**Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)**
Vol. 4, No. 2, September 2021, pp. 72 – 78
p-ISSN: 2614-3372 | e-ISSN: 2614-6150 

# A Testing of Case-Base Reasoning for Covid-19 Patient Status Confirmation

**[1]Salamun, [2]Diki Arisandi, [3]Luluk Elvitaria, [4]Liza Trisnawati**
[1,2,3,4] Program Studi Teknik Informatika, Fakultas Teknik, Universitas Abdurrab
Email: [1]salamun@univrab.ac.id, diki@univrab.ac.id, [3]luluk@univrab.ac.id, [4] liza.trisnawati@univrab.ac.id

**Article Info**

***Article history:***
Received Jan 13th, 2021
Revised Aug 6th, 2021
Accepted Sept 9th, 2021

***Keyword:***
Case-Based Reasoning (CBR)
Confirmation Status
K-NN Alghoritm
Pandemic
Protocols

**ABSTRACT**

The world is facing a global pandemic that attacks all countries. In Indonesia, there are three types of status for suspected patients: asymptomatic person, Person Under Supervision, and Patient Under Supervision. The statuses are issued by a paramedic, conducting medical examinations or direct interviews with patients with several criteria. We conducted several non-medical experiments to assist medical personnel in determining the asymptomatic. We exploit the Case-Based Reasoning (CBR) for determining the suspected patients, and the K-Nearest Neighbor (K-NN) for data grouping based on the level of similarity. The patients will be interviewed regarding their travel history, direct contact history, health status, and some other information for the past 14 days. This combination delivers the information of the similarity level from the given data and previous data. As a conclusion, the percentage level of similarity can be used by a medical officer to issue the status of patients and giving several recommendations to follow health protocols.



***Corresponding Author:***
Salamun,
Program Studi Teknik Informatika,
Universitas Abdurrab,
73 Riau Ujung, Pekanbaru, Payung Sekaki 28292, Riau, Indonesia.
Email: ademelisasuy@gmai.com

**DOI: http://dx.doi.org/10.24014/ijaidm.v4i2.11990**

## 1. INTRODUCTION

The CBR method is a method that can be used as a method to develop this research at the data test stage. Why is this method in the language in this simulation, namely how this method works, namely taking data that has existed before, while to diagnose COVID-19 patients, a doctor must know the history of the patient, namely travel history and disease history. Paramedics are currently encountering the Covid-19 outbreak. They were being being the frontliner of the fight of epidemic prevention and control, endure heavy work tasks, highly risk of infection, and work pressure [1].

Hospitals, health centers, and clinics are very strict about health protocols because these places are very vulnerable to virus transmission [2]. To determine a person with a particular status, the hospital will conduct several actions [3], one of them is a medical examination [4]. Before taking a medical examination, the patient will perform an interview with the paramedic [5], the aim is to obtain initial information for the next treatment [6]. Therefore, we proposed this method since the CBR [7],[8] and K-NN method [9],[10] can be exploited for data processing in medical research.

Case-based reasoning is a method that works from the activities that have been done before [11], but this method needs predefined basic data [12]. The method will compare the existing data with the current fact [13]. When there is any similarity between the data, then the algorithm will set it into the existing group. But if there is no congruity found in common with the previous data, then it will be saved as new cases or new rules. In testing the method, not only case-based reasoning but also data grouping based on the same data using the nearest neighbor algorithm [14].

## 2. RESEARCH METHOD

The method in data processing is the case base reasoning, it is the fundamental method to prepare new data based on existing data. For the process method, there are several steps that must be completed before proceeding to the data grouping. For the cycle method, it can be seen in Figure 2.

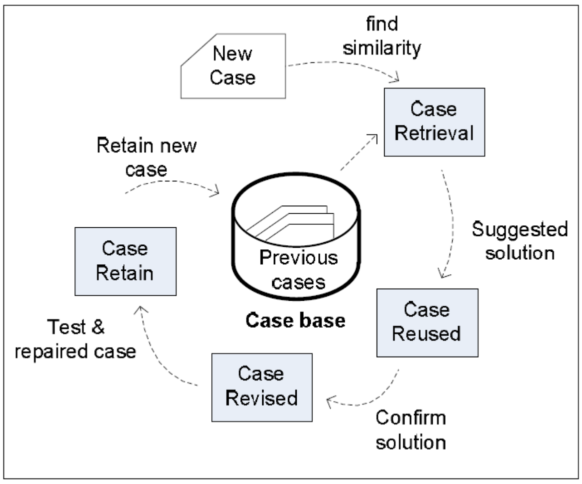


**Figure 2.** Case-Base Reasoning (CBR) Cycle

a. Retrieve
The process of identifying the cases, namely by identifying new problems. This is an important step in the CBR cycle that can be used to solve the target problem. Improving retrieval performance through more effective approaches to similarity assessment hasbeen the focus of a considerable amount of research.
b. Reuse
The reuse process in the CBR cycle is responsible for proposing a solution for a new problemfrom the solutions in the retrieved cases. Reusing a retrieved case can be as easy as returning the retrieved solution, unchanged, as theproposed solution for the new problem. This is often appropriate for classification tasks, whereeach solution (or class) is likely to be represented frequently in the case base, and therefore themost similar retrieved case, if sufficiently similar, is likely to contain an appropriate solution. But reuse becomes more difficult if there are significant differences between the new problem and theretrieved case’s problem. In these circumstances the retrieved solution may need to be adapted to account for these important differences. Medical decision making is one domain in which adaptation is commonly required. Adaptation becomes particularly relevant when CBR is used for constructive problem-solvingtasks such as design, configuration, and planning. For such tasks it is unlikely that each solution (design, configuration, or plan) will be represented in the case base. Thus the retrieved solution issimply an initial solution and any differences between the new problem and the retrieved case’sproblem are likely to alter the retrieved solution.
c. Retain
The Repaired Case may be retained in the case base as a newly Learned Case if it is likely to be useful for future problem-solving. Thus the primary knowledge source for CBR may be added to during problem-solving and is an evolving, self-adaptive collection of problem-solving experiences. In other word, the system saves new problems into the knowledge-base to solve future problems.
d. Revise
In this Reviese process, a review will be fixing the solutions that have been obtained to the problem. There are two main tasks from this stage, in this process the solutions that have been obtained from the Rause process will be re-evaluated. If successful, it will immediately proceed to the next process, namely the retain process. Otherwise, the system will improve again the case solution obtained from the retention process using a specific domain of knowledge.

## 3. RESULTS AND ANALYSIS

Case-Based Reasoning was chosen since the method able to deliver a solution from a collection of existing cases, so the new problem will adapt the solution from previous cases which have similarities with new cases. Each preexisting case is formatted as in table 1 below. The Patient Under Supervision will be categorized according to visible symptoms: fever, cough, shortness of breath, and sore throat.

**Table 1.** Case Factor

| Questions or Criteria for suspected patient |
| --- |
| The status of Patient |

We are defining question items or criteria for the Patient and assigning a certain weight to each item. The criteria in table 2 are obtained from the Covid-19 guidebook By the Indonesian Ministry of Health.

**Table 2.** List of Criteria

| Code | Question | Weight |
| --- | --- | --- |
| Q1 | Do you feeling unwell? | 2 |
| Q2 | Do you have cold symptoms? | 3 |
| Q3 | Do you have cough symptoms? | 3 |
| Q4 | Do you feel sore throat? | 3 |
| Q5 | Within 14 days, do you did a highly-risk duty? | 1 |
| Q6 | Before get ill, did you stay on the same house or having contact with COVID-19-suspected/ positive outside workplace? | 2 |
| Q7 | Before get ill, did you have travelled abroad / suspected country within 14 days? | 3 |
| Q8 | Is your colleague become reactively COVID-19 after having interaction with you? | 2 |
| Q9 | Have you perform Rapid test before? | 1 |
| Q10 | was your rapid test reactive? | 3 |
| Q11 | Do you work in the medical environment? | 1 |
| Q12 | Before get ill, did you have travelled abroad / suspected country within 14 days? | 1 |
| Q13 | Before get ill, did you stay on the same house or having contact with COVID-19-suspected / positive outside workplace? | 1 |

Furthermore, the status of a patients can be seen in Table 3 below.

**Table 3.** Patient Status

| Code | Status |
| --- | --- |
| S1 | Asymptomatic Person |
| S2 | Person Under Supervision |
| S3 | Patient Under Supervision |

It is necessary to determine a decision in a table to categorizing the criteria with patient status, the illustration can be seen in table 4 below.

**Table 4.** Decision Table

| Criteria | Status | | |
| --- | --- | --- | --- |
| | Asymptomatic Person | Person Under Supervision | Patient Under Supervision |
| Q1 | | | √ |
| Q2 | | | √ |
| Q3 | | | √ |
| Q4 | | | √ |
| Q5 | | | √ |
| Q6 | | | √ |
| Q7 | | | √ |
| Q8 | √ | | |
| Q9 | √ | | |
| Q10 | √ | | |
| Q11 | | √ | |
| Q12 | | √ | |
| Q13 | | √ | |

Table 4 categorizing each criteria according to its status. Table 4 must be determined and compute the data to discover the level of similarity of new cases to existing cases. After defining the data in table 1 - table 4, the next process is to perform an experiment on a sample case. As we mentioned before, Case-Based Reasoning method has four stages of the process: retrieving, reuse, revise, and retain. The system works based on the knowledge base that has been previously owned, which is sourced from real data from an expert or doctor, then it calculates the level of similarity by entering new cases obtained from answer of a new patient. The weighting is given according to the importance level of each criteria, where the highest weight is 3 (three), and the lowest is 1 (one). Weighting process should be obtained from medical officer/paramedic who has experience in COVID-19. This weight is called parameter (w). Where: Very important = 3, Important = 2, and Not so important=1.

$$\text{Similarity (problem, case)} = \frac{S1*w1+S2*w2.....+Sn*wn}{w1+w2+\cdots+wn} \quad (1)$$

Where :
S = similarity 1 (same) and 0 (different)
W = weight

The formula above could implemented on the case below :

**Table 5.** Illutration Of Existing Case

| Case | Criteria | Status |
|---|---|---|
| K1 | Q1, Q2, Q3, Q4, Q6, Q7 | Asymptomatic Person |
| K2 | Q11, Q12, Q13 | Person Under Supervision |
| K3 | Q8, Q9, Q10 | Patient Under Supervision |

Table 5 is a data on existing cases obtained from previous patient data that have been examined by the paramedic. The K1, K2 and K3 are codes for groups of existing cases. Next step is to create a new case table, it is a table that contains the criteria obtained from new patients examined a paramedic, which can be seen in table 6.

**Table 6.** New Case Illustration

| Case | Criteria | Weight |
|---|---|---|
| X | Q1 | 2 |
| | Q2 | 3 |
| | Q3 | 3 |
| | Q4 | 3 |
| | Q5 | 1 |
| | Q6 | 2 |
| | Q9 | 1 |

The table 6 above is a criteria that have been selected by the patient, X is the code for new cases that will be matched with the existing case according to the amount of weight. The calculation of cases will be computed to find similarities from new cases with existing cases by a weighting process using the Nearest Neighbor Retrieval formula.

Case I:

| caseX | | Existing Case (K1) |
|---|---|---|
| Q1 | → 2 | Q1 |
| Q2 | → 3 | Q2 |
| Q3 | → 3 | Q3 |
| Q4 | → 3 | Q4 |
| Q5 | → 1 | Q5 |
| Q6 | → 2 | Q6 |
| Q9 | → 0 | Q7 |

Similarity (X,K1):

$$= \frac{(1*2)+(1*3)+(1*3)+(1*3)+(1*1)+(1*2)+(0*1)}{2+3+3+3+1+2+1}$$

$$= \frac{2+3+3+3+1+2+0}{16} = \frac{14}{16}$$

$$= 0.87$$

Based on the process of case I (K1) above, it is explain the calculation of discovering the similarity of a new case using the Nearest Neighbor Retrieval formula based on criteria. As the example above, it can be seen that there are 6 criteria for new cases that are similar to the existing case. The results are 0.87 by the calculation of the first case.

Case II:

| caseX | | Existing Case (K2) |
|---|---|---|
| Q1 | → 2 | Q11 |
| Q2 | → 3 | Q12 |
| Q3 | → 3 | Q10 |
| Q4 | → 0 | - |
| Q5 | → 0 | - |
| Q6 | → 0 | - |
| Q9 | → 0 | - |

Similarity (X,K2):

$$= \frac{(0 * 2) + (0 * 3) + (0 * 3)}{2 + 3 + 3 + 3 + 1 + 2 + 1}$$

$$= \frac{0 + 0 + 0}{16} = \frac{0}{16}$$
$$= 0$$

Case II (K2) describes the process of discovering for the similarity of a new case to an existing case using the Nearest Neighbor Retrieval formula based on criteria mentioned above. It can be seen that there are no new case criteria which have similarities with the old case. Through the process above, the result is 0.

Case III:

| caseX | | Existing case (K3) |
|---|---|---|
| Q1 | → 2 | Q8 |
| Q2 | → 3 | Q10 |
| Q3 | → 0 | - |
| Q4 | → 0 | - |
| Q5 | → 0 | - |
| Q6 | → 0 | - |
| Q9 | → 1 | Q9 |

Similarity (X,K2):

$$= \frac{(0 * 2) + (0 * 3) + (1 * 1)}{2 + 3 + 3 + 3 + 1 + 2 + 1}$$

$$= \frac{0 + 0 + 1}{16} = \frac{1}{16}$$
$$= 0.06$$

Based on the process of case III (K3), the process of discovering the similarity of a new case using the Nearest Neighbor Retrieval formula based on criteria. From the case example above, it can be seen that a new case is similar to the existing case by 0.06 as the result. The output of the process above shown the lowest similarity is the K2 = 0 and K3 = 0.06. For the reuse process, the solution is the highest similarity solution between the existing case and the new case (K1 = 0.87). The output shows a confidence level was 0.87%, so the solution to the K1 is the most recommended to the patient, which is found in the existing case.

For the reuse process, The system uses the previous problem information based on the similarity weight of the most relevant to a new case, where it has similarities to solve new problems. When after the process is performed and there is no such similarity to the new case, a revise process will be conducted. The information obtained as an input criteria for new cases, where there is no resemblance to the knowledge base (rule). So it will be considered a revision / new case which will be evaluated to find a solution for the new case. In the retain process, the algorithm is integrating and extracting the new solution into the knowledge base. Furthermore, the new solution is stored in the knowledge base to solve future problems. After the revision process is complete and the correct solution has been found, then the paramedic of the COVID-19 starts adding rules by entering new case data that has been found in the knowledge base. This called the retain process. The criteria for case selection are cases that have the highest similarity, which will be advised to doctor to issue the status of patients.

**Table 7.** Status Acquisition of Patient

| New Cases | Result | Status |
|---|---|---|
| (K1,X) | 0.87 % | Patient Under Supervision |
| (K2,X) | 0 % | Person Under Supervision |
| (K3,X) | 0.06 % | Asymptomatic Person |

After discovering the similarity of the new case by the Case-Based Reasoning process using k-nearest neighbor, the results for the similarity of each case to the new case could be found in table 7. It can be seen that K1 obtained a higher similarity value of 0.87%. For K2 = 0%, because K2 does not have the same criteria as new cases, and K3 = 0.06% because K3 has the same criteria as new cases.

## 4. CONCLUSION

The Covid-19 outbreak is currently placing medical officers as frontliners. To assist medical officers in determining patient status, a combination of CBR and K-NN methods can be implemented based on the existing case. Based on the simulations, combining the two methods can identify the patient and determine the status, so the further action can be taken immediately.

## REFERENCES

[1] W. Wu *et al.*, "Psychological stress of medical staffs during outbreak of COVID-19 and adjustment strategy," *J. Med. Virol.*, vol. 92, no. 10, pp. 1962–1970, 2020.

[2] L. Rivett *et al.*, "Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission," *eLife*, vol. 9, pp. 1–20, 2020.

[3] H. Wang, S. Wang, and K. Yu, "COVID-19 infection epidemic: The medical management strategies in Heilongjiang Province, China," *Crit. Care*, vol. 24, no. 1, pp. 10–13, 2020.

[4] G. Soldati *et al.*, "Is There a Role for Lung Ultrasound During the COVID-19 Pandemic?," *J. Ultrasound Med.*, vol. 39, no. 7, pp. 1459–1462, 2020.

[5] H. Harenčárová, "Managing Uncertainty in Paramedics' Decision Making," *J. Cogn. Eng. Decis. Mak.*, vol. 11, no. 1, pp. 42–62, 2017.

[6] R. T. Davey *et al.*, "A randomized, controlled trial of ZMapp for ebola virus infection," *N. Engl. J. Med.*, vol. 375, no. 15, pp. 1448–1456, 2016.

[7] D. Gu, C. Liang, and H. Zhao, "A case-based reasoning system based on weighted heterogeneous value distance metric for breast cancer diagnosis," *Artif. Intell. Med.*, vol. 77, pp. 31–47, 2017.

[8] M. B. Bentaiba-Lagrid, L. Bouzar-Benlabiod, S. H. Rubin, T. Bouabana-Tebibel, and M. R. Hanini, "A case-based reasoning system for supervised classification problems in the medical field," *Expert Syst. Appl.*, vol. 150, p. 113335, 2020.

[9] N. C. Wong, C. Lam, L. Patterson, and B. Shayegan, "Use of machine learning to predict early biochemical recurrence after robot-assisted prostatectomy," *BJU Int.*, vol. 123, no. 1, pp. 51–57, 2019.

[10] Ž. Kovačević, L. Gurbeta Pokvić, L. Spahić, and A. Badnjević, "Prediction of medical device performance using machine learning techniques: infant incubator case study," *Health Technol. (Berl).*, vol. 10, no. 1, pp. 151–155, 2020.

[11] A. K. Goel and B. Diaz-Agudo, "What's hot in case-based reasoning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017, pp. 5067–5069.

[12] H. Y. A. Abutair and A. Belghith, "Using Case-Based Reasoning for Phishing Detection," *Procedia Comput. Sci.*, vol. 109, pp. 281–288, 2017.

[13] A. Rahman, C. Slamet, W. Darmalaksana, Y. A. Gerhana, and M. A. Ramdhani, "Expert System for Deciding a Solution of Mechanical Failure in a Car using Case-based Reasoning," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, no. 1, 2018.

[14] J. Maillo, S. Ramírez, I. Triguero, and F. Herrera, "kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data," *Knowledge-Based Syst.*, vol. 117, pp. 3–15, 2017.

[15] O. ten Cate, E. J. F. M. Custers, and S. J. Durning, *Principles and practice of case-based clinical reasoning education: a method for preclinical students*. Springer Nature, 2017.

[16] B. Shi and S. S. Iyengar, *Mathematical Theories of Machine Learning-Theory and Applications*. Springer, 2020.

[17] Q. Li *et al.*, "Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia," *N. Engl. J. Med.*, vol. 382, no. 13, pp. 1199–1207, 2020.

[18] P. Mehta, D. F. McAuley, M. Brown, E. Sanchez, R. S. Tattersall, and J. J. Manson, "COVID-19: consider cytokine storm syndromes and immunosuppression," *Lancet*, vol. 395, no. 10229, pp. 1033–1034, 2020.

**BIBLIOGRAPHY OF AUTHORS**

Salamun full name, Lecturer in the Department of Informatics Engineering, Universitas Abdurrab, my current concentration is mobile programming and web programming. research focus areas are health technology and educational technology.

Diki Arisandi, Lecturer in the Department of Informatics Engineering, Universitas Abdurrab. He also a Ph.D. student at Multimedia University, Melaka Campus, Malaysia. His research interest include networking and security, big data analytics, and IT Policy.

Luluk Elvitaria, Lecturer in the Department of Informatics Engineering, Universitas Abdurrab, my current concentration is programming logic. research focus areas are health technology and educational technology.

Liza Trisnawati, Lecturer in the Department of Informatics Engineering, Universitas Abdurrab, my current concentration is expert system. research focus areas are health technology and educational technology.