–982, 2014. https://doi.org/10.1007/s00521-013-1524-6
[27] L. G. May, "Classification".
[28] C. Dooley, "Data visualisation and machine learning web application with potential use in sports data analytics," 2017.
[29] D. A. S. Kali, "Pemetaan Sebaran Hujan Rancangan Menggunakan Interpolasi," vol. 04, no. 01, pp. 239–249, 2024.
[30] M. D. Muldani, S. S. Prasetiyowati, and Y. Sibaroni, "Air Pollution Classification Prediction Model with Deep Neural Network based on Time-Based Feature Expansion and Temporal Spatial Analysis," vol. 6, no. 2, pp. 867–877, 2024. https://doi.org/10.47065/bits.v6i2.5675
[31] K. Novak Zelenika, R. Vidaček, T. Ilijaš, and P. Pavić, "Application of deterministic and stochastic geostatistical methods in petrophysical modelling – a case study of upper pannonian reservoir in sava depression," *Geologia Croatica*, vol. 70, no. 2, pp. 105–114, 2017. https://doi.org/10.4154/gc.2017.10
[32] S. Pratama, "Prediksi Harga Tanah Menggunakan Algoritma Linear Regression," *Technologia : Jurnal Ilmiah*, vol. 7, no. 2, pp. 122–130, 2016. https://doi.org/10.31602/tji.v7i2.624
[33] M. L. Mu'tashim, T. Muhayat, S. A. Damayanti, H. N. Zaki, and R. Wirawan, "Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression," *Informatik : Jurnal Ilmu Komputer*, vol. 17, no. 3, p. 238, 2021. https://doi.org/10.52958/iftk.v17i3.3635
[34] A. Saiful, "Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 1, pp. 41–50, 2021. https://doi.org/10.35957/jatisi.v8i1.701
[35] W. Yustanti, "Nihru Nafi' Dzikrulloh, Indriati, Budi Darma Setiawan 2017," *Jurnal Matematika statistika dan komputasi*, vol. 9, no. 1, pp. 57–68, 2012.
[36] M. Ma, S. Mei, S. Wan, Z. Wang, and D. Feng, "Video summarization via nonlinear sparse dictionary selection," *IEEE Access*, vol. 7, no. December, pp. 11763–11774, 2019. https://doi.org/10.1109/ACCESS.2019.2891834
[37] V. Apostolidis-Afentoulis, "SVM Classification with Linear and Rbf Kernels," *ResearchGate*, no. July, pp. 0–7, 2015. https://doi.org/10.13140/RG.2.1.3351.4083
[38] S. Wang, J. Tang, and H. Liu, "Encyclopedia of Machine Learning and Data Science," *Encyclopedia of Machine Learning and Data Science*, no. October 2017, 2020. https://doi.org/10.1007/978-1-4899-7502-7
[39] Y. Akhiat and S. Amjad, "Feature Selection : A Review and Comparative Study Feature Selection : A Review and Comparative Study," no. May, 2022. https://doi.org/10.1051/e3sconf/202235101046
[40] T. Jurnal, S. Dan, F. Pohan, and R. Adi, "Ordinary kriging dalam penentuan lama penggalian tambang terbuka," vol. 15, no. 2, pp. 130–136, 2019.