Decision tree based algorithms for Indonesian Language Sign System (SIBI) recognition

Agil Zaidan Nugraha [a,1], Reni Fatrisna Salsabila [a,2], Anik Nur Handayani [a,3,*], Aji Prasetya Wibawa [a,4], Emanuel Hitipeuw [a,5], Kohei Arai [b,6]

[a] Department of Electrical Engineering and Informatic, Universitas Negeri Malang, Malang 65145, Indonesia
[b] Departemen of Information Science and Engineering, Saga University, Saga 840-8502, Japan
[1] agil.zaidan.2305348@students.um.ac.id; [2] reni.fatrisna.2305348@students.um.ac.id; [3] aniknur.ft@um.ac.id; [4] imanuel.hitipeuw.fip@um.ac.id; [5] aji.prasetya.ft@um.ac.id; [6] arai@cc.saga-u.ac.jp
* corresponding author

ARTICLE INFO

Article history
Received July 23, 2024
Revised July 28, 2024
Accepted August 07, 2024
Available online August 10, 2024

Keywords
Indonesian Sign Language System (SIBI)
Decision tree
C4.5

ABSTRACT

Indonesian Sign Language System (SIBI) recognition plays a crucial role in improving effective communication for individuals with hearing loss in Indonesia. To support automatic SIBI recognition, this research presents a performance analysis of two main algorithms, namely Decision Tree and C4.5, in the context of the SIBI recognition task. This research utilizes a rich SIBI dataset that includes a variety of SIBI signs used in everyday communication. Data pre-processing, model construction with both algorithms, and model performance evaluation using accuracy, precision, recall, and F1-score metrics are all part of the study. Regarding SIBI recognition accuracy, the experimental results demonstrate that the Decision Tree performs better than Decision Tree. The Decision Tree also makes models that are easier to understand, which is important for making communication systems based on SIBI.

1. Introduction

Human communication primarily occurs through spoken language, yet not all individuals, particularly those who are deaf, can effectively engage in verbal communication in social contexts. Consequently, they often rely on sign language. SIBI is a prevalent sign language system utilized in Indonesian communication systems [1]. However, difficulties arise when individuals who are deaf or have speech impairments interact with those who do not understand sign language. This lack of mutual understanding poses significant obstacles. Introducing a sign language translation system can mitigate these challenges [2]. This research aims to determine the best set of parameters to optimize accuracy in such a system. The Indonesian Sign Language System (SIBI) is widely used in the deaf community in Indonesia. It is the main form of communication for those who have trouble hearing and speaking, but not everyone can understand sign language [3].

A common challenge encountered in sign language communication arises when individuals without hearing impairments attempt to communicate with those who are deaf. These individuals often struggle to comprehend the sign language used by the deaf [4]. Communication is a fundamental human activity aimed at understanding the intentions and goals of others [5]. Sign language facilitates effective communication for individuals with hearing loss, fostering better inclusion, participation, and

accessibility across various facets of life [6]. In the realm of ever-evolving information technology, the automatic recognition of the Indonesian Sign Language System (SIBI) through computational means holds significant potential for enhancing communication capabilities and accessibility for its users.

However, this pursuit presents various complex technical challenges. As indicated by [7], the processing of hand signals often involves intricate complexities. Machine learning approaches have been a major topic of research interest in order to overcome these issues. Among the notable algorithms in pattern recognition and data classification tasks are the Decision Tree and Decision Tree C4.5 algorithms. These algorithms have been successfully applied in numerous applications, including the recognition of sign language. The decision tree algorithm, for instance, operates as a decision-making method represented in the form of a tree or hierarchy [8]. It entails a recursive process wherein a set of statistical units undergo successive divisions based on rules aimed at maximizing the homogeneity or purity of the response variable within each group [9]. According to [10], the C4.5 Decision Tree algorithm finds wide application in data classification research due to its interpretability. However, a notable drawback of the C4.5 algorithm is its susceptibility to overfitting, wherein it performs well during training but exhibits weaknesses when applied to unseen data. Additionally, the classification performance of the C4.5 decision tree algorithm is vulnerable to misclassification costs, often stemming from suboptimal attribute separation factors [11].

The utilization of both algorithms in various applications has sparked increasing interest in the field of machine learning and pattern recognition. Previous studies, such as "The Utilization of Data Mining for Predicting Timely Graduation with the C45 Algorithm" [12] and "C.45 Classification Model for Assessing Customer Service Quality at Bank BTN Pematangsiantar Branch" [13], highlight the significant potential of the C4.5 algorithm in data analysis for decision-making across various domains. Moreover, this research refers to other relevant studies such as "Using Deep Convolutional Networks for Gesture Recognition in American Sign Language" [14] and "An Efficient Hand Gesture Recognition System Based on Deep CNN" [15], which explore the application of image processing techniques and artificial neural networks (CNNs) in hand gesture recognition, potentially impacting SIBI gesture recognition. This study aims to analyze the performance of two main algorithms, namely Decision Tree and Decision Tree C4.5, within the context of the Indonesian Language Sign System (SIBI) recognition. Both algorithms will be applied and evaluated using a dataset comprising a variety of SIBI signs commonly used in everyday communicative interactions. The primary objective is to identify the most effective algorithm for accurately recognizing SIBI signs. Additionally, this research will encompass significant additional evaluation aspects. This evaluation will include assessing the extent to which the classification models generated by each algorithm can be interpreted. The presence of comprehensible interpretation is a critical factor in the context of developing SIBI-based communication technologies [16]. The outcomes of this research are expected to offer substantial insights into the capabilities of these algorithms within the context of SIBI, thereby fostering the development of technologies that support inclusive and user-friendly communication for the deaf community in Indonesia.

2. Method

This study employs a method that is structured and systematic in order to examine the effectiveness of the Decision Tree and Decision Tree C4.5 algorithms in recognizing the Indonesian Language Sign System (SIBI). We will describe the steps taken to evaluate both algorithms' performance in this method section. A flowchart depicting the comprehensive analysis procedure will be presented to serve as a visual guide. Research flowchart show as Fig. 1.

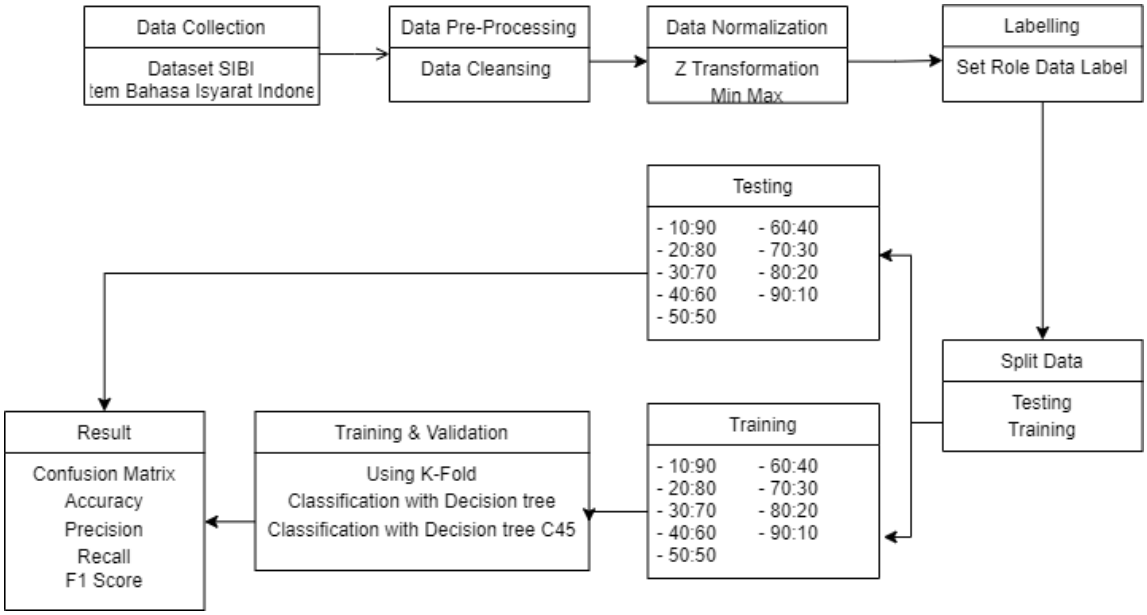

Fig. 1. Research Flowchart

2.1. SIBI Data Collection

Indonesian Sign Language System (SIBI) dataset containing hand signals used in everyday communication by the deaf community in Indonesia was collected. By distributing the form to responders, the data was collected [17]. This dataset contains a range of gestures that reflect the most often used words and phrases in the Indonesian Sign Language System. This data is used to test and train the SIBI recognition model [18].

2.2. Data Pre-Processing

Data pre-processing includes removing duplicate data, identifying inconsistent data, and fixing data problems. Another phase in this process is enrichment, which is the act of adding additional pertinent data to already-existing data [19]. Before the data can be used for model training, a number of data pre-processing steps are carried out to ensure good data quality.

2.3. Normalization of Data Z transform and Min Max Normalization

Two popular normalizing methods are Min-Max normalizing and the Z Transform (standardization). Through the use of the Z transform, the data are transformed into a normal distribution with a zero mean and a one standard deviation [20]. The formula provides the Z-transform of a discrete-time signal $x[n]$ is given by the formula.

$$X(z) = \sum_{n=-\infty}^{\infty} x|n|.z^{-n} \quad (1)$$

In this formula, $X(z)$ is the Z-transform of the sequence $x[n]$. The Z-transform is a mathematical technique used in digital signal processing to analyze and transform discrete-time signals. The Z-transform can be reached by using the product of each sequence $x[n]$ and z –n across all possible values of n. The variable z is a complex integer. The Z-transform is especially effective when examining the behavior of discrete-time systems in the frequency domain. It allows the representation of discrete-time signals and systems in terms of complex exponential functions. This formula essentially captures the relationship between the time-domain representation ($x[n]$) and its Z-transform ($X(z)$). The Z-transform is a powerful tool for analyzing the properties and behavior of discrete-time signals and

systems, providing insights into their frequency characteristics and facilitating the design of digital filters and systems.

While the MinMax Normalization method is a normalization method that changes the range of data values to be between 0 and 1 [21]. Min-Max normalization, also known as feature scaling or min-max scaling, is another technique commonly used to scale and normalize data. The formula for Min-Max normalization is as follows:

$$X_{normalized} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

X normalized is the normalized value of the variable. Two popular normalizing methods are Z Transform (standardization) and Min-Max normalizing. Using the Z transform, the data are transformed into a normal distribution with a mean of zero and a standard deviation of one. The Min-Max normalization is particularly useful when the data needs to be on a similar scale for certain algorithms or analyses, preventing features with larger scales from dominating those with smaller scales.

2.4. Labeling

The labeling process is a significant stage in data preparation that involves assigning a specific role to each attribute in the dataset [22]. This is done to identify the role and contribution of each attribute in the data analysis to be performed. In this context, the labeling process allows us to clearly define which attribute will be the target or label in our analysis, as well as the other attributes that will be used to understand, support, or classify that target. According to [23] the labeling process not only serves as a preliminary step in data analysis, but also helps in preparing the dataset well, so that we can optimize the use of certain attributes in our analysis. In other words, labeling is an important aspect that allows us to better understand the contribution of each attribute in the broader context of data analysis.

2.5. Split Data

Dividing the dataset into two primary subsets the training subset, which is used to train the model, and the testing subset, which is used to evaluate the model's performance is a crucial step in data analysis and machine learning [24]. The splitting ratio, such as 10:90, 20:80, etc., determines what percentage of the data is used for each purpose [25]. This process helps prevent overfitting and ensures the model can generalize to data it has never seen before, as well as assisting in assessing the extent to which the model can be used in real-world situations.

2.6. K-Fold

The k-fold process is a commonly used validation technique in data analysis and machine learning [26]. This method divides the dataset into k equal subsections. With one segment acting as a testing set and another as a training set, the model is repeatedly trained and tested. The results from each test are combined to provide a more stable and objective performance estimate. K-fold cross-validation helps in avoiding overfitting and ensures the model has good generalization ability to data that has never been seen before [27].

2.7. Classification

- Decision Tree

  Decision Tree is widely applied in many fields, such as classification and recognition [28]. A Decision Tree is a graphical representation resembling a tree, where nodes represent decision points based on selected attributes and associated questions, edges represent the possible answers, and

leaves signify the actual class labels [29]. The structure of the decision tree commences with a dataset (training set) that is divided at each node, leading to the formation of smaller subsets. This methodology adheres to a recursive partitioning approach. Accompanying the dataset is a collection of attributes. Objects may manifest as events, activities, or features that furnish information about the respective object. Each tuple within the dataset is associated with a class label, signifying the object's classification within a specific class [30]. measures how well an attribute splits a dataset into more homogeneous subsets based on the target class. The higher the better the attribute is at dividing the data.The formula is:

$$Gain(S,A) = E(S) - \sum v \in Values(A) \frac{|S_v|}{|S|}.E(S_v) \quad (3)$$

$E(S)$ is the entropy of set S, $alues(A)$ are the unique values of attribute A, $|S_v|$ is the number of instances in S with attribute A equal to v.

Entropy (E(S)) is calculated as:

$$E(S) = -\sum_{i=1}^{C} pi.log_2(pi) \quad (4)$$

where pi is the percentage of instances in set S with class i. After dividing the dataset based on attribute A, the reduction in entropy also known as uncertainty is measured by Information Gain

- C45

[30] introduced the C4.5 algorithm for constructing decision trees. The decision tree's framework initiates with a dataset (training set) that undergoes partitioning at each node, leading to smaller subsets. C4.5 is versatile, handling attributes of both discrete and continuous types. The algorithm selects attributes based on entropy, using information gain as a heuristic to identify the optimal subset of examples within a class. All attributes are treated as discrete value categories, necessitating the discretization of attributes with continuous values. Discretization aims to simplify the problem and enhance learning accuracy by grouping values according to predefined criteria [31]. In the C4.5 algorithm, attribute selection employs gain instead of information gain. A favorable attribute is one that results in the smallest decision tree size or effectively separates objects by class. Heuristically, the chosen attribute produces the cleanest node, measured by impurity level. Impurity is gauged using the concept of entropy, representing the impurity of a set of objects [32]. The goal is to maximize information gain by utilizing entropy values at each node. The formula is

$$Gain\ Ration(S,A) = \frac{Gain(S.A)}{SplitInfo(S.A)} \quad (5)$$

Gain (S,A) is the Information Gain, measured as the difference between the entropy before and after splitting the dataset based on attribute A. SplitInfo (S,A) is the information required to split the dataset based on attribute A, measured as the entropy of the distribution of attribute A in the dataset.

Where SplitInfo (S,A) is:

$$Gain(S,A) = E(S) - \sum v \in Values(A) \frac{|S_v|}{|S|}.E(S_v) \quad (6)$$

Where $E(S)$ is the entropy of set S, $Values(A)$ are the unique values of attribute A, S_v is the number of instances in S with attribute A equal to v. The formula for SplitInfo (S, A) is:

$$SplitInfo(S,A) = v \in Values(A) \frac{|S_v|}{|S|} . log_2 \left(\frac{|S_v|}{|S|}\right) \quad (7)$$

So it aims to ensure that attributes producing many values do not dominate attribute selection solely based on their high count

2.8. Result (Confusion Matrix)

After cross-validation, we measured the models' performance using different test data. We quantify accuracy, precision, recall, and F1-score using a confusion matrix to assess how effectively these models can identify SIBI signals. With regard to SIBI recognition, this assessment offers a more profound understanding of the effectiveness of both algorithms.

3. Results and Discussion

After cross-validation, we measured the models' performance using different test data. We quantify accuracy, precision, recall, and F1-score using a confusion matrix to assess how effectively these models can identify SIBI signals. With regard to SIBI recognition, this assessment offers a more profound understanding of the effectiveness of both algorithms.

3.1. Decision Tree Model Result

- Testing based on normalization as show in Table 1.

Table 1. Normalization

No.	Number of Nodes	Z-Transformation	Min Max
1.	10	6,7 %	23,3 %
2.	25	12,1 %	42 %
3.	50	9.9 %	34 %

In the normalization test table based on Z Transformation and Min-Max Normalization for different number of nodes (10, 25, and 50), several things can be observed:

- Z Transformation resulted in a higher percentage of normalization compared to Min-Max Normalization in all cases. This means that the data, after normalization, has a significant standard deviation from the initial mean.

  The percentage of normalization with Z Transformation increases as the number of nodes increases. In this test, the percentage of the Z Transformation test is 61.30% to 77.82% with the number of nodes to be tested. This shows that Z Transformation is able to produce data that is more dispersed and with greater variation.

- Min-Max Normalization results in a lower normalization percentage compared to Z Transformation in all cases. This means that the data after normalization remains within a more limited range (in this case, around 38.19% to 40.20%). Regardless of the number of nodes, the normalization percentage with Min-Max Normalization remains stable, which is around 38.19% to 40.20%.

The test results show that *Z* Transformation produces a higher normalization percentage than Min- Max Normalization in all cases. This means that *Z* Transformation is more effective in expanding the variety of data. In addition, it is seen that the normalization percentage with Min-Max Normalization tends to remain fixed, around 38.19% to 40.20%, regardless of the number of nodes. This shows that Min-Max Normalization keeps the data in a consistent range

- Accuracy

In experiments using the Decision Tree algorithm with different tree sizes, we evaluated the performance of the model using accuracy metrics. The following are the results and discussion of the experiments as show in Table 2.

Table 2. Accuracy Level of Split Data Decision Tree Model

No.	Number of Nodes	Accuracy
1.	10	23.21%
2.	25	36.65%
3.	50	47.02%
4.	100	50.35%
5.	250	77.82 %

In the first test with 10 nodes, the accuracy of the Decision Tree model was about 23.21%. However, this result shows that this model has a relatively low and inconsistent performance in recognizing SIBI cues. In the second test with 25 nodes, there was a significant increase in the accuracy of the model to about 36.65%. This result shows that this model has a relatively low and inconsistent performance in recognizing SIBI cues. The third test with 50 nodes saw a significant increase in model accuracy to about 47.02%. This result shows that this model has a relatively low and inconsistent performance in recognizing SIBI cues. The fourth test with 100 nodes saw a significant increase in model accuracy to about 50.35%. This result shows that this model has a relatively low and inconsistent performance in recognizing SIBI cues. In the last test with 250 nodes, the accuracy remained at the same level, which was about 77.82%. This shows that although larger trees have the capacity to capture more patterns, in this context, there is no significant improvement in the performance of the model.

The experimental results show that in SIBI sign recognition using the Decision Tree algorithm, the model with 25 nodes performs the best with an accuracy of about 77.82%. Increasing the number of nodes above 25 did not result in a significant improvement in accuracy. The variability in accuracy between experiments of about 0.77% indicates that there are several factors that can affect the performance of the model, such as diversity i the dataset or algorithm settings.

- Decision tree model evaluation result

The decision tree model with the best accuracy of 77.82% with the number of nodes 250. The model is then tested and the results of precision, recall, and F1 Score are obtained as show in Table 3:

Table 3. Decision Tree Model Evaluation Results

No.	Decision Tree Model Evaluation Results			*F1-Score*
	Letter	*Precision*	*Recall*	
1.	A	80.09%	79.06%	79.57%
2.	B	84.58%	85.47%	85.02%
3.	C	77.93%	77.64%	77.79%
4.	D	70.68%	69.27%	69.97%
5.	E	74.27%	72.29%	73.26%
6.	F	84.73%	84.79%	84.76%
7.	G	74.20%	73.52%	73.85%
8.	H	91.76%	92.70%	92.23%
9.	I	72.01%	69.77%	70.88%
10.	K	72.33%	68.20%	70.25%
11.	L	90.44%	88.93%	89.67%
12.	M	65.66%	67.45%	66.54%
13.	N	65.97%	67.63%	66.79%
14.	O	72.13%	73.84%	72.97%
15.	P	93.68%	94.52%	94.10%
16.	Q	93.91%	92.33%	93.12%
17.	R	74.09%	74.29%	74.19%
18.	S	58.30%	58.79%	58.55%
19	T	70.50%	69.30%	69.89%
20.	U	72.33%	74.56%	73.44%
21.	V	79.94%	81.07%	80.49%
22.	W	90.62%	91.14%	90.88%
23.	X	66.88%	68.33%	67.60%
24.	Y	84.00%	85.26%	84.62%

This table helps the reader to understand the performance of the model or system in recognizing specific SIBI letters. High values of precision, recall, and F1-Score indicate good performance, while low values may indicate that the model needs to be improved in the recognition of certain SIBI letters. Precision measures how many of the positive predictions made by the model are actually correct. In this context, precision measures how accurate the model is in recognizing each letter. The higher the precision value, the fewer false positive prediction errors made by the model. Recall measures the extent to which the model is able to detect all true positive cases. In this context, recall measures how well the model captures all possible letters that actually exist. The higher the recall value, the fewer positive cases the model misses. F1-Score is a measure that combines precision and recall into a single value. It provides a holistic picture of the model's performance, seeking a balance between accuracy and detectability. The higher the F1-Score value, the better the model performs in recognizing letters in SIBI. For example, the letter "P" has a Precision of 93.68%, a Recall of 94.52%, and an F1-Score of 94.10%. This shows that the model has a very good performance in recognizing the letter "P". In contrast, the letter "S" has a Precision of 58.30%, a Recall of 58.79%, and an F1- Score of 58.55%, which indicates that the model has a lower performance in recognizing the letter "S".

- Confusion matrix decision tree

  Here is the confusion matrix of the decision tree as show in Fig. 2:

ConfusionMatrix:

True:	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
A:	1038	0	0	4	26	5	12	1	4	3	12	38	45	1	0	0	4	23	66	0	0	0	0	15
B:	2	1136	8	4	5	75	4	2	4	15	2	1	1	4	0	0	11	2	4	18	6	36	2	0
C:	3	16	1044	34	9	8	70	12	3	7	7	1	1	55	5	8	1	7	9	2	2	3	29	4
D:	5	5	41	933	7	5	54	7	25	16	17	4	6	56	0	1	15	20	16	18	10	3	52	4
E:	27	6	4	11	973	5	5	2	43	4	4	45	26	5	0	1	7	69	19	5	3	0	23	23
F:	3	79	10	9	6	1155	4	0	12	11	4	3	1	4	0	0	11	14	1	11	3	15	7	0
G:	6	5	70	54	7	2	969	27	3	2	27	6	9	46	4	7	3	11	12	2	0	1	27	6
H:	1	0	7	8	0	0	22	1259	0	11	1	3	1	7	8	1	14	5	2	4	8	0	9	0
I:	7	4	3	16	43	13	7	2	853	10	6	17	16	8	1	0	10	49	28	5	2	1	28	55
K:	0	12	6	16	3	12	7	8	12	907	6	5	9	4	0	0	71	15	17	63	57	11	12	5
L:	15	1	2	21	4	1	24	2	5	4	1172	1	2	4	0	1	5	8	11	2	0	0	2	9
M:	47	0	0	3	59	1	8	1	27	3	5	889	174	3	1	0	4	69	27	2	3	0	7	21
N:	47	1	1	10	51	0	1	1	22	3	3	183	917	7	0	1	1	65	33	7	2	0	22	12
O:	1	2	70	59	6	1	52	7	20	6	4	4	5	968	4	8	3	18	22	1	2	3	71	5
P:	0	0	3	2	0	0	8	3	0	0	2	0	0	4	1586	81	0	1	1	0	0	0	1	1
Q:	0	0	10	1	0	0	4	0	0	0	1	0	0	6	62	1373	0	2	0	0	0	0	1	2
R:	2	8	1	14	5	14	3	7	7	86	10	5	6	3	1	0	1099	23	12	86	57	5	29	0
S:	30	0	12	23	77	12	15	2	46	23	6	64	52	19	1	0	19	766	68	7	3	2	44	23
T:	59	4	10	12	10	4	14	2	21	34	10	18	31	18	2	2	12	66	937	6	3	1	39	14
U:	1	5	1	18	6	15	1	3	2	70	6	4	8	1	0	0	102	11	11	1021	87	20	16	1
V:	0	3	2	9	3	6	1	4	4	77	3	1	7	2	0	0	54	8	1	77	1150	18	5	1
W:	0	30	3	10	2	24	2	0	1	14	0	0	1	3	0	0	9	0	1	13	12	1274	7	0
X:	3	11	31	68	22	3	22	5	28	23	2	8	22	80	2	1	30	40	38	14	5	6	947	5
Y:	16	0	7	8	22	0	9	0	82	1	9	18	16	3	1	2	1	11	16	0	1	0	6	1198

Fig. 2. Confusion Matrix Decision Tree

This Confusion Matrix provides information on how well the model can recognize each letter in the Indonesian Sign Language System (SIBI). For example, for the letter 'A', the model made correct predictions 1038 times, but also made several mispredictions such as predicting 'A' as 'E' 26 times, 'A' as 'F' 38 times, and so on.

- Split Result

  After conducting a test to determine the best value, then we conduct another experiment by splitting the data by dividing the dataset into two subsets, namely (training) that has been tested and (testing) that has been trained. The context of the Indonesian Sign Language System (SIBI) dataset in RapidMiner, the splitting process is divided into several inputs, namely 10:90, 20:80, 30:70, 40:60 and 50:50. Later on each input will be analyzed which formation is best for the SIBI dataset. The following are the results of the split data experiment on the Decision Tree algorithm as show in Table 4.

Table 4. Decision Tree Algorithm Split Data Results

Ratio	Decision Tree Split Data Experiment Results			F1-Score
	Accuracy	*Precision*	*Recall*	
10:90	73.81 %	75.22%	73.39%	74.29%
20:80	59.04 %	59.27%	58.48%	58.87%
30:70	57.38 %	57.67%	56.82%	57.24%
40:60	60.99 %	62.37%	60.50%	61.42%
50:50	63.78 %	64.10%	63.34%	63.71%

The Decision Tree Split Data experiment results indicate varying model performance depending on the data split ratio. Generally, accuracy tends to decrease with an increase in the proportion of training data. There is a pattern showing issues of overfitting as the proportion of training data increases. Precision and recall provide additional insight into the model's predictive quality. The experiment results show that the balance between precision and recall varies depending on the data split ratio. This research leads to the result that a model's capacity to recognize Indonesian Sign Language (SIBI) is influenced by the data split ratio. However,

these results provide valuable insights for the development of more effective models in SIBI recognition.

3.2. Decision Tree Model Result C45

- Accuracy

In experiments using the C45 algorithm with different tree sizes, we evaluated the performance of the model using accuracy metrics. The experiments' result and commentary are provided as show in Table 5:

Table 5. Accuracy Level of C45 Model

No.	Number of Nodes	Accuracy
1.	10	18.75%
2.	25	33.91%
3.	50	45.59%
4.	100	50.59%
5.	250	57.69%
6.	500	69.66%
7.	750	71.90%

In the first test with 10 nodes, the accuracy of the C4.5 Decision Tree model was about 18.75%. However, the variability in accuracy between trials was relatively high, with 10 nodes having low and inconsistent performance in recognizing SIBI cues. In the second test with 25 nodes, there was a significant increase in the accuracy of the model to about 33.91%. This result shows that models with more nodes have better performance in SIBI sign recognition. The third test with 50 nodes resulted in a further improvement in accuracy to about 45.59%. This shows that the addition of nodes after reaching 25 nodes still provides a meaningful improvement in model performance. The fourth test with 100 nodes resulted in an accuracy of approximately 50.59%. This test showed that the model with 100 nodes continued to improve its performance, although the improvement was not as great as in the previous stage. The fifth test with 250 nodes resulted in a significant improvement in accuracy to about 57.69%. This shows that increasing the capacity of the model by adding more nodes has a positive impact on performance. The sixth test with 500 nodes resulted in an accuracy of about 69.66%, which is a significant improvement from before. This shows that increasing the model capacity continues to contribute to improved performance. The seventh test with 750 nodes resulted in an accuracy of approximately 71.90%, which is a further improvement over the previous test.

The experimental results show that in SIBI sign recognition using the C4.5 algorithm, the performance of the model improves significantly with an increase in the number of nodes in the decision tree. The model with 750 nodes has the highest accuracy of about 71.90%. The low variability of accuracy between trials indicates that models with a larger number of nodes are more stable and consistent in SIBI sign recognition. This indicates that the C4.5 Decision Tree algorithm, when configured with an appropriate number of nodes, can be a good choice in Indonesian Sign Language (SIBI) recognition. However, it should be kept in mind that the use of models with many nodes may increase the complexity and computational requirements, so it is necessary to consider the trade-off between performance and efficiency. In addition, further

experiments and comparisons with other algorithms are needed to confirm these results and determine the most effective algorithm in the context of SIBI recognition

- C45 Decision tree Model Evaluation Results

  In the C45 decision tree model with the best accuracy of 71.90% with 750 nodes. The model is then tested and the results of precision, recall, and F1 Score are obtained as show in Table 6:

Table 6. Evaluation of C45 Model

No	Decision Tree Model Evaluation Results			F1-Score
	Letter	*Precision*	*Recall*	
1.	A	77.30%	74.71%	75.96%
2.	B	82.47%	81.85%	82.16%
3.	C	72.43%	73.40%	72.91%
4.	D	66.81%	68.30%	67.54%
5.	E	62.57%	62.48%	62.52%
6.	F	80.75%	82.00%	81.36%
7.	G	66.30%	67.91%	67.09%
8.	H	92.13%	92.34%	92.23%
9.	I	65.40%	64.54%	64.96%
10.	K	58.20%	60.83%	59.50%
11.	L	85.60%	84.76%	85.18%
12.	M	59.68%	57.28%	58.45%
13.	N	62.08%	60.62%	61.34%
14.	O	65.23%	63.39%	64.29%
15.	P	92.75%	92.25%	92.50%
16.	Q	90.89%	91.26%	91.07%
17.	R	63.51%	64.06%	63.78%
18.	S	46.49%	53.88%	50.12%
19	T	63.00%	59.69%	61.29%
20.	U	67.32%	67.96%	67.64%
21.	V	74.59%	70.90%	72.71%
22.	W	87.90%	88.28%	88.09%
23.	X	55.17%	53.10%	54.11%

This table helps the reader to understand the performance of the model or system in recognizing specific SIBI letters. High values of precision, recall, and F1-Score indicate good performance, while low values may indicate that the model needs to be improved in the recognition of certain SIBI letters. Precision measures how many of the positive predictions made by the model are actually correct. In this context, precision measures how accurate the model is in recognizing each letter. The higher the precision value, the fewer false positive prediction errors made by the model. Recall measures the extent to which the model is able to detect all true positive cases. In this context, recall measures how well the model captures all possible letters that actually exist. The higher the recall value, the fewer positive cases the model misses. F1-Score is a measure that combines precision and recall into a single value. It provides a holistic picture of the model's performance, seeking a balance between accuracy and detectability. The higher the F1-Score value, the better the model's performance in recognizing letters in SIBI

- Confusion Matrix C45

Here is the confusion matrix of the C45 as show in Fig. 3.

ConfusionMatrix:

True:	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
A:	1038	0	0	4	26	5	12	1	4	3	12	38	45	1	0	0	4	23	66	0	0	0	0	15
B:	2	1136	8	4	5	75	4	2	4	15	2	1	1	4	0	0	11	2	4	18	6	36	2	0
C:	3	16	1044	34	9	8	70	12	3	7	7	1	1	55	5	8	1	7	9	2	2	3	29	4
D:	5	5	41	933	7	5	54	7	25	16	17	4	6	56	0	1	15	20	16	18	10	3	52	4
E:	27	6	4	11	973	5	5	2	43	4	4	45	26	5	0	1	7	69	19	5	3	0	23	23
F:	3	79	10	9	6	1155	4	0	12	11	4	3	1	4	0	0	11	14	1	11	3	15	7	0
G:	6	5	70	54	7	2	969	27	3	2	27	6	9	46	4	7	3	11	12	2	0	1	27	6
H:	1	0	7	8	0	0	22	1259	0	11	1	3	1	7	8	1	14	5	2	4	8	0	9	0
I:	7	4	3	16	43	13	7	2	853	10	6	17	16	8	1	0	10	49	28	5	2	1	28	55
K:	0	12	6	16	3	12	7	8	12	907	6	5	9	4	0	0	71	15	17	63	57	11	12	5
L:	15	1	2	21	4	1	24	2	5	4	1172	1	2	4	0	1	5	8	11	2	0	0	2	9
M:	47	0	0	3	59	1	8	1	27	3	5	889	174	3	1	0	4	69	27	2	3	0	7	21
N:	47	1	1	10	51	0	1	1	22	3	3	183	917	7	0	1	1	65	33	7	2	0	22	12
O:	1	2	70	59	6	1	52	7	20	6	4	4	5	968	4	8	3	18	22	1	2	3	71	5
P:	0	0	3	2	0	0	8	3	0	0	2	0	0	4	1586	81	0	1	1	0	0	0	1	1
Q:	0	0	10	1	0	0	4	0	0	0	1	0	0	6	62	1373	0	2	0	0	0	0	1	2
R:	2	8	1	14	5	14	3	7	7	86	10	5	6	3	1	0	1099	23	12	86	57	5	29	0
S:	30	0	12	23	77	12	15	2	46	23	6	64	52	19	1	0	19	766	68	7	3	2	44	23
T:	59	4	10	12	10	4	14	2	21	34	10	18	31	18	2	2	12	66	937	6	3	1	39	14
U:	1	5	1	18	6	15	1	3	2	70	6	4	8	1	0	0	102	11	11	1021	87	20	16	1
V:	0	3	2	9	3	6	1	4	4	77	3	1	7	2	0	0	54	8	1	77	1150	18	5	1
W:	0	30	3	10	2	24	2	0	1	14	0	0	1	3	0	0	9	0	1	13	12	1274	7	0
X:	3	11	31	68	22	3	22	5	28	23	2	8	22	80	2	1	30	40	38	14	5	6	947	5
Y:	16	0	7	8	22	0	9	0	82	1	9	18	16	3	1	2	1	11	16	0	1	0	6	1198

Fig. 3. Confusion Matrix C45

The Confusion Matrix above provides an overview of the performance of the C45 Decision Tree model in classifying letters in the Indonesian Language Sign System (SIBI). For example, the letter 'A' has a recall value of 74.71%, which means most of the actual data labeled 'A' has been found by the model. However, the model also made some prediction errors, such as predicting 'A' as 'E' 26 times, and so on. In addition, F1 Score, which is a combined measurement of precision and recall, is also provided for each letter. For example, the letter 'A' has an F1 Score of 75.96%. Overall, this Confusion Matrix provides an overview of the extent to which the C45 Decision Tree model can recognize letters in the Indonesian Language Signing System (SIBI) and provides a better understanding of the model's performance in the task.

After experimenting based on the number of nodes in the C45 Decision Tree, the node value of 750 is obtained as the best value. Like the Decision Tree algorithm, we experiment again by splitting the data to compare the results with the algorithm. Just like before the Indonesian Sign Language System (SIBI) dataset is divided into several inputs, namely 10:90, 20:80, 30:70, 40:60 and 50;50. Later on each input will be analyzed which formation is best for the SIBI dataset. The following are the results of the split data experiment on the C45 Decision Tree algorithm as show in Table 7.

Table 7. Decision Tree Results C45 Split Data

Ratio	Decision Tree Split Data Experiment Results			F1-Score
	Accuracy	*Precision*	*Recall*	
10:90	22.39%	21.58%	20.95%	21.26%
20:80	20.44%	29.97%	19.29%	23.47%
30:70	14.53%	16.19%	13.12%	14.49%
40:60	19.44%	30.19%	18.22%	22.72%
50:50	11.77%	16.73%	10.35%	12.78%

The research exhibited notable strengths in evaluating various decision tree models for the recognition of indonesian sign language system (SIBI) letters. The comprehensive assessment

using metrics like accuracy, precision, recall, and f1-score provided a robust understanding of model performance. The detailed analysis of different configurations with varying numbers of nodes showcased a systematic exploration of the decision tree and c4.5 algorithms, elucidating their impact on recognition accuracy. Additionally, the experimentation with different data split ratios contributed valuable insights into the models' sensitivity to training-testing set distributions. This thorough examination highlights the research's strength in meticulously exploring multiple facets of model performance, thereby offering a rich understanding of the models' efficacy in sibi letter recognition.

However, the research also presents certain limitations. Despite the extensive evaluation of decision tree models, the focus predominantly on decision tree algorithms might limit the scope of comparison with other machine learning approaches specifically tailored for pattern recognition tasks. Moreover, while metrics like precision, recall, and F1-score provide insights into model performance, they might not fully capture the real-world application challenges or account for nuances in recognizing sign language gestures, where subtle variations could significantly impact interpretation. Furthermore, the research lacks an in-depth exploration of techniques to handle imbalanced data, which is often prevalent in sign language datasets and could affect model generalization and performance.

Moving forward, addressing these challenges presents itself as a pivotal direction for future research. Exploring a wider array of machine learning algorithms, including deep learning architectures like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), could provide comparative insights into their effectiveness in SIBI recognition tasks. Additionally, integrating domain-specific features or temporal information into the models might enhance their ability to capture the intricacies of sign language gestures. Moreover, devising strategies to mitigate the impact of imbalanced data and enhancing model interpretability could further improve the practical applicability of these models in real-world scenarios. Embracing interdisciplinary collaborations with experts in sign language linguistics or cognitive science could offer valuable perspectives in refining models to better align with the nuances and complexities of sign language interpretation, fostering more accurate and inclusive communication technologies

4. Conclusion

Conclusion for Decision Tree and Decision Tree C4.5 Comparison on Indonesian Sign Language (SIBI) Recognition. Decision Tree, Decision Tree models have varying levels of accuracy depending on the number of trees used. In this case, the highest accuracy was achieved at 25 trees with an accuracy of about 77.82% the evaluation results using Precision, Recall, and F1-Score metrics show a fairly good performance in recognizing SIBI letters, with some letters having higher scores than others overall, Decision Tree can be used as a model for letter recognition in the Indonesian Sign Language System (SIBI) with quite good performance, especially with the optimal number of decision trees. Decision Tree C4.5, Decision Tree model has a lower accuracy rate than the Decision Tree. The accuracy rate increases as the number of decision trees increases, with the best accuracy achieved at 750 trees with an accuracy of about 71.90%. Evaluation results using Precision, Recall, and F1-Score metrics show that this model has lower performance than Decision Tree in recognizing SIBI letters although Decision Tree C4.5 has a lower accuracy, it may still be used as one of the alternatives in letter recognition in the Indonesian Language Sign System (SIBI) if needed. Overall, based on the test results, the Decision Tree has a better

performance in letter recognition in the Indonesian Language Sign System (SIBI) compared to the C4.5 Decision Tree, especially in terms of accuracy and most other evaluation metrics. However, the choice between these two algorithms can also be affected by other factors, such as the speed of model training and available resources. However, the study's concentration on decision tree algorithms limits comparisons with other machine learning approaches tailored for pattern recognition, and the reliance on metrics may overlook the intricacies of sign language gestures, posing challenges in real-world applications. Future endeavors should encompass a broader spectrum of machine learning techniques, such as deep learning architectures, incorporate domain-specific features, address imbalanced data issues, and foster interdisciplinary collaborations to enhance model interpretability and align better with the complexities of sign language, advancing more inclusive and accurate communication technologies.

Acknowledgment

The completion of this research and the preparation of this journal would not have been possible without the guidance, support, encouragement from various parties. I would like to express my gratitude to the Department of Electrical Engineering and the lecturers for their valuable guidance and support during this research process.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] L. N. Hayati, A. N. Handayani, W. S. G. Irianto, R. A. Asmara, D. Indra, and M. Fahmi, "Classifying BISINDO Alphabet using Tensorflow Object Detection API," *Ilk. J. Ilm.*, vol. 15, no. 2, pp. 358–364, 2023, doi: 10.33096/ilkom.v15i2.1692.358-364.

[2] A. N. Handayani, M. I. Akbar, H. Ar-Rosyid, M. Ilham, R. A. Asmara, and O. Fukuda, "Design of SIBI Sign Language Recognition Using Artificial Neural Network Backpropagation," *2022 2nd Int. Conf. Intell. Cybern. Technol. Appl.*, pp. 192–197, 2022, doi: 10.1109/ICICyTA57421.2022.10038205.

[3] C. Suardi, A. N. Handayani, R. A. Asmara, A. P. Wibawa, L. N. Hayati, and H. Azis, "Design of Sign Language Recognition Using E-CNN," in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, Apr. 2021, pp. 166–170, doi: 10.1109/EIConCIT50028.2021.9431877.

[4] B. Majoro Nthabiseng Eugenia, "Challenges of using sign language interpreting to facilitate teaching and learning for learners with hearing impairment," National University of Lesotho, p. 106, 2021. [Online]. Available at: https://repository.tml.nul.ls/handle/20.500.14155/1619.

[5] E. Warren and J. Call, "Inferential Communication: Bridging the Gap Between Intentional and Ostensive Communication in Non-human Primates," *Front. Psychol.*, vol. 12, p. 718251, Jan. 2022, doi: 10.3389/fpsyg.2021.718251.

[6] M. P. Moeller *et al.*, "Family-Centered Early Intervention Deaf/Hard of Hearing (FCEI-DHH): Foundation Principles," *J. Deaf Stud. Deaf Educ.*, vol. 29, no. SI, pp. SI53–SI63, Feb. 2024, doi: 10.1093/deafed/enad037.

[7] A. Sani and S. Rahmadinni, "Image Processing Based Hand Gesture Detection," *J. Rekayasa Elektr.*, vol. 18, no. 2, pp. 115–124, Jul. 2022, doi: 10.17529/jre.v18i2.25147.

[8] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[9] C. J. Taylor *et al.*, "A Brief Introduction to Chemical Reaction Optimization," *Chem. Rev.*, vol. 123, no. 6, pp. 3089–3126, Mar. 2023, doi: 10.1021/acs.chemrev.2c00798.

[10] R. Muttaqien, M. G. Pradana, and A. Pramuntadi, "Implementation of Data Mining Using C4.5 Algorithm for Predicting Customer Loyalty of PT. Pegadaian (Persero) Pati Area Office," *Int. J. Comput. Inf. Syst.*, vol. 2, no. 3, pp. 64–68, Aug. 2021, doi: 10.29040/ijcis.v2i3.36.

[11] H.-B. Wang and Y.-J. Gao, "Research on C4.5 algorithm improvement strategy based on MapReduce," *Procedia Comput. Sci.*, vol. 183, pp. 160–165, Jan. 2021, doi: 10.1016/j.procs.2021.02.045.

[12] G. S. Reddy and S. Chittineni, "Entropy based C4.5-SHO algorithm with information gain optimization in data mining," *PeerJ Comput. Sci.*, vol. 7, p. e424, Apr. 2021, doi: 10.7717/peerj-cs.424.

[13] K. R. Tsaniyah Zulkarnain, H. Nurramdhani Irmanda, A. Muliawati, and R. Wirawan, "The Effect of Particle Swarm Optimization Feature Selection in Predicting Blood Donor Eligibility Classification at UTD Bekasi City Using Decision Tree C4.5," in *2023 International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS)*, Nov. 2023, pp. 131–136, doi: 10.1109/ICIMCIS60089.2023.10348997.

[14] A. Sadeghzadeh and M. B. Islam, "Triplet Loss-based Convolutional Neural Network for Static Sign Language Recognition," in *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Sep. 2022, pp. 1–6, doi: 10.1109/ASYU56188.2022.9925490.

[15] A. S. M. Miah, J. Shin, M. A. M. Hasan, and M. A. Rahim, "BenSignNet: Bengali Sign Language Alphabet Recognition Using Concatenated Segmentation and Convolutional Neural Network," *Appl. Sci.*, vol. 12, no. 8, p. 3933, Apr. 2022, doi: 10.3390/app12083933.

[16] M. Muhsi, S. Suprapto, and R. Rofiuddin, "Node Selection Method for Split Attribute in C4.5 Algorithm Using the Coefficient of Determination Values for Multivariate Data Set," *J. Penelit. Pendidik. IPA*, vol. 9, no. 7, pp. 5574–5583, Jul. 2023, doi: 10.29303/jppipa.v9i7.4031.

[17] A. N. Handayani, H. W. Herwanto, K. L. Chandrika, and K. Arai, "Recognition of Handwritten Javanese Script using Backpropagation with Zoning Feature Extraction," *Knowl. Eng. Data Sci.*, vol. 4, no. 2, p. 117, Dec. 2021, doi: 10.17977/um018v4i22021p117-127.

[18] T. Wahyudi and T. Silfia, "Implementation of Data Mining Using K-Means Clustering Method to Determine Sales Strategy In S&R Baby Store," *J. Appl. Eng. Technol. Sci.*, vol. 4, no. 1, pp. 93–103, Sep. 2022, doi: 10.37385/jaets.v4i1.913.

[19] E. Cappelli, "Big biomedical data modeling for knowledge extraction with machine learning techniques," Università degli studi Roma Tre, pp, 1-168, 2020. [Online]. Available at: https://arcadia.sba.uniroma3.it/handle/2307/40921.

[20] R. G. Whendasmoro and J. Joseph, "Analysis of the Application of Data Normalization Using Z-Score on the Performance of the K-NN Algorithm," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 872, Aug. 2022, doi: 10.30865/jurikom.v9i4.4526.

[21] I. Izonin, R. Tkachenko, N. Shakhovska, B. Ilchyshyn, and K. K. Singh, "A Two-Step Data Normalization Approach for Improving Classification Accuracy in the Medical Diagnosis Domain," *Mathematics*, vol. 10, no. 11, p. 1942, Jun. 2022, doi: 10.3390/math10111942.

[22] A. Onan, "SRL-ACO: A text augmentation framework based on semantic role labeling and ant colony optimization," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 7, p. 101611, Jul. 2023, doi: 10.1016/j.jksuci.2023.101611.

[23] V. R. Joseph and A. Vakayil, "SPlit: An Optimal Method for Data Splitting," *Technometrics*, vol. 64, no. 2, pp. 166–176, Apr. 2022, doi: 10.1080/00401706.2021.1921037.

[24] H. Wang, M. Yang, and J. Stufken, "Information-Based Optimal Subdata Selection for Big Data Linear Regression," *J. Am. Stat. Assoc.*, vol. 114, no. 525, pp. 393–405, Jan. 2019, doi: 10.1080/01621459.2017.1408468.

[25] L. Mardiana, D. Kusnandar, and N. Satyahadewi, "Discriminant Analysis With K Fold Cross Validation For Water Quality Classification In Pontianak City," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 11, no. 1, pp. 97–102, Jan. 2022, doi: 10.37403/sultanist.v11i1.505.

[26] I. D. Wijaya, A. G. Putrada, and D. Oktaria, "The Use of K-fold Method for Imbalance Data in Hwe and Qpq Classification in Sexual Harassment Tweet Crimes," *eProceedings Eng.*, vol. 8, no. 5, pp. 1–10, Oct. 2021. [Online]. Available at: https://openlibrarypublications.telkomuniversity.ac.id.

[27] M. Li, H. Xu, and Y. Deng, "Evidential Decision Tree Based on Belief Entropy," *Entropy*, vol. 21, no. 9, p. 897, Sep. 2019, doi: 10.3390/e21090897.

[28] V. Matzavela and E. Alepis, "Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100035, Jan. 2021, doi: 10.1016/j.caeai.2021.100035.

[29] F. Hajjej, M. A. Alohali, M. Badr, and M. A. Rahman, "A Comparison of Decision Tree Algorithms in the Assessment of Biomedical Data," *Biomed Res. Int.*, vol. 2022, no. 1, pp. 1–9, Jul. 2022, doi: 10.1155/2022/9449497.

[30] C. Azad, B. Bhushan, R. Sharma, A. Shankar, K. K. Singh, and A. Khamparia, "Prediction model using SMOTE, genetic algorithm and decision tree (PMSGD) for classification of diabetes mellitus," *Multimed. Syst.*, vol. 28, no. 4, pp. 1289–1307, Aug. 2022, doi: 10.1007/s00530-021-00817-2.

[31] A. Rahmadeyan, Mustakim, R. Ocviani, S. Sarah, and F. Ardiansyah, "Classification of COVID-19 Data Using Decision Tree Optimized with Particle Swarm Optimization and Genetic Algorithm," in *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS)*, Jan. 2024, pp. 1–5, doi: 10.1109/ICETSIS61505.2024.10459540.

[32] M. C. Bagaskoro, F. Prasojo, A. N. Handayani, E. Hitipeuw, A. P. Wibawa, and Y. W. Liang, "Hand image reading approach method to Indonesian Language Signing System (SIBI) using neural network and multi layer perseptron," *Sci. Inf. Technol. Lett.*, vol. 4, no. 2, pp. 97–108, Nov. 2023, doi: 10.31763/sitech.v4i2.1362.