International Journal Software Engineering and Computer Science (IJSECS) 3 (3), **2023, 261-269**
Published Online December 2023 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN : 2776-4869, E-ISSN : 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v3i3.1709.

K-Nearest Neighbor for Gorontalo City Chili Price Prediction Using Feature Selection, Backward Elimination, and Forward Selection

Abdul Yunus Labolo
Universitas Ichsan Gorontalo, Gorontalo City, Gorontalo Province, Indonesia.
E-mail: abdulyunuslabolo@gmail.com

Siti Andini Utiarahman
Universitas Ichsan Gorontalo, Gorontalo City, Gorontalo Province, Indonesia.
E-mail: andiniutiarahman@gmail.com

Moh. Efendi Lasulika
Universitas Ichsan Gorontalo, Gorontalo City, Gorontalo Province, Indonesia.
E-mail: fendidsn.ui@gmail.com

Ivo Colanus Rally Drajana
Universitas Pohuwato, Gorontalo City, Gorontalo Province, Indonesia.
E-mail: ivocolanusrally@gmail.com

Andi Bode *
Universitas Ichsan Gorontalo, Gorontalo City, Gorontalo Province, Indonesia.
E-mail: andibode22@gmail.com

Received: 28 October 2023; Accepted: 17 November 2023; Published: 1 December 2023.

Abstract: This study addresses chili price volatility, an important concern that impacts the national economy and societal welfare. Fluctuations in chili prices in the retail market greatly influence market demand, thereby influencing farming decisions, especially chili cultivation. To help make better decisions, Researchers use forecasting, which is defined as the projection of future trends based on the analysis of historical data, using statistical methods. The K-Nearest Neighbor (K-NN) algorithm is used because of its resistance to high noise on large training datasets. However, challenges arise in determining the optimal value of 'k' and selecting related attributes. To overcome this, Feature Selection is applied to refine the model by removing irrelevant features, resulting in a significant reduction in the model error rate. This improvement indicates an increase in the efficiency of the K-NN algorithm with the incorporation of Feature Selection. Our findings show that the model, with backward elimination in Feature Selection, achieves a Root Mean Square Error (RMSE) of 0.202, outperforming the model using forward selection. The prediction accuracy of this model reaches an average of 78.86%, which is much higher than the baseline data of 50%. This shows the success of the proposed method in predicting chili prices.

Keywords: Chili Price Prediction; K-Nearest Neighbor (K-NN); Feature Selection; Market Demand; Forecasting Accuracy.

1. Introduction

Cayenne pepper, a vegetable with significant demand in both domestic and international markets, has a particularly strong market in Gorontalo. This demand drives farmers to adapt their cultivation to align with consumer preferences, embodying the principle of market-oriented product planning as emphasized by Lim *et al*. [1]. Farmers rely on market demands to decide which variety to cultivate, predominantly choosing between cayenne pepper and its local variants. Notably, the choice often extends to varieties with longer shelf lives at room temperature, which aligns with market dynamics in terms of pricing and accessibility. However, the chili market in Gorontalo City is marked by considerable price instability, often showing a tendency to increase, which carries negative implications. This volatility is somewhat analogous to fuel markets, where price changes in retail markets affect consumer demand. Additionally, supply-side factors, including transportation challenges, fuel consumption, and purchasing power, also play a significant role. This is evidenced by the lower prices in production areas compared to consumption zones.

Table 1. Chili Price 2018-2020.

Month	Year		
	2018	2019	2020
January	-	Rp. 34.000	Rp. 75.000
February	Rp. 65.000	Rp. 50.000	Rp. 80.000
March	Rp. 60.000	Rp. 66.000	Rp. 60.000
April	Rp. 68.000	Rp. 120.000	Rp. 80.000
May	Rp. 65.000	Rp. 80.000	Rp. 40.000
June	Rp. 55.000	Rp. 60.000	Rp. 40.000
July	Rp. 80.000	Rp. 70.000	Rp. 35.000
August	Rp. 68.000	Rp. 70.000	Rp. 40.000
September	Rp. 68.000	Rp. 90.000	Rp. 40.000
October	Rp. 60.000	Rp. 120.000	Rp. 58.000
November	Rp 60.000	Rp. 120.000	Rp. 50.000
December	-	Rp. 70.000	-

Table 1 presents the fluctuating chili prices in Gorontalo City over the last three years (2018-2020), highlighting the difficulty faced by the Gorontalo City Food Service in making accurate price predictions for future months. The use of forecasting, an integral part of the decision-making process, involves statistical methods to predict future trends based on historical data [2]. Several methods, including Linear Regression, C-45, Naïve Bayes, K-NN, and SVM, are used in forecasting. This study specifically employs the K-Nearest Neighbor (K-NN) algorithm, known for its stability and effectiveness in handling large, noisy datasets. However, K-NN faces challenges in determining the optimal 'k' value and in effective feature selection, which is critical for enhancing its predictive accuracy [3]. Wanto and Windarto (2017) define forecasting as a scientific approach to estimate or predict future events based on historical data [4]. Several studies have applied different methods to predict various commodity prices. For instance, Fikri Nur Hardiansyah (2017) used the ARIMA method for chili price prediction, achieving varied levels of accuracy based on RMSE and MAPE values [5]. Similarly, Fatkhuroji *et al*. (2019) employed the Support Vector Machine method with forward selection for soybean price prediction, demonstrating improvements in prediction accuracy [6]. Prabowo B.U., *et al*. (2019) combined the K-NN method with linear regression to predict gold prices, resulting in improved predictive performance [7]. M. Efendi L and Andi Bode (2021) compared feature selection methods in corn price prediction, finding that feature selection significantly reduced the RMSE values [8]. These studies highlight the importance and effectiveness of feature selection in forecasting. This research aims to extend this approach to chili price prediction in Gorontalo City, comparing the impact of Forward Selection and Backward Elimination on the K-NN model's performance. The findings will provide valuable insights for the Gorontalo City Food Service, enhancing their capacity to anticipate market changes and plan accordingly.

2. Research Method

The K-Nearest Neighbor method is one of the ten most popular K-NN algorithms for finding the set of k nearest objects from the training data and organizing the data into groups of objects with a high degree of similarity. The method of determining similarity is based on the results of calculating the smallest distance between objects [9]. Various feature selection methods are used to select irrelevant features, such as forward selection, backward elimination, and optimize selection. The Backward Elimination method aims to improve the performance of the model through a working system by selecting the last one, another option is done by checking all types before selection and eliminating incompatible types [10]. The first option is used to select the inactive part when you start the loop, then this part is added to the selected part [11].

2.1. Research Methods

This study employs the K-Nearest Neighbor (K-NN) algorithm, recognized as one of the top ten algorithms for classifying data based on similarity. The K-NN method involves identifying a set of 'k' nearest objects from the training data and grouping them based on their high similarity. The determination of similarity within this method is based on calculating the minimal distance between the data objects. To enhance the K-NN algorithm's performance, various feature selection methods are utilized. These include Forward Selection, Backward Elimination, and Optimize Selection. Specifically, Backward Elimination focuses on improving the model by initially including all features and subsequently eliminating those that are less compatible or irrelevant. In contrast, Forward Selection starts by selecting an inactive subset of features and iteratively adding to this subset.

2.2. Proposed Methods

The primary objective of this research is to apply the K-Nearest Neighbor algorithm, integrating the feature selection techniques of Backward Elimination and Forward Selection, to predict chili prices in Gorontalo City. The dataset for this study is sourced from the Gorontalo City Food Service.

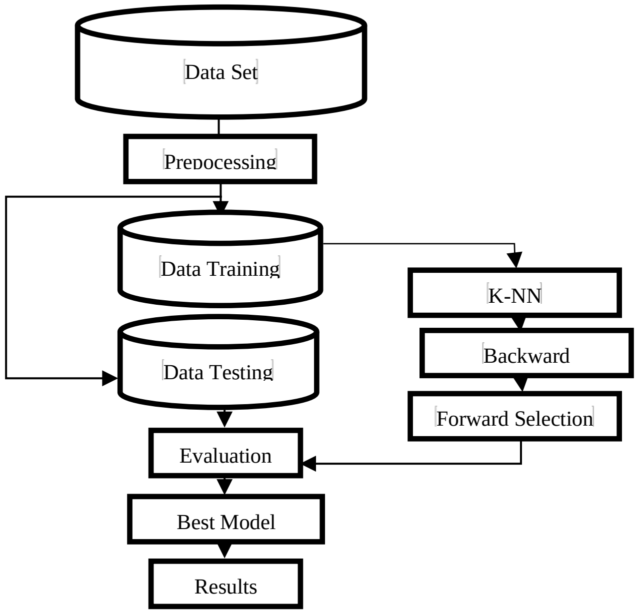

Figure.1. Experimental Design

From the flow of the experiment, the first is data collection, then the data set obtained will then be preprocessed, after preprocessing, the data is then divided into two data sets, namely training data and test data, after dividing the data, the next step is to enter the K-NN algorithm process stage, then K-NN using Backward Elimination and Forward Selection, after which the best model evaluation will be used to obtain prediction results. In the process of K-NN and K-NN - Backward Elimination and Forward Selection algorithms, the training data trials start from 1 to 4 periods, this is to get a good model so that a smaller RMSE value is obtained.

3. Result and Discussion

3.1 Results

The initial stage involves preprocessing the data collected from the Gorontalo City Food Service, which comprises daily data from 2018-2020. This data is first transformed from monthly to daily multivariate time series format, facilitating easier processing in RapidMiner tools.

Chili Price Data

KOMODITI	SATUAN	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑	↓
CABE RAWIT	KG	17-Feb-2018		18-Feb-2018		19-Feb-2018		20-Feb-2018		21-Feb-2018		22-Feb-2018		23-Feb-2018		24-Feb-2018	
		60.000	50.000	**55.000**	**50.000**	**55.000**	50.000	**55.000**	50.000	55.000	50.000	60.000	56.000	60.000	56.000	**60.000**	56.000
CABE RAWIT	KG	25-Feb-2018		26-Feb-2018		27-Feb-2018		28-Feb-2018									
		60.000	**55.000**	**60.000**	55.000	**65.000**	60.000	65.000	65.000								
CABE RAWIT	KG	1-Mar-2018		2-Mar-2018		3-Mar-2018		4-Mar-2018		5-Mar-2018		6-Mar-2018		7-Mar-2018		8-Mar-2018	
		60.000	60.000	60.000	55.000	**60.000**	50.000	**60.000**	**50.000**	**60.000**	**50.000**	**55.000**	50.000	60.000	54.000	55.000	50.000
CABE RAWIT	KG	9-Mar-2018		10-Mar-2018		11-Mar-2018		12-Mar-2018		13-Mar-2018		14-Mar-2018		15-Mar-2018		16-Mar-2018	
		55.000	50.000			**60.000**	**50.000**	**55.000**	**50.000**	55.000	50.000	**50.000**	45.000	**45.000**	**40.000**	50.000	45.000
CABE RAWIT	KG	17-Mar-2018		18-Mar-2018		19-Mar-2018		20-Mar-2018		21-Mar-2018		22-Mar-2018		23-Mar-2018		24-Mar-2018	
		50.000	40.000	50.000	40.000	50.000	40.000	40.000	36.000	36.000	36.000	40.000	35.000	40.000	36.000	40.000	38.000
CABE RAWIT	KG	25-Mar-2018		26-Mar-2018		27-Mar-2018		28-Mar-2018		29-Mar-2018		30-Mar-2018		31-Mar-2018			
		40.000	38.000	40.000	35.000	40.000	35.000	**40.000**	35.000	**50.000**	**40.000**			**60.000**	55.000		
CABE RAWIT	KG	1-Apr-2018		2-Apr-2018		3-Apr-2018		4-Apr-2018		5-Apr-2018		6-Apr-2018		7-Apr-2018		8-Apr-2018	
		60.000	55.000	56.000	50.000	50.000	50.000	**60.000**	55.000	**55.000**	**55.000**	**68.000**	60.000	**65.000**	65.000	60.000	56.000
CABE RAWIT	KG	9-Apr-2018		10-Apr-2018		11-Apr-2018		12-Apr-2018		13-Apr-2018		14-Apr-2018		15-Apr-2018		16-Apr-2018	
		60.000	58.000	54.000	50.000	55.000	50.000	**55.000**	54.000	**55.000**	**54.000**			**55.000**	54.000		

Figure 2. Sample data of chilli price

The next step involves converting the time series data into a multivariate format. This process begins with reordering the data from ascending to descending order using the Sort & Filter feature in Microsoft Office Excel. Subsequently, the data is structured into 4 period variables, as shown in Table 2.

Table 2. 4 Period Multivariate

Xt-4	Xt-3	Xt-2	Xt-1	Xt
0,333333	0,333333	0,483333	0,416667	0
0,283333	0,333333	0,333333	0,483333	0,416667
0,333333	0,283333	0,333333	0,333333	0,483333
0,333333	0,333333	0,283333	0,333333	0,333333
0,666667	0,333333	0,333333	0,283333	0,333333
0,5	0,666667	0,333333	0,333333	0,283333
0,666667	0,5	0,666667	0,333333	0,333333
0,625	0,666667	0,5	0,666667	0,333333
0,583333	0,625	0,666667	0,5	0,666667
1	0,583333	0,625	0,666667	0,5

From the table above, we can see that the data has passed the data normalization stage. In the database technology center or database normalization aims to avoid the occurrence of various errors and data inconsistencies. Then normalization is also done to find data with the minimum size to represent the original data without losing the characteristics of the data. The normalization equation can be seen below:

$$\text{Normalisation} = \frac{(\text{X}-\text{Min})}{(Max-Min)} \quad (1)$$

Where:
x = Data
Min = Minimum Data
Max = Maximum Data

Then, to restore the data to its original size, the data set is denormalized. Denormalization is applied to the experimental results of the data analysis in the form of chili price estimates. The denormalization equation is as follows :

$$\text{Denormalisasi} = Y(Max - Min) + Min \quad (2)$$

Where:
Y = Output Results of the Training
Min = Minimum Data
Max = Maximum Data

The source of data set in this research was taken from Gorontalo City Food Service, Jl. Nani Wartabone No. 03, Gorontalo 96133, Gorontalo City. The data collected is univariate time series daily quantitative data. This type of data is chili price data from 2018-2020.

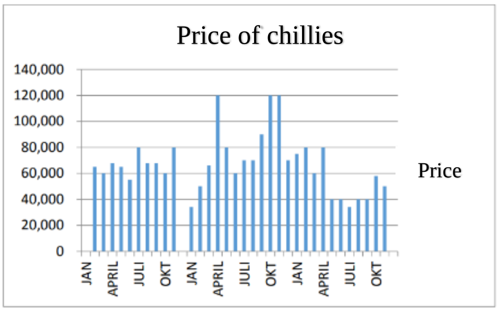

Figure 3. Chili Price Chart

In the next step, after the data is ready to be processed, the data processing experiment is ready to be run. For this data processing the RapidMiner tools are used.

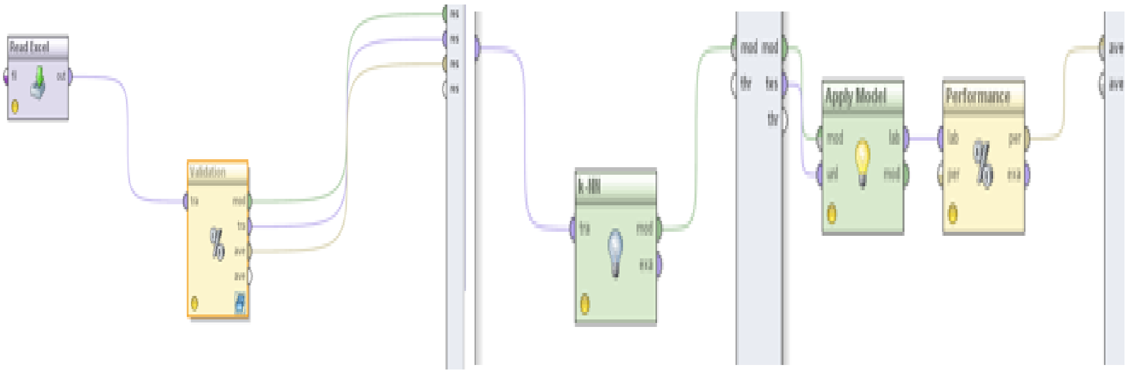

Figure 4. Display of RapidMiner Tools

3.1.1. K-Nearest Neighbour Parameters

At this point, find the shortest distance to k or the number of nearest neighbors, in the training method, that is, separate the training data from the test data, then test the method.

Table 3. Test Results on K-NN Algorithm

Period	Validation	K	RMSE
1	10	3	0,228
2	10	3	0,223
3	10	3	0,238
4	10	3	0,249

The table above is the result of experiments using the k-nearest neighbor model at cross-validation 10 and k-value 3, which is seen based on the root mean square error (RMSE) value. From the results, the best model based on the least error is in variable period 2 with an RMSE value of 0.223.

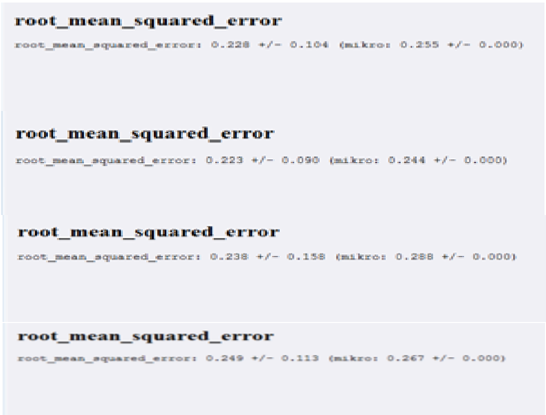

Figure 5. Display of K-NN model processing results

3.1.2. K-Nearest Neighbor Parameters Using Backward Elimination

At this stage, experiments were conducted using the k-nearest neighbor model selection function with backward elimination.

Table 4. Test Results on K-NN and Backward Elimination Algorithms

Period	Validation	K	RMSE
1	10	3	0,228
2	10	3	0,202
3	10	3	0,220
4	10	3	0,202

The table above is the result of experiments using the k-nearest neighbor model with backward elimination on cross-validation 10 and k-value 3, which is seen based on the root mean square error (RMSE) value. From the results, the best model based on the least error is in the two period variables, namely periods 2 and 4 with the same RMSE value of 0.202.

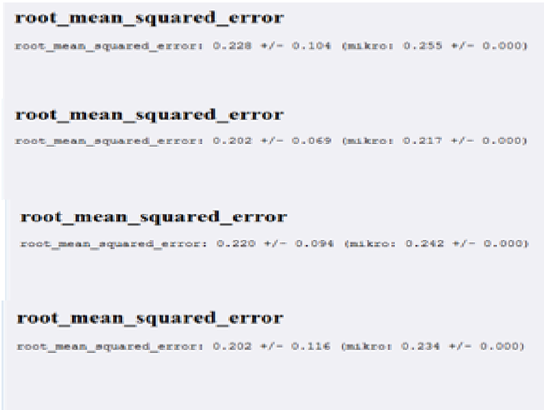

Figure 6. Display of K-NN and Backward Elimination Model Processing Results

3.1.3. K-Nearest Neighbor Parameters Using Forward Selection

At this stage, experiments were conducted using the k-nearest neighbor feature selection model with forward selection.

Table 5. Test Results on K-NN and Forward Selection Algorithms

Period	Validation	K	RMSE
1	10	3	0,218
2	10	3	0,215
3	10	3	0,208
4	10	3	0,213

The table above is the result of experiments using the k-nearest neighbor model with forward selection on cross-validation 10 and k-value 3, which is seen based on the root mean square error (RMSE) value. From the results, the best model based on the least error is the variable period 3 with an RMSE value of 0.208.

Figure 7. Display of K-NN Model Processing Results and Forward Selection

3.1.4. Evaluation

In the evaluation stage, the best model is selected based on the root mean square error (RMSE) value obtained. Based on the results of the experiments conducted, the least error value is used to predict the price of chili.

Table 6. Comparison of K-NN and Feature Selection

Model	Period	Validation	RMSE
K-NN	2	3	0,223
K-NN, Backward Elimination	3	3	0,202
K-NN, Forward Selection	4	3	0,208

The table above summarizes the comparison results of the K-Nearest Neighbor model using the Forward Selection and Backward Elimination selection features. The table explains that the addition of selection features can improve the performance of the algorithm by producing the least error value compared to without the use of selection features. From the existing results, it explains that the best model based on the least error value is the K-NN model using Backward Elimination with the least RMSE value of 0.202.

3.1.5. Implementation

At this implementation stage, the best k-nearest neighbor model is applied using the backward elimination selection function. From the test experiments conducted using the 5-record model to predict chili prices, the prediction results can be seen in the following table.

Table 7. Prediction Results for the Next Month

Month	Price
January	0,333333
February	0,522222
March	0,522222
April	0,288889
May	0,522222

According to the table above, here are the results of the forecasts for the next 5 months January - May 2021. The comparison of the percentage of actual data and prediction results can be seen in the table below:

Table 8. Percentage Comparison of Jan-May 2021

Month	Prediction Result	Actual Data	Percentage
Jan	0,333333	0,5	66,67%
Feb	0,522222	0,566667	92,15%
Mar	0,522222	0,566667	92,15%
Apr	0,288889	0,666667	43,33%
Mei	0,522222	0,458333	100%

According to the table above, the results of comparing the prediction results of chili prices with the chili price data at Gorontalo City Food Service January-May 2021. The prediction results produce an average accuracy percentage value of 78.86%, thus the prediction results are declared successful seeing from the accuracy of the resulting percentage above 50%. At the time of data preprocessing, the existing dataset has gone through the normalization data stage, so to find out the predicted price of chili prices in terms of rupiah (original data), the next step is to restore the data with the denormalization process. Data denormalization data can be seen in the table below:

Table 9. Prediction Results in Rupiah

Month	Price
January	Rp. 40.000
February	Rp. 62.667
March	Rp. 62.667
April	Rp. 34.000
May	Rp. 62.667

The table above is the result of predictions for the next 5 months in 2021, which have been denormalized so that the data can be seen in the form of rupiah or actual data.

3.2 Discussion

In this study, the K-Nearest Neighbor (K-NN) algorithm, along with feature selection techniques such as Backward Elimination and Forward Selection, was employed to predict chili prices in Gorontalo City. The results demonstrate the effectiveness of these methods in handling the challenges posed by the volatile nature of chili prices. The integration of feature selection methods significantly enhanced the K-NN algorithm's performance. Backward Elimination proved to be more effective, as evidenced by the lowest RMSE values obtained. This suggests that the process of starting with a full set of features and then eliminating the least significant ones allows for a more optimized model. Conversely, Forward Selection, which begins with no features and adds them iteratively, also improved the model's accuracy but to a lesser extent than Backward Elimination. The comparative analysis between Backward Elimination and Forward Selection highlights the importance of feature selection in predictive modeling. While both methods improved the accuracy of the K-NN algorithm, Backward Elimination consistently resulted in a lower RMSE, indicating a more precise prediction. This could be attributed to the method's ability to consider the full scope of features before making reductions, potentially capturing more relevant information for the model. The findings of this study are particularly relevant for the Gorontalo City Food Service. By employing the optimized K-NN model with Backward Elimination, the Food Service can achieve more accurate price predictions, aiding in better decision-making processes. This could lead to more effective strategies for price regulation, supply chain management, and overall market analysis. One of the challenges encountered in this research was the inherent volatility of the chili price data, which can be influenced by a variety of unpredictable factors like weather conditions, market demand, and supply chain disruptions. Future research could explore the integration of additional variables that capture these external factors, potentially further enhancing the predictive accuracy of the model.

Another area for future exploration could be the application of other machine learning algorithms and comparing their performance with the K-NN algorithm in predicting chili prices. This could provide a broader understanding of the most effective methods in this specific.

4. Related Work

The study of chili price prediction in Gorontalo City using the K-Nearest Neighbor (K-NN) algorithm with feature selection techniques aligns with and builds upon a range of existing research in the fields of agricultural economics, machine learning, and predictive analytics. This section discusses relevant studies and how they relate to the current research. Previous research has extensively explored the use of various machine learning and statistical methods for agricultural price prediction. Studies such as those by Hardiansyah (2017) and Fatkhuroji *et al*. (2019) have demonstrated the application of ARIMA and Support Vector Machine (SVM) methods in forecasting commodity prices, including chilies, soybeans, and other agricultural products. These studies lay the groundwork for the current research, highlighting the feasibility and effectiveness of machine learning techniques in predicting agricultural prices. The K-NN algorithm has been a subject of interest in many studies due to its simplicity and effectiveness in classification and regression tasks. Its application in various domains, including retail, finance, and now agriculture, underscores its versatility. Research by Prabowo B.U. *et al*. (2019) and others has illustrated the algorithm's adaptability and efficacy when combined with other methods, such as linear regression, to enhance predictive performance. The role of feature selection in improving the accuracy of predictive models has been a critical area of research. The comparative efficacy of Backward Elimination and Forward Selection, as explored in this study, is supported by previous work in the field. Studies by M. Efendi L and Andi Bode (2021) and similar research have shown that the appropriate selection of features can significantly reduce error rates in predictive models, a finding that aligns with the outcomes of the current research. Recent research has increasingly focused on the integration of machine learning algorithms into agricultural decision-making processes. The ability of these algorithms to handle large datasets and produce accurate predictions makes them invaluable tools for agricultural planning and management. This research contributes to this growing body of work by providing insights specific to chili price prediction in Gorontalo City. The current research on chili price prediction using the K-NN algorithm with Backward Elimination and Forward Selection contributes to the evolving landscape of predictive analytics in agriculture. It corroborates the findings of previous studies regarding the effectiveness of machine learning techniques in price prediction while also offering new insights into the comparative performance of different feature selection methods. This work not only adds to the academic discourse but also has practical implications for agricultural stakeholders, particularly in regions like Gorontalo City.

5. Conclusion

The research conducted on chili price forecasting in Gorontalo City using the K-Nearest Neighbor (K-NN) algorithm has yielded significant insights. It has been demonstrated that the inclusion of feature selection techniques, specifically Backward Elimination, considerably enhances the accuracy of the K-NN model. This is evident from the lowest Root Mean Square Error (RMSE) value of 0.202 achieved with the application of Backward Elimination, surpassing the performance of Forward Selection. The practical implications of these findings are substantial for policymakers at the Gorontalo City Food Service. By integrating the optimized K-NN model into their decision-making processes, they can achieve more accurate and reliable chili price forecasts, which are essential for effective market regulation and strategic planning. It is noteworthy that the addition of feature selection methods to the K-NN algorithm not only reduces the error rate but also refines the overall predictive performance of the model. The superiority of Backward Elimination in this context suggests its suitability for modeling chili price data in Gorontalo City, offering a robust tool for addressing the complexities of agricultural market predictions. As a recommendation for future research, exploring other predictive models and feature selection techniques could be beneficial. This exploration might uncover methods that yield even lower error values and higher data accuracy percentages, further advancing the field of predictive analytics in agriculture. Such research would continue to build upon the foundation laid by this study, contributing to more sophisticated and precise agricultural forecasting tools.

References

[1] Soewignyo, F. and Simatupang, N., 2020. PENGARUH PERUBAHAN HARGA KOMODITAS PERTANIAN TERHADAP KESEJAHTERAAN PETANI DI PROPINSI SULAWESI UTARA. *Klabat Accounting Review*, *1*(1), pp.14-26. DOI: https://doi.org/10.60090/kar.v1i1.454.14-26.

[2] Suma, D.V., 2020. Data mining based prediction of demand in Indian market for refurbished electronics. *Journal of Soft Computing Paradigm*, *2*(2), pp.101-110.

[3] Bode, A., 2017. K-nearest neighbor dengan feature selection menggunakan backward elimination untuk prediksi harga komoditi kopi arabika. *ILKOM Jurnal Ilmiah*, *9*(2), pp.188-195. DOI: https://doi.org/10.33096/ilkom.v9i2.139.188-195.

[4] Wanto, A. and Windarto, A.P., 2017. Analisis prediksi indeks harga konsumen berdasarkan kelompok kesehatan dengan menggunakan metode backpropagation. *Sinkron: jurnal dan penelitian teknik informatika*, *2*(2), pp.37-43.

[5] Hadiansyah, F.N., 2017. Prediksi Harga Cabai dengan Menggunakan pemodelan Time Series ARIMA. *Indonesia Journal on Computing (Indo-JC)*, *2*(1), pp.71-78. DOI: https://doi.org/10.21108/INDOJC.2017.2.1.144.

[6] Fatkhuroji, F., Santosa, S. and Pramunendar, R.A., 2019. Prediksi Harga Kedelai Lokal Dan Kedelai Impor Dengan Metode Support Vector Machine Berbasis Forward Selection. *Jurnal Cyberku*, *15*(1), pp.61-76.

[7] Utomo, P.B., Utami, E. and Raharjo, S., 2019. Implementasi Metode K-Nearest Neighbor Dan Regresi Linear Dalam Prediksi Harga Emas. *Informasi Interaktif*, *4*(3), pp.155-159.

[8] Lasulika, M.E. and Bode, A., 2021. Komparasi Algoritma Data Mining Menggunakan Forward Selection pada Prediksi Harga Jagung. *JURNAL TECNOSCIENZA*, *5*(2), pp.157-172. DOI: https://doi.org/10.51158/tecnoscienza.v5i2.392.

[9] Karo, I.M.K., Khosuri, A., Septory, J.S.I. and Supandi, D.P., 2022. Pengaruh Metode Pengukuran Jarak pada Algoritma k-NN untuk Klasifikasi Kebakaran Hutan dan Lahan. *Jurnal Media Informatika Budidarma*, *6*(2), pp.1174-1182. DOI: http://dx.doi.org/10.30865/mib.v6i2.3967.

[10] Bode, A., 2019. Perbandingan metode prediksi support vector machine dan linear regression menggunakan backward elimination pada produksi minyak kelapa. *Simtek: jurnal sistem informasi dan teknik komputer*, *4*(2), pp.104-107. DOI: https://doi.org/10.51876/simtek.v4i2.57.

[11] Harafani, H. and Al-Kautsar, H.A., 2021. Meningkatkan Kinerja K-Nn Untuk Klasifikasi Kanker Payudara Dengan Forward Selection. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *18*(1), pp.99-110. DOI: https://doi.org/10.23887/jptk-undiksha.v18i1.29905.