Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Prediction of Stunting in Toddlers Combining the Naive Bayes Method and the C4.5 Algorithm

Sri Melyani[1)*], Syaiful Zuhri Harahap[2)], Irmayanti[3)]
[1)2)3)]Universitas Labuhanbatu, Indonesia,
[1)] srimelyaniii99@gmail.com, [2)]syaifulzuhriharahap@gmail.com, [3)] irmayantiritonga2@gmail.com,

Submitted :April 1, 2024 | **Accepted** : April 20, 2024 | **Published** : April 25, 2024

Abstract: Research conducted to predict the incidence of stunting in toddlers, using data mining methods such as Naive Bayes and the C4.5 algorithm has been applied to analyze health data. The main aim of this research is to develop a predictive model that can identify toddlers who are at high risk of stunting, based on variables that have been collected from medical records and health surveys. The use of the Naive Bayes and C4.5 methods in this research aims to compare the effectiveness of the two methods in dealing with complex and unbalanced classification problems. This research involves a series of crucial stages starting from data selection, data pre-processing, data mining model design, data mining model testing, to method evaluation. In this study, the sample used consisted of 200 toddlers, of which 159 were diagnosed as having stunting and 41 others were not. The classification results show significant effectiveness in both methods used. The accuracy results of both methods are very encouraging, with both methods showing success rates of more than 90%. This shows that both Naive Bayes and C4.5 are very effective in identifying patterns related to the risk of stunting among toddlers. These highly accurate results not only demonstrate the power of data mining techniques in the field of public health but also provide insights that health practitioners can use to intervene earlier in at-risk populations.

Keywords: C4.5 Algorithm; Classification; Data Mining; Metode Decision Tree; Metode Naïve Bayes

INTRODUCTION

Stunting is a condition of failure to thrive in children caused by chronic malnutrition, especially in the first 1000 days of life, starting from pregnancy until the child is two years old (Harjanto, Vatresia, & Faurina, 2021). This condition is characterized by a height that is lower than their age standard. Stunting not only affects a child's physical growth, but also has a long-term impact on brain development, cognitive abilities and future productivity. The main cause of stunting is a lack of adequate nutritional intake during critical periods of growth. Other contributing factors include recurrent infections, limited access to clean water and proper sanitation, and inadequate parenting practices. In many developing countries, stunting is often exacerbated by poor economic conditions and lack of access to adequate health services. Stunting in toddlers is a serious condition that indicates chronic malnutrition during an important period of their growth and development. This condition occurs when a toddler has a height that is significantly shorter than the growth standard for children his age. Stunting is not just a physical problem involving body growth; More than that, stunting also affects children's cognitive development and learning capacity later in life. The main causes of stunting include inadequate nutritional intake, poor sanitation conditions, and low levels of access to health services. The impact of stunting can be very long-term, affecting the quality of life of children and adults, which emphasizes the importance of early nutritional and health interventions.

Stunting in children under five is a serious global health problem, because it has long-term implications that are detrimental to both individuals and society (Kaputama et al., 2021). Children who experience stunting generally have a much shorter height than children their age who grow up in conditions of adequate nutrition. More than just a physical problem, stunting can hinder brain development, which significantly reduces cognitive and learning abilities (Pratistha & Kristianto, 2024). This leads to lower school achievement and reduces their potential productivity as adults. Stunting not only affects academic and economic performance later in life, but also has serious health consequences. Children who are stunted are more susceptible to infections and diseases because their immune systems are not well developed. In addition, they also have a higher risk of experiencing health

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

complications in adulthood, such as diabetes and heart disease. These negative effects not only harm individuals affected by stunting but also burden a country's health and economic systems.

Therefore, the author will conduct research to predict toddlers who will be affected by stunting. The aim of this research is to identify toddlers who are at high risk of experiencing stunting so that intervention can be carried out early to prevent the condition from developing further. Stunting not only impacts children's physical growth, but also their cognitive development and long-term health, making this research particularly important in a public health context. This research will use data mining techniques to analyze and predict stunting incidents. The two methods that will be applied are Naive Bayes and the C4.5 Algorithm. Naive Bayes, which is a simple probability classification method, will be used to predict the probability of stunting based on a historical data set that includes nutritional information, health history and socio-economic conditions. As in research (Mulyanto et al., 2024), the Naive Bayes method was used to conduct stunting research and gave very high accuracy results, namely 95.08%. The C4.5 algorithm, on the other hand, is a machine learning technique that produces decision trees. This decision tree will help in identifying the most significant patterns and variables that contribute to the risk of stunting. This method was also used by (Hakim, Putrawansyah, & Syahri, n.d.). in his research on stunting and the accuracy obtained was also very high, namely 88.20%. The use of these two methods is expected to provide comprehensive and accurate insight into the factors that influence stunting. The data that will be used in this research includes, but is not limited to, anthropometric data, health records, as well as nutritional and socio-economic surveys from various households. It is hoped that the results of this research will not only enrich the academic literature on stunting prediction but also help health practitioners and policy makers in designing and implementing more effective intervention programs.

LITERATURE REVIEW

Data mining is the process of analyzing large sets of data to find patterns, trends and relationships that are not easily seen with manual analysis (Sinaga, Marpaung, Tarigan, & Tania, 2023) (Aji & Devi, 2023). As carried out in research (S. A. Hasibuan, Sihombing, & Nasution, 2023) that data mining can be used to predict the level of satisfaction of Lazada application users and also carried out in research (Pratama, Yanris, Nirmala, & Hasibuan, 2023) that data mining can be used to determine the level of visitor satisfaction at a ride. This technique is often used in various industries such as finance, marketing, healthcare, and retail to make more informed and efficient decisions. Data mining combines statistics, predictive analysis, and artificial intelligence technology to process and analyze data from various sources (Bustomi, Nugraha, Juliane, & Rahayu, 2023) (Saputra, Hindarto, & Haryono, 2023). This process involves several main steps, including data collection, data cleaning, data transformation, data mining, evaluation, and application of results (Abas et al., 2023). This technique allows companies to understand customer needs, improve marketing strategies, optimize operations, and reduce risks.

The Naive Bayes method is a probability-based classification technique based on Bayes' theorem with the assumption that predictors are independent of each other within a class (Alam, Alana, & Juliane, 2023) (Rahman & Fauzi Abdulloh, 2023) (Apriyani, Maskuri, Ratsanjani, Pramudhita, & Rawansyah, 2023). In research conducted (Nasution, Dar, & Nasution, 2023) the naive Bayes method was used to classify students' interest in using gaming laptops and the accuracy obtained was 90%, this means that the naive Bayes method can be used to carry out classification well (F. F. Hasibuan, Dar, & Yanris, 2023). This method is very effective and efficient for large datasets, making it popular in the fields of natural language processing and pattern recognition (Supendar, Rusdiansyah, Suharyanti, & Tuslaela, 2023) (Siregar, Irmayani, & Sari, 2023) (Tanjung, Tampubolon, Panggabean, & Nandrawan, 2023). Naive Bayes is easy to implement and requires a smaller amount of training data to estimate parameters, so it is effective for applications that require fast responses (Madjid, Ratnawati, & Rahayudi, 2023) (Anam, Rahmiati, Paradila, Mardainis, & Machdalena, 2023). Although simple, this method can be very effective if the assumption of independence between features is precise enough, and even if this assumption is violated, its performance is often still quite good (Lubis & Chandra, 2023) (Saleh, Dharshinni, Perangin-Angin, Azmi, & Sarif, 2023).

The C4.5 algorithm is a development of the ID3 algorithm created by Ross Quinlan. This algorithm is used to produce decision trees and is included in the category of machine learning algorithms (Maizura, Sihombing, & Dar, 2023). C4.5 makes decisions by converting training data into a decision tree. Decision trees are built by selecting the most effective attributes to divide the data set into subsets based on class homogeneity. This algorithm uses the concept of information entropy to measure how well an attribute separates training classes. One of the advantages of C4.5 is its ability to handle missing data as well as attributes with many classes, and it also automates tree pruning, which helps avoid overfitting the training data.

METHOD

This research was carried out using the Naive Bayes method and the C4.5 algorithm in data mining. So with this method, the data will be classified, this is because the 2 methods used are methods with classification models. In carrying out this research, there are several stages that will be carried out, namely as follows.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

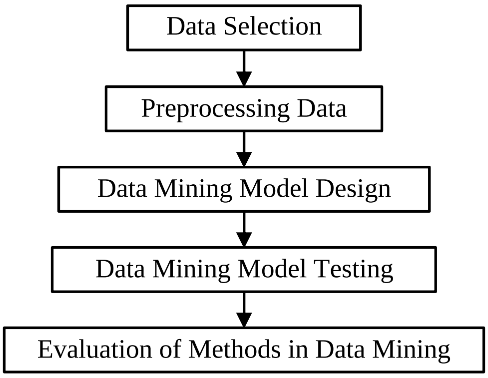

Fig 1. Research Stages

In the work flow above are the stages that the author will carry out so that this research can be carried out well and get good classification results and also get good evaluation results. The author will also explain the flow above so that each stage can be understood.

Data Selection	:	This stage will later be used to collect data that will be used as training and testing data.
Preprocessing Data	:	At this stage the data will be cleaned or selected for data that is suitable for use and the data will be arranged in a format that can be used in this research.
Data Mining Model Design	:	At this stage, we design the model that will be used to classify data in data mining.
Data Mining Model Testing	:	This stage is the result of the data classification process, so this stage will later provide classification results from the previously designed model design.
Evaluation of Methods in Data Mining	:	This stage is the stage that will be used to evaluate the method used to obtain performance results of the method used.

Confusion Matrix

Table 1
Confusion Matrix

		Prediction Class	
	Class	**True**	**False**
Attribute Class	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

From the table above for the explanation below.

1. TP (True Positive), namely the amount of positive data that has a true value.
2. TN (True Negative), namely the amount of negative data that has a true value.
3. FN (False Negative), namely the amount of negative data but which has the wrong value.
4. FP (False Positive), namely the amount of data that is positive but has the wrong value.

$$\boldsymbol{Acuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FN+FP}} \times \mathbf{100\%}$$ (Sari, Yanris, & Hasibuan, 2023)

$$\boldsymbol{Presisi} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \times \mathbf{100}\%$$ (S. A. Hasibuan et al., 2023)

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \times \mathbf{100}\%$$ (Pratama et al., 2023)

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

Data Selection

At this stage, the author will collect data to be used as a research data set. For the research data sets used, there are 2 data sets, namely training data and testing data. The data sets obtained were 250 toddler data sets. For training data, 50 data sets were used and for testing data, 200 toddler data were used. This research process will be carried out by designing a model in data mining using the Naïve Bayes method and the C4.5 algorithm, then after that training data is needed which will later help train the testing data so that it can be classified, therefore there are categories or statuses in the training data namely the aim is to help train testing data so that it can be classified in a previously designed model. For the data that the author used in this research, the author will not explain all the data used, but the author will only provide sample data so that the reader can understand and comprehend it.

Data Training

Training data is data used to assist the data classification process in data mining. For the training data used were 50 toddler data.

Table 2. Data Training

Toddler Names	Age (Months)	Weight	Height	Weight/Age	Height/Age	Status
Aldo Kurniawan	53	13.7	92.8	Normal	Short	Stunting
Andi Palguna	12	9.7	81.1	Normal	Normal	Not Stunting
Bella Arum	18	9.5	81.4	Normal	Short	Not Stunting
Bella Saphira	57	14	96.5	Normal	Short	Stunting
Candra Wilis	37	7.5	75.8	Very less	Very Short	Stunting
Cinta Dewi	11	12.1	80.3	Normal	Normal	Not Stunting
Dedek Mawar	10	12.3	80.2	Normal	Normal	Not Stunting
Dewi	41	10.4	85.5	Not enough	Very Short	Stunting
Dini Hari	23	7.8	74	Not enough	Very Short	Stunting
Elang Rajawali	42	10.4	85.5	Not enough	Very Short	Stunting

The data in the table above is training data consisting of 50 toddler data which will later be used to help with the data processing process.

Data Testing

Testing data is research sample data that will be used to classify data. So this testing data will later be processed and classified using the Naive Bayes method and the C4.5 algorithm.

Table 3. Data Testing

Toddler Names	Age (Months)	Weight	Height	Weight/Age	Height/Age
Adi Saputra	12	10.5	77	Normal	Short
Aji Pangestu	47	12	93.7	Not enough	Short
Akbar Tani	44	13	90	Normal	Short
Aldi Fajar Maulana	11	10.8	79.1	Normal	Normal
Aldo Kurniawan Lubis	15	12.1	90.1	Normal	Normal
Alif Yoga Pratama	5	11	87.2	Normal	Normal
Ana Mariana	35	10.4	85.5	Not enough	Very Short
Andi Baskoro Sulistyo	38	12	89	Normal	Short
Andi Palguna	24	9.8	79	Normal	Short
Anto Surya Wibowo	17	9.5	79.8	Normal	Short

In the table above is research sample data totaling 200 toddler data which will be used as testing data or research sample data. The data above will later be used to make predictions.

Preprocessing Data

At this stage the data that has been obtained will be further selected for its suitability for use, this is done so that the data can be used. Then, after the data has been selected, the data will be arranged in a good form and format. For this research, data was compiled using the Microsoft Office Excel application.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Mining Model Design

This stage is the process of designing a model in data mining so that sample data can be classified. This model will later be used to classify data in data mining using the Naïve Bayes method and the C4.5 algorithm.

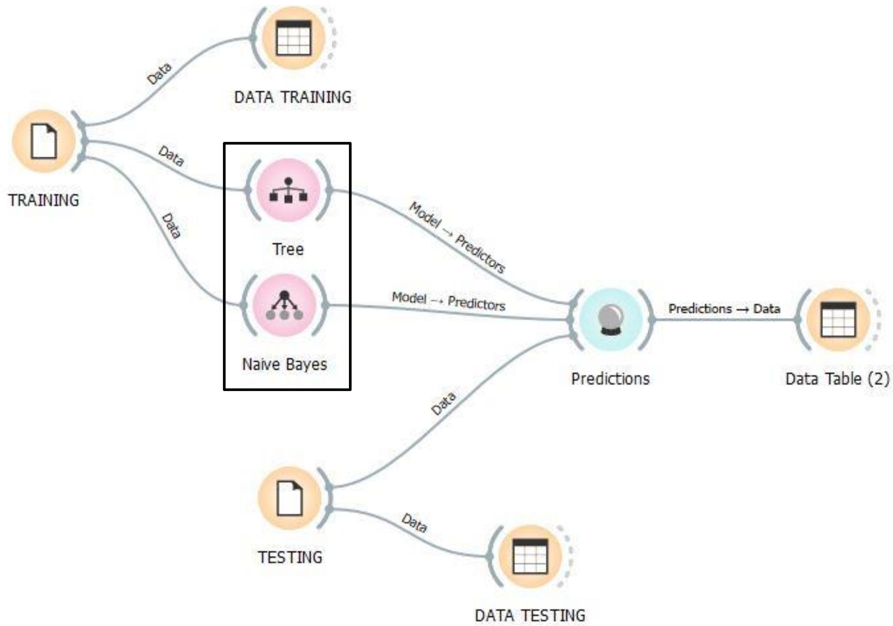

Fig 2. Data Mining Design Model

In the image above is a model designed to classify data in data mining using the Naïve Bayes method and the C4.5 algorithm. The method used can be seen in the image above on the widget in the black box. For the widget above, it is a decision tree method which is used as the C4.5 algorithm, so the C4.5 algorithm is also a decision tree method which carries out classification to obtain a decision tree. Meanwhile, the widget below is the Naïve Bayes method.

Data Mining Model Testing

This test is the result obtained from classifying sample data in data mining using the c4.5 algorithm and the naïve Bayes method. The classification results can be seen in the table below.

Table 4. Classification Results

Toddler Names	Age (Months)	Weight	Height	Weight/Age	Height/Age	Status
Adi Saputra	12	10.5	77	Normal	Short	Not Stunting
Aji Pangestu	47	12	93.7	Not enough	Short	Stunting
Akbar Tani	44	13	90	Normal	Short	Stunting
Aldi Fajar Maulana	11	10.8	79.1	Normal	Normal	Not Stunting
Aldo Kurniawan Lubis	15	12.1	90.1	Normal	Normal	Not Stunting
Alif Yoga Pratama	5	11	87.2	Normal	Normal	Not Stunting
Ana Mariana	35	10.4	85.5	Not enough	Very Short	Stunting
Andi Baskoro Sulistyo	38	12	89	Normal	Short	Stunting
Andi Palguna	24	9.8	79	Normal	Short	Stunting
Anto Surya Wibowo	17	9.5	79.8	Normal	Short	Stunting

In the table above are the results of data classification obtained from designing models in data mining using the Naive Bayes method and the C4.5 algorithm. The classification results were obtained from 200 sample data of toddlers, 159 sample data of toddlers experienced stunting and 41 sample data of toddlers who did not experience stunting.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Evaluation of Methods in Data Mining

In this evaluation the author will also design a model that will be used to evaluate the model that will be used to see the effectiveness of the method for classifying data. In this evaluation there will be 2 methods evaluated, namely the Naive Bayes method and the C4.5 algorithm. The model can be seen in the image below.

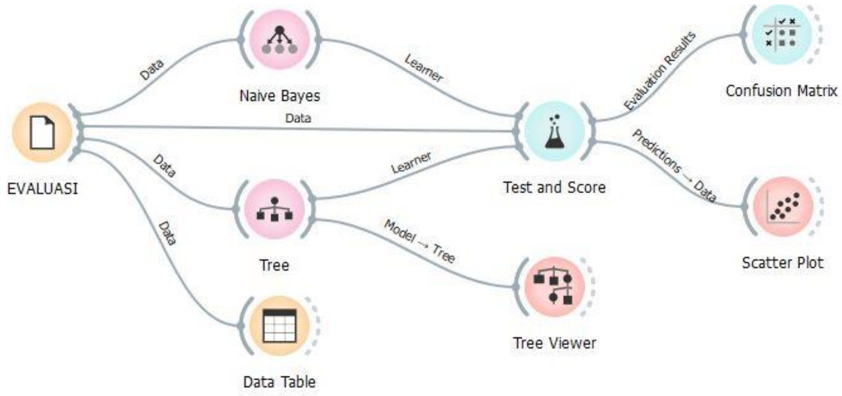

Fig 3. Method Evaluation Model

In the picture above is the evaluation model carried out in this research. The aim is to see the effectiveness of the method used. There are several widgets used to determine the effectiveness of the method used, namely test and score and confusion matrix. For Decision trees, the author uses the Tree Viewer widget and one more widget, namely Scatter Plot.

Test Results and Scores

Table 5. Result of Test and Score

Model	AUC	CA	F1	Precision	Recall	MCC
Naïve Bayes	0.994	0.985	0.985	0.985	0.985	0.854
Algoritma C4.5	0.963	0.985	0.985	0.985	0.985	0.854

In the table above are the evaluation results provided by the test and score widget. The accuracy results above provide quite good results. This is because the results exceed 90%.

Confusion Matrix Results
Confusion Matrix Results of the Naïve Bayes Method

Table 6. Confusion Matrix Results in the Naïve Bayes Method

		Predicted		
		Stunting	Not Stunting	∑
Actual	Stunting	159	0	159
	Not Stunting	3	38	41
	∑	162	38	200

In the table above is the composition of the confusion matrix results obtained from evaluating the Naïve Bayes method in data mining. For the results, there are True Positive (TP) results which are 159 data, for True Negative (TN) results which are 38 data, for False Positive (FP) results which are 0 and for False Negative (FN) results which are 3 data. For these results, it is not possible to directly measure the accuracy value, the data above must be calculated first using the formula in the confusion matrix, which is as follows.

$\boldsymbol{Accuracy = \frac{159+38}{159+38+0+3} + 100\%}$ Then the Accuracy value = 98%

$\boldsymbol{Presisi = \frac{159}{159+0} + 100\%}$ Then the Precision value = 100%

$\boldsymbol{Recall = \frac{159}{159+3} + 100\%}$ Then the Recall value = 98%

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Confusion Matrix Results of the C4.5 Algorithm

Table 7. Confusion Matrix Results in the C4.5 Algorithm

		Predicted		
		Stunting	Not Stunting	∑
Actual	Stunting	159	0	159
	Not Stunting	3	38	41
	∑	162	38	200

In the table above is the composition of the confusion matrix results obtained from evaluating the C4.5 Algorithm in data mining. For the results, there are True Positive (TP) results which are 159 data, for True Negative (TN) results which are 38 data, for False Positive (FP) results which are 0 and for False Negative (FN) results which are 3 data. For these results, it is not possible to directly measure the accuracy value, the data above must be calculated first using the formula in the confusion matrix, which is as follows.

$\boldsymbol{Accuracy} = \frac{\mathbf{159+38}}{\mathbf{159+38+0+3}} + \mathbf{100\%}$ Then the Accuracy value = 98%

$\boldsymbol{Presisi} = \frac{\mathbf{159}}{\mathbf{159+0}} + \mathbf{100\%}$ Then the Precision value = 100%

$\boldsymbol{Recall} = \frac{\mathbf{159}}{\mathbf{159+3}} + \mathbf{100\%}$ Then the Recall value = 98%

Tree Viewer Results

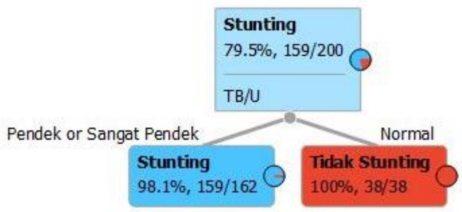

Fig 4. C4.5 Algorithm Decision Tree Results

The results in the image above are the results of the decision tree obtained from the C4.5 method or algorithm.

Scatter Plot Results

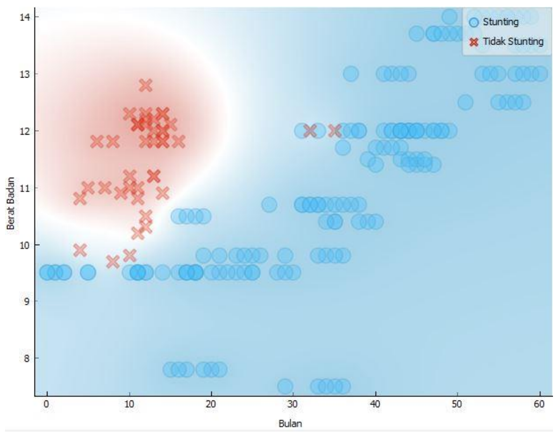

Fig 5. Scatter Plot Results

In the picture above is the result of comparing the number of data on toddlers affected by stunting and toddlers who are not affected by stunting. The results above explain that more toddlers are affected by stunting.

DISCUSSIONS

In recent research aimed at identifying toddlers affected by stunting, two popular data mining methods, namely Naive Bayes and the C4.5 algorithm, have been applied and provided very satisfactory results. This study uses the

*name of corresponding author

DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

'Test and Score' widget to assess the performance of the two models in classifying health data for toddlers. From the analysis carried out, the Naive Bayes method achieved very high accuracy, namely 99%, while the C4.5 algorithm also showed impressive performance with an accuracy of 96%. These two methods have proven to be very effective in detecting cases of stunting among toddlers, with results exceeding the 90% threshold. This shows the reliability of these techniques in extracting relevant insights from complex and large data. Although there is a difference of a few percent between the two methods, the comparison can be considered 1:1 as both offer almost the same advantages in this context. The success of both methods in achieving high accuracy not only confirms their effectiveness in clinical data analysis but also promises great potential for similar applications in other areas of public health. These results provide strong support for the continued use of advanced data mining techniques to aid in early detection and intervention in cases of stunting, which is an important step in combating child malnutrition in many countries.

CONCLUSION

This research confirms that data mining techniques, especially the Naive Bayes method and the C4.5 algorithm, are very effective in predicting the incidence of stunting in toddlers. Both methods show very high levels of accuracy, with Naive Bayes reaching 99% and C4.5 reaching 96%. These results prove that both methods are reliable in identifying children under five who are at risk of stunting, thereby enabling earlier intervention that can mitigate the long-term impacts of this nutritional problem. The use of Naive Bayes and C4.5 in this research also highlights the importance of choosing the right method for the specific data and analysis goals. Although the two methods have differences in accuracy levels, the differences are not very significant, indicating that both are equally efficient in the context of toddler health data. Their success indicates great potential in the application of similar predictive models in the health sector to address various public health issues. Overall, this research makes a significant contribution to efforts to improve the quality of children's health in the world. With the ability to accurately detect cases of stunting early, health care providers can intervene more quickly, which in turn can reduce the prevalence of stunting and improve long-term health outcomes for future generations. This research also paves the way for further studies that could explore the application of these models in other clinical scenarios or the development of newer models with even higher accuracy.

REFERENCES

Abas, M. I., Ibrahim, I., Syahrial, S., Lamusu, R., Baderan, U. S., & Kango, R. (2023). Analysis of Covid-19 Growth Trends Through Data Mining Approach As Decision Support. *Sinkron*, *8*(1), 101–108. https://doi.org/10.33395/sinkron.v8i1.11861

Aji, G. W., & Devi, P. A. R. (2023). Data Mining Implementation For Product Transaction Patterns Using Apriori Method. *Sinkron*, *8*(1), 421–432. https://doi.org/10.33395/sinkron.v8i1.12071

Alam, A., Alana, D. A. F., & Juliane, C. (2023). Comparison Of The C.45 And Naive Bayes Algorithms To Predict Diabetes. *Sinkron*, *8*(4), 2641–2650. https://doi.org/10.33395/sinkron.v8i4.12998

Anam, M. K., Rahmiati, R., Paradila, D., Mardainis, M., & Machdalena, M. (2023). Application of Naïve Bayes Algorithm for Non-Cash Food Assistance Recipients in Kampar Regency. *Sinkron*, *8*(1), 433–441. https://doi.org/10.33395/sinkron.v8i1.12032

Apriyani, M. E., Maskuri, R. A., Ratsanjani, M. H., Pramudhita, A. N., & Rawansyah, R. (2023). Digital Forensic Investigates Sexual Harassment on Telegram using Naïve Bayes. *Sinkron*, *8*(3), 1409–1417. https://doi.org/10.33395/sinkron.v8i3.12514

Bustomi, Y., Nugraha, A., Juliane, C., & Rahayu, S. (2023). Data Mining Selection of Prospective Government Employees with Employment Agreements using Naive Bayes Classifier. *Sinkron*, *8*(1), 1–8. https://doi.org/10.33395/sinkron.v8i1.11968

Hakim, R. X., Putrawansyah, F., & Syahri, R. (n.d.). *Penerapan Algoritma C4 . 5 Untuk Prediksi Anak Stunting Di Kota Pagar Alam*. *18*(2), 269–279.

Harjanto, T. D., Vatresia, A., & Faurina, R. (2021). *STUNTING MENGGUNAKAN METODE*. *9*(1), 30–42.

Hasibuan, F. F., Dar, M. H., & Yanris, G. J. (2023). Implementation of the Naïve Bayes Method to determine the Level of Consumer Satisfaction. *SinkrOn*, *8*(2), 1000–1011. https://doi.org/10.33395/sinkron.v8i2.12349

Hasibuan, S. A., Sihombing, V., & Nasution, F. A. (2023). Analysis of Community Satisfaction Levels using the Neural Network Method in Data Mining. *Sinkron*, *8*(3), 1724–1735. https://doi.org/10.33395/sinkron.v8i3.12634

Kaputama, S., Data, A., Mendapatkan, T., Ekslusifdan, A. S. I., Kurang, J., Makanandan, A., … Stunting, A. (2021). *Data Mining Pengelompokan Anak Stunting Berdasarkan Usia , Penyebab dan Pekerjaan Orang Tua Dengan Menggunakan Metode Clustering (Studi Kasus : Dinas Kesehatan Kabupaten Langkat)*.

Lubis, A. I., & Chandra, R. (2023). Forward Selection Attribute Reduction Technique for Optimizing Naïve Bayes Performance in Sperm Fertility Prediction. *Sinkron*, *8*(1), 275–285. https://doi.org/10.33395/sinkron.v8i1.11967

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 2, April 2024
DOI : https://doi.org/10.33395/v8i2.13651
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Madjid, F. M., Ratnawati, D. E., & Rahayudi, B. (2023). Sentiment Analysis on App Reviews Using Support Vector Machine and Naïve Bayes Classification. *Jurnal Dan Penelitian Teknik Informatika*, *8*(1), 556–562. Retrieved from https://doi.org/10.33395/sinkron.v8i1.12161

Maizura, S., Sihombing, V., & Dar, M. H. (2023). Analysis of the Decision Tree Method for Determining Interest in Prospective Student College. *SinkrOn*, *8*(2), 956–979. https://doi.org/10.33395/sinkron.v8i2.12258

Mulyanto, Y., Idifitriani, F., Wati, A., Sumbawa, U. T., Mining, D., & Tano, K. P. (2024). *Vol 7 No 2 , September 2024 KLASIFIKASI DATA MINING UNTUK PENENTUAN STUNTING*. *7*(2), 129–135.

Nasution, R. F., Dar, M. H., & Nasution, F. A. (2023). Implementation of the Naïve Bayes Method to Determine Student Interest in Gaming Laptops. *Sinkron*, *8*(3), 1709–1723. https://doi.org/10.33395/sinkron.v8i3.12562

Pratama, H. A., Yanris, G. J., Nirmala, M., & Hasibuan, S. (2023). *Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels*. *8*(3), 1832–1851.

Pratistha, R. N., & Kristianto, B. (2024). *Implementasi Algoritma K-Means dalam Klasterisasi Kasus Stunting pada Balita di Desa Randudongkal Abstrak*. *5*(2), 1193–1205.

Rahman, R., & Fauzi Abdulloh, F. (2023). Performance of Various Naïve Bayes Using GridSearch Approach In Phishing Email Dataset. *Sinkron*, *8*(4), 2336–2344. https://doi.org/10.33395/sinkron.v8i4.12958

Saleh, A., Dharshinni, N., Perangin-Angin, D., Azmi, F., & Sarif, M. I. (2023). Implementation of Recommendation Systems in Determining Learning Strategies Using the Naïve Bayes Classifier Algorithm. *Sinkron*, *8*(1), 256–267. https://doi.org/10.33395/sinkron.v8i1.11954

Saputra, A. D. S., Hindarto, D., & Haryono, H. (2023). Supervised Learning from Data Mining on Process Data Loggers on Micro-Controllers. *Sinkron*, *8*(1), 157–165. https://doi.org/10.33395/sinkron.v8i1.11942

Sari, M., Yanris, G. J., & Hasibuan, M. N. S. (2023). Analysis of the Neural Network Method to Determine Interest in Buying Pertamax Fuel. *SinkrOn*, *8*(2), 1031–1039. https://doi.org/10.33395/sinkron.v8i2.12292

Sinaga, B., Marpaung, M., Tarigan, I. R. B., & Tania, K. (2023). Implementation of Stock Goods Data Mining Using the Apriori Algorithm. *Sinkron*, *8*(3), 1280–1292. https://doi.org/10.33395/sinkron.v8i3.12852

Siregar, A. P., Irmayani, D., & Sari, M. N. (2023). Analysis of the Naïve Bayes Method for Determining Social Assistance Eligibility Public. *SinkrOn*, *8*(2), 805–817. https://doi.org/10.33395/sinkron.v8i2.12259

Supendar, H., Rusdiansyah, R., Suharyanti, N., & Tuslaela, T. (2023). Application of the Naïve Bayes Algorithm in Determining Sales Of The Month. *SinkrOn*, *8*(2), 873–879. https://doi.org/10.33395/sinkron.v8i2.12293

Tanjung, J. P., Tampubolon, F. C., Panggabean, A. W., & Nandrawan, M. A. A. (2023). Customer Classification Using Naive Bayes Classifier With Genetic Algorithm Feature Selection. *Sinkron*, *8*(1), 584–589. https://doi.org/10.33395/sinkron.v8i1.12182

*name of corresponding author