Ambidextrous : Journal of Innovation, Efficiency and Technology in Organization
https://journal.takaza.id/index.php/ambidextrous

Vol. XX, No. XX, 2023, pp. 122-127
E-ISSN: 3031-7002

E-mail: ambidex@takaza.id

Classification of Symptoms of Disease in Early Childhood Using the Decision Tree Algorithm

Nissa Albantaniyah[1], Dede Brahma Arianto[2]
[12]Faletehan University
Corresponding Author e-mail: nissaalbantaniyah@gmail.com

Article History:
Received: 27-04-2026
Revised: 12-05-2026
Accepted: 12-05-2026

Keywords: *disease classification, early childhood symptoms, decision tree, data mining*

Abstract: *Early childhood is highly vulnerable to infectious diseases due to immature immune systems, while similar symptoms often hinder early diagnosis. This study aims to classify disease symptoms in early childhood using the Decision Tree algorithm. The research employs a quantitative approach with a data mining method. The population consists of early childhood disease cases, with samples derived from primary data through interviews with healthcare workers and secondary data from medical records and literature. Data were collected using documentation and interview instruments. Data analysis involved preprocessing, model training and testing, and evaluation using accuracy, precision, recall, and F1-score based on a confusion matrix. The results show that the Decision Tree model is able to classify disease symptoms effectively, with good performance across all evaluation metrics, indicating reliable predictive capability.The conclusion indicates that the Decision Tree algorithm is appropriate for early disease detection systems, as it provides interpretable, systematic classification and supports faster clinical decision-making.*

How to Cite: Nissa albantaniyah, Dede Brahma Arianto(2026). *Classification of Symptoms of Disease in Early Childhood Using the Decision Tree Algorithm*.*4*(2). Pp.122-127 https://doi.org/10.61536/ambidextrous.v4i2.494

https://doi.org/10.61536/ambidextrous.v4i2.494

Introduction

Early childhood (0–5 years) is the most vulnerable group to various diseases because their immune systems are still developing and have not been able to fight various infections effectively. Diseases that often affect early childhood include upper respiratory tract infections (ARI), diarrhea, fever, and digestive and skin diseases. Delays in detecting symptoms of the disease at an early stage can result in serious complications or require more intensive medical care. (World Health Organization, 2013).

One of the causes of this delay is the lack of knowledge of parents or caregivers in recognizing the early symptoms of the disease. Many parents are only aware of their child's illness when the condition is severe. This shows the need for a system that can help parents in detecting the symptoms of the disease early so that prevention and treatment steps can be carried out faster. (Ministry of Health of the Republic of Indonesia., 2018).

With the rapid development of information technology, especially mobile devices, there is an opportunity to develop applications that utilize children's health data in real-time. Mobile applications can provide quick and practical access for parents to monitor their children's condition anytime and anywhere. (Kumar, S., et al, 2015).

To process data on children's disease symptoms, a data mining method is needed that is able to perform automatic and accurate classification. One of the popular and easy-to-implement algorithms is Decision Tree. This algorithm can build a decision tree based on the entered symptom data, resulting in a quick and easy-to-understand classification of disease types. With this approach, parents or early medical personnel can obtain predictive information about the possibility of diseases experienced by children. (Kaur, H., & Wasan, S. K., 2016).

Research Methods

This study uses a technology-based empowerment approach with a focus on developing a mobile application for the classification of disease symptoms in early childhood. The developed system involves the active participation of parents or caregivers from the planning, testing, to evaluation stages. With this approach, the application is expected to be used independently by users after the research is completed.

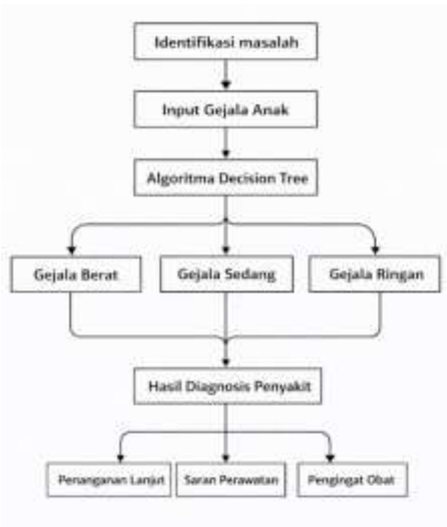

Figure 1. Research Methods

The process of this research is described in the form of a flowchart as shown in Figure 1. This flowchart describes the flow of the process of classifying disease symptoms in early childhood using the Decision Tree algorithm. The stages shown include problem

https://doi.org/10.61536/ambidextrous.v4i2.494

identification, symptom data collection, design and application of the Decision Tree model, and evaluation of classification results to produce diagnosis and treatment recommendations.

Result and Discussion

Data Collection

Data collection was carried out systematically to obtain data on symptoms and types of diseases in early childhood The data sources in this study came from primary data and secondary data. Primary data was obtained through interviews with health workers at health centers or child health clinics to find out the types of diseases and symptoms that are commonly experienced in early childhood. Secondary data were obtained from medical literature, scientific journals, and documentation of medical records that have been anonymized without including the patient's identity.

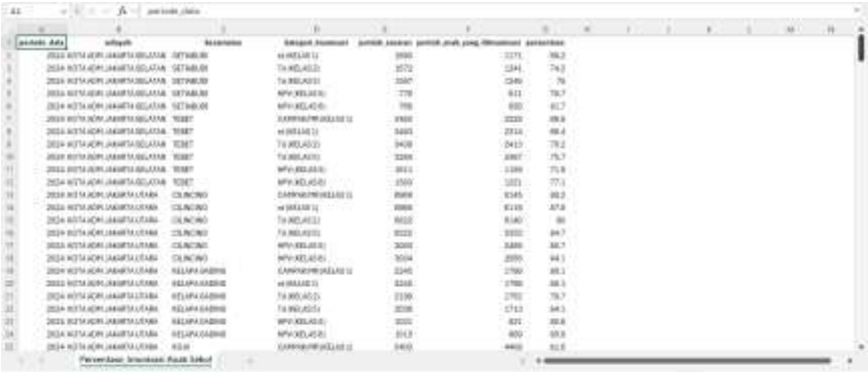

Figure 2 Child dataset

Modeling Design and Training

The Decision Tree algorithm was used to build a model for the classification of disease symptoms in early childhood. The datasets that had gone through the pre-processing stage were divided into training data and test data. The training data was used to train the model to be able to study the pattern of relationship between the symptoms experienced by the child and the type of disease classified.

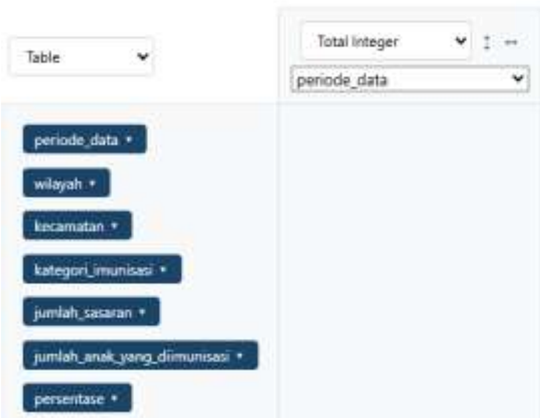

Figure 3 Planning

https://doi.org/10.61536/ambidextrous.v4i2.494

The training process is carried out by forming a decision tree structure that represents classification rules based on symptom attributes, so that the model is able to produce classification decisions in a systematic and easy-to-understand manner. Mobile applications are designed with one main actor which is user (orang tua) with the following features:

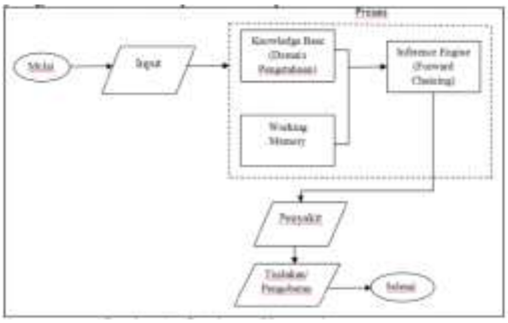

Figure 4 Modeling

Model Evaluation

The model performance evaluation was carried out to find out how well the Decision Tree algorithm classifies disease symptoms in early childhood. The model assessment was carried out using several evaluation metrics, namely accuracy, precision, recall, and F1-score. These metrics were calculated based on a confusion matrix consisting of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

1. Accuracy

   Accuracy indicates the level of accuracy the model has in classifying all early childhood disease symptom data, both correctly detected as sick and correctly detected as not sick. The accuracy value provides an overview of how well the model is performing overall. However, accuracy can be less representative if the data is unbalanced (e.g., the amount of healthy data is much higher than the sick data), because the model can appear "accurate" just by guessing the majority class.

   Accuracy=TP+TNTP+TN+FP+FNAccuracy=\frac{TP+TN}{TP+TN+FP+FN}Accuracy=TP+TN+FP+FNTP+TN

   Description:

   1. TP (True Positive): The detected illness data is correct

   2. TN (True Negative): Detected non-sick data is correct

   3. FP (False Positive): Data are not sick but detected sick

   4. FN (False Negative): Sick but undetected data

Precision

Precision indicates the degree of accuracy of the model in predicting a disease. That is, of all the "positive" predictions (e.g. a child is detected sick), how many are actually sick. Precision is especially important in cases where errors in positive predictions must be

https://doi.org/10.61536/ambidextrous.v4i2.494

minimized, for example so that incorrect diagnoses of the disease do not occur in the child.

Precision=TPTP+FPPrecision=\frac{TP}{TP+FP}Precision=TP+FPTP

The higher the precision value, the fewer errors in giving a positive diagnosis.

Recall (Sensitivity)

Recall shows the model's ability to detect all cases of disease that actually occurs. This means that of all children who are really sick, how many are successfully recognized by the system. Recall is very important in the health field because the mistake of not detecting the disease (false negative) can be fatal.

Recall=TPTP+FNRecall=\frac{TP}{TP + FN}Recall=TP+FNTP

A high recall value means that the system is able to catch most cases of the disease that exists.

F1 Score

F1-score is a harmonious average value between precision and recall. It is used to measure the balance between the two, especially when there is a data imbalance. F1-score will be high only if precision and recall are both high, so this metric is suitable for evaluating model performance more fairly.

F1-score=2×Precision×RecallPrecision+RecallF1\text{-}score = \frac{2 \times Precision \times Recall}{Precision + Recall}F1-score=Precision+Recall2×Precision×Recall

The F1-score is particularly useful in disease diagnosis systems because it considers two important aspects at once: the accuracy of diagnosis and the ability to detect disease.

Conclusion and Recommendation

Based on the results of the study, it can be concluded that the Decision Tree algorithm can be effectively used to classify disease symptoms in early childhood with satisfactory performance. The developed model is capable of identifying patterns and relationships between symptoms and specific types of diseases in a systematic manner while remaining simple and easy to interpret, which is one of the main advantages of the Decision Tree method. Furthermore, the evaluation results based on accuracy, precision, recall, and F1-score metrics indicate that the model performs fairly well in the classification process and is able to provide reliable predictions. Therefore, this system has strong potential to assist the early diagnosis of diseases in young children more quickly, improve efficiency in health services, and support decision-making processes in the healthcare sector.

https://doi.org/10.61536/ambidextrous.v4i2.494

References

Anzullah, M., & Saputri, R. S. (2019). Application of decision tree algorithm in symptom-based disease classification system. Journal of Information Technology and Computer Science, 6(3), 321–328.

Arifin, Z., & Setiawan, D. (2020). Decision support system for the diagnosis of children's diseases using the decision tree method. Journal of Health Informatics, 8(2), 115–123.

Han, J., Kamber, M., & Pei, J. (2018). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.

Kurniawan, A., & Prasetyo, E. (2021). Implementation of data mining for disease classification using a decision tree algorithm. Journal of Information Systems, 17(1), 45–54.

Nugroho, Y., & Lestari, S. (2019). Mobile-based disease diagnosis application using the decision tree method. Journal of Informatics and Computers, 4(2), 89–97.

Putri, R. A., & Wibowo, A. (2022). Symptom-based analysis of child disease classification using a decision tree algorithm. Journal of Information Technology and Systems, 10(1), 1–9.

Quinlan, J. R. (2019). C4.5: Programs for Machine Learning. Morgan Kaufmann.

Sari, M., & Hidayat, R. (2023). Mobile application-based early childhood disease classification system. Journal of Computer Science and Applications, 11(2), 134–142.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann.

Yuliana, D., & Santoso, B. (2020). Application of decision tree algorithms for symptom-based disease diagnosis. Journal of Systems Engineering and Information Technology, 5(1), 23–31.

https://doi.org/10.61536/ambidextrous.v4i2.494