



Visitor satisfaction prediction of the ‘pantai pohon cinta’ beach tourism using the backpropagation algorithm with particle swarm optimization feature selection

Annahl Riadi^{a, 1}; Marniyati H. Botutihe^{a, 2, *}

^a *Fakultas Ilmu Komputer Universitas Pohuwato, Jl. Trans Sulawesi No 147 Kec. Marisa, Gorontalo 96266, Indonesia*

¹ *phyloanna@gmail.com*; ² *marniyati.h.botutihe@gmail.com*;

* *Corresponding author*

Article history: Received February 19, 2021; Revised July 04, 2021; Accepted July 04, 2021; Available online August 07, 2021

Abstract

This study focuses on the visitors of Pohon Cinta beach tourist area. This beach is one of the potential tourism objects in Pohuwato Regency. The main problem that frequently occurs is that many visitors cannot directly convey their impression when visiting and enjoying the beauty of the Pohon Cinta beach. The government needs to know the level of visitor satisfaction to attempt to improve and develop the Pohon Cinta beach tourist attraction. Thus, to solve the problem above, a method that can help predict visitor satisfaction is needed. This study aims to measure visitor satisfaction through predictions using the Backpropagation algorithm and PSO feature selection to assist the government in developing tourism potential in Pohuwato Regency. The method used is the backpropagation algorithm for prediction and Particle Swarm Optimization which is considered effective in overcoming optimization problems. This algorithm is considered capable of solving problems in the backpropagation algorithm. The accuracy value of the backpropagation algorithm model is 84.67%, the accuracy value of the PSO-based backpropagation algorithm model is 85.00%, and the difference in accuracy is 0.33. The results of the application of the Backpropagation algorithm and Particle Swarm Optimization can increase the predictive accuracy value of visitor satisfaction at the Cinta Tree Beach tourist attraction.

Keywords: Prediction; Visitors; Tour; Backpropagation; PSO

Introduction

Tourism development is basically an effort to develop and utilize tourist objects and attractions. Tourism is an important part of the tourism industry, and one of the reasons tourists travel. [1]. Tourist attraction managers will continue to strive to increase tourism potential by improving the satisfaction level of tourism object visitors. One of the areas in Gorontalo Province that is trying to develop its tourism potential is Pohuwato Regency, which offers many potential tourist attractions. Pohon Cinta Beach is one of the well-known and potential tourism targets with potential development so that it becomes the main tourist destination in Pohuwato Regency. [2]. The existence of the Pohon Cinta Beach area is a matter of pride for the people of Pohuwato, in addition to its strategic location in the middle of the city of Pohuwato district, precisely in the Marisa area and is a landmark in this district [3].

The purpose of this study is to measure the level of visitor satisfaction through predictions so that it can assist the government in developing tourism potential in Pohuwato Regency. The results of the application of the Backpropagation algorithm and Particle Swarm Optimization are able to increase the predictive accuracy value of visitor satisfaction at the Cinta Tree Beach tourist attraction.

The main problem that frequently occurs is that many visitors cannot directly convey their impression when visiting and enjoying the beauty of the Pohon Cinta beach. The government needs to know the level of visitor satisfaction to attempt to improve and develop the Pohon Cinta beach tourist attraction. Thus, to solve the problem above, a method that can help predict visitor satisfaction is needed.

One of the methods that is often used in prediction is Neural Network with Backpropagation algorithms, such as research that has been done using the Back-propagation Neural Network Algorithm Based on Particle Swarm Optimization, with 77.96% accuracy results and 0.814 AUC [4]. Research with Neural Network Backpropagation Method, with the final result of this research, is the best accuracy value [6][7][8]. Neural networks have good performance and have advantages in non-linear prediction and the ability to tolerate errors [9]. The backpropagation algorithm is also believed to be able to reduce the error rate by adjusting the weight according to the target and the expected output. However, the weakness of the backpropagation algorithm is that the optimization used is less efficient and the need for large training data [10]. One effective algorithm in overcoming optimization problems is

Particle Swarm Optimization (PSO), which is considered capable of solving problems in the backpropagation algorithm [11].

Method

The method used in this research is the Backpropagation method. Training data and testing data are divided by the composition of 70% training data and 30% testing data from the total data obtained.

A. Neural Network

An artificial neural network is a parallel distributed processor, consisting of simple units, has the ability to store knowledge obtained through experiments, and is ready for various purposes (S. Haykin, 1999) [12]. This neural network simulates human thinking from a perspective:

- Knowledge is acquired from the environment through a network learning process.
- The strength of the connection between units is the synaptic weight, which is used to store the knowledge acquired by the network.

In 1943, Mc. Culloch and Pitts introduce a mathematical model show in **Figure 1** that simplifies the actual structure of a nerve cell

$$y=f(\sum_{i=1}^n [W_i X_i]) \tag{1}$$

The correlations around the three components in the above formula are: the vector shape of the signal x with size n (X_1, X_2, \dots, X_n) T will be strengthened by synapses w (W_1, W_2, \dots, W_n) T. In addition, the accumulation of this reinforcement will be altered by the activation function f. If the cumulative amplification rate of the signal has exceeded a certain limit, the f function will be monitored, and the neuron cell initially at the "0" state will emit a "1" signal. According to the output value (y), the neuron can be in two states: "0" or "1". A neuron is said to be in the firing state if it produces an output of "1".

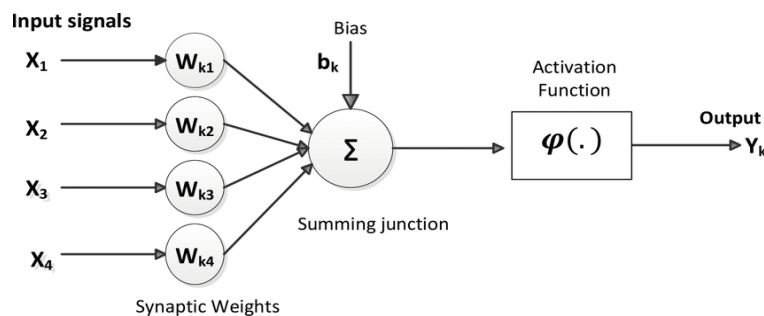


Figure 1. McCulloch and Pitts Neuron Model [13]

B. Metode Backpropagation

The backpropagation method is a guided training method on an artificial neural network that has the characteristic to minimize errors in the output obtained from the network. In the following figure, the input units are represented by X, the hidden units are represented by Z, and the output units are represented by Y. The weights between X and Z are represented by v, and the weights between Z and Y are represented by w. Backpropagation Network Architecture show in **Figure 2**.

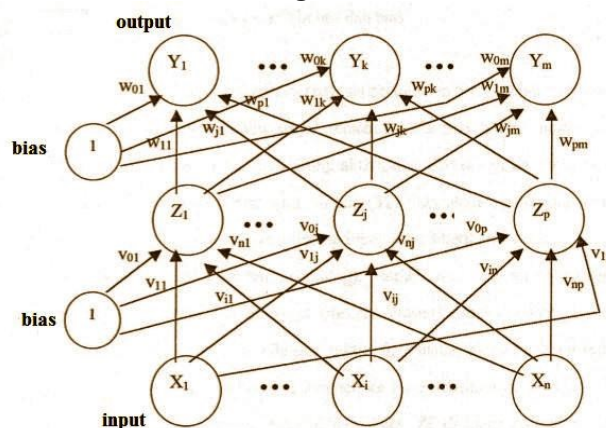


Figure 2. Backpropagation Network Architecture [13]

The implementation of a backward propagation network includes two stages:

1. Training stage, providing some data and training objectives
2. The test or evaluation stage is carried out after the training stage is complete

Basically, the backpropagation method training includes three steps, including:

1. Input data to network input (feedforward)
2. Computing related errors and back propagation
3. Weight and bias adjustment.

To make predictions in the Backpropagation Neural Network System, data is required as the input in the processing, in order to generate output [14].

C. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is an optimization method based on the behavior of a collection of animals. A particle in space has a position and each position in the search space is an alternative solution that can be evaluated using an objective function. Each particle can adjust its own position and velocity in a way that each particle conveys its best information to the other particles. Therefore, each particle has a tendency to go towards the position that is considered the best. The particle itself is each individual in a group [15].

D. Confusion Matrix

This is the method used to measure the performance of a predictive method and is used in calculating accuracy in data mining or decision support systems. The confusion matrix contains information on the prediction results [15].

E. AUC (Area Under Curve)

AUC shows the relationship between the test and the accuracy measurement results. A point at a value of 1 is said to be a True Positive level (TP), a point at a value of 0 is said to be a False Positive (FP) level. The point (0.1) is a good prediction classification because both positive and negative cases are said to be true (True). While for (1.0), all prediction classifications are said to be incorrect (False). **Figure 3.** is a display of the auc curve. This curve will display the test results and accuracy.

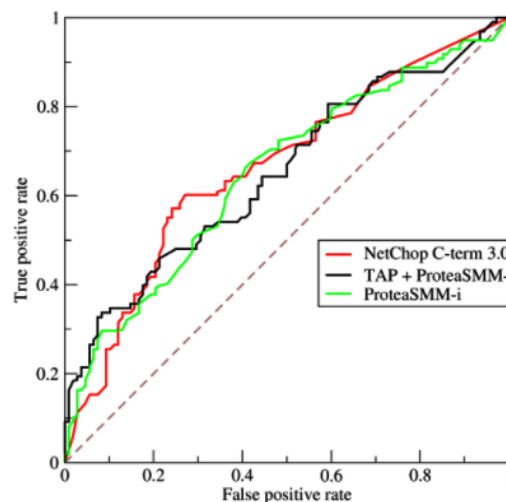


Figure 3. AUC curve

This study uses a descriptive research method as it is considered suitable for the problem under study. The research was conducted by taking samples from respondents or visitors around the Pohon Cinta Beach tourism object. The algorithm used is the Backpropagation Algorithm and PSO. The main steps of the backpropagation algorithm include input, tracking errors, and then setting the weights. The following are the stages of the research that will be carried out:

1. The first stage of data collection by taking an initial sample of 100 respondents to be used as a dataset
2. The second stage is preprocessing by normalizing and continuing with experiments until testing
3. The third stage is evaluating and validating the research, followed by writing a report on the results

F. Research Procedure

The procedure of this research can be seen in **Figure 4.**

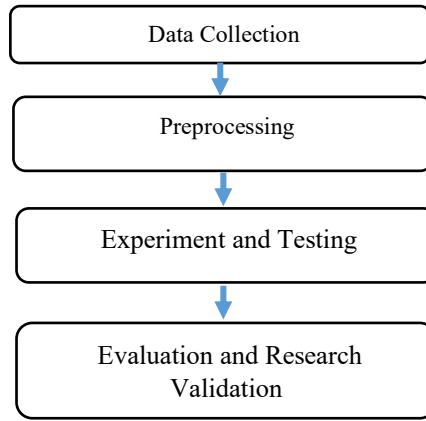


Figure 4. Research Procedure

G. Experimental Model

The following is the experimental model used in this study show in Figure 5.

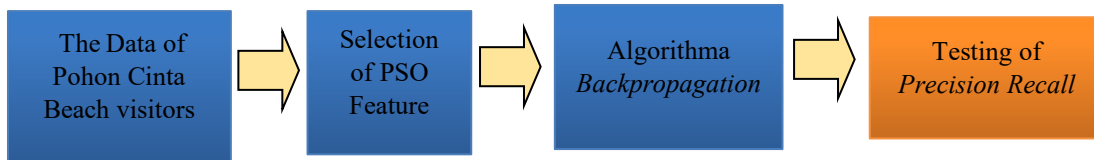


Figure 5. Experimental Model

Figure 6 is a description of the experimental model used in the study

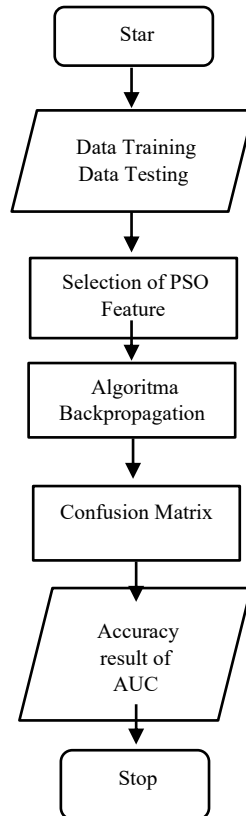


Figure 6. Flowchart Analysis

Results and Discussion

A. Data Collection

The processed data is data obtained from questionnaires distributed online via a link via Google Form <https://forms.gle/J9fmScqpCHKvyyhf6> so that 56 records and 25 fields are obtained.

B. Preprocessing

The data obtained from the questionnaire was processed from a CSV file into an xlsx file and then separates between training data and testing data.

C. Processing

a) Experimental Design

This study implements the backpropagation algorithm model using the particle swarm optimization (PSO) feature selection. The tourist attractions dataset comes from a CSV file that has been converted to xlsx, then imported into Rapidminer to be used by PSO to select features, and uses the backpropagation algorithm for classification, in order to produce test results, as shown in the following **Figure 7**.

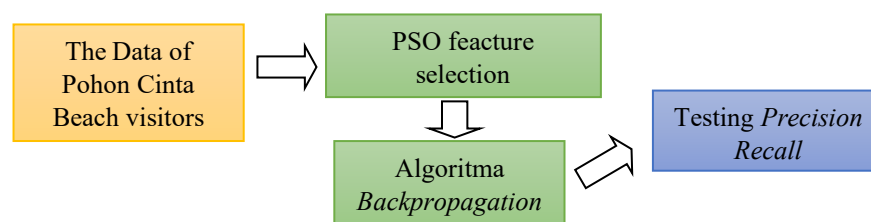


Figure 7. Experimental Model

b) Experiment and Testing Method

In this study, the value of the training cycle, momentum, and learning rate was determined by conducting experiments, by entering the value of the training period range of 500, the value of learning speed 0.3, and the momentum value of 0.2. **Table 1** is the results of the experiments that have been carried out.

Table 1. Results of the experiments that have been carried out

Parameter BPNN		Optimize Weights PSO	
Training Cycles	: 500	Population Size	: 5
Learning Rate	: 0.3	Maximum Number of Generation	: 30
Momentum	: 0.2	Inertia Weight	: 1.0
X-Validation		Local Best Weight	: 1.0
Number of Validation	: 10	Global Best Weight	: 1.0
Sampling Type	: Shuffled Sampling	Min Weight	: 0.0

The results of the complete backpropagation model testing are the level of measurement accuracy and AUC (Area Under Curve). In data mining, the assessment of prediction results uses AUC. The higher the AUC result, the better the prediction result. Based on the Confusion Matrix of the processed training data, the following results were obtained:

accuracy: 84.67% +/- 15.79% (mikro: 83.93%)			
	true Ya	true Tidak	class precision
pred. Ya	47	9	83.93%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

Figure 8. Backpropagation Accuracy Value

Figure 8 shows the results of the backpropagation accuracy values obtained from the experiments that have been carried out.

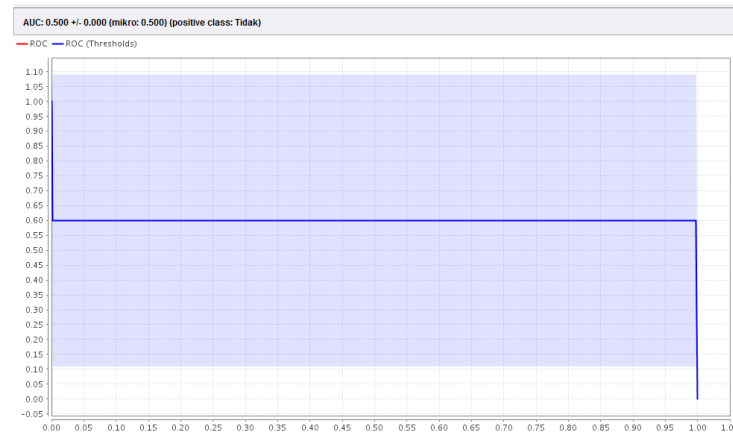


Figure 9. AUC (Area Under Curve) Backpropagation value

Figure 9 is the display of the AUC backpropagation values obtained from the experiments that have been carried out.

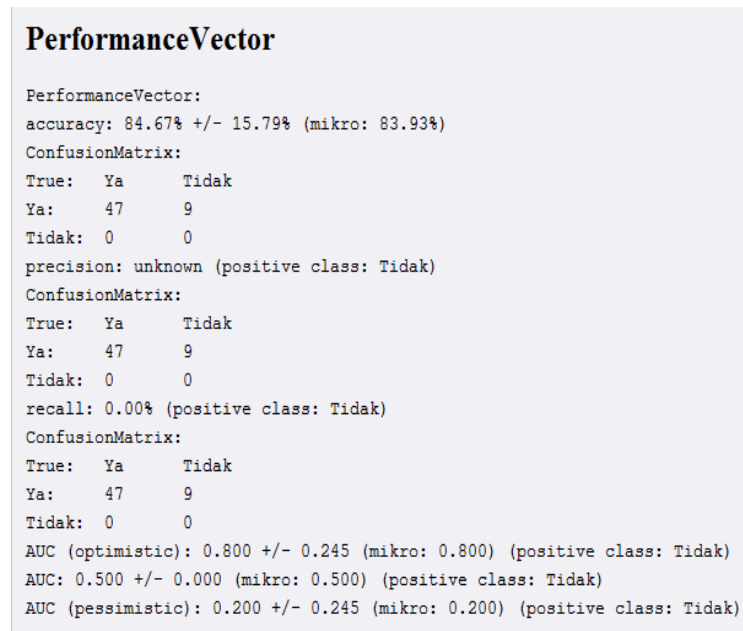


Figure 10. Performance Vector Backpropagation

Figure 10 shows an explanation of the results of the experiment and back-propagation testing

c) *Particle Swarm Optimization (PSO) Based Backpropagation Test Results*

The results of the completed model test are measurement accuracy and AUC (Area Under Curve) level. Confusion Matrix according to the processed data obtained the following results:

accuracy: 85.00% +/- 17.40% (mikro: 83.93%)			
	true Ya	true Tidak	class precision
pred. Ya	47	9	83.93%
pred. Tidak	0	0	0.00%
class recall	100.00%	0.00%	

Figure 11. Backpropagation accuracy value with PSO

Figure 11 is a display of the results of the backpropagation accuracy value using the PSO feature selection obtained from the experiments that have been carried out.

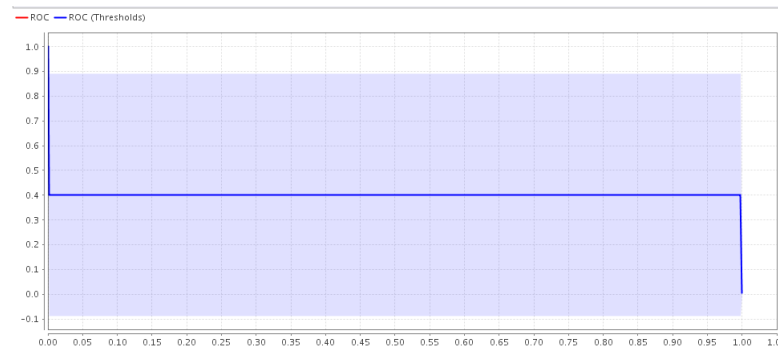


Figure 12. Value of AUC (Area Under Curve) Backpropagation with PSO

Figure 12 is the display of the results of the AUC backpropagation value using the PSO feature selection obtained from the experiments that have been carried out.

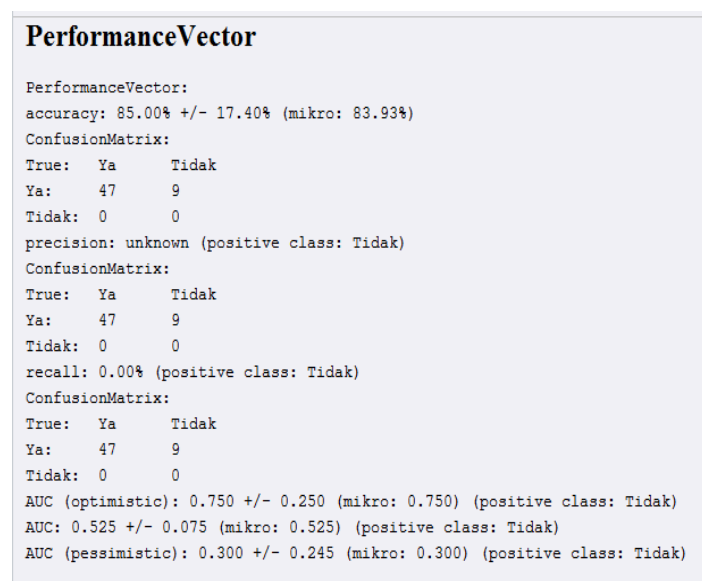


Figure 13. Performance Vector Backpropagation with PSO

Figure 13 shows an explanation of the results of the experiment and backpropagation testing with PSO feature selection.

D. Evaluation and Validation of Results

In this study, the implementation of the Backpropagation algorithm is carried out by determining the value of the training cycle, learning rate, and previous momentum. After getting the maximum accuracy and AUC values, the maximum value of the training period, the learning speed, and momentum will determine the hidden layer's size. So, the application of the backpropagation algorithm with PSO is based on the value of the training cycle, learning rate, and momentum in the algorithm in increasing accuracy and precision and recall values for evaluation.

Evaluation analysis and mode analysis, when viewed from the test results above, evaluation using the confusion matrix and ROC curve proves that the results of testing the PSO-based neural network algorithm have a higher accuracy value than the neural network algorithm. As shown in Figure 8 and **Figure 11**, the results obtained different accuracy due to feature selection was used to improve the effectiveness and efficiency of the algorithm. The accuracy value of the neural network algorithm model is 84.67% and the accuracy value of the PSO-based neural network algorithm model is 85.00% with an accuracy difference of 0.33%, can be seen in the **Table 2**.

Table 2. Accuracy Value

	Accuracy	AUC
Neural Network Backpropagation	84.67%	0.500
Neural Network Backpropagation with PSO	85.00%	0.525

Based on the experiment, it can be concluded that PSO can solve the optimization problem of the backpropagation algorithm in predicting tourist satisfaction with the attraction of Pohon Cinta Beach. From the increase in accuracy, it can be seen that the backpropagation algorithm generates an accuracy of 84.67%, and AUC of 0.500. The reason for the increase is that the PSO optimization method will find for the best solution until all particles have the same solution or reach the maximum number of iterations, which can increase the accuracy value.

Conclusion

The research concludes that there is an increase in accuracy using Backpropagation with PSO. The accuracy value of the back-propagation algorithm model is 84.67%, the accuracy value of the PSO-based back propagation algorithm model is 85.00%, and the difference in accuracy is 0.33. With this research, the Pohuwato Regency government can easily find out the level of visitor satisfaction of Pohon Cinta Beach, so that the Pohuwato Regency Government can quickly improve and attempt to develop Pohon Cinta Beach Tourism Object.

References

- [1] Nawangsari, D. Muryani, C. dan Utomowati, R, "Pengembangan Wisata Pantai Desa Watu Karung dan Desa Sendang Kabupaten Pactan," *Jurnal GeoEco*. Volume 4, Nomor 1, 31-40, 32. 2018.
- [2] Alan. "Layak Jadi Destinasi Andalan Gorontalo," *Pohuwato: Portal Pohuwato*. 2016. 24 Januari, 2020. <https://pohuwatokab.go.id/> [akses 24 Januari, 2020]
- [3] Nurmiah. Rahmawati, "Persepsi Pengunjung Terhadap Visual Fungsional Kawasan Pasca Revitalisasi," *Gorontalo: Journal of Infrastructure & Science Engineering* Vol 2 Nomor 2 Oktober 2019
- [4] Rizky. Aulia, "Prediksi Kemacetan Angsuran Leasing Motor Menggunakan Algoritma Backpropagation Neural Network Berbasis Particle Swarm Optimization," *Banjarmasin: Jurnal Technologia* Vol 8, No.4, Oktober-Desember 2017
- [5] Pujianto. Ade. Kusri. Andi Sunyoto, "Perancangan Sistem Pendukung Keputusan Untuk Prediksi Penerima Beasiswa Menggunakan Metode Neural Network Backpropagation," *Yogyakarta: Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* Vol. 5, No. 2, Mei 2018
- [6] Sari, Febrina. "Implementasi Data Mining Dalam Menganalisis Tingkat Kepuasan Pelanggan Menggunakan Metode Rough Set," *Jurnal Buana Informatika* Volume 8 Nomor 1. 2017
- [7] Dika. Harry dkk, "Implementasi Algoritma C4.5 Terhadap Kepuasan Pelanggan," *Bandung: Universitas Islam Bandung*. 2016
- [8] Darmawan. Agus. Dkk, "Implementasi Data mining Menggunakan Model SVM Untuk Prediksi Kepuasan Pengunjung Taman Tabebuaya," *Indraprasta : Jurnal String* Vol. 2 No. 3 April 2018
- [9] X. Shao. 2011, "Based on Two Swarm Optimized Algorithm of Neural Network to Prediction the Switch's Traffic of Coal," *2011 International Symposium on Computer Science and Society*, pp. 299–302, Jul. 2011.
- [10] A. Idris, M. Rizwan, and A. Khan, "Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies," *Computers & Electrical Engineering*, vol. 38, no. 6, pp. 1808–1819, Nov. 2012R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.*, in press.
- [11] Asih. Imelda, dkk, "Penerapan Algoritma Backpropagation dalam Memprediksi Persentase Penduduk Buta Huruf di Indonesia," *Jurnal Informatika Upgris*. DOI: 10.26877/jiu.v4i2.2423. 2018
- [12] Handayanto. Rahmadya Trias dan Herlawati, "Machine Learning Berbasis Desktop dan Web dengan Metode Jaringan Syaraf Tiruan Untuk Sistem Pendukung Keputusan," *Jakarta: Jurnal Komtika (Komputasi dan Informatika)*, Vol. 4 No. Mei 2020
- [13] Khademi. F, dan Jamal, S. M, "RESEARCH PAPERS PREDICTING THE 28 DAYS COMPRESSIVE STRENGTH OF", *I-Manager's Journal on Civil Engineering*, 6(August), 1–7. 2016
- [14] Asril, "Jaringan Syaraf Tiruan Backpropagation Untuk Prediksi Jumlah Pengunjung Kolam Renang Kolam Renang", *Jurnal SIMTIKA* Volume 2 No 1. 2019
- [15] Wang, X., Wen, J., Zhang, Y., & Wang, Y. (2014). Optik Real estate price forecasting based on SVM optimized by PSO. *Optik - International Journal for Light and Electron Optics*, 125(3), 1439–1443.
- [16] Apandi. Tri Herdiawan. Castaka Agus Sugianto, "Algoritma Naive Bayes untuk Prediksi Kepuasan Pelayanan Perekaman e-KTP," *Bandung: JUITA: Jurnal Informatika e-ISSN: 2579-9801; Volume 7, Nomor 2, November 2019*