BAREKENG: Journal of Mathematics and Its Applications
September 2023 Volume 17 Issue 3 Page 1483–1494
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss3pp1483-1494

CLASSIFICATION OF TODDLER'S NUTRITIONAL STATUS USING THE ROUGH SET ALGORITHM

Izzati Rahmi [1*], Yana Wulandari[2], Hazmira Yozza[3], Mahdhivan Syafwan[4]

[1,2,3,4] *Department of Mathematics and Data Science, Faculty of Mathematics and Natural Science, Andalas University*
Limau Manis, Padang, 25175, Indonesia

*Corresponding author's e-mail: ** ***izzatirahmihg@sci.unand.ac.id***

Article History:

Received: 28th February 2023
Revised: 3rd August 2023
Accepted: 9th August 2023

Keywords:

Nutritional status;
Rough set algorithm;
Decision rules;
Inconsistencies

ABSTRACT

The health and nutrition of children at the age of five are very important aspects in the children's growth and development. An assessment of the nutritional status of toddlers that is commonly used is anthropometry. This study aims to obtain the decision rules used to classify toddlers into nutritional status groups using the rough set algorithm and determine the level of classification accuracy of the resulting decision rules. The index used in this study is the weight-for-age index. Attributes used in this study were the mother's education level, mother's level of knowledge, the status of exclusive breastfeeding, history of illness in the last month, and nutritional status of toddlers. The results of the analysis show that there are 21 decision rules. In this study, the resulting decision rules experience inconsistencies. The selection of decision rules that experience inconsistencies is based on each decision rule's highest strength value. The rough set algorithm can be used for the classification process with an accuracy rate of 86.36%.

How to cite this article:

I. Rahmi, Y. Wulandari, H. Yozza and M. Syafwan., "CLASSIFICATION OF TODDLER'S NUTRITIONAL STATUS USING THE ROUGH SET ALGORITHM," *BAREKENG: J. Math. & App.*, vol. 17, iss. 3, pp. 1483-1494, September, 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

The toddler period is often referred to as the golden age in a person's developmental range. At this time, children experience extraordinary growth physically, motorically, emotionally, cognitively, and psychologically [1]. The health and nutrition of children at the age of five are crucial aspects of children's growth and development. Children's nutritional intake must be completed and balanced to maintain their health and optimize their growth.

In Indonesia, the common method for assessing nutritional status is anthropometry. The use of anthropometry to measure nutritional status can be done by measuring several indicators of a single dimension of the human body. Anthropometric indicators that are usually used to assess nutritional status are weight-for-age, height-for-age, and weight-for-height. The anthropometric indicator of weight-for-age is the most often used because it is easier and more quickly understood by the general public. The weight-for-age indicator classifies the nutritional status of toddlers into severely underweight, underweight, normal, and overweight [2].

Each region in Indonesia has different factors that affect toddler nutrition. One of them is Semen Village in Ngawi Regency, East Java Province. The results of research by Berlina state that four factors affect the nutritional status of toddlers in Semen Village: the mother's level of education, the mother's level of knowledge, the status of exclusive breastfeeding, and the history of illness in the last month [3].

Determining the nutritional status of children under five is usually done manually by comparing standard nutritional status and by measurement result data [4]. Determining nutritional status manually is less practical and susceptible to an accuracy error. Therefore, there is needed a system that can classify the nutritional status of children under five quickly and accurately, even by health workers or the wider society.

One way that can be used to predict whether a toddler is classified as severely underweight, underweight, normal, or overweight is by using a data mining classification approach. Some methods used in data mining classification are Naive Bayes, Support Vector Machines, Decision Trees, Rough Sets, etc [5]. In recent years, the research and applications of rough set theory have attracted more and more researchers' attention. And it is one of the hot issues in artificial intelligence [6]. RST has been used in various research fields, including the internet [7], medicine [8], education [9], economics and business [10], and engineering [11].

The rough set is a mathematical method developed by Zdislaw Pawlak in 1982, which is mainly used to solve data classification problems [12]. The purpose of the rough set algorithm is to obtain decision rules that can be used to classify objects into object classes. The decision rules are presented in a table called the decision table. Several previous studies have shown that the rough set algorithm has the ability to classify objects with high accuracy [13]. In this study, the rough set algorithm was used to classify the nutritional status of toddlers at one of the maternal and children health services (posyandu) in Semen Village, Paron District, Ngawi Regency.

2. RESEARCH METHODS

2.1 Data Sources

This study used secondary data obtained from the results of the research by Berlina [3] about factors that influence toddler's nutritional status at the posyandu in Semen Village, Paron District, Ngawi Regency. The data consisted of 87 observations, of which 75% of the dataset (65 data) was used as training data, and 25% of the dataset (22 data) was used as testing data.

2.2 Research Attributes

This study used condition attributes and decision attributes. The decision attribute used was the nutritional status of children under five based on the weight-for-age index denoted by K. The nutritional status of children under five was divided into four categories: severely underweight, underweight, normal, and overweight. The training data consisted of 1 severely underweight, 19 underweight, 43 normal, and 2 overweight.

The condition attributes used are factors suspected to influence the toddler's nutritional status. This study consisted of two ordinal condition attributes, namely the mother's level of education (A) and the mother's level of knowledge (B), and two nominal conditions attribute namely the status of exclusive breastfeeding (C), and the history of illness in the last month (D).

2.3 Rough Set Algorithm

The rough set can be used in intelligent data analysis and data mining, and as a tool to deal with problems of vagueness and imprecision, which is mainly used to solve data classification problems [14]. The data used in the rough set is usually categorical data. If the data is numeric, then the data must be transformed into categorical data before being analyzed using rough set algorithm. The purpose of the rough set algorithm is to get an estimate of the decision rules for classifying objects stated in a table called the decision table. The process of determining such decision rules is known as reduct discovery. Through the rough set algorithm, reduct as a decision table pattern can build classifiers to categorize new objects.

The rough set algorithm is as follows:

a. Data Selection

b. Decision table formation

The decision table is a table where each row shows the research object, while the columns show the research's conditional and decision attributes. A decision table can also be defined as

$$IS = (U, At = C \cup D, \{V_a | a \in At\}, I_a | a \in At),$$

where U is a finite non-empty set of n objects $\{x_1, x_2, \dots, x_n\}$; At is a finite non-empty set of attributes in which there are two types of attributes, namely condition attribute C, which describes the object, and decision attribute D, which identifies the class of the object; $V_a$is a non-empty set of values $a \in At$; $I_a: U \to V_a$ is a function that maps objects from U to exactly one value in V_a [15].

Generally, when the decision table is the result of observation or measurement results, the data in the table will likely experience inconsistencies. The decision table is said to be consistent if all object pairs with the same conditional value in C also have the same decision value in D. Conversely, the decision table is said to be inconsistent if there is one object pair that has the same conditional value in C but has a different decision value in D [16].

c. Equivalent class formation from each object indiscernibility relation.

In the decision table, an object can have the same value for a condition attribute. This relationship is called indiscernibility (cannot be distinguished). Suppose IS is a decision table and $A \subseteq C$, then the indiscernibility relation $IND(A) \subseteq U \times U$ is defined as

$$IND(A) = \{(x, y) \in U \times U | \forall a \in A, I_a(x) = I_a(y)\}.$$

The indiscernibility relation of the set of all condition attributes is called the equivalent class, which is then used to obtain a discernibility matrix [17], [18].

d. Discernibility matrix dan discernibility function formation

Two objects are said to be discernible if the values of the two objects differ in at least one attribute. The $M(x, y)$discernibility matrix is a symmetric matrix that is, $M(x, y) = M(y, x)$ and $M(x, y) = \emptyset$. Each element of $M(x, y)$ consists of a set of attributes that distinguish object x from object y. So, $M(x, y)$is defined as

$$M(x, y) = \{c \in C | [I_c(x) \neq I_c(y)] \wedge [I_D(x) \neq I_D(y)]\},$$

for $I_D(x) = I_D(y)$, then $M(x, y) = \emptyset$ [19].

From each column of the discernibility matrix, a Boolean function can be formed, which is known as the discernibility function. The discernibility function f_{IS} is defined as

$$f_{IS} = \bigwedge \left\{ \bigvee (M(x, y)) \mid \forall x, y \in U, M(x, y) = \emptyset \right\}$$

e. Formation of a reduct by simplifying the discernibility function

The discernibility function of each column of the discernibility matrix is simplified by using the laws of Boolean algebra. Simplifying the discernibility function means finding another form of equivalent function with a smaller number of terms or operations. The terms resulting from simplifying the discernibility function are called reducts.

f. Formation of decision rules

Decision rules are statements of the form "*if f then g*" represented as $f \to g$. f part is the value of the condition attribute and g part is the value of the decision attribute. In the rough set, decision rules can be drawn from the resulting reduct by observing the table of equivalent classes formed [20].

If the decision table experiences data inconsistencies, the resulting decision rules also experience inconsistencies. To overcome these problems, a quality measure is used to select decision rules that experience inconsistencies. Quality measures are divided into support, strength, accuracy, and coverage [20]. In this study, strength is used as a reference to determine the decision to be used. Suppose *IS* is a decision table. The strength of the decision rule is defined as

$$Strength = \frac{support}{card(U)} \times 100\%,$$

where support is the number of objects that match the condition and decision attributes of the decision rule, while *card(U)* is the total number of objects. Sometimes it is found that several inconsistent rules have the same strength. In this condition, another quality measure is needed and for this purpose the coverage will be used in this study The Coverage of the decision rule is defined as

$$Coverage = \frac{support}{card(g)} \times 100\%,$$

where card(g) is the number of objects in the dataset that satisfy only values decision attribute.

2.4 Classification Performance Measurement

Classification performance measurement is needed to find out how much the level of accuracy of a classification system. Accuracy is the model's true or false level of predictions through a data set [21]. In this study, the classification performance measurement method was carried out by forming a confusion matrix. The confusion matrix compares the predicted classification with the actual classification. The form of the confusion matrix with decision attribute consisting of 4 classes is shown in **Table 1.**

Table 1. Confusion Matrix

Testing Data	Prediction Results			
	1	2	3	4
1	X_{11}	X_{12}	X_{13}	X_{14}
2	X_{21}	X_{22}	X_{23}	X_{24}
3	X_{31}	X_{32}	X_{33}	X_{34}
4	X_{41}	X_{42}	X_{43}	X_{44}

Based on **Table 1**, X_{ij} implies that testing data shows class (i) while the prediction result shows class (j). The observations are predicted correctly when i=j, for i & j = 1, 2, 3, 4 that are observations that located on the diagonal matrix, namely X_{11}, X_{22}, X_{33}, and X_{44}. In contrast, other observations are failed to predict correctly.

3. RESULTS AND DISCUSSION

3.1 Results

a. Decision Table

The decision table in this study is shown in **Table 2**.

Table 2. Decision Table of Training Data

Toddler	A	B	C	D	K
P1	SHS	Good	No	Yes	Severely Underweight
P2	SHS	Good	No	No	Normal
P3	ES	Deficient	Yes	Yes	Underweight
P4	SHS	Good	No	Yes	Underweight
P5	SHS	Good	No	Yes	Underweight
⋮	⋮	⋮	⋮	⋮	⋮
P65	SHS	Deficient	Yes	No	Underweight

Note: SHS= Senior High School ES = Elementary School

Each row in the decision table provides information about the object. For example, toddler P1 in Table 2 informs that if the mother's last education is high school, the knowledge level of the mother is good, the toddler does not get exclusive breastfeeding, and the toddler has a history of illness in the last month, then the nutritional status of toddler P1 is severely underweight.

a. Decision Class Table

After forming the decision table, the next step is to find the indiscernibility relation of the set of condition attributes. The following presents the indiscernibility relation of the condition attribute set.

$$IND(\{A,B,C,D\}) = \{(P1,P4,P5),(P2,P16,P18,P20),(P47,P55,P56,P65),(P11,P63),(P3,P62),(P13,P14,P17,P40,P61),(P6,P8),(P48,P49,P51,P64),(P9),(P12,P21,P23,P54),(P24,P28,P30,P32,P36,P37,P38,P41,P43,P46),(P39,P44),(P60),(P10),(P45,P53),(P50,P52,P58,P59),(P15,P22),(P19),(P25,P31,P33,P35,P42),(P27),(P29,P34,P57),(P7),(P26)\}$$

Based on the indiscernibility relations, 23 equivalent classes were obtained, which are presented in Table 3.

Table 3. Decision Class Table

EC	Toddler	A	B	C	D	K
EC1	$(P1, P4, P5)$	SHS	Good	No	Yes	(Severely Underweight, Underweight)
EC2	$(P2, P16, P18, P20)$	SHS	Good	No	No	(Normal, Underweight)
EC3	$(P47, P55, P56, P65)$	SHS	Deficient	Yes	No	(Normal, Underweight)
EC4	$(P11, P63)$	JHS	Good	Yes	Yes	(Normal, Underweight)
⋮	⋮	⋮	⋮	⋮	⋮	⋮
EC10	$(P12, P21, P23, P54)$	SHS	Good	Yes	No	(Normal)
⋮	⋮	⋮	⋮	⋮	⋮	⋮
EC23	$(P26)$	ES	Sufficient	Yes	No	(Normal)

Table 3 shows objects that have the same condition attribute values but can have different decision attribute alleged values. For example, EC1 consists of toddlers P1, P4, and P5, which are equivalent classes with the same condition attributes, namely toddlers with the mother's last education is high school, the knowledge level of the mother is good, the toddler does not get exclusive breastfeeding and the toddler has a history of illness in the last month, with a suspected nutritional status of being underweight or severely underweight. Furthermore, different class equivalents have at least one different condition attribute value. For example, EC1 and EC2 have one different condition attribute, namely attribute D.

b. Discernibility Matrix

The discernibility matrix is a matrix that contains differences in condition attribute values between one equivalent class and another equivalent class. Table 4 presents the discernibility matrix of equivalent classes.

Table 4. Discernibility Matrix

	EC1	EC2	EC3	...	EC10	...	EC23
EC1	$\emptyset$	$\{D\}$	$\{B,C,D\}$	...	$\{C,D\}$	...	$\{A,B,C,D\}$
EC2	$\{D\}$	$\emptyset$	$\{B,C\}$	...	$\{C\}$	...	$\{A,B,C\}$
EC3	$\{B,C,D\}$	$\{B,C\}$	$\emptyset$	...	$\{B\}$	...	$\{A,B\}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
EC10	$\{C,D\}$	$\{C\}$	$\{B\}$	...	$\emptyset$	...	$\{A,B\}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
EC23	$\{A,B,C,D\}$	$\{A,B,C\}$	$\{A,B\}$	...	$\{A,B\}$	...	$\emptyset$

Based on **Table 3**, the EC1 and EC2 pairs have different condition attribute values, namely attribute D. Therefore, in the first row and second column of the discernibility matrix in **Table 4**, it is written as {D}. The EC10 and EC23 pairs have the same decision attribute values, namely, toddlers have good nutritional status, so the 10th row and 23rd column are written ∅. In the 1st row, the 1st column is written ∅ because EC1 cannot be distinguished from EC1 itself.

c. Discernibility Function and Its Simplification

The next step is to form the reduct by constructing the discernibility function and simplifying by using the laws of Boolean algebra. **Table 5** presents the results of the discernibility function formation and the discernibility function simplification called reduct.

Table 5. Discernibility Function and Reduct

EC	Discernibility Function	Reduct
EC1	$(D) \wedge (B \vee C \vee D) \wedge (A \vee C) \wedge (A \vee B \vee C) \wedge (A \vee C \vee D) \wedge (A \vee B \vee C \vee D) \wedge (C) \wedge (C \vee D) \wedge (B \vee D) \wedge (A \vee B \vee D) \wedge (A \vee D) \wedge (A \vee B)$	$(A \wedge C \wedge D)$, $(B \wedge C \wedge D)$
EC2	$(D) \wedge (A \vee C \vee D) \wedge (C \vee D) \wedge (C) \wedge (B \vee C) \wedge (B) \wedge (A \vee B \vee C \vee D) \wedge (A \vee C) \wedge (A) \wedge (A \vee B) \wedge (A \vee B \vee D) \wedge (A \vee B \vee C)$	$(A \wedge B \wedge C \wedge D)$
EC3	$(B \vee C \vee D) \wedge (A \vee B \vee D) \wedge (B \vee D) \wedge (B) \wedge (B \vee C) \wedge (C) \wedge (A \vee D) \wedge (A) \wedge (A) \wedge (A \vee B \vee C) \wedge (A \vee C) \wedge (A \vee C \vee D) \wedge (A \vee B)$	$(A \wedge B \wedge C)$
$\vdots$	$\vdots$	$\vdots$
EC23	$(A \vee B \vee C \vee D) \wedge (B \vee D) \wedge (A \vee B) \wedge (B) \wedge (A \vee B \vee D) \wedge (A \vee B \vee C) \wedge (A \vee C) \wedge (B \vee C) \wedge (B \vee C \vee D)$	$(A \wedge B)$, $(B \wedge C)$

For example, for EC1 in **Table 5**, the discernibility function of EC1 is the conjunction of (M(x, y)) in column EC1, where (M(x, y)) is the disjunction of all the attributes in EC1. EC1's discernibility function can be written as follows:

$$f_{EC1} = (D) \wedge (B \vee C \vee D) \wedge (A \vee C) \wedge (A \vee B \vee C) \wedge (A \vee C \vee D) \wedge (A \vee C) \wedge (A \vee B \vee C \vee D) \wedge (C) \wedge (C \vee D) \wedge (B \vee C \vee D) \wedge (B \vee D) \wedge (B \vee D) \wedge (A \vee B \vee C) \wedge (A \vee B \vee D) \wedge (A \vee B \vee C \vee D) \wedge (A \vee C \vee D) \wedge (A \vee D) \wedge (A \vee B \vee C \vee D) \wedge (A \vee B \vee D) \wedge (A \vee B \vee D) \wedge (A \vee B) \wedge (A \vee B \vee C \vee D)$$

$$= (D)(B+C+D)(A+C)(A+B+C)(A+C+D)(A+C)\ (A+B+C+D)(C)(C+D)(B+C+D)(B+D)(B+D)(A+B+C)\ (A+B+D)(A+B+C+D)(A+C+D)(A+D)(A+B+C+D)\ (A+B+D)(A+B+D)(A+B)(A+B+C+D)$$

$$= (D)\ (B+C+D)\ (A+C)\ (A+B+C)\ (A+C+D)\ (A+B+C+D)\ (C)\ (C+D)\ (B+D)\ (A+B+D)(A+D)(A+B);\ \textit{Idempotent low}$$

$$= (DB+DC+DD)\ (A\ (A+B+C)+C\ (A+B+C))\ (A\ (A+B+C+D)+C\ (A+B+C+D)+D\ (A+B+C+D))\ (C)\ (C+D)\ (B\ (A+B+D)+D\ (A+B+D))\ (A+D)\ (A+B);\ \textit{Distributive low}$$

$$= (DB+DC+DD)\ ((AA+AB+AC)+(CA+CB+CC))\ ((AA+AB+AC+AD)+(CA+CB+CC+CD)+(DA+DB+DC+DD))\ (C)\ (C+D)\ ((BA+BB+BD)+(DA+DB+DD))\ (A+D)\ (A+B);\ \textit{Distributive low}$$

$$= (DB+DC+D)\ ((A+AB+AC)+(CA+CB+C))\ ((A+AB+AC+AD)+(CA+CB+C+CD)+(DA+DB+DC+D))\ (C)\ (C+D)\ ((BA+B+BD)+(DA+DB+D))\ (A+D)\ (A+B);\ \textit{Idempotent low}$$

$$= (D)\ (A+C)\ (A+C+D)\ (C)\ (C+D)\ (B+D)\ (A+D)\ (A+B);\ \textit{Absorption law}$$

$$= (D)\ (A+C+D)\ (A+C)\ (C)\ (C+D)\ (B+D)\ (A+D)\ (A+B);\ \textit{Commutative low}$$

$$= (DA+DC+DD)\ (A+C)\ (C)\ (C+D)\ (B+D)\ (A+D)\ (A+B);\ \textit{Distributive low}$$

$$= (DA+DC+D)\ (A+C)\ (C)\ (C+D)\ (B+D)\ (A+D)\ (A+B),\ \textit{Idempotent low}$$

$$= (D)(C)\ (B+D)\ (A+D)\ (A+B);\ \textit{Absorption law}$$

$$= (C)(D)\ (B+D)\ (A+D)\ (A+B);\ \textit{Commutative low}$$

$$= (C)(D)\ (A+B);\ \textit{Absorption law}$$

$$= (CDA)+(CDB);\ \textit{Distributive low}$$

$= (ACD) + (BCD)$; *Commutative low*
$= (A \wedge C \wedge D) \vee (B \wedge C \wedge D)$

d. Decision Rules

After obtaining the reduct for each equivalent class, the next step is determining the decision rules. From 23 equivalent classes, 21 decision rules are obtained as follows.

1. $(A \wedge C \wedge D)$
   a. Following are the decision rules obtained from EC1.
   "If the mother's education level = SHS and status of exclusive breastfeeding = no and history of illness in the last month = yes, then the nutritional status of the toddler = severely underweight or underweight."
   b. Following are the decision rules obtained from EC5.
   "If the mother's education level = ES and status of exclusive breastfeeding = yes and history of illness in the last month = yes, then the nutritional status of the toddler = underweight or normal."
   c. Following are the decision rules obtained from EC9.
   "If the mother's education level = SHS and status of exclusive breastfeeding = yes and history of illness in the last month = yes, then the nutritional status of the toddler = underweight."
   d. Following are the decision rules obtained from EC16, EC17, and EC19.
   "If the mother's education level = JHS and status of exclusive breastfeeding = yes and history of illness in the last month = no, then the nutritional status of the toddler = normal."
2. $(B \wedge C \wedge D)$
   a. Following are the decision rules obtained from EC1.
   "If the mother's level of knowledge = good and status of exclusive breastfeeding = no and history of illness in the last month = yes, then the nutritional status of the toddler = severely underweight or underweight."
3. $(A \wedge B \wedge C \wedge D)$
   a. Following are the decision rules obtained from EC2.
   "If the mother's education level = SHS and the mother's level of knowledge = good and status of exclusive breastfeeding = no and history of illness in the last month = no, then the nutritional status of the toddler = normal or underweight."
   b. Following are the decision rules obtained from EC10.
   "If the mother's education level = SHS and the mother's level of knowledge = good and status of exclusive breastfeeding = yes and history of illness in the last month = no, then the nutritional status of the toddler = normal."
4. $(A \wedge B \wedge C)$
   a. Following are the decision rules obtained from EC8.
   "If the mother's education level = ES and the mother's level of knowledge = deficient and status of exclusive breastfeeding = yes, then the nutritional status of the toddler = underweight or normal."
   b. Following are the decision rules obtained from EC18.
   "If the mother's education level = JHS and the mother's level of knowledge = good and status of exclusive breastfeeding = no, then the nutritional status of the toddler = normal."
   c. Following are the decision rules obtained from EC20.

"If the mother's education level = JHS and the mother's level of knowledge = sufficient and status of exclusive breastfeeding = no, then the nutritional status of the toddler = underweight."

d. Following are the decision rules obtained from EC3.

"If the mother's education level = SHS and the mother's level of knowledge = deficient and status of exclusive breastfeeding = yes, then the nutritional status of the toddler = underweight."

5. $(A \wedge B \wedge D)$

a. Following are the decision rules obtained from EC14.

"If the mother's education level = JHS and the mother's level of knowledge = deficient and history of illness in the last month = yes, then the nutritional status of the toddler = underweight."

b. Following are the decision rules obtained from EC4.

"If the mother's education level = JHS and the mother's level of knowledge = good and history of illness in the last month = yes, then the nutritional status of the toddler = normal or overweight."

c. Following are the decision rules obtained from EC17 and EC18.

"If the mother's education level = JHS and the mother's level of knowledge = good and history of illness in the last month = no, then the nutritional status of the toddler = normal."

6. $(A \wedge D)$

a. Following are the decision rules obtained from EC7.

"If the mother's education level = university and history of illness in the last month = yes, then the nutritional status of the toddler = normal and underweight."

b. Following are the decision rules obtained from EC6.

"If the mother's education level = university and history of illness in the last month = no, then the nutritional status of the toddler = overweigh or normal."

7. $(A \wedge B)$

a. Following are the decision rules obtained from EC11 and EC12.

"If the mother's education level = SHS and mother's level of knowledge = sufficient, then the nutritional status of the toddler = normal."

b. Following are the decision rules obtained from EC23.

"If the mother's education level = ES and mother's level of knowledge = sufficient, then the nutritional status of the toddler = normal."

8. $(B \wedge C)$

a. Following are the decision rules obtained from EC13, EC15, EC21, and EC22.

"If the mother's level of knowledge = deficient and status of exclusive breastfeeding = no, then the nutritional status of the toddler = underweight."

b. Following are the decision rules obtained from EC11, EC19, and EC23.

"If the mother's level of knowledge = sufficient and status of exclusive breastfeeding = yes, then the nutritional status of the toddler = normal."

9. $(A \wedge C)$

a. Following are the decision rules obtained from EC21 and EC22.

"If the mother's education level = ES and status of exclusive breastfeeding = no, then the nutritional status of the toddler = underweight".

Some relations, such as $(C \wedge D)$, do not appear in the decision rule because they do not appear as a result of the reduct. Besides, there are several decision rules that experience inconsistencies. Therefore, it is necessary to calculate the strength of each decision rule so that it can choose a good decision rule in classifying data based on the highest strength. **Table 6** presents the strength of each decision rule.

Table 6. Strength of Decision Rule

No	Decision Rule	Strength	No	Decision rule	Strength
1	1.a(severely underweight)	1,54%	16	4.d(Normal)	4,62%
2	1.a(Underweight)	3,08%	17	4.d(Underweight)	1,54%
3	1.b(Underweight)	1,54%	18	5.a(Underweight)	1,54%
4	1.b(Normal)	1,54%	19	5.b(Normal)	1,54%
5	1.c(Underweight)	1,54%	20	5.b(Overweigh)	1,54%
6	1.d(Normal)	16,92%	21	5.c(Normal)	4,62%
7	2.a(severely underweight)	1,54%	22	6.a(Normal)	1,54%
8	2.a(Underweight)	3,08%	23	6.a(Underweight)	1,54%
9	3.a(Normal)	3,08%	24	6.b(Overweigh)	1,54%
10	3.a(Underweight)	3,08%	25	6.b(Normal)	6,15%
11	3.b(Normal)	6,15%	26	7.a(Normal)	18,46%
12	4.a(Underweight)	4,62%	27	7.b(Normal)	1,54%
13	4.a(Normal)	4,62%	28	8.a(Underweight)	10,77%
14	4.b(Normal)	1,54%	29	8.b(Normal)	24,62%
15	4.c(Underweight)	1,54%	30	9.a(Underweight)	6,15%

For example, strength of 1.b (Normal) can be obtained as follow

$$Strength\ (1.b\ Normal) = \frac{support}{card(U)} \times 100\% = \frac{1}{65} \times 100\% = 1{,}54\%$$

There are some inconsistent rules in **Table 6** and for each inconsistent decision rule, a rule with a higher strength will be selected. For example, writing 1.a (severely underweight) means that decision rule 1.a is severely underweight and 1.a (underweight) means that decision rule 1.a is underweight. Because 1.a (underweight) has a higher strength value than 1.a (severely underweight), the decision rule used in object classification is rule 1a(underweight). Furthermore, some inconsistent rules have some strength, so it is necessary to identify the coverage of these decision rules. **Table 7** presents the coverage of decision rules that have same strength. The rule with a higher coverage will be selected as final decision rule. For instance, the coverage for decision rule 1b (underweight) can be calculated as follow

$$Coverage(\ 1.b\ Normal) = \frac{support}{card(g)} \times 100\% = \frac{1}{43} \times 100\% = 2.33\%$$

Table 7. The coverage of decision rules that have same strength.

Decision Rule	Strength	Coverage
1.b(Underweight)	1,54%	5,26%
1.b(Normal)	1,54%	2,33%
3.a(Normal)	3,08%	4,65%
3.a(Underweight)	3,08%	10,53%
4.a(Underweight)	4,62%	15,79%
4.a(Normal)	4,62%	6,98%
5.b(Normal)	1,54%	2,33%
5.b(Overweigh)	1,54%	0,50%
6.a(Normal)	1,54%	2,33%
6.a(Underweight)	1,54%	5,26%

After the decision rules are obtained, the classification accuracy will be calculated using testing data. Classification accuracy measures how well the decision rules obtained classify new data. **Table 8** presents the confusion matrix based on the nutritional status of toddlers at the Posyandu in Semen Village, Paron District, Ngawi Regency.

Table 8. **Confusion Matrix**

Testing Data	Prediction Result			
	Overweigh	Normal	Underweight	Severely Underweight
Overweigh	0	0	1	0
Normal	0	16	1	0
Underweight	0	1	3	0
Severely Underweight	0	0	0	0

Based on **Table 8**, the estimation accuracy value can be determined as follows

$$\begin{aligned} Hit_{Ratio} &= \frac{0+16+3+0}{22} \times 100\% \\ &= \frac{19}{22} \times 100\% \\ &= 86.36\%. \end{aligned}$$

The accuracy of data classification is 86.36%, meaning that out of 22 data testing, 86.36% of the data had been classified correctly.

Discussion

The decision table in this study is the decision table that experiences inconsistencies. This is caused by the value of the condition attribute under five being the same but having a different decision attribute value. One of the cases that caused the inconsistency of the decision table was toddlers P1, P4, and P5 in **Table 2**. Toddlers P1, P4, and P5 all had mothers with Senior High School education levels, good knowledge, no history of exclusive breastfeeding, and a history of illness in the last month. However, toddlers P1 have a decision value of severely underweight, while toddlers P4 and P5 have a decision value of underweight.

Since the decision table belongs to the decision table that experiences inconsistency, the resulting decision rules will also experience inconsistency. This is indicated by decision rules 1.a, 1.b, 2a, 3.a, 4.a, 4.d, 5.b, 6.a, and 6.b. With the inconsistency of decision rules, it will be difficult to determine the nutritional status of children under five. For example, there are toddlers whose mothers have a senior high school education level, have a good knowledge level, have no history of exclusive breastfeeding, and have a history of illness in the last month. By paying attention to the attribute value of the toddler's condition the value of the condition attribute of the toddler, the toddler can be classified as having severely underweight or underweight status by decision rule 1.a. To overcome this problem, the highest strength value of decision rule 1.a will be considered. By observing **Table 6**, the toddler is classified as underweight because it has the highest strength value of 3.08%. There are several cases in the classification process, the value of the condition of children under five satisfies several different decision rules. For example, there are toddlers with mothers with senior high school education levels, deficient mothers' level of knowledge, history of exclusive breastfeeding, and history of illness in the last month. By taking into account the value of the condition attribute of the toddler, the toddler can be classified as having an underweight status by decision rule 1.c or can be classified as having normal or underweight by decision rule 4.d. To overcome this problem, the strength of each decision rule that satisfies will be calculated. By observing **Table 6**, the toddler is classified as having normal nutritional status because it has the highest strength value of 4.62%.

The accuracy results after applying the rough set algorithm to classify the nutritional status of toddlers at the posyandu in Semen Village, Paron District, Ngawi Regency, obtained a quite high accuracy of 86.36%. This shows that the rough set algorithm can give a good classification of toddlers at the posyandu in Semen Village, Paron District, Ngawi Regency, into groups of nutritional status. However, an Overweight person is predicted an underweght person. It is very significant and this condition may occur because of unbalanced data class. RST studies on unbalanced data can be developed into the next research topic.

4. CONCLUSIONS

Applying the rough set algorithm to classify toddlers at the posyandu in Semen Village, Paron District, Ngawi Regency, into groups of nutritional status produces 21 decision rules. The decision table in this study is inconsistent, so the resulting decision rules are also inconsistent. Of the 21 decision rules, nine decision rules experience inconsistency. The selection of decision rules that experience inconsistencies is based on each decision rule's highest strength value. The rough set algorithm can classify toddlers at the posyandu in Semen Village, Paron District, Ngawi Regency, into groups with good classification. This is indicated by the results of the accuracy of 86.36%, which is quite high.

ACKNOWLEDGMENT

The authors would like to thank the Faculty of Mathematics and Natural Sciences, Andalas University, for supporting this research

REFERENCES

[1] Y. Asih, "Hubungan pemberian stimulasi perkembangan balita di posyandu," *J. Ilm. Keperawatan Sai Betik*, vol. XI, no. 2, pp. 211–215, 2015.
[2] E. Budianita and Novriyanto, "Klasifikasi Status Gizi Balita Berdasarkan Indikator Antropometri Berat Badan Menurut Umur Menggunakan Learning Vector Quantization," *Semin. Nas. Teknol. Informasi, Komun. dan Ind.*, vol. 7, no. November, pp. 213–220, 2015.
[3] E. Berlina, "Faktor Yang Mempengaruhi Status Gizi Pada Balita Di Posyandu Desa Semen Kecamatan Paron Kabupaten Ngawi," STIKES Bhakti Husada Mulia Madiun, 2021.
[4] E. N. Candra, I. Cholissodin, and R. C. Wihandika, "Klasifikasi Status Gizi Balita menggunakan Metode Optimasi Random Forest dengan Algoritme Genetika (Studi Kasus: Puskesmas Cakru)," *J. Pengemb. Teknol. dan Ilmu Komput.*, vol. 6, no. 5, pp. 2188–2197, 2022.
[5] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, "Metode-metode Klasifikasi," *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
[6] Z. Qinghua, X. Qin, and W. Guoyin, "A survey on rough set theory and its application," *A Surv. rough set theory its Appl. CAAI Trans. Intell. Technol.*, vol. 1, no. 4, pp. 323–333, 2016, doi: 10.1016/j.trit.2016.11.001.
[7] A. Amin, S. Anwar, A. Adnan, M. A. Khan, and Z. Iqbal, "Classification of cyber attacks based on rough set theory," *2015 1st Int. Conf. Anti-Cybercrime, ICACC 2015*, pp. 1–6, 2015, doi: 10.1109/Anti-Cybercrime.2015.7351952.
[8] A. Burney and Z. Abbas, "Applications of Rough Sets in Health Sciences and Disease Diagnosis," *Recent Res. Appl. Comput. Sci.*, pp. 153–161, 2015, doi: 10.1166/jctn.2021.9720.
[9] E. Ramentol, J. Madera, and A. Rodríguez, "Early detection of possible undergraduate drop out using a new method based on probabilistic rough set theory," in *Studies in Fuzziness and Soft Computing*, vol. 377, Springer International Publishing, 2019, pp. 211–232.
[10] J. Becker, A. Radomska-Zalas, and P. Ziemba, "Rough set theory in the classification of loan applications," *Procedia Comput. Sci.*, vol. 176, pp. 3235–3244, 2020, doi: 10.1016/j.procs.2020.09.125.
[11] J. Qu, X. Bai, J. Gu, F. Taghizadeh-Hesary, and J. Lin, "Assessment of rough set theory in relation to risks regarding hydraulic engineering investment decisions," *Mathematics*, vol. 8, no. 8, p. 1308, 2020, doi: 10.3390/MATH8081308.
[12] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, 1982, doi: 10.1007/BF01001956.
[13] J. G. Bazan, H. S. Nguyen, S. H. Nguyen, P. Synak, and J. Wróblewski, "Rough Set Algorithms in Classification Problem," in *Rough Set Methods and Applications*, L. Polkowski, S. Tsumoto, and T. Y. Lin, Eds. Heidelberg: Physica, 2000, pp. 49–88.
[14] R. Kalaivani, M.V. Suresh, and A. Srinivasan, "Study of Rough Sets Theory and It's Application Over Various Fields," *J. Appl. Sci. Eng.*, vol. 3, no. 1, pp. 447-455, 2017.
[15] D. Q. Miao, Y. Zhao, Y. Y. Yao, H. X. Li, and F. F. Xu, "Relative reducts in consistent and inconsistent decision tables of the Pawlak rough set model," *Inf. Sci. (Ny).*, vol. 179, no. 24, pp. 4140–4150, 2009, doi: 10.1016/j.ins.2009.08.020.
[16] X. Li, "Attribute Selection Methods in Rough Set Theory," San Jose State University, 2014.
[17] S. M. Abbas, K. A. Alam, and S. Shamshirband, "A soft-rough set based approach for handling contextual sparsity in context-aware video recommender systems," *Mathematics*, vol. 7, no. 8, pp. 1–36, 2019, doi: 10.3390/math7080740.
[18] J. Zhang, X. Zhang, and W. Xu, "Lower Approximation Reduction Based on Discernibility Information Tree in Inconsistent Ordered Decision Information Systems," *Symmetry.*, vol. 10, no. 7, pp. 1–16, 2018, doi: 10.3390/sym10120696.
[19] A. Skowron and C. Rauszer, "The Discernibility Matrices and Functions in Information Systems," *Intell. Decis. Support*, vol. 11, pp. 331–362, 1992, doi: 10.1007/978-94-015-7975-9_21.
[20] S. Chen, Z. Jiang, C. Duan, and X. Dai, X. Mi, "A Decision Analysis Method Based on Rough Set," Advances in Intelligent Systems Research., vol. 161, pp. 50-56, 2018, doi: 10.2991/tlicsc-18.2018.8
[21] M. Awad and R. Khanna, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Berkeley: Apress, 2015.