Journal of Modern Applied Statistical Methods

Volume 21 | Issue 2 Article 4

Implementation of K-Nearest Neighbors Algorithm on Regional Food Security Classification in Indonesia

Wardhani Utami Dewi
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Khoirin Nisa*
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia, Khoirin.nisa@fmipa.unila.ac.id

Mustofa Usman
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Netti Herawati
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Recommended Citation

Wardhani Utami Dewi, Khoirin Nisa, Mustofa Usman, Netti Herawati (2022). Implementation of K-Nearest Neighbors Algorithm on Regional Food Security Classification in Indonesia. Journal of Modern Applied Statistical Methods, 21(2), https://doi.org/10.56801/Jmasm.V21.i2.4

Journal of Modern Applied Statistical Methods
2022, Vol. 21, No. 2,
Doi: 10.56801/Jmasm.V21.i2.4

Copyright © 2022 JMASM

ISSN 1538 – 9472

Implementation of K-Nearest Neighbors Algorithm on Regional Food Security Classification in Indonesia

Wardhani Utami Dewi
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Khoirin Nisa
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Mustofa Usman
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Netti Herawati
Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Classifying regions that are vulnerable to food security in the face of a global recession is critical. Thus, this research aims to classify regions in Indonesia based on food security characteristics. The sample used is 34 provinces grouped by geographical area in Indonesia. The food security variables used include food availability, access, and absorption. The K-nearest neighbors (KNN) classification was applied to the data using two nearest neighbors. All provinces in the region of Java and Sulawesi are predicted to have high food security.
In contrast, several provinces in Sumatera, Kalimantan, Nusa Tenggara, Maluku, and Papua still have relatively low food security. Papua has the lowest food security rate for availability, access, and utilization. In conclusion, the KNN model in this study has a fairly good performance with an accuracy rate of 76.47%. This is obtained by considering the amount of data, the size of the training data, and the model complexity of the many features.

Keywords: Machine Learning, KNN, Food Security.

1. Introduction

The k-nearest neighbor (KNN) algorithm is a simple yet effective machine learning algorithm for classification and regression tasks. This paper applies KNN to the regional classification of food security in Indonesia. Food security is critical for countries worldwide, including Indonesia (Widiana et al., 2022). However, food security in Indonesia is vulnerable to various factors, including significant economic changes (Mustafa et al., 2019; Setiadi et al., 2022). 2023 is predicted to be a year of economic recession, impacting people's economic well-being, including access to adequate and nutritious food (Montolalu et al., 2022; Organization, 2022; Woodhill et al., 2022). It is important to classify Indonesian regions based on similar characteristics of food security to provide insight for policymakers in identifying vulnerable areas and designing appropriate strategies to increase food security.

Some studies related to food security have been shown in literature, e.g., the classification of Ghanaian, Liberian, and Senegalese households for comparison and interpretation of food security using decision tree algorithm (Wardle et al., 2022), classification using logistic regression to assess factors related to food security (Wardle et al., 2022), and developing cybersecurity systems based on machine learning and deep learning algorithms to protect food safety systems: industrial control systems (Alkahtani & Aldhyani, 2022). However, until recently, studies using the KNN algorithm to classify Indonesian regions based on similar food security characteristics are none in the literature.

Therefore, this article aims to provide another approach to the classification of regional food security in Indonesia by using the KNN algorithm based on similar characteristics of regional food security. In addition, in this article, we use appropriate cross-validation techniques and evaluation metrics to measure the accuracy and reliability of the resulting classification model. The results of this research are expected to give useful information to policymakers in assessing and improving food security in Indonesia.

2. Research Methods

To support this research, the R Studio 4.3.1 software was employed. The data consists of food security characteristics from 34 provinces grouped into 7 regions based on Indonesia's geography: Sumatera, Java, Kalimantan, Sulawesi, Nusa Tenggara, Maluku, and Papua. Data were obtained from the Badan Pusat Statistik (BPS) and the Center for Agricultural Data and Information Systems, Ministry of Agriculture. The dimensions of food security include food availability, accessibility, and food absorption. Then, more specifics of each dimension are presented in Table 1.

According to mohanraj et al. (2021), the knn method is included in the supervised learning algorithm family. this algorithm is used to classify objects based on their attributes (bhatti et al., 2021; borman et al., 2021; wang et al., 2021). a useful approach for identifying patterns and predicting classification based on similarity to nearest neighbors (henderi et al., 2021; himeur et al., 2021).

The procedures applied for data analysis in research are given in stages as follows.

a. Before implementing the KNN method, it is necessary to preprocess data to ensure consistency and quality (Erol et al., 2022; Tao et al., 2021). Preprocessing steps may include cleaning the data from missing or invalid values and data normalization to adjust the variable scale (Joshi & Patel, 2021).

b. Dividing the observational data into two parts for the training process of the testing process by looking at the highest level of accuracy

c. The size of the dataset is the main consideration in choosing a k value. This study's dataset was relatively small, so k=1, k=2, and k=3 were chosen. This is adjusted so that the nearest significant neighbor can be identified. The shortest distance between the data and the centroid generally uses the Euclidean distance as in equation (1) (Gupta & Chandra, 2022).

$$d_{m,n} = \sqrt{\sum_{i=1}^{n} (X_{mi} - Y_{ni})^2} \quad (1)$$

d. Evaluate the model using the confusion matrix. A confusion matrix is used to describe the performance of the classification model by comparing the predictions of the model with the actual value of the observed data (Chicco & Jurman, 2020; Wexler et al., 2019). According to (Hachmi et al., 2019), there are four main elements in the confusion matrix, namely true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In the context of KNN, these elements are used to measure how well the KNN model performs classification (Abu Alfeilat et al., 2019; Shah et al., 2020). The following is the confusion matrix, which can be seen in Table 2.

Table 1. Food Security Variable Dataset

Variable	Description
Availability	Agricultural commodity production
Poor	Percentage of poor population by province
Electricity	Percentage of households with a PLN electricity source by province
Distribution	The Main Chain of Rice Distribution
Consumption	The average percentage of household expenditure used for non-food consumption by province
Stunt	Percentage of children under five who are stunted by province
Underweight	Percentage of children under five who are underweight by province

Source: (badan pusat statistika, 2022; kementerian pertanian, 2022)

Table 2. Cross tables

Actual Value	Value	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Based on Table 2., the level of accuracy in the classification process by measuring accuracy, sensitivity, and specificity.

$$\text{Acuracy} = \frac{(\text{TP+TN})}{(\text{TP+TN+FP+FN})} \text{ x } 100\% \quad (2)$$

$$\text{Sensitivity} = \frac{(\text{TP})}{(\text{TP+FN})} \text{ x } 100\% \quad (3)$$

$$\text{Specificity} = \frac{(\text{TN})}{(\text{TN+FP})} \text{ x } 100\% \quad (4)$$

3. Results and discussion

3.1 Preprocessing Data

Before implementing the KNN method, data preprocessing steps need to be carried out to ensure data consistency and quality. The results of this data preprocessing include data cleaning from missing or invalid values and data normalization to adjust the variable scale presented in Table 3.

Table 3. Data preprocessing results

Availability	Poor	Electricity	Distribution	Consumption	Stunt	Underweight
-0.0446	0.4862	-0.6991	-0.3594	-1.1188	-1.2205	0.0499
0.5888	0.3890	-0.1978	0.1332	-0.9674	0.3416	0.4212
-0.3959	0.4008	-0.0412	-1.1985	0.2593	-0.2925	0.2217
-0.3140	0.0453	0.2617	0.3130	-0.6014	0.9758	0.3879
-0.1301	0.3586	0.0528	-0.1489	-0.4251	0.8211	0.5320
⋮	⋮	⋮	⋮	⋮	⋮	⋮
-0.0027	-0.1854	-0.2605	-0.2164	-0.1565	-1.8083	0.0000
-1.5480	-0.1560	1.9222	-0.4627	0.3890	-0.4317	0.8541
-1.2112	-0.3690	1.0136	-0.1642	0.2821	-0.4317	0.5983
-1.1093	-1.1045	2.2982	-0.7641	-0.2903	-1.0349	0.1643
-2.9749	-5.0766	1.6925	-4.1551	-3.9321	-1.7464	0.9982

In the context of data preprocessing on a dataset that contains the percentage of underweight variables in all provinces in Indonesia, the results of the data preprocessing are “positive”. Referring to the fact that the percentage values of underweight for each province are not negative. This means that each value of the percentage of underweight in the dataset has a value greater than or equal to zero. This is to be expected because the underweight percentage measures the percentage of children who weigh less than the standard expected for their age and height. Therefore, the percentage underweight values are always positive (or zero if no child is included in the underweight category).

3.2 Determination of k

Different training data sizes can influence the best k in the KNN algorithm because it can affect the quality of predictions and overall model performance. In determining the best k in the KNN algorithm, the level of accuracy can be an important factor in choosing the optimal k value. Table 4 presents several alternative training data sizes for 10% to 90% intervals, with the proposed k values being 1, 2, and 3.

Table 4. The KNN accuracy for different training data sizes using k = 1,2,3.

Training data size	Accuracy value of KNN		
	k=1	k=2	k=3
10%	0.6129	0.5161	0.4839
20%	0.5000	0.6071	0.6071
30%	0.5833	0.4167	0.5000
40%	0.5238	0.4762	0.5714
50%	0.6235	0.7056	0.6523
60%	0.5714	0.4286	0.5000
70%	0.4545	0.5455	0.5455
80%	0.5714	0.5714	0.2857
90%	0.2500	0.2500	0.5000

Based on Table 5, the training data of 50% indicates that half of the dataset is used to train the model, while the remaining half is used for testing. Choosing the value of k = 2 has the highest level of accuracy, which is equal to 0.7056, indicating that considering only the two nearest neighbors gives the most accurate results in that context. Another reason is the relatively minimum number of observations used. However, choosing the best k depends on the accuracy level and involves considering other factors, such as model complexity. The number of features in the dataset affects the complexity of the KNN model. This study uses 7 features. Many features exist, causing increased dimensions and increased data processing complexity. In addition, data normalization or scaling has been carried out at the preprocessing stage to make it easier to find relevant nearest neighbors.

3.3 Model Evaluation with Confusion Matrix

The confusion matrix is arranged as a two-dimensional table, with the horizontal axis representing the predicted labels generated by the model and the vertical axis representing the actual labels of the test data. In each cell in the table, the number of samples that fall into a certain combination of classifications, such as true positive, true negative, false positive, and false negative, is recorded. Following are the results of the matrix confusion shown in Table 5.

Table 5. Results of the KNN classification matrix Confusion

Test Labels	Label predictions		Row Totals
	Low	High	
Low	6 0.667 0.857 0.353	3 0.333 0.500 0.176	9 0.529
High	1 0.125 0.143 0.058	7 0.875 0.70 0.412	8 0.471
Columns Total	7 0.412	10 0.588	17

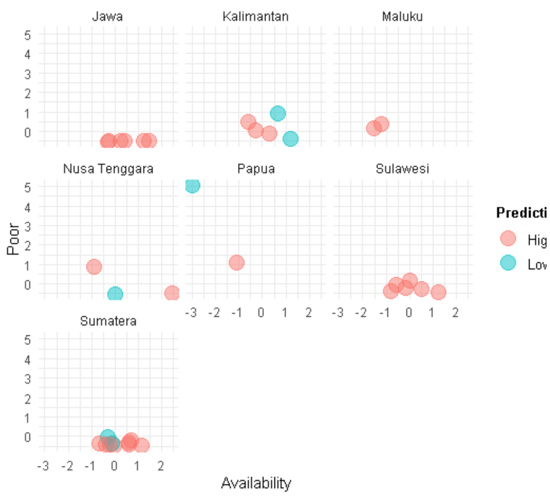
Based on Table 5, the KNN model has been classified into two classes: high = 1 and low = 2. This evaluation provides more detailed information about the model's performance in classifying low and high classes. In the classification for low class, out of 9 low samples, the model correctly classified 6 samples (True Positive/TP). However, 3 samples were incorrectly classified as high (False Positive/FP). In the classification for high class, out of 8 high samples, the model correctly classified 7 samples (True Positive/TP). However, there was 1 sample was incorrectly classified as low (False Negative/FN).

From these values, we can calculate some commonly used evaluation metrics. The accuracy of the KNN model in this classification is 76.47%, which means that the model correctly classifies 76.47% of the total sample. The precision for the low class is 85.70%, which means the percentage of samples that are correctly classified as low from the total low prediction given by the model. The recall or sensitivity for the high class is 40%, which means the percentage of high samples correctly classified by the model out of the total number of true high samples. The KNN model's F1 value (F1-Score) in this classification is 54.54%, which combines precision and recall into one score that describes the balance between the two.

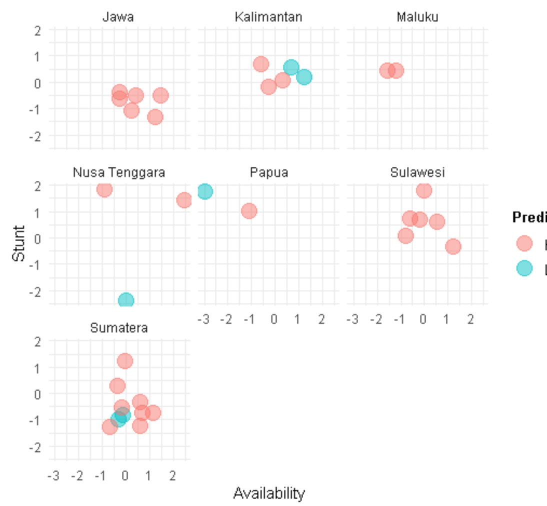
Based on this evaluation, we can conclude that the KNN model performs fairly well classifying low and high classes, with an accuracy of 76.47%. However, some misclassifications, such as false positives in the low and false negatives in the high classifications, need attention. This evaluation provides a more detailed understanding of model performance and can be used to make necessary improvements or adjustments to improve model performance in a specific classification context.

3.4 Region classification based on food security variables

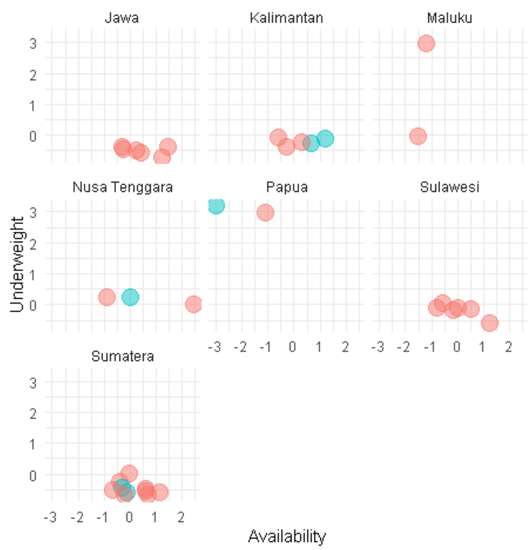
This study used 34 provinces in Indonesia, which were grouped into their respective region. So, the 7 regions in Indonesia are used as levels and the status of food security. Through plots using the KNN algorithm, regions have been grouped based on food security variables, as shown in Figure 1.



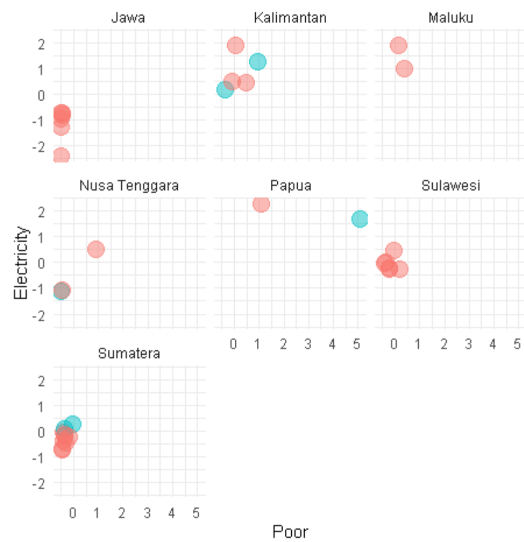
(a) Region and status grouping based on availability variables with poor



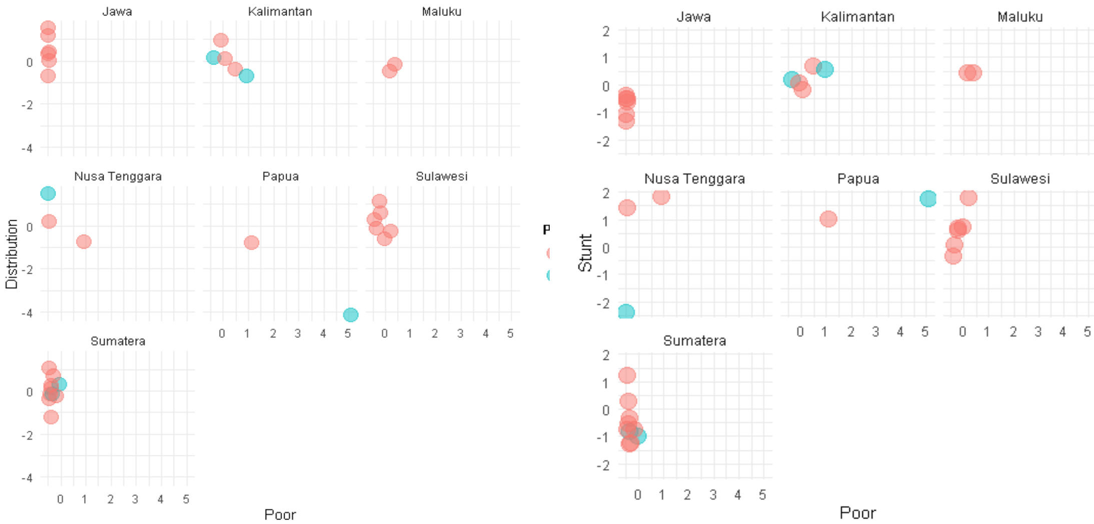
(b) Region and status grouping based on availability and stunt variables



(c) Region and status grouping based on availability variable with underweight



(d) Region and status grouping based on poor variables with electricity

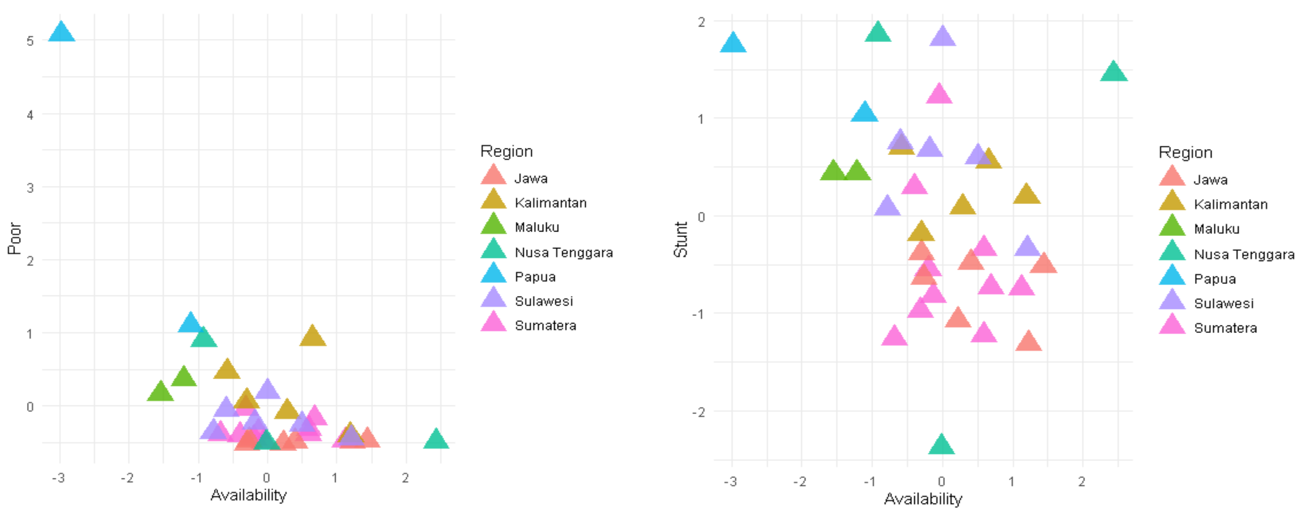


(e) Region and status grouping based on poor variables with distribution

(f) Region and status grouping based on poor and stunt variables

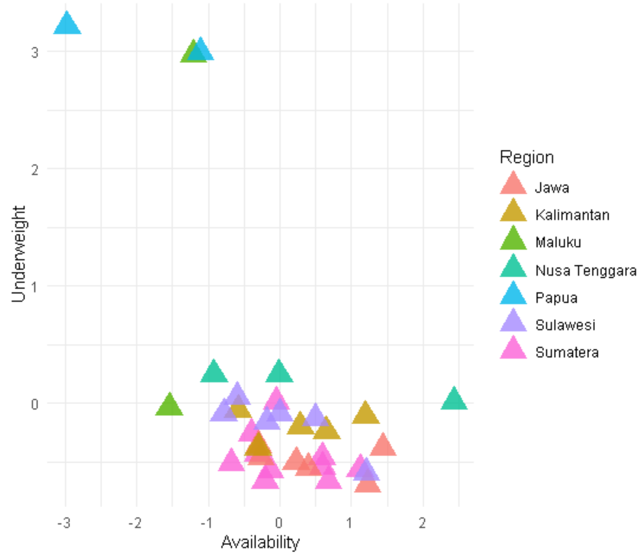
Figure 1. Grouping of regions and status based on all variables.

Based on Figure 1. for all provinces in the region of Java and Sulawesi shows a high prediction of food security based on all variables. Meanwhile, several provinces in the region of Kalimantan, Sumatera, Papua, and Nusa Tenggara have low predictions of food security. Then, it also examines the relationship between food security variables, which represent all regions in Indonesia, which are presented in Figure 2 and Figure 3.



a) The relationship between Availability and Poor variables

b) The relationship between Availability and Stunt variables



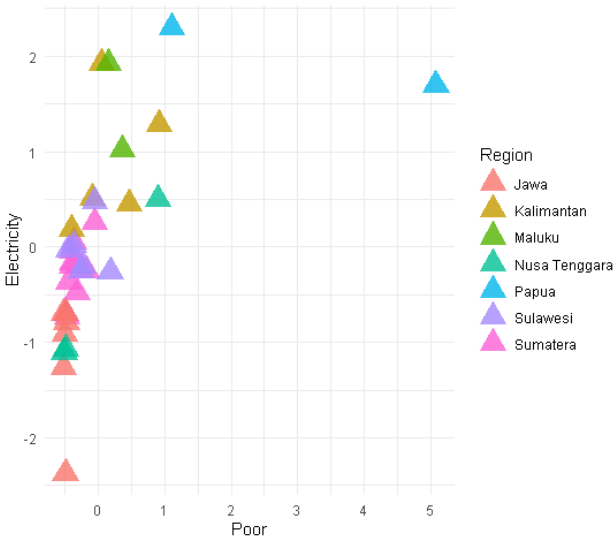
c) Relationship between Availability and Underweight variables

Figure 2. The plot of the relationship between Availability variables and each Poor, Stunt, and Underweight variables

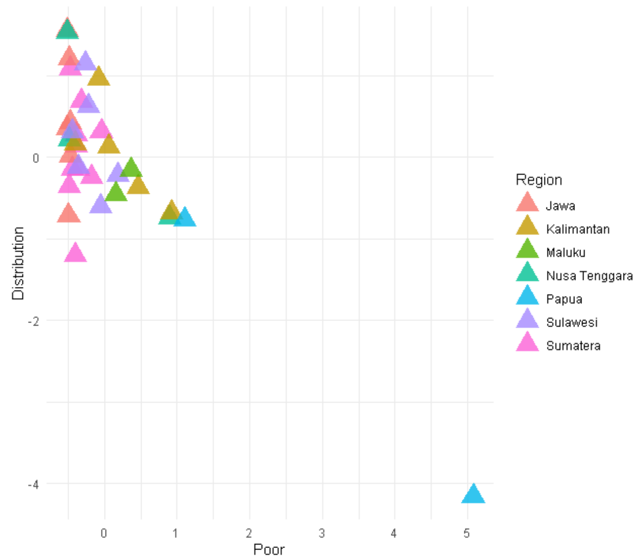
---

Figure 2. Visualizing the relationship between the Availability variable and several other variables. As shown in Figure 3, point a), the greater the agricultural commodity production, the smaller the percentage of the poor population in all regions in Indonesia. The plot shows that one of the provinces in Papua has the highest percentage of the poor population, so agricultural commodity production is lower. At the same time, the other 6 regions tend to be grouped but still have differences in the levels of each variable. Then, for Figure 2. point b), all regions tend to spread out and are difficult to analyze based on the relationship between agricultural commodity production and the high stunting rate. Continue in Figure 2. Point c) shows something similar to point a) in which there is a tendency between equal regions or groups in the relationship between agricultural commodity production and the number of underweight children. The greater the agricultural commodity production, the lower the percentage value is relatively small. One of Maluku's provinces, Papua, has the highest percentage of underweight children.

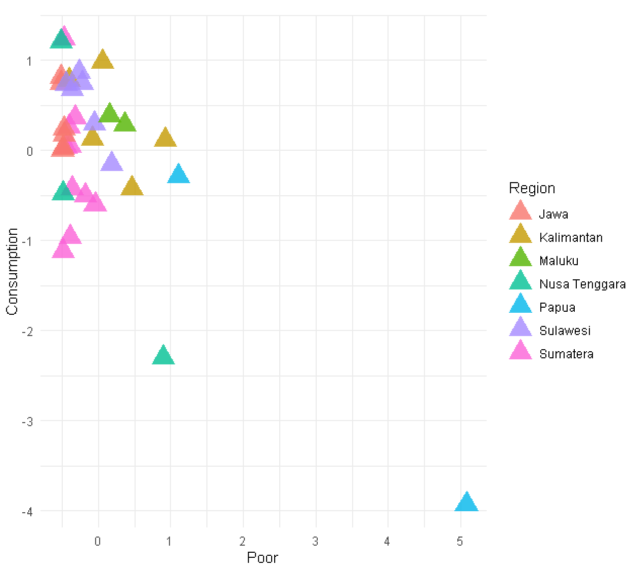
Based on Figure 3. point a), the lower percentage of the poor population indicates that many residents already use PLN electricity. Papua Island has the highest poverty rate. Figure 3. points b) and c) shows that one of the provinces in Papua has the highest percentage of the poor population, so the percentage of the main chain of rice distribution and the average percentage of household expenditure used for non-food consumption and consumption is very low. Meanwhile, other regions have low poverty rates because the main chain of rice distribution is quite high. Figure 3. Point d) The average area in Indonesia with a low poverty rate still has a fairly high percentage of children experiencing stunting. Meanwhile, Figure 5. Point e) The low population of poor people also shows that very few children are underweight.



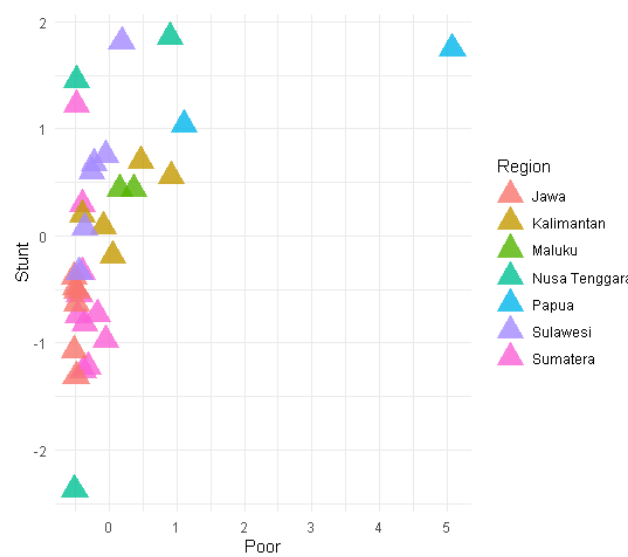
a) Relationship between Poor and Electricity



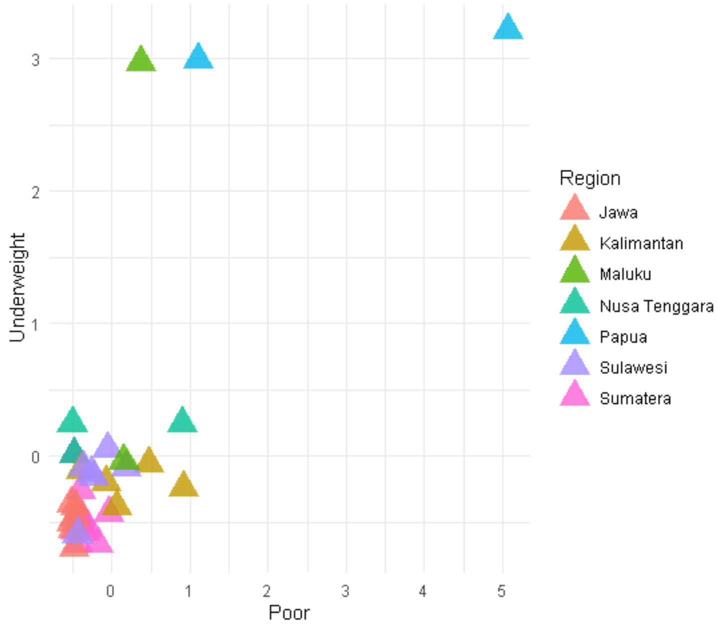
b) The relationship between the Poor and Distribution



c) Relationship between Poor and Consumption



d) Relationship between Poor and Stun



e) Relationship between Poor and Underweight

Figure 3. The plot of the relationship between the Poor and each variable variable Electricity, Distribution, Consumption, and Underweight

---

The KNN model performs fairly well in classifying low and high classes with an accuracy of 76.47%. However, some misclassifications, such as false positives in the low and false negatives in the high classifications, need attention. The results of regional grouping based on food security variables show that all provinces in Java and Sulawesi have high predictions of food security. Then, when viewed based on the relationship between food security variables, Papua is the island with the lowest food security rating, both in terms of the poor population, the ratio of normative consumption per capita, the large number of stunted children, the low use of PLN electricity, as well as of distribution and consumption. Used is still not feasible.

The advantage of this research is using the KNN algorithm, which can provide detailed information about the model's performance in classifying low and high food security classes in 7 regions in Indonesia. In addition, this study also visually illustrates the relationship between various food security variables through plots that provide a more detailed understanding of the relationship between variables. However, a drawback of this study is the limited sample size, i.e., 34 provinces in Indonesia, which may not cover all the country's food security variations. In addition, this study also did not consider other factors that might affect food security, such as climate, culture, and accessibility.

This research aligns with previous research that has applied the KNN algorithm in classification and prediction. Rahim et al. (2022), The k value that works best will depend on the specific problem being solved and the number of characteristics of the analyzed data. So, Chen et al. (2022) said that KNN is more suitable for medium-sized data sets; in their research, the dataset used was small, but the performance of KNN was better. Likewise, in their research, Hu et al. (2020) said that very large k causes a global flattening of class boundaries. This can lead to underfitting, in which the KNN model becomes too general and performs poorly on training and test data. Considering this, this study chose a small k because it was based on high accuracy and because the sample used was a minimum. Ramezan et al. (2021) The efficiency of the KNN algorithm is also affected by the size of the training set. In this research, the size of the training data and the number of k values affect the accuracy of the KNN model so that the size of the training data ranges from 10% to 90% with the number of k = 1, k = 2, k = 3 and k is selected with the highest degree of accuracy.

The results of this research imply that it can provide useful information for the government and stakeholders in making decisions related to food security policies in various regions in Indonesia. With a more detailed understanding of the factors affecting food security, steps can be taken to increase food security in areas where this is still low. In addition, the results of this study can also be the basis for further research and development in the field of food security and regional classification based on related variables.

4. Conclusions

The value of k = 2 was chosen because it has the highest level of accuracy with a training data size of 50%. So, this study only uses the two nearest neighbors for classification because considering the minimum amount of data, the number of variables used is the higher the complexity of the model. Based on the results of the regional grouping, it was found that all provinces in Java and Sulawesi had high food security predictions. In contrast, several provinces in Kalimantan, Sumatera, Papua, and Nusa Tenggara had low food security predictions. In addition, it was found that Papua has a high poverty level and various aspects of food security that still need improvement. The KNN model in this study has a fairly good performance with an accuracy rate of 76.47%.

Suggestions for further research, more samples, and representative data can increase the validity and generalizability of research results. This study only uses data from 34 provinces in Indonesia, so it does not cover all variations in food security in this country. Further research can involve other methods, such as logistic regression and cluster analysis, to compare results with the KNN method and gain a more comprehensive understanding of food security in Indonesia.

Acknowledgement

Thank you to the Directorate of Research, Technology, and Community Service for the research funds provided in the context of the Operational Assistance Program for

State Universities Research and Community Service Program for Fiscal Year 2023 (Number 0536/E5/PG.02.00/2023 dated 30 May 2023). The financial support provided by this Directorate has become a strong foundation for the success of this research.

References

Abu Alfeilat, H. A., Hassanat, A. B. A., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., Eyal Salman, H. S., & Prasath, V. B. S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big Data, 7(4), 221–248.

Alkahtani, H., & Aldhyani, T. H. H. (2022). Developing cybersecurity systems based on machine learning and deep learning algorithms for protecting food security systems: industrial control systems. Electronics, 11(11), 1717.

Badan Pusat Statistika. (2022). Konsumsi Kalori dan Protein Penduduk Indonesia dan Provinsi. Badan Pusat Statistika Indonesia.

Bhatti, U. A., Yu, Z., Chanussot, J., Zeeshan, Z., Yuan, L., Luo, W., Nawaz, S. A., Bhatti, M. A., Ain, Q. U., & Mehmood, A. (2021). Local similarity-based spatial–spectral fusion hyperspectral image classification with deep CNN and Gabor filtering. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–15.

Borman, R. I., Napianto, R., Nugroho, N., Pasha, D., Rahmanto, Y., & Yudoutomo, Y. E. P. (2021). Implementation of PCA and KNN Algorithms in the Classification of Indonesian Medicinal Plants. 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), 46–50.

Chen, B., Chen, C., Hu, J., Sayeed, Z., Qi, J., Darwiche, H. F., Little, B. E., Lou, S., Darwish, M., & Foote, C. (2022). Computer Vision and Machine Learning-Based Gait Pattern Recognition for Flat Fall Prediction. Sensors, 22(20), 7960.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics, 21(1), 1–13.

Erol, G., Uzbaş, B., Yücelbaş, C., & Yücelbaş, Ş. (2022). Analyzing the effect of data preprocessing techniques using machine learning algorithms on the diagnosis of COVID-19. Concurrency and Computation: Practice and Experience, 34(28), e7393.

Gupta, M. K., & Chandra, P. (2022). Effects of similarity/distance metrics on k-means algorithm with respect to its applications in IoT and multimedia: A review. Multimedia Tools and Applications, 81(26), 37007–37032.

Hachmi, F., Boujenfa, K., & Limam, M. (2019). Enhancing the accuracy of intrusion detection systems by reducing the rates of false positives and false negatives through multi-objective optimization. Journal of Network and Systems Management, 27, 93–120.

Henderi, H., Wahyuningsih, T., & Rahwanto, E. (2021). Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. International Journal of Informatics and Information Systems, 4(1), 13–20.

Himeur, Y., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors. International Journal of Intelligent Systems, 36(6), 2865–2894.

Hu, T., Tang, T., Lin, R., Chen, M., Han, S., & Wu, J. (2020). A simple data augmentation algorithm and a self-adaptive convolutional architecture for few-shot fault diagnosis under different working conditions. Measurement, 156, 107539.

Joshi, A. P., & Patel, B. V. (2021). Data preprocessing: the techniques for preparing clean and quality data for data analytics process. Oriental Journal of Computer Science and Technology, 13(0203), 78–81.

Kementerian Pertanian. (2022). Statistik Ketahanan PanganTahun 2022.

Mohanraj, T., Yerchuru, J., Krishnan, H., Aravind, R. S. N., & Yameni, R. (2021). Development of tool condition monitoring system in end milling process using wavelet features and Hoelder's exponent with machine learning algorithms. Measurement, 173, 108671.

Montolalu, M. H., Ekananda, M., Dartanto, T., Widyawati, D., & Panennungi, M. (2022). The Analysis of Trade Liberalization and Nutrition Intake for Improving Food Security across Districts in Indonesia. Sustainability, 14(6), 3291.

Mustafa, S. N., Kakakhel, S. J., & Shah, F. A. (2019). The moderating effect of entrepreneurial culture and government support on the relationship between entrepreneurial orientation and firm performance. In Abasyn Journal of Social Sciences (Vol. 12, Issue 2). Abasyn University. https://doi.org/10.34091/ajss.12.2.04

Organization, W. H. (2022). The state of food security and nutrition in the world 2022: Repurposing food and agricultural policies to make healthy diets more affordable (Vol. 2022). Food & Agriculture Org.

Rahim, R., Ahmar, A. S., & Hidayat, R. (2022). Cross-Validation and Validation Set Methods for Choosing K in KNN Algorithm for Healthcare Case Study. JINAV: Journal of Information and Visualization, 3(1), 57–61.

Ramezan, C. A., Warner, T. A., Maxwell, A. E., & Price, B. S. (2021). Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. Remote Sensing, 13(3), 368.

Setiadi, R., Artiningsih, A., Sophianingrum, M., & Satriani, T. (2022). The dimension of rural-urban linkage of food security assessment: An Indonesian case study. Asian Geographer, 39(2), 113–131.

Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A comparative analysis of logistic regression, random forest and KNN models for the text classification. Augmented Human Research, 5, 1–16.

Tao, X., Chi, O., Delaney, P. J., Li, L., & Huang, J. (2021). Detecting depression using an ensemble classifier based on Quality of Life scales. Brain Informatics, 8, 1–15.

Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognition Letters, 141, 61–67.

Wardle, J. A., Sagan, V., & Mohammed, F. (2022). Using open data cube on the cloud to investigate food security by means of cropland changes in djibouti. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 1039–1044.

Wexler, J., Pushkarna, M., Bolukbasi, T., Wattenberg, M., Viégas, F., & Wilson, J. (2019). The what-if tool: Interactive probing of machine learning models. IEEE Transactions on Visualization and Computer Graphics, 26(1), 56–65.

Widiana, A., Wijaya, C., & Atmoko, A. W. (2022). The Challenges of Food Security Policy in Indonesia: Lesson Learned from Vietnam, India, and Japan. Technium Soc. Sci. J., 33, 1.

Woodhill, J., Kishore, A., Njuki, J., Jones, K., & Hasnain, S. (2022). Food systems and rural wellbeing: challenges and opportunities. Food Security, 14(5), 1099–1121.