

# Implementation of Machine Learning Algorithm for Classification of Public Interest in Covid-19 Vaccine

*by* Scientific Publication Manager

---

**Submission date:** 23-Nov-2021 03:46AM (UTC+0200)

**Submission ID:** 1710703814

**File name:** 7\_Public\_Interest\_-\_Turnitin.docx (507.94K)

**Word count:** 2267

**Character count:** 12116

# Implementation of Machine Learning Algorithm for Classification of Public Interest in Covid-19 Vaccine

**Abstract.** Rough Set is a machine learning algorithm that analyses and determines important attributes based on an uncertain data set. The purpose of this study is to classify public interest in the Covid-19 vaccine. Vaccination is one of the solutions from the government that is considered the most appropriate to reduce the number of Covid-19 cases. Data collection was taken through a questionnaire distributed to the village community in Air Manik Village, Padang-West Sumatra, randomly as many as 100 respondents. The assessment attributes in this study are Vaccine Understanding (1), Environment (2), Community Education (3), Vaccine Confidence (4), and Cost (5), while the target attribute is the result that contains the community's interest or not to participate in vaccination. The analysis process is assisted using the Rosetta application. This study resulted in 3 reductions with 58 rules based on 100 respondents. This study concludes that the Rough Set algorithm can be used to classify public interest in the Covid-19 vaccine. Based on this research, it is hoped that it can provide information and input for local governments to be more aggressive in urging and encouraging the public to be vaccinated.

## 1. Introducing

Until now, the Coronavirus pandemic has not ended. To suppress the increasing number of cases, the administration of the COVID-19 vaccine has far begun [1]. The government also recommends that everyone get it. So, why should everyone get vaccinated against COVID-19? Currently, the COVID-19 vaccine is being distributed to all Indonesian people [2]. Giving this vaccine is the most appropriate solution to reduce the number of cases of infection with the SARS-CoV-2 virus that causes COVID-19 disease [3]. Since the introduction of the COVID-19 vaccine in Indonesia, only a small number of people have disagreed with the government's suggestion that they receive the vaccine. In order to protect the population from COVID-19 and to help restore the social and economic conditions of countries that have been devastated by the pandemic, it is critical that this vaccine be made available. In order to make a person's immune system more capable of recognizing and fighting germs or viruses that cause infection, vaccination or immunization is performed [4]. The COVID-19 vaccine has been developed with the purpose of reducing morbidity and death associated with this virus [5]. Despite the fact that this vaccine is not 100 percent effective in protecting a person from Coronavirus infection, it can minimize the likelihood of severe symptoms and consequences caused by COVID-19. It is also hoped that the COVID-19 immunization will help to encourage herd immunity to

be established [6]. This is critical because some people are unable or unwilling to receive vaccinations for certain reasons. Children and adolescents under the age of 18 years, as well as persons with certain disorders such as diabetes or uncontrolled hypertension, are not recommended to receive immunizations or are not a priority for the COVID-19 vaccine. This paper will discuss the classification of public interest in the Covid-19 vaccine using the machine learning rough set algorithm. This algorithm is used because it can analyze and determine important attributes based on uncertain data sets [7]–[9].

Several previous studies that underlie this research include rough set algorithms and genetic algorithms to classify heart failure problems. The results of this study are 77 rules and eight reductions from the use of 20 samples [10]. The following research is for the classification of heart disease based on medical data obtained from the characteristics of the ECG signal using the Rough Set. The proposed technique finds the main factors that influence the decision. A simple set quantitative measure, degree of dependence ( $k$ ), was performed to organize the attribute set according to their significance for making decisions. In addition, identical objects in the training dataset are reduced to avoid unnecessary analysis. Currently, three types of ECG data, namely Normal, Inferior MI, and Anterior MI, have been tested. Therefore, the experimental results show that the proposed method extracts specialized knowledge accurately. This study resulted in 206 rules and 42 reductions [11]. The following research uses rough sets to identify learning styles in the teaching and learning process in schools. Rough sets create rules that can be used to attribute decision-making in prediction and categorization. However, due to the method's reliance on categorical data, caution should be exercised while defining attribute categories. As a result, this study examines numerous attribute categorizations with the purpose of identifying learning styles. The findings indicated that the strategy resulted in a more accurate prediction of learning strategies. Different categories produce varying results in terms of accuracy, the amount of data omitted, the number of forgotten attributes, and the number of generated rule criteria. It can be examined for decision-making purposes by balancing these characteristics [12].

Based on the studies that have been put forward, this paper proposes the use of the Rough Set method to classify the extent to which the public is willing and interested in vaccines in the hope of reducing the number of Covid-19 cases.

## 2. Method

### 2.1. Research Dataset

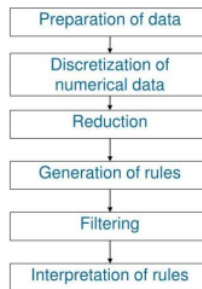
Data was collected through a questionnaire distributed to the village community in Air Manik Village, Padang-West Sumatra, randomly as many as 100 respondents. The assessment attributes in this study are Vaccine Understanding (1), Environment (2), Community Education (3), Vaccine Confidence (4), and Cost (5), while the target attribute is the result that contains the public's interest or not to participate in vaccination.

**Table 1.** Questionnaire Data

No	Vaccine Understanding	Environment	Community Education	Vaccine Confidence	Cost	Results
1	Do not Understand	Do not support	Never	Yes	No	Not interested
2	Understand	Support	Never	Yes	No	Interested
3	Do not Understand	Do not support	Never	Yes	No	Not interested
4	Understand	Support	Ever	No	No	Interested
5	Understand	Support	Never	No	No	Not interested
6	Understand	Support	Ever	Yes	No	Interested
7	Understand	Support	Never	No	No	Not interested
8	Understand	Support	Never	No	No	Not interested
9	Understand	Do not support	Never	Yes	No	Not interested
10	Do not Understand	Do not support	Never	Yes	No	Not interested
11	Understand	Support	Ever	Yes	No	Interested
12	Understand	Support	Never	No	No	Not interested
13	Understand	Support	Ever	No	No	Interested

No	Vaccine Understandlege	Environment	Community Education	Vaccine Confidence	Cost	Results
14	Do not Understand	Support	Never	Yes	No	Not interested
15	Understand	Support	Never	No	No	Not interested
16	Understand	Support	Ever	Yes	No	Interested
17	Understand	Do not support	Never	Yes	No	Not interested
18	Understand	Support	Ever	No	No	Interested
...	...	...	...	...	...	...
82	Understand	Support	Never	Yes	Yes	Interested
83	Understand	Do not support	Never	No	Yes	Not interested
84	Understand	Do not support	Never	Yes	No	Not interested
85	Understand	Do not support	Never	Yes	No	Not interested
86	Do not Understand	Do not support	Ever	Yes	No	Not interested
87	Understand	Support	Never	Yes	Yes	Interested
88	Understand	Support	Never	Yes	Yes	Interested
89	Do not Understand	Do not support	Never	No	Yes	Not interested
90	Understand	Support	Ever	No	Yes	Interested
91	Understand	Support	Never	Yes	No	Not interested
92	Understand	Do not support	Never	No	Yes	Not interested
93	Understand	Do not support	Ever	No	No	Not interested
94	Understand	Support	Never	No	No	Not interested
95	Understand	Support	Ever	No	No	Interested
96	Do not Understand	Do not support	Never	No	Yes	Not interested
97	Understand	Support	Never	No	No	Not interested
98	Understand	Support	Never	No	No	Not interested
99	Understand	Do not support	Ever	No	No	Not interested
100	Understand	Support	Ever	Yes	Yes	Interested

## 2.2. Research Flowchart

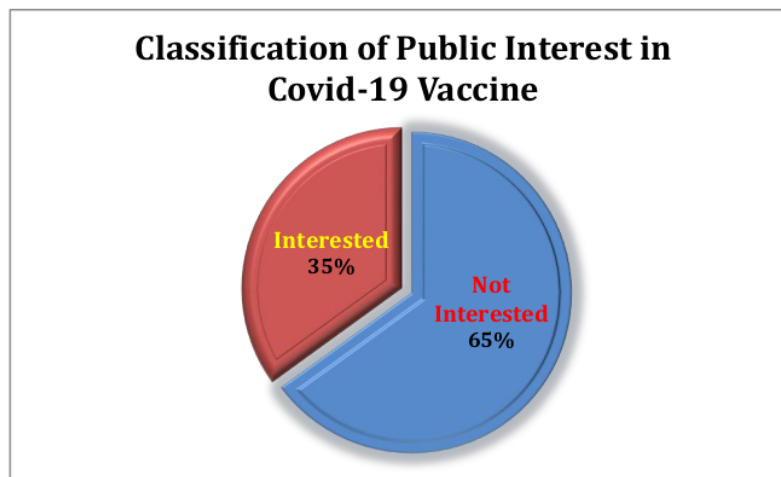


**Figure 1.** Research Flowchart Using Rough Set [13]

Sample designs using discrete-continuous variables to enable the execution of a logical sequence of operations. Three levels of categorization are used to classify the design variables in this section. Each level is associated with a distinct value range for the design parameters and objective function, corresponding to the minimum, middle, and maximum ranges of the levels 1, 2, and 3, respectively. When it comes to objective functions, clusters can be regarded of as discrete categories in place of these levels. A deterministic rule is applied to each design, which describes both the conditions (design variables) and the outcomes (results) (objective functions and clusters). All of the data is changed into a collection of rule sets as a result of this transformation. Because there are as many conditions as design variables in this rule set, it is difficult to understand. Due to the fact that some design variables have no effect on the outcome or choice, it is feasible to reduce the number of design variables required to get the same result. The operation of 'subtraction' is used to acquire the smallest possible set of circumstances for determining the desired decision characteristic, allowing for a more straightforward rule with fewer conditions. Set operations provide reduction. The reduced rule set is then filtered by frequency to determine the dominant rule set. Finally, the filtered set of rules' meaning is interpreted. ROSETTA open source software is utilized to perform the necessary calculations.



Based on the results of the generated rules, it shows 58 new rules or knowledge composed of the constituent attributes. Based on the analysis from the rules obtained, the number of occurrences of the vaccine understanding attribute is 42 times, and the environment is 16 times. So it can be seen that the most influential attribute in the classification of public interest in the Covid-19 vaccine is the vaccine understanding attribute because it has the highest number of occurrences or around 65%



**Figure 4.** Classification of Public Interest in Covid-19 Vaccine

By utilizing the Rosetta 1.4.41 tools, the resulting knowledge can be seen clearly, both the determining parameters called Reducts and obtaining General Rules from more detailed research results.

#### **7** 4. Conclusion

Based on the research conducted, it can be concluded that the classification of public interest in the Covid-19 vaccine is beneficial in processing data that has been categorized as vaccine understanding, environment, community education, vaccine confidence, and cost. To produce decision support rules that can later be used as information about public interest in implementing vaccines, as one of the anticipatory steps in suppressing the development of COVID-19. Implementing the rough set method with Rosetta Tools can answer the problem of public interest in implementing vaccines. Based on the rules found, it can be seen that the most influential condition attributes are vaccine understanding and environment.

#### **References**

- [1] T. Toharudin *et al.*, "National vaccination and local intervention impacts on covid-19 cases," *Sustainability (Switzerland)*, vol. 13, no. 15, 2021.
- [2] H. Harapan *et al.*, "Acceptance of a COVID-19 Vaccine in Southeast Asia: A Cross-Sectional Study in Indonesia," *Frontiers in Public Health*, vol. 8, no. July, pp. 1–8, 2020.
- [3] D. V. Mehrotra *et al.*, "Clinical endpoints for evaluating efficacy in covid-19 vaccine trials," *Annals of Internal Medicine*, vol. 174, no. 2, pp. 221–229, 2021.
- [4] D. Calina *et al.*, "Towards effective COVID-19 vaccines: Updates, perspectives and challenges (Review)," *International journal of molecular medicine*, vol. 46, no. 1, pp. 3–16, 2020.
- [5] E. C. Dinleyici, R. Borrow, M. A. P. Safadi, P. van Damme, and F. M. Munoz, "Vaccines and routine immunization strategies during the COVID-19 pandemic," *Human Vaccines and Immunotherapeutics*, vol. 17, no. 2, pp. 400–407, 2021.
- [6] S. H. Hodgson, K. Mansatta, G. Mallett, V. Harris, K. R. W. Emary, and A. J. Pollard, "What

- defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2," *The Lancet Infectious Diseases*, vol. 21, no. 2, pp. e26–e35, 2021.
- [7] M. Riaz, B. Davvaz, A. Firdous, and A. Fakhar, "Novel concepts of soft rough set topology with applications," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 4, pp. 3579–3590, 2019.
- [8] M. Hu and Y. Yao, "Structured approximations as a basis for three-way decisions in rough set theory," *Knowledge-Based Systems*, vol. 165, no. xxxx, pp. 92–109, 2019.
- [9] W. Wei and J. Liang, "Information fusion in rough set theory: An overview," *Information Fusion*, vol. 48, pp. 107–118, 2019.
- [10] S. Andini, R. Sitanggang, A. Wanto, H. Okprana, G. S. Achmad Daengs, and S. Aryza, "Utilization of Rough Sets Method with Optimization Genetic Algorithms in Heart Failure Cases," *Journal of Physics: Conference Series*, vol. 1933, no. 1, p. 012038, 2021.
- [11] B. Halder, S. Mitra, and M. Mitra, "Development of cardiac disease classifier using rough set decision system," *Advances in Intelligent Systems and Computing*, vol. 813, pp. 775–785, 2019.
- [12] D. S. S. Sahid, R. Efendi, E. H. Putra, and M. Wahyudi, "Categorizing attributes in identifying learning style using rough set theory," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 1, pp. 292–298, 2020.
- [13] S. Obayashi, S.-K. Jeong, K. Shimoyama, K. Chiba, and H. Morino, "Multi-Objective Design Exploration and its Applications," *International Journal of Aeronautical and Space Sciences*, vol. 11, no. 4, pp. 247–265, 2010.

# Implementation of Machine Learning Algorithm for Classification of Public Interest in Covid-19 Vaccine

## ORIGINALITY REPORT

10%

SIMILARITY INDEX

4%

INTERNET SOURCES

7%

PUBLICATIONS

1%

STUDENT PAPERS

## PRIMARY SOURCES

- 1 "Emerging Technologies in Data Mining and Information Security", Springer Science and Business Media LLC, 2019  
Publication 2%
- 2 Dadang Syarif Sihabudin Sahid, Riswan Efendi, Emansa Hasri, Muhammad Wahyudi. "Categorizing Attributes in Identifying Learning Style using Rough Set Theory", International Journal of Advanced Computer Science and Applications, 2020  
Publication 2%
- 3 Submitted to University of Oklahoma  
Student Paper 1%
- 4 [b.3cdn.net](http://b.3cdn.net)  
Internet Source 1%
- 5 B. Halder, S. Mitra, M. Mitra. "Chapter 68 Development of Cardiac Disease Classifier Using Rough Set Decision System", Springer Science and Business Media LLC, 2019  
Publication 1%



6	<a href="http://www.siemens-healthineers.com">www.siemens-healthineers.com</a> Internet Source	1 %
7	A N Islami, T Tazkiaturrizki, A Rinanti. "The effect of pH-temperature on plastic allowance for Low-Density Polyethylene (LDPE) by Thiobacillus sp. and Clostridium sp.", Journal of Physics: Conference Series, 2019 Publication	1 %
8	<a href="http://id.nii.ac.jp">id.nii.ac.jp</a> Internet Source	1 %
9	Sepyan Purnama Kristanto, Reza Syehma Bahtiar, Meriksa Sembiring, Hidayatulah Himawan et al. "Implementation of ML Rough Set in Determining Cases of Timely Graduation of Students", Journal of Physics: Conference Series, 2021 Publication	<1 %
10	Toni Toharudin, Resa Septiani Pontoh, Rezzy Eko Caraka, Solichatus Zahroh et al. "National Vaccination and Local Intervention Impacts on COVID-19 Cases", Sustainability, 2021 Publication	<1 %
11	<a href="http://jcreview.com">jcreview.com</a> Internet Source	<1 %
12	<a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a> Internet Source	<1 %

---

Exclude quotes Off

Exclude matches Off

Exclude bibliography On