Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 7 No. 5 (2023) 1009 – 1018 ISSN Media Electronic: 2580-0760

Classification of Secondary School Destination for Inclusive Students using Decision Tree Algorithm

Rizal Prabaswara[1], Julianto Lemantara[2]*, Jusak Jusak[3]
[1,2] Information System Department, Faculty of Technology and Informatics, Universitas Dinamika, Surabaya, Indonesia
[3]Computer Engineering Department, Faculty of Technology and Informatics, Universitas Dinamika, Surabaya, Indonesia
[1]18410100131@dinamika.ac.id, [2]julianto@dinamika.ac.id, [3]jusak@dinamika.ac.id

Abstract

Inclusive student education has become one of the most important agendas for UNESCO and the Indonesian government. Developing inclusive children's education is critical to adjust their abilities while attending school. However, most parents and educators who assist students in selecting their future secondary school after finishing primary school are frequently not aware of their genuine ability. The problem is mainly because the decision is not based on objective assessments like IQ, average and mental scores. In this study, we aims to create a school-type decision support system using data mining as a factor analytic approach in extracting rules for the knowledge model. The system uses some variables as the basic principles for building school-type classification rules using the ID3 decision tree method. This system can also assist educators in making decisions based on existing graduate data. Evaluation showed that the proposed system produced an accuracy of 90% by allocating 75% of data for training and 25% for testing. The accuracy value from evaluation phase stated that the ID3 Decision tree algorithm have a good peformance. This system also can dynamically create new decision trees based on newly added datasets. Further research is expected to have more variable and more dynamic system that can have more accurate result for the inclusive student classification of secondary school.

Keywords: *inclusive student; education; decision support system; ID3 algorithm; classification*

1. Introduction

Inclusive education for children with special needs is a specific service provided to students with special needs due to the implementation of Article 31, paragraph 1 of the 1945 Constitution and Law Number 20 of 2003 Concerning the National Education System. In the 2020/2021 school year, the number of children with special needs enrolled in special schools (SLB) reached 144,621. One approach to helping students stay in their special education is a stronger foundation of content knowledge, academic skills, and non-cognitive skills. Unfortunately, most of the schools are not ready to provide adequate infrastructure and environmental conditions for them, and the community surrounding the school cannot create a welcoming environment for these children. There is not even a proper profiling of students in most schools to determine student knowledge from the school database. [1]–[4]. Then, the requirements of a suitable learning environment, as well as teachers who focus on children with special needs in learning and actively invite them to have a more positive influence on inclusive child brain development and cause the child to hone their abilities to the fullest[4]. In other country like Philippines, 265,000 people with disabilities in the Philippines did not continue their education[5]. In China, the government is paying more attention to inclusive children by training more professional teachers/educators. The Chinese government has used this successfully, with around 63.19% of inclusive children receiving higher and further education [5], [6].

Education development for inclusive children is critical in adjusting the child's talents, interests, and abilities while attending school. The classification is subjective from the parent's point of view and does not objectively view the factors that influence the development of their talents. This would cause the student not meet the best education properly[4], [7]. Based on data from the Indonesian Ministry of Education and Culture [8] there has been an increase in inclusive dropout students from the 2017/2018 school year to 2020/2021 in the East Java region where we conduct the research as shown in Figure 1. The reason for the high number of inclusive primary school students who do not continue their education is the lack of parental concern for handling ABK (47.27%). Followed by, parents' lack of

Accepted: 09-05-2023 | Received in revised: 15-06-2023 | Published: 30-08-2023

understanding about ABK (41.21%), parents feel embarrassed so they want their children to go to public schools (3.64%), tolerance from parents of regular students towards ABK is lacking (3.64%), parents are illiterate (2, 42%), parents are impatient with ABK (1.21%), and single parenting (0.61%) [9]. Based on the abovementioned issues, a software system that can provide solutions for the secondary school appropriate for the child is highly required to assist parents and teachers/educators in determining the best decisions for children based on their abilities and talents demonstrated in primary school.

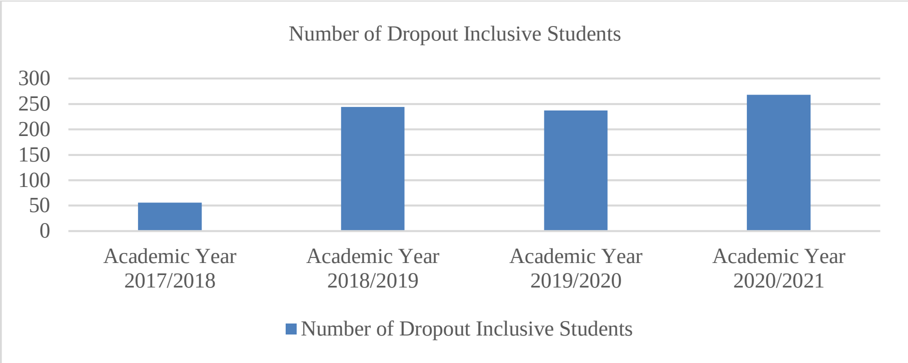

Figure 1. Dropout Inclusive Student

There are several algorithms that can be used for classification, but many studies have shown that the ID3 algorithm has the advantage of being able to provide a high level of accuracy. For the example, research conducted by Mahendra Putra et al. focused on the classification system for beneficiaries of bidikmisi assistance using the ID3 algorithm, with an accuracy rate of 98.3%[10]. Another study using the ID3 algorithm by Fiarni C. et al. to choose an information system sub-major showed an accuracy rate of 70% [11]. Another study by Wiyono S. et al.[12], highlighted comparing the Decision tree, KNN, and SVM for predicting student performance data. It showed that decision tree algorithms performed better, with an accuracy of 92% for predicting the data. The fourth study by Aldino A. and Sulistiani H [13], classified tuition aid grant programs using decision trees have an 87% accuracy of the model knowledge. Meanwhile, Bansod et al. proved that the decision tree and naïve bayes have a high accuracy for predicting the child development compare to the others. Which, ID3 decision tree has more high accuracy in 85% and naïve bayes just 57% [14]. A different research topic conducted by sifaunajah et al. tested the ID3 algorithm in dealing with datasets of prospective students with a predicted value of 82% [15]. Others research conducted by Fadillah Rahman et al. stated that ID3 algorithm have a high accuracy value in 96% to predict recipient of social assistance [16].

Many studies on the child/student development have been conducted in recent years. However, research on the classification of secondary school for inclusive children that utilize machine learning with ID3 algorithm in the form of website application or decision support system has not been found in Indonesia. Several studies by researchers showed that the ID3 algorithm is a reliable and stable method because it can produce research with the highest predictive accuracy of other classification algorithms with a small amount of data. This research compared the validity and accuracy of the proposed system with other methods such as C4.5, Naïve Bayes, KNN (K-Nearest Neighbour), and SVM (Support Vector Machine) algorithms to make sure that the proposed system has high accuracy. Based on that background, researchers developed a decision support system website for assisting inclusive primary school students in deciding their future secondary school level using adaptive ID3 algorithm. This adaptive ID3 algorithm can change rules according to the conditions/data loaded into the application. So, the rules that are created will follow the conditions of the educational environment that can change, especially in terms of determination the secondary school. The study result can be considered an assistive technology to support parents and inclusive students to make the best possible decisions for their future school.

2. Research Methods

To achieve the objectives, we employed research stages as shown in Figure 2. It comprises of six stages including business understanding, data understanding, data preparation, modeling, evaluation, and system deployment.

2.1 Data Identification

The data identification process is collecting data from observations and interviews as well as literature studies. The researcher conducted this data collection stage using data provided by teachers in inclusive schools in Surabaya city. Utilizing this information we label the data into two groups, i.e., SMP/Inclusive School and special school /SLB. The data will be used in the data mining process to build a system knowledge model. The

DOI: https://doi.org/10.29207/resti.v7i5.5081

process of extracting and identifying helpful information and related knowledge from databases by employing statistics, mathematics, artificial intelligence, and machine learning techniques is known as data mining [17]–[19].

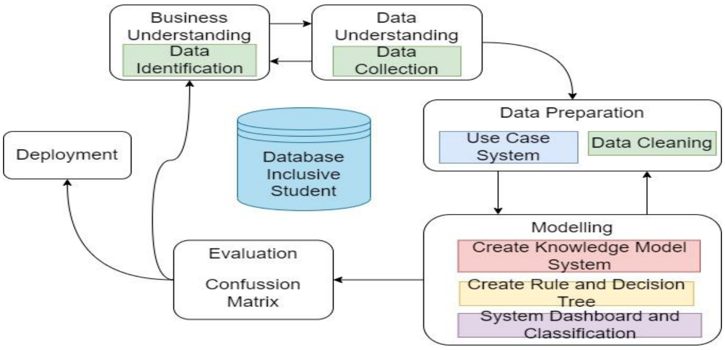

Figure 2. Research Method

The primary goal of data mining is to find patterns in the data in the system so that previously unseen information/patterns can be obtained [19]–[21]. From the data identification activity, 6 attributes contribute to the decision-making for selecting the secondary school, as in **Error! Not a valid bookmark self-reference.**.

Table 1. Data Attributes

No	Variable	Data Type
1	IQ	Numeric
2	Grades	Numeric
3	Active	Category: Bad Good Very Good
4	Discipline	Category: Bad Good Very Good
5	Cooperation	Category: Bad Good Very Good
6	Disability	Category: Physical Mental Academic
7	Type of School	Category: SMP / Inclusive School SLB / Special School

2.2 Data Collection

The researcher conducted this stage of data collection by using data provided by one of inclusive school in Surabaya city and conducting an interview with the Special Guidance Teacher of inclusive student, which includes an explanation which the label of the classification divided into two group which is SMP/Inclusive School and special school /SLB.

The amount of the inclusive student that become our data process is 39 students. From that data there are 9 record data with Special school class type and 30 record with Inclusive school class type. The data is presented in

Table 2.

Table 2. Data Training

No	IQ	Grade	Active	Discipline	Cooperation	Disability	School
1	65	<75	Good	Good	Good	Mental	SMP
2	79	>=75	Very Good	Good	Bad	Academic	SMP
3	65	<75	Bad	Very Good	Bad	Academic	SLB
n	..n	..n	..n	..n	..n	..n	..n
39	69	>=75	Very Good	Good	Bad	Academic	SMP

2.3. Data Cleaning

Data cleaning was performed to improve the quality of the data. In this study, data cleaning included deleting duplicate data, data anomalies, or abnormal data and deleting data attributes that were not relevant in describing student school patterns.

2.4 Modelling

The decision tree method is one of many classification methods for building a predicting system [22]–[24]. The decision tree method converts a large number of facts into a decision tree that represents the rules. The rules that are developed must be simple to understand in everyday language. A decision tree is a structure that can be used to divide large data sets into smaller sets of

DOI: https://doi.org/10.29207/resti.v7i5.5081

records by enforcing a set of decision rules developed during calculations [25]. The attribute specifies the parameters that will be used to create the tree. The decision tree process converts data (tables) into tree models, then tree models into rules, and the final rules are simplified. A decision tree can be built using a variety of algorithms, including ID3, CART, and C4.5. The C4.5 algorithm evolved from the ID3 algorithm [26], [27].

A decision tree is a structure that can help people apply a set of decision rules developed during calculations [28]. The highest attribute value is used during calculations to select an attribute as root [29]. The first formula used to calculate the highest value in root finding is to find the entropy value, as shown in Formula 1.

$$H\ (s) = \sum_{j=1}^{k} - P_j * log2\ P_j \qquad (1)$$

Where H (s) is the entropy of each case set, k is the total number partition of H (s), P_j is the proportion of Si to S, S is the total sample, and Si is the total sample data for specific criteria.

$$P_j = \frac{the\ number\ of\ class\ attribute\ data\ Si}{the\ total\ amount\ of\ attribute\ data}$$

Then calculate Gain value as shown in Formula 2.

$$Gain\ (S, A) = H\ (S) - \sum_{i=1}^{n} P_t * H(t) \qquad (2)$$

Whetre A is the gain of each case set on every attribute, n is the total number partition, H (t) is the entropy of the attribute data.

$$P_t = \frac{the\ total\ amount\ of\ attribute\ data}{the\ total\ amount\ of\ data}$$

There is flowchart that conclude all the steps that our program used to make classification rule from retrieve the data until the decision tree was made. The flowchart of calculate the ID3 can be seen in Figure 3.

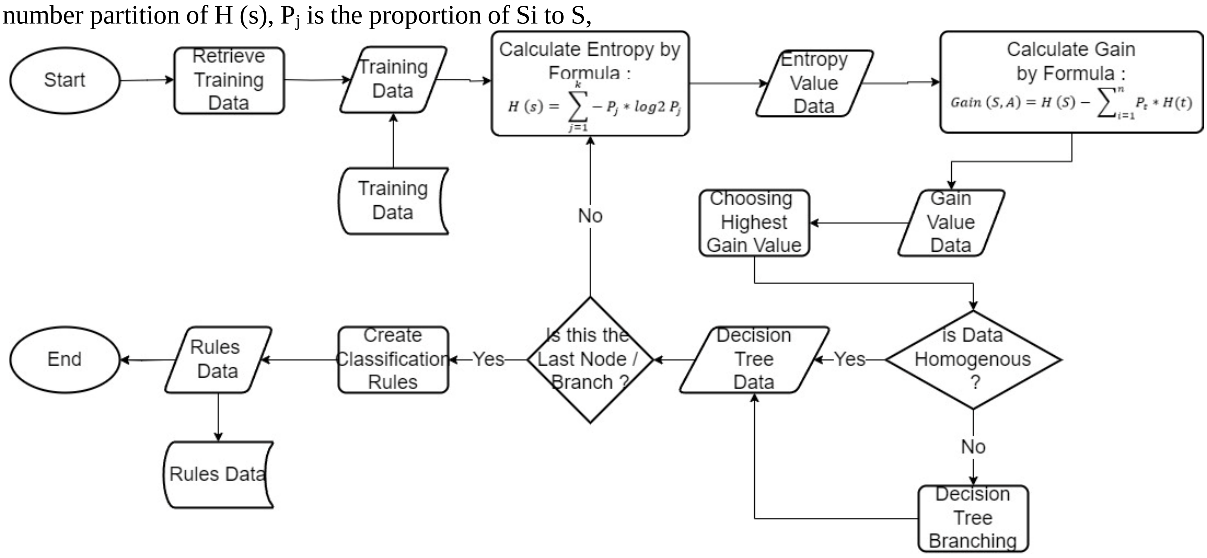

Figure 3. Flowchart Calculate the ID3 Algorithm

2.5 Evaluation

At this point, the researcher used a confusion matrix to evaluate the decision support system's performance in classifying the secondary school for inclusive students. Accuracy, precision, and recall can be used to evaluate the performance of decision support systems [30]. Precision is calculated by dividing the optimistic accurate class predictions by the overall positive class predictions. Then, recall is calculated by dividing the total TP and FN by the actual positive. Finally, accuracy is the ratio of correct and incorrect predictions to total data. The researcher calculated the accuracy of the algorithm used for school classification based on the Confusion Matrix table researcher created in Table 3.

Table 3. Confussion Matrix

Predict	Actual	
	True	False
True	True Positive (TP)	False Negative (FN)
False	False Positive (FP)	True Negative (TN)

The success rate of the predicted value with the actual value is the test method with accuracy. Accounting is the sum of all negative and positive data values divided by the sum of all actual data values. The accuracy formula is Formula 3.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}\ x\ 100\% \qquad (3)$$

Precision is a measure that indicates the extent to which a classification system, under stable conditions, obtains the same results when there are repetitions. Precision is determined using the f precision calculation in Formula 4.

$$Precision = \frac{TP}{TP+FP}\ x\ 100\% \qquad (4)$$

DOI: https://doi.org/10.29207/resti.v7i5.5081

The recall is the test's ability to identify the ratio of optimistic outcome predictions from several truly positive data. The recall is determined using the recall calculation in Formula 5.

$$Recall = \frac{TP}{TP+FN}\ x\ 100\% \qquad (5)$$

2.6 Deployment

During the deployment stage, the evaluated and refined application will be implemented to classify the secondary school for children with special needs from elementary/primary school graduates based on influential attributes.

There will be some interface of the system that researchers will raise to provide an idea of how the application's shape will be.

3. Results and Discussions

3.1 ID3 Method Calculation Process

The following is an example of manual calculations using the same training data as the applications and objects we studied. The amount of initial data is 39 data and a split ratio of 75:25 is carried out, where 75% of all initial data will be used as training. The entropy calculation formula can be seen in Formula 1 and the gain calculation formula in Formula 2. The data that use for calculation can be seen in Table 4.

Table 4. Case Examples Calculation of ID3

No	Attribute	Value	Total Case	SMP / Inclusive Class	SLB / Special Class
1	Total		29	21	8
2	IQ				
		<75	20	12	8
		>=75	9	9	0
3	Grades				
		<75	11	7	4
		>=75	18	14	4
4	Active				
		Bad	11	6	5
		Good	13	10	3
		Very Good	5	5	0
5	Discipline				
		Bad	11	4	7
		Good	12	11	1
		Very Good	6	6	0
6	Cooperation				
		Bad	9	4	5
		Good	13	10	3
		Very Good	7	7	0
7	Disability				
		Pyshical	0	0	0
		Mental	14	9	5
		Academic	15	12	3

Calculation Example

$$\boldsymbol{Total\ Entropy} = (-\left(\frac{21}{29}\right) x\ \log_2 \frac{21}{29}) + (-\left(\frac{8}{29}\right) x\ \log_2 \frac{8}{29}) = \mathbf{0,85}$$

$$\boldsymbol{Entropy\ IQ < 75} = (-\left(\frac{12}{20}\right) x\ \log_2 \frac{12}{20}) + (-\left(\frac{8}{20}\right) x\ \log_2 \frac{8}{20}) = \mathbf{0,971}$$

$$\boldsymbol{Entropy\ IQ \geq 75} = (-\left(\frac{9}{9}\right) x\ \log_2 \frac{9}{9}) + (-\left(\frac{0}{9}\right) x\ \log_2 \frac{0}{9}) = \mathbf{0}$$

$$\boldsymbol{Entropy\ Discipline\ BAD} = (-\left(\frac{4}{11}\right) x\ \log_2 \frac{4}{11}) + (-\left(\frac{7}{11}\right) x\ \log_2 \frac{7}{11}) = \mathbf{0,946}$$

$$\boldsymbol{Entropy\ Discipline\ GOOD} = (-\left(\frac{11}{12}\right) x\ \log_2 \frac{11}{12}) + (-\left(\frac{1}{12}\right) x\ \log_2 \frac{1}{12}) = \mathbf{0,414}$$

$$\boldsymbol{Entropy\ Discipline\ VERY\ GOOD} = (-\left(\frac{6}{6}\right) x\ \log_2 \frac{6}{6}) + (-\left(\frac{0}{6}\right) x\ \log_2 \frac{0}{6}) = \mathbf{0}$$

Calculate Gain

$$Gain\ IQ = (0{,}85 - (\left(\frac{20}{29}\right) x\ 0{,}971) + (\left(\frac{9}{29}\right) x\ 0) = 0{,}18$$

$$Gain\ Discipline = (\ 0{,}85 - \left(\left(\frac{11}{29}\right) x\ 0{,}946\right) + \left(\left(\frac{12}{29}\right) x\ 0{,}414\right) + \left(\left(\frac{6}{29}\right) x\ 0\right)) = 0{,}319$$

From that calculation the discipline attribute will be chosen to become the branch of decision tree. This process will always happen until there are no more case from one of the class. After that basically each of attribute will be have their own branch depending how much the value of the attributes. In this study we limit the value of each attribute just 3 so the rule model knowledge has more accuracy and can represent every type of training data that exists.

3.2 Use Case System

Each user has different interactions in the system. An example function for the parents is adding student competencies and looking for recommendations for the secondary school that follows the values and competencies of their inclusive children. The function for the teachers is the dashboard system that displays information in the form of comparisons between inclusive students with different types of school results, comparisons of each variable value that impact the model knowledge, and the classification status of every inclusive child. The proposed system usecase system explained in this study is shown in Figure 4.

3.3 System Interface

DOI: https://doi.org/10.29207/resti.v7i5.5081

The users of this proposed system are teachers and their parents. They each have some access to the proposed system's features, with the inclusive student teacher having 8 system features and student parents having 3 system features. The proposed system's interface design is the next step in system development. The process of defining how the system will interact with the user so that the user understands how the decision support system works is known as the system interface. The system has a dashboard with some charts describing the classification information. As shown in **Error! Reference source not found.**, this can add information for teachers to know the actual situation in the field.

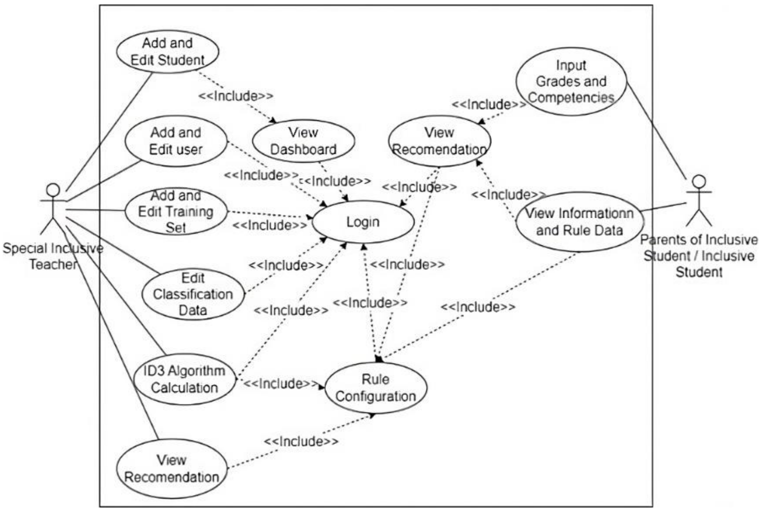

Figure 4. Use Case Diagram

This dashboard contains information, including a bar chart with comparative data between inclusive students that enter SLB/Special Schools and SMP/inclusive schools. If the user selects the detail data button, the system will display the student data, such as names and other information, based on the user's choices.Then there is a bar chart showing every inclusive student's classification status. Finally, six pie charts contain information about comparing each value in the six variables used in the decision support system's knowledge model. The classification process with ID3 begins with an automatic calculation of ID3 by the program we created by taking the training data we have prepared. Admin can just go to the "Perhitungan Decision Tree" menu to start the automatic calculation.

Figure 5. Dashboard System 1

```
(Kedisiplinan='Baik') AND (Kerja_Sama='Sangat Baik') data ini bernilai homogen sehingga menunjukkan bahwa hasil keputusan nantinya adalah SMP
LEAF Keputusan = SMP
================================
cabang (Kedisiplinan='Sangat Baik')

(Kedisiplinan='Sangat Baik') data ini bernilai homogen sehingga menunjukkan bahwa hasil keputusan nantinya adalah SMP
LEAF Keputusan = SMP
================================
cabang (Kedisiplinan='Kurang')

cabang (Kedisiplinan='Kurang') data ini bernilai heterogen sehingga menunjukkan bahwa hasil keputusan nantinya bercabang
0.946Entropy Total
0.764 = Entropy1
0 = Entropy2
Gain IQ = 0.321
0.722 = Entropy1
1 = Entropy2
Gain Rata_rata_Nilai_US = 0.072
1 = entropy 1
0 = entropy 2
0 = Entropy 3
Gain Keaktifan = 0.401
0.722 = entropy 1
0.811 = entropy 2
0 = Entropy 3
Gain Kerja_Sama = 0.323
0.863 = Entropy1
1 = Entropy2
Gain Ketunaan = 0.033
Atribut terpilih = Keaktifan, dengan nilai gain = 0.401

================================
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Baik')
```

Figure 5. The ID3 Calculation Page (a)

```
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Kurang')

(Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Kurang')data ini bernilai homogen sehingga menunjukkan
bahwa hasil keputusan nantinya adalah SLB
LEAF Keputusan = SLB
================================
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Sangat Baik')

(Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Sangat Baik')data ini bernilai homogen sehingga menunjukkan
bahwa hasil keputusan nantinya adalah SMP
LEAF Keputusan = SMP
================================
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Baik')

(Kedisiplinan='Kurang') AND (Keaktifan='Baik') AND (Kerja_Sama='Baik')data ini bernilai homogen sehingga menunjukkan
bahwa hasil keputusan nantinya adalah SMP
LEAF Keputusan = SMP
================================
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Kurang')

(Kedisiplinan='Kurang') AND (Keaktifan='Kurang')data ini bernilai homogen sehingga menunjukkan
bahwa hasil keputusan nantinya adalah SLB
LEAF Keputusan = SLB
================================
cabang (Kedisiplinan='Kurang') AND (Keaktifan='Sangat Baik')

(Kedisiplinan='Kurang') AND (Keaktifan='Sangat Baik')data ini bernilai homogen sehingga menunjukkan
bahwa hasil keputusan nantinya adalah SMP
LEAF Keputusan = SMP
================================
```

Figure 6. The ID3 Calculation Page (b)

The example of the entropy and gain calculation page for each attribute to create a rule tree, can be seen at Figure 5 and Figure 6. The decision tree and rules obtained by the system by calculating the existing training data as the knowledge base of the decision support system for selecting the secondary school for children with special needs are depicted in Figure 7 and Figure 8. The data set consists of 39 distinct and unique datasets separated into two, 75% for the training set and 25% for the testing set. The system will compute the entropy and gain values automatically.

After obtaining the highest gain value from the attribute, that attribute will be the first node.The system will calculate all the attributes again from the beginning, ignoring the attribute previously that have been a node leaf. When no attributes can be calculated, the system will return the decision tree rule result shown in Figure 7 and Figure 8. Figure 89 is the decision tree's result if the dataset changes.

After completing the rules of the knowledge model used for classification, the user can go to the classification menu and enter the included student data which will be classified according to the data entered by the user. The data that have user to be inserted into the program is consist of 6 attribute that each attribute have their own value to represent the condition that similar to the inclusive student. The data entry page is shown in **Error! Reference source not found.**.

(Discipline='Very Good') Then Secondary School = **SMP (1)**
(Discipline='Bad' / Discipline='Good') : **?**
--- (Discipline='Good') : **?**
--- --- (Cooperation='Very Good') Then Secondary School = **SMP (2)**
--- --- (Cooperation='Good') Then Secondary School = **SMP (3)**
--- --- (Cooperation='Bad') : **?**
--- --- --- (IQ < 75) Then Secondary School = **SLB (4)**
--- --- --- (IQ >= 75) Then Secondary School = **SMP (5)**
--- (Discipline='Bad') : **?**
--- --- (Active='Very Good') Then Secondary School = **SMP (6)**
--- --- (Active='Bad') Then Secondary School = **SLB (7)**
--- --- (Active ='Good') : **?**
--- --- --- (Cooperation='Very Good') Then Secondary School = **SMP (8)**
--- --- --- (Cooperation='Good') Then Secondary School = **SMP (9)**
--- --- --- (Cooperation='Bad') Then Secondary School = **SLB (10)**

Figure 7. Decision Tree Rule for First Dataset (39 data)

(Discipline='Very Good') Then Secondary School = **SMP (1)**
(Discipline='Bad' / Discipline='Good') : **?**
--- (Discipline='Good') : **?**
--- --- (Cooperation='Very Good') Then Secondary School = **SMP (2)**
--- --- (Cooperation='Good') Then Secondary School = **SMP (3)**
--- --- (Cooperation='Bad') : **?**
--- --- --- (Active='Very Good') Then Secondary School = **SMP (4)**
--- --- --- (Active='Bad') Then Secondary School = **SLB (5)**
--- --- --- (Active ='Good') Then Secondary School = **SMP (6)**
--- (Discipline='Bad') : **?**
--- --- (Active='Very Good') Then Secondary School = **SMP (7)**
--- --- (Active='Bad') Then Secondary School = **SLB (8)**
--- --- (Active ='Good') : **?**
--- --- --- (Cooperation='Very Good') Then Secondary School = **SMP (9)**
--- --- --- (Cooperation='Good') Then Secondary School = **SMP (10)**
--- --- --- (Cooperation='Bad') Then Secondary School = **SLB (11)**

Figure 8. Decision Tree Rule for Second Training Set (55 data)

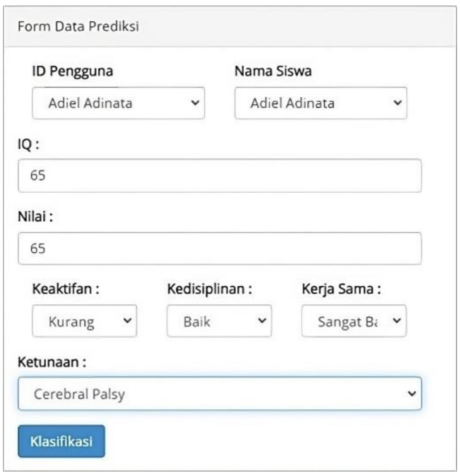

Figure 10. The Classification Page

After entering the student details, the user clicks on the "Classify" button and the system automatically matches the rules with the entered data and presents the user with the best recommendations shown in Figure 9.

Figure 9. The Result Recommendation Page

The decision support system showed a 90.00% accuracy for the 39 datasets, with good precision and recall value for each label. Especially when using the RapidMiner application's calculation like that stated in Table 5. Based on these findings, it is possible to conclude that the system developed can make accurate recommendations.

Table 5. Classification accuracy result with Rapidminer using ID3

Class	Accuracy	Precision	Recall
SMP	90.00%	100.00%	88.89%
SLB		50.00%	100.00%

3.4 *Comparison with other methods*

In addition to the ID3 algorithm, this study employs four other classification machine learning methods to compare their levels of accuracy and validity to the ID3 algorithm. The four methods used for comparison are the C4.5 algorithm, Naive Bayes, KNN (K-Nearest Neighbours), and SVM (Support Vector Machine). For a comparison of the ID3 algorithm and the other 4 methods commonly used in machine learning classification, as described in Table 6 and Table 77 . Table 6 shows the initial data's accuracy, precision, and recall (39 data). Furthermore, Table 7 shows the accuracy, precision, and recall of the changing value data (55 data).

Table 6. Algorithm Comparison for 39 data

Algorithm	Accuracy %	Precision % (SMP/SLB)	Recall % (SMP/SLB)
ID3	90.00	100.00	88.89
		50.00	100.00
C4.5	90.00	100.00	88.89
		50.00	100.00
Naïve Bayes	100.00	100.00	100.00
		100.00	100.00
KNN	90.00	90.00	100.00
		0.00	0.00
SVM	100.00	100.00	100.00
		100.00	100.00

Table 7. Algorithm comparison for 55 data

Algorithm	Accuracy %	Precision % (SMP/SLB)	Recall % (SMP/SLB)
ID3	100.00	100.00	100.00

DOI: https://doi.org/10.29207/resti.v7i5.5081

Algorithm	Accuracy %	Precision % (SMP/SLB)	Recall % (SMP/SLB)
		100.00	100.00
C4.5	85.71	100.00	83.33
		50.00	100.00
Naïve Bayes	100.00	100.00	100.00
		100.00	100.00
KNN	100.00	100.00	100.00
		100.00	100.00
SVM	100.00	100.00	100.00
		100.00	100.00

As shown in the two comparison Table 6 and 7, the ID3 algorithm has consistent accuracy, precision, and recall values that tend to increase even as the data changes/increases. However, the accuracy value of the Naïve Bayes and SVM methods must be admitted to be relatively high. The KNN method, based on the results, has an increased accuracy value even when the data changes. The C4.5 method does not have high accuracy compared with other methods. Another downside for the C4.5 is that the accuracy results are slightly decreased compared to when using the initial data, so there is a possibility that it will decrease again if more data is added, causing instability despite the high accuracy. The comparison can also be seen at Figure 102.

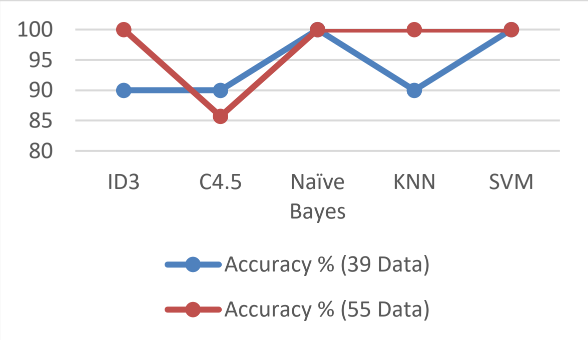

Figure 10. Accuracy comparison between each algorithm

Other comparison for precision and recall value from each algorithm is depicted in Figure 11 and Figure 124. From the comparison we can get information that the best algorithm that fits the research is Naïve Bayes. However, the ID3 algorithm and the SVM algorithm are also suitable because based on the graphs, the two algorithms produce positive values and are not much different from the values generated by the Naïve Bayes algorithm. This result is in line with the results of research conducted by Wiyono et al. [12] ,Bansod et al. [14], and Fitrianah et al. [31] where in their research they also revealed that ID3 (Decision Tree) and Naïve Bayes scores tended to be superior to the others. Information can be taken that the id3 algorithm does have a high level of accuracy when used in the data mining classification process.

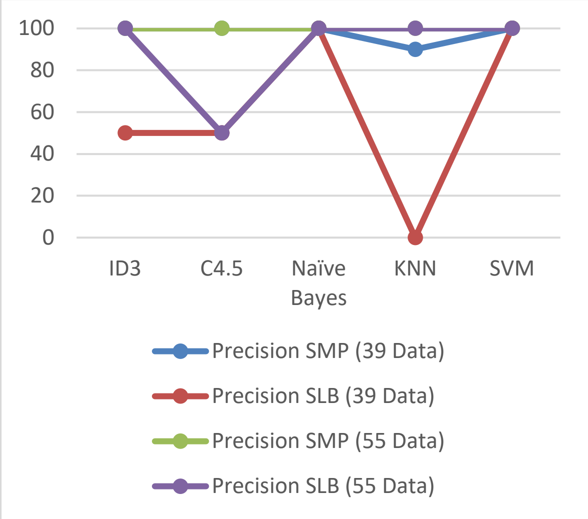

Figure 11. Precision comparison between each algorithm

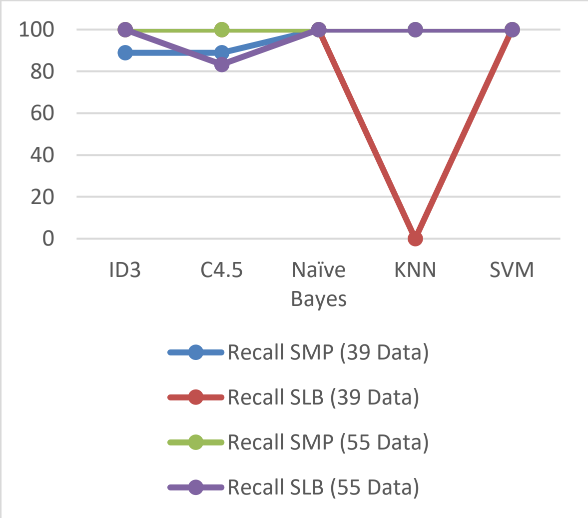

Figure 12. Recall comparison between each algorithm

Furthermore, the system automatically creates and calculates decision trees, making it simple even for non-technologists like primary school teachers. After cleaning and modification by the system, the new rule/decision tree can be used directly by the user for classification. Even when the decision tree changes, the system's accuracy is relatively reasonable and can be used as a benchmark for classifying. This accuracy may not be as high as the many studies mentioned in the introduction, where some studies can achieve more than 90% accuracy. However, this system provides flexibility and some good functionality for classification, as well as a high accuracy for updated data. This system also uses the ID3 algorithm that can adapt to new incoming data by simply changing the existing decision tree and still have a high accuracy compared to the other methods.

DOI: https://doi.org/10.29207/resti.v7i5.5081

4. Conclusion

In this study, we created program that have knowledge system to classify by extracting the existing primary inclusive student data on ability and academic history to assist inclusive students selecting secondary level school, In this study, we trained 39 data set which include records of cognitive, affective, and psychomotor in the form of academic history. Our evaluation showed the proposed academic decision support system produced an accuracy of 90.00%. In terms of precision and recall, our evaluation is able to provide 100% precision and 88.89% recall for the normal school class/inclusive class, while for the special school it can provide only 50% precision and 100% recall. Therefore, to improve the overall results, we trained the system using 55 data set. With this amount of data, it shows the system is able to provide precision 100% and recall 100% for both cases.

This study is another critical step toward expanding research in the field of academic data mining. Our goal and hope are that this work can be used as a model for future research in other fields of educational data mining. Furthermore, the decision support system application can be improved to be more dynamic. It is hopes that this research can be use as reference to provide better classification in data mining algorithm especially in the aspect of education for the inclusive children.

References

[1] T. Hamim, F. Benabbou, and N. Sael, “Survey of Machine Learning Techniques for Student Profile Modelling,” *International Journal of Emerging Technologies in Learning*, vol. 16, no. 4, pp. 136–151, 2021, doi: 10.3991/ijet.v16i04.18643.

[2] S. M. Stehle and E. E. Peters-Burton, “Developing student 21st Century skills in selected exemplary inclusive STEM high schools,” *Int J STEM Educ*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40594-019-0192-1.

[3] A. Kart and M. Kart, “Academic and social effects of inclusion on students without disabilities: A review of the literature,” *Educ Sci (Basel)*, vol. 11, no. 1, pp. 1–13, Jan. 2021, doi: 10.3390/educsci11010016.

[4] B. Lu and F. Zheng, “Influence of Experiential Teaching and Itinerary Assessment on the Improvement of Key Competencies of Students,” *International Journal of Emerging Technologies in Learning*, vol. 17, no. 2, pp. 173–188, 2022, doi: 10.3991/IJET.V17I02.29007.

[5] R. Faragher *et al.*, “Inclusive Education in Asia: Insights From Some Country Case Studies,” *J Policy Pract Intellect Disabil*, vol. 18, no. 1, pp. 23–35, Mar. 2021, doi: 10.1111/jppi.12369.

[6] C. S. Yu *et al.*, “Predicting metabolic syndrome with machine learning models using a decision tree algorithm: Retrospective cohort study,” *JMIR Med Inform*, vol. 8, no. 3, Mar. 2020, doi: 10.2196/17110.

[7] I. M. K. Karo, M. Y. Fajari, N. U. Fadhilah, and W. Y. Wardani, “Benchmarking Naïve Bayes and ID3 Algorithm for Prediction Student Scholarship,” *IOP Conf Ser Mater Sci Eng*, vol. 1232, no. 1, p. 012002, Mar. 2022, doi: 10.1088/1757-899x/1232/1/012002.

[8] Kementrian Pendidikan dan Kebudayaan, “Data Siswa Inklusi di Indonesia,” 2019. [Online]. Available: www.pusdatin.kemdikbud.go.id

[9] Haromain, “Pengembangan Program Layanan Sekolah Inklusi di Kota Mataram,” *Jurnal Realita*, vol. 5, no. 1, pp. 102–110, 2020, doi: https://doi.org/10.33394/realita.v5i1.2904.

[10] G. Mahendra Putra and V. Zahrotun Kamila, “Evaluasi klasifikasi penerima bidikmisi menggunakan algoritma interative dichotomiser 3 (ID3),” *Sains, Aplikasi, Komputasi dan Teknologi Informasi*, vol. 2, no. 1, pp. 1–10, 2020, doi: http://dx.doi.org/10.30872/jsakti.v2i1.

[11] C. Fiarni, E. M. Sipayung, and P. B. T. Tumundo, “Academic Decision Support System for Choosing Information Systems Sub Majors Programs using Decision Tree Algorithm,” *Journal of Information Systems Engineering and Business Intelligence*, vol. 5, no. 1, p. 57, Apr. 2019, doi: 10.20473/jisebi.5.1.57-66.

[12] S. Wiyono, D. S. Wibowo, M. F. Hidayatullah, and D. Dairoh, “Comparative Study of KNN, SVM and Decision Tree Algorithm for Student’s Performance Prediction,” *International Journal of Computing Science and Applied Mathematics*, vol. 6, no. 2, p. 50, Aug. 2020, doi: 10.12962/j24775401.v6i2.4360.

[13] H. Sulistiani and A. A. Aldino, “Decision Tree C4.5 Algorithm for Tuition Aid Grant Program Classification (Case Study: Department of Information System, Universitas Teknokrat Indonesia),” *Edutic - Scientific Journal of Informatics Education*, vol. 7, no. 1, pp. 40–50, Nov. 2020, doi: 10.21107/edutic.v7i1.8849.

[14] J. Bansod, M. Amonkar, T. Vaz, M. Sanke, and S. Aswale, “Prediction of Child Development using Data Mining Approach,” *Int J Comput Appl*, vol. 177, no. 44, pp. 975–8887, 2020, doi: http://dx.doi.org/10.5120/ijca2020919955.

[15] A. Sifaunajah and R. D. Wahyuningtyas, “Penggunaan Algoritma ID3 Untuk Klasifikasi Data Calon Perserta Didik,” *CSRID (Computer Science Research and Its Development Journal)*, vol. 14, no. 2, p. 103, Sep. 2022, doi: 10.22303/csrid.14.2.2022.103-112.

[16] M. Fadillah Rahman and Fadilah, “Klasifikasi Penerima Program Bantuan Beras Miskin Menggunakan Algoritma Iterative Dichotomiser 3,” *Progresif: Jurnal Ilmiah Komputer*, vol. 18, no. 1, pp. 101–114, 2022, doi: http://dx.doi.org/10.35889/progresif.v18i1.797.

[17] M. V. Sebt, E. Komijani, and S. S. Ghasemi, “Implementing a data mining solution approach to identify the valuable customers for facilitating electronic banking,” *International Journal of Interactive Mobile Technologies*, vol. 14, no. 15, pp. 157–174, 2020, doi: 10.3991/IJIM.V14I15.16127.

[18] G. Qiang, L. Sun, and Q. Huang, “ID3 algorithm and its improved algorithm in agricultural planting decision,” in *IOP Conference Series: Earth and Environmental Science*, Institute of Physics Publishing, May 2020. doi: 10.1088/1755-1315/474/3/032025.

[19] K. Ramya, Y. Teekaraman, and K. A. Ramesh Kumar, “Fuzzy-based energy management system with decision tree algorithm for power security system,” *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1173–1178, 2019, doi: 10.2991/ijcis.d.191016.001.

[20] Z. Zheng and K. S. Na, “A Data-Driven Emotion Model for English Learners Based on Machine Learning,” *International Journal of Emerging Technologies in Learning*, vol. 16, no. 8, pp. 34–46, 2021, doi: 10.3991/ijet.v16i08.22127.

[21] K. Sunday, P. Ocheja, S. Hussain, S. S. Oyelere, O. S. Balogun, and F. J. Agbo, “Analyzing student performance in programming education using classification techniques,” *International Journal of Emerging Technologies in Learning*, vol. 15, no. 2, pp. 127–144, 2020, doi: 10.3991/ijet.v15i02.11527.

[22] P. C. Saibabu, H. Sai, S. Yadav, and C. R. Srinivasan, “Synthesis of model predictive controller for an identified model of MIMO process,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 2, pp. 941–949, 2019, doi: 10.11591/ijeecs.

[23] N. Jahan and R. Shahariar, “Predicting fertilizer treatment of maize using decision tree algorithm,” *Indonesian Journal of*

DOI: https://doi.org/10.29207/resti.v7i5.5081

Electrical Engineering and Computer Science, vol. 20, no. 3, pp. 1427–1434, Dec. 2020, doi: 10.11591/ijeecs.v20.i3.pp1427-1434.

[24] Y. An and H. Zhou, "Short term effect evaluation model of rural energy construction revitalization based on ID3 decision tree algorithm," *Energy Reports*, vol. 8, pp. 1004–1012, Jul. 2022, doi: 10.1016/j.egyr.2022.01.239.

[25] S. Tangirala, "Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm*," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 2, pp. 612–619, 2020, doi: 10.14569/IJACSA.2020.0110277.

[26] I. S. Damanik, A. P. Windarto, A. Wanto, Poningsih, S. R. Andani, and W. Saputra, "Decision Tree Optimization in C4.5 Algorithm Using Genetic Algorithm," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Sep. 2019. doi: 10.1088/1742-6596/1255/1/012012.

[27] I. D. Mienye, Y. Sun, and Z. Wang, "Prediction performance of improved decision tree-based algorithms: A review," in *Procedia Manufacturing*, Elsevier B.V., 2019, pp. 698–703. doi: 10.1016/j.promfg.2019.06.011.

[28] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[29] I. M. K. Karo, M. Y. Fajari, N. U. Fadhilah, and W. Y. Wardani, "Benchmarking Naïve Bayes and ID3 Algorithm for Prediction Student Scholarship," *IOP Conf Ser Mater Sci Eng*, vol. 1232, no. 1, p. 012002, Mar. 2022, doi: 10.1088/1757-899x/1232/1/012002.

[30] S. T. Ahmed, R. Al-Hamdani, and M. S. Croock, "Developed third iterative dichotomizer based on feature decisive values for educational data mining," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 1, pp. 209–217, 2019, doi: 10.11591/ijeecs.v18.i1.pp209-217.

[31] D. Fitrianah, W. Gunawan, and A. Puspita Sari, "Studi Komparasi Algoritma Klasifikasi C5.0, SVM dan Naive Bayes dengan Studi Kasus Prediksi Banjir Comparative Study of Classification Algorithm between C5.0, SVM and Naive Bayes with Case Study of Flood Prediction," *Techno.COM*, vol. 21, no. 1, pp. 1–11, 2022, doi: https://doi.org/10.33633/tc.v21i1.5348.

DOI: https://doi.org/10.29207/resti.v7i5.5081