

## MEASUREMENT OF CLASSIFICATION PERFORMANCE WITH THE LEARNING VECTOR QUANTIZATION METHOD ON COVID-19 VACCINATION DATA AT THE PARUMPANAI HEALTH CENTER

ADHIYAKSA PRANANDA RS<sup>a</sup>, SISWANTO<sup>b,\*</sup>, SRI ASTUTI THAMRIN<sup>c</sup>, A. MUH. AMIL  
SIDDIK<sup>d</sup>

<sup>a,b,c</sup> Department of Statistics, Hasanuddin University,

<sup>d</sup> Department of Mathematics, Hasanuddin University.

email : aksakiseki@gmail.com, siswanto@unhas.ac.id, tuti@unhas.ac.id, amilsiddik@unhas.ac.id

Diterima ..... Direvisi ..... Dipublikasikan .....

**Abstract.** *In the midst of the COVID-19 pandemic, various countries are always trying their best to restore global stability. One effective way is the discovery of several vaccines to prevent transmission of the virus. Indonesia is one of the countries that is aggressively implementing the COVID-19 vaccination. The vaccination process which has been carried out from February 2021 until the end of 2021 has covered approximately 160 million people or 76.83% of the target set by the government. Vaccine recipients have criteria to be able to get vaccinated to avoid side effects or complications. So it is necessary to classify groups that can receive vaccines and also delay vaccination. This research aims to determine the performance of the learning vector quantization classification method. Learning vector quantization method classification produces 95% accuracy, 97% precision, and 96% sensitivity. From these performance measurements, it can be concluded that the learning vector quantization method is very good and can be used in the classification of COVID-19 vaccination recipients at the Parumpanai Public Health Center, East Luwu Regency.*

*Keywords:* classification, COVID-19, learning vector quantization, vaccination

### 1. Introduction

The COVID-19 pandemic, which first hit China caused emergencies in various countries. Indonesia is one of the countries most affected by the COVID-19 pandemic. According to the Emergency Committee, the spread of the COVID-19 virus can be prevented by carrying out protection, detection, isolation, and rapid treatment [1]. One of the focuses of the Indonesian government in stopping the spread of the COVID-19 virus is by vaccinating the community.

\*corresponding author

Vaccines are one of the most effective and economical ways to prevent the transmission of diseases caused by viruses [2]. Countries worldwide are trying to accelerate the research and development of COVID-19 vaccines. More than 160 vaccine candidates have been developed and 20 of them are in the clinical evaluation stage [3]. The important thing to consider in developing a vaccine is to ensure the safety of the benefits provided. One of the potential risks in developing a COVID-19 vaccine is whether the immune responses elicited by a vaccine could enhance COVID-19 acquisition or make the disease worse when infection occurs after vaccination [4]. Report to the official website covid19.go.id, the vaccination process which has been carried out from February 2021 until the end of 2021 covers around 160 million Indonesians in giving the first dose of the COVID-19 vaccine. Specifically in South Sulawesi Province, according to data from the website vaksin.kemkes.go.id, the vaccination process until mid-September 2021 recorded 2.1 million people who had vaccinated against COVID-19 with a percentage of 30.14% of the government's target. However, public perception of the safety and effectiveness of vaccines is a major obstacle to vaccination implementation in Indonesia [5]. This is commonly known as vaccine hesitancy. Vaccine hesitancy is defined as the delay in acceptance or refusal of vaccination despite the availability of vaccination service [6]. Another thing that hinders vaccination in Indonesia is access to vaccine dose distribution which is difficult to reach in remote areas. The Parumpanai Public Health Center in East Luwu Regency is one of the many vaccine distribution sites that have difficulty reaching people in remote areas. The hilly area and poor transportation access make the area difficult to access for health workers who want to distribute vaccines. One way to make the COVID-19 vaccination process more even, especially in East Luwu Regency, is to classify vaccination recipients in the area.

Classification is a process of finding a model that describes and differentiates data classes so that they can be used to predict classes [7]. Classification problems have been widely studied and have a big role in analyzing data and making decisions. The learning vector quantization method is one type of classification method of artificial neural networks based on competitive learning supervised [8]. The way the artificial neural network works is the same as the structure of the neurons in the brain of humans who are related to each other. The neural network consists of the number of units joined in the same pattern. This unit is separated into parts, namely the input layer, the output layer, and the hidden layer [9]. This study uses the Learning Vector Quantization method to classify vaccination recipients at the Parumpanai Public Health Center, East Luwu Regency. This study aims to obtain the performance of the classification model so that it can be determined that the Learning Vector Quantization method can classify vaccination recipients at the Parumpanai Public Health Center, East Luwu regency correctly.

## **2. Metodology**

### **2.1. Data**

The data used in this research is secondary data obtained from the medical records of monthly vaccinations at the Parumpanai Public Health Center, East Luwu Re-

gency. The data consists of 1 dependent variable and 5 independent variables. The data used in this research starts from June-August 2021 with a total of 500 data. Then the population will be divided into training data and test data with a proportion of 80:20 which produces 400 training data and 100 test data. The independent and dependent variables used in this research are shown in Table 1.

Table 1. Variables in this research

	Variable		Variable Name
$Y$	Vaccination Status	1	Continue vaccination
$X_1$	Age	2	Postpone vaccination The requirement to receive the COVID-19 vaccination is someone who is over 12 years old
$X_2$	Blood pressure	0	Blood pressure $\geq 140/90$ mmHg
		1	Blood pressure $< 140/90$ mmHg
$X_3$	Disease History	0	Have a history of disease
		1	Have no history of disease
$X_4$	Pregnant/ Breast-Feed	0	Have pregnant/breastfeeding
		1	Have no pregnant/breastfeeding
$X_5$	Body Temperature	0	Body temperature $> 37,5$ °C
		1	Body temperature $\leq 37,5$ °C

## 2.2. Research Method

Data analysis was carried out with the help of python software. The steps of data analysis carried out in this research are as follows.

- (1) Normalize the data using the Min-Max Normalization method on the  $X_1$  variable. In this method, the data is usually scaled to a fixed range, usually in binary 0 and 1 with the formula [10]:

$$X^* = \frac{X_i - X_{min}}{X_{max} - X_{min}}$$

Information:

$X^*$  : Normalized data

$X_i$  : Preliminary data,  $i = 1, 2, 3, \dots, n$

$X_{min}$  : Minimum data

$X_{max}$  : Maximum data

Normalization aims to change the value in  $X_1$  so that it is in a binary form like other variables.

- (2) Labelling the  $X_1$  variable. Labeling aims to assign a value of 0 to  $X_1(\text{normalization}) < 0.5$  and 1 to  $X_1(\text{normalization}) \geq 0.5$ .
- (3) Create an LVQ network architecture with predefined variables.

- (4) Initiating the learning rate parameter ( $\alpha$ ) = 0.05, Dec  $\alpha$  = 0.1 and Maxepoch = 50.
- (5) Entering data and classification targets.
- (6) Run the LVQ algorithm.
- (7) Calculate the Euclidean distance between the data and the predefined weights.
- (8) Calculate the target value to get the final weight.
- (9) Calculate the accuracy, precision, and sensitivity of the LVQ classification with 4 possible test records generated in the classification process, namely true positive, true negative, false positive, and false negative with the formula [11]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Information:

$TP$  : True Positive

$TN$  : True Negative

$FP$  : False Positive

$FN$  : False Negative

$$Precision = \frac{TP}{TP + FP}$$

Information:

$TP$  : True Positive

$FP$  : False Positive

$$Sensitivity = \frac{TP}{TP + FN}$$

Information:

$TP$  : True Positive

$FN$  : False Negative

### 3. Results and Discussion

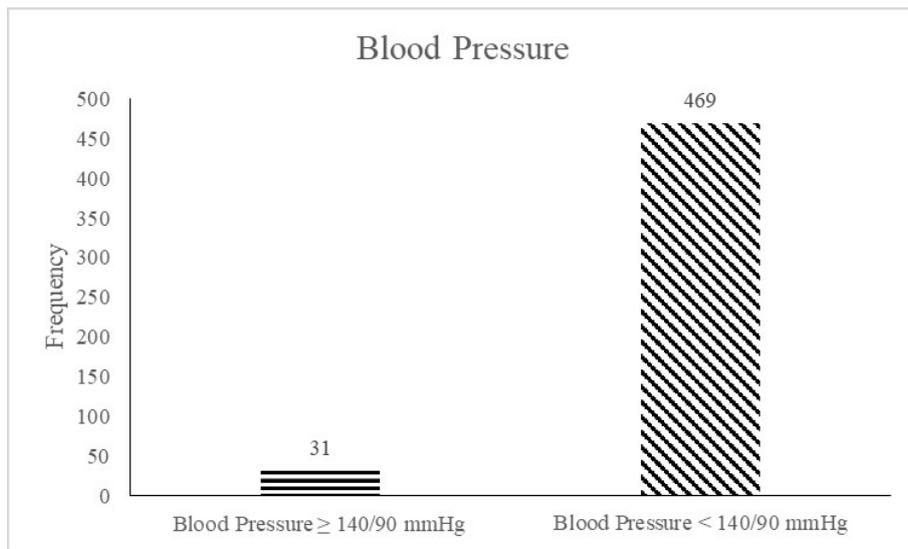
#### 3.1. Descriptive Statistics

Table 2. Description of the  $X_1$  Variable

Mean	Median	Min	Max	Standar Deviation
33.912	34	14	56	10.234

Based on Table 2, the highest age of recipients of COVID-19 vaccination in June-August 2021 at the Parumpanai Public Health Center, East Luwu Regency is

56 years old, while the lowest age is 14 years old. The average age of 500 COVID-19 vaccination recipients at the Parumpanai Public Health Center, East Luwu Regency is 33.912 years, the median of the age variable is 34 years and the standard deviation is 10.234 years which represents the distribution of the data to the average value.



Picture 1. Bar chart of  $X_2$  (Blood Pressure) variable

Based on Picture 1, the number of vaccination participants who can receive the vaccine is 469 people who are in the blood pressure group less than 140/90 mmHg, while 31 other people have blood pressure more than and equal to 140/90 mmHg, which means that they have to postpone giving the vaccine. Things that need to be done so that blood pressure is normal are regular exercise, not smoking, and dealing with stress. Foods that can maintain normal blood pressure such as vegetables, milk, fruits, and nuts should also be consumed.

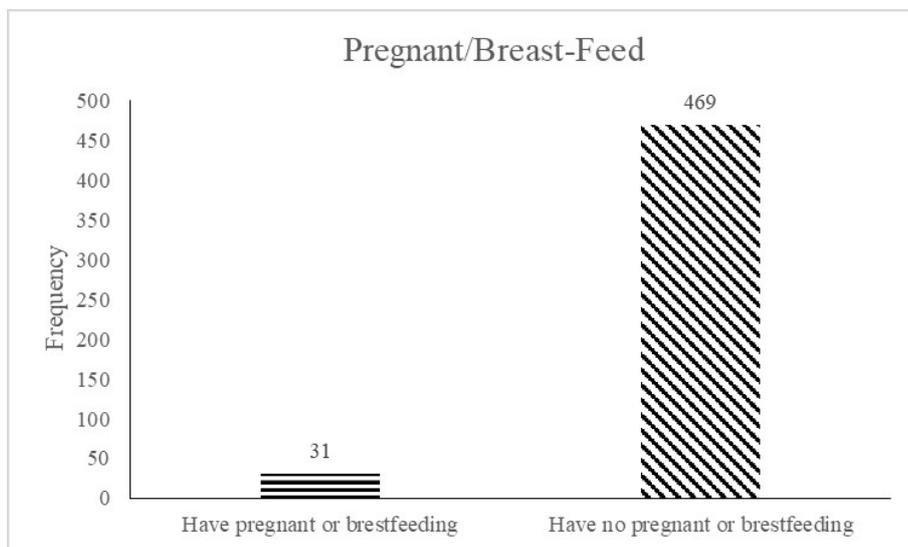
Based on Picture 2, it is known that the number of vaccination participants who are eligible to continue receiving the vaccine is 458 people who are in the group without a history of the disease, while 42 others have a history of diseases such as hypertension, diabetes mellitus and other blood disorders that must delay or even not be allowed to receive the vaccine. Both of these diseases can be prevented by maintaining a lifestyle such as eating regularly, getting enough rest, and exercising.

Based on Picture 3, it is known that the number of vaccination participants who meet the requirements to continue receiving the vaccine is 469 people who are in the non-pregnant/breastfeeding group, while 31 others are in the group of pregnant/breastfeeding women, which means they have to postpone giving the vaccine until the breastfeeding time is over.

Based on Picture 4, it is known that the number of vaccination participants who are eligible to continue receiving the vaccine is 452 people who are in the group



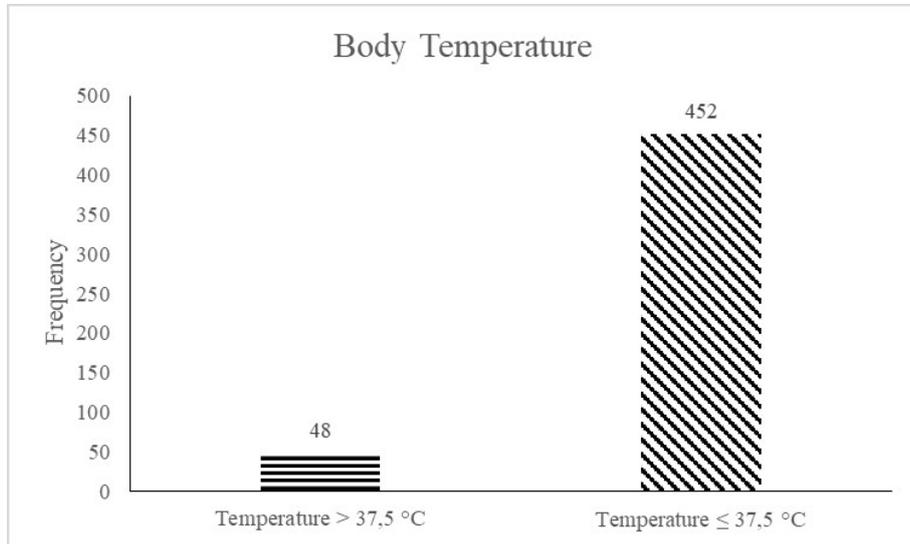
Picture 2. Bar chart of  $X_3$  (Disease History) variable



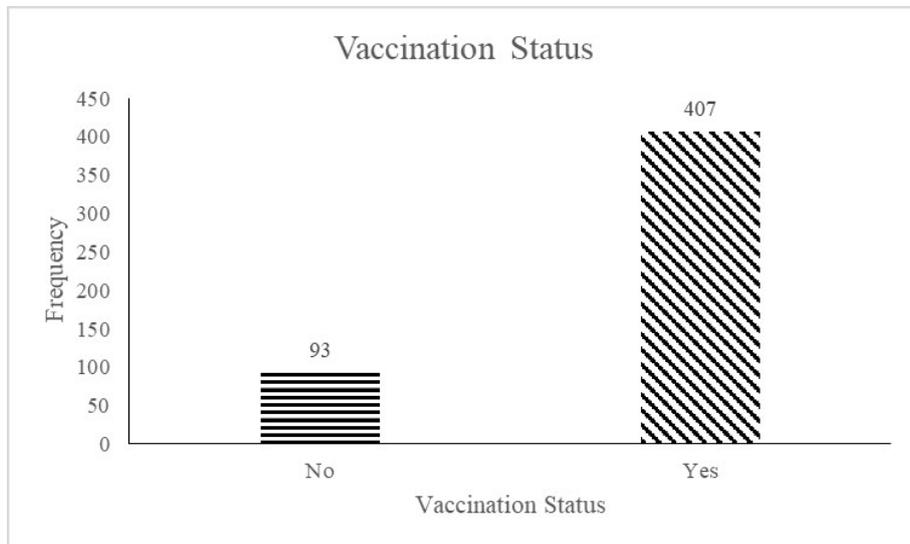
Picture 3. Bar chart of  $X_4$  (Pregnant/Breast-Feed) variable

with body temperature less than and equal to  $37.5^{\circ}\text{C}$ , while 48 other people have body temperature more than  $37.5^{\circ}\text{C}$  which means should delay the administration of the vaccine until the body temperature drops to less than and equal to  $37.5^{\circ}\text{C}$ . Efforts that can be made to maintain a normal body temperature are to drink lots of mineral water to avoid dehydration.

Based on Picture 5, it is known that the number of vaccination participants who



Picture 4. Bar chart of  $X_5$  (Body Temperature) variable



Picture 5. Bar chart of  $Y$  (Vaccination Status) variable

are eligible to continue receiving the vaccine is 407 people out of 500 COVID-19 vaccination participants or equivalent to 81% of the total registered in June-August 2021 at the Parumpanai Public Health Center, East Luwu Regency, while 93 people of 500 COVID-19 vaccination participants or 19% of participants had to delay giving the vaccine. This is because there are vaccination indicators that are not met, so vaccination must be postponed.

### 3.2. Data Normalization

In this step, only the  $X_1$  (Age) variable will be normalized. The normalization used is the min-max normalization method. Normalization is done with the aim that the  $X_1$  (Age) variable has a binary scale of 0 and 1 to facilitate the classification process.

Table 3. Result of normalization with min-max normalization method

$X_{1*}$	0.428	0
$X_{2*}$	0.476	0
$X_{3*}$	0.166	0
$\vdots$	$\vdots$	$\vdots$
$X_{499*}$	0.714	1
$X_{500*}$	0.666	1

### 3.3. Classification with Learning Vector Quantization Method

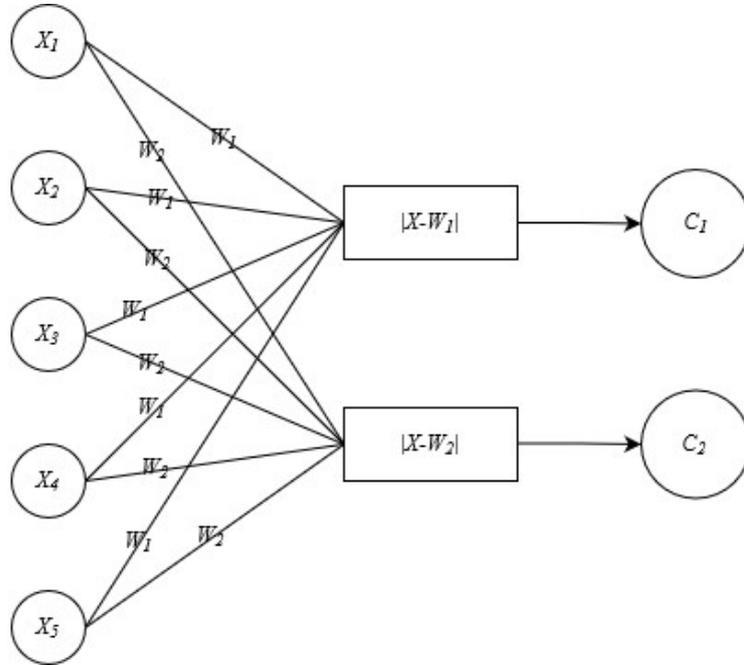
Learning vector quantization is a technique on artificial neural networks in classifying data into the right class. The input layer consists of 5 input units taken from the number of variables in the research. The outer layer consists of 2 output units taken from the number of classes in the dataset consisting of classes 1 (Continued) and 2 (Delayed). Before starting the classification, the LVQ network architecture in this research will be described as follows.

After creating the LVQ architecture network, the LVQ algorithm will then be run with the help of python software and predetermined parameters. After the algorithm stops, the following results are obtained.

Table 4. Result of normalization with min-max normalization method

Epoch	Aproprate Data
1	85
2	88
$\vdots$	$\vdots$
49	95
50	95

From Table 4, the algorithm of the LVQ method stops at a maximum of 50 epochs which results in 95 correctly classified data from 100 test data. With details of TP totaling 78 and TN totaling 17 and the rest are predicted wrongly.



Picture 6. Learning Vector Quantization network architecture

**3.4. Classification Performance from Learning Vector Quantization Method**

Performance measurement is an illustration of how accurately the model formed can predict the data. Classification performance is measured by 3 measurement metrics, namely accuracy, precision, and sensitivity.

Table 5. Confussion Matrix from LVQ classification method

Actual Class	Predict Class	
	Continued	Delayed
Continued	78	2
Delayed	3	17

It can be seen in Table 5 that the value of TP = 78, TN = 17, FP = 2, and FN = 3. TP shows that of the 100 test data used, there are 78 data that are predicted to be positive and correct in the actual class placement, while FN is 3 data that are predicted to be negative and incorrect in its actual class assignment. The FP shows that of the 100 test data used, there are 2 data that are predicted to be positive and wrong in the actual class placement, while TN is 17 data that are predicted to be negative and true to the actual data placement. Furthermore, the classification performance of the LVQ method will be measured using accuracy, precision, and sensitivity. The following are the results of the three performance measurements.

$$Accuracy = \frac{78 + 17}{78 + 17 + 2 + 3} = 0.95 = 95\%$$

$$Precision = \frac{78}{78 + 2} = 0.97 = 97\%$$

$$Sensitivity = \frac{78}{78 + 3} = 0.96 = 96\%$$

Based on results of accuracy, precision, and sensitivity on the LVQ method, it is found that the model built is very good because the results of measuring performance with these 3 metrics are very high or greater than 90%. An accuracy of 95% is the percentage of vaccination recipients who predicted continued and delayed correctly. The precision of 97% is the percentage of correctly predicted advanced vaccination recipients of the total predicted advanced vaccination recipients. The sensitivity of 96% is the percentage of vaccination recipients with the advanced status that is correctly predicted from all vaccination recipients whose actual data are at advanced status.

#### 4. Conclusion

From the 100 data used as test data, the model formed by the LVQ classification can predict 95 data with accuracy, precision and sensitivity values of 95%, 97% and 96%, so it can be concluded that the performance of the model from LVQ classification is very good and can be used to classify COVID-19 vaccination data at the Parumpanai Public Health Center, East Luwu Regency. With classification results with these conditions, the Parumpanai Public Health Center in East Luwu Regency can maintain the performance of medical workers in correctly recording vaccination recipients so that there are no unwanted errors when administering vaccination doses.

#### Daftar Pustaka

- [1] Makmun, A., Hazhiyah, S. F., 2020, Tinjauan Terkait Pengembangan Vaksin Covid 19, *Molucca Medica*, **2020** : 52 – 59
- [2] Febriyanti, N., Choliq, M. I., and Mukti, A. W., 2021, Hubungan tingkat pengetahuan dan kesediaan vaksinasi covid-19 pada warga kelurahan dukuh menanggal kota surabaya, *SNHRP*, **2021** : 36 – 42
- [3] Wang, J., Jing, R., Lai, X., Zhang, H., Lyu, Y., Knoll, M. D., and Fang, H., 2020, Acceptance of COVID-19 Vaccination during the COVID-19 Pandemic in China, *Vaccines*, **8(3)** : 482
- [4] Haynes, B.F., Corey, L., Fernandes, P., Gilbert, P.B., Hotez, P.J., Rao, S., Santos, M.R., Schuitemaker, H., Watson, M. and Arvin, A., 2020, Prospects for a safe COVID-19 vaccine, *Science translational medicine*, **12(568)** : eabe0948
- [5] Ichsan, D.S., Hafid, F., Ramadhan, K. and Taqwin, T., 2021, Determinan kesediaan masyarakat menerima vaksinasi Covid-19 di Sulawesi Tengah, *Poltekita: Jurnal Ilmu Kesehatan*, **15(1)** : 1 – 11

- [6] Soares, P., Rocha, J.V., Moniz, M., Gama, A., Laires, P.A., Pedro, A.R., Dias, S., Leite, A. and Nunes, C., 2021, Factors associated with COVID-19 vaccine hesitancy, *Vaccines*, **9(3)**: 300
- [7] Han, J., Pei, J. and Tong, H., 2022, *Data mining: concepts and techniques*, Morgan kaufmann
- [8] Tomia, S., Leleury, Z.A. and Aulele, S.N., 2017, Perbandingan Metode Jaringan Saraf Tiruan Backpropagation Dan Learning Vector Quantization Dalam Deteksi Hama Pengerek Batang, *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, **11(1)**: 13 – 26
- [9] Heo, S. and Lee, J.H., 2018, Fault detection and classification using artificial neural networks, *IFAC-PapersOnLine*, **51(18)**: 470 – 475
- [10] Jo, J.M., 2019, Effectiveness of normalization pre-processing of big data to the machine learning performance, *The Journal of the Korea institute of electronic communication sciences*, **14(3)** : 547 – 552
- [11] Khan, S.A and Rana, Z.A., 2019, Evaluating Performance of Software Defect Prediction Models Using Area Under Precision-recall Curve (AUC-PR), *2nd International Conference on Advancements in Computational Sciences (ICACS)*, pp. 1-6