Journal of Engineering, Electrical and Informatics
Vol.3, No.2 Juni 2023
e-ISSN: 2809-8706; p-ISSN : 2810-055, page 01-16
DOI: https://doi.org/10.55606/jeei.v3i1.1468

Enhancing Special Needs Identification for Children: A Comparative Study on Classification Methods Using ID3 Algorithm and Alternative Approaches

Fathul Hafidh
Department of Informatics Engineering, Faculty of Information Technology,
Islamic University of Kalimantan Muhammad Arsyad Al Banjari Banjarmasin, South Kalimantan Indonesia

Address: Jl. Adhyaksa, Jl. Kayu Tangi 1 Jalur 2 No.2, Sungai Miai, North Banjarmasin District, Banjarmasin City, South Kalimantan 70123
Author's correspondence: *hafidh@uniska-bjm.ac.id*

***Abstract**. This research aims to compare the performance of classification methods in identifying special needs in children. The dataset used consists of identifications of various types of special needs, such as ADHD, autism, mild cerebral palsy, mild intellectual disability, moderate intellectual disability, and hearing impairment. The methods compared include ID3 (previous study), Naive Bayes, Random Forest, k-NN, and Gradient Boosting. The comparison results show that ID3 achieves an accuracy rate of 91.81%. The new alternative methods show better performance, with Naive Bayes achieving an accuracy of 95.28%, Random Forest 95.14%, k-NN 95.28%, and Gradient Boosting 83.47%. Although Random Forest does not outperform Naive Bayes and k-NN, it has the advantage of forming decision trees that align with symptom attributes and predict disability labels. However, in the implementation of the Gradient Boosting algorithm, there is a low model probability, especially in identifying ADHD. The conclusion of this research provides insights for researchers in selecting appropriate classification methods for identifying special needs in children, considering accuracy, efficiency, and handling imbalanced data.*

***Keywords**: Special Needs Children, Identification, Classification Methods, ID3, Naïve Bayes, Random Forest, k-NN, Gradient Boosting.*

BACKGROUND

The significance of identifying special needs in Children with Special Needs (CSN) is crucial in providing appropriate attention and support for children with specific requirements. The process of identifying special needs serves as a vital initial step in determining tailored learning programs (Haryono, Syaifudin, & Widiastuti, 2015). In the context of identifying special needs in CSN, classification methods are employed to categorize children into appropriate disability categories based on their symptoms or characteristics. One commonly used method is the ID3 Algorithm, which generates decision trees for data classification (Hafidh, 2021).

Received Maret 20, 2023; Revised April 22, 2023; Accepted Mei 26, 2023
*Fathul Hafidh, hafidh@uniska-bjm.ac.id

However, in the effort to enhance the accuracy of special needs identification, it is essential to compare the performance of classification methods with alternative approaches. The alternative methods that can be evaluated include Naïve Bayes, Random Forest, k-NN, and Gradient Boosting. This comparison is crucial to understand whether there are other classification methods that can achieve better accuracy than ID3 (Nalatissifa, Gata, Diantika, & Nisa, 2021).

Through this case study, the performance of classification methods in identifying specific types of disabilities, such as ADHD, Autism, Hearing Impairment, and Motor Impairment, can be evaluated. Previous studies have mainly focused on the ID3 algorithm, leaving a gap in the exploration of alternative methods and their performance. This research fills this gap by investigating the effectiveness of Naive Bayes, Random Forest, k-NN, and Gradient Boosting in classifying disabilities. Accurate identification of disabilities early on is crucial to ensure appropriate interventions and support for CSN. Furthermore, a comprehensive comparison of multiple classification algorithms yields accuracy and adequacy in handling imbalanced data in the context of disability identification. These findings will contribute to the development of decision-making processes and improve efficiency in identifying special needs in CSN (Anam & Rusdiana, 2020).

The alternative methods under investigation, namely Naive Bayes, Random Forest, k-NN, and Gradient Boosting, exhibit higher classification accuracy compared to ID3 on an imbalanced dataset. Naive Bayes achieves an accuracy of 95.28%, followed by k-NN with the same accuracy, and Random Forest with an accuracy of 95.14%. Notably, Random Forest, although achieving comparable accuracy to Naive Bayes and k-NN, possesses the advantage of modeling with diverse decision trees tailored to the symptom attributes. However, Gradient Boosting shows a lower accuracy rate of 83.47%. This research highlights the significance of comparing classification methods for the identification of special needs and emphasizes the novelty of exploring alternative approaches beyond ID3.

THEORETICAL BACKGROUND

Related Studies

According to (Vishal, Singh, Jinila,, Shyry, & Jabez, 2022) Autism Spectrum Disorder (ASD) is a neurological condition that affects an individual's mental, social, and physical well-being. Classical approaches to identifying autism in individuals are time-consuming and costly.

Data mining approaches have paved the way for intelligent diagnosis. This research focuses on identifying specific features that aid in automating the diagnosis process and conducting a comparative analysis of various machine learning algorithms such as K-Nearest Neighbour, Logistic Regression, SVM, and Naïve Bayes to predict the occurrence of autism spectrum disorders. Experimental analysis demonstrates that the Naïve Bayes algorithm achieves superior accuracy of 99.6% compared to other algorithms.

According to (Sethu, Navya, & Vyas, 2020) this research concludes that ADHD is a neurodevelopmental disorder that affects the social and personal characteristics of children. Diagnosis of ADHD involves various assessment scales and MRI. In this study, SVM and ANN have been found to provide accurate diagnoses, and the use of genetic programming-based algorithms produces better prediction models. These findings will assist researchers in developing improved treatments for children with neurodevelopmental disorders such as ADHD.

Meanwhile, according to (Nalatissifa, Gata, Diantika, & Nisa, 2021) this study aims to predict workplace absenteeism based on the Absenteeism at work dataset using the Weka 3.8 application and the Naïve Bayes, SVM, and Random Forest algorithms. The research results show that the Random Forest algorithm achieves the highest accuracy, precision, and recall compared to Naïve Bayes and SVM, with an accuracy of 99.38%, precision of 99.42%, and recall of 99.39%. This indicates that the Random Forest algorithm can be used as an effective prediction method for addressing workplace absenteeism issues.

Children with special needs (CSN)

CSN refers to children who have limitations in physical or psychological abilities. They require special attention and interventions due to the disorders or developmental abnormalities they experience. Common characteristics of CSN include delayed growth and development, such as toddlers who start walking at the age of three, and specific behavioral deficiencies, such as speech impairment.

Several terms are used to describe CSN, including disabilities, disorders, and limitations. According to the World Health Organization (WHO), disability refers to limitations or lack of ability to perform activities according to rules or within normal limits, usually used at an individual level. Disorders can occur with the loss or abnormality in psychological aspects, anatomical structures, or their functions. Whereas limitations are the disadvantages faced by individuals due to impairments or disabilities that restrict or hinder the fulfillment of normal roles (Federici, Bracalenti, Meloni, & Luciano, 2017).

CSN is a group that requires special attention and support in their education and daily environment. It is important for society to understand and support children with special needs so that they can grow and develop according to their potential (Nurfadhillah, 2021).

Early detection of Children with Special Needs (CSN)

Early detection in CSN is an important initial step in gathering relevant information with the aim of addressing related issues. In early detection, we observe the physical and psychological development of children with the goal of providing appropriate treatment and timely interventions. Through effective early detection, we can identify early signs of limitations or disorders in CSN, allowing them to receive appropriate support and adaptive learning programs.

Early detection differs from assessment, which involves a more in-depth evaluation process. Early detection focuses on initial observations of possible limitations in children, while assessment involves a more detailed assessment of the characteristics and individual needs of children with special needs. Through a comprehensive and responsive early detection approach, we can provide the necessary attention and support for CSN to develop optimally and overcome the challenges they face (Nuryati, 2022).

Early detection of Children with Special Needs is intended as an effort by individuals (teachers) to screen children who experience abnormalities/deviations as early as possible, in order to provide appropriate educational services and prevent learning problems (Mursanib, 2014).

Algoritma-algoritma Klasifikasi

The ID3 (Iterative Dichotomiser 3) algorithm is a machine learning algorithm used for tree-based data classification. This algorithm works by identifying the most informative attribute in separating data into different classes. ID3 measures the importance of attributes using entropy, which quantifies the randomness or disorder in the data. By making decisions based on attributes with the lowest entropy, ID3 constructs a decision tree that can be used to classify new data based on relevant attributes. The ID3 algorithm is particularly useful in rule-based decision making and can be applied in various fields, including data mining, artificial intelligence, and data analysis (Bhatt, Mehta, & D'mello, 2015).

The Naive Bayes algorithm is a machine learning algorithm used for data classification. This algorithm is based on Bayes' Theorem and assumes that each feature in the data is independent of one another. In other words, the Naive Bayes algorithm assumes no correlation between the existing features. During the classification process, the Naive Bayes algorithm calculates the probability of each class based on the features present and then selects the class with the highest probability as the prediction. The Naive Bayes algorithm has a simple and fast implementation and performs well in classifying data with many features. Due to these advantages, the Naive Bayes algorithm is often used in various applications, such as text classification, spam filtering, and recommendation systems (Jiang, Zhang, Li, & Wu, 2018).

The Random Forest algorithm is a classification method that utilizes an ensemble (collection) of decision trees. In this algorithm, multiple decision trees are built randomly, and the combination of prediction results from each tree is used to generate the final prediction. The model-building process starts by dividing the training data into random subsets and constructing a decision tree for each subset. Each decision tree is built using a different subset and by randomly selecting attributes at each node split (Agustiani, Arifin, Junaidi, Wildah, & Mustopa, 2022).

The k-NN (k-Nearest Neighbors) algorithm is a classification method based on the principle of data proximity. In this algorithm, the classification of a new data point is done by finding the k nearest data points to that data in feature space. The value of k is a predefined parameter. Then, the majority class among the k nearest data points becomes the prediction for the new data point (Kurnia, Kurniawan, Fahmi, & Monalisa, 2019).

The classification process in the k-NN algorithm can be performed using distance metrics such as Euclidean distance or Manhattan distance to measure the proximity between data points. Data points with closer distances will have a higher likelihood of belonging to the same class. In the case of multi-class classification, if there are multiple classes that dominate the k nearest data points, majority voting is usually used to determine the predicted class.

The Gradient Boosting algorithm is an ensemble learning method used for predictive modeling. Essentially, this algorithm combines several weak prediction models, such as simple decision trees, to form a strong predictive model. The main process in the Gradient Boosting algorithm involves sequentially building new models that focus on reducing the prediction errors from the previous models (Ismanto & Novalia, 2021).

RESEARCH METHODS

The research method employed in this study is quantitative research with a confirmatory approach. This approach aims to test hypotheses and validate existing theories through systematic data collection and analysis (Wahid, 2004). The study follows several stages:

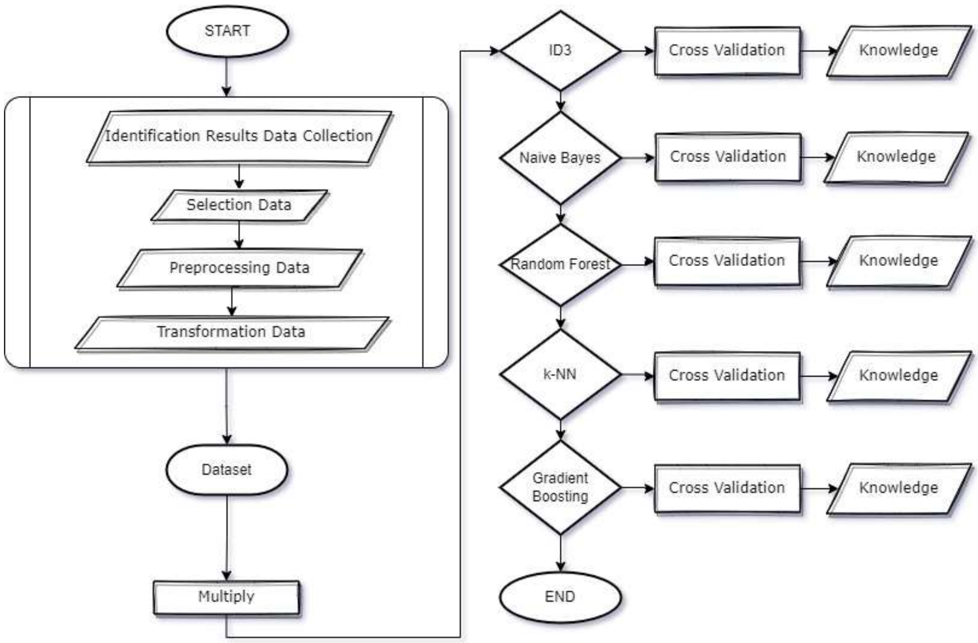

Figure 1. Research Methodology

1. Data collection stage

The dataset used consists of a collection of data obtained from the identification of children with special needs. The data then undergoes selection, preprocessing, and transformation stages to prepare it in a suitable format for analysis.

2. Data multiplication stage.

Data multiplication is performed to enable one data point to be accessed by multiple classification algorithms that will be applied in the study. This facilitates the comparison and evaluation of the performance of various classification methods.

3. Implementation of classification methods

The implementation of classification methods such as ID3, Naive Bayes, Random Forest, k-NN, and Gradient Boosting. Each method is applied to the prepared dataset. The ID3 method is used to build a decision tree based on the attributes in the dataset. The Naive Bayes method performs probabilistic classification assuming attribute independence. The Random Forest method utilizes ensemble decision trees to generate accurate predictions. The k-NN method classifies based on the nearest neighbors using distance. Meanwhile, the Gradient Boosting method combines multiple weak models to form a stronger model.

4. Comparative analysis

Finally, a comparative analysis is conducted among the applied methods. This comparison includes evaluating the performance and accuracy of each method in identifying children with special needs. The results of this comparison provide insights into which method performs best and can serve as a basis for selecting the optimal classification method in identifying children with special needs.

RESULT AND DISCUSSION

Dataset

The data collection process yielded a dataset of symptom instrument data for children with special needs. The data included labels assigned by experts indicating the type of disability. Preprocessing was performed to remove noise and repair any corrupted or disruptive data.

A total of 84 instrument data samples were collected and processed using data mining methods. The processed data for data mining consisted of various types of disabilities, such as ADHD Hyperactive, Autism, Mild Cerebral Palsy, Mild Intellectual Disability, Moderate Intellectual Disability, and Hearing Impairment. The following is an example of the processed source data sample:

Table 1. Sampel data kajian terdahulu

Disability	G031	G032	G033	G034	G035	G036	G037	G038	G039	G040	G041
Autism	No	No	No	No	No	No	No	No	No	No	No
Autism	No	No	No	No	No	No	No	No	No	No	No
Autism	No	No	No	No	No	No	No	No	No	No	No
Autism	No	No	No	No	No	No	No	No	No	No	No
Mild Cerebral Palsy	No	No	Yes	No	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	Yes	No	Yes	No	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No

Mild Cerebral Palsy	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Mild Cerebral Palsy	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Mild Intellectual Disability	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Mild Intellectual Disability	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Mild Intellectual Disability	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Mild Intellectual Disability	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Mild Intellectual Disability	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes

From the entire sample data, a total of 84 samples were obtained for classification, with 6 samples of ADHD, 18 samples of Autism, 8 samples of mild cerebral palsy, 14 samples of mild intellectual disability, 23 samples of moderate intellectual disability, and 15 samples of hearing impairment. This can be seen in the following graph:

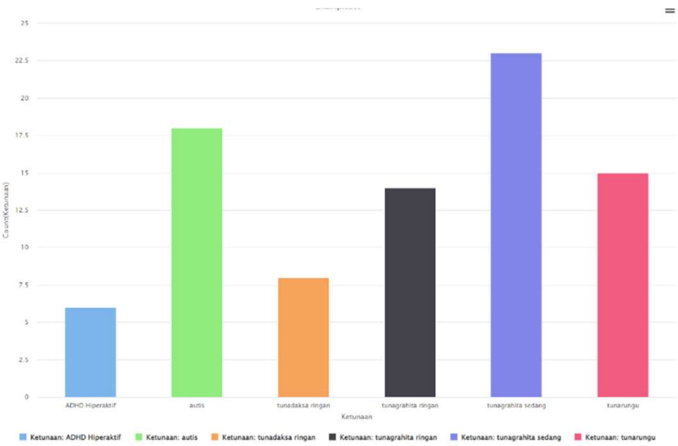

Figure 1. Dataset Graph with Special Needs Label

ID3 Algorithm

The entire dataset was processed using the ID3 algorithm, resulting in a decision tree as follows:

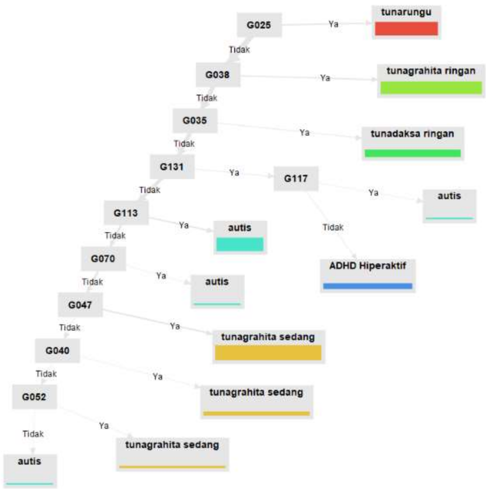

Figure 2. ID3 Algorithm Applied on Previous Study

From the above graph (Figure 2), which represents the results of the previous study the identification process starts by considering the condition G025, which is whether the individual has difficulty hearing clearly. If the answer is "Yes," it can be concluded that the individual is identified as hearing impaired. However, if the answer is "No," the next step is to check the condition G038, which is whether the individual has an IQ range between 50-70 based on WISC. If so, the individual is identified as having mild intellectual disability. But if the answer is "No," the attention shifts to the condition G035, which is whether the individual experiences stiffness, weakness, paralysis, or lethargy in body movements. If this is the case, the individual is identified as having mild cerebral palsy.

Algoritma Naïve Bayes

Table 2. Naïve Bayes Algorithm Results

No	Label	Count	Fraction	Prediction	Prediction Fraction
1	Moderate Mental Retardation	23	0.274	25	0.298
2	Autism	18	0.214	15	0.179
3	Mild Mental Retardation	15	0.179	14	0.167
4	Deafness	14	0.167	14	0.167
5	Mild Physical Disability	8	0.095	9	0.107
6	ADHD Hyperactive	6	0.071	7	0.083

From the above table, it can be observed that the application of the Naïve Bayes algorithm has nearly perfect model probabilities for each category. However, there is a slight difference in the fraction for the "Autism" label, where the predicted fraction is 0.179 instead of the expected 0.214. The following graph illustrates the difference between the labels and the predicted results:

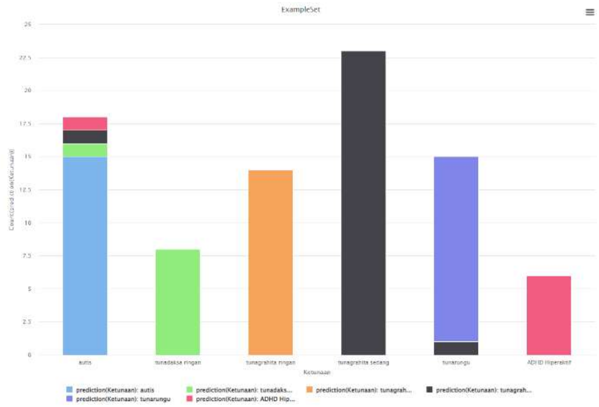

Figure 3. Naïve Bayes Prediction Results

Random Forest Algorithm

The application of the Random Forest method begins with testing 100 decision trees. Eventually, the optimal number of trees is determined to be 20, which yields the shortest decision tree with 2 attributes and 3 labels. Any number of trees beyond this would result in less effective decision trees, such as having only 1 symptom attribute but multiple labels.

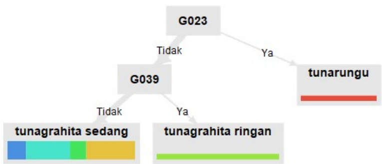

Figure 4. Shortest Decision Tree from Random Forest Method

On the other hand, the longest decision tree is obtained from the second tree, which has 9 attributes and 10 labels, as shown in the following image:

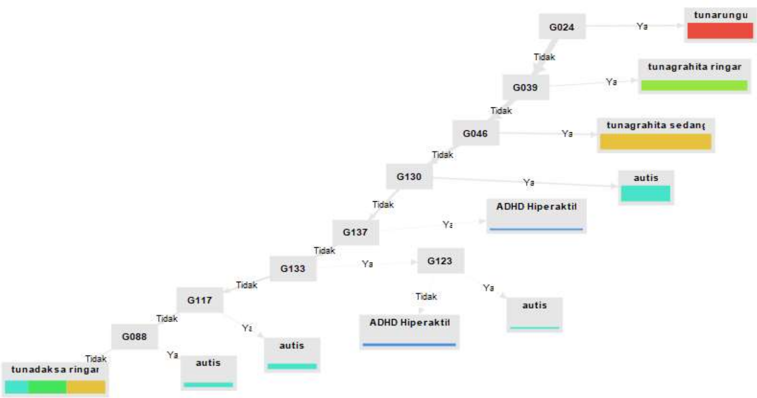

Figure 5. Longest Decision Tree from Random Forest Method

Table 3. Hasil Algoritma Random Forest

No	Label	Count	Fraction	Prediction	Prediction Fraction
1	Moderate Mental Retardation	23	0.274	25	0.298
2	Autism	18	0.214	15	0.179
3	Mild Mental Retardation	15	0.179	15	0.179
4	Deafness	14	0.167	14	0.167
5	Mild Physical Disability	8	0.095	8	0.095
6	ADHD Hyperactive	6	0.071	7	0.083

From the above table, it can be observed that the application of the Random Forest algorithm has nearly perfect model probabilities for each category. However, similar to the Naïve Bayes results, there is a slight difference in the fraction for the "Autism" label, where the predicted fraction is 0.179 instead of the expected 0.214. The following graph illustrates the difference between the labels and the predicted results:

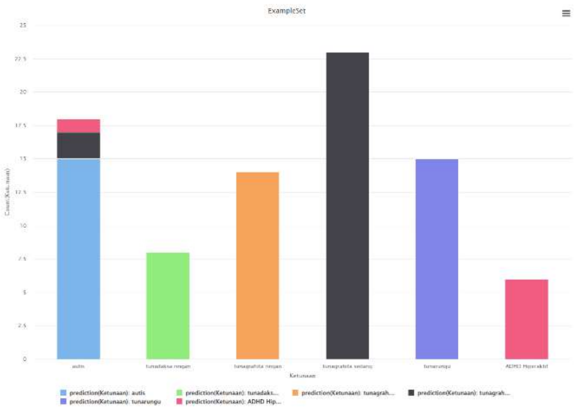

Figure 6. Random Forest Prediction Results

k-NN Algorithm

Table 4. k-NN Algorithm Results

No	Label	Count	Fraction	Prediction	Prediction Fraction
1	Moderate Mental Retardation	23	0.274	25	0.298
2	Autism	18	0.214	15	0.179
3	Mild Mental Retardation	15	0.179	14	0.167
4	Deafness	14	0.167	14	0.167
5	Mild Physical Disability	8	0.095	9	0.107
6	ADHD Hyperactive	6	0.071	7	0.083

From the above table, it can be observed that the application of the k-NN algorithm has the same model probabilities for each category as the Naïve Bayes method. However, similar to the Naïve Bayes and Random Forest results, there is a significant difference in the fraction for the "Autism" label, where the predicted fraction is 0.179 instead of the expected 0.214. The following graph illustrates the difference between the labels and the predicted results:

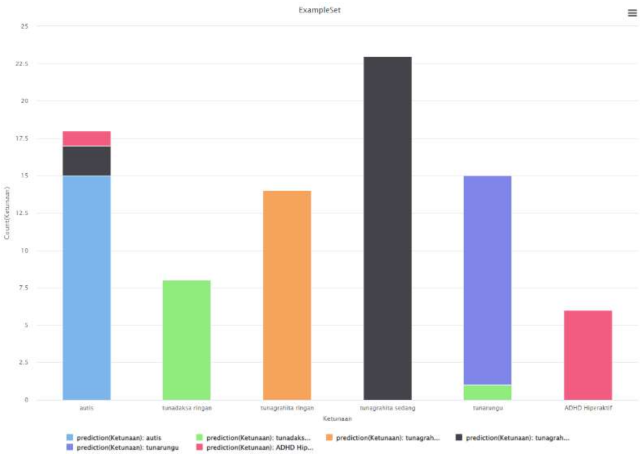

Figure 7. k-NN Prediction Results

Algoritma Gradient Boosting

Table 5. Gradient Boosting Algorithm Results

No	Label	Count	Fraction	Prediction	Prediction Fraction
1	Moderate Mental Retardation	23	0.274	22	0.262
2	Autism	18	0.214	20	0.238
3	Mild Mental Retardation	15	0.179	14	0.167
4	Deafness	14	0.167	14	0.167
5	Mild Physical Disability	8	0.095	14	0.107
6	ADHD Hyperactive	6	0.071	0	-

From the above table, it can be observed that the application of the Gradient Boosting algorithm does not have optimal model probabilities. The prediction results fail to identify any instances of ADHD, which should have 6 labels in the dataset. The following graph illustrates the difference between the labels and the predicted results:

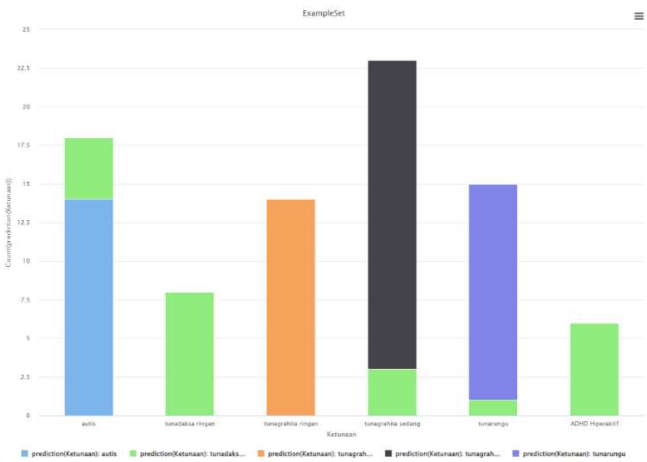

Figure 8. Gradient Boosting Prediction Results

Hasil Pengujian

Model data mining has been tested for its accuracy using cross-validation, which is a performance evaluation method involving the division of the dataset into segments for training and testing. The 10-fold cross-validation method divides the dataset into 10 equally sized segments, where each segment is used alternately as the testing data in 10 iterations of training and testing processes (Kurniawan & Rosadi, 2017). The testing results can be seen in the image below:

Table 6. Accuracy Comparison Table

Method	Accuracy	Advantages	Disadvantages
ID3	91.81%	- Easy to interpret - Efficient for small datasets	- Prone to overfitting - Does not handle continuous attributes
Naive Bayes	95.28%	- Simple and efisien - Handles continuous attributes	- Assumes feature independence, which is not always met
Random Forest	95.14%	- Handles overfitting - Robust against imbalanced data - Produces predictions with variations from multiple decision trees	- Requires more time and resources for training and prediction
k-NN	95.28%	- Does not assume data distribution - Handles continuous attributes	- Prone to outliers - Requires proper parameter tuning
Gradient Boosting	83.47%	- Handles continuous attributes - Handles imbalanced data	- Prone to overfitting - Requires more time for training and prediction

In the above table, it can be observed that each method has its own advantages and disadvantages. ID3 is easy to interpret and efficient for small datasets but is prone to overfitting and does not handle continuous attributes. Naive Bayes is simple and efficient but assumes feature independence, which may not always hold true. Random Forest can handle overfitting and imbalanced data but requires more time and resources. k-NN does not assume data distribution, handles continuous attributes, but is sensitive to outliers and requires proper parameter tuning. Gradient Boosting can handle continuous attributes and imbalanced data but is prone to overfitting and requires more time. When selecting a method, it is important to consider the advantages and disadvantages that align with the research context and objectives.

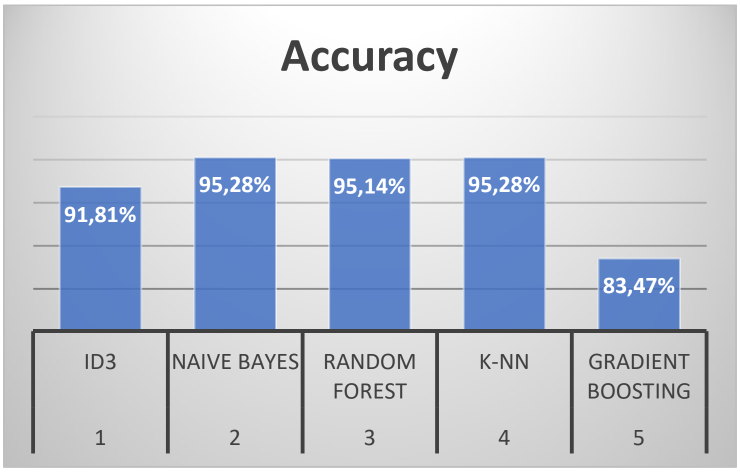

Figure 9. Comparison Graph of Alternative Methods to the Previous Study

CONCLUSION AND RECOMMENDATIONS

Conclusion

Based on the research conducted in this study, the following conclusions can be drawn:

1. The previous study's classification method using the ID3 algorithm resulted in an accuracy rate of 91.81% in identifying children with special needs. This indicates that ID3 can be used as a reasonably good method for classifying the disability of special needs children.
2. The alternative methods investigated, namely Naive Bayes, Random Forest, k-NN, and Gradient Boosting, showed higher accuracy rates compared to ID3. Naive Bayes and k-NN achieved an accuracy of 95.28%, followed by Random Forest with an accuracy of 95.14%. However, Gradient Boosting exhibited a lower accuracy rate of 83.47%.

3. The analysis results indicate that Gradient Boosting has suboptimal model probabilities, as its predictions did not identify any cases of ADHD despite the dataset containing 6 ADHD labels. This highlights a weakness in Gradient Boosting's ability to identify ADHD cases.
4. Random Forest, despite having comparable accuracy to Naive Bayes and k-NN, has the advantage of modeling with multiple decision trees that align with symptom attributes. Additionally, its prediction results provide useful disability labels for identifying children with special needs.

Recommendations

Based on the comparison of classification methods for identifying children with special needs, the following recommendations are given:

1. If accuracy is the primary factor in the research, Naive Bayes, Random Forest, or k-NN methods can be considered. Naive Bayes has the advantage of simplicity and efficiency, while Random Forest can handle overfitting, and k-NN can handle continuous attributes. The choice among these methods should be based on assumptions that align with the data and research objectives.
2. To select the appropriate method, it is important to consider other factors such as model complexity, interpretation of results, and available resources. Furthermore, conducting cross-validation and further experiments can assist in choosing the most suitable method for the research context and objectives.

REFERENCES

Agustiani, S., Arifin, Y. T., Junaidi, A., Wildah, S. K., & Mustopa, A. (2022). Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram. *Jurnal Komputasi, 10*(1), 65-74.

Anam, C., & Rusdiana, N. (2020). Analisis Pemeringkatan Kualitas Klasifier Pada Dataset Tidak Seimbang. *JIMP (Jurnal Informatika Merdeka Pasuruan), 5*(1), 38-44.

Bhatt, H., Mehta, S., & D'mello, L. R. (2015). Use of ID3 decision tree algorithm for placement prediction. *Int J Comput Sci Inform Technology, 6*(5), 4785-4789.

Federici, S., Bracalenti, M., Meloni, F., & Luciano, J. V. (2017, Nov 6). World Health Organization disability assessment schedule 2.0: An international systematic review. *Disability and rehabilitation*, pp. 2347-2380.

Hafidh, F. K. (2021). Identifikasi ketunaan anak berkebutuhan khusus dengan algoritma iterative dichotomiser 3 (id3). *Jurnal Buana Informatika, 12*(2), 1278-87.

Haryono, H., Syaifudin, A., & Widiastuti, S. (2015). Evaluasi pendidikan inklusif bagi anak berkebutuhan khusus (ABK) di Provinsi Jawa Tengah. *Jurnal Penelitian Pendidikan, 32*(2), 119-126.

Ismanto, E., & Novalia, M. (2021). Komparasi Kinerja Algoritma C4. 5, Random Forest, dan Gradient Boosting untuk Klasifikasi Komoditas. *Techno. Com, 20*(3), 400-410.

Jiang, L., Zhang, L., Li, C., & Wu, J. (2018). A correlation-based feature weighting filter for naive bayes. *IEEE transactions on knowledge and data engineering, 31*(2), 201-213.

Kurnia, F., Kurniawan, J., Fahmi, I., & Monalisa, S. (2019). Klasifikasi Keluarga Miskin Menggunakan Metode K-Nearest Neighbor Berbasis Euclidean Distance. *In Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI), 11*, 230-239.

Kurniawan, M. Y., & Rosadi, M. E. (2017). Optimasi Decision Tree Menggunakan Particle Swarm Optimization Pada Data Siswa Putus Sekolah. *Jurnal Teknologi Informasi Universitas Lambung Mangkurat (JTIULM), 2*(1), 7-14.

Mursanib, M. (2014). Meningkatkan keterampilan Identifikasi Anak Berkebutuhan Khusus Bagi Mahasiswa Prodi. PG/PAUD FKIP Universitas Tadulako. *Tri Sentra Jurnal Ilmu Pendidikan, 2*(4), 20-30.

Nalatissifa, H., Gata, W., Diantika, S., & Nisa, K. (2021). Perbandingan Kinerja Algoritma Klasifikasi Naive Bayes, Support Vector Machine (SVM), dan Random Forest untuk Prediksi Ketidakhadiran di Tempat Kerja. *Jurnal Informatika Universitas Pamulang, 5*(4), 578-584.

Nurfadhillah, S. (2021). *Pendidikan Inklusi Pedoman bagi Penyelenggaraan Pendidikan Anak Berkebutuhan Khusus.* Sukabumi: CV Jejak (Jejak Publisher).

Nuryati, N. (2022). *Pendidikan Bagi Anak Berkebutuhan Khusus.* Yogyakarta: Unisa Press.

Sethu, Navya, & Vyas, R. (2020). Overview of machine learning methods in ADHD prediction. *Advances in Bioengineering.*, (pp. 51-71). Singapore.

Vishal, V., Singh, A., J. Y., Shyry, S., & Jabez, J. (2022, April). A Comparative Analysis of Prediction of Autism Spectrum Disorder (ASD) using Machine Learning. *In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, (pp. 1355-1358). IEEE.

Wahid, F. (2004). Metodologi penelitian sistem informasi: sebuah Figurean umum. *Media Informatika, 2*(1), 69-81.