JRSSEM 2023, Vol. 02 No. 9, 1968 – 1988
E-ISSN: 2807 - 6311, P-ISSN: 2807 - 6494

NUTRITION CLASSIFICATION IN TODDLERS AT UPTD PUSKESMAS TIGARAKSA USING A COMPARISON OF SUPPORT VECTOR MACHINE (SVM) AND K-NEAREST NEIGHBOR (KNN) METHODS

Amelia Sholikhaq[1]
Gerry Firmansyah[2]
Bob Tjahjono[3]
Habibullah Akbar[4]
[1,2,3,4] *Master of Computer Science Study Program, Faculty of Computer Science*
Esa Unggul University Jakarta, Indonesia
*Email: ameliasholihah8@gmail.com, *gerry@esaunggul.ac.id
budi.tjahjono@esaunggul.ac.id, habibullah.akbar@esaunggul.ac.id
*Correspondence: ameliasholihah8@gmail.com

Submitted: March 27th 2023 *Revised*: April 12th 2023 *Accepted*: April 20th 2023

Abstract: Toddlers are a group of people who are vulnerable to nutritional problems. If the incidence of malnutrition is not addressed, it will hurt children under five, malnutrition is a condition experienced by a person due to a lack of nutritional intake of the number of nutrients consumed below. Health centers are required to improve and organize health services as well as possible therefore researchers conduct research at the UPTD Tigaraksa Health Center by doing a comparison of classification results on toddler nutritional data using the Support Vector Machine and K-Nearest Neighbor methods using WEKA Tools. Based on the result of a comparison between the Support Vector Machine and K-Nearest Neighbor methods using WEKA Tools by carrying out 5 (five) stages of testing namely: Use Training Set, 4 Cross-Validation, 8 Cross-Validation, 50% Percentage Split dan 80% Percentage Split, the results show that the Support Vector Machine method Kernel Radial Basis Function (RBF) is an average accuracy value of 100% higher than the K-Nearest Neighbor Euclidean Distance algorithm with an average accuracy of 93%.

Keywords: Toddler Nutrition; Data Mining; Classification; Support Vector Machine and K-Nearest Neighbor methods; WEKA Tools.

INTRODUCTION

Poor nutrition is a condition experienced by a person due to a lack of nutritional intake or the number of nutrients consumed below standard. Nutrients needed include carbohydrates, proteins, and calories. One of the most important and common nutritional problems experienced by infants under 5 years old (toddlers) is a lack of protein energy. It is associated with the economic level of society. In addition, parents lack knowledge about the importance of nutrition for children's growth and development. The nutritional status of toddlers can be determined through laboratory examination or anthropometry. Anthropometric measurements are measurements used to determine a person's nutritional state.

Puskesmas is a technical implementation unit of the district/city health office that is responsible for organizing health development in a work area. Puskesmas are required to improve and provide health services as well as possible. UPTD Puskesmas Tigarakasa is a Puskesmas located in Tigaraksa District, Tangerang Regency, which is located at Jl. Kongsi No.12, Tigaraksa Village, Tigaraksa District, Tangerang Regency, Banten Province, Puskesmas work area. Tigaraksa Health Center covers several villages, namely Bantar Panjang, Cileles, Kadu Agung, Margasari, Sodong, Tapos, and Tigaraksa.

In this study, the goal to be achieved is to improve the accuracy of results using method comparison with the comparison of Support Vector Machine and K-Nearest Neighbors methods in classifying nutrition in toddlers (Arsi & Waluyo, 2021). Proving that using a comparison of the Support Vector Machine and K-Nearest Neighbors methods can increase the percentage of accuracy so that it is more optimal in the early detection of nutritional status in toddlers quickly, precisely, and accurately (Sugara & Subekti, 2019).

MATERIALS AND METHODS

Nutritional Status of Toddlers

Nutritional Status is a measure of success in fulfilling nutrition for children indicated by the child's weight and height. Nutritional status can be defined as the health status produced by the balance between nutritional needs and inputs. A child can be said to be a toddler when he is between 0 to 5 years old. Because the term BALITA is an abbreviation of Infants Under Five Years". The age of toddlers is the age of growth, which is when a toddler must be active and energetic in acting. Toddlers are active and energetic in doing actions because curiosity about something they meet arises in their minds. The nutritional status of toddlers can be classified as follows: Poor Nutrition, Lack of Nutrition, Good Nutrition, and More Nutrition (Hariri &; Pamungkas, 2016).

Data Mining

Data mining is the process by which statistical, mathematical, artificial intelligence and machine learning techniques are used to extract and identify useful information and related knowledge in large databases. Data mining is not an entirely new field. One of the difficulties in

identifying data mining is the fact that data mining has long roots in fields such as artificial intelligence, machine learning, statistics, databases, and information retrieval (Nikmatun & Waspada, 2019).

Tools WEKA

WEKA has useful tools for data preprocessing, classification, regression, clustering, association rules, and visualization. Can be used to preprocess data, enter into a learning schema, and analyze the classification generated by its performance, done without writing program code. Examples of using WEKA by applying a learning method to a dataset and analyzing the results to obtain information about the data, or applying the method and comparing its performance to be selected.

Classification

Classification is a data mining method that can be used for the process of searching for a set of data models (functions) that can explain and distinguish data classes or data concepts, which has the aim that the model can be used to predict class objects that have labels whose value is unknown or used to predict the tendency of data that often appears in the future.

The model in classification is defined in detail as a working model whose process of conducting training / requires a learning model of the target function, which is usually interpreted as a place to receive input (training data), then be able to think about the input, and provide answers as an output of the results of his thoughts. The model is used to predict classes from test data. The process of work in classification can be seen in Figure 1.

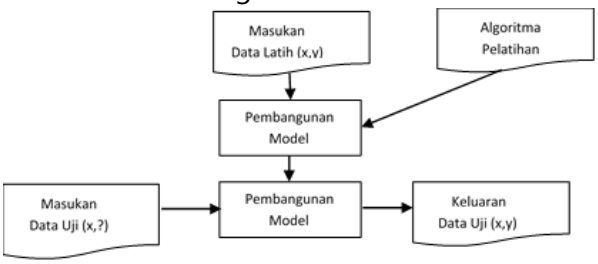


Figure 1. The process of classification work

In Figure 1. The model that is already built at the time of training input can then be used to predict class labels from new data that does not yet know the class labels. In building a model during the training process, an algorithm is needed to build it, which is called a training algorithm (learning algorithm).

Comparisons

In the Big Indonesian Dictionary (KBBI) Comparison is the difference (difference) between similarities and similarities, then the comparison is an attempt to observe the differences or similarities possessed by two or more objects that have a certain similarity (Nasution & Hayaty, 2019; Pratama & Salamah, 2022).

Metode Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the existing classification methods (Fadilah et al., 2020). SVM was developed by Boser, Guyon, and Vapnik, and was first presented in 1992 at the Annual Workshop on Computational Learning Theory (Permana & Sahara, 2019). The basic concept of SVM is a harmonized combination of computational theories of money that had existed decades earlier, such as the hyperplane margin (Duda and Hart in 1973, cover in 1965, Vapnik in 1964,

etc.), the kernel introduced by Aronszajn in 1950, and so with other supporting concepts (Wati et al., 2020). However, until 1992, there had never been an attempt to assemble these components (Mase et al., 2018).

Metode K-Nearest Neighbors (KNN)

The K-Nearest Neighbors (KNN) method is a method for classifying based on the proximity of the location (distance) of one data with other data (Pamungkas & Kharisudin, 2021). K-Nearest Neighbors includes supervised learning algorithms (Saidah et al., 2019). The working principle of K-Nearest Neighbors (KNN) itself is to find the closest distance between the evaluated data and (K) its closest neighbor in the training data. Before finding the closest distance between the evaluated data, the K-Nearest Neighbors algorithm must be preprocessed or normalized first (Budianto et al., 2019).

Testing and Evaluation

The test of the analysis aims to determine the level of accuracy of the comparison between the support vector machine method and the k-nearest neighbor in the classifier of determining nutritional status in toddlers at UPTD Puskesmas Tigaraksa (Baita et al., 2021). Accuracy testing with classification is carried out in several experiments to get more accurate results (Hakim et al., 2020).

In processing the classification, researchers use the Weka Tool to determine a better level of accuracy between the Support Vector Machine method and K-Nearest Neighbor with several stages in processing datasets in each classification method using the Support Vector Machine and K-Nearest Neighbor (Ichwan & Dewi, 2018), namely:

a. Use Training Set, at this stage WEKA Tools uses the previously inputted training data as testing data. In other words, the training and testing process uses the same data.

b. Cross-Validation, the training data will randomly be divided into k parts at this stage. Furthermore, the k-1 part is used as training data, and one part is used as test data. The process is repeated so that each part has the opportunity to become test data. In WEKA Tools, the default value is 10.

c. Percentage Split, the data inputted in the previous step will be divided into training data and test data based on a certain percentage. The default value in WEKA Tools is 66 %, where the input data is divided into 66% training data and 34% test data.

Data Collection

In collecting toddler data used in this study, data obtained from UPTD Puskesmas Tigaraksa data on toddlers born in 2018 to 2022 aged 0 months to 60 months, focusing on variables or features that are numerical in nature with four parameters, namely Age in months, Gender, Weight in kg, Height in cm.

Table 1. Data on Toddlers UPTD Puskesmas Tigaraksa toddler age aged 5 to 59 months

ID	NAMA	JK	USIA-BLN	ATT-BB/U	ATT-TB/U	ATT-TB/BB
1	AWAL	Laki-laki	5	-0,07	-1,30	-0,04
2	ROBI	Laki-laki	8	-1,13	-1,85	-0,24
3	ALIY	Laki-laki	7	-2,24	-1,91	-2,94
4	DEFR	Laki-laki	10	-2,20	-1,52	-2,08
5	ERIN	Laki-laki	5	1,50	-2,00	1,01
6	GEND	Laki-laki	11	2,60	-1,30	1,07
7	ARAM	Laki-laki	6	2,60	-1,06	2,12
8	LUTH	Laki-laki	8	2,66	-0,54	2,18
9	MUHA	Laki-laki	10	2,79	-1,17	3,97
10	RIFAT	Laki-laki	9	3,20	-1,63	3,17
...	.....	....	....	....	....	....
...	.....	....	....	....	....	....
...	.....	....	....	....	....	....
91	ALFA	Perempuan	56	-1,69	-1,61	-1,1
92	RINA	Perempuan	55	-1,85	-1,56	-1,3
93	ADIB	Perempuan	52	-3,77	-2,32	-2,4
94	AFIF	Perempuan	51	-2,46	-2,13	-2,3
95	ARSY	Perempuan	50	1,24	0,86	1,5
96	ASMA	Perempuan	53	1,17	2,56	1,5
97	AURE	Perempuan	58	2,45	-0,65	2,2
98	AYUN	Perempuan	50	2,55	-0,87	2,2
99	CHAT	Perempuan	58	3,34	-1,22	3,4
100	FITR	Perempuan	52	3,22	-1,11	3,2

RESULTS AND DISCUSSION

Nutritional Status Data for Toddlers Using Z Score

In inputting anthropometric data for toddlers aged between 1 to 59 months carried out by UPDT Puskesmas Tigaraksa which includes seven villages, namely tigaraksa, kaduagung, marga sari, song, tapos, bantarpanjang, and cells, which have been calculated using the z-score formula to determine the nutritional status of each toddler who has been calculated according to age, gender, weight (kg) and length/height (cm).

Table 2. UPTD Toddler Data Puskesmas Tigaraksa

ID	NAMA	ATT-JK	USIA-BLN	ATT-BB/U	BB/U	ATT-TB/U	TB/U	ATT-TB/BB	STATUS
1	AWAL	Laki-laki	5	-0,07	BB Normal	-1,30	TB Normal	-0,04	Gizi Baik
2	ROBI	Laki-laki	8	-1,13	BB Normal	-1,85	TB Normal	-0,24	Gizi Baik
3	ALIY	Laki-laki	7	-2,24	BB Kurang	-1,91	TB Normal	-2,94	Gizi Kurang
4	DEFR	Laki-laki	10	-2,20	BB Kurang	-1,52	TB Normal	-2,08	Gizi Kurang
5	ERIN	Laki-laki	5	1,50	BB Resiko Lebih	-2,00	TB Normal	1,01	Resiko Gizi Lebih
6	GEND	Laki-laki	11	2,60	BB Resiko Lebih	-1,30	TB Normal	1,07	Resiko Gizi Lebih
7	ARAM	Laki-laki	6	2,60	BB Lebih	-1,06	TB Normal	2,12	Gizi Lebih
8	LUTH	Laki-laki	8	2,66	BB Lebih	-0,54	TB Normal	2,18	Gizi Lebih
9	MUHA	Laki-laki	10	2,79	BB Obesitas	-1,17	TB Normal	3,97	Obesitas
10	RIFAT	Laki-laki	9	3,20	BB Obesitas	-1,63	TB Normal	3,17	Obesitas
11	ELSA	Perempuan	5	-0,19	BB Normal	-1,06	TB Normal	0,35	Gizi Baik
12	LALA	Perempuan	8	0,08	BB Normal	-1,56	TB Normal	0,71	Gizi Baik
13	FELI	Perempuan	7	-2,76	BB Kurang	-1,77	TB Normal	-2,20	Gizi Kurang
14	DITA	Perempuan	10	-2,85	BB Kurang	-1,89	TB Normal	-2,11	Gizi Kurang
15	AKIL	Perempuan	5	2,20	BB Resiko Lebih	-1,52	TB Normal	1,08	Resiko Gizi Lebih
16	HAND	Perempuan	11	2,48	BB Resiko Lebih	-1,20	TB Normal	1,10	Resiko Gizi Lebih
17	ANIA	Perempuan	6	2,60	BB Lebih	-0,39	TB Normal	2,18	Gizi Lebih
18	ROTU	Perempuan	8	2,66	BB Lebih	-0,84	TB Normal	2,19	Gizi Lebih
19	ZIHA	Perempuan	11	3,20	BB Obesitas	-1,88	TB Normal	3,67	Obesitas
20	RANI	Perempuan	9	3,26	BB Obesitas	-1,85	TB Normal	3,19	Obesitas
...	....	....	....	....	.....	....	.....	....	.....
...	....	....	....	....	.....	....	.....	....	.....
...	....	....	....	....	.....	....	.....	....	.....
81	BILA	Laki-laki	58	-1,04	BB Normal	2,57	TB Tinggi	0,1	Gizi Baik
82	ALFI	Laki-laki	50	-1,33	BB Normal	2,20	TB Tinggi	-0,5	Gizi Baik
83	CHAR	Laki-laki	53	-2,68	BB Kurang	-2,32	TB Pendek	-2,5	Gizi Kurang
84	ALTA	Laki-laki	50	-2,55	BB Kurang	-2,26	TB Pendek	-2,4	Gizi Kurang
85	AZZA	Laki-laki	49	1,18	BB Resiko Lebih	2,22	TB Tinggi	1,5	Resiko Gizi Lebih
86	EDWA	Laki-laki	56	1,21	BB Resiko Lebih	-0,48	TB Normal	1,5	Resiko Gizi Lebih
87	EVAN	Laki-laki	58	2,22	BB Lebih	-1,27	TB Normal	2,9	Gizi Lebih
88	FARZA	Laki-laki	59	2,80	BB Lebih	-1,35	TB Normal	2,5	Gizi Lebih
89	GIBR	Laki-laki	49	3,45	BB Obesitas	-1,09	TB Normal	3,3	Obesitas
90	LINT	Laki-laki	55	3,33	BB Obesitas	-0,06	TB Normal	4,2	Obesitas
91	ALFA	Perempuan	56	-1,69	BB Normal	-1,61	TB Normal	-1,1	Gizi Baik
92	RINA	Perempuan	55	-1,85	BB Normal	-1,56	TB Normal	-1,3	Gizi Baik
93	ADIB	Perempuan	52	-3,77	BB Kurang	-2,32	TB Pendek	-2,4	Gizi Kurang
94	AFIF	Perempuan	51	-2,46	BB Kurang	-2,13	TB Pendek	-2,3	Gizi Kurang
95	ARSY	Perempuan	50	1,24	BB Resiko Lebih	0,86	TB Normal	1,5	Resiko Gizi Lebih
96	ASMA	Perempuan	53	1,17	BB Resiko Lebih	2,56	TB Tinggi	1,5	Resiko Gizi Lebih
97	AURE	Perempuan	58	2,45	BB Lebih	-0,65	TB Normal	2,2	Gizi Lebih
98	AYUN	Perempuan	50	2,55	BB Lebih	-0,87	TB Normal	2,2	Gizi Lebih
99	CHAT	Perempuan	58	3,34	BB Obesitas	-1,22	TB Normal	3,4	Obesitas
100	FITR	Perempuan	52	3,22	BB Obesitas	-1,11	TB Normal	3,2	Obesitas

In the table above the toddler data of UPTD Puskesmas Tigaraksa, researchers took samples of 100 toddler data for the August 2022 period aged 1 to 59 months. Among them are males and females, whose grouping can be seen in the table below.

Table 3. Toddler data grouped by gender

No	Gender	Sum	Status
1	Man	10	Undernutrition
		10	Good Nutrition
		10	More Nutritional Risks
		10	More Nutrition
		10	Obesity
2	Woman	10	Undernutrition
		10	Good Nutrition

		10	More Nutritional Risks
		10	More Nutrition
		10	Obesity

Table 4. Data on the nutritional status of toddlers before normalization are grouped according to age

ID	ATT-JK	USIA-BLN	ATT-BB/U	ATT-TB/U	STATUS
1	Laki-laki	5	-0,07	-1,30	Gizi Baik
2	Laki-laki	8	-1,13	-1,85	Gizi Baik
3	Laki-laki	7	-2,24	-1,91	Gizi Kurang
4	Laki-laki	10	-2,20	-1,52	Gizi Kurang
5	Laki-laki	5	1,50	-2,00	Resiko Gizi Lebih
6	Laki-laki	11	2,60	-1,30	Resiko Gizi Lebih
7	Laki-laki	6	2,60	-1,06	Gizi Lebih
8	Laki-laki	8	2,66	-0,54	Gizi Lebih
9	Laki-laki	10	2,79	-1,17	Obesitas
10	Laki-laki	9	3,20	-1,63	Obesitas
...	.....	.....	.....	.....	.....
...	.....	.....	.....	.....	.....
...	.....	.....	.....	.....	.....
81	Laki-laki	58	-1,04	2,57	Gizi Baik
82	Laki-laki	50	-1,33	2,20	Gizi Baik
83	Laki-laki	53	-2,68	-2,32	Gizi Kurang
84	Laki-laki	50	-2,55	-2,26	Gizi Kurang
85	Laki-laki	49	1,18	2,22	Resiko Gizi Lebih
86	Laki-laki	56	1,21	-0,48	Resiko Gizi Lebih
87	Laki-laki	58	2,22	-1,27	Gizi Lebih
88	Laki-laki	59	2,80	-1,35	Gizi Lebih
89	Laki-laki	49	3,45	-1,09	Obesitas
90	Laki-laki	55	3,33	-0,06	Obesitas
91	Perempuan	56	-1,69	-1,61	Gizi Baik
92	Perempuan	55	-1,85	-1,56	Gizi Baik
93	Perempuan	52	-3,77	-2,32	Gizi Kurang
94	Perempuan	51	-2,46	-2,13	Gizi Kurang
95	Perempuan	50	1,24	0,86	Resiko Gizi Lebih
96	Perempuan	53	1,17	2,56	Resiko Gizi Lebih
97	Perempuan	58	2,45	-0,65	Gizi Lebih
98	Perempuan	50	2,55	-0,87	Gizi Lebih
99	Perempuan	58	3,34	-1,22	Obesitas
100	Perempuan	52	3,22	-1,11	Obesitas

In the table above, toddler nutrition data has been normalized with the following information:

1. Gender (male = 1, female = 2).
2. Age is calculated in months according to the main data of the toddler table.
3. The Weight Score uses the z-score value of body weight against age (BB/U).

4. The Height Value still uses the z-score value of height against age (TB/U).

Nutritional status results from the calculation of z-score body weight to height (BB / TB).

Classification of Toddler Nutrition Data Using Support Vector Machine (SVM) Method and K-Nearest Neighbor Using Weka Tool

a. Using the Support Vector Machine (SVM) method.

Classification of the nutritional status of toddlers using the support vector machine (SVM) method with Radial Basis Function (RBF) kernels is carried out with 5 tests, namely:

- Use Training Set (data testing with the same training data)
- 4 &; 8 Cross-Validation (dividing data into k-subsets. For example Folds are used 10, 9 will be used as training data, and 1 as testing data until all data)
- 50% Percentage Split. (Splits the data according to the parameters that will be the data training).
- 80% Percentage Split. (Splits the data according to the parameters that will be the data training).

The following is the result of the support vector machine (SVM) classification using the WEKA tool:

1. Results of Support Vector Machine Classification Using WEKA Tool (Use Data Training)

```
Time taken to test model on training data: 0.02 seconds

=== Summary ===

Correctly Classified Instances         100               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 20  0  0  0  0 |  a = Gizi Baik
  0 20  0  0  0 |  b = Gizi Kurang
  0  0 20  0  0 |  c = Gizi Lebih
  0  0  0 20  0 |  d = Obesitas
  0  0  0  0 20 |  e = Resiko Gizi Lebih
```

Gambar 2. Klasifikasi Support Vector Machine (SVM) use data training (Naufal et al., 2020)

The picture above is the result of the classification of support vector machines in the WEKA tool using a use training set which shows the results of 100 correct predictions with an accuracy of 100% and 0 incorrect predictions with a percentage of 0% with a classification time of 0.02 seconds (Muhammad Yusuf Ramadan, 2019).

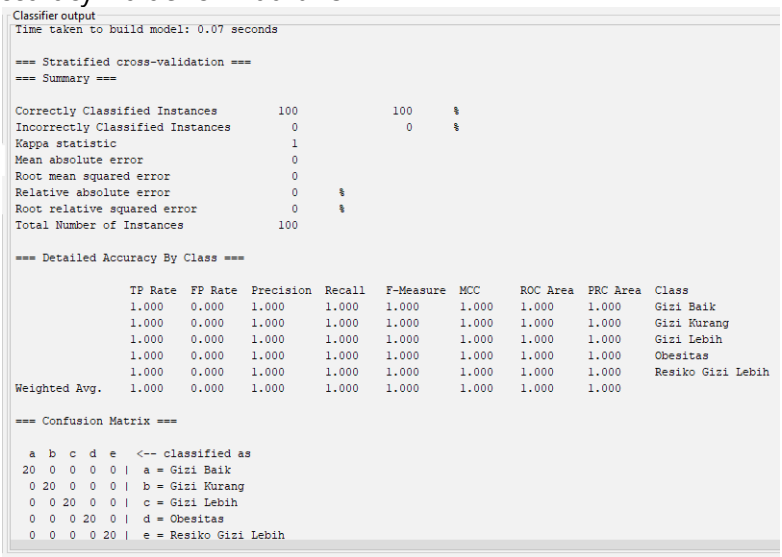
Table 5. Confusion Matrix on WEKA Use Data Training Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity

Good Nutrition	20	0	0	0	0
Undernutrition	0	20	0	0	0
More Nutritional Risks	0	0	20	0	0
More Nutrition	0	0	0	20	0
Obesity	0	0	0	0	20

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 100 data on the nutritional status of toddlers. Therefore, an accuracy value of 100% is obtained.

2. Results of Support Vector Machine Classification Using WEKA Tool (4 Cross-Validation)



```
Classifier output
Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         100               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 20  0  0  0  0 |  a = Gizi Baik
  0 20  0  0  0 |  b = Gizi Kurang
  0  0 20  0  0 |  c = Gizi Lebih
  0  0  0 20  0 |  d = Obesitas
  0  0  0  0 20 |  e = Resiko Gizi Lebih
```

Figure 3. Support Vector Machine Classification (4 Cross-Validation)

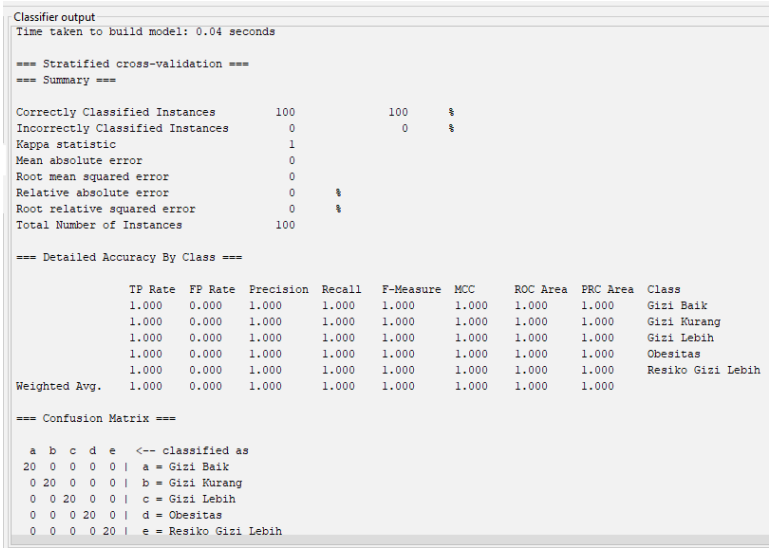
The picture above is the result of the classification of the support vector machine in the WEKA tool using 4 Cross-Validation which shows the results of 100 correct predictions with an accuracy of 100% and 0 incorrect predictions with a percentage of 0% with a classification time of 0.07 seconds.

Table 6. Confusion Matrix on 4 Cross-Validation Tools WEKA

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	20	0	0	0	0
Undernutrition	0	20	0	0	0
More Nutritional Risks	0	0	20	0	0
More Nutrition	0	0	0	20	0
Obesity	0	0	0	0	20

Based on the table above, it can be seen that the nutritional status of toddlers can be calcified correctly, namely 100 data from 100 data on the nutritional status of toddlers. Therefore, an accuracy value of 100% is obtained.

3. Results of Support Vector Machine Classification Using WEKA Tool (8 Cross-Validation)



```
Classifier output
Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         100               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 20  0  0  0  0 |  a = Gizi Baik
  0 20  0  0  0 |  b = Gizi Kurang
  0  0 20  0  0 |  c = Gizi Lebih
  0  0  0 20  0 |  d = Obesitas
  0  0  0  0 20 |  e = Resiko Gizi Lebih
```

Figure 4. Support Vector Machine Classification (8 Cross-Validation)

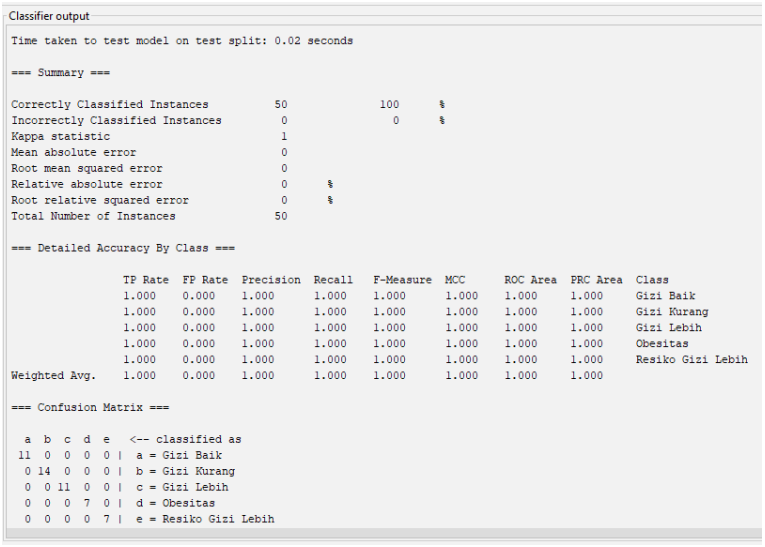
The picture above is the result of the support vector machine classification in the WEKA tool using 8 Cross-Validation which shows the results of 100 correct predictions with an accuracy of 100% and 0 incorrect predictions with a percentage of 0% with a classification time of 0.04 seconds.

Table 7. Confusion Matrix on 8 WEKA Cross-Validation Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	20	0	0	0	0
Undernutrition	0	20	0	0	0
More Nutritional Risks	0	0	20	0	0
More Nutrition	0	0	0	20	0
Obesity	0	0	0	0	20

Based on the table above, it can be seen that the nutritional status of toddlers can be calcified correctly, namely 100 data from 100 data on the nutritional status of toddlers. While none of the data is classified less precisely or differently from the original data. Therefore, an accuracy value of 100% is obtained.

4. Support Vector Machine classification results using WEKA Tool (50% Percentage Split).



```
Classifier output

Time taken to test model on test split: 0.02 seconds

=== Summary ===

Correctly Classified Instances          50               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances               50

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.  1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 11  0  0  0  0 |  a = Gizi Baik
  0 14  0  0  0 |  b = Gizi Kurang
  0  0 11  0  0 |  c = Gizi Lebih
  0  0  0  7  0 |  d = Obesitas
  0  0  0  0  7 |  e = Resiko Gizi Lebih
```

Figure 5. Support Vector Machine Classification (50% Percentage Split)

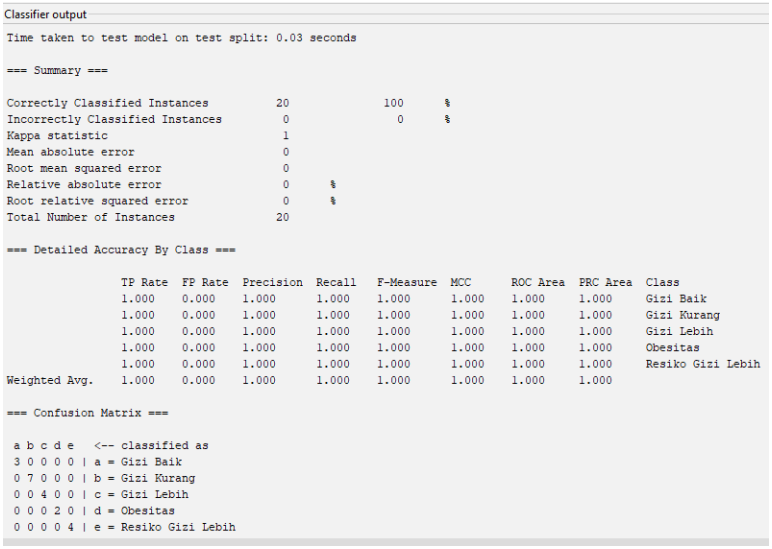
The picture above is the result of the support vector machine classification in the WEKA tool using a 50% Percentage Split which shows the results of 50 correct predictions with 100% accuracy and 0 incorrect predictions with a percentage of 0% with a classification time of 0.02 seconds.

Table 8. Confusion Matrix at 50% Percentage Split with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	11	0	0	0	0
Undernutrition	0	14	0	0	0
More Nutritional Risks	0	0	11	0	0
More Nutrition	0	0	0	7	0
Obesity	0	0	0	0	7

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 50 data from 50 nutritional status data of toddlers. While none of the other data is classified less precisely or differently from the original data. Therefore, an accuracy value of 100% is obtained.

5. Support Vector Machine classification results using WEKA Tool (80% Percentage Split).



```
Classifier output

Time taken to test model on test split: 0.03 seconds

=== Summary ===

Correctly Classified Instances          20               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0
Root mean squared error                  0
Relative absolute error                  0      %
Root relative squared error              0      %
Total Number of Instances               20

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

 a b c d e   <-- classified as
 3 0 0 0 0 | a = Gizi Baik
 0 7 0 0 0 | b = Gizi Kurang
 0 0 4 0 0 | c = Gizi Lebih
 0 0 0 2 0 | d = Obesitas
 0 0 0 0 4 | e = Resiko Gizi Lebih
```

Figure 6. Support Vector Machine Classification (80% Percentage Split)

The picture above is the result of the support vector machine classification in the WEKA tool using an 80% Percentage Split which shows the results of 20 correct predictions with 100% accuracy and no wrong predictions with a percentage of 0% with a classification time of 0.03 seconds.

Table 9. Confusion Matrix at 80% Percentage Split with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	3	0	0	0	0
Undernutrition	0	7	0	0	0
More Nutritional Risks	0	0	4	0	0
More Nutrition	0	0	0	2	0
Obesity	0	0	0	0	4

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 20 data from 20 nutritional status data of toddlers. While no other data is classified less precisely or different from the original data. Therefore, an accuracy value of 100% is obtained.

b. Comparison Results of Support Vector Machine Accuracy Evaluation from Nutritional Status Data in Toddlers

After analyzing the Support Vector Machine classification in the WEKA tool using the Use Training Set, 4 Fold Cross Validation, 8 Fold Cross Validation, 50% Percentage Split, and 80% Percentage Split, the accuracy obtained in each test has the same value, namely with an accuracy percentage of 100% for Correctly Classified Instances and 0% for Incorrectly Classified Instances. The comparison can be seen in Table 10.

Table 10. Support Vector Machine Accuracy Evaluation Comparison

Evaluation Model	Accuracy	Number of Toddlers	Percentage
Use Training Set	Correctly Classified Instances	100	100%
	Incorreclty Classified Instances	0	0%
4 Fold Cross-Validation	Correctly Classified Instances	100	100%
	Incorreclty Classified Instances	0	0%
8 Fold Cross-Validation	Correctly Classified Instances	100	100%
	Incorreclty Classified Instances	0	0%
50% Percentage Split	Correctly Classified Instances	50	100%
	Incorreclty Classified Instances	0	0%
70% Percentage Split	Correctly Classified Instances	20	100%
	Incorreclty Classified Instances	0	0%

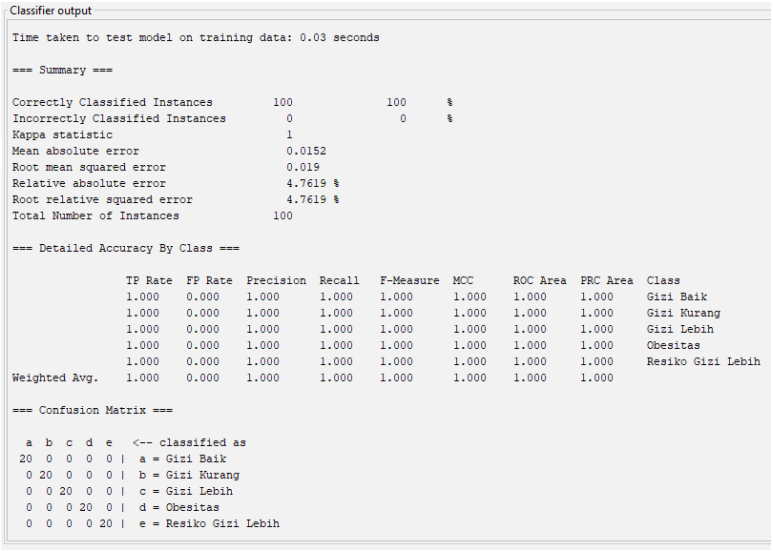
c. Menggunakan Metode K-Nearest Neighbour (KNN).

Classification of nutritional status of toddlers using K-Nearest Neighbor with Euclidean Distance algorithm carried out 5 tests (Nikmatun & Waspada, 2019), namely:

- Use Training Set (data testing with the same training data)
- 4 &; 8 Cross-Validation (dividing data into k-subsets. For example Folds used 10.9 will be used as training data and 1 as testing data up to all data)
- 50% &; 80% Percentage Split. (Splits the data according to the parameters that will be the data training).

The following results from the K-Nearest Neighbor classification using the WEKA (Use Data Training) Tool.

1. Hasil Klasifikasi K-Nearest Neighbour Menggunakan Tool WEKA (Use Data Training).



```
Classifier output

Time taken to test model on training data: 0.03 seconds

=== Summary ===

Correctly Classified Instances         100               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
Mean absolute error                      0.0152
Root mean squared error                  0.019
Relative absolute error                  4.7619 %
Root relative squared error              4.7619 %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Baik
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Resiko Gizi Lebih
Weighted Avg.    1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 20  0  0  0  0 |  a = Gizi Baik
  0 20  0  0  0 |  b = Gizi Kurang
  0  0 20  0  0 |  c = Gizi Lebih
  0  0  0 20  0 |  d = Obesitas
  0  0  0  0 20 |  e = Resiko Gizi Lebih
```

Gambar 7. Klasifikasi K-Nearest Neighbour (Use Training Set)

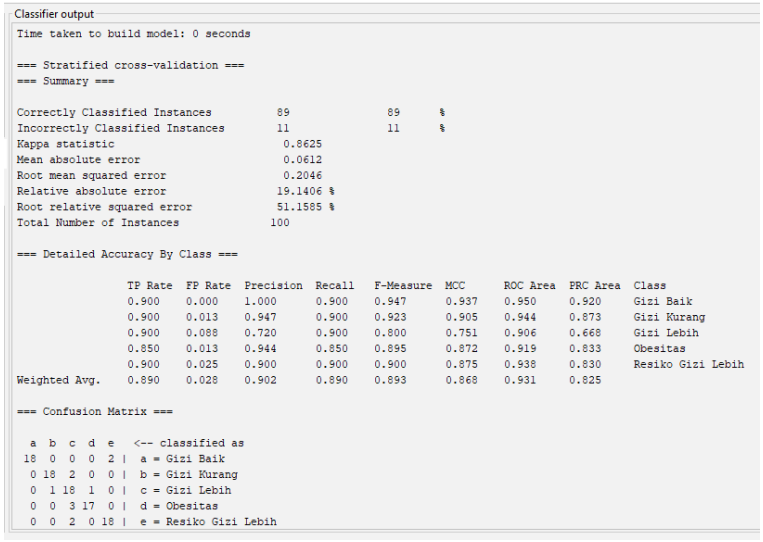
The picture above is the result of the K-Nearest Neighbor classification in the WEKA tool using a use training set which results in 100 correct predictions with 100% accuracy and no wrong predictions with a percentage of 0% with a classification time of 0.03 seconds (Amalia et al., 2021).

Table 11. Confusion Matrix on Using Data Training with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	20	0	0	0	0
Undernutrition	0	20	0	0	0
More Nutritional Risks	0	0	20	0	0
More Nutrition	0	0	0	20	0
Obesity	0	0	0	0	20

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 100 data from 100 nutritional status data of toddlers. While no other data is classified less precisely or different from the original data. Therefore, an accuracy value of 100% is obtained.

2. K-Nearest Neighbor Classification Results Using WEKA Tool (4 Cross-Validation)



```
Classifier output
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          89               89      %
Incorrectly Classified Instances        11               11      %
Kappa statistic                          0.8625
Mean absolute error                      0.0612
Root mean squared error                  0.2046
Relative absolute error                 19.1406 %
Root relative squared error             51.1585 %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.900    0.000    1.000      0.900    0.947      0.937    0.950     0.920     Gizi Baik
                 0.900    0.013    0.947      0.900    0.923      0.905    0.944     0.873     Gizi Kurang
                 0.900    0.088    0.720      0.900    0.800      0.751    0.906     0.668     Gizi Lebih
                 0.850    0.013    0.944      0.850    0.895      0.872    0.919     0.833     Obesitas
                 0.900    0.025    0.900      0.900    0.900      0.875    0.938     0.830     Resiko Gizi Lebih
Weighted Avg.    0.890    0.028    0.902      0.890    0.893      0.868    0.931     0.825

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 18  0  0  0  2 |  a = Gizi Baik
  0 18  2  0  0 |  b = Gizi Kurang
  0  1 18  1  0 |  c = Gizi Lebih
  0  0  3 17  0 |  d = Obesitas
  0  0  2  0 18 |  e = Resiko Gizi Lebih
```

Gambar 8. Klasifikasi K-Nearest Neighbour (4 Cross-Validation)

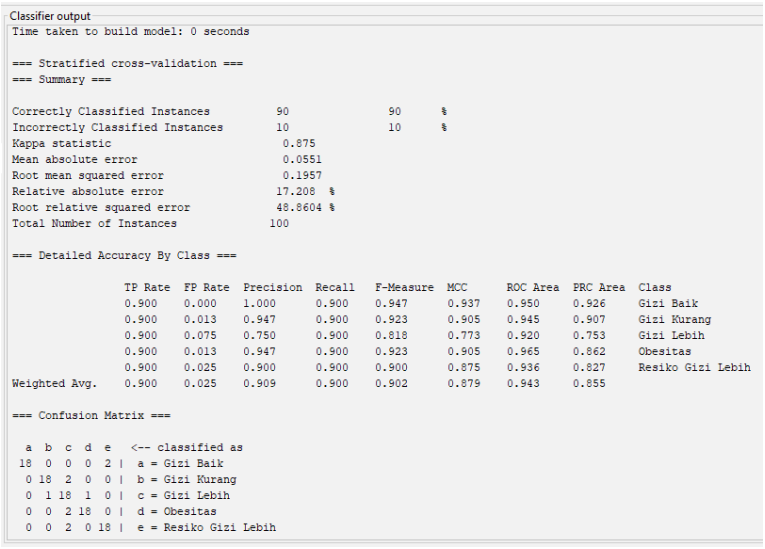
The picture above is the result of the K-Nearest Neighbor classification in the WEKA tool using 4 Cross-Validation which shows the results of 89 correct predictions with an accuracy of 89% and 11 incorrect predictions with a percentage of 11% with a classification time of 0 seconds.

Table 12. Confusion Matrix in 4 Cross-Validation with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	18	0	0	0	0
Undernutrition	0	18	2	0	0
More Nutritional Risks	0	1	18	1	0
More Nutrition	0	0	3	17	0
Obesity	0	0	2	0	18

Based on the table above, it can be seen that the nutritional status of toddlers can be calcified correctly, namely 89 data from 100 nutritional status data of toddlers. While the other 11 data are classified incorrectly or differently from the original data. Therefore, an accuracy value of 89% was obtained.

3. The results of the K-nearest neighbor classification using the WEKA (8 Cross-Validation) tool.



```
Classifier output
Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          90               90      %
Incorrectly Classified Instances        10               10      %
Kappa statistic                          0.875
Mean absolute error                      0.0551
Root mean squared error                  0.1957
Relative absolute error                 17.208  %
Root relative squared error             48.8604 %
Total Number of Instances              100

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.900    0.000    1.000      0.900    0.947      0.937    0.950     0.926     Gizi Baik
                 0.900    0.013    0.947      0.900    0.923      0.905    0.945     0.907     Gizi Kurang
                 0.900    0.075    0.750      0.900    0.818      0.773    0.920     0.753     Gizi Lebih
                 0.900    0.013    0.947      0.900    0.923      0.905    0.965     0.862     Obesitas
                 0.900    0.025    0.900      0.900    0.900      0.875    0.936     0.827     Resiko Gizi Lebih
Weighted Avg.    0.900    0.025    0.909      0.900    0.902      0.879    0.943     0.855

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 18  0  0  0  2 |  a = Gizi Baik
  0 18  2  0  0 |  b = Gizi Kurang
  0  1 18  1  0 |  c = Gizi Lebih
  0  0  2 18  0 |  d = Obesitas
  0  0  2  0 18 |  e = Resiko Gizi Lebih
```

Gambar 9. Klasifikasi K-Nearest Neighbour (8 Cross-Validation)

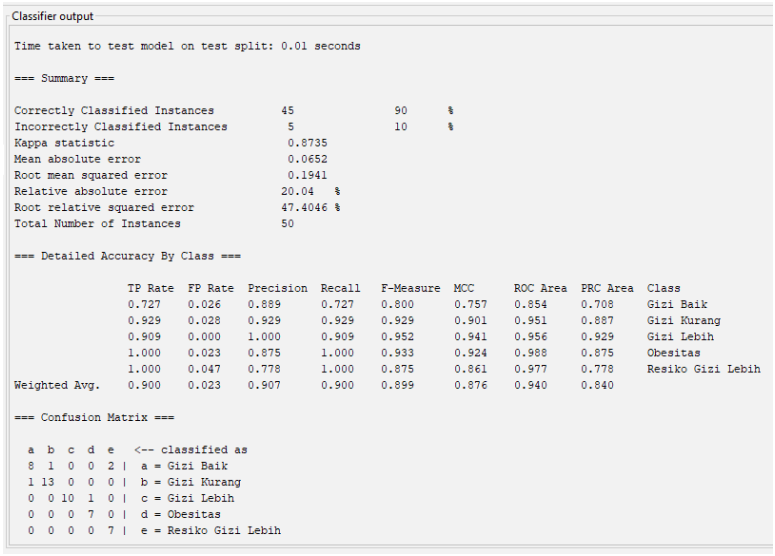
The picture above is the result of the K-Nearest Neighbor classification in the WEKA tool using 8 cross-validations which shows the results of 90 correct predictions with an accuracy of 90% and 10 incorrect predictions with a percentage of 10% with a classification time of 0 seconds.

Table 13. Confusion Matrix in 8 Cross-Validation with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	18	0	0	0	2
Undernutrition	0	18	0	0	0
More Nutritional Risks	0	1	18	1	0
More Nutrition	0	0	2	18	0
Obesity	0	0	2	0	18

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 90 data from 100 nutritional status data of toddlers. While the other 10 data are classified incorrectly or differently from the original data. Therefore, an accuracy value of 90% is obtained.

4. K-Nearest Neighbor Classification Results Using WEKA Tool (50% Percentage Split)



```
Classifier output

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances          45               90      %
Incorrectly Classified Instances         5               10      %
Kappa statistic                          0.8735
Mean absolute error                      0.0652
Root mean squared error                  0.1941
Relative absolute error                 20.04   %
Root relative squared error             47.4046 %
Total Number of Instances               50

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.727    0.026    0.889      0.727    0.800      0.757    0.854     0.708     Gizi Baik
                 0.929    0.028    0.929      0.929    0.929      0.901    0.951     0.887     Gizi Kurang
                 0.909    0.000    1.000      0.909    0.952      0.941    0.956     0.929     Gizi Lebih
                 1.000    0.023    0.875      1.000    0.933      0.924    0.988     0.875     Obesitas
                 1.000    0.047    0.778      1.000    0.875      0.861    0.977     0.778     Resiko Gizi Lebih
Weighted Avg.    0.900    0.023    0.907      0.900    0.899      0.876    0.940     0.840

=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
  8  1  0  0  2 |  a = Gizi Baik
  1 13  0  0  0 |  b = Gizi Kurang
  0  0 10  1  0 |  c = Gizi Lebih
  0  0  0  7  0 |  d = Obesitas
  0  0  0  0  7 |  e = Resiko Gizi Lebih
```

Gambar 10. Klasifikasi K-Nearest Neighbour (50% Percentage Split)

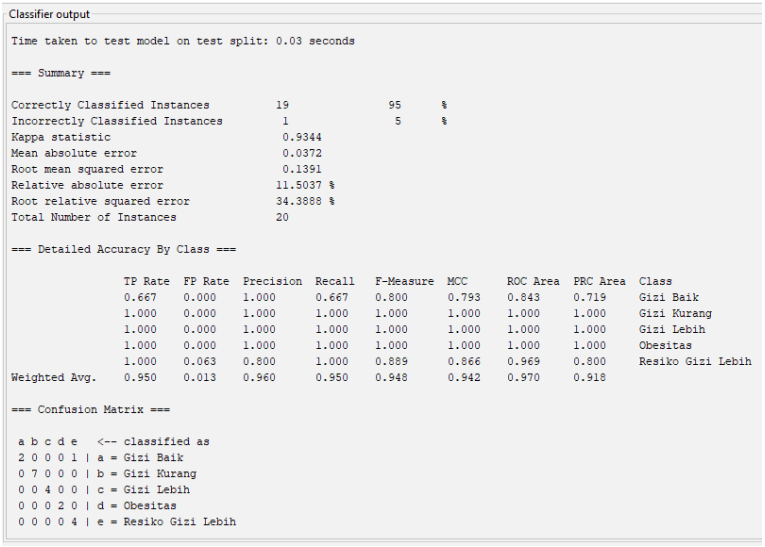
The picture above is the result of the K-Nearest Neighbor classification in the WEKA tool using a 50% Percentage Split which shows the results of 45 correct predictions with an accuracy of 90% and 5 incorrect predictions with a percentage of 10% with a classification time of 0.01 seconds.

Table 14. Confusion Matrix at 50% Percentage Split with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	8	1	0	0	2
Undernutrition	1	13	0	0	0
More Nutritional Risks	0	0	10	1	0
More Nutrition	0	0	0	7	0
Obesity	0	0	0	0	7

Based on the table above, it can be seen that the nutritional status of toddlers can be calcified correctly, namely 45 data from 50 nutritional status data of toddlers. While the other 5 data are classified incorrectly or differently from the original data. Therefore, an accuracy value of 90% is obtained.

5. K-Nearest Neighbor Classification Results Using WEKA Tool (80% Percentage Split)



```
Classifier output

Time taken to test model on test split: 0.03 seconds

=== Summary ===

Correctly Classified Instances          19               95      %
Incorrectly Classified Instances         1                5      %
Kappa statistic                          0.9344
Mean absolute error                      0.0372
Root mean squared error                  0.1391
Relative absolute error                 11.5037 %
Root relative squared error             34.3888 %
Total Number of Instances               20

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.667    0.000    1.000      0.667    0.800      0.793    0.843     0.719     Gizi Baik
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Kurang
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Gizi Lebih
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Obesitas
               1.000    0.063    0.800      1.000    0.889      0.866    0.969     0.800     Resiko Gizi Lebih
Weighted Avg.  0.950    0.013    0.960      0.950    0.948      0.942    0.970     0.918

=== Confusion Matrix ===

 a b c d e   <-- classified as
 2 0 0 0 1 | a = Gizi Baik
 0 7 0 0 0 | b = Gizi Kurang
 0 0 4 0 0 | c = Gizi Lebih
 0 0 0 2 0 | d = Obesitas
 0 0 0 0 4 | e = Resiko Gizi Lebih
```

Gambar 11. Klasifikasi K-Nearest Neighbour (80% Percentage Split)

The picture above is the result of the K-Nearest Neighbor classification in the WEKA tool using an 80% Percentage Split which shows the results of 19 correct predictions with an accuracy of 95% and 1 wrong prediction with a percentage of 5% with a classification time of 0.03 seconds.

Table 15. Confusion Matrix at 80% Percentage Split with WEKA Tools

Predictions	Good Nutrition	Undernutrition	More Risk	More Nutrition	Obesity
Good Nutrition	2	0	0	0	1
Undernutrition	0	7	0	0	0
More Nutritional Risks	0	0	4	0	0
More Nutrition	0	0	0	2	0
Obesity	0	0	0	0	4

Based on the table above, it can be seen that the nutritional status of toddlers can be classified correctly, namely 19 data from 20 nutritional status data of toddlers. While the other 1 data is classified incorrectly or differently from the original data. Therefore, an accuracy value of 95% is obtained.

d. Comparison of K-nearest neighbor accuracy evaluation results from nutritional status data in toddlers

After analyzing the classification of the K-Nearest Neighbor with the Euclidean Distance algorithm in the WEKA tool using the Use Training Set, 4 Fold Cross Validation, 8 Fold Cross Validation, 50% Percentage Split, and 80% Percentage Split, the highest accuracy was obtained using the Use Training Set with an accuracy percentage of 100% for Correctly Classified Instances and 0% for Incorrectly Classified Instances. The comparison can be seen in Table 16.

Table 16. K-Nearest Neighbor Accuracy Evaluation Comparison

Evaluation Model	Accuracy	Number of Toddlers	Percentage
Use Training Set	Correctly Classified Instances	100	100%
	Incorreclty Classified Instances	0	0%
4 Fold Cross-Validation	Correctly Classified Instances	89	89%
	Incorreclty Classified Instances	11	11%
8 Fold Cross-Validation	Correctly Classified Instances	90	90%
	Incorreclty Classified Instances	10	10%
50% Percentage Split	Correctly Classified Instances	45	90%
	Incorreclty Classified Instances	5	10%
70% Percentage Split	Correctly Classified Instances	19	95%
	Incorreclty Classified Instances	1	5%

e. Comparison of Support Vector Machine Method Classification Results with K-Nearest Neighbor

Comparison of classification results using 2 methods, namely Support Vector Machine and K-Nearest Neighbor can be seen in the table below (Faruk & Nafi'iyah, 2020) (Iskandar & Nataliani, 2021).

Table 17. Comparison of SVM and KNN Classification Results

5 tes	SVM	KNN
Test Results	Correctly Classified Instances	Correctly Classified Instances
Use Training Set	100%	100%
4 Fold Cross-Validation	100%	89%
8 Fold Cross-Validation	100%	90%
50% Percentage Split	100%	90%
80% Percentage Split	100%	95%
Percentage Average	**100%**	**93%**

CONCLUSIONS

Based on the results of research that we have conducted using data mining, namely the classification of nutrition in toddlers at UPTD Puskesmas Tigaraksa

using a comparison of the support vector machine (SVM) and k-nearest neighbor (KNN) methods, a conclusion can be drawn, as follows:

a. Classification using the Support Vector Machine (SVM) method with the Radial Basis Function (RBF) kernel doing 5 tests can be obtained the accuracy value of the five tests has the same accuracy, which is 100%, where as many as 100 toddler data are declared correct (Correctly Classified Instances).
b. Classification using the K-Nearest Neighbor (KNN) method with the Euclidean Distance algorithm by conducting 5 tests using only training set data obtained an accuracy value of 100%, where as many as 100 toddler data were declared correct (Correctly Classified Instances) but the next four tests could not produce perfect accuracy scores.

The highest accuracy value in classifying toddler nutrition data using comparisons can be seen in that the Support Vector Machine method provides a more accurate accuracy value with the same percentage, which is 100% on each test tested.

REFERENCES

Amalia, B. S., Umaidah, Y., & Mayasari, R. (2021). Analisis Sentimen Review Pelanggan Restoran Menggunakan Algoritma Support Vector Machine Dan K-Nearest Neighbor. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, *19*(1), 28–34.

Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(1), 147.

Baita, A., Pristyanto, Y., & Cahyono, N. (2021). Analisis Sentimen Mengenai Vaksin Sinovac Menggunakan Algoritma Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN). *Information System Journal*, *4*(2), 42–46.

Budianto, A., Ariyuana, R., & Maryono, D. (2019). Perbandingan K-Nearest Neighbor (KNN) Dan Support Vector Machine (SVM) Dalam Pengenalan Karakter Plat Kendaraan Bermotor. *Jurnal Ilmiah Pendidikan Teknik Dan Kejuruan*, *11*(1), 27–35.

Fadilah, W. R. U., Agfiannisa, D., & Azhar, Y. (2020). Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine. *Fountain Informatics J*, *5*(2), 45.

Hakim, I., Nugroho, A., Sukmana, S. H., & Gata, W. (2020). Sentimen Analisis Stay Home menggunakan metode klasifikasi Naive Bayes, Support Vector Machine, dan k-Nearest Neighbor. *Paradigma-Jurnal Komputer Dan Informatika*, *22*(2020), 169–174.

Ichwan, M., & Dewi, I. A. (2018). Klasifikasi Support Vector Machine (SVM) Untuk Menentukan TingkatKemanisan Mangga Berdasarkan Fitur Warna. *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, *3*(2), 16–23.

Iskandar, J. W., & Nataliani, Y. (2021). Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5* (6), 1120–1126.

Mase, J., Furqon, M. T., & Rahayudi, B. (2018). Penerapan Algoritme Support Vector Machine (SVM) Pada Pengklasifikasian Penyakit Kucing. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN*, *2548*, 964X.

Muhammad Yusuf ramadan, D. S. T. (2019). "Implementasi Metode Klasifikasi Support Vector Machine (SVM) Terhadap Pemakaian Minyak Goreng. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *3* (2).

Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *J. Inform*, *6* (2), 226–235.

Naufal, S. A., Adiwijaya, A., & Astuti, W. (2020). Analisis Perbandingan Klasifikasi Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) untuk Deteksi Kanker dengan Data Microarray. *JURIKOM (Jurnal Riset Komputer)*, *7*(1), 162–168.

Nikmatun, I. A., & Waspada, I. (2019). Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, *10*(2), 421–432.

Pamungkas, F. S., & Kharisudin, I. (2021). Analisis Sentimen dengan SVM, NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter. *PRISMA, Prosiding Seminar Nasional Matematika*, *4*, 628–634.

Permana, R. A., & Sahara, S. (2019). Metode Support Vector Machine Sebagai Penentu Kelulusan Mahasiswa pada Pembelajaran Elektronik. *Jurnal Khatulistiwa Informatika*, *7*(1).

Pratama, I. H., & Salamah, U. (2022). Perbandingan Algoritma K-Nearest Neighbor Dan Support Vector Machine Untuk Menentukan Prediksi Produk-Produk Terlaris Pada Toko Madura Kecamatan Pondok Aren. *JTIK (Jurnal Teknik Informatika Kaputama)*, *6*(2), 846–858.

Saidah, S., Adinegara, M. B., Magdalena, R., & Caecar, N. K. (2019). Identifikasi Kualitas Beras Menggunakan Metode k-Nearest Neighbor dan Support Vector Machine. *TELKA-Jurnal Telekomunikasi, Elektronika, Komputasi Dan Kontrol*, *5*(2), 114–121.

Sugara, B., & Subekti, A. (2019). Penerapan Support Vector Machine (Svm) Pada Small Dataset Untuk Deteksi Dini Gangguan Autisme. *Jurnal Pilar Nusa Mandiri*, *15*(2), 177–182.

Wati, R. A., Irsyad, H., & Al Rivan, M. E. (2020). Klasifikasi Pneumonia Menggunakan Metode Support Vector Machine. *J. Algoritm*, *1*(1), 21–32.