Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

Optimization of C4.5 Algorithm Performance Using Particle Swarm Optimization in Predicting Stunting Risk

Khaerul Ma'mur
Informatics Engineering, Faculty of Computer Science, Pamulang University, South Tangerang, Banten, Indonesia

Article Info
Keywords:
Prediction,
C4.5 algorithm,
PSO,
Stunting.

ABSTRACT
Stunting is a serious global health problem, especially in developing countries. It is caused by chronic malnutrition in children, especially toddlers, which inhibits physical and cognitive growth. Stunting also has the potential to reduce quality of life and productivity in the future. Therefore, early detection of stunting risk is crucial so that appropriate interventions can be provided. Currently, data mining-based classification methods, such as the C4.5 algorithm, have been widely used to predict stunting risk. However, the performance of the C4.5 algorithm in terms of accuracy and efficiency is still lacking, especially in attribute selection and parameter settings. This research aims to improve the accuracy of the C4.5 algorithm in predicting stunting risk by implementing Particle Swarm Optimization (PSO) as an optimization technique. PSO is chosen because of its ability to find optimal solutions quickly and efficiently through the principles of particle social behavior. By using PSO, this research is expected to optimize the attribute selection process and parameter settings in the C4.5 algorithm, so as to produce a more accurate classification model in detecting stunting risk. The result of this research is a significant increase in prediction accuracy compared to the use of the C4.5 algorithm without optimization, so that the resulting model can be a more reliable tool for governments, health institutions, and other policy makers in designing interventions and strategies to overcome stunting.

Corresponding Author:
Khaerul Ma'mur
Pamulang University, South Tangerang, Banten, Indonesia
dosen00844@unpam.ac.id

INTRODUCTION

Stunting is a condition in which children under five experience stunted growth, usually of short stature, which is a serious problem in children's health (Putra & Muhammad, 2024). It is defined as impaired growth in children under five due to chronic malnutrition and recurrent infections that inhibit physical growth usually in the first 1000 days of life (Budiastutik & Nugraheni, 2018) .

Based on the World Health Organization (WHO) definition, stunting is measured by a child's height that is significantly lower than the age-appropriate growth standard. Stunting in children under five has a serious impact on their future, as it can affect cognitive abilities, mental development, and productivity as adults. In Indonesia in particular, the problem of stunting continues to be a serious concern for the government and society, with a high

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

prevalence in some regions. Because stunting can increase the risk of death and various other health problems (Anwar et al., 2022) .

The main problem with stunting is the difficulty of early detection. Generally, stunting is only recognized after children experience significant growth disturbances. Therefore, a preventive approach through early prediction is very important so that appropriate interventions can be provided to reduce the adverse effects of stunting. So it is necessary to develop an accurate prediction model to detect the risk of stunting based on various related factors. One of them is the use of the C4.5 algorithm as a modeling guide.

In data-based research, the C4.5 algorithm has become one of the classification algorithms that is often used to build decision trees from available datasets (Sulistiyanto, 2018) . This algorithm was first developed by Ross Quinlan in 1993 (Xsanal Hakim et al., 2024) as an improvement of the ID3 algorithm. The advantage of this algorithm is its ability to work with discrete data as well as numerical data (Saleh, 2020) , as well as being able to handle missing values and pruning to reduce overfitting. Despite its popularity, the C4.5 algorithm has disadvantages such as dependence on attribute selection and sensitivity to unbalanced data distribution, which can decrease accuracy and increase complexity.

To overcome this limitation, various optimization methods have been proposed in the literature, one of which is the use of a metaheuristic-based optimization algorithm, namely Particle Swarm Optimization (PSO). PSO optimization uses an optimization technique that utilizes a group of particle populations to find and determine the optimal solution to a given problem (Sinaga et al., 2024) . In its optimization, PSO can be used to find the optimal combination of attributes and parameters in the C4.5 algorithm, resulting in better accuracy.

Various previous studies have examined the use of the C4.5 algorithm in various classification models and applications. Studies conducted by (Prasetya et al., 2022) . using the C4.5 algorithm to predict Hepatitis C disease optimized with the PSO algorithm resulted in an optimal accuracy rate of 99.67%. Other research by (Hayadi & Damanik, 2022) with the Machine Learning approach is able to analyze understanding in website programming and the accuracy results are quite good at 83%.

Research by (Fanani Rudi & Fikriah Katul, 2023) also shows that the combination of C4.5 and PSO successfully improves the accuracy of the violence type classification model. Another study by (Nurcahyo et al., 2023) also showed similar results in the case of breast cancer disease prediction. The result is that the C4.5 algorithm integrated with PSO is able to increase the accuracy of breast cancer detection up to 10% better.

Through PSO optimization, the model formed is able to improve the selection of more relevant attributes and reduce prediction errors that often occur in the C4.5 algorithm without optimization. With higher prediction accuracy (Br. Tarigan et al., 2020) , the results of this study can make a significant contribution in mapping the risk of stunting in children under five, so that it can be used as a tool in more effective early intervention.

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

METHODS

Type of Research

This research uses a quantitative approach, which focuses on numerical data and systematic analysis of measurable phenomena (Rohman et al., 2020) . This approach was chosen because it is suitable for comparing the performance of two algorithms objectively. Of the three main types in quantitative-inferential, experimental, and simulation. This research adopts a comparative experimental design, which compares the results of the C4.5 and C4.5 methods that have been optimized using PSO to determine the best level of accuracy.

Research Instruments

The instruments in this research are algorithm performance testing tools, namely datasets, processing tools, and evaluation metrics.

a. The dataset used is sourced from the Kaggle public repository, taken from previous research.
b. The data consists of 6500 random samples with two class labels, presented in tabular format.
c. The analysis process was conducted using RapidMiner software to make modeling and evaluation more efficient and structured.

Research Methods

Research methods in terms of data mining refer to a set of techniques and approaches used to extract useful knowledge or patterns from large and complex data sets (Ma'mur et al, 2024) . In general, the research methods that can be used to solve this research problem are as follows:

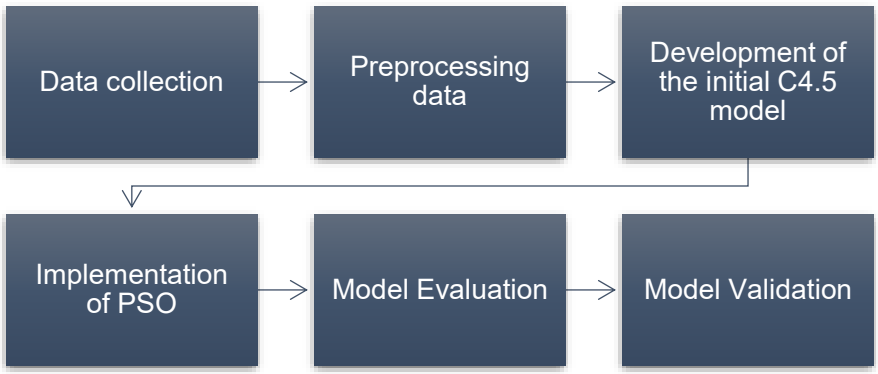


Figure 1. Research Methods

In general, the stages in the research method as shown in the chart above are arranged to facilitate the flow of research completion. Where the initial stage is data collection (Anggita & Ikmah, 2021) , then preprocessing the data, after that developing an initial model to see the decision tree model, which is then implemented PSO to increase performance, after which it is evaluated, and finally validated using Confusion Matrix and T-test.

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

RESULTS AND DISCUSSION

Dataset

In this research, secondary data derived from Kaggle open data is used. There are 6500 samples presented randomly with ownership of 2 decision prediction class labels, namely Yes as many as 3312, and No as many as 3188. The data samples obtained are listed in the table below.

Table 1. Research dataset

Gender	Age	Birth Weight	Birth Length	Body Weight	Body Length	Breastfeeding	Stunting
female	56	2.9	50	11	90	Yes	No.
female	20	3.3	49	11.1	80.5	No.	No.
male	4	2.8	48	6.5	63	No.	No.
female	14	2	49	7	71	Yes	No.
male	32	3.2	49	11	88.7	Yes	No.
male	30	2.3	50	12	90	Yes	No.
male	2	2.9	49	8.5	74.2	Yes	No.
male	33	2.5	49	10	91.5	No.	Yes
male	33	3	50	15	96	Yes	No.
female	15	4	51	5.9	58.3	Yes	No.
male	16	2.4	48	6.6	76	Yes	Yes
female	47	2.5	49	10	91.5	No.	Yes
female	31	3.3	49	9.1	90	Yes	No.
female	25	2.6	49	9.8	90	No.	No.
female	4	2.7	47	5.8	69.6	Yes	No.
female	12	3.6	50	8.2	70.5	Yes	No.
female	3	2.8	48	6	54	No.	Yes
female	55	3	49	8.5	81	No.	No.
female	12	2.9	49	5.8	69.5	Yes	No.
....	....	.....	....	....	....	....	....
male	53	2.9	49	15	96	No.	Yes
male	9	2.9	50	7.3	62	No.	Yes
female	20	1.8	48	7.3	73	Yes	Yes
male	11	2.9	49	7.7	66	No.	Yes
female	14	2.9	49	6.5	66	No.	Yes

The table above lists several attributes that have different roles in determining the label results. An explanation of each attribute can be seen in the table below.

Table 2. Dataset Attribute Explanation

No.	Attributes	Explanation
1	Gender	Gender of respondent
2	Age	Respondent's age at screening (month)
3	Birth Weight	Weight at first birth (grams)
4	Birth Length	Height at first birth (cm)

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

No.	Attributes	Explanation
5	Body Weight	Body weight at examination (grams)
6	Body Length	Height at examination (cm)
7	Breastfeeding	Consumption of breast milk during infancy
8	Stunting	Stunting concordance prediction results

Data Preprocessing

At this stage, none of the values in the attributes are indicated as missing values, so the dataset presented does not require clearing, normalizing, or correcting the data. One of the reasons why data preprocessing is not done during the search is because this algorithm has a strong built-in ability to handle raw or minimally preprocessed data.

C4.5 can effectively handle issues such as missing values, mixed data types, and diverse scales without requiring complex preprocessing steps. By using raw data directly, it can take advantage of the diversity and overall information contained in the dataset without the risk of losing details or affecting the analysis results. So that it can be continued to the next stage.

Initial Model Building

This stage is carried out modeling in the form of a decision tree which is determined based on the calculation of entropy, information gain, and gain ratio using the Rapidminer tool. Each attribute is calculated for the total information gain, where the highest gain will be the initial node in the formation of the decision tree. This step is done to determine the number of classes that are declared Yes and No.

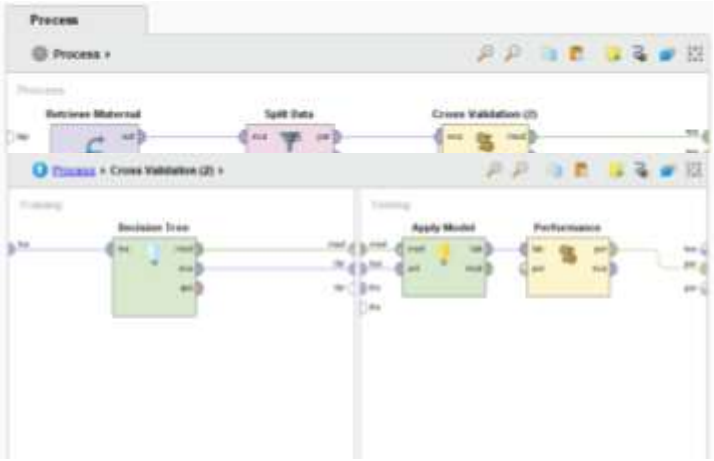


Figure 2. C4.5 algorithm modeling process using Rapidminer

As shown in the figure, the dataset that has been entered into the worksheet is linked to split data with the intention that the training data and testing data can be automatically separated at a ratio of eight to two. Then it is connected back to a validation operator called cross validation. Then the next step in the Training work box is entering the Decision Tree (C4.5) modeling operator. After that, enter the scoring operator, namely Apply Model, followed by the Validation operator, namely Performance in the Testing box to see how much

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

the accuracy value is. Then click the play/Run icon. So as to produce a decision tree model, rules, and performance from using the C4.5 algorithm without PSO optimization. The following is an image of the decision tree results.

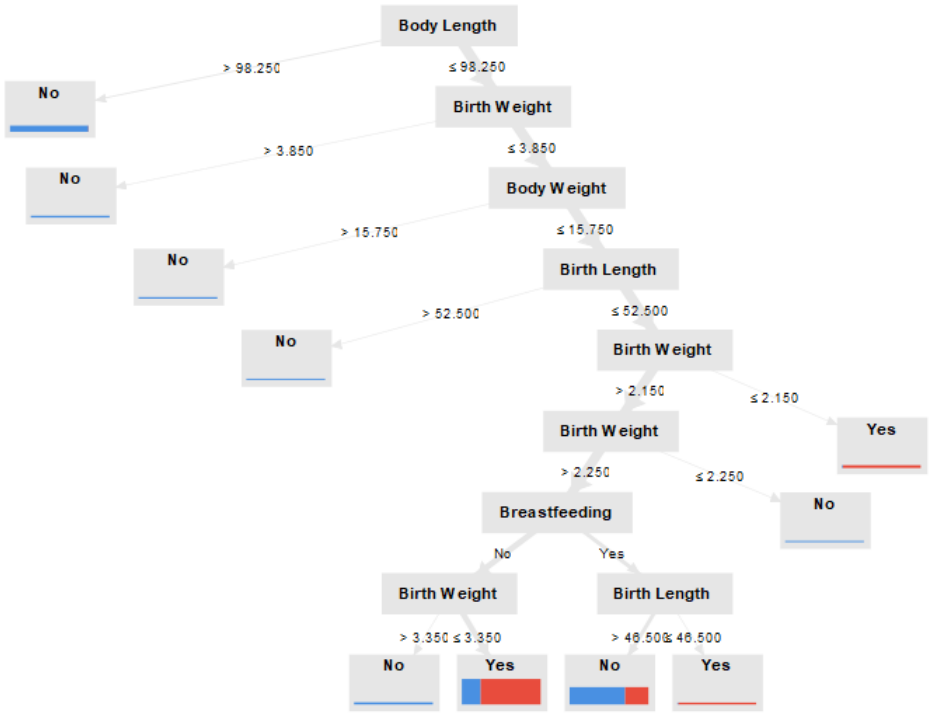


Figure 3. The result of the decision tree with the C4.5 algorithm

The main purpose of data analysis using the Decision Tree algorithm is to obtain rules that can be used in decision making on new data (Azuaje, 2006) . From the decision tree, a rule is obtained in predicting the risk of stunting based on the dataset that has been processed as follows.

1. IF Body Length > 98,250 THEN no
2. IF Body Length > 98,250 AND Birth Weight > 3,850 THEN no
3. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight > 15,750 THEN no
4. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length > 52,500 THEN no
5. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight > 2,150 AND Birth Weight > 2,250 AND Breastfeeding = No AND Birth Weight > 3,350 THEN no
6. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight > 2,150 AND Birth Weight > 2,250 AND Breastfeeding = No AND Birth Weight ≤ 3,350 THEN yes

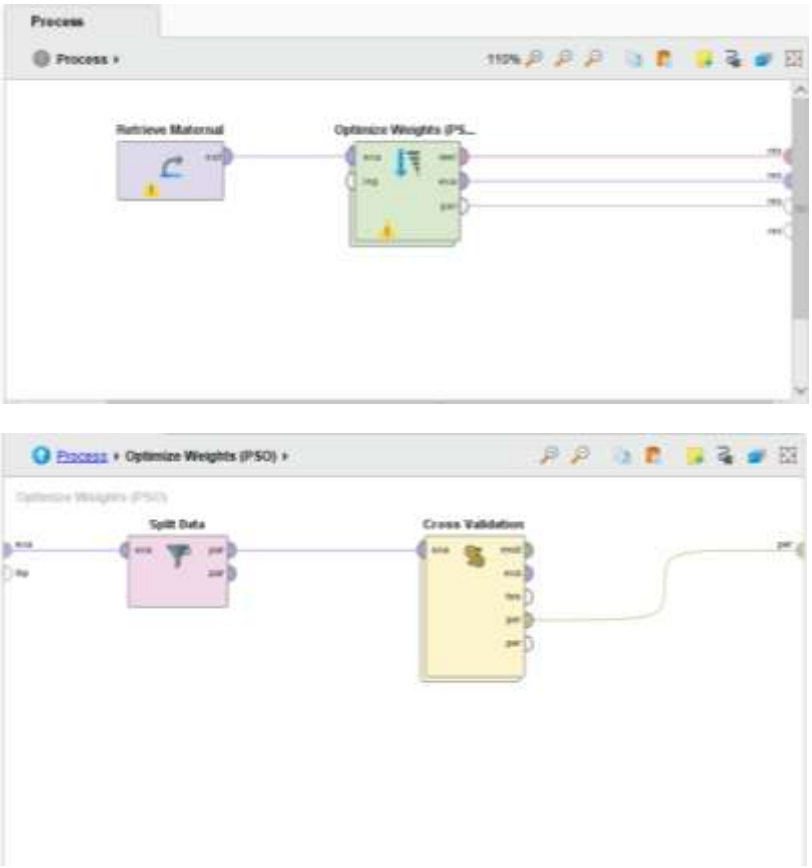
Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

7. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight > 2,150 AND Birth Weight > 2,250 AND Breastfeeding= Yes AND Birth Length > 46,500 THEN no
8. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight > 2,150 AND Birth Weight > 2,250 AND Breastfeeding = Yes AND Birth Length ≤ 46,500 THEN yes
9. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight > 2,150 AND Birth Weight ≤ 2,250 THEN no
10. IF Body Length ≤ 98,250 AND Birth Weight ≤ 3,850 AND Body Weight ≤ 15,750 AND Birth Length ≤ 52,500 AND Birth Weight ≤ 2,150 THEN yes

At this stage, the percentage level of accuracy obtained using the cross validation method that was included during the formation of the previous decision tree model is also seen. This is done as a determinant of the initial number of accuracy levels to be compared. The results of the accuracy level are discussed in the evaluation point.

PSO Optimization

At this stage, particle swarm optimization (PSO) optimization is added to the C4.5 algorithm processing. This stage is carried out to provide an increase in the level of accuracy in determining the risk of stunting according to the existing dataset. Optimization is done by still using the rapidminer tool by adding the Optimize Weight (PSO) operator to the work area. The following flow is applied to the PSO optimization process as shown below.



Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

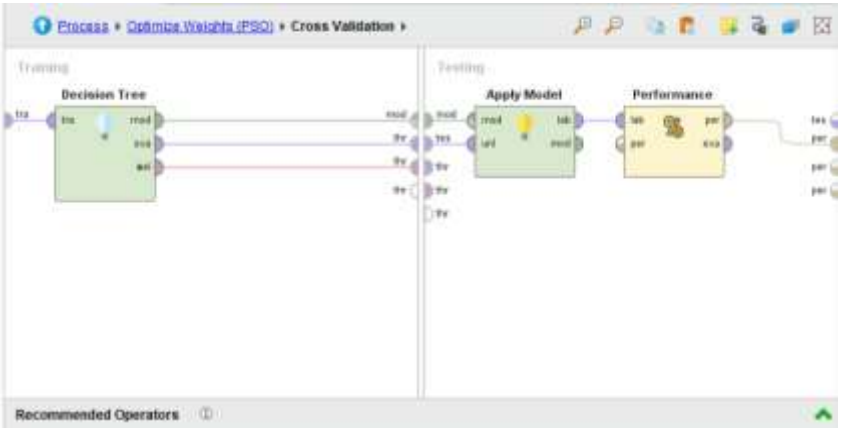


Figure 4. Flow of PSO optimization addition

In the initial worksheet as shown in the figure, the process of connecting the operator with the dataset source is carried out. In the first flow, the dataset is connected to the PSO operator as an optimization tool. Then in the PSO operator, it will open a second worksheet filled with split data operators that have the same ratio ratio as the initial modeling stage. Then it is also connected to the cross validation operator. After connecting, the third worksheet will appear which is divided into a training page and a testing page. On the training page, the C4.5 algorithm model operator is given, namely Decision Tree. While the testing page is given the apply model and performance operators. The results of the additional optimization are presented in the evaluation point.

Model Evaluation

At the model evaluation stage, also known as the classification stage, testing is carried out to observe the results of the accuracy value. This test involves observing the accuracy results of the classification process using the C4.5 algorithm and the C4.5 algorithm that has been optimized by Particle Swarm Optimization (PSO), then the evaluation method uses a confusion matrix.

Confusion matrix is an evaluation tool used to assess the performance of classification models, including the C4.5 algorithm, which is a machine learning algorithm used to create decision trees. In RapidMiner, the confusion matrix helps understand the extent to which the trained model is able to classify data correctly and identify misclassifications.

The confusion matrix model at this stage produces a 2x2 matrix according to the number of class labels. The data set that has been calculated before the addition of PSO optimization in the previous stage is displayed in the form of a confusion matrix to see the accuracy value and available tuples. The resulting data is shown in the figure below:

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

accuracy: 80.63% +/- 2.01% (micro average: 80.63%)

	true No	true Yes	class precision
pred. No	1943	400	82.93%
pred. Yes	607	2250	78.75%
class recall	76.20%	84.91%	

Figure 5. Confusion matrix results on the C4.5 algorithm

From the figure presented, to calculate the accuracy value that will be compared using a formula or equation. Accuracy is calculated as the percentage of correct predictions compared to the total predictions that have been determined according to the ratio. The equation is:

$$\text{Accuracy} = \frac{\text{number of correct predictions}}{\text{Total prediction}}$$

Thus, the calculation of the accuracy value based on the confusion matrix presented is:

$$\text{Accuracy} = \frac{1943 + 2250}{1943 + 400 + 607 + 2250}$$

$$\text{Accuracy} = \frac{4193}{5200}$$

$$\text{Accuracy} = 80{,}63$$

The calculation results above show the accuracy value. Then the accuracy value obtained using the C4.5 algorithm without Particle Swarm Optimization (PSO) optimization is 80.63. Furthermore, the accuracy value of the C4.5 algorithm using PSO optimization is calculated again as a comparison by displaying the results of confusion matrix performance.

accuracy: 85.94% +/- 1.09% (micro average: 85.94%)

	true No	true Yes	class precision
pred. No	1975	156	92.68%
pred. Yes	575	2494	81.26%
class recall	77.45%	94.11%	

Figure 6. Confution matrix results of the C4.5 algorithm using PSO optimization

From the figure above, the accuracy value is calculated again using the equation as before. Thus, the calculation of the accuracy value based on the confusion matrix presented is:

$$\text{Accuracy} = \frac{1975 + 2494}{1975 + 156 + 575 + 2494}$$

$$\text{Accuracy} = \frac{4469}{5200}$$

$$\text{Accuracy} = 85{,}94$$

Based on the results of the above calculations, the accuracy value generated in predicting the risk of stunting using the C4.5 algorithm with PSO optimization is 85.94.

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

Model Validation

Based on the model evaluation conducted, looking at the confusion matrix results shows that the C4.5 classification algorithm using Particle Swarm Optimization (PSO) has a higher accuracy than the conventional C4.5 classification algorithm. The C4.5 algorithm has an accuracy of 80.63%, while the accuracy of the PSO-based C4.5 algorithm is 85.94%, with an accuracy difference of 5.31%, as shown in the table below.

Table 3 . Comparison results of C4.5 classification algorithm testing

Algorithm	Accuracy
C4.5	80,63
PSO-based C4.5	85,94

For further testing, a statistical test was conducted using T-Test. This test compares two algorithms, namely the C4.5 algorithm and the PSO-based C4.5 algorithm, resulting in the following data:

Table 4 . T-test statistical test results

	C4.5	PSO-based C4.5
C4.5		0,049
PSO-based C4.5	0,049	

The results of Table 4.4 show the results of T-Test testing that compares two algorithms in turn, as shown in the figure below using the rapidminer tool.

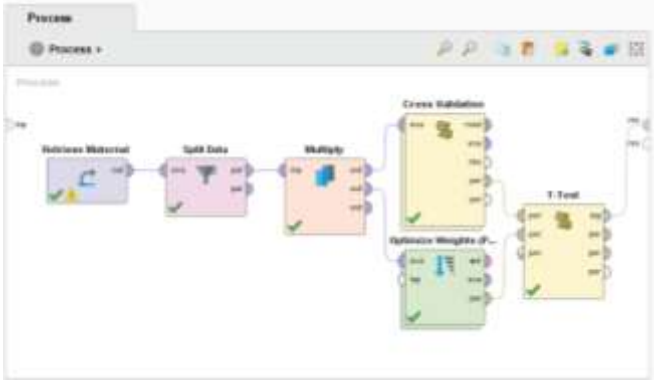


Figure 7. Statistical testing model of t-test

When the model as shown above is run, it will get the statistical calculation results as shown below.

Pairwise t-Test (T-Test)

A	B	C
	0.806 +/- 0.020	0.851 +/- 0.015
0.806 +/- 0.020		0.000
0.851 +/- 0.015		

Figure 8. Results of t-test statistics

Jurnal Info Sains : Informatika dan Sains
Volume 15, Number 01, 2025, DOI 10.54209/infosains.v15i01
ESSN 2797-7889 (Online)
https://ejournal.seaninstitute.or.id/index.php/InfoSains

Based on Figure 4.7, there is a significant difference between group C and groups A and B, indicated by a p-value of 0.000 ($p < 0.05$). This indicates that the mean score in group C (0.851 ± 0.015) is statistically higher than A (0.806 ± 0.020) and B (0.805 ± 0.020). Meanwhile, the comparison between groups A and B resulted in a p-value of 0.800 ($p > 0.05$), which means there is no significant difference between them. So it can be analyzed that the PSO-based C4.5 algorithm shows a significant difference in value with a probability of <0.05 compared to the C4.5 algorithm. This shows that the PSO-based C4.5 algorithm has better accuracy in predicting stunting risk.

CONCLUSION

Based on the evaluation of the C4.5 algorithm *optimization* results using *particle swarm optimization* (PSO) without pruning, it can be concluded that the C4.5 algorithm model that has been optimized with PSO achieves an accuracy rate of 85.94%. This result shows an increase in performance compared to the standard C4.5 algorithm model which only achieves an accuracy rate of 80.63%. The difference between the two models is 5.31%, which shows the superiority of the PSO optimization approach in improving the prediction accuracy of the C4.5 algorithm. So the application of PSO optimization techniques has proven effective in increasing the accuracy value of the C4.5 classification algorithm. This shows that the results of this research have the potential to be implemented as a system in the future.

REFERENCE

Anggita, S. D., & Ikmah, I. (2021). Implementation of Pso for Attribute Weight Optimization in C4.5 Algorithm in Predicting Student Graduation. *JIPI (Scientific Journal of Informatics Research and Learning)*,*6* (2), 416-423. https://doi.org/10.29100/jipi.v6i2.2440

Anwar, S., Winarti, E., & Sunardi, S. (2022). Systematic Review of Risk Factors, Causes and Impacts of Stunting in Children. *Journal of Health Sciences*,*11* (1), 88. https://doi.org/10.32831/jik.v11i1.445

Azuaje, F. (2006). Data Mining: Practical Machine Learning Tools and Techniques 2nd edition. *BioMedical Engineering OnLine*,*5* (1). https://doi.org/10.1186/1475-925x-5-51

Br. Tarigan, D. M., Rini, D. P., & Samsuryadi. (2020). Feature Selection in Blood Sugar Disease Classification Using Particle Swarm Optimization (PSO) on C4.5 Algorithm. *RESTI Journal (System Engineering and Information Technology)*, *4*(3), 569–575.

Budiastutik, I., & Nugraheni, S. A. (2018). Determinants of Stunting in Indonesia: A Review Article. *International Journal Of Healtcare Research*, *1*(2), 43–49.

Fanani Rudi, M., & Fikriah Katul, F. (2023). PSO Feature Selection for Violence Type Classification with C4.5 Algorithm. *Smart Comp: The Journal of Computer Smart People*,*12* (1). https://doi.org/10.30591/smartcomp.v12i1.4407

Hayadi, B. H., & Damanik, A. R. (2022). Machine Learning Approach Using Pso-Based C4.5 Algorithm in Analyzing Website Programming Understanding. *Journal of Informatics and Applied Electrical Engineering*, *10*(3).

Ma'mur, K., Maulana, A. E., Informatics, T., Pamulang, U., Raya, J., & No, P. (2024). *PSO Optimization to Improve the Performance of C4. 5 algorithm in predicting pregnancy*

health risks. *9*(4).

Nurcahyo, R., Fanani, A. Z., Affandy, A., & Aziz, M. I. (2023). Improvement of C4 Algorithm. 5 algorithm based on PSO on breast cancer disease. *Journal of Budidarma Informatics Media*,*7* (4), 1758-1765. https://doi.org/10.30865/mib.v7i4.6841

Prasetya, F. D., Nugroho, H. W., & Triloka, J. (2022). Data Mining Analysis for Hepatitis C Disease Prediction Using Decision Tree C.45 Algorithm with Particle Swarm Optimization. *National Seminar on Research Results and Community Service*, *April 1989*, 199-209. http://archive.ic

Putra, B., & Muhammad, A. H. (2024). Prediction of Stunting Prevalence in Indonesia with Ordinary Least Square (OLS). *G-Tech: Journal of Applied Technology*,*8* (3), 1890-1900. https://doi.org/10.33379/gtech.v8i3.4623

Rohman, R. S., Saputra, R. A., & Firmansaha, D. A. (2020). Comparison of pso and ga-based c4.5 algorithm for stroke disease diagnosis. *CESS (Journal of Computer Engineering System and Science)*, *5*(1), 155–161.

Saleh, H. (2020). Analysis of Factors Causing Stunting Using the C4.5 Algorithm. *ScientiCO: Computer Science and Informatics*, *3*(1), 11–17.

Sinaga, A. S., Ramen, S., & Mulyani, S. (2024). Predicting the Success of Stunting Handling Using PSO Feature Selection with SaaS Cloud Computing. *Journal of SAINTIKOM (Journal of Informatics and Computer Management Science)*,*23* (1), 87. https://doi.org/10.53513/jis.v23i1.9561

Sulistiyanto. (2018). Application of C4. 5 Based Particle Swarm Optimization (PSO) in Predicting Students Passing College Selection. *Techno Nusa Mandiri*,*7* (2), 162-170. eprints.dinus.ac.id/16925/1/journal_16115.pdf%0Ahttps://www.ojs.amikom.ac.id/index.php/semnasteknomedia/article/view/908

Xsanal Hakim, R., Putrawansyah, F., & Syahri, R. (2024). Application of C4.5 Algorithm for Prediction of Stunting Children in Pagar Alam City. *JATI (Journal of Informatics Engineering Students)*,*8* (2), 2469-2478. https://doi.org/10.36040/jati.v8i2.9301