Jurnal Infotel
Vol. 16, No. 1, February 2024, pp. 96–104. doi:10.20895/infotel.v16i1.974

Research Article

Combination of Binary Particle Swarm Optimization (BPSO) and Multilayer Perceptron (MLP) for Survival Prediction of Heart Failure Patients

Sutikno

Department of Computer Science/Informatics, Diponegoro University, Semarang 50275, Indonesia

*Corresponding email: sutikno@lecturer.undip.ac.id

Received: May 22, 2023; Revised: October 28, 2023; Accepted: November 17, 2023.

Abstract: A serious illness known as heart failure prevents the heart from adequately pumping blood, which can be fatal. Globally, heart failure is a leading cause of death, which causes both society and the government to have major worries. Even this is the most significant cost burden. To improve this treatment it needs methods to predict patient survival. The addition of filter feature selection to machine learning algorithms has increased accuracy. However, this feature selection has a drawback: it produces relatively lower accuracy than wrapper feature selection. This paper proposed combining wrapper feature selection, namely binary particle swarm optimization (BPSO) and a multilayer perceptron (MLP) classifier called BPSO-MLP. BPSO is used to determine the most relevant feature subset, and MLP is used to calculate its fitness. In addition, we also compared 6 MLP training algorithms. The study used a public dataset. The experiment's findings demonstrated that the suggested strategy might generate results with an accuracy of up to 91.11 %. The proposed method can increase accuracy by 8.89 % compared to MLP (without BPSO). The addition of this wrapper feature selection has a significant influence on the accuracy results. The proposed method and selected features can be used to develop an expert system for predicting the survival of kidney failure patients.

Keywords: BPSO, MLP, heart failure prediction

---

1 Introduction

When the heart cannot efficiently pump blood, it suffers from heart failure, a dangerous condition that can be deadly [1]. Globally, heart failure is a leading cause of death, which causes both society and the government to have major worries. Even this is the most significant cost burden. Based on data from the Social Security Administration for Health, Indonesia, in 2021, the most numerous health financing for this disease is IDR 7.7 trillion.

Errors in predicting heart failure will have fatal consequences for the patient's survival. Several researchers have used this patient survival prediction method to help diagnose heart failure and therapy. These include machine learning [2–10], deep and wide neural networks (DWNN) [11], machine learning and deep learning combined [12], deep pyramid convolutional neural networks, and extreme gradient boosting [13]. Furthermore, several researchers changed one of the machine learning methods, including the two support vector machines (SVM) [14], the boosting SVM [15], the hybrid of a random forest and decision tree [16], and the hybrid random forest with a linear model [17].

To improve performance, Dissanayake and Johan [18] compared several filter feature selections. According to the trial findings, backward feature selection produces the highest accuracy, 88.52 %. However, it has the disadvantage of making relatively lower accuracy than wrapper feature selection [19, 20]. For this reason, this study proposed combining wrapper feature selection, namely binary particle swarm optimization (BPSO), with a multilayer perceptron (MLP). BPSO is used to determine the most relevant feature subset, and MLP is used to calculate its fitness. The gap in this research is that previous research used filter feature selection, while this study uses wrapper feature selection. The novelty of this research is combining BPSO and MLP. This combination aims to increase the accuracy of survival prediction of heart failure patients.

2 Research Method

This section discusses binary particle swarm optimization, the proposed binary particle swarm, and the evaluation method.

2.1 Binary Particle Swarm Optimization (BPSO)

BPSO is an algorithm used to solve binary optimization problems. It was developed from particle swarm optimization (PSO) [21]. In BPSO, each particle in the swarm represents potential solutions. BPSO has been able to optimize various problems in engineering and computer science.

BPSO involves steps to reach an optimal solution in the binary search space. A population of particles representing potential solutions was first randomly initialized. The objective function is then used to assess each person's fitness. The population's particles then update their location and velocity. Particle velocity is affected by previous velocity, individual experience (pbest), and global experience (gbest). The velocity updates the particle's position in the binary search space. After updating the position, it re-evaluated the fitness of the new particle, and if the new position has a better value, the pbest is updated.

Assume that $S = 0, 1^D$ is the search space. The objective function f is to maximize max $f(x)$. The ith swarm particle could be illustrated as $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})^T$, $x_{id} \in 0, 1$, $d = 1, 2, \ldots, D$. This particle's velocity may be expressed by $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})T$, $v_{id} \in [-V_{\max}, V_{\max}]$, $V_{\max}$ is highest velocity. The optimum position of the ith particle visited before could be expressed as $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})^T$, $p_{id} \in 0, 1$. If g is the best index in the swarm and p_{gd} is the best swarm, the velocity is updated by (1).

$$v_{id} = v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \tag{1}$$

The position is updated using (2).

$$x_{id} = \begin{cases} 1 & \text{if } U(0,1) < \text{sigm}(v_{id}), \\ 2 & \text{otherwise} \end{cases} \tag{2}$$

where $i = 1, 2, \ldots, N$. N is the swarm's size; c_1 and c_2 are constants for the social and cognitive scaling factors; r_1 and r_2 are random values with a uniform distribution in $[0, 1]$. A random number symbol with a distribution of 0 to 1 is called $U(a, b)$. A sigmoid limiting transformation is defined by (3) as sigm(v_{id}).

$$\text{sigm}(v_{id}) = \frac{1}{1 + \exp -\lambda . v_{id}} \tag{3}$$

where λ is the steepness controlling value of the sigmoid function.

2.2 The proposed Binary Particle Swarm Optimization - Multilayer Perceptron (BPSO-MLP)

The proposed method for predicting heart failure survival is BPSO-MLP, a combination of BPSO and MLP, as shown in Figure 1. The best feature subset is chosen using BPSO, and the value fitness is determined using MLP.

1. Initialize population particles randomly
   The population's total count of particles (N) and the total count of dimensions (D) each particle possesses are counted in the first stage. In this study, N equals 30, and D equals the number of features, which is 12, making a total of 12. Each dimension is either 0 or 1, which is randomly chosen. The linked feature is picked if the value is more than 0 and is not selected if the value is less than 1.
2. Initialize personal best and global best
   A created particle is a personal best, but the fitness particle is a world record. Equation (4) calculates the fitness value [22].

   $$\text{Fitness} = \alpha \gamma_R(D) + \beta \frac{|R|}{|C|} \tag{4}$$

   where $|R|$ is the number of selected features, $|C|$ is the sum of all the features, γ_R is the classification error, α is the weight of error classification, and β is the weight of selected features. The α value is 0.99, and the β value is 0.01 [22].
   As illustrated in Figure 2, the MLP classifier has several input layers based on the feature subset, one hidden layer, and one output layer. The amount of neurons in the hidden layer is calculated using (5) [23].

   $$n_H = \sqrt{n_i + n_0} + l \tag{5}$$

   where l is between 1 and 10, ni and no are the neuron's numbers in the input and output layers. The choices for the training parameters are learning rate = 0.05, maximum

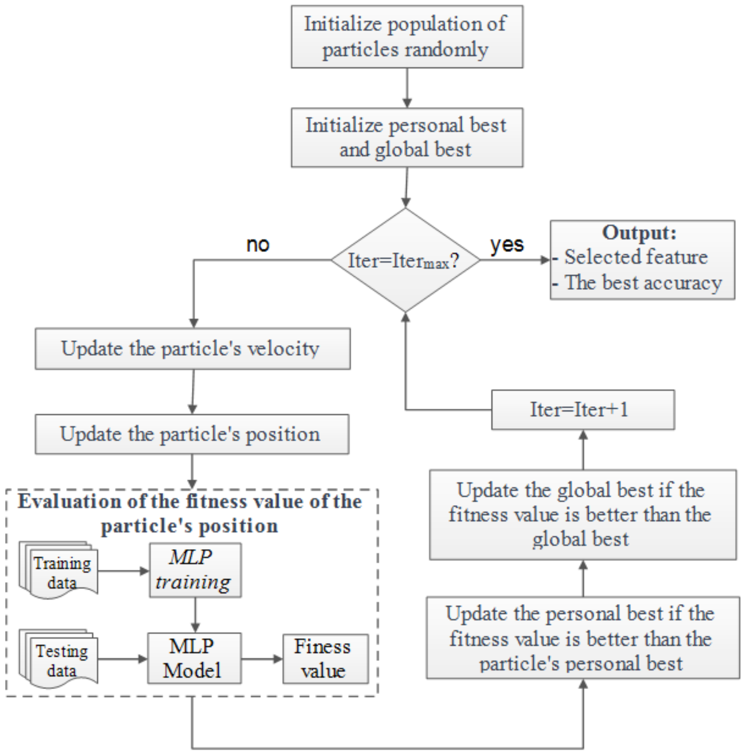

Figure 1: The proposed BPSO-MLP for predicting heart failure survival.

epochs = 1000, and error limit = 0.001 [24]. Six training algorithms were compared in the experiment: traingd (gradient descent backpropagation), traingdx (gradient descent with momentum and adaptive learning rate backpropagation), traingda (gradient descent with adaptive learning rate backpropagation), and traincgp (conjugate gradient backpropagation with Polak-Ribiere update).

3. Update the particle's velocity
   Equation (1) updates each dimension according to the particle velocity.
4. Update the particle's position
   Particle positions update each dimension using (2).
5. Calculate the fitness of the particle's position
   The fitness at every particle position is evaluated using (4).
6. Update the personal best
   If the current fitness is less than the personal best fitness (pbest), the particle fitness (pbest) and position change.
7. Update the global best
   The comparison of all particles' fitness levels is made during the global best update procedure. The location and the fitness value are updated if a better particle is found.
8. Check stop condition
   The algorithm will stop if the iteration equals the maximum iteration. In this study, the maximum iteration (Iter_{max}) is 333.

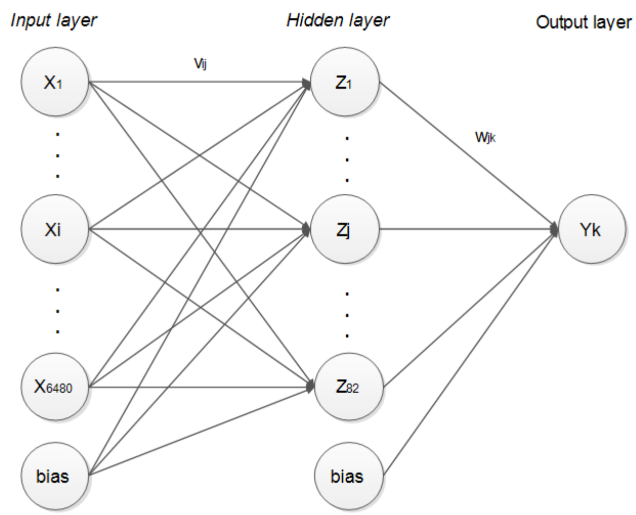

Figure 2: The MLP architecture.

2.3 Evaluation Method

The accuracy determined using (6) is the parameter used to assess the performance of the suggested approach.

$$\text{Accuration} = \frac{T_{pos} + T_{neg}}{T_{pos} + T_{neg} + F_{pos} + F_{neg}} \tag{6}$$

$T_{pos}, T_{neg}, F_{pos}$, and F_{neg} denote True Positive, True Negatif, False Positif, and False Negative [25].

3 Result

A public dataset from 299 heart failure patients who were treated at Allied Hospital in Faisalabad, Pakistan, in 2015 was utilized in the experiment [2]. It is available in [26]. The 13 variables of this dataset consist of age (f-1), anemia (f-2), high blood pressure (f-3), diabetes (f-5), ejection fraction (f-6), platelets (f-7), sex (f-8), serum creatinine (f-9), serum sodium (f-10), smoking (f-11), time (f-12), and death event (f-13). The others were utilized as input, with the death event feature acting as a goal. 70 % and 30 %, respectively, of the data were used in the training and testing task. This mixture was also utilized in an earlier investigation [14].

The experiments were carried out without and with feature selection to look for this influence. In addition, we also compared 6 MLP training algorithms.

3.1 Experiment without Feature Selection

Table 1 displays the outcomes of the experiment using the six MLP training techniques without the use of feature selection. The most excellent accuracy was attained (82.22 %)

using traingd. With an accuracy score of 81.11 %, traingdx has a slightly lower accuracy rating than traingd.

Table 1: Experiment results using MLP with a variety of training algorithms

Training algorithm	**Accuracy** (%)
traincgf	74.44
trainlm	76.67
traingd	**82.22**
traingdx	81.11
traingda	75.56
traincgp	75.56

3.2 Experiment with Feature Selection

The experiment results using BPSO-MLP are shown in Table 2. The best accuracy is achieved using traincgp, which is 91.11 %. The lowest accuracy is achieved using traingd, which is equal to 88.90 %, while the others produce the same accuracy, which is 90 %. The proposed method can significantly increase the accuracy, namely 8.89 % compared to the previous experiment (without feature selection). Therefore, it can be concluded that adding the BPSO feature selection process can increase accuracy. In addition, it can reduce the number of features so that computing time is reduced.

Table 2: Experiment results using BPSO-MLP with a variety of training algorithms

Training Algorithm	**Accuracy** (%)	**Fitness**	**Feature Number**	**Feature**
traincgf	90.00	0.1023	4	f-5, f-7, f-11
trainlm	90.00	0.104	6	f-1, f-3, f-7, f-9, f-10, f-12
traingd	88.89	0.1125	3	f-5, f-10, f-12
traingdx	90.00	0.1023	4	f-3, f-6, f-11, f-12
traingda	90.00	0.1023	4	f-3, f-8, f-10, f-12
traincgp	**91.11**	0.093	6	f-3, f-4, f-5, f-9, f-10, f-12

4 Discussion

The experiment used 12 input features, namely f-1 to f-12. The proposed method searches for the best combination of features. Six features (f-3, f-4, f-5, f-9, f-10, and f-12) were combined to provide the most accurate results. For this reason, these features significantly influence the prediction of patient survival.

As demonstrated in Table 3, the proposed approach contrasted with earlier studies. In this study, [4] used a decision tree classifier, [2] used stratified logistic regression, [5] combined balanced random forest (BRF) and chi-square, [7] used random forest (RF), and [8] combined RF with SMOTE-ENN. The proposed method is the most accurate compared to earlier research. Previous research has used machine learning methods. It also combines machine learning and filter feature selection, while the proposed method combines machine learning methods (MLP) and wrapper feature selection (BPSO). Wrapper feature selection can select relevant features and delete irrelevant ones to increase accuracy.

Table 3: Comparison of the proposed method with previous studies

Method	**Accuracy** (%)
[4]	80.00
[2]	83.80
[5]	76.25
[7]	86.20
[8]	90.00
Proposed method	91.11

5 Conclusion

This paper proposed the BPSO-MLP method to predict survival in renal failure patients. It was a merger of BPSO and MLP. The testing results showed that the suggested strategy might generate an accuracy of up to 91.11 %. The proposed method can increase accuracy by 8.89 % compared to MLP (without feature selection). The addition of the BPSO wrapper feature selection significantly influences the accuracy results. In future research, the proposed method and selected features can be used to develop an expert system for predicting the survival of kidney failure patients.

References

[1] L. Kusumawati and D. D. Lukitasari, "Kepatuhan minum obat dengan interval waktu kejadian rawat inap ulang pasien gagal jantung di rumah sakit x," *Java Health Jounal*, vol. 10, no. 01, pp. 52–60, 2023.

[2] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," *BMC Medical Informatics and Decision Making*, vol. 20, p. 16, Dec. 2020.

[3] A.-Z. S. Bin Habib and T. Tasnim, "An ensemble hard voting model for cardiovascular disease prediction," in *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, (Dhaka, Bangladesh), pp. 1–6, IEEE, Dec. 2020.

[4] A. A. Almazroi, "Survival prediction among heart patients using machine learning techniques," *Mathematical Biosciences and Engineering*, vol. 19, no. 1, pp. 134–145, 2022.

[5] A. Newaz, N. Ahmed, and F. Shahriyar Haq, "Survival prediction of heart failure patients using machine learning techniques," *Informatics in Medicine Unlocked*, vol. 26, p. 100772, 2021.

[6] T. Admassu Assegie, T. Karpagam, S. Subramanian, S. M. Janakiraman, J. Arumugam, and D. O. Ahmed, "Prediction of patient survival from heart failure using a cox-based model," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, p. 1550, Sept. 2022.

[7] T. A. Assegie, V. Elanangai, J. S. Paulraj, M. Velmurugan, and D. F. Devesan, "Evaluation of feature scaling for improving the performance of supervised learning methods," *Bulletin of Electrical Engineering and Informatics*, vol. 12, pp. 1833–1838, June 2023.

[8] M. Muntasir Nishat, F. Faisal, I. Jahan Ratul, A. Al-Monsur, A. M. Ar-Rafi, S. M. Nasrullah, M. T. Reza, and M. R. H. Khan, "A comprehensive investigation of the performances of different machine learning classifiers with smote-enn oversampling technique and hyperparameter optimization for imbalanced heart failure dataset," *Scientific Programming*, vol. 2022, pp. 1–17, Mar. 2022.

[9] S. G. A, R. Prabha, M. Razmah, T. Veeramakali, S. S, and Y. R, "Machine learning heart disease prediction using knn and rtc algorithm," in *2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, (Chennai, India), pp. 1–5, IEEE, Dec. 2022.

[10] N. Mohan, V. Jain, and G. Agrawal, "Heart disease prediction using supervised machine learning algorithms," in *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, (Mathura, India), pp. 1–3, IEEE, Oct. 2021.

[11] B. Wang, K. He, Y. Bai, Z. Yao, J. Li, W. Dong, Y. Tu, W. Xue, Y. Tian, and Y. Wang, "A multi-task neural network architecture for renal dysfunction prediction in heart failure patients with electronic health records," *IEEE Access*, vol. 7, pp. 178392–178400, 2019.

[12] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–11, July 2021.

[13] Y. Chen, X. Qin, L. Zhang, and B. Yi, "A novel method of heart failure prediction based on dpcnn-xgboost model," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 495–510, 2020.

[14] L. Ali, A. Niamat, J. A. Khan, N. A. Golilarz, X. Xingzhong, A. Noor, R. Nour, and S. A. C. Bukhari, "An optimized stacked support vector machines based expert system for the effective prediction of heart failure," *IEEE Access*, vol. 7, pp. 54007–54014, 2019.

[15] E. Owusu, P. Boakye-Sekyerehene, J. K. Appati, and J. Y. Ludu, "Computer-aided diagnostics of heart disease risk prediction using boosting support vector machine," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–12, Dec. 2021.

[16] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart disease prediction using hybrid machine learning model," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, (Coimbatore, India), pp. 1329–1333, IEEE, Jan. 2021.

[17] L. Chandrika and K. Madhavi, "A hybrid framework for heart disease prediction using machine learning algorithms," *E3S Web of Conferences*, vol. 309, p. 01043, 2021.

[18] K. Dissanayake and M. G. Md Johar, "Comparative study on heart disease prediction using feature selection techniques on classification algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2021, pp. 1–17, Nov. 2021.

[19] J. Zhang, Y. Xiong, and S. Min, "A new hybrid filter/wrapper algorithm for feature selection in classification," *Analytica Chimica Acta*, vol. 1080, pp. 43–54, Nov. 2019.

[20] B. Remeseiro and V. Bolon-Canedo, "A review of feature selection methods in medical applications," *Computers in Biology and Medicine*, vol. 112, p. 103375, Sept. 2019.

[21] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, (Perth, WA, Australia), pp. 1942–1948, IEEE, 1995.

[22] Y. Hou, J. Li, H. Yu, and Z. Li, "Biffoa: a novel binary improved fruit fly algorithm for feature selection," *IEEE Access*, vol. 7, pp. 81177–81194, 2019.

[23] X.-w. Chen, Z.-h. Liu, and Z.-k. Zhang, "The measurement of planning surface roughness by neural networks based on image," in *2010 Sixth International Conference on Natural Computation*, vol. 2, pp. 705–708, Aug. 2010. ISSN: 2157-9563.

[24] H. Mustafidah, S. Hartati, R. Wardoyo, and A. Harjoko, "Selection of most appropriate backpropagation training algorithm in data pattern recognition," *International Journal of Computer Trends and Technology*, vol. 14, pp. 92–95, Aug. 2014. arXiv:1409.4727 [cs].

[25] M. Paliwal and U. A. Kumar, "Neural networks and statistical techniques: A review of applications," *Expert Systems with Applications*, vol. 36, pp. 2–17, Jan. 2009.

[26] D. Chicco, "Heart failure clinical records," 2020.