COVID-19 Incidence Prediction Model Based on Community Behavior With Neural Networks

Victor Hulu1,2, RNS Fransiska3, Widya Yanti Sihotang4, Suharni Sinaga5, Frans Judea Samosir1, Astaria Br Ginting2, Risika Wani Eka Putri1, Lam Murni Br Sagala1, Yuni Vivi Santri1, Purwadita Wahyuni2, Putranto Manalu1

1Department of Public Health, Faculty of Medicine, Dentistry and Health Sciences, Universitas Prima Indonesia, Medan, Indonesia; 2Department of Midwifery, Sekolah Tinggi Ilmu Kesehatan Senior, Medan, Indonesia; 3Midwifery Study Undergraduate Program, Sekolah Tinggi Ilmu Kesehatan Mitra Husada Medan, Medan, Indonesia; 4Department of Nursing, Akademik Keperawatan Kesdam I/BB, Pematangsiantar, Indonesia; 5Department of Nursing, Faculty of Nursing, Sekolah Tinggi Ilmu Kesehatan Mumi Teguh, Medan, Indonesia

Abstract

BACKGROUND: The COVID-19 pandemic has created a global health emergency that requires a public health response to prevent the spread of the virus.

AIM: The purpose of this study was to determine the prediction model for the incidence of COVID-19 based on the community behavior.

METHODS: This study used a cross-sectional study design. The study population was all people aged above 18 years in Medan City. The samples were 395 people recruited by convenience sampling technique. The research instrument used was a questionnaire in Google form, then, we transferred the data from the survey to a computer program using Microsoft Office Excel. Furthermore, the data were analyzed using the neural networks method. Later, the features importance was calculated using the random forest with mean decrease impurity method.

RESULTS: The results showed that, based on the confusion matrix, the prediction value for those who did not suffer from COVID-19 was correct from negative data = 8, the correct prediction value for COVID-19 from positive data = 8. While the incorrect prediction value for machines that predicted negative results but the actual data was positive = 2, and predicted a positive result but the actual data was negative = 4. Thus, based on the neural net classification method, the accuracy value was 72%. The results of this study indicate that poor preventive behavior in participants with adequate information about COVID-19 have increased the risk of infection.

CONCLUSION: Poor community behavior, such as not limiting their interaction or contact with other people, not exercising frequently, leaving the house without keeping a safe distance, and not washing hands regularly can all contribute to the COVID-19 transmission in the community.

Introduction

The COVID-19 disease originated in Wuhan City, Hubei Province, China. It has affected much of the world and created a global health emergency that requires an effective public health response including the role of society in preventing the spread of the virus [1, 2]. COVID-19 is caused by a new virus officially named severe acute respiratory syndrome COVID-19 2 which caused a global pandemic in 2020 [2, 3]. COVID-19 infection is mainly transmitted through respiratory droplets. It can also occur by touching surfaces or objects that contain the virus, then one touches their mouth, nose, and possibly eyes [4, 5]. This disease is reported to have high morbidity and mortality rate with clinical presentations such as fever, fatigue, dry cough, malaise, and difficulty in breathing [6, 7].

Globally, the number of COVID-19 cases has reached 177,588,390 people in the world with a death rate of 3,845,051 people. The total population that has received the vaccine is 2,525,891,976 people [8]. There have been more than 1.9 million cases of COVID-19 in Indonesia, 1.6 million have recovered and 54,043 people have died [9]. North Sumatra became the province with the 13th highest number of positive COVID-19 cases, which amounted to 33,313 people. Unfortunately, the city of Medan has become the largest contributor to 17,205 new cases, 15,745 people recovered and 617 people died [10].

Preliminary studies show that positive attitudes and good practices are more likely to appear in participants with adequate information about COVID-19 [11]. Individuals who believe in their efficacy are more likely to wear a face mask. Furthermore, hand hygiene is more common in individuals who have higher...
perceived susceptibility and efficacy beliefs. Individuals with greater efficacy beliefs avoid crowded areas to prevent COVID-19 [12]. However, avoiding touching one’s face and using a tissue or handkerchief to cover one’s mouth when sneezing or coughing are only practiced occasionally [7]. Therefore, the best possible intervention is prevention because COVID-19 has lately emerged as a global hazard, with significant health, economic, and societal consequences [13]. However, from these various studies, the presentation of data is only limited to describing the value of the proportion in the research sample and the relationship between variables. It cannot predict the incidence of COVID-19 based on people’s behavior. Meanwhile, in this study, a COVID-19 prediction model is carried out based on people’s behavior by applying the neural network method. Prediction of COVID-19 incidence is crucial in detecting the cause and effect of COVID-19 cases in the community so that appropriate planning and strategy development can be carried out to overcome the COVID-19 problem in Medan City. To make accurate predictions, the right method must be needed. Thus, neural network or also known as an artificial neural network is a computational method that can be used to predict incidence.

The previous study of Almeida and Azkune used neural network to predict human behavior [14]. The research of Melin et al. also used neural network to predict confirmed patients in China, Belgium, France and other countries [15]. Sánchez et al. made the prediction approach of the COVID-19 virus pandemic behavior and they indicated that neural network could be an efficient predictor for each country [16]. These researches have become evidences to use neural network as a model to predict human behavior. Although numerous prior studies have been undertaken to forecast community behavior in the incidence of a COVID-19 viral pandemic, researching the COVID-19 problem through the perspective of the people in Medan City is critical since the community behavior is inextricably linked to COVID-19 transmission. However, behavior such as not using a mask, leaving the house without keeping a safe distance, and not washing hands regularly might result in a large increase of COVID-19 transmission.

Therefore, this study aims to determine the prediction of COVID-19 based on people’s behavior with the neural network method. To the best of our knowledge, this is the first study in Indonesia that used neural network as a model to predict the behavior of the community during pandemic COVID-19.

Methods

Study design

This study used an analytical observational method with a cross-sectional study design. It collected information related to COVID-19 predictions based on community behavior in Medan city from June 7 to June 11, 2021.

Population

The population of this study was all people aged above 18 years in Medan City. Sample size was determined using a 95% confidence level assuming a single population proportion, 50% predicted prevalence (absence of estimates from related studies), and 5% precision. According to this assumption, the sample size obtained was 395 people recruited by convenience sampling technique. Furthermore, the inclusion criteria in this study included people aged above 18 years, willing to be participants, having an Android cellphone and installing WhatsApp and Telegram applications. On the other hand, the exclusion criteria included people who were sick or being treated at a health service place, were not willing to be a participant, and did not have an Android cellphone. The research flow chart is shown in Figure 1 below. The flow chart describes the steps of the research process carried out in this study.

Measurement

In this study, measurements were carried out using a questionnaire. The research questionnaire has 60 questions consisting of ten questions for knowledge, ten questions for intention and 40 questions for COVID-19 preventive behavior. The knowledge question consists of two choices, namely, Yes = 1 and No = 0. The intention question consists of five choices, namely strongly agree = 5, agree = 4, neutral = 3, disagree = 2, and strongly disagree = 1. Preventive behavior questions COVID-19 consists of two choices, namely, Yes = 1 No = 0. To improve the validity of the
questions, all research question items were pre-tested through groups of respondents similar to the anticipated respondents. From the results of testing the question items with validity and reliability tests, it showed that all corrected item total correlation values were valid with the value of \( r \) count > \( r \) table = 0.361. Then, the reliability value based on Cronbach’s alpha was above 0.7. The independent variables in this study are the variables of knowledge, intentions, and COVID-19 preventive behavior including social distancing, washing hands, wearing masks, staying at home, consuming nutritious food, physical activity, avoiding mobility, and crowd. The dependent variable is the incidence of the COVID-19.

**Data collection**

Data collection was carried out online using a Google form. We distributed the Google form link containing information on characteristics (e.g., gender, age, and occupation), informed consent, knowledge questionnaires, intentions, and behaviors to prevent COVID-19 through WhatsApp and Telegram.

**Data processing**

Data collected from the field were examined, either as a list of questions or as the responses to surveys that have been filled out by participants in the study. Then, we put a response code to the questionnaire that the respondent filled out during the research. In addition, we used Microsoft Office Excel to transfer the responses of the respondents and then imported the data into a computer program package. Then, using Microsoft Office Excel, we transferred the data from the survey to a computer program. Graphs, frequency distribution tables, and cross tables were used to summarize the data in the final step.

**Data analysis**

Coding of the software was made using the python programming language. The multi-layer perceptron (MLP) algorithm was then built using a neural network method to predict COVID-19 incidence based on community behavior. Then, the features importance was calculated using the random forest with mean decrease impurity (RF-MDI) method.

**Results**

**Characteristics of respondents**

Figure 2 shows that the gender of the respondents is mostly in the female category (74.4%) and the age of the respondents is mostly in the 19–29 year age category (66.3%). However, in this study the frequency of the age above 60 years was not found because the distribution of questionnaires through internet surveys through Google forms allowed them not to contribute to this study because they did not have an Android cellphone with the WhatsApp and Telegram applications installed. The respondent’s occupation is mostly in the student category (46.3%).

**Neural nets classification**

Table 1 from the classification process done, the training accuracy and the split data is shown in Table 1. The model used in this research is the MLP model with 50 training data, 22 test data and training accuracy value of 72%.

<table>
<thead>
<tr>
<th>Model</th>
<th>Training data</th>
<th>Test data</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>50</td>
<td>22</td>
<td>0.7272</td>
</tr>
</tbody>
</table>

MLP: Multi layer perceptron.

**Confusion matrix**

The result and performance of the classification is shown in Table 2 in form of confusion matrix. From 22 test data by comparing the dependent test data with the
predicted model data, it can be seen that the prediction value for those who do not suffer from COVID-19 is correct from negative data = 8, the correct prediction value for COVID-19 suffers from positive data = 8. While the machine’s erroneous prediction value predicts negative results but the actual data is positive = 2, and predicts positive results but the actual data are negative = 4.

Table 2: Confusion matrix

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

**Parameter design of neural networks algorithm**

The parameter of the model used is listed on Table 3 in which consist of solver='lbfgs', alpha=1e-5, hidden_layer_sizes=(150, 10), random_state=1.

Table 3: Parameter design of neural networks algorithm

<table>
<thead>
<tr>
<th>Model</th>
<th>Solver</th>
<th>Alpha</th>
<th>Hidden layer sizes</th>
<th>Random state</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP</td>
<td>lbfgs</td>
<td>1e-5</td>
<td>150, 10</td>
<td>1</td>
</tr>
</tbody>
</table>

**Random forest with mean decrease in impurity**

In this study, the search for features importance data was processed using the RF-MDI algorithm technique. The results were obtained in the form of Figure 3 below. Figure 3 shows a comparison of the influence of community behavior on the incidence of COVID-19. Feature 0 until 94 is the alternative name for every column survey question after the process of encoding. The graphs show how high the importance score of each feature that affect the process of classification of the data dependent/target.

**Discussion**

In this study, our findings have indicated that respondents have adequate knowledge about COVID-19, including virus transmission through droplets, transmission through droplets contaminated objects or places around infected people. They were also aware of common symptoms of COVID-19 infection such as fever, dry cough, and shortness of breath. The results of the study show that the knowledge of most family members was good (71.1%). However, these respondents do not have adequate intentions about COVID-19. Overall, 67.9% of respondents still have intentions in the weak category. From the responses of the respondents to the questionnaire through Google form (online), it shows that they do not agree if religious meetings and events are temporarily canceled. Moreover, they do not limit themselves to interact with other people such as in parties or crowds.

Although some knowledge of the respondents seem adequate, their behavior in preventing the transmission of COVID-19 infection is still not good. Overall, most of the COVID-19 preventive behavior is not good (62.3%). The lack of proper behavior from respondents can be seen from their answers to the questionnaire through Google form, indicating that they do not often wash their hands regularly with soap and running water for 40–60 s. This hand washing action is carried out only when leaving the house and going shopping. However, when going out to meet friends and other people, the act of washing hands regularly is not carried out despite the facility found in their destination. Nevertheless, regular hand washing is very important in efforts to prevent the transmission of COVID-19 infection.

In addition, respondents do not limit themselves to interact or contact with other people because they feel bored at home. Moreover, most respondents are still relatively young and students. They are more often mobilizing, like going out of the house with friends to visit malls, cafes, and other places for recreation. From the results of the study, the majority age group...
of respondents was 19–29 years (66.3%) and 30–40 years (18.5%). This is in line with the research of [17] that indicated the people who are most infected with COVID-19 are above 30 years and belong to the age group of 30–39 years. While the elderly are less exposed to COVID-19, but deaths from exposure to COVID-19 are higher in old age [18]. The high risk of death globally occurs among those above 50 years, but above 40 years in Indonesia [19]. This can happen because as we age, the body will experience various functional declines due to the aging process, including a decrease in the immune system [20].

Hadjidemetriou et al. specifically investigated the correlation between human mobility and the number of deaths caused by COVID-19. The results show that reducing human mobility during a pandemic significantly reduces COVID-19-related deaths. When there is a reduction in mobility, there is also a decrease in deaths due to COVID-19 cases [21]. Shao et al. also concluded based on the same thing that human mobility is positively related to the rate of rapid transmission of COVID-19 [22]. However, in his research, he did not compare it with other factors that might have a strong correlation with the spread of COVID-19.

Some of the actions that are predicted to reduce mobility of people are travel restrictions and social distancing, as well as restrictions on the number of passengers on public transport and stations. Population density, age, and the previous health conditions are factors that are considered for the high number of COVID-19 cases [21]. Rader et al. analyzed and demonstrated that the intensity of the COVID-19 epidemic could be shaped by density. This makes it easier for infectious disease epidemics to spread and last longer in areas with large populations [23].

This social restriction has also been imposed in Indonesia for a long time and even quarantine is implemented as an effort to prevent the spread of the virus. However, there are still a few notes that make the implementation of intervention disrupted and potentially less effective. Research by Orgilés et al. showed that the quarantine imposed made 52% of people feel bored [24]. It can be concluded that social restrictions for the community has contributed to boredom. This feeling of boredom tends to make people violate the rules of social isolation in various ways such as holding social gatherings and being in close proximity to more people [25], [26].

The incidence of COVID-19 in the community is also related to work. From the results of the study, the occupation of respondents varies, including working as employees, lecturers, teachers, doctors, nurses, midwives, businessman or entrepreneurs, as well as working as traders in the tax sector which requires them to go to the tax office to sell their products such as vegetables, onions, chilies, and so on. Based on work as a trader, maintaining a minimum distance of one meter is rare and difficult. COVID-19 is estimated to have a higher risk for workers who do not experience work from home because they will continue to undergo social interaction with other people. Based on Yunus research, the anxiety level of workers increased by 28.9% and they also found that cases of exposure to COVID-19 were higher in health workers compared to other occupations [27]. This can happen because working as health care workers are so close to patients exposed to the virus and examining cases [28].

The government suggestion such as staying at home or self-quarantine has been a heavy economic burden for most of the population, especially for workers who earn money daily. Their needs require them to go out and do their work every day. Thus, it could be the reason why the mobility of the community is still high.

As a result of being busy at work every day, people do not practice physical activities such as exercise for at least 30 minutes per day. It could worsen their stamina. In fact, we are aware that doing sports regularly every day can increase endurance during the COVID-19 pandemic. However, respondents always wear masks when leaving the house and avoid shaking hands with other people. Overall, this study shows that people’s behavior in preventing COVID-19 is not fully carried out by respondents when they leave the house.

Prevention of disease transmission by the individual can drastically impact the infection rate in a population [29]. Although the COVID-19 vaccination program has started, people must continue their healthy behavior. Therefore, individual with behavior change program has started, people must continue their healthy behavior. Thereafter, individual with behavior change has been promoted as an effective and inexpensive method of controlling the pandemic [30], [31].

As the COVID-19 virus spread widely, large-scale social distancing interventions are essential. Population mobility data or predictions about the patterns of changing human movement can help refine existing interventions and overcome COVID-19.

**Limitations and strengths**

This study has several limitations. First, since the distribution of questionnaires through internet surveys using Google Forms, people aged above 60 years did not contribute to this study. This can be seen from the results of the study that the frequency of people aged above 60 is unknown. Second, it is challenging to make questions with long sentences. We only relied on simple and clear questions related to people’s behavior in efforts to prevent COVID-19. Third, since the research is only conducted in Medan city, it cannot predict the incidence of COVID-19 based on the behavior of the community outside the given area. However, through our findings in this study, we would like to emphasize the importance of additional investigation of this issue with more specific steps going forward.

On the other hand, this study has several strengths. First, to our knowledge that the research with
the given topic using neural network method has never been carried out by previous researchers in Medan city, so this research can be used as a basis for conducting further research. Second, using the neural network method in this study, it can provide us with reliable information related to the incidence of COVID-19 based on the behavior of the people in Medan City.

Conclusion

The findings of the study suggest that the community’s behavior in terms of COVID-19 prevention is not fully implemented when people leave the house or go to work. Poor community behavior, such as not limiting their interaction or contact with other people, not exercising frequently, leaving the house without keeping a safe distance, and not washing hands regularly, can all contribute to the COVID-19 transmission in the community. Barriers related to community behavior in efforts to prevent COVID-19 are important to be managed more thoroughly with more specific steps so that the transmission of COVID-19 can be suppressed through changes in community behavior.

Acknowledgment

We would like to thank Universitas Prima Indonesia and the entire community of Medan City for their support and participation in this research.

References


PMid:3338486


PMid:33240167


PMid:32758711


PMid:33651707

PMid:32374718

PMid:22043275