



# Classification of COVID-19 Monthly Cases Using Artificial Neural Network (ANN) Method

Christina Natalia\*, Aniq A. Rohmawati, Sri Suryani Prasetyowati

Informatics, School of Computing, Telkom University, Bandung  
Jl. Telekomunikasi. 1, Terusan Buahbatu - Bojongsoang, Telkom University, Sukapura, Kec. Dayeuhkolot, Kabupaten Bandung, Jawa Barat, Indonesia

E-mail: <sup>1,\*</sup>christinanatalia@student.telkomuniversity.ac.id, <sup>2</sup>aniqatiqi@telkomuniveristy.ac.id, <sup>3</sup>srisuryani@telkomuniveristy.ac.id

Correspondence Author Email: christinanatalia@student.telkomuniversity.ac.id

Submitted: 02/09/2022; Accepted: 17/10/2022; Published: 31/10/2022

**Abstract**—This study proposes the MLP type ANN method with the influence of Cross-Validation evaluation using three K-Fold tests, namely 3, 5, and 8. The data used are the data that relate to COVID-19 cases issued from the Bandung Public Health Office, climate report from Bandung Meteorological, Climatological, and Geophysical Agency (BMKG), population data from Bandung Population Office, citizen's educational history from the Bandung Education Office and the West Java Open Data website. The data also was gathered from 151 sub-districts in Bandung City, with a total of 22 attributes collected from November 2020 to December 2021. The ANN method is included in the deep learning process. Therefore, the number of hidden layers utilized has a significant impact on the performance of the model being constructed. The implementation of Cross-Validation evaluation with K=8 results in an accuracy value of up to 98% and an error metric measurement of 0.3404 for MAE and 0.5994 for RMSE. This study's objective is to provide information on the optimal K-Fold Cross Validation parameters used in this ANN method to provide better performance during building a classification model for confirmed Covid-19 patients.

**Keywords:** COVID-19; ANN; Classification; K-Fold; Model Evaluation

## 1. INTRODUCTION

Coronavirus, also known as COVID-19, is a respiratory disease caused by the Coronavirus, with symptoms including fever, shortness of breath, cough, headache, dry throat, and even diarrhea [1]. COVID-19 patients have a loss of taste, loss of smell, fatigue, flu, and cough. These symptoms, however, are not restricted to COVID-19 [1]. The COVID-19 outbreak has occurred in numerous countries, culminating in an alarming number of deaths. On Monday, March 2, 2020, the president of the Republic of Indonesia revealed the first two cases of COVID-19 in Indonesia [2]. As of early 2021, the outbreak has reached the Bandung City area. Several factors influence the virus's spread rate, including the city's population size, population density, temperature, humidity, sunlight, and citizens' vaccination rate and protocol compliance [3], [4], [5]. DKI Jakarta has the highest number of COVID-19 growth rates in Indonesia with a percentage of cases of 22.0%. West Java reported it with a percentage of cases of 18.3% [6]. Several urban villages in Bandung City with the highest active cases of COVID-19 include Kebon Jeruk, Cigadung, Cigelereng, Antapani Kidul, Margasari, Manjaleg, Cipamokolan Sarijadi, Dago, and Turangga [7]. As a result, the public expects the government to develop a system that uses information technology to predict classification to prevent early Covid-19 cases. Although these predictions are not always correct, this program may serve to alert people and raise awareness in the future.

The machine learning method is one of the information technologies that has been used in the prediction of the Covid-19 classification in Indonesia. In 2022, research [8] discusses utilizing machine learning methods to classify COVID-19, normal pneumonia, and viruses using chest X-ray images. This research compares individual methods such as SVM, ANN, and LR with the combined Stacking method. The results show that the stacking model is able to produce better performance than other individual methods, with respective accuracy levels of 90.2%, 96.2%, 96.7%, and 96.9%. In 2021 a similar study also investigated the detection of the coronavirus initial phases as explained in [9] using the SVM method with cross-validation evaluation of k=2, k=5, and k=10. The highest accuracy was achieved at 99.98% with cross-validation using k=10. Based on a study published in 2022 [10], computed tomography (CT) images of the chest can be used as samples for COVID-19 detection tests because they show bilateral changes in patients infected with the virus. This approach, however, is a relatively difficult task because it predicts inaccurate bilateral changes results. Thus, the Deep Learning method was used in this study to categorize COVID-19 patients. In this study, the ANN architecture of Resnet50 was used to diagnose COVID-19 in patients. Furthermore, the hybrid model developed for the diagnosis of COVID-19 produced an accuracy rate of up to 96.30%.

On the other hand, this ANN method has also been used in the classification of Covid-19 data in research studies [11], [12]. In this 2022 study [11], the focus is on developing a rainfall classification model using the hybrid voting and stacking method through the use of a combination of five machine learning methods and the SMOTE oversampling approach. One of the individual classification methods in this study was ANN with 23 hidden layer parameters. The results show that the Stacking combination method produces 99.71% better accuracy model performance than all other methods tested. In this study, the ANN method produced a model performance of 88.69% when combined with the SMOTE technique. The weakness in this study is the ANN model which is not

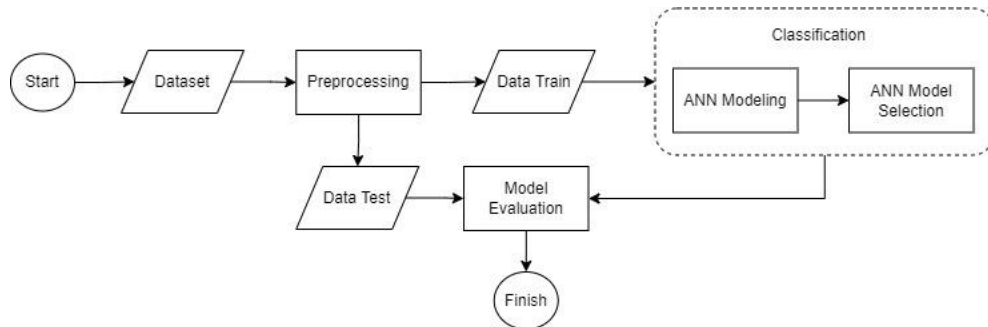
able to produce an accuracy of more than 90%. A similar study on climate classification was also conducted [12] in 2021, but only for weather cases. To calculate the weight of the ANN network, ANN is combined with backpropagation. The study focuses on developing a classification prediction model with two scenarios, namely using five factors and four factors over a six-year period. Based on the experiments, 5 factors produce the highest accuracy value of 83.33%, with a confidence interval of 92%. The weakness of this study is that the accuracy of the ANN model is not able to reach more than 85%.

Based on several previous studies, there has been no research on the classification of COVID-19 by [8], [9], [10] and research [11], [12] on the ANN method, and no research on the classification of COVID-19 based on the number of COVID-10 cases. As a result, this study will construct a model that can classify COVID-19 using the ANN method by comparing cross-validation evaluations using K-Fold = 3, 5, and 8. The data used is obtained from the city of Bandung, which has 151 urban villages and 22 attributes. The contribution of this research is to create a system to create a classification model of confirmed COVID-19 cases from each sub-district based on class so that early diagnosis can be made to control the outbreak in Bandung. Although the ANN method in research [11] and [12] has an accuracy rate of less than 90%, it is capable of producing more optimal performance.

## 2. RESEARCH METHODOLOGY

### 2.1 Research Stages

The system that will be built in this study is a classification of the spread of COVID-19 using the Artificial Neural Network (ANN) method in the Bandung City area, which consists of 151 Kelurahan.



**Figure 1.** System Planning

Based Figure 1. is a system design flow. Where the initial state starts from dataset collection, the second stage is preprocessing, where the collected data is processed into two, namely test data and train data, the third stage of classification has two processes, namely ANN modelling and ANN model selection based on train data. Then the test data from the preprocessing results along with the results of the ANN classification, will be carried out in the modeling evaluation phase of the parameters and at the end of the finish stage.

#### 2.1.1 Dataset

In this study, the data used were COVID-19 data collected from the Bandung City Public Health Office, climate reports from the Bandung City BMKG, the Bandung City Population Service, citizens' education history, and the West Java Open Data website. The data collected is from the 151 sub-districts in Bandung City, consisting of 22 attributes between November 2020 to December 2021. Attribute names are denoted by X1... X(n). Table 1 shows the results of the attribute name notation and a description of each attribute.

**Table 1. Dataset**

Notation	Description
X1	Total of Male
X2	Total of Female
X3	Unfinished Elementary School
X4	Elementary School
X5	Junior High School
X6	Senior High School
X7	1-year and 2-years Associate Degree (D1/D2)
	3-years Associate Degree (D3)
X8	Bachelor/Undergraduate (D4/S1)
X9	Postgraduate (S2)
X10	Doctorate (S3)
X11	Not in school

Notation	Description
X12	Rainfall
X13	Sunlight
X14	Average Temperature
X15	Maximum Temperature
X16	Minimum Temperature
X17	1 <sup>st</sup> Dose Vaccination
X18	2 <sup>nd</sup> Dose Vaccination
X19	3 <sup>rd</sup> Dose Vaccination
X20	Social Distancing Compliance
X21	Wearing Mask Compliance
X22	Class
Y	

### 2.1.2 Preprocessing

Preprocessing is an important part of machine learning to produce quality data results. At this stage, the data is divided in two groups: data train and data test [13]. In this study, the data preprocessing was conducted by changing the data with a missing value to a value of 0, and changing the data from several attributes into integers. The dataset used in this study was divided into 80% of data training to train the model and 20% of data test to observe and determine the performance of the model that has been obtained at the training stage. In this study, the target attribute were the class of the COVID-19 cases which have been classified based on the COVID-19 confirmed cases.

#### 2.1.2.1 Data Labeling

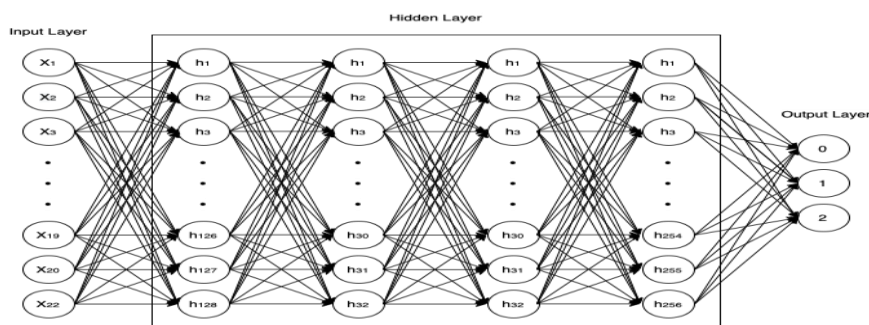
In this study, the target of the data set is the confirmed case class attribute, but it is not used in this study as an attribute. Confirmed cases of COVID-19 consist of 3 classes, namely class 0 is a case of COVID-19 outbreak in the low risk category with an interval of  $DA < 52$ , class 1 is an outbreak of COVID-19 in the medium risk category with an interval of  $52 \geq DA < 196$ , and class 2 is the case of the spread of COVID-19 in the high category with a DA interval of 196. In this study, the labeling of the data set class in this study used the binning method. As stated in research [14], the binning method is a method that functions to group stored data according to the surrounding information. Values are sorted into smaller pieces called bins. Because this method relies on local data grouping and neighboring results, the minimum and maximum bin value limits for each bin value are larger where the value interval range for each bin is the same.

#### 2.1.3 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a computational system inspired by the human learning process, where there are neurons that are interconnected to perform tasks based on their own experiences. In general, ANN consists of three layers, including input layer, hidden layer and output layer [15]. The ANN concept has several advantages, one being able to learn quickly to find solutions to various high-level problems [16]. A simple form of ANN called perceptron is a neural network builder, where each input is multiplied by the appropriate edges and then combined into a mathematical function called neuron activation while the other functions calculate the output value [17]. A set of perceptron in a network, for example, if the input to the  $i$  number of perceptron in a network is denoted by  $X_{1i}, \dots, X_{ni}$ , assuming the sum is used to calculate the output value, which is denoted by  $Z_i$ . The following is the mathematical formula [17]:

$$Z_i = \sum_{j=1}^m x_{ij}w_{ij} + b_i \tag{1}$$

Where  $n$  is the input value,  $m$  is the neuron value in the current layer,  $w_{ij}$  is the neuron edges, where inputs to- $j$  to cells to- $i$  and  $b_i$  are bias [17].



**Figure 1.** MLP Neural Network Structure

In this study, the artificial neural network model used is the multilayer perceptron (MLP) neural network. MLP class feed forwards the artificial neural network by having one or more hidden layers between the input and output layers. This model is used in supervised machine learning classification [18]. The attributes used in the input layer dramatically affect the performance of the targeted model [18]. The more attribute data used, the better the resulting model is. Figure 2 is the MLP structure used in this study, with 22 layers of input layer attributes starting from the attributes  $X_1$  as the total of male respondents,  $X_2$  the total of female respondents,  $X_3$  respondents unfinished from elementary school,  $X_4$  elementary school graduates,  $X_5$  junior high school graduates,  $X_6$  senior high school graduates,  $X_7$  1-year and 2-year associate diploma,  $X_8$  3-year associate diploma,  $X_9$  undergraduates,  $X_{10}$  post-graduates,  $X_{11}$  doctorate,  $X_{12}$  No school,  $X_{13}$  as the rainfall rate,  $X_{14}$  sunlight,  $X_{15}$  as the average temperature,  $X_{16}$  the maximum temperature,  $X_{17}$  minimum temperature,  $X_{18}$  as the 1<sup>st</sup> dose vaccination status,  $X_{19}$  2<sup>nd</sup> dose vaccination status,  $X_{20}$  3<sup>rd</sup> dose vaccination status,  $X_{21}$  as the respondent's compliance to social distancing, and  $X_{22}$  as the respondent's adherence to wear masks. The structure has 4 hidden layers, each with 128, 32, 32 and 256 amount of neuron units, and 3 layers of output layer which are the classification of confirmed cases of COVID-19. MLP class, which feed forwards the artificial neural network by having one or more hidden layers between the input and output layers, is used in the classification of supervised machine combustion [17]. For testing the predictive value of this study, these two evaluation measures are used:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |O_i - P_i| \quad (3)$$

### 2.1.3.1 Determining the Best ANN Model

The way cross-validation works are by dividing the dataset into several folds according to the k-value each fold is carried out through a training process and model testing [19]. The study applied K-fold to the data training, using 3-fold, 5-fold and 10-fold cross-validation. The result with the highest accuracy value becomes the best predictive model.

### 2.1.3.2 Classifying the Best ANN Model

Classification is grouping the data according to its similar characteristics [20]. This is applied in this study to group the data of the confirmed cases of COVID-19 spread risk based on the level of the spreading risks. This study classified the classes of confirmed cases of COVID-19 by developing an artificial neural network model. The overview of the modelling structure in this study can be seen in Figure 2.

### 2.1.4 Model Evaluation

In measuring the performance of the classification model for the spread of COVID-19 cases, the error value must be considered, hence the necessity for a confusion matrix [21]. The parameters to be evaluated were accuracy, precision, recall, and F1-score. Accuracy is the ability of the model to predict the class of the COVID-19 cases with mathematical formulas calculated with equation (4). Precision is the number of confirmed cases of COVID-19 class that match the actual data, with the mathematical formulas calculated using equation (5). Recall is the ratio of true positive predictions with all the correct data, with the mathematical formulas calculated using equation (6). F1-Score is the number of categories that can be predicted by the system correctly, with mathematical formulas that can be calculated using equation (7). The mathematical formulation of the confusion matrix is [22].

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% \quad (4)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \times 100\% \quad (5)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \times 100\% \quad (6)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (7)$$

## 3. RESULT AND DISCUSSION

In this study, the dataset used is the COVID-19 spread dataset. After the preprocessing stage, the dataset was then divided into 80% datasets for the ANN Model Training and 20% datasets for model testing. The classification model is created with several scenarios:

### 3.1 Data Collection

This study used a dataset of the spread of COVID-19 with 3 categories. These categories represent classes with interval values based on the number of confirmed cases of COVID-19. As seen on Table 2, the low risk category

has 704 confirmed COVID-19 classes, the medium risk category has 705 confirmed COVID-19 classes, and the high risk category has 705 confirmed COVID-19 classes.

**Table 2.** Class Number on Original Data

Category	Class	Interval	Total
Low Risk	0	DA < 52	704
Medium Risk	1	52 ≥ DA < 196	705
High Risk	2	DA ≥ 196	705

### 3.2 Building ANN Classification Model

In this study, building an artificial neural network model used Tensor Flow to implement and apply a heterogeneous system machine learning model available in Python. Furthermore, the Keras library tuner is utilized to find the finest ANN architecture. Keras tuner is a library that can automatically tune the ANN model by doing a RandomSearch of the predetermined parameters, to get the optimal hypermeter. Keras tuner is used to explore all possible parameters which can improve the model’s accuracy. This study performed tuning on the unit size for each hidden layer and the learning rate parameter on the optimizer. The ANN structure in this study can be seen in Figure 2.

The process of forming the ANN model using training data is then implemented further testing to achieve the best classification model and is evaluated using cross-validation. By using 3 scenarios with cross-validation parameters, namely K = 3, K = 5 and K = 8, the performance of the model is then calculated using the RMSE (root mean square error) and MAE (microwave assisted extraction) metric measurements as shown in Table 2.

**Table 3.** Metric Measurement Test Results

K-Fold	Average Accuracy	MAE	RMSE
3	66%	0.3687	0.6376
5	63%	0.3877	0.6450
8	67%	0.3404	0.5994

#### 3.2.1 Monthly Classification

After achieving the best model, the classification was implemented in 14 months. Then, the cross-validation evaluation was conducted to classify the models each month. The K parameters used were 3, 5, and 8. Consecutively, the three parameters were analysed to measure the model from each month by calculating the confusion matrix, such as: accuracy, precision, recall, and f1-score. The dataset used are 14 months of data, starting from November 2020 to December 2021.

### 3.3 Test Result

To measure the performance of the ANN model, a confusion matrix is created which consists of calculating accuracy, precision, recall and f1-score. The confusion matrix is used for its technique of evaluating the model with the output of classifications. Table 3 illustrates the details.

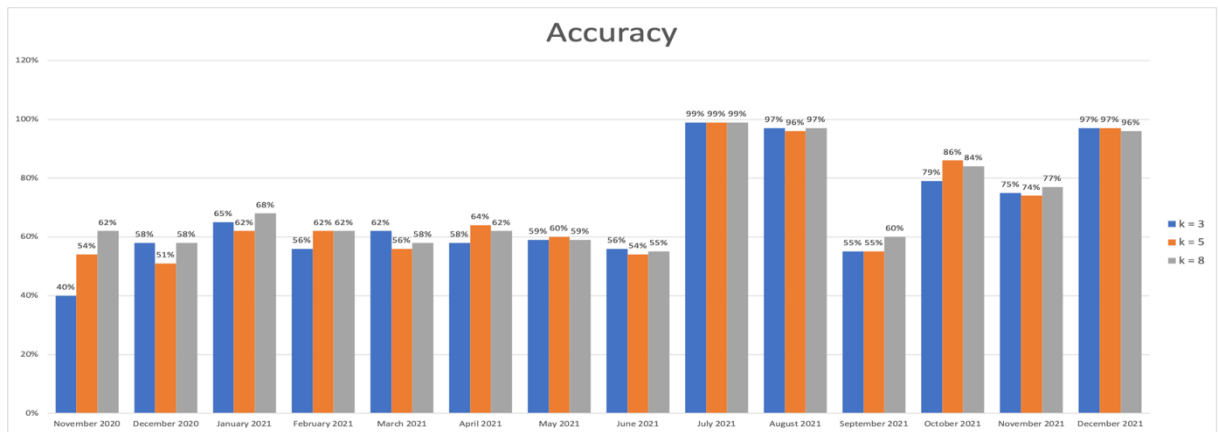
**Table 4.** ANN Model Performance with Cross Validation Evaluation and Confusion Matrix Calculation

Month	K=3				K=5				K=8			
	Accur acy	Precis sion	Rec all	F1- Sco re	Accur acy	Precis sion	Rec all	F1- Sco re	Accur acy	Precis sion	Rec all	F1- Sco re
Novembe r 2020	40%	61%	40	29	54%	56%	54	52	62%	68%	62	62
Decembe r2020	58%	56%	58	50	51%	34%	51	38	58%	61%	58	52
January 2021	65%	62%	65	62	62%	68%	62	61	68%	69%	68	68
February 2021	56%	61%	56	51	62%	71%	62	61	62%	61%	62	60
March 2021	62%	60%	62	60	56%	66%	56	54	58%	63%	58	57
April 2021	58%	64%	58	56	64%	68%	64	61	62%	67%	62	58
May 2021	59%	63%	59	56	60%	61%	60	59	59%	61%	59	53

Month	K=3				K=5				K=8			
	Accur acy	Precis ion	Rec all	F1- Sco re	Accur acy	Precis ion	Rec all	F1- Sco re	Accur acy	Precis ion	Rec all	F1- Sco re
June 2021	56%	60%	56	47	54%	55%	54	54	55%	52%	55	98
			50	98			99	98			99	98
July 2021	99%	97%	97	97	99%	98%	99	98	99%	97%	99	98
			97	97			96	96			97	97
August 2021	97%	96%	97	97	96%	96%	96	96	97%	96%	97	97
			55	48			55	53			60	59
September 2021	55%	56%	55	48	55%	53%	55	53	60%	61%	60	59
			79	81			86	81			84	81
October 2021	79%	85%	79	81	86%	88%	86	81	84%	80%	84	81
			75	69			74	62			77	70
November 2021	75%	71%	75	69	74%	54%	74	62	77%	78%	77	70
			97	96			97	96			96	95
December 2021	97%	95%	97	96	97%	95%	97	96	96%	95%	96	95
			97	96			97	96			96	95

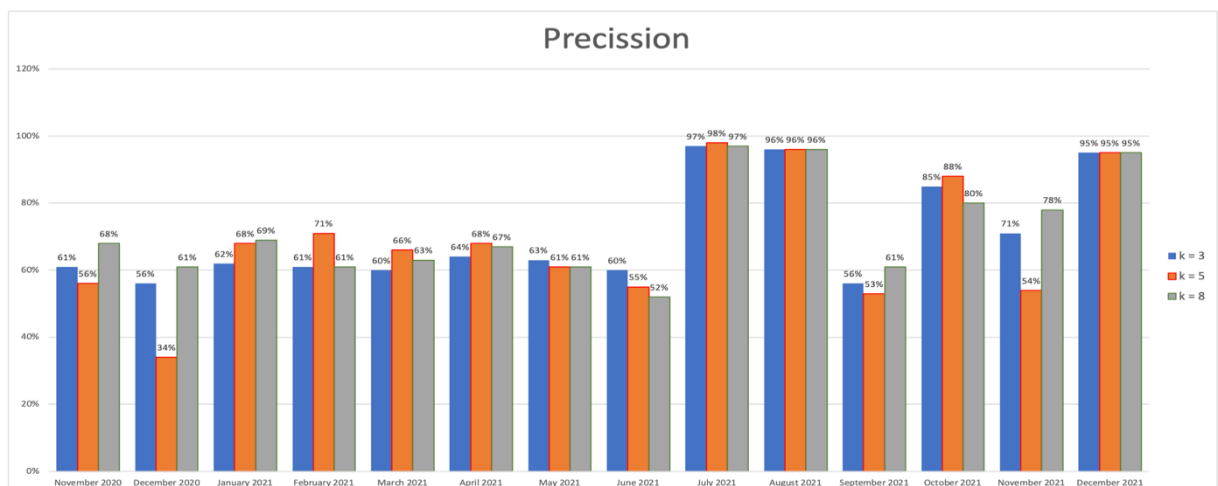
### 3.4 K-Fold Test Result Implementation

To facilitate the analysis, the model evaluation results table is illustrated in bar charts. The goal is to make it easier to perform a comparison analysis of the K parameters used in cross-validation with a monthly perspective.



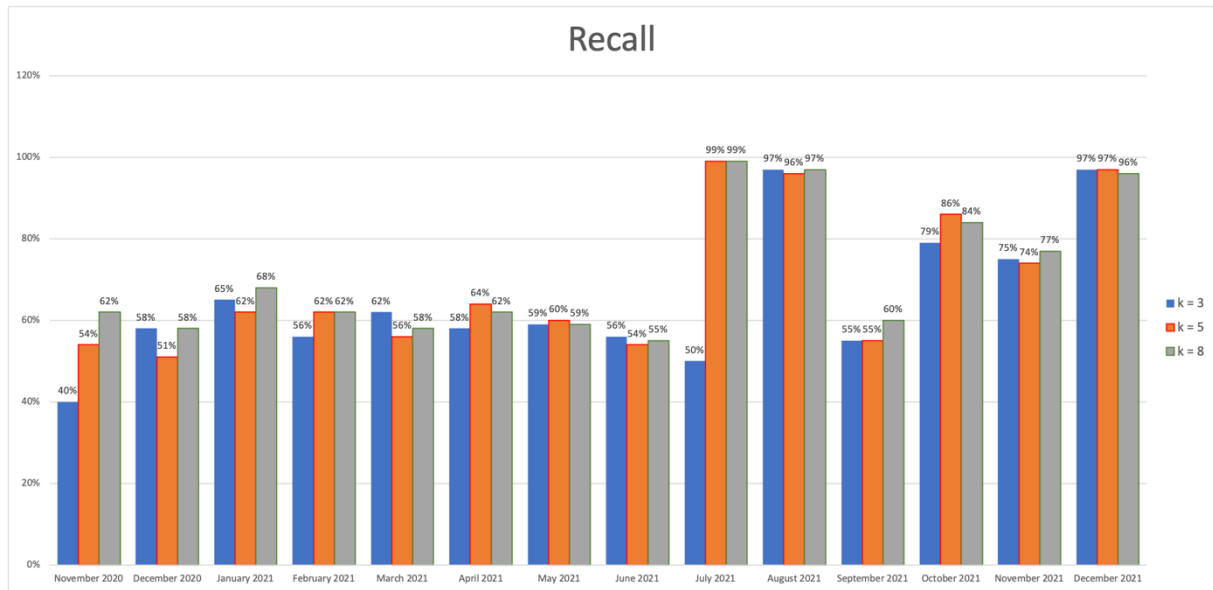
**Figure 2.** Evaluation of ANN model accuracy based on K-Fold

Accuracy is the level of proximity of the predicted value to the actual value. Based on the accuracy value displayed in Figure 3, each month has an average value of 60% with all three scenarios. However, the accuracy escalated significantly in July, August and December 2021. The highest accuracy was in July 2021 with all the proposed K parameters. Then in August 2021 and December 2021, the accuracy values were relatively similar. The K = 8 parameter provides a slightly superior accuracy throughout the months.



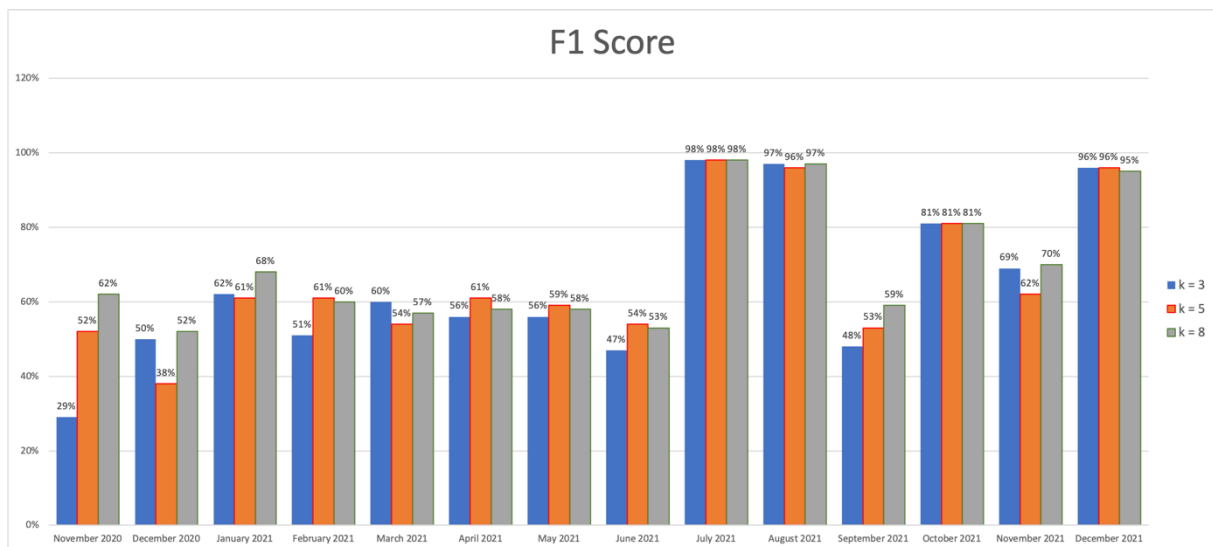
**Figure 3.** Evaluation of ANN model precision based on K-Fold

Precision is a class measure of the confirmed cases of COVID-19. In measuring the performance of the developed model, accuracy is not the only focus, precision measurement is necessary as well. The higher the precision value, the better the classification model is. The best precision evaluation was in July 2021 with K = 5 and 8 that reached 99% precision value. The performance of the built model is less efficient in other months, such as December 2020 and June 2021 due to the inability of the precision measurement value to climb higher than 60%. Further details are portrayed in Figure 4.



**Figure 4.** Evaluation of ANN models recall based on K-Fold

Recall is the ratio of the correct positive class prediction with the correct data used. In this model, recall measures the positive class model and marks whether the class classification results conform to the actual data. Recall and Precision generally have similarities, furthermore, both are parameters used to measure how well the classification model is developed. As seen on Figure 5, the highest recall values were in July 2021, which were soaring up to 99% when using K=5 and K=8. While the lowest were in June 2021, with recall values lesser than 60% in all 3 test scenarios.



**Figure 5.** Evaluation of F1-Score model using K-Fold

F1-Score was used to measure the number of categories the system can predict correctly. F1-Score is resulted from the process of Precision and Recall, and is usually used to measure model performance with imbalanced data. Figure 6 displays that the lowest F1-Score risk occurred in November 2020 when risk value only reached 29% using the K=3 model. This happens because the precision and recall values generated using K=3 in that month have the lowest values.

Based on the results, the ANN model with the effect of Cross-validation using K = 5 has better performance, which produced accuracy, precision, recall, and f1-score values of more than 98% in July 2021. But overall, the ANN model with the effect of cross-validation with K = 8 has a relatively superior performance



compared to the K=5 scenario. When measured with RMSE and MAE, the K=8 scenario has the smallest error values at 0.3404 and 0.5994. The ANN method proposed in this study has a better performance with the effect of cross-validation with K=5 and K=8, which is higher than 98%. The model in this study outperformed the previous models [11] and [12], which were only able to produce a performance of less than 90%.

## 4. CONCLUSION

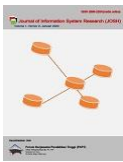
The artificial neural network method with the influence of cross-validation classification model proved successful to produce better performance, especially when using K=8. The type of ANN method used was the multilayer perceptron (MLP) with 22 attributes at the input layer, 4 hidden layers with a number of neuron units each (128, 32, 32 and 256), and 3 output layers of 0, 1, and 2 as the class. The application of K-Fold into the evaluation model with 3 evaluation scenarios of K=3, K=5, and K=8 resulted in presumably the best classification model. The K-Fold model = 8 was the one that produced the highest accuracy, precisely in July 2021, with an accuracy of 98%. Moreover, in testing the ANN classification model, the MAE value is 0.3404 while the RMSE value is 0.5994, indicating that the model's error value has a better performance value than the previous models. The results of the classification prediction set as an output in this study are the medium risk category with a class 1 label of 856, which is classified as a medium risk of COVID-19 spread.

## ACKNOWLEDGE

The author would like to thank as much as possible to Telkom University for its facilities and infrastructure. Moreover, the Bandung City Health Office, Bandung City BMKG, Bandung City Population Service, Education History Data, and the Open Data Website provide information and data related to this research. So that research can be completed.

## REFERENCES

- [1] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021, doi: 10.1080/07391102.2020.1788642.
- [2] E. A. Retaduari, "2 Maret 2020, Saat Indonesia Pertama Kali Dilanda Covid-19," 2020.
- [3] H. Coşkun, N. Yıldırım, and S. Gündüz, "The spread of COVID-19 virus through population density and wind in Turkey cities," *Sci. Total Environ.*, vol. 751, 2021, doi: 10.1016/j.scitotenv.2020.141663.
- [4] H. M. Zawbaa *et al.*, "Effect of mutation and vaccination on spread, severity, and mortality of COVID-19 disease," *J. Med. Virol.*, vol. 94, no. 1, pp. 197–204, 2022, doi: 10.1002/jmv.27293.
- [5] P. Mecnas, R. T. da Rosa Moreira Bastos, A. C. Rosário Vallinoto, and D. Normando, "Effects of temperature and humidity on the spread of COVID-19: A systematic review," *PLoS One*, vol. 15, no. September 9, pp. 1–21, 2020, doi: 10.1371/journal.pone.0238339.
- [6] Admin, "PETA SEBARAN."
- [7] Admin, "Detail Kasus COVID-19 Kota Bandung." <https://covid19.bandung.go.id/detail-kasus>
- [8] Y. S. Taspınar, I. Cinar, and M. Koklu, "Classification by a stacking model using CNN features for COVID-19 infection diagnosis," *J. Xray. Sci. Technol.*, vol. 30, no. 1, pp. 73–88, 2022, doi: 10.3233/XST-211031.
- [9] Ş. Öztürk, U. Özkaya, and M. Barstuğan, "Classification of Coronavirus (COVID-19) from X-ray and CT images using shrunken features," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 1, pp. 5–15, 2021, doi: 10.1002/ima.22469.
- [10] M. Yildirim and A. Cinar, "A deep learning based hybrid approach for covid-19 disease detections," *Trait. du Signal*, vol. 37, no. 3, pp. 461–468, 2020, doi: 10.18280/ts.370313.
- [11] A. Gumilar, S. S. Prasetyowati, and Y. Sibaroni, "Performance Analysis of Hybrid Machine Learning Methods on Imbalanced Data (Rainfall Classification)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 3 SE-Artikel Teknologi Informasi, Jul. 2022, doi: 10.29207/resti.v6i3.4142.
- [12] S. Shofura, S. Suryani M.Si, L. Salma, and S. Harini, "The Effect of Number of Factors and Data on Monthly Weather Classification Performance Using Artificial Neural Networks," *Int. J. Inf. Commun. Technol.*, vol. 7, no. 2, pp. 23–35, 2021, doi: 10.21108/ijoict.v7i2.602.
- [13] I. Ahmad and S. M. Asad, "Predictions of coronavirus COVID-19 distinct cases in Pakistan through artificial neural network," *Epidemiol. Infect.*, 2020, doi: 10.1017/S0950268820002174.
- [14] S. A. Alasadi and W. S. Bhaya, "Review of data preprocessing techniques in data mining," *J. Eng. Appl. Sci.*, vol. 12, no. 16, pp. 4102–4107, 2017, doi: 10.3923/jeasci.2017.4102.4107.
- [15] S. Chakraverty and S. K. Jeswal, *Applied artificial neural network methods for engineers and scientists: Solving algebraic equations*. 2021.
- [16] A. Neuron *et al.*, and *Machine Learning – ICANN 2018*. 2018.
- [17] A. Mollalo, K. M. Rivera, and B. Vahedi, "Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States," *Int. J. Environ. Res. Public Health*, vol. 17, no. 12, pp. 1–13, 2020, doi: 10.3390/ijerph17124204.
- [18] A. Prasetya Wibawa, W. Lestar, A. Bella Putra Utama, I. Tri Saputra, and Z. Nabila Izdihar, "Multilayer Perceptron untuk Prediksi Sessions pada Sebuah Website Journal Elektronik," *Indones. J. Data Sci.*, vol. 1, no. 3, pp. 57–67, 2020, doi: 10.33096/ijodas.v1i3.15.
- [19] Y. Zeng, K. Jiang, and J. Chen, "Automatic seismic salt interpretation with deep convolutional neural networks," *ACM Int. Conf. Proceeding Ser.*, pp. 16–20, 2019, doi: 10.1145/3325917.3325926.



- [20] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, “Metode-metode Klasifikasi,” *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
- [21] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, “The impact of class imbalance in classification performance metrics based on the binary confusion matrix,” *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.
- [22] M. A. Rahman and R. C. Muniyandi, “Feature selection from colon cancer dataset for cancer classification using Artificial Neural Network,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 8, no. 4–2, pp. 1387–1393, 2018, doi: 10.18517/ijaseit.8.4-2.6790.