INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN MULTIDISCIPLINARY EDUCATION

ISSN (print): 2833-4515, ISSN (online): 2833-4531

Volume 04 Issue 12 December 2025

DOI: 10.58806/ijirme.2025.v4i12n19 Impact factor- 6.748

Page No. 1354-1363

Comparative Analysis of Support Vector Machine, K-Nearest Neighbor, and Multilayer Perceptron for Nutritional Status Classification

Muhammad Satria Nugraha[1*],Erik Iman Heri Ujianto[2]

[1]Department of Informatics, University of Technology Yogyakarta, Yogyakarta, Indonesia

[2]Department of Master Information Technology, University of Technology Yogyakarta, Yogyakarta, Indonesia

ABSTRACT: Nutritional status plays an important role in describing the balance between nutritional intake and requirements, which can be measured through anthropometric indicators such as weight, height, and body mass index. Nutritional imbalances, whether deficiencies or excesses, can cause various serious health problems. This study aims to conduct a comparative analysis of three machine learning algorithms, namely Support Vector Machine, K-Nearest Neighbor, and Multilayer Perceptron, in classifying the nutritional status of adults based on anthropometric data from the RSUD Kota Banjar. The research stages included data preprocessing (data cleaning, normalization, standardization, and label encoding), dividing the data into train, validation, and test sets, and training the model with the best parameters for each algorithm. The testing was conducted using 10% of the test data, which consisted of 151 samples. The results of the experiment showed that the MLP model provided the best performance with the highest accuracy of 94.04%, followed by SVM and KNN with accuracies of 93.38% each. These results indicate that MLP is superior in recognizing non-linear patterns in anthropometric data and has better generalization capabilities than the other two algorithms. Thus, the MLP model can be used as an effective alternative in machine learning-based nutritional status classification systems to support analysis and early detection of nutritional problems in the health sector.

KEYWORDS: Nutritional Status, SVM, KNN, MLP, Classification

I. INTRODUCTION

Nutritional status serves as an indicator of the adequacy of daily food intake and reflects the body's balance, which can be expressed in the form of certain variables [1]. Nutritional imbalance, whether deficiency or excess, can lead to serious diseases. Normal nutritional status contributes to increased life expectancy, while nutritional deficiencies or excesses in adulthood can increase the risk of disease and reduce productivity. Common nutritional problems include stunting and overnutrition [2]. Therefore, an accurate and efficient nutritional status classification system is needed to support early diagnosis.

Machine learning-based classification approaches are now one of the most widely used alternatives in the data classification process [3]. Machine learning is a field of machine learning in which systems are given a set of specific algorithms so that they can learn from data and perform tasks automatically without having to be explicitly programmed [4]. In its development, machine learning techniques have been widely used for data analysis, with algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Multi-Layer Perceptron (MLP). SVM has good generalization capabilities on small datasets, KNN excels due to its simplicity, while MLP is capable of capturing complex patterns through hidden layer optimization. However, comprehensive comparative studies between these three methods for classifying the nutritional status of adults are still rare.

Previous studies have compared various machine learning models, including research conducted by Febrayanti (2025) which discussed the comparative performance of Support Vector Machine (SVM), Random Forest, and Logistic Regression algorithms in performing classification of stunting status in toddlers. The dataset was divided using a 70:30 and 80:20 ratio with. stratified sampling and validated using 5-fold cross-validation. The results showed that SVM provided the best performance with an accuracy of 92%, precision of 91%, recall of 99%, f1-score of 95%, and AUC of 99%, followed by Random Forest with an accuracy of 91% and AUC of 98%, and Logistic Regression with an accuracy of 92% and AUC of 97% [5].

Another study conducted by Safira (2025) focused on developing a prediction model for the nutritional status of children aged 0–5 years to detect the risk of malnutrition early. The SVM model with a linear kernel showed the best performance with an accuracy of 98.10% after applying the feature selection technique [6].

Another study conducted by Loka (2025) used mass weighing data of toddlers in Solok City with attributes of Gender, Age, Weight, Height, and Weight/Height Status as labels (total 1,089 data; 989 training and 100 testing) showed that the KNN algorithm produced

an accuracy of 96.10%, while the NBC algorithm obtained 90.94%. Thus, KNN proved to be superior to NBC in classifying the nutritional status of toddlers based on the BB/TB parameter [7].

This study aims to find the most suitable model for the in-depth comparative nutritional status dataset between SVM, KNN, and MLP in classifying the nutritional status of adults using anthropometric data from the RSUD Kota Banjar. This study evaluates the performance of each model based on various metrics such as accuracy, precision, recall, and F1-score to determine the most effective method.

To achieve this objective, the proposed approach integrates data preprocessing, normalization, and the division of train, testing, and validation data to ensure a balanced evaluation. The MLP model is designed with hyperparameter optimization, including learning rate and number of epochs. The results of this comparison aim to highlight the advantages and disadvantages of each algorithm when applied to real anthropometric data.

II. METHOD

Research methods are the stages of the research process that are arranged based on interrelated steps to solve a problem [8]. The methods in this study will include several systematic stages, including a nutritional status classification dataset, data preprocessing, including data cleaning, data normalization, feature scaling, label encoding, data splitting, and continued application to the MLP, SVM, and KNN *machine learning* models, followed by model testing.

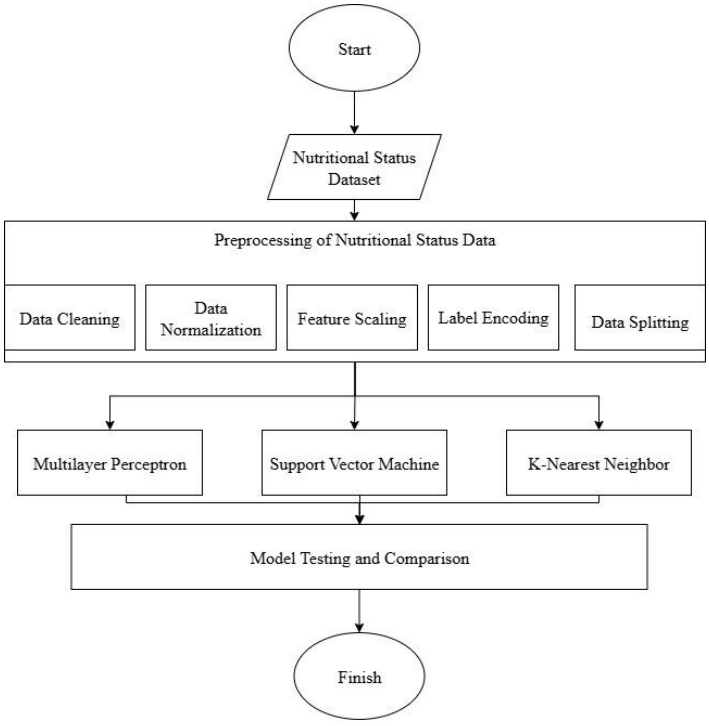

Figure 1: Comparison of MLP, SVM, and KNN models for nutritional status classification

Figure 1 illustrates that the methods used include entering the nutritional status dataset, which involves entering hardcopy data into an Excel file and converting it to CSV format. Data preprocessing consists of data cleaning, data normalization, feature scaling, label encoding, and data splitting. The results of data preprocessing are then fed into the model for the training stage of the MLP, SVM, and KNN models. Finally, each model is tested and compared to determine which model performs best in processing nutritional status classification data.

A. Nutritional Status Dataset

The data source used in this study was obtained from the Banjar City Regional General Hospital, which is one of the health facilities in the city of Banjar. The dataset is the result of manual nutritional status classification performed by nutritionists on their patients. The data used includes patient anthropometric information, such as weight, height, Body Mass Index, and Nutritional Status. The data was obtained from the medical records of adult patients during the period from January 2024 to March 2025, totaling 1504 data points.

Table 1. Nutritional status classification dataset table

No.	BB	Height	BMI	Nutritional Status
1	55	1.55	22.89	NORMAL
2	60	1.60	23.44	NORMAL
3	50	1.50	22.22	NORMAL
4	50	1.65	18.37	NORMAL
...	...	...	...	...
1504	82	1.58	32.85	OBESITY

Table 1 shows an example of the data used in the training process for the SVM, KNN, and MLP models. The data contains anthropometric information on patients, consisting of weight, height, Body Mass Index (BMI), and nutritional status as classification labels. This dataset then undergoes a pre-processing stage to ensure data quality and consistency before being used in model training. Through this process, the model is expected to be able to recognize data patterns well so that it can make accurate predictions on new data.

B. Data Preprocessing

The data preprocessing stage is the process of converting raw data into a well-organized dataset so that it can be used effectively in the model training process [9]. The preprocessing stage is carried out before the data is used in the model training, validation, and testing processes [10]. This stage includes cleaning the data by removing empty values and duplicate data, as well as standardizing the decimal format in the Body Mass Index column. The labels in the Nutritional Status column are also standardized by removing spaces and capitalizing letters for consistency. Next, feature standardization is performed using Standard Scaler to normalize the numerical variables of weight, height, and BMI so that they have a comparable scale. The label encoding process is applied to convert nutritional status categories into a numerical form that can be recognized by machine learning algorithms. After that, the dataset was divided into three parts, namely training data (70%), validation data (20%), and test data (10%) using the stratified sampling method so that the proportion of each class remained balanced. Finally, the scaler and label encoder objects were saved in .pkl format to maintain consistency in the training and prediction stages of the model.

C. SVM, KNN, and MLP Model Architecture

Model architecture describes the process of processing, sending, and analyzing data in machine learning algorithms, which play an important role in determining the efficiency and effectiveness of models in solving problems [11]. Each model has a different approach and working principle in mapping the relationship between features (input) and labels (output). In this study, three main algorithms, namely SVM, KNN, and MLP, were used to classify nutritional status.

Support Vector Machine (SVM) is one of the most widely used classification algorithms due to its ability to apply various types of kernel functions to separate data based on its characteristics [12]. This approach has proven effective in various fields, such as classification, regression, time series forecasting, and estimation processes [13]. This algorithm utilizes kernel functions such as linear, polynomial, and Radial Basis Function (RBF) to handle non-linear data by mapping features to a higher-dimensional space. In this study, SVM was used to classify nutritional status based on anthropometric data such as weight, height, and Body Mass Index (BMI).

K-NN is a classification algorithm in data mining that works by grouping new data based on its proximity to training data that already has a specific label or class [14]. In the KNN regression method, Euclidean distance calculations are performed first, then the degree of proximity between the data is analyzed to determine the relationship between the data points [15]. In this method, new data will be classified into the majority class of its K nearest neighbors, which are calculated using distance measures such as Euclidean Distance. The K value is an important factor because it determines the sensitivity of the model to the data; a K value that is too small can cause overfitting, while a value that is too large can blur the boundaries between classes. In this study, KNN was used to group individuals' nutritional status based on the similarity of their anthropometric values, so that similar data distribution patterns produced consistent nutritional class predictions. The advantages of KNN lie in its non-parametric nature and ease of implementation, although its computation time increases as the amount of training data increases.

The MLP model has good capabilities in classifying data with more than two categories or multi-class [16]. MLP is able to learn nonlinear relationships between input and output variables through feedforward and backpropagation processes, where network weights are updated based on prediction errors to minimize the loss function. Activation functions such as ReLU, sigmoid, or tanh are used to introduce non-linearity into the model. In this study, MLP was applied to predict nutritional status based on weight, height, and BMI data, with the aim of improving the model's ability to recognize complex patterns that cannot be captured by linear algorithms.

D. Model Testing and Comparison

Testing is conducted to determine how well the model classifies the nutritional status dataset. Testing will also help in comparing the SVM, KNN, and MLP models in terms of accuracy, precision, and generalization ability. Testing also serves to evaluate the performance of each SVM, KNN, and MLP model.

A confusion matrix is a table commonly used to describe or explain the performance of a model in machine learning [17]. The confusion matrix is used as a tool to evaluate model performance by grouping prediction results and actual data into four main categories, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [18]. Based on this matrix, several evaluation metrics are calculated, such as accuracy, precision, recall, and F1-score, to assess the accuracy of the model in classifying nutritional status. These metrics are used to compare the performance of the SVM, KNN, and MLP algorithms, so that the model that provides the most optimal results for the anthropometric data used can be identified. Precision is a quantitative measure that evaluates the accuracy of data by considering the quality of information and the model's ability to produce accurate predictions [20]. The results of this evaluation provide an overview of how well the system classifies the dataset [19].

III. RESULTS AND DISCUSSION

All paragraphs must be indented. All paragraphs must be justified.

A. Dataset

The dataset section presents data visualization in the form of graphs to provide an initial understanding of the characteristics and patterns of the dataset used. The purpose of this visualization is to show the relationship between the variables of weight, height, and Body Mass Index, as well as to display the distribution of nutritional status classes as labels in the classification process. Through the relationship between variables, a tendency for an increase in BMI values can be seen along with an increase in weight and height. The class distribution graph shows the proportion of data in each nutritional status category, such as normal, overweight, obese, underweight, and malnourished. This visual analysis provides an overview of the data balance and its potential influence on the performance of the machine learning model in the next stage of analysis.

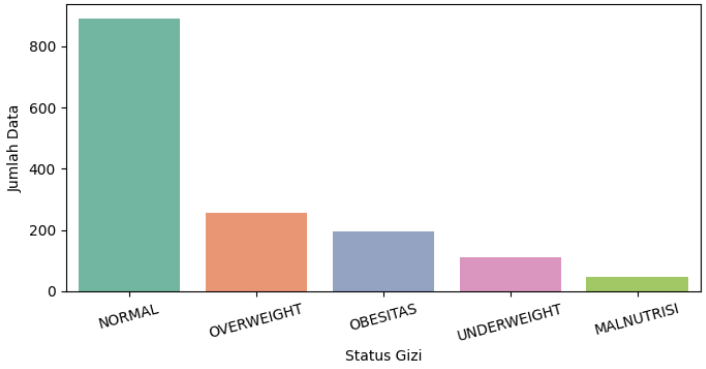

Figure 2. Distribution of nutritional status classes

Figure 2 illustrates the distribution of data numbers in each nutritional status category. Based on the graph, it can be seen that the Normal class dominates the amount of data compared to other categories, followed by Overweight and Obesity, while Underweight and Malnutrition have relatively little data. This imbalance in the number of samples indicates class imbalance in the dataset, which has the potential to affect the performance of the classification model, especially in algorithms that are sensitive to differences in proportions between classes.

B. Data Preprocessing

The data preprocessing stage was carried out to ensure that the dataset used in this study was clean, consistent, and ready for model training. This process included several steps, such as data cleaning, normalization, feature scaling, label encoding, and data division to support the machine learning classification process. Data cleaning is used to clean the data from empty data attributes and strange data attributes, for example in data number 8.

Table 2. Dataset table no. 8

No.	BB	TB	IMT	Status Gizi
8			#DIV/OI	NORMAL

Therefore, after data cleaning, dataset No. 8 will be deleted so that it does not interfere with the model's performance in classifying the dataset.

Data normalization is performed to equalize the scale between numerical variables so that each feature has a balanced contribution in the model training process. This process uses the StandardScaler method, which converts the value of each feature into a distribution with a mean of zero and a standard deviation of one. Thus, the variables of weight, height, and Body Mass Index are on the same scale, so that the model is not biased towards features with a larger range of values.

Table 3. Data before standard scaler

Weight (kg)	Height (m)	BMI
55	1.55	22.89
60	1.60	23.44
50	1.50	22.22
50	1.65	18.37

Before feature scaling, the values of each numerical variable, such as weight, height, and body mass index, still have different units and ranges. Differences in scale between variables can cause machine learning models to give greater weight to features with larger numerical values, thereby affecting learning outcomes and causing bias in the classification process.

Table 4. Data after standard scaler

BB	TB	BMI
-0.228510	-0.273162	-0.138406
0.177278	0.582891	-0.008899
-0.634299	-1.129215	-0.296170
-0.634299	1.438944	-1.202723

After performing feature scaling using the Standard Scaler method, all numerical variables have been converted to a standard scale with a mean of 0 and a standard deviation of 1. This transformation ensures that each feature has an equal contribution in the model training process and helps algorithms such as SVM, KNN, and MLP to calculate distances and optimize parameters more efficiently. After all numerical features have undergone the scaling process, the next step is to perform label encoding on the categorical variable NUTRITIONAL STATUS. This process aims to convert categorical data into numerical form so that it can be processed by machine learning algorithms. Each nutritional status category, which was originally in text form, such as NORMAL, OBESE, or UNDERWEIGHT, is converted into a number that represents a specific class.

Table 5. Nutritional status attribute data before and after encoding.

No	NUTRITIONAL STATUS	NUTRITIONAL_STATUS_ENC
1	NORMAL	1
2	OVERWEIGHT	3
3	MALNUTRITION	0
4	OBESITY	2
5	UNDERWEIGHT	4

Based on the results in Table 5, each nutritional status category was successfully converted into a unique numerical value. These values were used as class representations in the model training and testing process. With this label encoding process, the model can recognize numerical patterns between classes more efficiently without losing the meaning of each nutritional status category. In addition, the encoding results also ensure label consistency throughout the training, validation, and prediction stages of the classification model**.**

After the label encoding process is complete, the next step is to divide the dataset into three subsets, namely training data, validation data, and test data. This division aims to evaluate the model's performance objectively and avoid overfitting. The stratified sampling method is used so that the proportion of each class in the NUTRITIONAL STATUS label remains balanced in the three subsets. The visualization in the following figure shows the label distribution after the data division process.

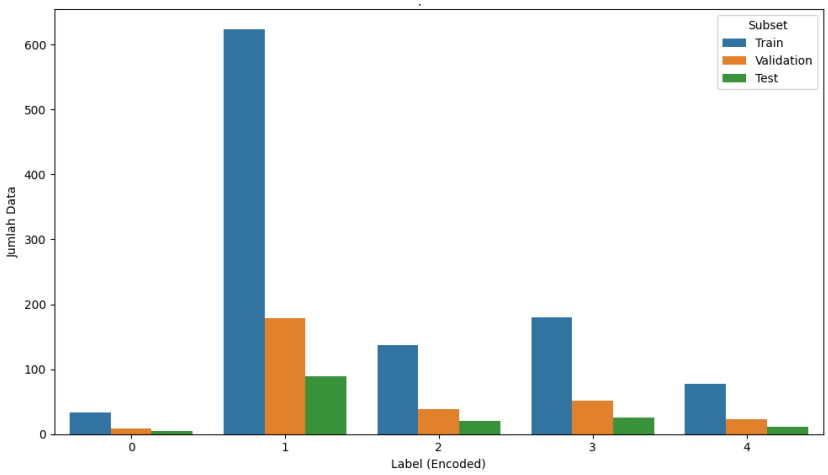

Figure 3. Data Splitting

Figure 3 visualizes the comparison of data used in the model learning process. The data is divided into three main parts: training data (70% or 1051 samples), validation data (20% or 301 samples), and test data (10% or 151 samples) from a total of 1503 data. Dividing the dataset into these three parts is important to avoid overfitting and ensure that the model performs well on new data.

C. Support Vector Machine Results

The results of testing the SVM model with various parameter combinations are shown in the following table. Testing was conducted to determine the combination of kernel, C, and gamma that produced the best performance on the training and validation data.

Table 6. Model Testing with Several Combinations

No	Kernel	C	Gamma	Train Accuracy	Val Accuracy
0	rbf	10.0	scale	94.67	91.69
1	rbf	10.0	auto	94.67	91.69
2	rbf	1.0	auto	93.63	90.70
3	rbf	1.0	scale	93.53	90.70
4	linear	10.0	auto	94.58	90.37
5	linear	10.0	scale	94.58	90.37
6	linear	0.1	scale	92.29	89.04
7	linear	0.1	auto	92.29	89.04
8	linear	1.0	scale	93.82	88.37
9	linear	1.0	auto	93.82	88.37
10	poly	10.0	auto	89.15	87.04

Based on the results in Table 6, the SVM model with the RBF kernel provides the best performance compared to other kernel types. The combination of parameters C = 10.0 and gamma = scale produces a training accuracy value of 94.67% and a validation accuracy of 91.69%, indicating excellent generalization capabilities. The linear kernel also shows quite competitive performance with a validation accuracy above 90%, tends to be slightly lower than the RBF kernel. Meanwhile, the polynomial and sigmoid kernels produce lower accuracy, ranging from 81% to 87% and 61% to 75%, respectively, indicating that these two kernels are less optimal in mapping nonlinear patterns in anthropometric data. Overall, these results indicate that the RBF kernel is most effective for nutritional status classification, as it is better able to capture the nonlinear relationship between weight, height, and BMI variables

Table 7. Classification Results for Each SVM Model Class

No	Precision	Recall	F1-Score	Support
0	0.67	0.80	0.73	5
1	0.99	1.00	0.99	89
2	0.89	0.85	0.87	20
3	0.85	0.88	0.87	26
4	0.89	0.73	0.80	11

ased on the results in Table 7, the SVM model shows excellent overall classification performance. The class with label 1 Normal has the highest precision, recall, and f1-score values, namely 0.99, 1.00, and 0.99, respectively, indicating that the model is able to

recognize the normal nutrition category with a very high level of accuracy. Classes 2 Obesity and 3 Overweight also show stable performance with f1-scores above 0.85, indicating a balance between detection capability and prediction accuracy. Meanwhile, performance in classes 0 Malnutrition and 4 Underweight was slightly lower, especially in terms of recall, indicating that the model still had difficulty recognizing classes with relatively little data. Overall, these results show that the SVM model is classifying nutritional status well, especially in majority classes with high precision and generalization.
Recommended font sizes are shown in Table 1.

D. K-Nearest Neighbor Results

The results of testing the KNN model with various parameter combinations are shown in the following table. Testing was conducted to determine the combination of n_neighbors, weights, and metrics that produced the best performance on the training and validation data.

Table 8. Testing of the KNN Model with Several Combinations

No	n_neighbors	weights	metric	Train Accuracy	Val Accuracy
0	3	distance	Manhattan	96.29	92.36
1	9	distance	Euclidean	97.05	92.03
2	11	distance	Euclidean	97.05	92.03
3	7	distance	Manhattan	97.05	92.03
4	11	distance	Manhattan	97.05	92.03
5	9	distance	Manhattan	97.05	92.03
6	3	uniform	Manhattan	95.05	91.69
7	3	distance	Euclidean	96.29	91.69
8	7	distance	Euclidean	97.05	91.69
9	5	distance	Manhattan	97.05	91.69

Based on the results in Table 8, the KNN configuration with n_neighbors = 3 and distance weighting using the Manhattan metric provides the best validation results with an accuracy of 92.36%. In general, all parameter variations show stable performance with validation accuracy above 91%, indicating that the KNN model is classifying nutritional status data well. The combination of distance weighting has been proven to improve prediction accuracy compared to the uniform weighting method, because gives greater weight to neighbors that are closer to the test point.

Table 9. Classification Results for Each KNN Model Class

No	Precision	Recall	F1-Score	Support
0	0.67	0.80	0.73	5
1	0.99	1.00	0.99	89
2	0.83	0.95	0.88	20
3	0.91	0.81	0.86	26
4	0.89	0.73	0.80	11

Based on the results in Table 9, the K-Nearest Neighbor (KNN) model shows good classification performance with high accuracy and consistency in most classes. Class 1 (Normal) has the highest precision, recall, and f1-score values of 0.99, 1.00, and 0.99, respectively, indicating that the model is very accurate in recognizing data with normal nutritional status. Classes 2 (Obesity) and 3 (Overweight) also show fairly stable performance with f1-scores of 0.88 and 0.86, respectively, indicating a balance between the model's ability to accurately detect and predict these categories. Meanwhile, performance in classes 0 (Malnutrition) and 4 (Underweight) is slightly lower, with recall values below 0.80, due to the relatively small amount of data and the similarity of anthropometric patterns between classes. Overall, these results indicate that the KNN model is capable of classifying nutritional status well, especially in the majority classes, and demonstrates good generalization ability without significant overfitting.

E. Multilayer Perceptron Results

The results of testing the MLP model with various parameter combinations are shown in the following table. Testing was conducted to determine the combination of learning rate and epoch that produced the best performance on the training and validation data.

Table 10. Classification Results for Each MLP Model Class

Epoch	Learning Rate	Train Accuracy	Train Loss	Val Accuracy	Val Loss
20	0.1000	90.1	0.2825	87.38	0.3678
20	0.0100	92.39%	0.1842	90.7	0.3210

20	0.0010	89.91%	0.2509	90.03%	0.2984
20	0.0001	71.93%	0.9185	70.76%	0.8982
50	0.1000	89.82%	0.3694	89.04%	0.3406
50	0.0100	92.96%	0.1637	91.69%	0.2936
50	0.0010	93.05%	0.1834	90.37%	0.2969
50	0.0001	83.25%	0.5111	83.39%	0.4861
100	0.1000	88.11%	0.3151	86.38%	0.3059
100	0.0100	94.58%	0.1399	87.71%	0.4388
100	0.0010	93.91%	0.1531	91.03%	0.3160
100	0.0001	87.73%	0.3229	88.04%	0.3527

Based on the results in Table 10, it can be seen that changes in the learning rate and number of epochs have a significant effect on the performance of the MLP. The model with a learning rate of 0.0100 and 50 epochs produced the best performance with a validation accuracy of 91.69% and a validation loss of 0.2936, indicating a good balance between convergence speed and generalization ability. A learning rate that is too large, such as 0.1, causes higher loss and fluctuations in the learning process, while a value that is too small, such as 0.0001, slows down the convergence process and reduces accuracy. Overall, these results show that selecting the appropriate learning rate and number of epochs is very important to achieve optimal performance in the MLP model for nutritional status classification.

Table 11. Classification Results for Each Class of the MLP Model

No	Precision	Recall	F1-Score	Support
0	0.80	0.80	0.80	5
1	0.99	1.00	0.99	89
2	0.89	0.85	0.87	20
3	0.85	0.88	0.87	26
4	0.90	0.82	0.86	11

Based on the results in Table 11, the Multilayer Perceptron (MLP) model shows excellent overall classification performance. Class 1 (Normal) has the highest precision, recall, and f1-score values of 0.99, 1.00, and 0.99, respectively, indicating that the model is able to recognize normally labeled data with near-perfect accuracy. Classes 2 (Obesity) and 3 (Overweight) also show stable results with an f1-score of 0.87, indicating a good balance between accuracy and detection success in these classes. Classes 0 Malnutrition and 4 Underweight had slightly lower performance with f1-scores below 0.85, which was likely influenced by the smaller amount of data in these categories. Overall, the MLP model achieved a good level of generalization and was able to capture the non-linear relationship patterns between anthropometric variables such as weight, height, and Body Mass Index (BMI) in the nutritional status classification process.

F. Comparison

The comparison was made by testing the accuracy using test data. The test results show that each model has different performance characteristics. The results are visualized in Figure 4.

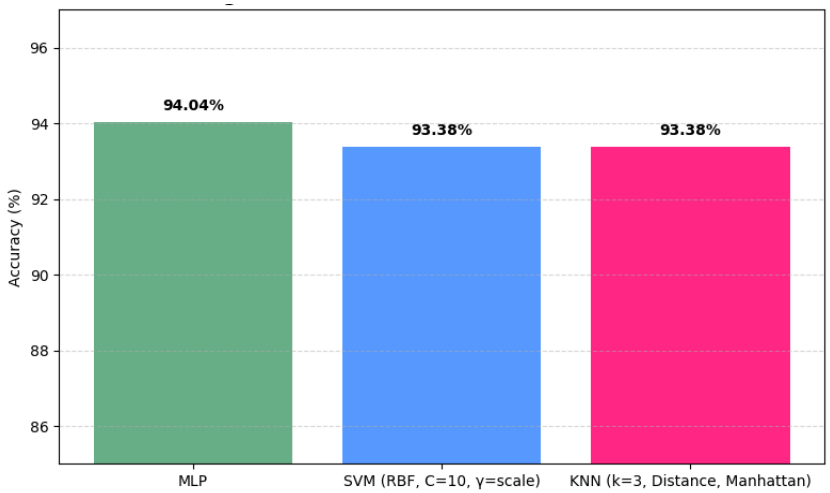

Figure 4. Comparison of test data accuracy

Figure 4 shows the results of comparing the accuracy levels of the three classification models, namely MLP, SVM, and KNN in the combined train and validation training process. Based on the test results, the MLP model showed the best performance with the highest accuracy of 94.04%, followed by SVM and KNN, which each achieved 93.38%. This difference in accuracy values indicates that MLP has better generalization capabilities in recognizing non-linear patterns in anthropometric data, such as the relationship between weight, height, and Body Mass Index. Meanwhile, SVM and KNN still show competitive results, with advantages in prediction stability and ease of application. Overall, all three algorithms performed well, but MLP proved to be the most optimal in producing a nutritional status classification model with consistent accuracy.

Table 12. Comparison of test data accuracy

Model	Accuracy	Best Performing Model
SVM	93.38	Parameters C = 10.0 and gamma = 'scale' were used with the RBF kernel
KNN	93.38	n_neighbors = 3 and distance-based weighting using the Manhattan distance metric
MLP	94.04	Learning rate 0.0100 and 50 epochs

Based on the results shown in Table 12, the performance of each model was evaluated using test data to assess its generalization ability. The MLP model achieved the highest test accuracy of 94.04%, surpassing the SVM and KNN models, which achieved accuracies of 93.38% respectively. The best configuration for the MLP model was achieved with a learning rate of 0.0100 and 50 epochs, which showed an effective balance between convergence speed and accuracy. In comparison, the SVM model achieved optimal performance with a combination of parameters C = 10.0 and gamma = 'scale' using the RBF kernel, while the KNN model showed the best results with n_neighbors = 3 and distance-based weighting using the Manhattan metric. Overall, these results confirm that the MLP model provides the best combination of learning stability and prediction accuracy when tested on previously unseen data.

IV. CONCLUSION

This study aims to find the most suitable model for the nutritional status dataset through an in-depth comparison between the SVM, KNN, and MLP algorithms in classifying the nutritional status of adults based on anthropometric data from the RSUD Kota Banjar. The comparison was conducted using 10% of the total dataset as test data, consisting of 151 samples. The results showed that the MLP model provided the best performance with the highest accuracy of 94.04%, followed by SVM and KNN, which achieved 93.38% accuracy. These results prove that MLP has a better ability to recognize non-linear patterns between the variables of weight, height, and Body Mass Index, resulting in a high level of generalization. These findings indicate that neural network-based models can be used as an effective alternative in nutritional status classification systems to support the diagnosis process in the health sector. For further research, it is recommended to expand the amount of data and implement deep learning or reinforcement learning methods to determine which model can classify nutritional status even better.

ACKNOWLEDGMENT

The author would like to thank Yogyakarta University of Technology for providing facilities and a conducive academic environment during this research. Thanks are also extended to the Banjar City Regional General Hospital for granting permission and providing the anthropometric data used in this study. The support from both parties was instrumental in the implementation of this research on nutritional status classification.

REFERENCES

1) N. R. Aulia, "The Role of Nutrition Knowledge on Energy Intake, Nutritional Status, and Attitudes Towards Nutrition in Adolescents," Journal of Nutrition and Health (JIGK), vol. 2, no. 02, pp. 31–35, Feb. 2021, doi: 10.46772/jigk.v2i02.454. Wulandari et al., "The Relationship Between Nutrition Knowledge Level and Nutritional Status in Students at Ibn Khaldun University Bogor," Tropical Public Health Journal, vol. 1, no. 2, pp. 72–75, Sep. 2021, doi: 10.32734/trophico.v1i2.7266.
2) *Baharuddin and A. Tjahyanto, "Improving Machine Learning Classification Performance Through Comparison of Machine Learning Methods and Dataset Enhancement," Jurnal Sisfokom (Information Systems and Computers), vol. 11, no. 1, pp. 25–31, Mar. 2022, doi: 10.32736/sisfokom.v11i1.1337.*
3) M. Y. Saputra and I Wayan Santiyasa, "Optimizing Machine Learning Performance with The Naive Bayes Classifier Process in Smart Farming," Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), vol. 13, no. 2, pp. 382–393, Jul. 2024, doi: 10.23887/janapati.v13i2.76926.
4) N. R. Febriyanti, K. Kusrini, and A. D. Hartanto, "Comparative Analysis of SVM, Random Forest, and Logistic Regression Algorithms for Predicting Stunting in Toddlers," Edumatic: Journal of Informatics Education, vol. 9, no. 1, pp. 149–158, Apr. 2025, doi: 10.29408/edumatic.v9i1.29407.

5) R. Safira and I. P. Sari, "Prediction Model of Child Growth and Development for the Prevention of Malnutrition Risk Using Support Vector Machine (SVM) with Backward Elimination Feature Selection," Blend Sains Jurnal Teknik, vol. 4, no. 1, pp. 1–11, Jul. 2025, doi: 10.56211/blendsains.v4i1.812.
6) S. K. P. Loka and A. Marsal, "Comparison of K-Nearest Neighbor and Naïve Bayes Classifier Algorithms for Classifying the Nutritional Status of Toddlers," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 3, no. 1, pp. 8–14, May 2023, doi: 10.57152/malcom.v3i1.474.
7) W. Harma, F. Hadi, and D. Kartika, "Business Digitalization in the Marketing Strategy of BSF Maggots in Anak Nagari Agribusiness Using the Mamdani Fuzzy Method," Jurnal KomtekInfo, pp. 11–17, Mar. 2024, doi: 10.35134/komtekinfo.v11i1.497.
8) T. A. Alghamdi and N. Javaid, "A Survey of Preprocessing Methods Used for Analysis of Big Data Originated From Smart Grids," IEEE Access, vol. 10, pp. 29149–29171, 2022, doi: 10.1109/ACCESS.2022.3157941.
9) Kadek Eka Sapta Wijaya, Gede Angga Pradipta, and Dadang Hermawan, "The Implementation of Bayesian Optimization for Automatic Parameter Selection in Convolutional Neural Network for Lung Nodule Classification," Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), vol. 13, no. 3, pp. 438–449, Dec. 2024, doi: 10.23887/janapati.v13i3.82467.
10) J. Kufel et al., "What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine," Diagnostics, vol. 13, no. 15, p. 2582, Aug. 2023, doi: 10.3390/diagnostics13152582.
11) D. Isla-Cernadas, M. Fernández-Delgado, E. Cernadas, M. S. Sirsat, H. Maarouf, and S. Barro, "Closed-Form Gaussian Spread Estimation for Small and Large Support Vector Classification," IEEE Trans Neural Netw Learn Syst, vol. 36, no. 3, pp. 4336–4344, Mar. 2025, doi: 10.1109/TNNLS.2024.3377370.
12) Bunga Tiara, A. M. Siregar, D. S. K. Kusumaningrum, And T. Rohana, "Bank Customer Segmentation Model Using Machine Learning," National Journal Of Informatics Education (Janapati), Vol. 13, No. 1, Mar. 2024, Doi: 10.23887/Janapati.V13i1.75233.
13) wulandari, w. j. sari, z. alfian, l. legito, and t. arifianto, "implementation of naïve bayes classifier and k-nearest neighbor algorithms for chronic kidney disease classification," malcom: indonesian journal of machine learning and computer science, vol. 4, no. 2, pp. 710–718, apr. 2024, doi: 10.57152/malcom.v4i2.1229.
14) R. Yunis, Sudarto, and I. Adiputra Pardosi, "Enhancing Rice Production Prediction: A Comparative Machine Learning Analysis of Climate Variables," Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), vol. 13, no. 1, Mar. 2024, doi: 10.23887/janapati.v13i1.71527.
15) Y.-H. Cheng, M.-H. Shih, and C.-N. Kuo, "Early Prediction of Student Suspension and Dropout for Counseling Resource Optimization Using Multilayer Perceptron," IEEE Access, vol. 13, pp. 119810–119819, 2025, doi: 10.1109/ACCESS.2025.3587484.
16) K. Kristiawan and A. Widjaja, "Comparison of Machine Learning Algorithms in Assessing a Retail Store Location," Journal of Informatics and Information Systems Engineering, vol. 7, no. 1, Apr. 2021, doi: 10.28932/jutisi.v7i1.3182.
17) N.-C. Yang and K.-L. Sung, "Non-Intrusive Load Classification and Recognition Using Soft-Voting Ensemble Learning Algorithm With Decision Tree, K-Nearest Neighbor Algorithm and Multilayer Perceptron," IEEE Access, vol. 11, pp. 94506–94520, 2023, doi: 10.1109/ACCESS.2023.3311641.
18) M. S. Nugraha and F. I. Sanjaya, "Classification of Nutritional Status Using the Fuzzy Mamdani Method: Case Study at RSUD Kota Banjar," Journal of Applied Informatics and Computing, vol. 9, no. 4, pp. 1498–1505, Aug. 2025, doi: 10.30871/jaic.v9i4.9893.
19) Merinda Lestandy, Abdurrahim Abdurrahim, Amrul Faruq, M. Irfan, and Novendra Setyawan, "Ensembled Machine Learning Methods and Feature Extraction Approaches for Suicide-Related Social Media," Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), vol. 13, no. 2, pp. 192–203, Jul. 2024, doi: 10.23887/janapati.v13i2.70016.