Received: February 14, 2023. Revised: March 13, 2023.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Rules Determination Based on Time-Series Data to Classify Unsupervised Cases Based on Fuzzy Expert System

Feby Artwodini Muqtadiroh[1] **Tsuyoshi Usagawa[2]** **Riris Diana Rachmayanti[3]**
Supeno Mardi Susiki Nugroho[4] **Eko Mulyanto Yuniarno[4]** **Mauridhi Hery Purnomo[4]***

[1]Department of Electrical Engineering, Dept. of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
[2]Department of Computer Science and Electrical Engineering, Kumamoto University, Kumamoto, 860-8555, Japan
[3]Department of Health Promotion and Behavioral Sciences, Faculty of Public Health, Universitas Airlangga, Surabaya, 60115, Indonesia
[4]Department of Electrical Engineering, Dept. of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
* Corresponding author's Email: hery@ee.its.ac.id

Abstract: Classification of uncertain conditions requires computational modeling to obtain exact non-vague results for making the right decision, such as opening and closing school cases during a pandemic. We cannot rely solely on normative and textual government regulations because of numerous constraints and uncertainty in implementation. Unsupervised classification techniques can deal with such issues without needing prior references that contain definitive hesitancy. This motivates us to use a fuzzy system based on knowledge-based composition rules for complex problems such as the dynamics of COVID-19 because of its ability to adapt to changes and uncertainties. Therefore, we construct rules based on knowledge about COVID-19 to the issue of opening/closing schools using three fuzzy approaches: conventional fuzzy, intuitionistic fuzzy system (IFS), and fuzzy c-means (FCM). We can demonstrate a correlation between the number of school openings and the COVID-19 dynamics by utilizing the fuzzy approach to reduce the degree of hesitance. Experiments on available public time-series datasets demonstrate that the IFS is more efficient in forming rigidly distinct two classes. The results indicate that the accuracy of IFS is 99.47%, FCM is 91.28, and conventional FS is 84.33%, including the IFS silhouette score, which is higher than the others, at 0.91 or closer to 1, indicating excellent classification results. IFS is less superior in running time, while FCM is the fastest. This is because there are multiple stages in the IFS by considering non-membership functions.

Keywords: Unsupervised, Classification, Fuzzy system, Time-series data, Rules determination, Hesitant degree.

1. Introduction

Uncertainty always attracts the attention of researchers to resolve using computation-based robust and advanced techniques, one of which is the fuzzy logic approach. To reduce the uncertainty, the fuzzy logic theorem can precisely formulate the natural language, which utilizes mathematic-based formal tools for efficient information processing, including modeling complex phenomena [1]. For instance, the COVID-19 phenomenon has significantly disrupted life in all aspects. In this study, we investigate the incident of learning loss caused by school closures during the COVID-19 pandemic [2].

It is true that schools have opened up all around, but the trend is based on dogma or government statements that appear to be a test run with uncertain future developments. If there are numerous cases of death, it will be closed. The opening and closing formulas have not been thoroughly researched up to this point. Therefore, we are interested in modeling the opening and closing status using the massive COVID-19 data. We see not only the current potential but also the possibility of an uncertain future condition if it occurs suddenly and has a high-risk

Received: February 14, 2023. Revised: March 13, 2023.

impact.

Up until now, discussions about opening schools have been presented as a framework with standardized practices and action plans from UNICEF's guidelines [3]. So, when exactly and which schools in which areas can be opened during the outbreak have not been determined with certainty. We will predict the region covering several schools to declare safe reopening based on the actual value of the movement of COVID-19 cases. We call this an unsupervised classification because there is no prior reference or definitive uncertainty. We will develop a deterministic rule to represent new knowledge about school openings based on the characteristics of COVID-19 dynamics rather than simply implementing a fuzzy approach.

Fuzzy systems have been widely used in research across a wide range of fields, including image segmentation, text classification, signal processing, and successful time-series prediction [4]. Fuzzy logic is effective, as reported by [5, 6], especially for time-series classifications showing outperforming results. However, the research by [7] contested that conventional fuzzy logic could not cope with problems of uncertainty in the real world, while intuitionistic fuzzy logic accurately modeled such uncertainty and offered better performance. Ascertained by [8] in his brief explanation, which proves that the intuitionistic fuzzy approach contributes more accurately to describing expert knowledge. Typically, in the case of complicated problems with high hesitancy and ambiguity under unprecedented conditions, a collection of intuitionistic fuzzies offers better strengths compared to those of the conventional fuzzy approach [9].

However, the use of algorithms based on fuzzy systems (FS), particularly extended in the form of intuitionistic fuzzy systems (IFS), is mostly studied and implemented in the medical field, where accuracy is required (low error rate/low uncertainty). As in the case of [10], They investigated the use of intuitionistic fuzzy for grouping breast cancer. [11] is also conducting research in the health sector that uses IFS for brain image segmentation. Whereas [12] conducted a study demonstrating the queuing model using the fuzzy technique in two classes. The findings indicate that the intuitionistic fuzzy approach is more adaptable.

Unfortunately, all the studies mentioned above use well-defined rules, explicit regulations, and quantifiable value boundaries. Even though fuzzy is well-known for its model accuracy due to the use of a rule-based system as a knowledge base in making the proper decisions, fuzzy is not without its flaws [1]. As a result, we propose a rule formulation in fuzzy encapsulation in this study to form an empirical grouping of two classes. Furthermore, because few studies use time-series data for cases that have never occurred, there is no generalizable deterministic approach.

Another fuzzy-based unsupervised classification technique is well-known for its effectiveness in forming well-separated groups: Fuzzy C-Means (FCM) [13]. FCM has shown to perform better than K-Means for breast cancer image segmentation [14]. Analysis by [15] also strengthened the justification that FCM can perform optimal partitioning, as indicated by the lower squared error criterion function compared to K-Means. However, the rule definition is not disclosed in the paper.

In terms of fuzzy-based rule formulation, it was first revealed by [16] in 1995, who classified fuzzy controllers into two classes based on linear linguistic rules. The study demonstrated nonlinear aspects by providing general mathematical formulas as a substitute for linguistic tables. Unfortunately, the developed rules are not used with real-world data. As a result, the impact of established rules has yet to examine.

The study by [17] also claimed that fuzzy rules could provide the appropriate functional relationship between input and output variables. The study operated six types of real data to build rule settings. As a result, the approach can produce satisfactory results without overfitting. This demonstrates that fuzzy has enormous power, allowing for modifying fuzzy rules. However, most of the studies only focus on the membership function (MF). It is also important to consider the non-membership function, which has the ability to significantly reduce the degree of hesitation, as revealed by [12]. We aim to compare the effectiveness of both conventional fuzzy systems and intuitionistic fuzzy systems for this reason.

A study by [18] proposed rule-based reasoning on intuitionistic fuzz to create a more optimal automated trading system. The research objective is nearly identical to our research objective of assisting in the decision to buy or sell stocks. Unfortunately, the observed investment period is only up to $t+9$. Our study will group the school opening and closing status for long-term prediction. This is because we use a large amount of time-series data from March 23, 2020 to October 31, 2022, so the training process is significant to provide more optimal fuzzy-based modeling results that are believed to be true.

This drives us to conduct experiments to create rules for unsupervised binary classification using the fuzzy technique (both conventional fuzzy and intuitionistic fuzzy). We will analyze three approaches; conventional fuzzy expert systems, intuitionistic fuzzy methods, and FCM, which are trusted to minimize the degree of hesitancy to support anticipative decision-making processes.

Therefore, the contribution of our study is as follows:

1. For a situation with no rigid standards, we can develop deterministic rules using a fuzzy system for unsupervised classification spanning from expert knowledge conceptualization to government regulations.
2. By employing fuzzy logic definitions, we can provide expected results in grouping large time-series data to separate significantly into two classes.
3. We can demonstrate a correlation between the number of school openings and the COVID-19 dynamics by utilizing the fuzzy concept to reduce the degree of hesitance so that the predicted outcomes can approximate the expected result.
4. We believe that our work in unsupervised classification and descriptive analysis will be useful in making the right decision in the future.

2. Preliminaries

Uncertainty, ambiguity, misrepresentation of information, and partial truth can all be mathematically represented using fuzzy set theory. The main component in fuzzy set theory is the membership function defining an object's degree of closeness to specific attributes [19-21].

2.1 Fuzzy rules of inference

The structure of a fuzzy system generally consists of a fuzzifier block, a fuzzy reasoning block that involves the knowledge base as the basis for decision-making, and a defuzzification block to convert linguistic to numerical values.

The knowledge representation comes from the systematic rule of fuzzy logic. First, we determine the membership function as Definition 1.

Definition 1. X is a set. A fuzzy subset A of X is characterized by a fuzzy membership function of $\mu_A[X] \rightarrow [0,1]$ related to each point $x \in X$ with a membership degree of $\mu_A(x) \in [0,1]$. We define $\mu_A(x)$ as a membership function of a fuzzy subset of real numbers close to zero. The membership function to represent a set of real numbers close to a certain number α shown in Eq. (1):

$$\mu_A(x) = \frac{1}{1+10(x-\alpha)^2} \tag{1}$$

While FS simply establishes a membership function, IFS theory inquiries about a membership function, a non-membership function, and a degree of hesitation [20, 22]. IFS is often called a hesitant fuzzy set applicable to decision-making problems. Therefore, we also propose the IFS approach to generalize a more robust solution to the uncertainty and become attractive to researchers [23, 24].

Definition 2. The value of variable X is fixed. The IFS of A in X is defined in the following formulation:

$A = \{(x,\ \mu_A(x),\ \lambda_A(x)) \mid x \in X\}$, where function $\mu_A(x)$ is a membership function, while $\lambda_A(x)$ represents the non-membership degree of the element $x \in X$ to set A. In general, we present the function as $\mu_A : X \rightarrow [0,1]$ and $\lambda_A : X \rightarrow [0,1]$, while the degree of hesitation or degree of uncertainty is denoted as:

$$\pi_A(x) = 1 - \mu_A(x) - \lambda_A(x) \tag{2}$$

As a result, we have the components of a membership degree $\mu_A(x)$, non-membership degree $\lambda_A(x)$, and uncertainty degree $\pi_A(x)$.

The total calculation representing IFS is formulated as $\mu_A(x) + \lambda_A(x) + \pi_A(x) =$ 1/10/100/1000 and so forth, consistently. Researchers can choose any total value (1 or 10 or 1,000, or other values) based on their preference to obtain the expected final weight.

The membership function in this study is a list of schools ready to open, while the non-membership function is a list of schools to close. Suppose we chose ten (10) for the overall fuzzy value based on reviews pertinent to our investigation [25]. Thus, the degree of hesitation will be as follows:

$$\pi_A(x) = 10 - \mu_A(x) - \lambda_A(x) \tag{3}$$

We determine that X_1, X_2, and X_n are linguistic variables from the fuzzy set A (premises) with labels L_{A1}, L_{A2}, … L_{An}, and L_B is the linguistic variable from the output (conclusion). Then the fuzzy rule determination can be stated as follows:

If (X_1: L_{A1}) **AND** (X_2: L_{A2})…**AND** (X_n: L_{An}) THEN (Y_1: L_{B1})

The **AND** conjunction for the premise of a fuzzy rule represented by t-norm T in a membership function is then written as follows:

$$\mu_A(x) = \mu_{A_1}(x_1) \star T\ \mu_{A_2}(x_2) \star T\ ...\ \mu_{A_n}(x_n) \star T \tag{4}$$

Definition 3 is a way to alter the membership of the intuitionistic fuzzy system. *X* is a universe set. We write $\beta \in f$, where *f* is a collection of fuzzy sets in *X*, while θ : $X \rightarrow [0,1]$ and β: $X \rightarrow [0,1]$, so *f*: $[0,1]^2 \times [0,1] \rightarrow T^\star$.

Therefore,

$$f(x,\theta,\beta) = (f_\mu(x,\theta,\beta), f_\lambda(x,\theta,\beta)) \text{ and}$$
$$f_\mu(x,\theta,\beta) = (x - x\theta\beta) = x(1 - \theta\beta)$$
$$f_\lambda(x,\theta,\beta) = 1 - \theta\beta - f_\mu(x,\theta,\beta)$$

Following the definition of the rules, the solution is presented in the form of a triangular function. By implementing Eq. (3), we set a fuzzy value of 10. Thus, the membership function is formulated using the linguistic rules listed below:

- If the value is very low/very bad: 1, 8, 1
- If the value is low/bad: 3, 6, 1
- If the value is fair: 5, 4, 1
- If the value is good: 7, 2, 1
- If the value is very good: 9, 0, 1

2.2 Fuzzy C-means

This section reviews another unsupervised classification technique called fuzzy C-means (FCM). FCM offers the benefit of automatically and iteratively grouping points into clusters. It returns a low membership value when the point is far from the centroid, implying that the cost function is low [26]. FCM is a configurable algorithm for grouping large amounts of sequential data due to its effectiveness in mapping the input space to the output space. FCM is one of the most well-known approaches [13].

We must identify the factors that influence the formation of opening/closing classes. The variables in this study are factors that contribute to the growth of the COVID-19 dynamics. For the iteration of cluster formation is based on the number of zones for COVID-19 transmission movement. This study becomes interesting by predicting the movement of COVID-19 on the decision to open school.

As for the weighted fuzzy values: 0, 0.5, 1 (closed, partially opened, and opened). The weighted data can then be placed into the X matrix to perform the FCM processing with the following parameters: the number of clusters (c), power (e), maximum iteration is the number of the observed zone where the school is located (*Max*), smallest expected error (*err*), the initial objective function (OF = 0), and initial iteration (t = 1).

Thus, the following step is to create a *U* matrix with the elements of the membership function (μ_{ik}); where i is the quantity of observation data (daily COVID-19 cases), and k is equal to 2. The cluster center (V_{kj}) will be determined using the newly generated *U* matrix. Members of each cluster can be determined by calculating the distance from the data point to the cluster center:

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^e * X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^e} \tag{5}$$

$$\mu_{ik} = \frac{\left[\sum_{k=1}^{c}(X_{ij}-V_{kj})^2\right]^{\frac{-1}{e-1}}}{\sum_{k=1}^{c}\left[\sum_{k=1}^{c}(X_{ij}-V_{kj})^2\right]^{\frac{-1}{e-1}}} \tag{6}$$

The iteration will continue to repeat until the intended value is fulfilled: if ($|OF_t$ - $OF_{t-1}| <$ *err*) or (t>*Max*).

3. Research method

Fig. 1 shows the unsupervised classification model we propose. We focus on how the method works, especially by suggesting a deterministic rule for controlling schools' safe opening and closing during a pandemic. We summarize fuzzy-based rule-setting for unsupervised classification in Algorithm 1. We believe this algorithm can be applied to analyze the decision-making regarding the categorization issues that do not yet have a structured notion of the dynamic and large amounts of time-series data.

We will demonstrate that there must be a correlation between the number of school openings and the COVID-19 dynamics by utilizing the fuzzy approach, which has the capability to reduce hesitancy so that the predicted outcomes can approximate the expected result.

3.1 Accumulative data processing

First, we focus on consistently processing the daily into weekly time-series data based on the standards of the Indonesian government [27]. Pre-processing tasks allow us to reduce the number of features in dimension and size by extracting noisy or anomalous data. This is to ensure that the data inputs in the fuzzy process are free of waste. We implement algorithms for detecting outliers and preventing data irregularities using if-then rules in Python programming. The public data are taken from the official website of the Surabaya City of Indonesia, consisting of the daily cases in 154 zones from March 23, 2020, to October 31, 2022 [28]. The essential features to be further processed as variables in the fuzzy unsupervised classification process are:

Algorithm 1: Rule determination on fuzzy unsupervised classification

Input: Recovery Increase Rate (RI) and Case Decrease Rate (CD)
Output: Status of Classification

Define: the degree of closed (*op_closed*), degree of partially opened (*op_partopened*), degree of opened (*op_opened*)
for t = 1 to end of the data, do
    Calculate the degree of membership function
    for each rule evaluation
        If *f_case* == low AND *f_recov* == low THEN partially opened (cr_1)
        If *f_case* == low AND *f_recov* == medium THEN opened (cr_2)
        If *f_case* == low AND *f_recov* == high THEN opened (cr_3)
        If *f_case* == medium AND *f_recov* == low THEN closed (cr_4)
        If *f_case* == medium AND *f_recov* == medium THEN partially opened (cr_5)
        If *f_case* == medium AND *f_recov* == high THEN opened (cr_6)
        If *f_case* == high AND *f_recov* == low THEN closed (cr_7)
        If *f_case* == high AND *f_recov* == medium THEN closed (cr_8)
        If *f_case* == high AND *f_recov* == high THEN partially opened (cr_9)
    end for

    R_{closed} = cr_8 OR cr_7 OR cr_4
    R_{opened} = cr_6 OR cr_3 OR cr_2
    $R_{partopened}$ = cr_9 OR cr_5 OR cr_1

    ag_closed = R_{closed} AND (((cr_7 OR cr_4) OR cr_8) OR *op_closed*)
    ag_partopened = $R_{partopened}$ AND (((cr_1 OR cr_5) OR cr_9) OR *op_partopened*)
    ag _opened = R_{opened} AND (((cr_2 OR cr_3) OR cr_6) OR *op_opened*)

    Aggregate *ag_closed* AND *ag_ partopened* AND *ag_opened* → aggregate_result

    Defuzzification on aggregate_result → readiness_level
    If readiness_level <= 2.5 THEN 0
    Else 1
end for

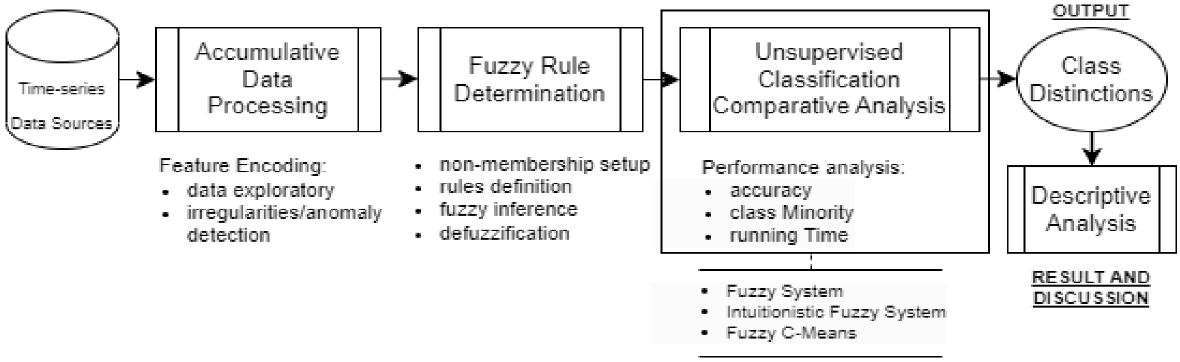

Figure. 1 Proposed method

1. *f_case*: a decrease in cases in the last three consecutive weeks.
2. *f_mortal*: a decrease in the number of deaths in the last three consecutive weeks.
3. *f_hosp*: a decrease in the number of hospitalizations in the last two consecutive weeks.
4. *f_susp*: a decrease in the number of suspected cases in the last two consecutive weeks.
5. *f_recov*: an increase in the number of daily recovered.

Having obtained accumulative sequential data from unstructured sources to be well classified and consistent, we then move to the classification task for school opening and closing.

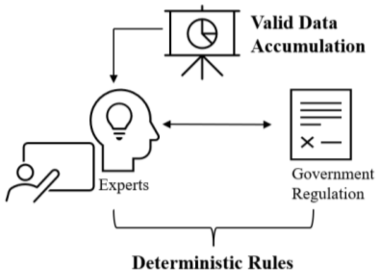

Figure. 2 An overview to construct the rule determination

3.2 Fuzzy rule determination

We emphasize that the rules we utilized are based on expert knowledge aligned with the accumulated data that has been analyzed and has been mapped against government regulations [29] through focus group discussion (FGD). The rules are then mapped to three categories: closed, partially opened, and fully opened, as shown in Fig. 2. We determined two dominant variables/factors: case decrease (CD) during the last three successive weeks and recovery increase (RI), which are very influential in outbreak instability based on the extensive deliberation process with the experts in the previous research [30].

There are nine rules for further processing with fuzzy using either conventional fuzzy classification, intuitionistic fuzzy, or FCM. The nine rules are depicted in Fig. 3.

Next is the rule aggregation process, which employs the AND and OR operators. The AND, OR, and NOT operations provide a fuzzy way of interpreting crisp values by considering the distance of data points or objects to the membership set space. De Morgan demonstrated the utility of AND and OR operators for measuring minimum and maximum values. The decision function can be realized using the AND and OR aggregation and the fuzzy geometry interpretation. Then, the defuzzification calculations are used to optimize the decision.

The defuzzification process changes the set of fuzzy outputs into non-vague values. We adopt the centroid method of the center of area (CoA), commonly known as the center of gravity (CoG).

While the center of sums (CoS) to detect overlapping occurs twice [31]. In contrast, the CoA can effectively detect overlapping in a single run despite its computationally complex calculation, as presented in Eq. (7). The study also clearly demonstrated that the final calculation using the CoA/CoG method offered better performance.

$$A^* = \frac{\int_{\alpha}^{n} A.max_n \, \mu_A{}^n A \; da}{\int_{\alpha}^{n} max_n \, \mu_A{}^n A \; da} \tag{7}$$

In this process, we implement the rules under the Indonesian National Agency for Disaster Management [32] to confirm the classification results:

- High-Risk Zone: 0-2.5
- Low-Risk Zone: > 2.5

The high-risk zone is consequently prohibited from engaging in outdoor activities or mobility, including learning activities.

3.3 Comparative analysis

For the last step, we conduct a comparative analysis of the three fuzzy system-based approaches. We measure the accuracy of the unsupervised classification result, the class minority formation, the silhouette score, and the running time to execute each algorithm. We also reveal the evidence as a result of the establishment of government policies. We compare the real-world conditions over a certain period based on our time-series data modeling.

The measurements to calculate the accuracy of conventional fuzzy or intuitionistic fuzzy as follow:

1. First, we compute the error value by subtracting the reference value (real value) from the measurement value (predicted value).
2. The percentage error is calculated by comparing the error value to the real value and multiplying it by 100%.
3. To determine accuracy, subtract the percentage error value from 100%.

When the accuracy over the observation period is close to the actual value, the results of the two class classifications in school opening and closing decisions can be declared relevant/accurate.

Meanwhile, the silhouette score (SS) calculation will be used to compare the outcomes of the FCM to the two fuzzy techniques [33]. The first step in calculating SS is to find the average distance between the *i-th* data and all data in the same cluster, known as $a(i)$. While $b(i)$ is the minimum distance value in i-th data mean to all data in different clusters.

$$SS\ (i) \coloneqq \frac{b(i)-a(i)}{\max\ \{a(i)-b(i)\}} \tag{13}$$

The value of $SS\ (i)$ ranges from -1 to 1,

$SS\ (i) \approx 1$ => getting closer to 1 indicates that the data can be properly classified.

$SS\ (i) \approx 0$ => getting closer to 0 indicates that the data is split between two clusters.

$SS\ (i) \approx -1$ => getting closer to -1 indicates that

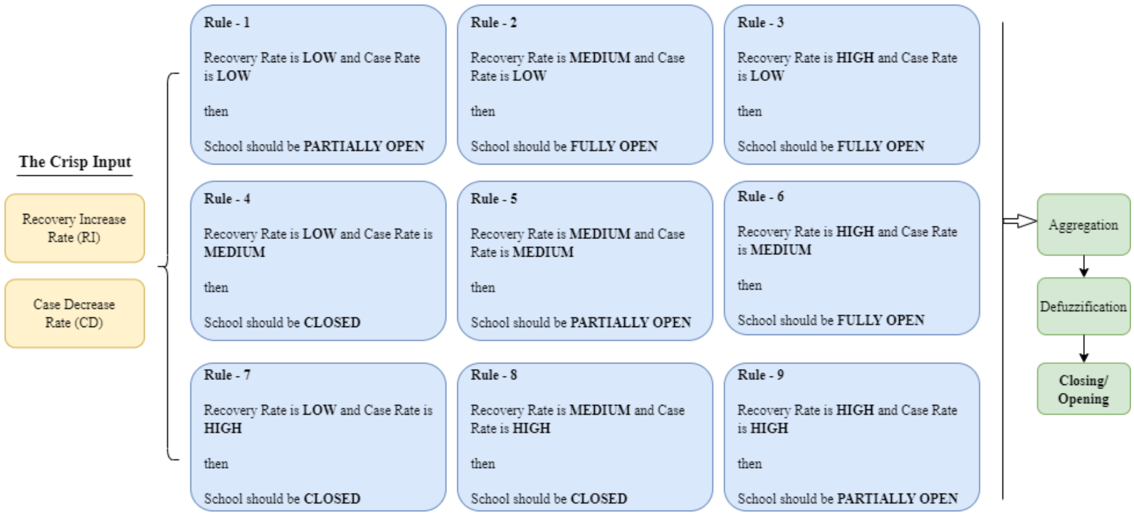

Figure. 3 The Fuzzy rule determination

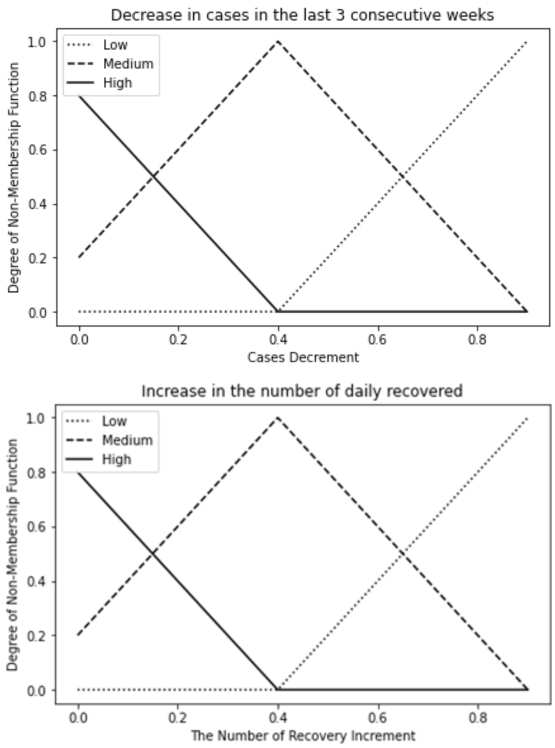

Figure. 4 Non-membership value of the fuzzy

the data is classified as weak.

4. Experiment result and discussion

This section will discuss the result of the unsupervised classification to the FS, IFS, and FCM. We determine the school opening and closing category in the range [0, 1]; when the value is 0, the school is closed, and when it is 1, it is opened.

We define the low-risk, medium-risk, and high-risk zones into triangular fuzzy values subject to rules with the linguistic value; low [0, 0, 0.5], medium [0, 0.5, 1], and high [0.5, 1, 1].

The linguistic values are applied in each variable.

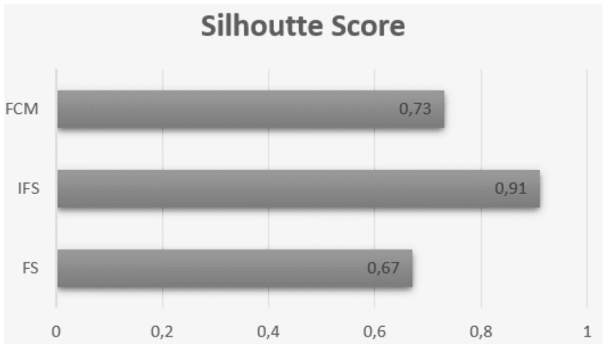

Figure. 5 Silhouette score result

Consequently, the fuzzy triangular numbers and the aggregation functions are presented as *opPos = fuzz.trimf(xop, [5, 10, 10])* to accommodate Eq. (3) for the defuzzification results are maximally in range 10 is similar to the previous research on ranking the region [25]. This is also agreed upon by experts.

Then we can implement Definition 3 to change the FS value into the IFS one; it describes non-membership values as presented in Fig. 4, showing an example IFS non-membership value for *f_case*: a decrease in cases in the last three consecutive weeks (upper) and *f_recov*: an increase in the number of daily recovered (lower).

Having calculated the fuzzification process and the aggregation rule, we must obtain the value to generate the final decision which complies with the categories as stated by the government [32] and described below:

- *f_case*: a decrease in cases in the last three consecutive weeks (weight: 7.5%).
  Formula 1 = Scoring * 0.075.
- *f_mortal*: a decrease in the number of deaths in the last three consecutive weeks (weight: 10%).
  Formula 2 = Scoring * 0.1.

Received: February 14, 2023. Revised: March 13, 2023.

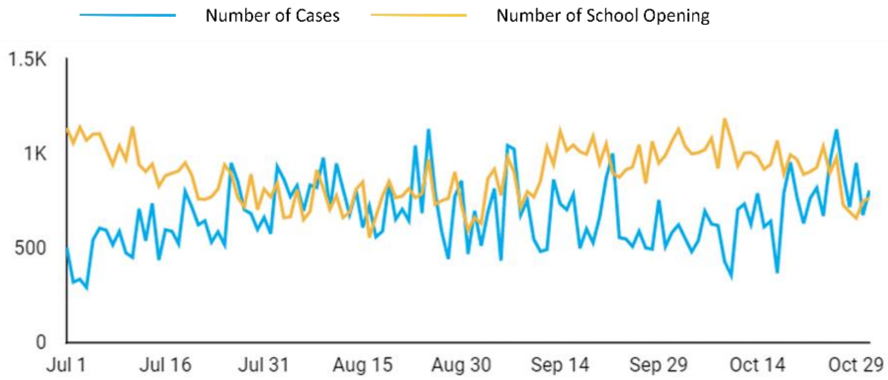

Figure. 6 The correlation between the number of school openings and the COVID-19 dynamics

Table 1. Performance evaluation

	Actual		FCM		FS		IFS	
Status	**Number of Schools**	**%**	**Number of Schools**	**%**	**Number of Schools**	**%**	**Number of Schools**	**%**
Opened	119,505	100%	109,084	91.28%	100,777	84.33%	118,871	99.47%
Closed	-	-	10,421	**8.72%**	18,728	**15.67%**	634	**0.53%**
Total	119,505	100%	119,505	100%	119,505	100%	119,505	100%
Average Value of Fuzzy	-	-	3.98		3.72		4.08	
Running Time	-	-	4.62 seconds		38.25 seconds		16.28 seconds	

Table 2. Opening-closing status derived from the IFS (the best approach)

Date	Zone	*f_case*	*f_mortal*	*f_hosp*	*f_susp*	*f_recov*	Value	Status
2022/07/19	Zone 1	94	15	2	264	0	4,5	Opened
				…				
2022/07/18	Zone 2	83	10	2	296	0	1,85	Closed
2022/07/15	Zone 3	94	17	5	115	2	3,65	Opened
2022/07/16	Zone 3	94	19	0	292	1	4,6	Opened
2022/07/17	Zone 3	78	12	3	263	1	1,96	Closed
				…				
2022/07/20	Zone 6	97	39	4	44	2	4,5	Opened
2022/07/21	Zone 6	88	12	2	302	0	1,79	Closed
2022/07/22	Zone 6	76	13	3	151	2	2,01	Closed
				…				
2022/10/30	Zone 154	46	4	3	113	23	3.55	Closed

- *f_hosp*: a decrease in the number of hospitalizations in the last two consecutive weeks (weight: 5%.) Formula 3 = Scoring * 0.05.
- *f_susp*: a decrease in the number of suspected cases in the last two consecutive weeks (weight: 7.5%).
  Formula 4 = Scoring * 0.075
- *f_recov*: an increase in the number of daily recovered (weight: 7.5%).
  Formula 5 = Scoring * 0.075

The formula multiplication for all variables is then summed to categorize COVID-19 risk in each school in the specific zone. The school is opened if the result is > 2.5. We generate 1,048,575 data rows of information about school opening and closing per day among all schools in 154 zones all over Surabaya City, Indonesia.

Classification formation and Silhouette score

In April 2022, the Mayor of Surabaya city, Indonesia, explicitly supported the format of face-to-face learning on full quota. Yet, the schools are instructed to consistently maintain the health protocols when ready to 100% reopen.

Following those instructions, we present our prediction results for classifying school opening and closing from the fuzzy-based modeling; FS, IFS, and FCM.

Therefore, Table 1 shows the performance indicating the total number of schools that may open or close based on implementing the classification algorithm against the actual conditions (119,505 schools in Surabaya city, Indonesia). Based on the CoA calculation, the low-risk values are those larger than 2.5 then the schools can reopen safely. The better the classification, the higher the average value. The IFS technique offers the highest average value among others, which is 4.08, shown in the second line from the last row of Table 1.

Meanwhile, we utilize the SS calculation with the outcomes shown in Fig. 5 to evaluate the effectiveness of the two classes (opening and closing) created by the fuzzy-based rule determination. The IFS achieves the best grouping compared to other approaches, with an SS value of 0.91 or closer to 1, indicating excellent classification outcomes.

Running time

We also highlight the running time to generate the school opening and closing status, as shown in the last row of Table 1. Interestingly, FCM yielded the fastest running time, only 4.62 seconds. While the IFS running time is about 16.82 seconds, and the FS needs 38.25 seconds. When comparing IFS and FCM, FCM has a faster processing time because the intuitive modeling process on IFS still requires several stages compared to ordinary fuzzy by considering non-membership functions.

Minority class

For the minority class (bold value in Table 1), the IFS method generates a value of 0.53% or close to 0, meaning that the number of schools closed (minor) decreases or face-to-face learning can proceed. While the FS method shows the percentage of the class minority is still high, approximately 15.6%. The FS result is even worse than FCM, in which the class minority of FCM is 8.72%.

Model accuracy

The rule formulation of unsupervised classification we propose based on the fuzzy expert algorithm produces 99.47% accuracy of school openings by utilizing the IFS approach. This means that the rule we determine by mapping the expert knowledge and the government regulation provides a good result close to the actual condition for expecting 100% face-to-face learning, as directed by the Surabaya city Government.

Those are some performance parameters for an unsupervised model based on a fuzzy expert system derived from defining deterministic rules. We further strengthen the modeling results in Fig. 6, showing the correlation between the number of school openings and the fluctuation of daily COVID-19 cases from July-October 2022. As further information, the Indonesian academic year begins in July each year.

Fig. 6 depicts the number of schools that can be opened negatively correlates with case growth. This implies that fewer schools are opening as cases grow in number. On the other hand, more schools will open if the number of cases falls dramatically. In short, education decision-makers must understand the dynamics of cases very well to initiate face-to-face learning. Our result can be a tool to monitor the right time to reopen school safely.

Let us analyze some examples of school opening and closing formulations in each zone. Table 2 shows which zones can open schools safely in the middle of July 2022, which is based on the modeling of the best approach (IFS). We took some days in the middle of July to observe the dynamics of school opening and closing. For instance, zone 3 on July 17 was to be closed right after the day before it was opened (grey highlighted). A similar thing occurred in zone 6, where the circumstances for opening and closing schools vary day by day. The results are actually burdensome to implement, as the school formula for opening and closing is dynamic as the spread of COVID-19 fluctuation.

Since our study in formulating the deterministic rule reveals certainty from actual data processing, the result can be used as a reference. Consequently, when schools implement this findings, face-to-face learning can be safely improved to accelerate education and accommodate the prominence of the education sector.

5. Conclusion

The fuzzy expert system facilitates the results by adjusting our deterministic rules. The three fuzzy approaches prove that the IFS outperforms the others in accuracy, minority class, and silhouette score, but not the running time. The key is that the precise values produced by the rule formulation can present the results of a non-vague classification for two classes in a clear and close to real-world manner regarding schools' safe opening and closing.

We can conclude that the algorithm we implement in the experiments using the IFS is adaptable to the rules we define from expert knowledge and government regulation. It has been demonstrated that IFS's accuracy regarding school opening is close to 100%, as expected by all stakeholders.

For further development, we attempt to construct a fusion model of IFS and deep learning for conflicting high-risk cases that do not yet have a specific class. By establishing structured rules, the hesitant degree can be derived as accurately as possible as the claim of the intuitionistic fuzzy algorithm.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The first author is responsible for the research concept, methodology, experiment, and manuscript. The second and sixth authors checked the manuscript and performed quality control. The third author has approved the verification. The fourth author completed a writing review. While the fifth author is responsible for data interpretation and processing. The sixth author was in charge of the writing—review and validation.

Acknowledgments

The Indonesian government supported this work through the Lembaga Pengelola Dana Pendidikan (LPDP) RI Scholarship Scheme and the Indonesian Ministry of Education and Culture's 2023 National Research Program. Part of this work was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research 19H12251.

References

[1] R. Czabanski, M. Jezewski, and J. Leski, *Introduction to Fuzzy Systems*, 2017.

[2] Nasir Mustafa, "Impact of the 2019 – 20 coronavirus pandemic on education", *Int. J. Heal. Prefer. Res.*, pp. 1–36, 2020.

[3] N. Reuge, R. Jenkins, M. Brossard, B. Soobrayan, S. Mizunoya, J. Ackers, L. Jones, and W. G. Taulo, "Education response to COVID 19 pandemic, a special issue proposed by UNICEF: Editorial review", *Int. J. Educ. Dev.*, Vol. 87, p. 102485, 2021, doi: 10.1016/j.ijedudev.2021.102485.

[4] O. Yazdanbakhsh and S. Dick, "A systematic review of complex fuzzy sets and logic", *Fuzzy Sets Syst.*, Vol. 338, No. 1, pp. 1–22, 2018, doi: 10.1016/j.fss.2017.01.010.

[5] M. Benhari and R. Hossseini, "An Improved Fuzzy Deep Learning (IFDL) model for managing uncertainty in classification of pap-smear cell images", *Intell. Syst. with Appl.*, Vol. 16, No. October, p. 200133, 2022, doi: 10.1016/j.iswa.2022.200133.

[6] P. Ravikumar and V. Susheela Devi, "Fuzzy classification of time series data", In: *Proc. of IEEE International Conference on Fuzzy Systems*, pp. 1–6, 2013, doi: 10.1109/FUZZ-IEEE.2013.6622571.

[7] O. Castillo and P. Melin, "Towards Interval Type-3 Intuitionistic Fuzzy Sets and Systems", *Math. MDPI*, Vol. 10, No. 4091, pp. 1–13, 2022, doi: https://doi.org/10.3390/math10214091.

[8] V. Kreinovich, "Fuzzy , Intuitionistic Fuzzy , What Next ? Fuzzy Logic : A Brief Reminder", In: *Proc. of Fuzzy, Intuitionistic Fuzzy, What Next?*, pp. 1–28, 2015.

[9] M. Xia and Z. Xu, "Hesitant fuzzy information aggregation in decision making", *Int. J. Approx. Reason.*, Vol. 52, No. 3, pp. 395–407, 2011, doi: 10.1016/j.ijar.2010.09.002.

[10] C. L. Chowdhary and D. P. Acharjya, "Breast cancer detection using intuitionistic fuzzy histogram hyperbolization and possibilitic fuzzy c-mean clustering algorithms with texture feature based classification on mammography images", *ACM Int. Conf. Proceeding Ser.*, Vol. 12-13-Augu, 2016, doi: 10.1145/2979779.2979800.

[11] Y. K. Dubey and M. M. Mushrif, "Segmentation of brain MR images using intuitionistic fuzzy clustering algorithm", In: *ACM International Conference Proceeding Series*, pp. 1–6, 2012, doi: 10.1145/2425333.2425414.

[12] S. Aarthi and M. Shanmugasundari, "Comparison of Single Transmit Queuing System Including Proportions of Execution Using Fuzzy Queuing Model and Intuitionistic Fuzzy Queuing Model with Two Classes", *Int. J. Intell. Eng. Syst.*, Vol. 15, No. 5, pp. 172–183, 2022, doi: 10.22266/ijies2022.1031.16.

[13] Y. Li, X. Hu, W. Pedrycz, F. Yang, and Z. Liu, "Multivariable fuzzy rule-based models and their granular generalization: A visual interpretable framework", *Appl. Soft Comput.*, Vol. 134, p. 109958, 2023, doi: 10.1016/j.asoc.2022.109958.

[14] M. Y. Kamil and A. M. Salih, "Mammography images segmentation via Fuzzy C-mean and K-mean", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 1, pp. 22–29, 2019, doi: 10.22266/IJIES2019.0228.03.

[15] J. C. Dunn, "Well-Separated Clusters and Optimal Fuzzy Partitions", *J. Cybern.*, Vol. 4, No. 1, pp. 95–104, 2008.

[16] A. El Hajjaji and A. Rachid, "Analytic

formulation of linguistic rules for fuzzy controller", *Fuzzy Sets Syst.*, Vol. 73, No. 2, pp. 219–225, 1995, doi: 10.1016/0165-0114(94)00316-Y.

[17] K. Bahani, M. Moujabbir, and M. Ramdani, "Fuzzy rule learning with linguistic modifiers", In: *ACM International Conference Proceeding Series*, pp. 1–6, 2018, doi: 10.1145/3289402.3289533.

[18] K. Kaczmarek, L. Dymova, and P. Sevastjanov, "Intuitionistic fuzzy rule-base evidential reasoning with application to the currency trading system on the Forex market", *Appl. Soft Comput.*, Vol. 128, pp. 1–26, 2022, doi: 10.1016/j.asoc.2022.109522.

[19] R. Coppi, M. A. Gil, and H. A. L. Kiers, "The fuzzy approach to statistical analysis", *Comput. Stat. Data Anal.*, Vol. 51, pp. 1–14, 2006, doi: 10.1016/j.csda.2006.05.012.

[20] B. Davvaz, I. Mukhlash, and S. Soleha, "Fuzzy Sets and Rough Sets", *Limits J. Math. Its Appl.*, Vol. 18, No. 1, p. 79, 2021, doi: 10.12962/limits.v18i1.7705.

[21] N. Rijati, D. Purwitasari, S. Sumpeno, and M. Purnomo, "A decision making and clustering method integration based on the theory of planned behavior for student entrepreneurial potential mapping in Indonesia", *Int. J. Intell. Eng. Syst.*, Vol. 13, No. 4, pp. 129–144, 2020, doi: 10.22266/ijies2020.0831.12.

[22] K. Atanassov, "Intuitionistic fuzzy logics as tools for evaluation of Data Mining processes", *Knowledge-Based Syst.*, Vol. 80, pp. 122–130, 2015, doi: 10.1016/j.knosys.2015.01.015.

[23] C. Cornelis, G. Deschrijver, and E. E. Kerre, "Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: Construction, classification, application", *Int. J. Approx. Reason.*, Vol. 35, No. 1, pp. 55–95, 2004, doi: 10.1016/S0888-613X(03)00072-0.

[24] Z. Xu and R. R. Yager, "Dynamic intuitionistic fuzzy multi-attribute decision making", *Int. J. Approx. Reason.*, Vol. 48, No. 1, pp. 246–262, 2008, doi: 10.1016/j.ijar.2007.08.008.

[25] W. Hayuningtyas, M. H. Purnomo, and A. D. W. Wibawa, "Performance Evaluation of 198 Village Governments using Fuzzy TOPSIS and Intuitionistic Fuzzy TOPSIS", *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, Vol. 4, No. 2, pp. 111–120, 2022, doi: 10.22219/kinetik.v7i2.1393.

[26] P. Singh, R. Mukundan, and R. De Ryke, "A comparative analysis of clustering algorithms for ultrasound image despeckling applications", In: *ACM International Conference Proceeding Series*, pp. 51–56, 2018, doi: 10.1145/3282286.3282290.

[27] Indonesia, Covid-19 Task Force, *Analysis of Recommendations Towards a Productive and Safe to COVID-19 in Indonesia*, June 7, 2020.

[28] P. K. Surabaya, "Againts COVID-19 in Surabaya City Government", *Pemerintah Kota Surabaya*, 2020. https://lawancovid-19.surabaya.go.id/ (accessed Feb. 17, 2021).

[29] R. I. Kemdikbud, "Guidelines for the Implementation of Learning During the Coronauurus Disease 2019 (COVID-19) Pandemic", *Minist. Educ. Cult. Repub. Indones.*, Vol. 1, No. 021, p. 28, 2020, [Online]. Available: https://bersamahadapikorona.kemdikbud.go.id/panduan-pembelajaran-jarak-jauh/.

[30] F. A. Muqtadiroh, D. Purwitasari, E. M. Yuniarno, A. P. Subriadi, and R. D. Rachmayanti, "Fuzzy Unsupervised Approaches to Analyze Covid- 19 Spread for School Reopening Decision Making", In: *Proc. of IECON 2021 - the 47th Annual Conference of the IEEE Industrial Electronics Society*, 2021, pp. 1-7, doi: 10.1109/IECON48115.2021.9589699.

[31] N. K. Arun and B. M. Mohan, "Modeling, stability analysis, and computational aspects of some simplest nonlinear fuzzy two-term controllers derived via center of area/gravity defuzzification", *ISA Trans.*, Vol. 70, pp. 16–29, 2017, doi: 10.1016/j.isatra.2017.04.023.

[32] B. N. P. B. Indonesia, "Grouping of COVID-19 Risk Criteria in Indonesia Regions", *Indonesia National Agency for Disaster Management*, 2020. https://bnpb.go.id/berita/pengelompokan-kriteria-risiko-covid19-di-daerah-berdasarkan-zonasi-warna (accessed Oct. 03, 2022).

[33] R. L. Hale, "Cluster analysis in school psychology: An example", *J. Sch. Psychol.*, Vol. 19, No. 1, pp. 51–56, 1981, doi: 10.1016/0022-4405(81)90007-8.