

## Improve of Multiobjective Model on the Classification Problem of Food Consumption Levels in Indonesia

Eka Susanti<sup>1\*</sup>, Novi Rustiana Dewi<sup>1</sup>, Arsi<sup>2</sup>, Anggelina Maria Putri Sihotang<sup>1</sup>,  
Dhea Adelina Lumban Raja<sup>1</sup>

<sup>1</sup>Mathematics Department, Universitas Sriwijaya, Indralaya, Indonesia

<sup>2</sup>Department of Plant Protection, Faculty of Agriculture, Universitas Sriwijaya, Indralaya, Indonesia

\*<sup>1</sup>email: [eka\\_susanti@mipa.unsri.ac.id](mailto:eka_susanti@mipa.unsri.ac.id)

### Abstract

Classification is the process of grouping objects based on similarities and differences. In this article, a multi-objective classification model is developed with three objective functions, namely the function that maximizes the values of accuracy, sensitivity and specificity. The developed model is applied to the problem of classifying meat, egg and fish consumption levels. The classification method used is K-Nearest Neighbor (KNN) with three objective functions and the addition of the GridSearchCV module to the KNN calculation. Completion of the multiobjective model using the weighting method and Particle Swam Optimization (PSO). Based on the data, with objective function weights of 1, 2 and 3 respectively being 0.7, 0.15 and 0.15, the results obtained for Rural Areas Meat, Fish and Egg Attributes of the model performance are in good criteria. for Urban Areas Attributes of Meat, Fish and Eggs the model's performance in the criteria is very good. Addition of the GridsearchCV module can facilitate the calculation of the KNN method classification because the model will provide the best k value without having to do repeated calculations.

**Keywords:** *Classification, Multiobjective, K-Nearest Neighbor, GridSearchCV*

### Abstrak

Klasifikasi adalah proses pengelompokan objek berdasarkan kesamaan dan perbedaan. Pada artikel ini, model klasifikasi multi-objektif dikembangkan dengan tiga fungsi objektif, yaitu fungsi yang memaksimalkan nilai akurasi, sensitivitas, dan spesifisitas. Model yang dikembangkan diaplikasikan pada permasalahan klasifikasi tingkat konsumsi daging, telur dan ikan. Metode klasifikasi yang digunakan adalah K-Nearest Neighbor (KNN) dengan tiga fungsi objektif dan penambahan modul GridSearchCV pada perhitungan KNN. Penyelesaian model multiobjektif menggunakan metode pembobotan dan Particle Swam Optimization (PSO). Berdasarkan data, dengan bobot fungsi objektif 1, 2 dan 3 berturut-turut adalah 0.7, 0.15 dan 0.15, diperoleh hasil kinerja model dengan kriteria baik untuk atribut daging, ikan dan telur wilayah pedesaan. Sedangkan atribut daging, ikan dan telur untuk wilayah perkotaan kinerja model dalam kriteria sangat baik. Penambahan modul GridsearchCV dapat mempermudah perhitungan klasifikasi metode KNN karena model akan memberikan nilai k yang terbaik tanpa harus melakukan perhitungan berulang-ulang.

**Kata kunci:** *Klasifikasi, Multiobjektif, K-Nearest Neighbor, GridSearchCV*

\*) Corresponding Author

Received: 17-11-2024, Accepted: 4-02-2025, Published: 17-02-2025

Citation: A. Author1, B. C. Author2, and D. E. Author3, "Title of the article", *KUBIK J. Publ. Ilm. Mat.*, Vol. 9, No. 1, pp. 1-9, 2024.

## Introduction

Food consumption classification will provide an overview of the level of consumption in an area. The level of consumption is one of the factors in determining the amount of product availability, so research related to classification is very necessary in an effort to ensure the availability of food products. Food is a national issue that is of concern to the government. Data classification is the categorization of data obtained based on the parts of the research that have been determined. Data classification aims to separate heterogeneous data into homogeneous data groups so that the dominant properties of the data can be easily seen. Classification can be done by applying the data mining method. The data mining method consists of Naïve Bayes (NB) [1], [2], [3], [4], *K-Nearest Neighbor* (KNN) [5], [6], [7], *Random Forest* (RF) [8], [9], *Decision Stump* (DS), *Decision Tree* (DT) [10], [11], *Rule Induction* (RI), *Linear Regression* (LR), *Linear Discriminant Analysis* (LDA), *Neural Network* (NN) [12], and *Support Vector Machine* (SVM).

The application of classification techniques in various fields has been carried out a lot. Application of PSO-based Naïve Bayes classification technique in the banking world [13]. Application to Covid 19 data analysis [14], [15]. Analysis of cancer [16], [17]. The classification method is also used in the problem of biomass estimation [18]. A study related to the classification of social media users [19]. The development of PSO-based classification methods has been widely developed with multiobjective problems and applied in various fields. Studies related to cancer with a PSO-based multiobjective concept are discussed by [20]. The combination of the KNN method with the MOPSO algorithm in classifying certain disease genes has been discussed by [21] and gives the result that there is an increase in accuracy value after the combination of the methods. The application of MOPSO to the classification of high-dimensional cancer disease data was discussed by [22], the result was obtained that the application of MOPSO provided the best model performance. In this study, a multiobjective classification model with a PSO-based one will be developed with the addition of a gridsearchCV module to the KNN classification process to obtain the best k value. to classify the average weekly per capita consumption according to food commodities and per capita expenditure groups. The novelty of the research is the addition of the GridsearchCV module to the KNN method. Addition of the GridsearchCV module can facilitate the calculation of the KNN method classification because the model will provide the best k value without having to do repeated calculations. The aim of this research is to obtain a classification of food product consumption levels in Indonesia with a good level of accuracy, sensitivity and specificity.

## Methods

The research conducted is a quantitative descriptive research. The development of a multiobjective classification model is applied to classify the level of consumption of the Indonesian population towards foodstuffs. The data used is quantitative data, namely data on the level of consumption of foodstuffs. The results of the classification will provide an overview of the consumption level of the Indonesian population from each province based on the criteria of high, medium and low consumption levels. The following are the classification stages.

### 1. Data

The data used is data on the level of per capita consumption per week based on the class of public expenditure in Indonesia for urban and rural areas. The data was obtained from the website of

the Central Statistics Agency of 34 Provinces in Indonesia. The classification attributes are Meat, Fish, Eggs.

## 2. Data Processing

The initial stage of the classification process is data labeling. The data is labeled with high, medium and low criteria. The determination of the label is based on Regulatory Badan Pengawas Obat dan Makanan (BPOM) Number 30 of 2018 concerning Food Consumption Figures.

## 3. Data Classification and Analysis

The classification technique used is the KNN method [5], [6], [7]. The solution of the multiobjective model uses the PSO algorithm. The following is given the function of the purpose of multiobjective problems

$$f_1 = \text{Maximize Accuracy} = \text{Maximize} \frac{TP+TN}{(TP+FP+FN+TN)} \quad (1)$$

$$f_2 = \text{Maximize sensitivity} = \text{Maximize} \frac{TP}{(TP+FN)} \quad (2)$$

$$f_3 = \text{Maximize specificity} = \text{Maximize} \frac{TN}{(TN+FP)} \quad (3)$$

Using the weighting method, the Multiobjective Models (1), (2) and (3) can be written as the following objective single models.

$$F = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 \quad (4)$$

Subject to

$$0 < \lambda_1 < 1 \quad (5)$$

$$0 < \lambda_2 < 1 \quad (6)$$

$$0 < \lambda_3 < 1 \quad (7)$$

where

$TP$  = True Positive

$TN$  = True Negative

$FP$  = False Positive

$FN$  = False Negative

Model performance measurements use accuracy, precision and recall values.

## Results and Discussion

### Data Labeling Results.

The following are the results of data labeling for rural and urban areas with meat, fish and egg attributes.

**Table 1. Weekly Per Capita Consumption in Rural Areas Attributes of Meat, Fish and Egg Attributes**

No	Province	Expense Group			
		IDR 500.000- IDR 749.999,-	IDR 750.000- IDR 999.999,-	IDR 1.000.000- IDR 1.499.999	$\geq$ IDR 1.500.000
1	Nanggroe Aceh Darussalam	High	High	High	High
2	West Sumatra	Medium	High	High	High
3	North Sumatra	High	High	High	High
4	South Sumatra	High	High	High	High
5	Lampung	High	High	High	High

6	Riau	High	High	High	High
7	Riau Islands	High	High	High	High
8	Jambi	Medium	High	High	High
9	Bangka Belitung Islands	Medium	High	High	High
10	Bengkulu	Medium	High	High	High
11	Banten	Medium	High	High	High
12	West Java	High	High	High	High
13	Central Java	High	High	High	High
14	East Java	High	High	High	High
15	DI Yogyakarta	High	High	High	High
16	Bali	High	High	High	High
17	West Nusa Tenggara	High	High	High	High
18	East Nusa Tenggara	Medium	Medium	Medium	Medium
19	West Kalimantan	High	High	High	High
20	South Kalimantan	High	High	High	High
21	Central Kalimantan	Medium	High	High	High
22	East Kalimantan	High	High	High	High
23	North Kalimantan	Medium	High	High	High
24	West Sulawesi	Medium	High	High	High
25	Southeast Sulawesi	Medium	High	High	High
26	South Sulawesi	High	High	High	High
27	Central Sulawesi	Medium	Medium	High	High
28	North Sulawesi	Medium	High	High	High
29	Gorontalo	Medium	Medium	High	High
30	Maluku	Medium	Medium	Medium	Medium
31	North Maluku	Medium	Medium	Medium	Medium
32	Papua	Medium	Medium	Medium	Medium
33	Papua Barat	Medium	Medium	Medium	Medium

**Table 2. Weekly Per Capita Consumption in Urban Areas Attributes of Meat, Fish and Egg Attributes**

No	Province	Expense Group			
		IDR 500.000- IDR 749.999,-	IDR 750.000- IDR 999.999,-	IDR 1.000.000- IDR 1.499.999	≥ IDR 1.500.000
1	Nanggroe Aceh Darussalam	High	Medium	Medium	Medium
2	North Sumatra	High	Medium	Medium	High
3	West Sumatra	High	Medium	Medium	Medium
4	Riau	High	Medium	Medium	Medium
5	Riau Islands	High	Medium	Medium	High
6	Jambi	High	Medium	Medium	Medium

7	South Sumatra	High	Medium	High	High
8	Bengkulu	High	Medium	Medium	High
9	Lampung	High	Medium	Medium	High
10	Bangka Belitung	High	Medium	Medium	High
11	DKI Jakarta (Special Capital Region of Jakarta)	High	High	High	High
12	West Java	High	Medium	High	High
13	Banten	High	Medium	Medium	High
14	Central Java	High	Medium	Medium	Medium
15	Special Region of Yogyakarta	High	Medium	Medium	Medium
16	East Java	High	Medium	Medium	High
17	Bali	High	High	High	High
18	West Nusa Tenggara	High	High	High	High
19	East Nusa Tenggara	Medium	Medium	High	High
20	West Kalimantan	High	Medium	Medium	High
21	Central Kalimantan	High	Medium	Medium	High
22	South Kalimantan	High	Medium	Medium	High
23	East Kalimantan	High	Medium	Medium	Medium
24	North Kalimantan	High	Medium	Medium	Medium
25	Southeast Sulawesi	Medium	High	High	High
26	North Sulawesi	Medium	Medium	High	High
27	Gorontalo	Medium	Medium	Medium	High
28	Central Sulawesi	Medium	High	High	High
29	West Sulawesi	Medium	High	High	High
30	South Sulawesi	High	High	High	High
31	Maluku	Medium	Medium	Medium	High
32	North Maluku	Medium	Medium	Medium	Medium
33	West Papua	Medium	Medium	Medium	Medium
34	Papua	Medium	Medium	Medium	High

Based on Tables 1 and 2, 80% of the total data and 20% of the test data were randomly taken. The process of random determination of training data and test data as well as the classification of test data uses python programming. By taking the weights of the accuracy values, sensitivity and specificity were 0.7, 0.15, and 0.15 respectively. The following are the results of the classification of test data.

**Table 3. The Classification Results of Weekly Per Capita Consumption Test Data in Rural Areas Attributes of Meat, Fish and Egg Attributes**

Expense Group	IDR 500.000 - 749.000	IDR 750.000 - 999.000	IDR 1.000.000 - 1.499.000	IDR > 1.500.000
BEST K	K=4	K=5	K=8	K=5

Test Data	Data	Prediction	Data	Prediction	Data	Prediction	Data	Prediction
Papua	Medium	High	Medium	High	Medium	High	High	High
Bali	High	High	High	High	High	High	High	High
Central Sulawesi	Medium	Medium	Medium	Medium	High	High	High	High
East Nusa Tenggara	Medium	Medium	Medium	High	Medium	High	High	High
Bangka Belitung	Medium	Medium	High	High	High	High	High	High
Bengkulu	Medium	Medium	High	High	High	High	High	High
South Kalimantan	High	High	High	High	High	High	High	High

**Table 4. Accuracy, Precision and *Recall* Values of Classification of Weekly Per Capita Consumption Test Data in Rural Areas Attributes of Meat, Fish and Egg Attributes**

Expense Group	BEST K	Weighted	Performance Vector Result	
IDR 500.000 - 749.000	K=4	0,7; 0,15; 015	Accuracy	86%
			Precision	90%
			Recall (Sensitivity)	86%
			Specificity	80%
IDR 750.000 - 999.000	K=5		Accuracy	71%
			Precision	80%
			Recall (Sensitivity)	71%
			Spesificity	33%
IDR 1.000.000 - 1.499.000	K=8		Accuracy	71%
			Precision	51%
			Recall (Sensitivity)	71%
			Spesificity	0%
IDR >1.500.000	K=5		Accuracy	100%
			Precision	100%
			Recall (Sensitivity)	100%
			Spesificity	-

Based on the results in Table 3 and Table 4 when viewed from the values of accuracy, precision and recall, the model performance is in good criteria.

**Table 5. Classification Results of Weekly Per Capita Consumption Test Data in Urban Areas Attributes of Meat, Fish and Eggs**

Expense Group	IDR 500.000 - 749.000		IDR 750.000 - 999.000		IDR 1.000.000 - 1.499.000		IDR >1.500.000	
BEST K	K=4		K=5		K=5		K=5	
Test Data	Data	Predictions	Data	Predictions	Data	Predictions	Data	Predictions
East Java	High	High	Medium	Medium	Medium	Medium	High	High

West Kalimantan	High	High	Medium	Medium	Medium	Medium	High	High
Central Sulawesi	Medium	Medium	High	Medium	High	High	High	High
Gorontalo	Medium	Medium	Medium	Medium	Medium	Medium	high	High
Lampung	High	High	Medium	Medium	Medium	Medium	High	High
South East Sulawesi	Medium	Medium	High	Medium	High	High	High	High
South Kalimantan	High	High	Medium	Medium	Medium	Medium	High	High

**Table 6 Accuracy, Precision and Recall Values Classification of Week Per Capita Consumption Test Data in Urban Areas Attributes of Meat, Fish and Eggs**

Expense Group	BEST K	Weighted	Performance Vector	
			Result	
IDR 500.000 - 749.000	K=4	0,7; 0,15; 015	Accuracy	100%
			Precision	100%
			Recall (Sensitivity)	100%
			Specificity	100%
IDR 750.000 - 999.000	K=5		Accuracy	71%
			Precision	51%
			Recall (Sensitivity)	71%
			Specificity	100%
IDR 1.000.000 - 1.499.000	K=5		Accuracy	100%
			Precision	100%
			Recall (Sensitivity)	100%
			Specificity	100%
IDR >1.500.000	K=5		Accuracy	100%
			Precision	100%
			Recall (Sensitivity)	100%
			Specificity	-

Judging from Table 5 and Table 6, in general, the model's performance in the criteria is very good. However, for the spending group of IDR. 750,000 - 999,000, the model performance is in good criteria.

## Conclusion

Based on the results of the calculation using population consumption level data for meat, fish and egg attributes in rural and urban areas in Indonesia, it can be concluded that the addition of the GridsearchCV module can facilitate the calculation of the KNN method classification because the model will provide the best k value without having to do repeated calculations. In general, the

performance of the multiobjective model is good in the criteria so that the multiobjective model developed can be applied to classification problems. For further research, the food consumption classification of districts in Indonesia can be considered so that the results better reflect the Indonesian region as a whole.

## Acknowledgments

"The Research/publication of this article was funded by DIPA of public Service Agency of Universitas Sriwijaya 2024. SP DIPA Number 023.17.2.677515/2024, on November 24, 2023. In accordance with the Rector's Decree Number : 0013/UN9/LP2M.PT/2024, On May 20, 2024"

## References

- [1] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, "Variable Selection for Naïve Bayes Classification," *Comput. Oper. Res.*, vol. 135, p. 105456, 2021.
- [2] S. Farhana, "Classification of Academic Performance for University Research Evaluation by Implementing Modified Naive Bayes Algorithm," *Procedia Comput. Sci.*, vol. 194, pp. 224–228, 2021.
- [3] M. Libnao, M. Misula, C. Andres, J. Mariñas, and A. Fabregas, "Traffic Incident Prediction and Classification System using Naïve Bayes Algorithm," *Procedia Comput. Sci.*, vol. 227, pp. 316–325, 2023.
- [4] M. Vishwakarma and N. Kesswani, "A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection," *Decis. Anal. J.*, vol. 7, no. January, p. 100233, 2023.
- [5] A. A. Hamidi, B. Robertson, and J. Ilow, "A new approach for ECG artifact detection using fine-KNN classification and wavelet scattering features in vital health applications," *Procedia Comput. Sci.*, vol. 224, pp. 60–67, 2023.
- [6] J. Zhang, Y. Li, F. Shen, Y. He, H. Tan, and Y. He, "Hierarchical text classification with multi-label contrastive learning and KNN," *Neurocomputing*, vol. 577, no. May 2023, p. 127323, 2024.
- [7] A. M. Vommi and T. K. Battula, "A hybrid filter-wrapper feature selection using Fuzzy KNN based on Bonferroni mean for medical datasets classification: A COVID-19 case study," *Expert Syst. Appl.*, vol. 218, no. May 2022, p. 119612, 2023.
- [8] S. Álvarez-Diez, J. S. Baixauli-Soler, G. Lozano-Reina, and D. Rodríguez-Linares Rey, "Subsidies for investing in energy efficiency measures: Applying a random forest model for unbalanced samples," *Appl. Energy*, vol. 359, no. January, pp. 1–15, 2024.
- [9] B. K. Meher, M. Singh, R. Birau, and A. Anand, "Forecasting stock prices of fintech companies of India using random forest with high-frequency data," *J. Open Innov. Technol. Mark. Complex.*, vol. 10, no. 1, p. 100180, 2024.
- [10] E. Carrillo, M. González, R. Parrilla, and A. Tarrega, "Classification trees as machine learning tool to explore consumers' purchasing decision pathway. A case-study on parent's perception of baby food jars," *Food Qual. Prefer.*, vol. 109, no. June, pp. 0–6, 2023.
- [11] A. Dhakar, B. Singh, and P. Gupta, "Fault diagnosis of air compressor set-up using decision tree based J48 classification algorithm," *J. Eng. Res.*, no. April, 2023.
- [12] B. Wang, Y. Sun, B. Xue, and M. Zhang, "Evolving deep neural networks by multi-objective particle swarm optimization for image classification," *GECCO 2019 - Proc. 2019 Genet. Evol. Comput. Conf.*, pp. 490–498, 2019.
- [13] Y. R. Religia, G. T. Pranoto, and I. M. Suwancita, "Analysis of the Use of Particle Swarm Optimization on Naïve Bayes for Classification of Credit Bank Applications," *JISA(Jurnal Inform. dan Sains)*, vol. 4, no. 2, pp. 133–137, 2021.
- [14] F. Manzella, G. Pagliarini, G. Sciavico, and I. E. Stan, "The voice of COVID-19: Breath and cough recording classification with temporal decision trees and random forests," *Artif. Intell. Med.*, vol. 137, no. February, p. 102486, 2023.



- [15] F. Özen, "Random forest regression for prediction of Covid-19 daily cases and deaths in Turkey," *Heliyon*, vol. 10, no. 4, pp. 1–19, 2024, doi: 10.1016/j.heliyon.2024.e25746.
- [16] K. Aguerchi, Y. Jabrane, M. Habba, and A. H. El Hassani, "A CNN Hyperparameters Optimization Based on Particle Swarm Optimization for Mammography Breast Cancer Classification," *J. Imaging*, vol. 10, no. 2, 2024.
- [17] T. M. T. A. Hamid, R. Sallehuddin, Z. M. Yunus, and A. Ali, "Ensemble Based Filter Feature Selection with Harmonize Particle Swarm Optimization and Support Vector Machine for Optimal Cancer Classification," *Mach. Learn. with Appl.*, vol. 5, no. December 2020, p. 100054, 2021.
- [18] X. Zhang *et al.*, "Improved random forest algorithms for increasing the accuracy of forest aboveground biomass estimation using Sentinel-2 imagery," *Ecol. Indic.*, vol. 159, no. February, p. 111752, 2024.
- [19] A. Huang, R. Xu, Y. Chen, and M. Guo, "Research on multi-label user classification of social media based on ML-KNN algorithm," *Technol. Forecast. Soc. Change*, vol. 188, no. August 2022, p. 122271, 2023.
- [20] M. Jafari, B. Ghavami, and V. S. Naeini, "A Decision Making Approach for Chemotherapy Planning based on Evolutionary Processing," pp. 1–11.
- [21] Y. Kowsari, S. Nakhodchi, and D. Gholamiangonabadi, "Gene selection from Microarray Expression Data: A Multi-objective PSO with Adaptive K-Nearest Neighborhood," *arXiv Prepr. arXiv2205.15020*, 2022.
- [22] V. Kumar, P. S. Singh Aydav, and S. Minz, "Multi-view ensemble learning using multi-objective particle swarm optimization for high dimensional data classification," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8523–8537, 2022.