IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2290~2299
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2290-2299

Combination of binary particle swarm optimization and random forest for stroke disease prediction

Sutikno[1], **Rismiyati**[1], **Khadijah**[1], **Abdul Karim**[2]
[1]Department of Informatics, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia
[2]Cerebrovascular Disease Research Center, Hallym University, Chuncheon, South Korea

Article Info

Article history:

Received Sep 19, 2024
Revised Feb 12, 2026
Accepted Apr 20, 2026

Keywords:

Binary particle swarm optimization
Integrated model
Prediction
Random forest
Stroke disease

ABSTRACT

Stroke is a leading cause of death and disability worldwide, making early risk prediction critical for prevention. Machine learning methods such as random forest (RF) have shown strong predictive performance, but accuracy can be further improved through effective feature selection. This research proposes an integrated model that combines binary particle swarm optimization (BPSO) for feature selection with RF for stroke risk classification. Experiments were conducted on two public datasets: the stroke prediction dataset (SPD) and the brain stroke dataset (BSD). Data preprocessing included handling missing values, normalization, and the synthetic minority oversampling technique (SMOTE) to mitigate the minority and majority classes. BPSO was employed to select the most informative features, followed by RF for classification. The BPSO-RF model delivered superior accuracies of 96.13% on the SPD and 96.07% on the BSD, outperforming competing classifiers and feature selection techniques. Important features such as gender, age, work type, residence type, average glucose level, body mass index (BMI), and smoking status were consistently identified as key predictors. These results indicate that integrating swarm intelligence with ensemble learning can effectively improve stroke risk prediction and support clinical decision-making.

Corresponding Author:

Sutikno
Department of Informatics, Faculty of Science and Mathematics, Diponegoro University
Prof. Jacob Rais Street, Semarang, Central Java, Indonesia
Email: sutikno@lecturer.undip.ac.id

1. INTRODUCTION

Stroke is deadly disease that significantly affects the quality of life by impairing physical, cognitive, and emotional well-being. The World Health Organization (WHO) reported that stroke represents the main cause of mortality and disability globally [1]. In prevention efforts, stroke risk prediction plays an important role and has been successfully implemented using machine learning algorithms over the past few decades.

Previous research proposed the support vector machine (SVM) [2]–[5] for stroke prediction. Compared with several machine learning algorithms, including naive Bayes (NB), logistic regression (LR), and decision tree (DT), the proposed method achieved higher accuracy. The random forest (RF) algorithm [6], [7] was also developed and showed superior performance compared to LR, k-nearest neighbor (KNN), SVM, extreme gradient boosting (XGBoost), adaptive boosting (AdaBoost), DT, and light gradient-boosting machine (LightGBM). Furthermore, the use of artificial neural network (ANN) [6], adaptive neuro-fuzzy inference system (ANFIS) [7], and convolution neural network (CNN) has been reported in stroke prediction efforts [8].

Several methods have been used to improve accuracy, including ensembling, oversampling, and feature selection. Ensembling methods applied include XGBoost [9]–[11], AdaBoost [12], and stacking

[13]–[15], with results showing significant improvements in accuracy. The oversampling method can also improve accuracy on unbalanced datasets. Srinivasu *et al.* [16] applied the synthetic minority oversampling technique (SMOTE) to balance the class distribution and used an ANN as the classifier. This method was also used in [17], who reported significantly superior performance. Meanwhile, Biswas *et al.* [18] used the random over sampling (ROS) method and an SVM classifier, resulting in a substantial increase in accuracy.

Another method to increase accuracy is to add a feature selection process to identify the best features in the dataset [19]. Lee *et al.* [20] proposed principal component analysis (PCA) feature selection and oversampling for stroke prediction. Meanwhile, Victor *et al.* [21] proposed using particle swarm optimization (PSO) for feature selection and federated learning (FL) as a classifier. The results showed that these methods could increase accuracy, with PSO as the wrapper feature selection method producing relatively superior performance. In some reports, RF [22], [23] and PSO [21] achieved better accuracy than other methods. Therefore, this research proposes a novel hybrid BPSO–RF approach, the first integrating binary particle swarm optimization (BPSO) and RF for stroke prediction, to enhance classification accuracy and feature relevance. The results were expected to provide valuable information on prediction efforts for faster and more effective care to be provided to patients at increased risk of stroke.

2. METHOD

The proposed research steps in stroke prediction are depicted in Figure 1. The input of this research is the datasets and the output is model evaluation. The main processes are preprocessing, oversampling, splitting data, feature selection, and classification.

Figure 1. The proposed research steps in stroke prediction

2.1. Dataset

This research used two public datasets: the stroke prediction dataset (SPD) and the brain stroke dataset (BSD), available on Kaggle [24], [25]. Both datasets contain data on stroke and non-stroke patients, each with 12 attributes. The stroke attributes, which take values 0 or 1, are the targets, while the others are features. A value of 0 indicates the patient did not have a stroke, while a value of 1 indicates the patient did. The SPD dataset contains 5,110 rows, with 4.87% stroke patients and 95.13% non-stroke patients. Meanwhile, the BSD dataset contains 4,981 rows, with 4.98% stroke patients and 95.02% non-stroke patients. It appears that the classes in both datasets are imbalanced. The data imbalance was a challenge for this research, alongside missing values in several attributes, underscoring the need for special handling.

2.2. Preprocessing

Preprocessing functions to prepare data and facilitate the machine learning algorithms, thereby improving data quality and enhancing model performance. In this research, preprocessing comprised three stages: handling missing values, transforming attributes, and normalization. Mean imputation is used handle missing values. The features that have this issue are gender, body mass index (BMI), and smoking status. After that, the ID column was removed because it did not affect the model's results. Subsequently, attribute transforming was performed by changing the string value to an integer. The five-string attributes were gender, ever married, work type, residence type, and smoking status. The last step in preprocessing was normalization to produce numeric values in the range 0 to 1, calculated using (1). Where x_{norm} is the normalized feature value, x is the actual value of the feature, x_{min} is lowest value, and x_{max} is highest value.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

2.3. Oversampling

The imbalance in positive samples biases the model toward the majority class and degrades detection performance, necessitating class-balancing techniques such as SMOTE to generate synthetic minority samples. SMOTE generates synthetic data points by interpolating among samples of the minority class. Given two samples, x_i and x_j, from the minority class, a new synthetic instance is generated as (2). Where x_{new} is the new artificial sample, x_i and x_j are two randomly selected minority-class samples.

The variable λ denotes a randomly sampled value in the interval [0, 1] that controls the interpolation. This process helps the model learn from underrepresented classes by generating additional training data, thus improving the balance between the positive and negative classes. Resampling techniques have been widely used to address class imbalance and have been empirically shown to improve the model's ability to learn underrepresented minority classes [26].

$$x_{new} = x_i + \lambda \times (x_j - x_i) \tag{2}$$

2.4. Data splitting

Data splitting was performed after oversampling. Data splitting was performed using 5-fold cross-validation. The final result was the average.

2.5. Feature selection

This research used BPSO feature selection to solve binary optimization problems. The algorithm is a development of PSO introduced by Kennedy and Eberhart [27]. In BPSO, each particle is expressed as a sequence of binary bits. The movement of particles is influenced by individual experience and the best experience of their neighbors in finding the optimal solution.

BPSO includes steps to reach an optimal solution in a binary search space. Initially, a population of particles is randomly established. This step is followed by evaluating each particle's fitness value using the objective function to be optimized. Furthermore, the particle in the population updates its velocity and position. The velocity is influenced by the previous results, individual experience (pbest), and global experience (gbest). The velocity update moves the particle in the binary solution space. Subsequently, the fitness value of the new particle is evaluated, and when the new position provides better fitness, the particle pbest is updated.

When the search space is $S = \{0,1\}^D$ and the objective function f aims to maximize, $max\, f(x)$, particle i can be formulated as a D-dimensional vector, $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T$, with $x_{id} \in \{0, 1\}$, $d = 1, 2, \dots, D$. The velocity is given by $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})^T$, with $v_{id} \in [-V_{max}, V_{max}]$, where V_{max} is the highest velocity value. The best position that particle i has ever visited is represented by $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})^T$, where $p_{id} \in \{0,1\}$. When g is defined as the best index in the swarm and p_{gd} is the best position, the velocity update is calculated using (3), where λ is a parameter that controls the steepness of the sigmoid function. The position update is then formulated by (4).

$$v_{id} = v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \tag{3}$$

$$x_{id} = \begin{cases} 1 & if\ U(0,1) < sigm(v_{id}), \\ 2 & otherwise, \end{cases} \tag{4}$$

The feature selection process is in multiple stages, as illustrated in Figure 2. This process combined two algorithms: BPSO, used to select the optimal feature subset; and RF, used to compute the fitness value. The results of this process are the best feature subset and the highest accuracy.

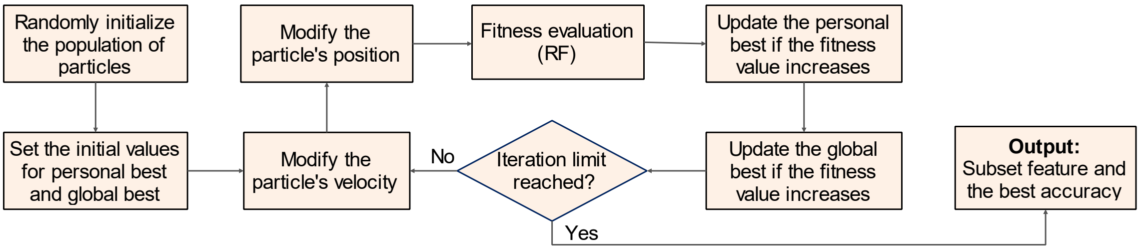


Figure 2. Feature selection process using BPSO and RF

The initial stage involves determining the number of particles (*N*) and the number of dimensions per particle (*D*). In this research, *D* is equal to the number of features. There are 11 features, with each dimension taking values 0 or 1. Specifically, 0 corresponds to a non-selected feature, and 1 to a selected feature. During initialization, these binary values are randomly determined.

A personal best is a particle that already exists, while a global best has the best fitness value. Fitness is calculated using (5) [28]. Where γ_R is the average classification error, $|R|$ represents the number of selected features, $|C|$ denotes the complete set of features, α is the weight for classification accuracy, and β is the weight for the selected features. KNN is often used as a classifier [29]. This research uses an RF classifier with $\alpha = 0.99$ and $\beta = 0.01$ [28].

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|C|} \tag{5}$$

2.6. Classification

RF is a tree-based ensemble algorithm that forms multiple DT from randomly selected samples and combines their predictions through majority voting. Its main advantages are robustness against overfitting, strong generalization performance, and the ability to identify important features [30]. RF applies bootstrap sampling to generate multiple datasets and randomly selects feature subsets at each node, where splits are determined using criteria such as Gini impurity or entropy. Gini impurity measures the imperfection of a split at a particular node using (6). The variable of p_i is proportion of samples with class i at that node, and C is number of classes. Another alternative for measuring irregularity at a node is entropy, as defined in (7).

$$Gini(D) = 1 - \sum_{i=1}^{C} p_i^2 \tag{6}$$

$$Entropy(D) = -\sum_{i=1}^{C} p_i \, log_2(p_i) \tag{7}$$

The splitting process continues until it meets a specific stop condition. After all trees are formed, the final step is voting. Each tree provides a class prediction, and the final prediction results from majority voting across all trees, as in (8). Where $h_t(x)$ is the prediction from tree t for sample x, and mode returns the most frequent class among all trees.

$$\hat{y} = mode\{h_t(x)\}_{t=1}^{T} \tag{8}$$

2.7. Model evaluation

The performance evaluation was conducted using the metrics of accuracy, precision, recall, and F1-score, calculated sequentially with (9) to (12) [30]. The values of true positive (T_{pos}), true negative (T_{neg}), false positive (F_{pos}), and false negative (F_{neg}) were derived from confusion matrix, as illustrated in Table 1.

$$Accuration = \frac{T_{pos}+T_{neg}}{T_{pos}+T_{neg}+F_{pos}+F_{neg}} \tag{9}$$

$$Precision = \frac{T_{pos}}{T_{pos}+F_{pos}} \tag{10}$$

$$Recall = \frac{T_{pos}}{T_{pos}+F_{neg}} \tag{11}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

Table 1. Confusion matrix of stroke prediction

		Actual values	
		Stroke	Non-stroke
Predicted values	Stroke	T_{pos}	F_{pos}
	Non-stroke	F_{neg}	T_{neg}

3. RESULTS AND DISCUSSION

Testing was conducted using several scenarios. In the first scenario, we tested the RF classifier without feature selection and then compared it with other classifiers, namely XGBoost, LightGBM, and categorical boosting (CatBoost). All tests were performed with hyperparameter tuning to achieve the best accuracy. The range of values for the tuned parameter is shown in Table 2. In the following scenario, we tested BPSO feature selection with the RF classifier and compared it with other feature selection methods, including mutual information (MI), genetic algorithms (GA), and Boruta. The parameters used from the feature selection are shown in Table 3.

Table 2. Range value of the parameter tuning experiment in the classifier

Classifier	Range value of the parameter
RF	Number of trees: 50, 100, and 150 Minimum samples per leaf: 1, 2, 3, and 4 Number of features per split: 2 to 10
XGBoost	Tree count: 50, 100, and 150 The depth limit of each tree: 3, 4, and 5
LigtGBM	Tree count: 50, 100, and 150 The depth limit of each tree: 3, 4, and 5
CatBoost	Tree count: 50, 100, 150 The depth limit of each tree: 3, 4, and 5

Table 3. The set of parameter ranges considered for tuning in the feature selection stage

Feature selection	Range value of the parameter
PSO	Number of particles: 12 Maximum iteration: 50
MI	Feature: 10% to 100% (increament:10%)
GA	Population size: 12 Number of generations: 50 Crossover rate: 0.8 Mutation rate: 0.05

The results of the first test used RF for classification without feature selection. It was then compared with classifiers XGBoost, LightGBM, and CatBoost. The comparison results of these classifiers are shown in Table 4. RF achieves higher accuracy, precision, recall, and F1-score than the other classifiers on both datasets. Therefore, RF is more suitable for stroke disease classification than the other three classifiers.

The next test involved adding a feature selection process. The proposed feature selection method is BPSO. The results of this test are shown in Table 5. The proposed method achieved an accuracy of 0.9613 on the SPD dataset and 0.9607 on the BSD dataset. Adding feature selection increased accuracy by 0.08% on the SPD dataset and 0.01% on the BSD dataset. The number of features in both datasets is 10. The BPSO feature selection yielded seven dominant features in the SPD dataset: gender, age, work type, residence type, average glucose level, BMI, and smoking status. Meanwhile, in the BSD dataset, eight dominant features were identified: gender, age, ever married, work type, residence type, average glucose level, BMI, and smoking status. Two features were not selected in the tests for both datasets, namely hypertension and heart disease.

Table 4. Comparison of RF classifier with XGBoost, LightGBM, and CatBoost

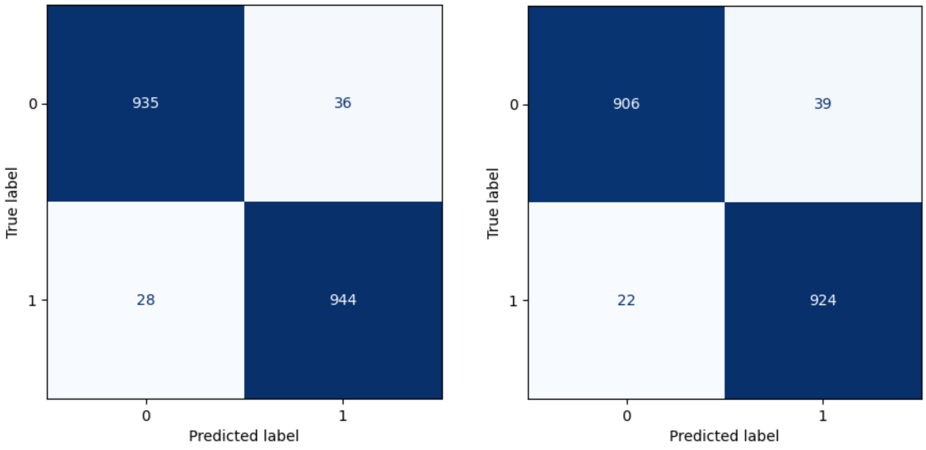
Method	SPD dataset				BSD dataset			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
XGBoost	0.9074	0.9218	0.8903	0.9057	0.9180	0.9247	0.9102	0.9174
LigtGBM	0.9212	0.9456	0.8940	0.9190	0.9275	0.9443	0.9087	0.9262
CatBoost	0.8582	0.8677	0.8454	0.8564	0.8679	0.8757	0.8578	0.8666
RF	0.9605	0.9538	0.9679	0.9607	0.9606	0.9553	0.9664	0.9608

Table 5. Results of the proposed method using BPSO feature selection and RF classifier

Fold number	SPD dataset					BSD dataset				
	Accuracy	Precision	Recall	F1-score	Time (s)	Accuracy	Precision	Recall	F1-score	Time (s)
1	0.9671	0.9633	0.9712	0.9672	647	0.9572	0.9482	0.9672	0.9576	2377
2	0.9619	0.9591	0.9650	0.9620	662	0.9677	0.9595	0.9767	0.9680	2396
3	0.9624	0.9572	0.9681	0.9626	658	0.9625	0.9553	0.9704	0.9628	2357
4	0.9599	0.9434	0.9784	0.9606	654	0.9635	0.9592	0.9683	0.9637	2419
5	0.9553	0.9511	0.9599	0.9555	654	0.9562	0.9482	0.9652	0.9567	2434
Average	0.9613	0.9548	0.9683	0.9615	655	0.9607	0.9538	0.9683	0.9610	2397

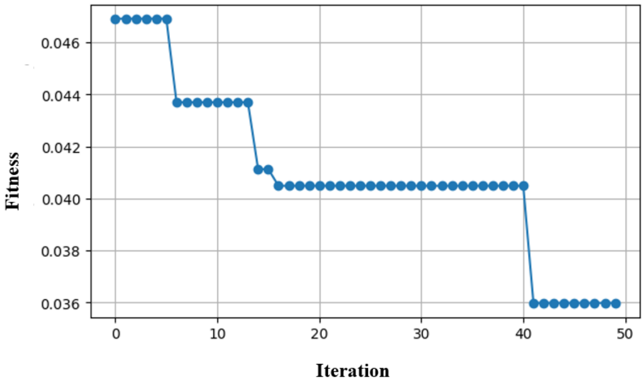
The confusion matrix of the experiment is shown in Figure 3, where Figure 3(a) shows results for the SPD dataset and Figure 3(b) shows results for the BSD dataset. Evaluation results on the best number of folds indicated that, for the SPD dataset, the model made 64 classification errors (3.29%), consisting of

36 false positives (1.85%) and 28 false negatives (1.44%). For the BSD dataset, there were 61 classification errors (3.23%), including 39 false positives (2.06%) and 22 false negatives (1.16%). Figure 4 presents the results of the fitness function for feature selection. Figure 4(a) presents the fitness obtained from the SPD dataset, while Figure 4(b) shows the fitness for the BSD dataset. The performance improvement obtained by integrating BPSO with RF is mainly due to BPSO's ability to select a more informative feature subset while eliminating redundant and irrelevant features. This feature reduction reduces noise and overfitting, enabling the RF to understand discriminative patterns better and improve generalization.

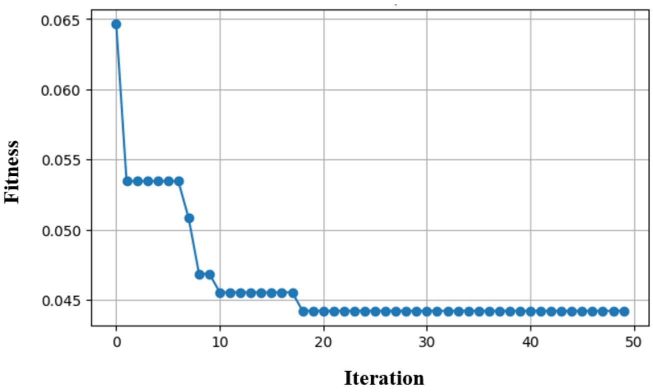


(a) (b)

Figure 3. Confusion matrix of the proposed method results for (a) SPD dataset and (b) BSD dataset



(a)



(b)

Figure 4. Convergence curve of BPSO feature selection for (a) SPD dataset and (b) BSD dataset

The proposed feature selection was then compared with other feature selection methods, namely Boruta, MI, and GA. The results of this comparison are shown in Table 6. In the SPD dataset, the proposed method outperforms the others in relation to accuracy, recall, and F1-score. Meanwhile, in the BSD dataset, MI performs better than the others, with the proposed method slightly below it. Boruta and MI require less time compared to GA and PSO. Boruta and MI are filter-based feature selection methods, whereas GA and PSO are wrapper-based approaches that generally require longer computational time.

Table 6. Performance comparison of the proposed method with other feature selection methods

Method	SPD dataset					BSD dataset				
	Accuracy	Precision	Recall	F1-score	Time (s)	Accuracy	Precision	Recall	F1-score	Time (s)
Boruta	0.9305	0.9295	0.9319	0.9306	**12**	0.9308	0.9280	0.9341	0.9310	**10**
MI	0.9566	**0.9593**	0.9537	0.9563	31	**0.9614**	0.9535	**0.9702**	**0.9618**	32
GA	0.9587	0.9522	0.9660	0.9590	1641	0.9600	**0.9541**	0.9666	0.9603	1624
Proposed	**0.9613**	0.9548	**0.9683**	**0.9615**	655	0.9607	0.9538	0.9683	0.9610	2397

A comparative analysis is conducted between the proposed method, previous research approaches, and deep learning models. These models are SVM [3]–[5], DT [23], KNN [23], NB [3], and CNN, as summarized in Table 7. The outcomes indicate that the proposed method yields better accuracy, precision, and F1-score than the others, although its recall is lower than that of the KNN method. Comparison of accuracy results between the proposed method and other methods is shown in Figure 5.

The study suggests that the proposed BPSO-RF model can support early stroke screening by identifying high-risk individuals and enabling preventive strategies through lifestyle monitoring and interventions based on key risk factors. However, this study has several limitations, including the use of a relatively small data set. Furthermore, stroke prediction was simplified as a binary classification task, whereas real-world clinical settings involve a variety of stroke subtypes and severity levels that were not considered.

Table 7. Comparative assessment of the proposed method and related studies

Method	SPD dataset				BSD dataset			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
SVM [3]–[5]	0.8329	0.8116	0.8670	0.8384	0.8412	0.8174	0.8793	0.8471
DT [23]	0.9251	0.9148	0.9376	0.9260	0.9253	0.9144	0.9387	0.9263
KNN [23]	0.9130	0.8562	**0.9930**	0.9195	0.9186	0.8631	**0.9951**	0.9244
NB [3]	0.7430	0.7503	0.7282	0.7390	0.7242	0.7439	0.6841	0.7126
CNN	0.8613	0.8259	0.9164	0.8686	0.8731	0.8394	0.9233	0.8793
Proposed	**0.9613**	**0.9548**	0.9683	**0.9615**	**0.9607**	**0.9538**	0.9683	**0.9610**

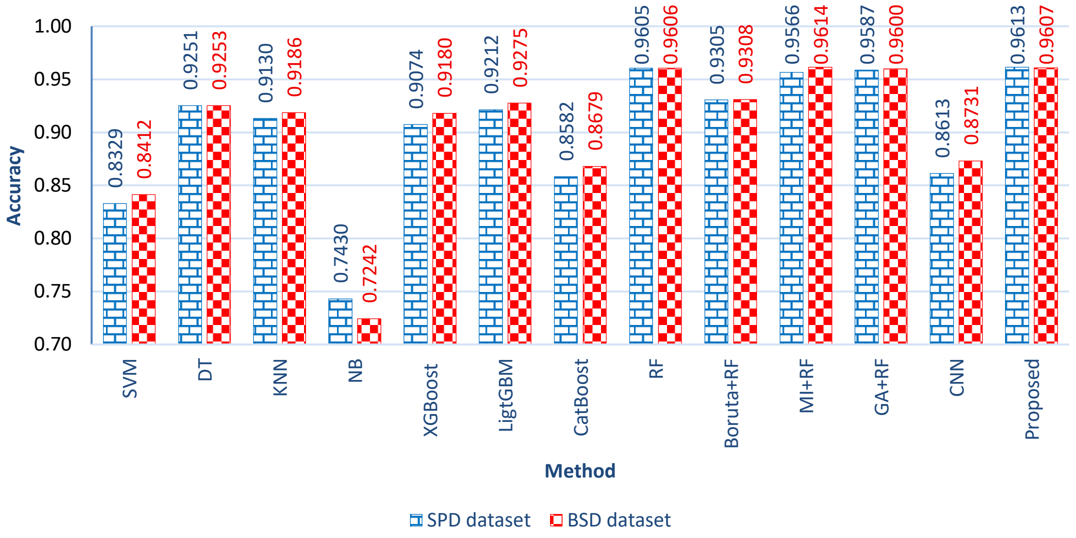


Figure 5. Comparison of accuracy between the proposed method and other methods

4. CONCLUSION

This study proposed combining BPSO and RF to predict stroke, representing the first systematic integration of BPSO with RF. The BPSO method selected features, and RF was applied to classify them. The results showed that adding BPSO could improve accuracy, achieving better results than previous research. The BPSO method identified important features that significantly affected stroke disease. These features were gender, age, work type, residence type, average glucose level, BMI, and smoking status for the SPD dataset; and gender, age, ever married, work type, residence type, average glucose level, BMI, and smoking status for the BSD dataset. Future studies should explore the use of larger and more diverse datasets, the integration of multimodal clinical data, and the adoption of explainable AI techniques to further enhance model robustness, clinical relevance, and interpretability. In addition, the proposed model can be extended toward integration into real-time hospital applications and clinical decision support systems to support healthcare professionals in early stroke risk assessment and timely intervention.

FUNDING INFORMATION

This research was funded from Faculty of Science and Mathematics at Diponegoro University through funding Dean's Decree Number 20/UN7.F8/HK/II/2024.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sutikno	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Rismiyati		✓		✓		✓			✓	✓	✓			✓
Khadijah			✓		✓				✓	✓	✓		✓	✓
Abdul Karim	✓	✓			✓	✓			✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

This paper presents two datasets that have been used and are available from Kaggle in the stroke prediction dataset, reference number [24] and in the brain stroke dataset, reference number [25].

REFERENCES

[1] V. L. Feigin *et al.*, “World stroke organization (WSO): global stroke fact sheet 2022,” *International Journal of Stroke*, vol. 17, no. 1, pp. 18–29, 2022, doi: 10.1177/17474930211065917.

[2] M. Altayeb and A. Arabiat, “Enhancing stroke prediction using the waikato environment for knowledge analysis,” *IAES International Journal of Artificial Intelligence*, vol. 13, no. 3, pp. 3010–3017, 2024, doi: 10.11591/ijai.v13.i3.pp3010-3017.

[3] I. Mishra and S. Mohapatra, “An enhanced approach for analyzing the performance of heart stroke prediction with machine learning techniques,” *International Journal of Information Technology*, vol. 15, no. 6, pp. 3257–3270, 2023, doi: 10.1007/s41870-023-01321-8.

[4] S. Mushtaq, K. S. Saini, and S. Bashir, “Machine learning for brain stroke prediction,” in *2023 International Conference on Disruptive Technologies (ICDT)*, 2023, pp. 401–408, doi: 10.1109/ICDT57929.2023.10151148.

[5] I. G. Ivanov, Y. Kumchev, and V. J. Hooper, “An optimization precise model of stroke data to improve stroke prediction,” *Algorithms*, vol. 16, no. 9, 2023, doi: 10.3390/a16090417.

[6] G. F. Issa, K. Shaalan, Y. Shaalan, H. A. Saeed, N. Fatima, and A. U. Rehman, “Brain stroke prediction using ANN,” in *2022 International Conference on Cyber Resilience (ICCR)*, 2022, pp. 1–9, doi: 10.1109/ICCR56254.2022.9995885.

[7] A. Lakkshmanan, S. R. Adaline, N. Priyanka, D. B. Anand, K. Nagaraju, and U. G. Naidu, “A novel machine learning based stroke prediction system using magnetic resonance imaging and adaptive new fuzzy inference system,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 10s, pp. 576–585, 2024.

[8] M. Ashrafuzzaman, S. Saha, and K. Nur, “Prediction of stroke disease using deep CNN-based approach,” *Journal of Advances in Information Technology*, vol. 13, no. 6, 2022, doi: 10.12720/jait.13.6.604-613.

[9] X. Ye, W. Xu, X. Ye, D. Long, Q. Yin, and B. Huang, “Stroke prediction using the trust evaluation with data leakage avoiding,” *Journal of Physics: Conference Series*, vol. 2560, no. 1, 2023, doi: 10.1088/1742-6596/2560/1/012051.

[10] M. G. Rao, K. P, K. R. Kotian, and D. P. Hegde, “Multimodal approach for predicting the brain stroke using the machine learning techniques,” in *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE)*, 2022, pp. 242–247, doi: 10.1109/AIDE57180.2022.10060424.

[11] U. Singh, A. K. Jena, and M. T. Haque, “An ensemble learning approach and analysis for stroke prediction dataset,” in *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, 2022, pp. 1–8, doi: 10.1109/ASSIC55218.2022.10088363.

[12] P. Pal, M. Nandal, S. Dikshit, A. Thusu, and H. V. Singh, “Harnessing the power of ensemble machine learning for the heart stroke classification,” *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, no. 1, 2023, doi: 10.4108/eetpht.9.4617.

[13] E. Dritsas and M. Trigka, “Stroke risk prediction with machine learning techniques,” *Sensors*, vol. 22, no. 13, 2022, doi: 10.3390/s22134670.

[14] S. A. Mostafa, D. S. Elzanfaly, and A. E. Yakoub, “A machine learning ensemble classifier for prediction of brain strokes,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 12, 2022, doi: 10.14569/IJACSA.2022.0131232.

[15] A. Rehman, T. Alam, M. Mujahid, F. S. Alamri, B. Al Ghofaily, and T. Saba, “RDET stacking classifier: a novel machine learning based approach for stroke prediction using imbalance data,” *PeerJ Computer Science*, vol. 9, 2023, doi: 10.7717/peerj-cs.1684.

[16] P. N. Srinivasu, U. Sirisha, K. Sandeep, S. P. Praveen, L. P. Maguluri, and T. Bikku, “An interpretable approach with explainable AI for heart stroke prediction,” *Diagnostics*, vol. 14, no. 2, 2024, doi: 10.3390/diagnostics14020128.

[17] H. M. -Solis, R. P. -Siguas, E. M. -Solis, L. M. -Zamudio, S. M. -Gomez, and J. F. V. -Molina, “Application of machine learning for the prediction of strokes in Peru,” *International Journal of Engineering Trends and Technology*, vol. 70, no. 10, pp. 54–60, 2022, doi: 10.14445/22315381/IJETT-V70I10P207.

[18] N. Biswas, K. M. M. Uddin, S. T. Rikta, and S. K. Dey, “A comparative analysis of machine learning classifiers for stroke prediction: a predictive analytics approach,” *Healthcare Analytics*, vol. 2, 2022, doi: 10.1016/j.health.2022.100116.

[19] E. Stenwig, P. S. Rossi, G. Salvi, and N. K. Skjærvold, “The impact of feature combinations on machine learning models for in-hospital mortality prediction,” *Scientific Reports*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-26611-y.

[20] J. Lee *et al.*, “Comparative analysis of neural network method on stroke prediction,” in *2023 10th International Conference on ICT for Smart Society (ICISS)*, 2023, pp. 1–6, doi: 10.1109/ICISS59129.2023.10291404.

[21] N. Victor, S. Bhattacharya, P. K. R. Maddikunta, F. M. Alotaibi, T. R. Gadekallu, and R. H. Jhaveri, “FL-PSO: a federated learning approach with particle swarm optimization for brain stroke prediction,” in *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)*, 2023, pp. 33–38, doi: 10.1109/CCGridW59191.2023.00020.

[22] Y. Miao, “Using machine learning method to predict stroke risk,” in *2021 2nd International Conference on Information Science and Education (ICISE-IE)*, 2021, pp. 665–668, doi: 10.1109/ICISE-IE53922.2021.00156.

[23] M. C. Das *et al.*, “A comparative study of machine learning approaches for heart stroke prediction,” in *2023 International Conference on Smart Applications, Communications and Networking (SmartNets)*, 2023, pp. 1–6, doi: 10.1109/SmartNets58706.2023.10216049.

[24] F. S. Palacios, “Stroke prediction dataset,” *Kaggle*. 2021. Accessed: Aug. 30, 2024. [Online]. Available: https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset

[25] G. Jillani, “Brain stroke dataset,” *Kaggle*, 2022. Accessed: Aug. 30, 2024. [Online]. Available: https://www.kaggle.com/datasets/jillanisofttech/brain-stroke-dataset

[26] Z. Zhao, T. Cui, S. Ding, J. Li, and A. G. Bellotti, “Resampling techniques study on class imbalance problem in credit risk prediction,” *Mathematics*, vol. 12, no. 5, 2024, doi: 10.3390/math12050701.

[27] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95 - International Conference on Neural Networks*, 1995, vol. 4, pp. 1942–1948, doi: 10.1109/ICNN.1995.488968.

[28] W.-L. Cheng *et al.*, “An adaptive and parallel metaheuristic framework for wrapper-based feature selection using arctic puffin optimization,” *Computers, Materials & Continua*, vol. 85, no. 1, pp. 2021–2050, 2025, doi: 10.32604/cmc.2025.064243.

[29] M. Al Ghamri, D. Ibrahim, R. Sihwail, and M. Shehab, “Whale optimization algorithm for feature selection enhances classification in malware datasets,” *Journal of Computational and Cognitive Engineering*, vol. 4, no. 3, pp. 387–396, 2024, doi: 10.47852/bonviewJCCE42024233.

[30] A. Gummaraju, A. K. B. Shenoy, and S. N. Pai, “Performance comparison of machine learning models for handwritten Devanagari numerals classification,” *IEEE Access*, vol. 11, pp. 133363–133371, 2023, doi: 10.1109/ACCESS.2023.3336912.

BIOGRAPHIES OF AUTHORS

Dr. Sutikno received a Doctor of Philosophy (Ph.D.) degree in Computer Science from the Faculty of Mathematics and Natural Sciences at the Universitas Gadjah Mada, Indonesia. Now, he is an assistant professor at the Department of Informatics, Faculty of Sciences and Mathematics, Universitas Diponegoro. His research interests include machine learning, computer vision, and artificial intelligence. He can be contacted at email: sutikno@lecturer.undip.ac.id.

Rismiyati received a Bachelor of Engineering (B.Eng.) from the School of Electrical and Electronics Engineering, Nanyang Technological University (NTU), in 2007 and a Master of Computer Science (M.Cs.) from the Universitas Gadjah Mada, Indonesia, in 2016. She has been a lecturer in the Department of Informatics at Universitas Diponegoro since 2017. Her main research interests are artificial intelligence, image processing, and machine learning. She can be contacted at email: rismiyati@live.undip.ac.id.

Khadijah received a Bachelor of Informatics Engineering (S.Kom.) from the Universitas Diponegoro, Indonesia, in 2011 and a Master of Computer Science (M.Cs.) from the Universitas Gadjah Mada, Indonesia, in 2014. She has been a lecturer in the Department of Informatics at Universitas Diponegoro since 2014. Her main research interests are artificial intelligence and machine learning. She can be contacted at email: khadijah@live.undip.ac.id.

Dr. Abdul Karim received his Ph.D. from Universitas Gadjah Mada, Yogyakarta, Indonesia. He is currently a research professor at Hallym University, South Korea. His research interests span a wide range of topics, including natural language processing, artificial intelligence, machine learning, deep learning, software engineering, and computer vision. He is actively working on large language models (LLMs) and their applications to various real-world problems, advancing AI technologies in both academic and practical domains. He can be contacted at email: abdullkarim@hotmail.com.