Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2024, 2024:25
https://doi.org/10.28919/cmbn/8429
ISSN: 2052-2541

PERFORMANCE TEST OF NAIVE BAYES AND SVM METHODS ON CLASSIFICATION OF MALNUTRITION STATUS IN CHILDREN

DEVIE ROSA ANAMISA*, ACHMAD JAUHARI, FIFIN AYU MUFARROHA

Department of Informatics, Faculty of Engineering, University of Trunojoyo Madura, Bangkalan 69162, Indonesia

Abstract: Paying attention to children's nutrition is very important because children's daily activity is relatively high. Therefore, one of the efforts that the Indonesian Government must make is to reduce the incidence of malnutrition because high malnutrition can increase the death rate. In recent years, Madura Island has experienced an increase in the number of children suffering from stunting, especially in the Sumenep area. Therefore, it is necessary to collect data to handle it quickly and validly. However, to explore data, classification techniques are needed. Given these problems, this research has carried out a comparative analysis with two classifier methods, namely Naïve Bayes and Support Vector Machine (SVM) to categorize malnutrition in children. Both methods have advantages. The Naïve Bayes method can be used to make predictions based on the probability of members of a malnutrition category class. Meanwhile, SVM can classify based on the kernel to form the best hyperplane on the input data. Based on 694 data on malnutrition in children, the SVM method has produced the best level of accuracy with a value of 89.76% with a Kernel Polynomial at a Cost (C) of 5 compared to the Naïve Bayes method of 86.31%. Thus, it can be concluded that the SVM method can classify malnutrition very well.

Keywords: classification; malnutrition; methods; Naïve Bayes; SVM.

2020 AMS Subject Classification: 92C15.

---

*Corresponding author
E-mail address: devros_gress@trunojoyo.ac.id
Received January 05, 2024

1. INTRODUCTION

Madura is the closest island to Java, but developments in various sectors are very far compared to developments in Java [1]. Therefore, the Government needs to pay special attention to the Madura region in developing community welfare, especially in the health sector, namely malnutrition. This is because nutritional problems impact the quality of human resources. Based on the results of primary health research in 2013, the stunting rate had reached 37.2%, with a suboptimal growth rate of 8.9 million Indonesian children [2]. This shows that cases of malnutrition are pretty high in Indonesia, especially in Sumenep, Madura. Apart from that, over a long period, there have been many cases of children experiencing malnutrition, which has had an impact on the amount of data piling up. Based on this case, health agencies in Madura experienced difficulties in classifying nutritional data for children with cases of malnutrition.

Therefore, in this research, we designed a mining technique to solve this problem using the Classifier method. In this research, Naive Bayes and SVM methods were applied. In this research, we have chosen these two methods because the Naïve Bayes method has advantages in determining probability values in predicting opportunities using previous experience data [3]. So, the Naive Bayes method is very suitable for categorizing malnutrition in children. The Naïve Bayes method is simple and has fast computing time for finding models [4]. The SVM method also has several advantages, including producing a level of transition accuracy that depends on the kernel function and parameters used [5]. Also, SVM is divided into two types based on characteristics: Linear Support Vector Machine and Non-Linear Support Vector Machine [6].

Based on previous research, these two methods have also been compared with other classification methods, such as research conducted by [7] regarding the comparison of Naïve Bayes and K-Nearest Neighbor (KNN) for the classification of Indonesian articles, where the Naive Bayes method has better performance than KNN. Research [8] about The Naïve Bayes method for predicting the smooth level of terrace rental payments for Micro, Small, and Medium Enterprises has produced an accuracy of 81.81%. Meanwhile, in research [9] regarding comparing the Naïve Bayes and Random Forest (RF) methods for regional language classification, the analysis results in this study produce the Naïve Bayes method as better than RF. Apart from the Naïve Bayes method, which has high accuracy, the SVM method produces a small error rate, as in research [10] regarding the comparison of SVM and Decision Tree for tourist attraction recommendation systems, where the SVM method has better performance than the Decision Tree method. It has also been proven in research [11] regarding comparing the KNN and SVM methods for air quality classification in Jakarta, and the SVM method is superior to KNN with 100 kernels.

Based on previous references and the problems above, this research tries to apply the Naïve Bayes and SVM methods to find the best and most effective method for classifying malnutrition in children. Apart from that, this system can assist health agencies in registering children with malnutrition so that they can immediately receive treatment by providing vitamins or counseling for the child's family.

2. PRELIMINARIES

Classification is grouping similar objects and also separating objects that are not the same [12]. This classification process is an essential part of an information organization to make it easier to access information. From the definition above, the problem in this research is everything related to data mining in the form of nutritional data from malnutrition in children. The data in this research was obtained from the Kalianget Community Health Center. This malnutrition data consists of 694 datasets obtained in 2016-2022 with six attributes involved, including gender, child's age, birth weight, sbirth height, and current weight, as in Table 1. Then, a data standardization process is carried out, dividing the data into training and testing data. Training data is data used to build a model while testing data is testing a model created with other data to determine the model's accuracy [13]. After that, a classification process will be carried out using the Naive Bayes and the SVM methods to compare the two. The method comparison model in this research can be seen in Figure 1.

2.1. Data Standardization

Data standardization is initial data processing before classification, where the data is normalized with all x values transformed into Z values [14]. In this study, Z-Score was used. The Z-Score value is a measure of data deviation from the average value (μ), which is measured in standard deviation units (σ) [15]. The aim is to standardize the dataset by adjusting the balance of comparison values between data before and after processing. The equation for calculating data standardization can be seen in equation (1).

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

2.2. Classification Process

Carry out the classification process using both Naïve Bayes and SVM. Using Naïve Bayes and SVM, the classification process aims to find precise and accurate results. At this modeling stage, accuracy values from the processed data will be produced. The naïve Bayes method is a data-mining classification method based on Bayes' theorem, used with independent probability and statistical methods [16]. Bayes' theorem has equation (2). The main characteristic of Classification with Naive Bayes is a very strong (naïve) estimate of the independence of each event [17]. The flow of the Naïve Bayes method in classifying malnutrition can be seen in Figure 2.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{2}$$

where P(H|X) is the probability of hypothesis H based on condition X (posteriori probability), P(X|H) is the Probability of X based on these conditions, P(H) is the probability of hypothesis H (prior probability), and P(X) is the Probability of X.

Table 1. Dataset Gizi Malnutrition

Data	Gender (M/F)	Age (Month)	Birth Weight (Kg)	Birth Height (cm)	Weight (kg)	Height (cm)
…	…	…	…	…	….	….
11	M	66	2.9	49	14.6	104.2
12	M	61	2.9	48	15.1	105.7
13	F	59	3.2	50	13.1	98.4
14	M	53	3.2	49	11.7	86.5
15	M	48	3.5	50	13.0	96.8
16	F	46	3.0	50	12.0	92.8
17	F	46	3.5	50	12.0	96.8
18	M	45	3.3	50	12.4	86.8
19	M	44	3.7	52	14.2	94.3
20	F	40	3.4	52	11.6	87.2
…	…	…	…	…	…	…
694	F	22	2.9	48	7.8	77.0

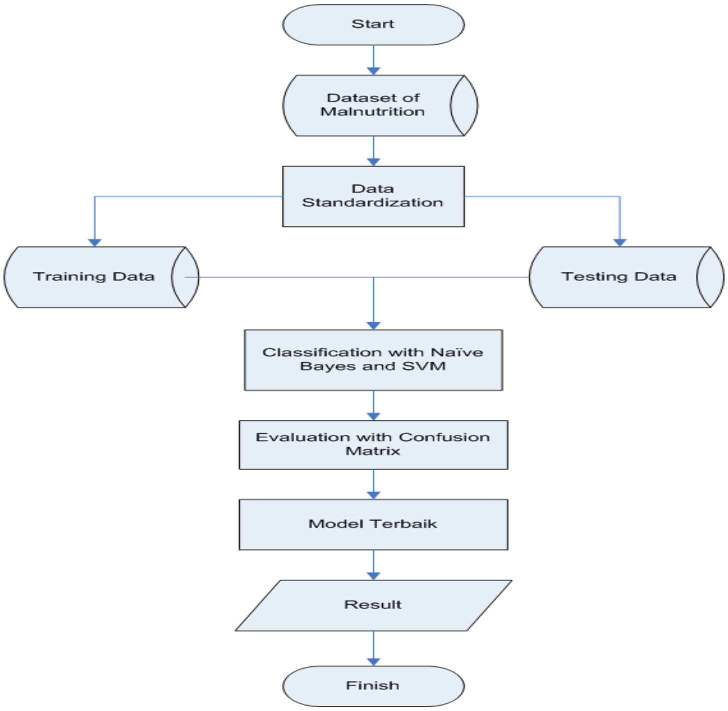

Figure 1. Comparative Design Model of Classification Methods for Malnutrition

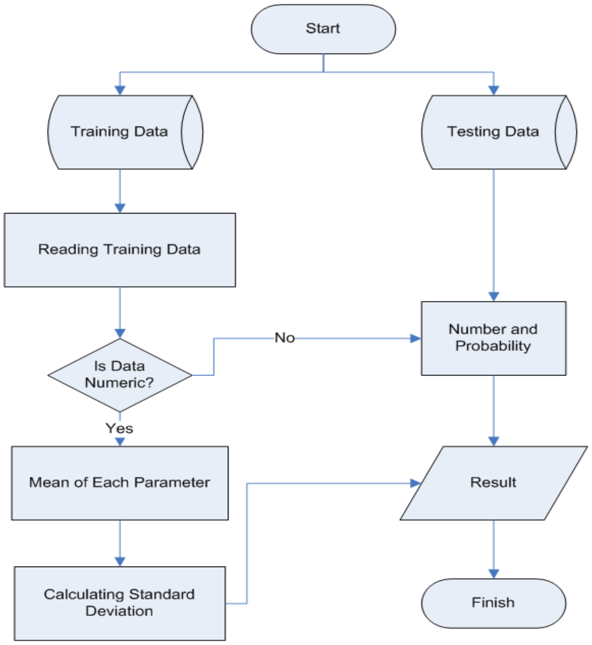

Figure 2. Naive Bayes Method Flow for Malnutrition Classification

The SVM method is a classification method that finds the best hyperplane that separates two classes in the input space [18]. This method uses hypotheses as linear functions in a high-dimensional feature space by implementing learning bias from statistical learning theory. In this research, we chose the SVM method because it has the advantage of determining distances using support vectors, so the computing process is fast. Meanwhile, in high-dimensional space, SVM can search for hyperplanes that maximize the distance (margin) between data classes [19]. The equation for determining the hyperplane can be seen in equation (3). To get the best hyperplane, look for a hyperplane in the middle between two class boundary planes or maximize the margin between two sets of objects from different classes, as seen in equation (4). After calculating the margin, the course of data to be predicted or testing data can be determined based on the function value in equation (3). The kernel function used to map the initial dimensions (lower dimensions) of the data set to new dimensions (relatively higher dimensions) can be seen in equation (5). Carry out calculating the Hessian matrix value using equation (6). Then, this process is repeated until the iteration value meets the maximum iteration limit or reaches the max value (|δα_i|) < ε (epsilon). So, the concept of the SVM method in classifying malnutrition status can be seen in Figure 3. The kernel function and regularization parameters, such as C and gamma, are the most critical hyperparameters.

$$f(x) = w^T x + b \tag{3}$$

$$d = \frac{2}{|w|} \tag{4}$$

$$K(x_i, x_j) = (1 + x^T . x_i)^2 \tag{5}$$

$$D_{ij} = y_i y_j (K(x_i, x_j) + \alpha^2 \tag{6}$$

$$E_i = \sum_{j=1}^{1} \alpha_i D_{ij}$$

$$\delta\alpha_i = \min\{\max[\gamma(1 - E_i) - \alpha_i], C - \alpha_i$$

$$\alpha_i = \alpha_i + \delta\alpha_i$$

where $[(w^T . x_i) + b \geq 1$ to $y_i = +1$ whereas $[(w^T . x_i) + b \leq 1$ to $y_i = -1$. Dan x_i is the training data set, i = 1, 2, ..., n whereas y_i = Class label of $x_{i,}$ $x_{i,}$ and x_j is a pair of two training data, α_i is Alpha parameter value, D_{ij} is the result of the Hessian matrix value, $\delta\alpha_i$ is the delta alpha parameter for i.

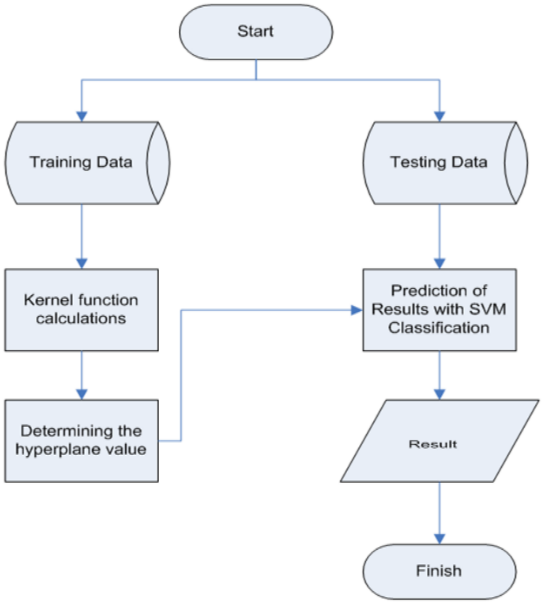

Figure 3. SVM Method of Classification Malnutrition Model

2.3. Evaluation Testing Stage

At the evaluation testing stage, ratio testing is carried out with a comparison of training data and testing data of 90%:10%, 80%:20%, 70%:30%, 60%:40%, 50%:50%, 40%:60%, 30%:70%, 20%:80%, and 10%:90%, Where this process is intended to determine the quality and accuracy values that the two methods have produced. So, the calculation of accuracy, precision, and recall values is made at this stage using the confusion matrix. Confusion Matrix is visual data containing several cases that have been classified well and correctly and cases that have been mistyped [18].

3. MAIN RESULTS

One of the reasons for the increase in child mortality is nutritional needs that are not met [20]. Therefore, there is a need for a system that can classify children's nutritional status early through intensified growth monitoring [21]. Risk factors for malnutrition consist of indirect and direct causes [22]. Indirect causal factors include a lack of quantity and quality of food consumed, which affects the weight-to-height ratio in normal children or decreases, children's linear growth decreases or stops and weight gain decreases. Meanwhile, direct factors such as poverty, low education, food availability, and job opportunities. Therefore, overcoming malnutrition requires cooperation from various related parties. The impact of malnutrition on children is classified into three classes, namely Marasmus, Kwashiorkor, and Marasmix-Kwashiorkor. Marasmus is a severe form of nutrition characterized by symptoms of looking very thin, thin, and dull hair. Kwashiorkor is a form of severe protein malnutrition, resulting in impaired growth and changes in mental status. Meanwhile, Marasmix-Kwashiorkor is a combination of these two classes [23].

The statistical model is one of the models that is trusted and very reliable as a support for decision-making [24]. The concept of probability is a form of statistical model. In this research, Naïve Bayes and SVM methods were used. However, before the classification process, a preprocessing process is carried out by standardizing the data. The results of data standardization can be seen in Table 2. After the data standardization process, the next process is to carry out classification using Naïve Bayes by determining the probability value of each variable by adding up the frequencies and combinations of values from the data set. In this method, all attributes will contribute to decision-making, with equally important attribute weights and each attribute is independent of each other, where the first step of Naive Bayes is to calculate the mean and standard deviation of each continuously valued variable in each category, such as in Table 3. Then calculate the probability of the nutritional status category for each category itself, as in Table 4.

Table 2. Data Standardization Results

Data	Gender	Age	Birth Weight	Birth Height	Weight	Height
μ	0.5186	37.7873	3.0587	49.1163	10.9765	83.4791
σ	0.50001	16.1521	0.4047	1.5415	2.6481	11.4305
…	…	…	…	…	…	…
11	-1.0373	1.7466	-0.392	-0.0754	1.3682	1.8127
12	-1.0373	1.4371	-0.392	-0.7241	1.5571	1.9439
13	0.9626	1.3132	0.3488	0.5732	0.8018	1.3053
14	-1.0373	0.9418	0.3488	-0.0754	0.2731	0.2642
15	-1.0373	0.6322	1.0901	0.5732	0.7640	1.1653
16	0.9626	0.5084	-0.145	0.5732	0.3864	0.8154
17	0.9626	0.5084	1.0901	0.5732	0.3864	1.1653
18	-1.0373	0.4465	0.5959	0.5732	0.5375	0.2905
19	-1.0373	0.3846	1.5842	1.8706	1.2172	0.9466
20	0.9626	0.1369	0.8430	1.8706	0.2354	0.3255
…	…	…	…	…	…	…
694	-1.0373	1.1894	-0.886	-0.724	0.3109	1.0078

Table 3. Mean and Standard Deviation of Variables in Every Category

Attribute	Mean		Standard Deviation	
	Malnutrition (M)	Normal (N)	Malnutrition (M)	Normal (N)
Gender	-0.0136	0.0134	0.9983	0.9986
Age	0.2146	0.1169	0.1939	0.1114
Birth Weight	0.7327	0.5666	0.2374	0.2803
Birth Height	0.5333	0.7233	0.1801	0.2984
Weight	0.6494	0.5620	0.2382	0.1817
Height	0.4578	0.4522	0.2219	0.1615

Table 4. Likelihood of Each Attribute for Each Category in Nutritional Status

Data		P(X|H)						Likelihood P(H|X)
		Gender	Age	Birth Weight	Birth Height	Weight	Height	
..	…	..	..	..	..	..	..	…
11	M	0.6075	0.7947	0.1922	0.0896	0.8163	0.5812	0,004
	N	0.7731	0.9684	0.4462	0.4587	0.7780	0.5299	0,063
12	M	0.6175	0.7521	0.5669	0.8444	0.8265	0.8444	0,155
	N	0.7831	0.9275	0.3549	0.5472	0.8414	0.5472	0,065
13	M	0.3994	0.5669	0.7364	0.8444	0.7364	0.4120	0,043
	N	0.2768	0.3549	0.6285	0.5472	0.6285	0.5190	0,011
14	M	0.6115	0.5304	0.1922	0.6497	0.8345	0.8763	0,030
	N	0.7631	0.6286	0.4462	0.7678	0.8516	0.6482	0,091
15	M	0.3994	0.8678	0.7115	0.8444	0.6497	0.6005	0,081
	N	0.2768	0.9678	0.4983	0.5472	0.7678	0.7231	0,041
16	M	0.6215	0.5304	0.7863	0.0896	0.8345	0.4120	0,008
	N	0.7231	0.6286	0.6318	0.4695	0.8414	0.5190	0,059
17	M	0.6117	0.8290	0.7863	0.0885	0.7567	0.5879	0,016
	N	0.7431	0.9894	0.6318	0.4587	0.6614	0.4809	0,068
18	M	0.6325	0.8148	0.7863	0.0896	0.6113	0.4120	0,009
	N	0.7531	0.7142	0.6318	0.4777	0.4578	0.5197	0,039
19	M	0.6325	0.7521	0.7863	0.0896	0.6497	0.5879	0,013
	N	0.7332	0.9275	0.6318	0.4587	0.7678	0.4769	0,072
20	M	0.6005	0.8628	0.7863	0.0896	0.6497	0.4120	0,010
	N	0.7231	0.8682	0.6318	0.4587	0.7678	0.5187	0,072
…	.	…	….	…	….	…	……	…
69	M	0.6355	0.8548	0.7651	0.8444	0.0203	0.4120	0,003
4	N	0.7321	0.8662	0.7231	0.5472	0.0184	0.5190	0,002

Meanwhile, the classification process with SVM begins with the process of forming a polynomial kernel to represent data in analysis when the data is not separated linearly. The results of kernel calculations for training data can be seen in Table 5. The parameters of the kernel polynomial function consist of C, and Degree (d). In this study, a polynomial kernel function was used with values d=1 and d=2 and parameters C=1, C=5, C=10, C=50, and C=100. Classification accuracy for each polynomial kernel function parameter can be seen in Figure 4. It can be seen in the graph that the best accuracy in training data classification is 87.36% for parameters d=2 and C=5. Meanwhile, the best accuracy data was 89.49% for testing data on parameters d=1 and C=5. From the accuracy results produced by kernel calculations, it has been shown that the polynomial kernel is the best kernel for the classification of malnutrition.

Table 5. Kernel Polynomial Calculation Results for Training Data

No-P	11	12	13	14	15	16	17	18	19	20
…	….	….	….	….	….	….	….	….	…..	…..
11	7.43	9.70	8.91	7.07	3.98	1.28	2.92	3.77	0.02	0.02
12	9.70	13.1	12.3	11.8	4.40	2.90	2.86	4.07	0.04	0.04
13	8.91	12.3	12.2	13.3	3.09	4.14	1.53	2.10	0.15	0.15
14	7.07	11.8	13.3	56.3	1.39	67.1	0.32	3.51	0.28	0.28
15	3.98	4.40	3.09	1.39	5.80	0.20	7.18	10.4	0.14	0.14
16	1.28	2.90	4.14	67.1	0.20	13.8	0.04	4.83	0.71	0.71
17	2.92	2.86	1.53	0.32	7.18	0.05	10.3	15.7	0.50	0.50
18	3.77	4.07	2.10	3.51	10.4	4.83	15.7	29.7	2.04	2.04
19	0.02	0.04	0.15	0.28	0.14	0.71	0.50	2.04	5.84	5.84
20	0.31	0.05	0.27	16.4	0.21	49.9	1.22	14.1	9.99	9.99
…	……	…..	……	……	……	……	…….	……	…..	……

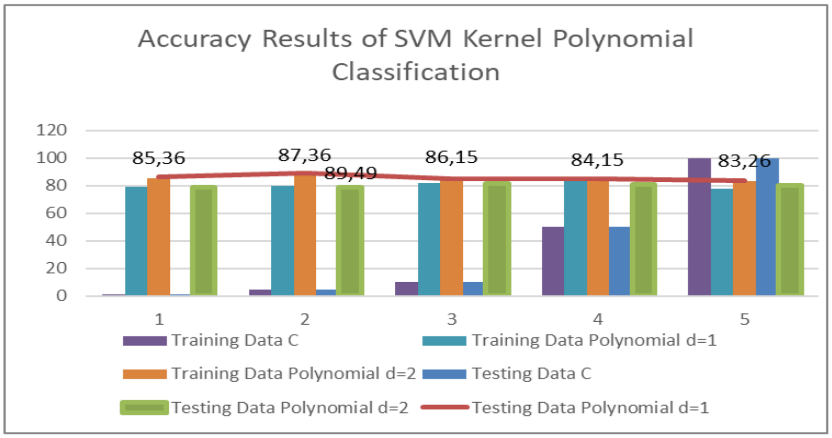

Figure 4. Graphic of Accuracy Kernel Polynomial with Training and Testing Data

Apart from determining the kernel in the SVM for classification, the next step is calculating the Hessian matrix. The Hessian matrix is the product of the polynomial kernel and the Y value. The Y value here is a value in the form of a vector containing the values 1 and -1. This calculation is carried out repeatedly until the maximum iteration state is reached or max $(|\delta \alpha i|) < \varepsilon\ (epsilon)$. The results of calculating the Hessian matrix for ten training data can be seen in Table 6. Then the final step in the classification process with SVM is calculating the margin. The results of the margin calculations on the training data can be seen in Table 7.

Table 6. Hessian Matrix Result of Training Data

No-P	11	12	13	14	15	16	17	18	19	20
….	….	….	….	….	….	….	….	…	…	…
11	4.68	4.46	4.85	-4.6	4.83	4.48	-4.6	-4.3	-5.2	-4.8
12	4.46	4.26	4.64	-4.4	4.63	4.29	-4.5	-4.1	-5.0	-4.6
13	4.85	4.64	5.10	-4.8	5.13	4.72	-4.9	-4.5	-5.5	-5.1
14	-4.6	-4.4	-4.8	6.70	-4.9	-4.5	5.70	6.43	6.34	5.87
15	4.83	4.63	5.13	-4.9	5.25	4.77	-5.0	-4.7	-5.6	-5.2
16	4.48	4.29	4.72	-4.5	4.77	4.42	-4.6	-4.3	-5.1	-4.7
17	-4.6	-4.4	-4.9	5.70	-5.0	-4.6	5.31	5.52	5.91	5.46
18	-4.3	-4.1	-4.5	6.43	-4.7	-4.3	5.52	6.35	6.05	5.65
19	-5.2	-5.0	-5.5	6.34	-5.6	-5.1	5.91	6.05	6.66	6.11
20	-4.8	-4.6	-5.1	5.87	-5.1	-4.7	5.46	5.65	6.11	5.63
…	…	….	…	…	…	…	…	…	….	….

Table 7. Calculation of Weights and Margin Values for Training Data

No-P	W.X+	W.X-
….	…………	……………
11	0.0000177215	0.0000174742
12	0.0000168541	0.0000166867
13	0.0000184346	0.0000184411
14	-0.0000174742	-0.0000258134
15	0.0000183368	0.0000186653
16	0.0000169547	0.0000170578
17	-0.0000176103	-0.0000218350
18	-0.0000162550	-0.0000247382
19	-0.0000199802	-0.0000243892
20	-0.0000182406	-0.0000225061
….	……………………..	………………………
Total	-0.0000012586	-0.0000309569
Margin	0.0000161077825509976	

Measurement of classification performance on original data and data resulting from the classification model is carried out using cross-tabulation (confusion matrix), which serves to analyze whether the classifier is good at recognizing tuples from different classes. Evaluation with the confusion matrix produces accuracy, precision, recall, and f1-score values. Accuracy in classification is the percentage of accuracy of data records that are classified correctly after testing the classification results [25]. The process of the classification analysis stage for the Naïve Bayes algorithm with Confusion matrix results can be seen in Table 8. Meanwhile, the analysis of classification with SVM uses the best parameters from SVM, namely parameters C = 5 and d = 2 for training data and d = 1 and C = 5 for testing data. Performance evaluation based on the confusion matrix table between predictions and actuals can be seen in Table 9. Based on table 9, is obtained from comparing these two methods to determine the method with the highest level of accuracy. From the results of the two algorithms, accuracy, precision, recall, and f1-score are obtained as in Table 10. Based on Table 10, the results of the comparison between the Naïve Bayes algorithm and SVM can be concluded with the accuracy graph in Figure 5. Based on Figure 5 shows that the values obtained From the test results by dividing the dataset, the best accuracy results were obtained by the SVM method with Kernel Polynomial and parameter C=5 of 89.76% for the malnutrition classification process in children.

Table 8. Confusion Matrix Table Naïve Bayes Method for Malnutrition Classification

Training Data (%)	Testing Data (%)	Prediction	Actual	
			M	N
90	10	M	320	55
		N	40	279
80	20	M	340	52
		N	49	253
70	30	M	279	35
		N	70	310
60	40	M	219	45
		N	65	365
50	50	M	214	45
		N	70	365
40	60	M	265	53
		N	66	310
30	70	M	260	56
		N	73	305
20	80	M	265	66
		N	73	290
10	90	M	255	60
		N	83	290

Table 9. Confusion Matrix Table SVM Method for Malnutrition Classification

Training Data (%)	Testing Data (%)	Prediction	Actual	
			M	N
90	10	M	338	56
		N	15	285
80	20	M	334	56
		N	25	279
70	30	M	344	53
		N	33	264
60	40	M	340	57
		N	36	261
50	50	M	350	50
		N	48	246
40	60	M	372	48
		N	58	216
30	70	M	362	58
		N	68	206
20	80	M	299	61
		N	75	259
10	90	M	289	56
		N	85	264

Table 10. Classification Model Performance Measurement Results

Performance Parameters	Training Data (%)	Naïve Bayes Method (%)	SVM Method (%)
Accuration	90	86.31	89.76
	80	85.44	88.32
	70	84.87	87.60
	60	84.14	86.59
	50	83.42	85.87
	40	82.85	84.72
	30	81.41	81.84
	20	79.97	80.40
	10	78.53	79.68
Precision	90	88.88	95.75
	80	87.40	93.03
	70	79.94	91.24
	60	77.11	90.42
	50	75.35	87.93
	40	80.06	86.51
	30	78.07	84.16
	20	78.40	79.94
	10	75.44	77.27
Recall	90	85.33	85.78
	80	86.73	85.64
	70	88.85	86.64
	60	82.95	85.64
	50	82.62	87.50
	40	83.33	88.57
	30	82.27	86.19
	20	80.06	83.05
	10	79.43	83.76
F1-score	90	87.07	90.49
	80	87.06	89.18
	70	84.16	88.88
	60	79.92	87.96
	50	78.82	87.71
	40	81.66	87.52
	30	80.12	85.17
	20	79.22	81.47
	10	77.38	80.38

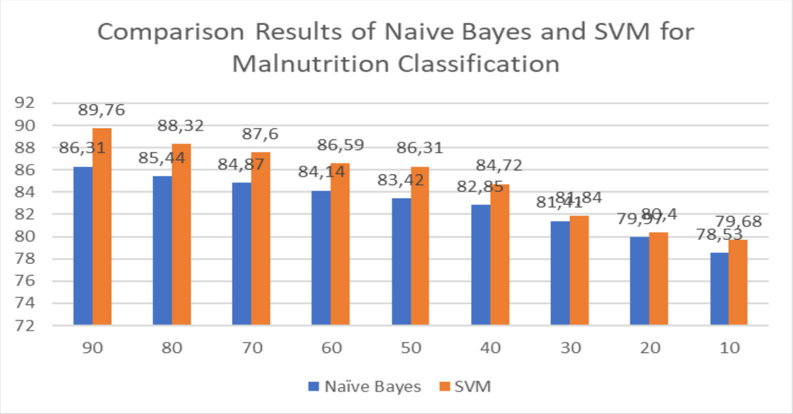

Figure 5. Graphic of Accuration to Comparison Results of Naive Bayes and SVM Classifier Methods.

4. CONCLUSIONS

From the results of research that has been carried out to classify the nutritional status of malnutrition in children by comparing the Naïve Bayes and SVM methods, it can be concluded that this classification process uses dataset division and then measures the performance of the two methods to determine the level of accuracy by evaluating the confusion matrix against several of the best experiments. The SVM method with a polynomial kernel has the highest accuracy value of 89.76%. This can be said to be more accurate when compared to the Naïve Bayes method, with an accuracy rate of 86.31%. Apart from that, for this research to be better in the future, it is necessary to develop the following research with the suggestion to combine attribute selection methods to increase the accuracy value in the classification process and to be able to compare SVM with several kernels.

ACKNOWLEDGEMENTS

The author would like to thank the Kalianget Community Health Center, Madura for its assistance in providing data in this research.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] D. Susilowati, S. Joko, R.A. Pramuja, Quality study of human development and poverty in Madura Island, J. Ekon. Pembang. 17 (2020), 175. https://doi.org/10.22219/jep.v17i2.10360.

[2] E.A. Suryana, M. Azis, The potential of economic loss due to stunting in Indonesia, J. Ekon. Kesehatan Indonesia. 8 (2023), 52. https://doi.org/10.7454/eki.v8i1.6796.

[3] J.P. Tanjung, F.C. Tampubolon, A.W. Panggabean, et al. Customer classification using Naive Bayes classifier with genetic algorithm feature selection, SinkrOn 8 (2023), 584-589. https://doi.org/10.33395/sinkron.v8i1.12182.

[4] T. Permatasari, Y. Sandy, C. Pratiwi, et al. Naive Bayes Classifier (NBC) application on the nutritional status of adolescents in Medan, in: Proceedings of the 4th Annual Conference of Engineering and Implementation on Vocational Education, ACEIVE 2022, 20 October 2022, Medan, North Sumatra, Indonesia, EAI, Medan, Indonesia, 2023. https://doi.org/10.4108/eai.20-10-2022.2328884.

[5] A. Sholikhaq, G. Firmansyah, B. Tjahjono, et al. Nutrition classification in toddlers at UPTD puskesmas tigaraksa using a comparison of support vector machine (SVM) and K-nearest neighbor (KNN) methods, JRSSEM. 2 (2023), 1968-1988. https://doi.org/10.59141/jrssem.v2i09.415.

[6] V. Viswanatha, A.C. Ramachandra, R. Mohammed, et al. Support vector machine implementation to separate linear and non-linear dataset, Saudi J. Eng. Tech. 8 (2023), 4-15. https://doi.org/10.36348/sjet.2023.v08i01.002.

[7] R.N. Devita, H.W. Herwanto, A.P. Wibawa, Perbandingan kinerja metode Naive Bayes dan K-nearest neighbor untuk klasifikasi artikel berbahasa Indonesia, J. Teknol. Inform. Ilmu Komputer. 5 (2018), 427-434. https://doi.org/10.25126/jtiik.201854773.

[8] R. Rachman, R.N. Handayani, Klasifikasi algoritma Naive Bayes dalam memprediksi tingkat kelancaran pembayaran sewa teras UMKM, J. Inform. 8 (2021), 111-122. https://doi.org/10.31294/ji.v8i2.10494.

[9] G.M. Momole, Perbandingan Naïve Bayes dan Random Forest Dalam Klasifikasi Bahasa Daerah, J. Teknik Inform. Sist. Inform. 9 (2022), 855-863. https://doi.org/10.35957/jatisi.v9i2.1857.

[10] R. Oktafiani, R. Rianto, Perbandingan algoritma support vector machine (SVM) dan decision tree untuk sistem rekomendasi tempat wisata, J. Nasional Teknol. Sist. Inform. 9 (2023), 113-121. https://doi.org/10.25077/teknosi.v9i2.2023.113-121.

[11] B.V. Jayadi, T. Handhayani, M.D. Lauro, Perbandingan knn dan svm untuk klasifikasi kualitas udara di jakarta, J. Ilmu Komputer Sist. Inform. 11 (2023), 1-7. https://doi.org/10.24912/jiksi.v11i2.26006.

[12] R. Syahputra, G.J. Yanris, D. Irmayani, SVM and Naïve Bayes algorithm comparison for user sentiment analysis on Twitter, SinkrOn. 7 (2022), 671-678. https://doi.org/10.33395/sinkron.v7i2.11430.

[13] E. Haerani, F. Syafria, F. Lestari, et al. Classification academic data using machine learning for decision making process, J. Appl. Eng. Technol. Sci. 4 (2023), 955-968. https://doi.org/10.37385/jaets.v4i2.1983.

[14] H.A. Prihanditya, Alamsyah, The implementation of Z-score normalization and boosting techniques to increase accuracy of C4.5 algorithm in diagnosing chronic kidney disease, J. Soft Comput. Explor. 1 (2020), 63-69. https://doi.org/10.52465/joscex.v1i1.8.

[15] S.J. Mukhyber, A.D. Majeed, Effect Z-score Normalization on Accuracy of classification of liver disease, Turk. J. Comput. Math. Educ. 12 (2021), 658-662.

[16] H. Henderi, 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, Int. J. Inform. Inform. Systems. 4 (2021), 13-20. https://doi.org/10.47738/ijiis.v4i1.73.

[17] A. Solikhatun, E. Sugiharti, Application of the Naïve Bayes classifier algorithm using N-gram and information gain to improve the accuracy of restaurant review sentiment analysis, J. Adv. Inform. Syst. Technol. 2 (2020), 1-12.

[18] N.H. Ovirianti, M. Zarlis, H. Mawengkang, Support vector machine using a classification algorithm, SinkrOn 7 (2022), 2103-2107. https://doi.org/10.33395/sinkron.v7i3.11597.

[19] P.N.M. Weking, I.N.W. Partha, A.I. Weking, Application of data mining with support vector machine (SVM) in selling prediction trend of spiritual goods, Int. J. Eng. Emerg. Technol. 4 (2019), 20-24.

[20] P. Shirisha, The importance of disaggregated data analysis of child undernutrition and its determinants - A district level analysis in the non-high focus state of India, Clinic. Epidemiol. Glob. Health 17 (2022), 101147. https://doi.org/10.1016/j.cegh.2022.101147.

[21] M. Monga, C. Sikorski, H. de silva, et al. Identifying underweight in infants and children using growth charts, lookup tables and a novel “MAMI” slide chart: A cross-over diagnostic and acceptability study, PLOS Glob. Public Health. 3 (2023), e0002303. https://doi.org/10.1371/journal.pgph.0002303.

[22] N. Faizi, S. Kazmi, Universal health coverage - There is more to it than meets the eye, J. Family Med. Prim. Care. 6 (2017), 169-170. https://doi.org/10.4103/jfmpc.jfmpc_13_17.

[23] S. Pingale, V.W. Patil, M. Hire, et al. Prevalence of kwashiorkor, marasmus, marasmic kwashiorkor and age wise distribution of malnourished tribal children of Town Dhadgaon, District-Nandurbar of Maharashtra State, India, Res. J. Pharm. Technol. 7 (2014), 59-63.

[24] B. Jeong, H. Cho, J. Kim, et al. Comparison between statistical models and machine learning methods on classification for highly imbalanced multiclass kidney data, Diagnostics 10 (2020), 415. https://doi.org/10.3390/diagnostics10060415.

[25] J. Kozak, B. Probierz, K. Kania, et al. Preference-driven classification measure, Entropy. 24 (2022) 531. https://doi.org/10.3390/e24040531.