19. Jurnal Mantik-The Activity Activation Function-2020

by Diperiksa Oleh Universitas Potensi Utama

Jurnal Mantik
Volume 4, Number 1, May 2020, pp. 555-561 E-ISSN 2685-4236
https://iocscience.org/ejournal/index.php/mantik

The Activity Activation Function Of Multilayer Perceptron - Based Cardiac Abnormalities

Mutiara S. Simanjuntak[1], Wanayumini[2], Rika Rosnelly[3], Teddy Surya Gunawan[4]

[1234] Master of Computer Science, Faculty of Engineering and Computer Science, Universitas Potensi Utama, Jl. K.L. Yos Sudarso KM.6,5 No. 3A, Tj. Mulia, Medan

Email: [1]sarahwaty.mutiara@gmail.com, [2]wanayumini@gmail.com, [3]rikarosnelly@gmail.com

ARTICLE INFO

Article history:

Received: 04/04/2020
Revised: 20/04/2020
Accepted: 30/05/2020

Keywords: Activation, Multilayer Perceptron, Electrocardiogram, Backpropagation, Bayesian Regularity

ABSTRACT

Cardiac disorders refer to irregular activity at the heart. Cardiac abnormalities sometimes do not exhibit any and unreasonable symptoms that can lead to sudden death due to heart-cracking functions. This article is to develop a program capable to detect cardiac abnormalities activity through the application of Multilayer Perceptron (MLP). A certain number of heart rate signal data from an electrocardiogram (EKG) will be used in this paper to train and to test the network performance of the MLP. MLP is trained by several techniques that Backpropagation (BP), Bayesian regularity (BR), and Levenberg-Marquardt (LM).

1. Introduction

The heart is an important organ in the human body and it is a myogenic contract without the signal of the nervous system. Pacemakers will generate signal waves for contracts and delays in AV nodes[1]. Then, the signal is passed to the peak of the heart and is scattered throughout the ventricle. This is a heart process for pumping oxygenated blood throughout the body. However, these organs still have failures or abnormal activity for some reason. Heart abnormalities will occur as blood pumps fail for the rest of the body in the heart beating normally[2]. This will happen when a blood clot is blocked by an artery connected to the heart. The presence of blood clots will make the arteries narrow and disrupt the transportation system in the human body. In a short period, the heart will be weak, and the blood containing oxygen pumped throughout the body will be lacking, causing the cells in the body to lack oxygen [3]. This will cause a heart attack.

Heart disease is a cardiovascular disease that is a major health problem throughout the country. Heart disease is currently one of the highest causes of death. Death from heart disease throughout the United States reaches 959,227 sufferers, i.e. 41.4% of all deaths or every day 2,600 inhabitants died from heart disease. In Indonesia, the death caused by heart disease reaches 9.7% with 5.1% is ischemic Heart sufferer [4]. Heart disease is a narrowing or blockage (arteriosclerosis) of a coroner artery caused by a buildup of fatty substances (cholesterol, triglycerides) that are getting longer and accumulate under the deepest layer (endothelium) of the artery wall. As a result of the blockage itself, the supply of chemical energy to the heart muscle decreases, so there is a disturbance of the balance between supply and blood requirements [5].

Coronary heart disease is a condition of a person who has an obstruction in the heart vein due to plaque buildup. One factor in the occurrence of coronary heart disease is the pattern of poor food consumption that increases the level of fat in the body[6]. Coronary heart disease (CHD) is a form of cardiovascular disease that has become the number one cause of death in the world [7].

Conventional Analytical techniques that have been used as long as it is not as effectively used to diagnose because this technique causes bias, errors, and excessive medical costs affecting the quality of services provided to patients. Along with the development of medical knowledge-based systems, the demand for the use of computer-based knowledge systems as an analytical technique in diagnosing disease becomes increasingly important. Therefore, it is time to develop a computer-based knowledge system that is modern, effective, and efficient in diagnosing diseases [8].

Heart defects can be detected by using ECG. EKG is a test to demonstrate the electrical system in the work of the heart. During ECG, the power leads are placed on the chest, arms, and legs. The contact is

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

detected as a small electrical signal and tracing on the graph paper to depict electrical signals throughout the heart muscle. Heart defects can be detected as electrical signals are described differently from normal. Data from the ECG will extract to detect cardiac abnormalities using the Multilayer Perceptron (MLP) neural network technique. P, QRS and T amplitude peaks will show heart abnormalities. An MLP consists of multiple layers with each layer fully connected to each. The project will train the MLP neural network using a learning algorithm to detect abnormal activity from the heart by using P, Q, R, S, and T amplitude and ration as input parameters. The learning algorithm used in the MLP network to train the neural networks, backpropagation (BP), Bayesian regularity (BR), and Levenberg-Marquardt (LM [9]. Multilayer Perceptron is the most common topology of JST, in which the connected Perceptronperceptron forms multiple layers. An MLP has an input layer, at least one hidden layer, and an outer layer (output layer). [10]

2. Overview

This section will talk about more details about how human heart functions and how impulses are generated. It will then elaborate on the EKG signal that will be able to see on the Holter Monitor. EKG signals will extract for use in MLP networks. The MLP network will be used BP, BR, and LM training algorithms that allow the algorithm to recognize the activity of heart abnormalities, human hearts, located behind the breastbone (breastbone) with the size of the fist, and consist of the heart muscle. It is a myogenic organ, which means it is not mastered by the brain. Relaxes to pump oxygen and deoxygenated blood to all parts of the body; Controlled by a pacemaker that produces electrical impulses. Heart contracting and relaxation in rhythmic cycles [11].

Deoxygenated Blood will enter the heart through the vein cava, then the blood will pump into the right atrium before entering the right ventricle. From the right ventricle, the blood will pump to the lungs through the pulmonary artery. Oxygenated blood from the lungs will enter the heart through the pulmonary vein to the left atrium. Then, the blood enters the left ventricle before the heart leaves through the aortic to be sent to all parts of the body. To ensure there is no blood flow, all four chambers have been called second valves and tricuspid. Due to blood pressure on the valve, it produces a ' club, dups ' sound rhythm[12]. Diastole atrial and ventricular take about 0.4 s, atrial take systole about 0.1 s, and ventricular systole takes about 0.3 s; Mean Systole and Diastole complete take cycle about 0.8 s. The average heart rate essentially takes about 0.8 s/beat or 75 beats/minute [13].

The race or SA node generated the signal to create an atrium contract simultaneously. The signal will be delayed in the AV node to ensure the completion of the atrium is completed and their space completely emptied. Then, the signal is done for the cardiac apex with Purkinje fibers and the bundle branch. When the spread of signals across the ventricle, it would be a contract from top to top. Signal results will be recorded as an EKG signal. ECG can be defined as a non-invasive device used to measure and monitor the electrical activity of the heart through the skin. A wave is produced by an electrocardiogram. This particular wave is in response to the electrical changes that occur in the liver. The electrical changes that occur in the liver that are used to detect problems begin with the environmental changes of the ionic liver. Followed by the decompression system of cardiac conduction with the breakdown of muscle fibers due to lack of blood supply then some heart disease detected. The waves produced by the monitoring electrocardiogram indicate the electrical changes occurring inside the liver.

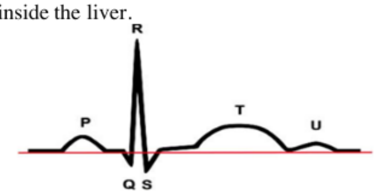

Fig 1. ECG Signal Normal Wave

By looking at Figure. 1, there are essentially a few waves that can be described as Wave P, the QRS complex (which often includes Q wave, R wave, and P wave), T wave and U wave. Each of these waves can be explained as follows;

TABLE 1
SIGNAL WAVES AND THEIR REPRESENTATION

Wave	Heart conditions

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

Wave	Heart conditions
Wave P	Atrial Depolarization3
QRS Complex	Ventricular depolarization
T Wave	Ventricular repolarization
Wave U	Repolarized Fiber Purkinje

An MLP is a logistical regression classifier where the first input is transformed using a non-linear learning transformation[14]. This transformation projects the input data into the space in which it becomes linear separated. This intermediate layer is referred to as a hidden layer. A single hidden layer is enough to make MLPs a universal approximator. Multilayer Perceptron coating illustrated in the figure. 2

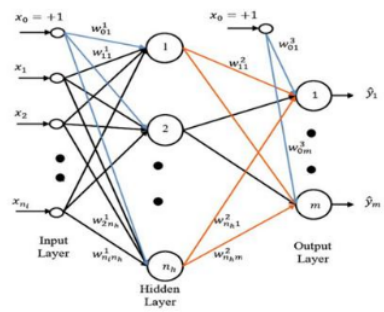

Fig 2. Multilayer Perceptron Coating

The most common algorithm for the Multilayer Perceptron network is The Backpropagation (BP) algorithm. The following discussion regarding the Backpropagation algorithm is based on the book by Haykin (1994)[15]. Backpropagation is a gradient descent procedure that attempts to reduce errors between the network output and the desired outcome. Bayesian Regularization (BR) is a powerful add-on for neural network training that allows automatic optimization of weight-loss parameters. Thus, both Levenberg-Marquardt (LM) and Bayesian Regularization (BR) are more efficient. A combination of LM and BR applies very little additional compute charges, as it exploits the approach to Hessian. Therefore, heavy decay can be updated easily after each training cycle [16].

3. Method

Neural networks are computing systems, inspired by the concept of biological nerve cells known as neurons. Neurons are small cells that our brains consist of. A biological neuron is defined as a collection of several billions of neurons[17]. Based on the concept, the proposed artificial neural network, which can model the structure of biological neurons both in terms of their architecture and operation. They serve as mathematical computing models for non-linear approaches, data classification, and clustering/parametric regression functions. Serves as a simulation for the behavior of the human model's biological neuron collection. Artificial neural networks are capable of providing performance reliability especially in decision making in placing the human brain[18].

Therefore, that can act as an alternative to intelligent systems. One of the most commonly used artificial neural networks for that purpose is the Multilayer Perceptron network. In 1958, Rosenblatt introduced the model of the Perceptron. Cascading of some models of perceptron in the layer generates a network as shown in the figure. 2. This neural network is referred to as the Multilayer Perceptron network. Based on the Fig. 2, the Multilayer Perceptron network with m output and hidden nodes and input nodes can be declared as.

$$\hat{Y}(t)=\sum_{j=1}^{nh} w_{jk}^{2}\, f\left(\sum_{i=1}^{1} w_{ij}^{1}\, x_1^0(t) + b_j^1\right) \text{for } 1 \leq k \leq m \text{ ....(1)}$$

Based on Equation 1, W_{ij}^1 Is that each shows a heavy connection between the input layer to the hidden layer while W_{jk}^2 Shows as the connection weights between the hidden layers to the output layer, respectively. b_j^1 Equation 1 represents the hidden threshold of nodes and x_1^0 Shows the input parameter that is fed to the input layer. In this study, the sigmoid activation function has been selected and signifies it as in Equation 1.

By referring to equations (1), the values W_{ij}^{1}, W_{ij}^{1}, and b_j^1 Determined based on the appropriate algorithm. In this study, three different training algorithms were used which were Backpropagation (BP), Lavenberg Marquardt (LM), and BAYESIAN (BR) algorithms.

Backpropagation, the most commonly used training algorithm, is the gradient of hereditary procedures that efficiently calculates derivative values, and modify the weights according to the parameters known as the level of learning [16]. Backpropagation is a decent steep type algorithm where the heavy relationships between j^{th} neuron are hidden layers and i^{th} with neurons, the input layer is each updated according to:

$$w_{ji}\ (t) = w_{ji}\,(t-1) + \Delta w_{ji}\,(t)$$
$$b_j\ (t) = b_j\,(t) + \Delta b_j\,(t)$$

$\Delta w_{ji}\,(t)$ and $\Delta b_j\,(t)$ given by:

$$\Delta w_{ji}\ (t) = \mu_w \rho_j\,(t) x_i(t) + \alpha_w \Delta w_{ji}\,(t)$$
$$\Delta b_j\ (t) = \mu_w \rho_j\,(t) x_i(t) + \alpha_w \Delta b_j\,(t)$$

That specifies the W and B effect change parameters αw and αb In the direction of the movement in the parameter space, μw, and μb and represents the level of learning and $\rho_j\,(t)$ is a signal error from a j^{th} neuron hidden layer that is re-deployed in the network. Due to the neuron activation function of linear output, the error signal on the output node:

$$\rho\ (t) = y_k\ (t) - \hat{Y}_k\,(t) \ldots\ldots(2)$$

Where $y_k(t)$ is the expected output. For neurons on hidden layers:

$$\rho_j\,(t) = F''\big(X_i\,(t)\big) \textstyle\sum_j \rho_j^k\ (t) w_{jk}^2\,(t-1) \ldots\ldots(3)$$

Where F'' (x_i(t)) is a derivative of F'' (x_i(t)) with x_i(t). Since the back of the Backpropagation algorithm is a kind of steep-worthy algorithm, the algorithm suffers from a slow convergence rate. The search for global minima may be stuck in the local minima and the algorithm can be sensitive to use parameters to be selected.

The Levenberg-Marquardt algorithm is a localized, highly deterministic optimization algorithm based on gradients. Once used to train the MLP model, the advantage of Levenberg-Marquardt over the traditional Back Propagation algorithm is that it can deliver (second-order) levels of convergence faster and remain relatively stable. Like the Kuasi-Newton method, the Levenberg-Marquardt algorithm is designed to approach the second-order pace of training without having to calculate the Hessian matrix[16]. When performance functions have the form of some squares (like typing in the Feed Training Network forward), then the Hessian matrix can be estimated as:

$$H=J^TJ$$

And gradients can be counted as:

$$g=J^T\rho$$

Where J is a Jacobian matrix that contains the first derivative of a network fault concerning weights and biases, and e is a vector of network errors. The Jacobian matrix can be calculated through a standard re-propagation technique that is much more complex than the Hessian matrix computing. The Levenberg-Marquardt algorithm uses this approach for the Hessian matrix as following Newton-like updates:

$$\Delta w = -[J^TJ + \mu I]^{-1} J^T \rho \ldots\ldots(4)$$

Where Δw is differential weights and μ is a control parameter. When scalar μ Zero, this is just Newton's method of using the approximate Hessian matrix. When large, this becomes a gradient descent with a small step size. The Newton method is faster and more accurate near the minimum error, so the goal is shifted towards Newton's method as quickly as possible. So, μ Decreases after each successful step (decreased performance function) and increases only when the tentative step improves performance functions. In this way, performance functions will always be reduced at every iteration of the algorithm. Remembering the Baye rules where:

$$P(\theta|D) = \left(\frac{Þ(\theta|D)}{P(D)}\right) \ldots\ldots(5)$$

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

In general, this will give the entire θ. This process is applied distribution through the network value and comes with a probability distribution over the network weight, w, considering the training data p (w | D). When locating the posterior distribution:

$$P(\theta|D)= (\frac{Þ(D|W)p(w)}{P(D)})= (\frac{Þ(D|W)p(w)}{\int P(D|w)p(w)dw})$$

In Bayesian formalism, learning weights means changing our beliefs about weights from before, p (w), to Posteriorly, p (w | D) As a consequence of viewing the data as illustrated by Figure. 3

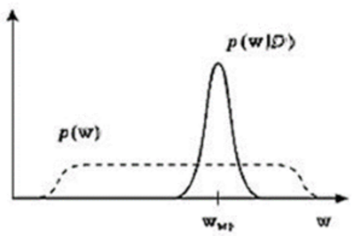

Fig 3. Changing weights before Posterior weight

Meaning a square fault (MSE) is an error that occurs in each of the data obtained each time the data iteration is added. The smaller the MSE value obtained, the more accurate the prediction is obtained and the smallest MSE value obtained when the MSE graph obtained is horizontal. At this time, neural network performance is constant and stable. The highest predicted accuracy performance with the lowest MSE value will determine the optimal amount of data iteration. However, the performance focus of neural networks is devoted to performance in the trial phase. Then, determining the number of hidden nodes is performed using the same method. At this time, the number of data set training iterations with the optimum value gained previously.

The accuracy of the detection data will be calculated based on the amount of correctly classified and divided by the amount of data in the class. For the determination of MSES, the data output deviation detection versus the actual output Data will be supported-squared and the result will be total to find the average deviation of each data. MSE will be calculated with the following equation:

$$MSE = \left(\sum_{n=1}\left(y(t) - \hat{Y}(t)^2\right)\right)/n \quad \text{..(6)}$$

With *y(t)* is a temporary data forecasting output $\hat{Y}$ *(t)* actual and nth output data is the amount of data used.

4. Result

To determine the suitability of a Multilayer Neural Network as a diagnostic system for cardiac abnormalities activity problems, fundamental analysis is required which is diagnostic performance analysis. Performance diagnostic analysis shows the accuracy measurements of Backpropagation *(BP)*, Levenberg-Marquardt *(LM)*, and *Bayesian regularization* (BR) training and testing. Table 2 shows the performance of *the* Multilayer Perceptron network Trained By Backptopagation, Levenberg-Marquardt, and Bayesian regularization. From the results, Bayesian Regularization trains algorithms Outperform Levenberg-Marquardt and backpropagation training algorithms in given the best accuracy predictions. The Regularization Bayesian training algorithm delivers 98.523% during the training phase while 98.305% at the test stage.

MSE performance shows 4.667, 4.067, and 0.237 BP, LM, and BR, respectively. From MSE results, BR can provide the best prediction to compare with LM and BP training algorithms. Results show the accuracy of prediction by using the BR training algorithm either normal patients or not. BP's training algorithms are capable of fast processing but cannot sustain good results since the pitfalls in local minima. LM delivers better results than BP's extended BP. Better results during performance predictions and MSE analysis is given. The BR training algorithm is based on performance analysis statistics show better on both predictions and MSE analysis.

TABLE 2
RESULTS IN THE ACCURACY OF TRAINING ALGORITHMS

Achievements	Algorithm Training		
	MR	LM	BR
Accuracy training, (%)	92.104	93.270	98.532
Testing accuracy, (%)	94.932	94.145	98.305
Mse	4.667	4.067	0237

5. Conclusion

This article analyzes the capabilities of the MLP network to classify the detection of cardiac abnormalities activities by using features from EKG signals. The MLP network demonstrates the reliability and ability to classify the heart abnormalities activity with high accuracy performance and low MSE reading. BR training algorithm is capable to be used as an MLP network.

6. Reference

[1] C. E. Boom and O. Widyapuspita, “Tatalaksana Perioperatif pada Pasien dengan Cardiac Implantable Electronic Devices (Cieds) atau Alat Elektronik Kardiovaskular Implan (Aleka),” *JAI (Jurnal Anestesiol. Indones.*, vol. 11, no. 2, p. 81, 2019, doi: 10.14710/jai.v11i2.24714.

[2] M. S. Simanjuntak, R. Wijaya, J. T. Informatika, and P. I. Komputer, “Melalui Gejala Dan Pola Hidup Menggunakan Metode,” vol. 3, no. 3, pp. 122–125, 2019.

[3] B. Rifai, “Algoritma Neural Network Untuk Prediksi,” *Techno Nusa Mandiri*, vol. IX, no. 1, pp. 1–9, 2013.

[4] Kemenkes RI, *Profil Kesehatan Republik Indonesia Tahun 2009*. 2009.

[5] M. S. Simanjuntak, R. Wijaya, and P. I. Komputer, “APLIKASI PENCARIAN GAMBAR DENGANALGORITMA CONTENT-,” vol. 3, no. 3, 2019.

[6] H. H. Rahma and R. B. Wirjatmadi, “Hubungan Asupan Zat Gizi Makro Dan Profil Lipid Dengan Kejadian Penyakit Jantung Koroner Pada Pasien Lansia Di Rumah Sakit Islam Jemursari Surabaya,” *Media Gizi Indones.*, vol. 12, no. 2, p. 129, 2018, doi: 10.20473/mgi.v12i2.129-133.

[7] I. Iskandar, A. Hadi, and A. Alfridsyah, “Faktor Risiko Terjadinya Penyakit Jantung Koroner pada Pasien Rumah kit Umum Meuraxa Banda Aceh,” *AcTion Aceh Nutr. J.*, vol. 2, no. 1, p. 32, 2017, doi: 10.30867/action.v2i1.34.

[8] F. R. Hashim, J. J. Soraghan, L. Petropoulakis, and N. G. N. Daud, “EMG cancellation from ECG signals using modified NLMS adaptive filters,” *IECBES 2014, Conf. Proc. - 2014 IEEE Conf. Biomed. Eng. Sci. “Miri, Where Eng. Med. Biol. Humanit. Meet,”* no. December, pp. 735–739, 2014, doi: 10.1109/IECBES.2014.7047605.

[9] F. R. Hashim, L. Petropoulakis, J. Soraghan, and S. I. Safie, “Wavelet based motion artifact removal for ECG signals,” *2012 IEEE-EMBS Conf. Biomed. Eng. Sci. IECBES 2012*, no. December, pp. 339–342, 2012, doi: 10.1109/IECBES.2012.6498019.

[10] Wanayumini, O. S. Sitompul, M. Zarlis, and S. Suwilo, “Automatic Detection of Chaos Phenomenon in Tornadoes Prediction Using Edge Detection,” *2019 3rd Int. Conf. Electr. Telecommun. Comput. Eng. ELTICOM 2019 - Proc.*, pp. 48–52, 2019, doi: 10.1109/ELTICOM47379.2019.8943845.

[11] S. Palaniappan and R. Awang, “Intelligent heart disease prediction system using data mining techniques,” *AICCSA - 6th IEEE/ACS Int. Conf. Comput. Syst. Appl.*, pp. 108–115, 2008, doi: 10.1109/AICCSA.2008.4493524.

[12] R. Rosnelly, L. Wahyuni, and J. Kusanti, “Optimization of Region of Interest (ROI) Image of Malaria Parasites,” *J. Appl. Intell. Syst.*, vol. 3, no. 2, pp. 87–95, 2018, doi: 10.33633/jais.v3i2.2060.

[13] F. R. Hashim, N. G. N. Daud, S. N. Mokhtar, A. F. Rashidi, J. Adnan, and K. A. Ahmad, “Optimization of ECG Peaks (Amplitude and Duration) in Predicting ECG Abnormality using Artificial Neural Network,” *Indian J. Sci. chnol.*, vol. 10, no. 12, pp. 1–5, 2017, doi: 10.17485/ijst/2017/v10i12/112970.

[14] M. Hamdan, O. O. Khalifah, and T. S. Gunawan, “Measuring the road traffic intensity using neural network with computer vision,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 10, no. 1, pp. 184–190, 2018, doi: .11591/ijeecs.v10.i1.pp184-190.

[15] W. Wanayumini, O. S Sitompul, M. Zarlis, S. Suwilo, and A. M H Pardede, “A Research Framework for Supervised Image Classification For Tornado Chaos Phenomena,” *Int. J. Eng. Technol.*, vol. 7, no. 4.15, p. 447, 18, doi: 10.14419/ijet.v7i4.15.25254.

[16] F. R. Bin Hashim, J. J. Soraghan, and L. Petropoulakis, “Multi-classify Hybrid Multilayered Perceptron (HMLP) network for pattern recognition applications,” *IFIP Adv. Inf. Commun. Technol.*, vol. 381 AICT, no. PART 1, pp. 19–27, 2012, doi: 10.1007/978-3-642-33409-2_3.

[17] J. Adnan *et al.*, “Heart abnormality activity detection using multilayer perceptron (MLP) network,” *AIP Conf. Proc.*, vol. 2016, no. September, 2018, doi: 10.1063/1.5055415.

Accredited “Rank 4”(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

[18] L. A. M. Pereira *et al*., "Multilayer perceptron neural networks training through charged system search and its Application for non-technical losses detection," *2013 IEEE PES Conf. Innov. Smart Grid Technol. ISGT LA 2013*, 2013, doi: 10.1109/ISGT-LA.2013.6554383.

19. Jurnal Mantik-The Activity Activation Function-2020

ORIGINALITY REPORT

13%	15%	10%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Ibn Khaldun Student Paper	6%
2	csrid.potensi-utama.ac.id Internet Source	4%
3	garuda.ristekbrin.go.id Internet Source	3%

Exclude quotes Off

Exclude bibliography Off

Exclude matches < 3%