Journal of Applied Science and Advanced Technology
Journal Homepage : https://jurnal.umj.ac.id/index.php/JASAT

Weather Forcast Optimization Using Learning Vector Quantization Methods with Genetic Algorithms

Siska Andriani[1*], Kotim Subandi[1]
[1]Department of Computer Science, Universitas Pakuan Bogor, Jl. Pakuan Ciheuleut Bogor, Indonesia

ARTICLE INFO

JASAT use only:
Received date : 15 August 2020
Revised date : 19 September 2020
Accepted date : 25 October 2020

Keywords:
Optimization
Weather forcast
Learning Vector Quantization
Genetic Algorithms

ABSTRACT

Weather forecasting is one of the important factors in daily life, as it can affect the activities carried out by the community. The study was conducted to optimize weather forecasts using artificial neural network methods. The artificial neural network used is a learning vector quantization (LVQ) methods and genetics algorithms (GA). BMKG weather data was originally modeled using the LVQ method, then also created the LVQ Method Optimization weather forecast model using GA. Data attributes consist of numeric and category. Numeric attributes as input parameters are: temperature, evaporation, sunlight, humidity and rainvol. While the categorical attributes are output from weather forecasts include: Cloudly (C), Partly Cloudly (PC), Sunny (S), Rain (R) and Cloudly rain (CR). Sample data used is 1096 data. Both models were tested so that they obtained 72% accuracy results for weather forecast models using the LVQ method and 73% of the weather forecast accuracy results that were optimized using GA. The results have not achieved the most optimal results because it turns out that citeko region weather data is not suitable for use in both methods. Because the data has an imbalance in the amount of data per class.

INTRODUCTION

Optimization is the process of finding the best value based on predetermined criteria. Forecast is an activity that predicts an event in the future. Weather is a state of the atmosphere which is expressed by the value of various parameters, including temperature, pressure, wind, humidity and various phenomena of rain, in a place or region during a short period of time, minutes, hours, days, months, seasons, years. Weather forecasting is an activity that predicts the weather in the future, weather forecasts can predict the potential for rain that will occur in the future [1].

By applying artificial technology, one of the artificial methods will be used to perform weather forecasts, namely the Artificial Neural Network (ANN) and Genetic Algorithm (AG) methods. Artificial Neural Network is a suitable method in terms of weather forecasting based on existing weather parameters. The ANN method used is Learning Vector Quantization (LVQ). The prediction classes to be made are cloudy, rainy, cloudy rain, sunny and cloudy. To find out the optimal prediction results in this study, prediction optimization will also be carried out using Genetic Algorithms, so that we can see the comparison of the prediction optimization results using the LVQ and the combined LVQ-GA method

Basically, the method mentioned above has been widely used in terms of prediction, one example is the optimization of predictions used for optimization of rainfall forecasting [2], in this study back propagation ANN and genetic algorithms were used. Another research that has been done previously is the optimization of the determination of the quality of cow's milk [3]. In this study, the LVQ method was used which the weight vector of the LVQ was optimized using a genetic algorithm. In addition, the combination of the LVQ method

*Corresponding author.
E-mail address: siska.andriani@unpak.ac.id
DOI: https://dx.doi.org/10.24853/JASAT.3.2.45-50

and genetic algorithm has been applied for the classification of electrical signal quality [4]. The results obtained in previous research have shown good results for the use of ANN and genetic algorithms, but whether the prediction results are good or not depends on the type of data used.

To make predictions, it requires a classification of the elements required in terms of prediction of the field. In predicting the weather, at least one year of data is needed, the data is taken every day with a set hour, the data is obtained from the Meteorology and Climatology Agency (BMKG) where this research was conducted. The data used to generate training data in this study were BMKG daily weather data for 3 years. ANN in modeling relies on previous data which is used as learning data which can then be used for future predictions because of learning from past data. The genetic algorithm performs optimization using the dominant factors used in weather prediction in Indonesia.

EXPERIMENTAL METHOD

This research was conducted using daily weather data of BMKG Class III Citeko Cisarua Bogor. The data used as training data generator is BMKG daily weather data for 3 years. The research stages carried out are: (1) data collection, (2) data sharing, (3) modeling using the LVQ method, (4) weather prediction optimization modeling using the LVQ method with genetic algorithms, (5) testing, analysis and evaluation. The flow diagram for the stages is shown in Figure 1.

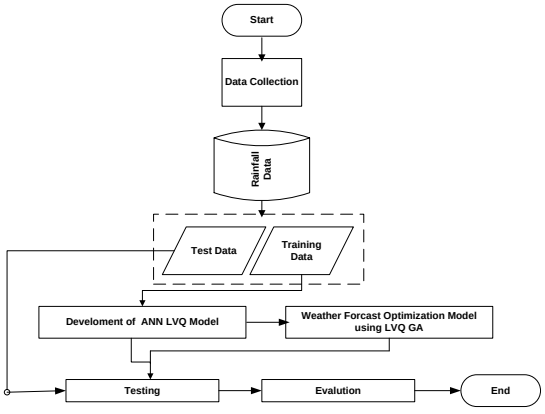

Fig.1. Flowchart of Research Stage

Data Collection

The data used in this study is the daily weather data of the BMKG Class III Citeko Bogor for 3 years. BMKG Class III Citeko is located in Citeko Village, Cisarua District, Bogor.

Data Sharing

Data sharing is done to divide weather data into training data and test data. The training data will be used to generate the model and the test data will be used for evaluating the performance of the predictive optimization model made. Data sharing uses a percentage split with the provisions of 70% for training data and 30% for test data.

Weather Forecast Modeling using LVQ

An Artificial Neural Network is a computer system that mimics the thinking of the human brain that is capable of completing a number of calculation processes. There are several types of neural networks but almost all of them have the same components. Neural network consists of several neurons like the human brain, and these neurons are related to one another or known as weights. LVQ is a method for training supervised competitive layers. The competitive layer will learn automatically to classify a given input vector. Figure 2 shows the LVQ network in the research carried out, namely with 6 units in the input layer, and 2 units (neurons) in the output layer [5]. Learning Vector Quantization (LVQ) is ANN with higher accuracy and faster computation time [6].

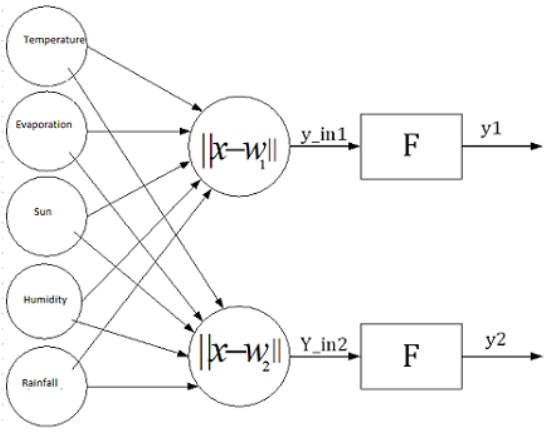

Fig. 2. Network Architecture LVQ

The Learning Vector Quantization (LVQ) method is divided into two stages, namely training and testing. The following are the steps for the algorithm:

a. Training Algorithm

- Set the initial weight of the j-input variable towards the i-class (cluster): Wij, with i = 1,2,…, K; and j = 1,2,…, m.
- Set Maximum epoh: MaxEpoh
- Set the learning rate Parameter: α
- Set learning rate reduction: Decα
- Set the allowed learning rate: Minα
- Enter input data:; with i = 1,2,…, n; and j = 1,2,…, m
- Enter targets in the form of classes:; with k = 1,2,…, n
- Set initial condition: epoh = 0
- Work if: (epoh ≤ MaxEpoh) and (α ≥ Minα)

a. epoh = epoh + 1

b. Work for i = 1 through n

1. Define J such that | - | minimum; where j = 1,2,…, K

2. Fix it with the following conditions:

If T = then, = + α (-)

If T ≠ then, = - α (-)

3. Reduce the α value (α reduction can be done by: α = α - Decα; or by: α = α - α * Decα)

- After training, the final weight (W) will be obtained, this weight will be used for the simulation or testing process [5].

b. Simulation (testing)Algorithm

- Enter the data to be tested, for example:; with i = 1,2, .., np; and j = 1,2,…, m.
- Work for i = 1 to np

1. Determine such J to Q% - Q% minimum; with j = 1,2,…, K

2. J is class for

The structure of the LVQ ANN used in this study is shown in **Table 1**.

Table 1. ST LVQ Structure Using

Characteristics	Specification
Architecture	2 Layer
Neuron Input	5
Neuron Output	1 (wheather prediction)
Epoch	1000
Learning rate	0.01 0.1 0.2
Lr_declay	0.2
Error Tolerance	0.001
Artificial Neural Network Algirithm	Learning Vector Quantization

Weather Forecast Optimization Modeling using the LVQ-GA

Genetic Algorithms are inspired by Darwin's theory of evolution. GA is a search algorithm based on natural system mechanisms, namely genetics and natural selection. In the GA application, the solution variable is coded into a string structure that represents a sequence of genes, which is a characteristic of the problem solution [2].

GA differs from conventional search techniques, GA is derived from a randomly generated set of solutions. This set is called the population. The individual represents a problem that has been coded (encoded) into chromosome form. Individuals will compete for reproduction, consisting of crossovers and mutations, to produce offspring that have better traits or genes than their parents. Furthermore, individuals and offspring will be selected based on the fitness that has been evaluated previously. Individuals who have good fitness will pass the selection to the next generation [7].

The selection used is elitism. Sorting fitness from largest to smallest then taking individuals with high fitness, this is based on what has been done by Arniantya [8] and Dwi [9].

At this stage, modeling the weather forecast by optimizing the weight of the LVQ using GA. Increasing the performance of LVQ can be done by optimizing the weight, one way is the genetic algorithm method [10].

Testing, Analysis and Evaluation

The test is done by comparing the actual value with the predicted value. Furthermore, the evaluation of the prediction results is carried out with a confusion matrix. Confusion matrix is a method for evaluating the success rate of a data when it is called against its data class [11].

RESULTS AND DISCUSSION

The results of the study are divided into 2, namely the results of forecast modeling using the LVQ method and the results of weather forecast optimization modeling using the LVQ-GA. The data attributes consist of numeric and categoric.

The numerical attributes as input parameters are: temperature, evaporation, sunlight,

 ISSN : 2622-6553 (Online)

humidity and rainvol. While the categorical attributes, namely as the output of the weather forecast, include: Cloudly (C), Partly Cloudly (PC), Sunny (S), Rain (R) and Cloudly rain (CR). The sample data used is 1096 data, the data for each class is shown in Figure 3.

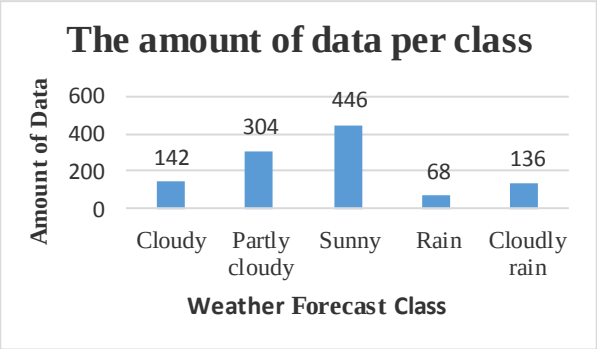

Fig. 3. The Number of sample for each class

Weather forecast using LVQ

Weather forecasting using the LVQ method is carried out by generating weather training data with the ANN structure shown in Table 1. After that, tests are carried out on the model that has been built so that the weather forecast results are obtained. Comparison of weather prediction results using the LVQ method with actual data can be seen in Figure 4.

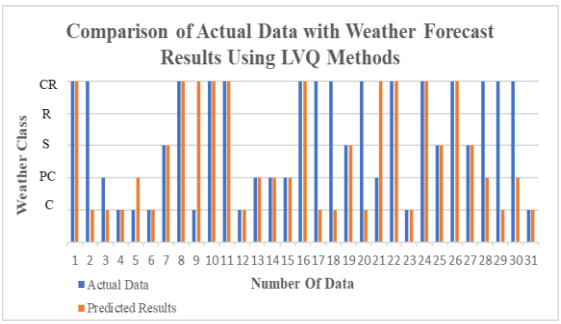

Fig. 4. Comparison of the predicted LVQ results with actual data

Figure 4 shows the comparison of the actual data with the predicted data. From the sample test data taken as many as 31 data, it can be seen that there are some results that are not in accordance with the actual data, namely the test data 2, 3, 5, 9, 17, 18, 20, 21, 28, 29 and 30.

Optimization of Weather Forecast using the LVQ method with GA

The weather prediction optimization modeling uses the LVQ-GA method, namely in this case GA plays a role in optimizing the best weight generated by LVQ so that it gets new weights which will later be used to perform weather forecasts.

The GA parameters used in this study are as follows:

1. Representation of solution: Binary String
2. Population: 60 individuals
3. Number of generations: 100 generations
4. Moving crossover: Simple crossover (0.6)
5. Mutation: Binary Mutation (0.0333)

This GA model is based on a model that had been done by Kusumoputro [12] and Kustiyo [2]. Comparison of weather prediction results using a combination of the LVQ-GA method with actual data can be seen in Figure 5.

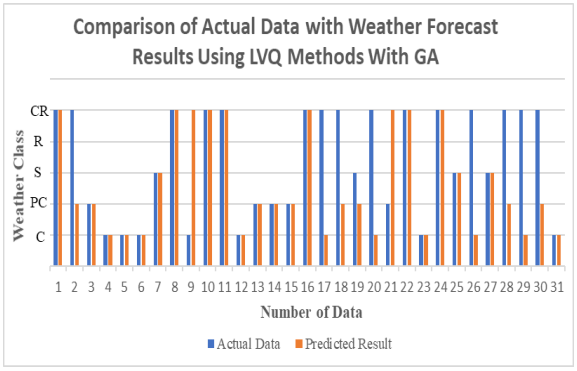

Fig. 5. Comparison of LVQ-GA prediction results with actual data

Figure 5 shows the comparison of the actual data with the predicted LVQ-GA data. From the sample test data taken as many as 31 data, it can be seen that there are some results that do not match the actual data, namely the test data 9, 17, 18, 19, 20, 21, 28, 29, 30 and 31.

The results showed that the BMKG's daily weather data for the Citeko Cisarua Bogor area were incompatible when modeled using the LVQ method or LVQ optimization with GA. Based on the results of calculating the accuracy using a confusion matrix, the resulting accuracy is 72% for weather forecasting using the LVQ method and 73% when the LVQ has been optimized using GA.

This research has succeeded in making a weather forecast model using the LVQ and LVQ-GA methods. The weather parameters used are temperature, evaporation, sunlight, humidity and rainvol. While the weather forecast class consists of 5 classes, namely Cloudly (C), Partly Cloudly (PC), Sunny (S), Rain (R) and Cloudly rain (CR). The weather data obtained previously was modeled using the LVQ method and also the LVQ optimization modeling was carried out using GA. After testing with 31 samples of test data, the accuracy value obtained is 72% for the LVQ method and 73% for the optimization of the LVQ-GA.

CONCLUSION

This research has succeeded in making a weather forecast model using the LVQ and LVQ-GA methods. The weather parameters used are temperature, evaporation, sunlight, humidity and rainvol. While the weather forecast class consists of 5 classes, namely Cloudly (C), Partly Cloudly (PC), Sunny (S), Rain (R) and Cloudly rain (CR). The weather data obtained previously was modeled using the LVQ method and also the LVQ optimization modeling was carried out using GA. After testing with 31 samples of test data, the accuracy value obtained is 72% for the LVQ method and 73% for the optimization of the LVQ-GA.

ACKNOWLEDGMENT

A big thank you to the National Research and Innovation Agency (BRIN) of the Indonesian Ministry of Research and Technology, for providing grants to the author to conduct this research and LPPM Pakuan University for helping a lot in this research process.

REFERENCES

[1] Gibbs. 1987. Pengertian Cuaca.

[2] A. Kustiyo, A. Buono, N. Apriyanti. 2006. Optimasi Jaringan Syaraf Tiruan dengan Algoritma Genetika untuk Peramalan Curah Hujan. Jurnaillmiah IImu Komputer. Edisi 6 Nol.4 No.1; 1-9.

[3] K. Widyawati, B.D. Setiawan, P.P Adikara. T. 2018. Optimasi Vektor Bobot Learning Vector Quantization Menggunakan Algoritme Genetika untuk Penentuan Kualitas Susu Sapi. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 2, No. 1, Januari 2018, hlm. 217-225.

[4] O. Sen, S. Zhengxiang, W. Jianhua, C. & Degui. 2002. Application of LVQ Neural Networks Combined with Genetic Algorithm in Power Quality Signals Classification, Shaanxi: School of Electrical Engineering.

[5] S. Kusumadewi. 2004. Membangun Jaringan Syaraf Tiruan menggunakan Matlab dan Excell Link. Penerbit Graha Ilmu.

[6] D. M. Wuryandari, I. Afrianto. 2012. Perbandingan Metode Jaringan Syaraf Tiruan Backpropagation Dan Learning Vector Quantization Pada Pengenalan Wajah.

[7] R. Sulistiyorini, W. F. Mahmudy. 2015. Penerapan Algoritma Genetika Untuk Permasalahan Optimasi Distribusi Barang Dua Tahap, Malang: Doro Jurnal.

[8] R. Arniantya, B.D. Setiawan, P.P Adikara. T. 2018. Optimasi Vektor Bobot Pada Learning Vector Quantization Menggunakan Algoritme Genetika Untuk Identifikasi Jenis Attention Deficit Hyperactivity Disorder Pada Anak. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 2, No.2, Februari 2018, hlm. 679-687.

[9] A. A. Dwi, C. Dewi, Indriati. 2019. Implementasi Algoritma Genetika untuk Optimasi LVQ pada Penentuan Kelayakan Kredit (Studi Kasus: Bank X). Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 3, No. 1, Januari 2019, hlm. 1114-1119.

[10] M. Hendry, R. M. Dayawati, A. T. Wibowo. 2009. Analisis dan Implementasi Optimasi Jaringan Saraf Tiruan dengan Menggunakan Algoritma Genetika untuk Pendiagnosaan Penyakit Stroke (Studi Kasus : RS. Dr. M. Djamil Padang Sumbar).

[11] J. Han, M. Kamber, J. Pei. 2012. Data Mining: Concept and Techniques. Ed ke-3. Amsterdam (NL): Elsevier, Morgan Kaufmann.

[12] B. Kusumoputro. 2004. Pengembangan Sistem Pengenal Wajah secara 3 Dimensi menggunakan Hemisphere

Structure of Neural Networks dan Optimasi Struktur menggunakan Algoritma Genetika. Makalah SNIKTI V. Fakultas lImu Komputer, VI, Depok.