Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

Prediction of Coronary Artery Disease Using Enhanced Feature Selection Using Firefly Based Optimization

Jayamol P James[1*], E. Chandra Blessie[2]

[1] Research Scholar, Department of Computer Science, Nehru College of Management, India
[2] Assistant Professor, Department of Computing(AIML), Coimbatore Institute of Technology, India
[1]cooljaimol@gmail.com

ABSTRACT

Coronary Artery Disease (CAD) is an uncontrolled deadly disease that causes unexpected sudden death and it is rapidly increasing among people. CAD is becoming most common in middle age and old age people. So far angiography is considered as a better method to diagnose CAD, but it leads to severe side effects and highly expensive. Multiple researches have been conducted with the techniques involving data mining and machine learning to predict and diagnose CAD. In this paper, Enhanced Feature Selection based on Firefly Optimization (EFS-PO) is proposed for increasing the prediction of CAD. In EFS-PO, the performance of feature selection is enhanced by firefly optimization that ends with enhanced classification accuracy. Weights of the feature are enhanced via firefly optimization towards better selection. EFS-PO has achieved better results in terms of all considered benchmark data mining performance metrics.

Keywords
Classification, Heart Disease, Cleveland, Z-Alizadeh Sani, Statlog

Introduction

With relation to computer use in medicine, the accomplishments are astounding. Generally, the confirmed diagnosis and physical signs of a patient are addressed by a physician (Abdar et al., 2019; JayaSree & Koteswara Rao, 2020). The doctor's perspective view show which patient diagnosis is the most accurate. Being in the healthcare sector is an incredibly arduous activity due to the rise of medical research, advancement, and creativity. Computers can potentially hold a lot of data when they use specialized technology (Głowacki et al., 2020).

For medical decision support systems to be built and deployed, various difficult obstacles must be addressed. The algorithms will develop the capacity for decision-making to tolerate inaccuracy and ambiguity (Kawasaki et al., 2020; Yang et al., 2020). Medical expertise is required for a correct diagnosis and clinical evaluation of symptoms. In their daily environment, medical professionals used sophisticated machine learning techniques to improve decision-making in various areas to leverage computerized intelligent systems. Before beginning the procedure, a physician has been using his or her prior experience to develop a better understanding of an illness (Alizadehsani et al., 2019; Mishra et al., 2020). The doctor then administers multiple experiments to understand the severity of the disease using various medical tests. The medical test reports can also be made easily by supporting the practitioner with computerized smart systems. The new patient will be diagnosed with adequate justification and history of electronic patients used in a broad archive for knowing all prior cases (Han et al., 2020). Instance: The biggest advantage of the use of these smart systems is that patients will be monitored at the same time with minimum time consumption and expense to obtain better diagnostic precision (Li et al., 2020).

According to the Centers for Disease Control research, the world is suffering from rising mortality rates from coronary artery disease. Machine learning algorithms have become mandatory for treating patients suffering from different kinds of heart disease (Lee et al., 2020). Owing to the delay in medical attention, proper identification of heart disease turns into a difficult and time-consuming process. Such syndromes may occur regularly in the circulatory system which does not involve the heart.

The main intention of this paper is to propose optimization based feature selection methodology that assist in enhancing the classification accuracy of CAD.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

Literature Review

Disease-Specific Feature Selection strategy (Zhang et al., 2014) was proposed for heartbeat classification in an automated manner towards predicting the cardiac attack. It holds the idea of a 1-vs-1 feature towards searching for the best feature, where it uses the support vector machine classification concept. The result showed that the feature selection is not suited for this classification, where the results came with false-negative rate got increased. The multi-Objective Classification method (Esmaelian et al., 2016) was proposed with the ensemble of particle swarm optimization and genetic algorithms to predict heart disease in an early stage. It calculated the coefficient of polynomial, also the limit of the threshold value which was set for the class and attributes. This calculation was made to decrease the error, but the misclassification error got increased a lot in classifying to the wrong class Fuzzy Classifier (Vafaie et al., 2014) was proposed to perform classification with dynamic electrocardiogram signals to predict heart disease in an early stage. It has worked with the dynamic unknown features resulting in very low accuracy. It was also analyzed that the algorithm can work well with known features only. Identification of Heart Disease with an Embedded System (de Carvalho et al., 2013) was proposed and analyzed viability. It takes the input as electrocardiogram signals for the initial stage clustering and finally used Gustafson Kessel based Fuzzy clustering algorithm to classify and correlate the signals. The result came with an increased false-negative rate.

Hybrid Intelligent Modeling Scheme (Shao et al., 2014) was proposed to find the various set of a descriptive variable to classify the diseases related to the heart. It is an ensemble of (i) logistic regression, (ii) artificial neural network, (iii) multivariate adaptive regression splines, and (iv) rough set method. Initially, it has eliminated the descriptive variables which were considered an important feature and utilize the other variables as input. Due to this, the classification accuracy became low. Firefly Classifier (Long et al., 2015) was proposed to predict the heart disease by incorporating the attributes that are filtered by rough sets concept. The fuzzy concept was used to filter the attributes even more. The clustering concept was applied before filtering the attributes for heart disease prediction. The result made an indication that it has very high true positive and zero false positives, which is unacceptable to proceed with the decision support system. Binary Classifier (de Menezes et al., 2017) was proposed to classify the risk level of coronary artery disease. It used the Binomial Boosting algorithm integrated with the maximum likelihood and the logistic regression models, to choose the better fit value towards prediction. Hence there arise mismatches between maximum likelihood and the logistic regression models leading towards providing inaccurate results. Decision Tree Classifier (Tayefi et al., 2017) was proposed to predict CAD. The evaluation was made by assessing the biomarkers of clinical data to establish the risk factors. The results showed that the classifier has poor performance towards classification accuracy.

Differential Evolution Algorithm based Feature Selection (Vivekanandan & Sriman Narayana Iyengar, 2017) was proposed to predict heart disease in an early stage. Initially pre-processing was done as a mandatory step for classification, and feature selection was performed for optimization. The optimization was used to increase the accuracy. The further fuzzy logic concept too was used to get more accuracy, but the result got tremendously increased false positive. Significant Feature Identification Method (Amin et al., 2019) was proposed to predict cardiovascular disease. It is an ensemble of naïve Bayes, decision trees, support vector machines, and neural network algorithms. During the classification process, misclassification gets arise between the classification algorithms and leads the result to low accuracy. Machine-Learning based Classifier (Wang et al., 2017) was proposed to segregate the noisy signals to give a possible way to build and maintain the vocabularies of specific diseases. As a initial stage information relevant to the features is studied, i.e., the occurrence frequency. Labeled datasets were used for the evaluation purpose, but the results came with an increased false-negative. Knowledge System based Classifier (Nilashi et al., 2017) was proposed to predict CAD by making use of clustering concept. Further, the noise removal concept was used to increase the prediction level. For the classification purpose, a classification and regression tree algorithm was used. The result indicates that the prediction cannot be used to treat the patients due to the false positive and false negative rates. Optimization (Ramkumar & Vadivel, 2020b, 2020a) started playing is significant role in different streams. In medical streams also it has registered its better performance towards enhancing the results.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

Enhanced Feature Selection Using Firefly Based Optimization

Assume C be a set of constant number of records (i.e., dataset) having D number of samples with P features. Intention of feature selection (FS) is to select c features from Q, where $c \leq P$ enhance the accuracy of classification via objective function $obf(\cdot)$. Shortly, the aim of FS is to find of features $M \subseteq Q$ having features with minimum count and it lead to enhanced performance.

This research work adopts binary way of encoding strategy to indicate the solution for FS:

$$M_a = \left(m_{a,1}, m_{a,2}, m_{a,3}, \dots, m_{a,P-1}, m_{a,P}\right), m_{a,b} \epsilon \{0,1\} \tag{1}$$

where $m_{a,b} = 1$ indicates that bth feature is chosen in ath feature subset M_a, $m_{a,b} = 0$ indicates unselect of the feature.
When coming to result, the issue of FS can be mathematically expressed as Eq.(2)

$$\begin{cases} \text{maximum } obf(M) \\ \text{subject to } M = (m_1, m_2, m_3, \dots, m_{P-1}, m_P), m_b \epsilon \{0,1\}, b = 1,2,3, \dots, P-1, P \\ P \geq |M| \geq 1 \end{cases} \tag{2}$$

where $|M|$ indicates the count of features present in M.

In firefly algorithm (FA), the value of brightness is utilized to evaluate the advantages of firefly and it involves the issues related to optimization. The value of brightness is inversely proportional to the resultant values of $obf(\cdot)$. Update regarding the position of firefly towards attractiveness is inversely proportional to the distance present between fireflies. When considering two fireflies, the firefly having minimum level of brightness follows the firefly having better brightness. Specifically, firefly is represented as $M_t = \left(m_{t,1}, m_{t,2}, m_{t,3}, \dots, m_{t,P-1}, m_{t,P}\right)$ and the current position are updated using Eq.(3)

$$m_{t,b} = m_{t,b} + B(m,1)\left(m_{s,b} - m_{t,b}\right) + a(rand - 0.5) \tag{3}$$

$$B(t,s) = B_0 e^{-yu_{t,s}^2} \tag{4}$$

$$r_{t,s} = \|M_t - M_s\| = \left(\sum_{a=1}^{P} \left(M_{t,a} - M_{s,a}\right)^2\right)^{0.5} \tag{5}$$

where M_s denotes the firefly M_t that attracts other firefly more and it can be of any firefly in the swarm having increased brightness. The coefficient of light absorption y is applied in controlling the intensity of light. $u_{t,s}$ represents the distant that is present in between M_t and M_s. $B(t,s)$ represents the M_t level of attract to M_s. B_o represents the level of attractiveness at $u_{t,s} = 0$. The parameter a denotes the random parameter and $rand$ indicates a function that selected a random value in the set $\{0,1\}$

To handle FS issues in data mining, binary classification is maintained. It transforms two terms into a probability vector in Eq. (3) with the function related to sigmoid.

$$m_{t,s} = \begin{cases} 1, & (1 + e^{Vt,b})^{-1} > rand \\ 0, otherwise \end{cases} \tag{6}$$

$$V_{t,s} = B(m,s)\left(m_{t,b} - m_{s,b}\right) + a(rand - 0.5) \tag{7}$$

During the later stage, below function is utilized in Eq.(4) to enhance the accuracy of binary classification.

$$tah\left(V_{t,b}\right) = \frac{e^{2\|V_{t,b}\|} + 1}{e^{2\|V_{t,b}\|} - 1} \tag{8}$$

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

To overcome the issues in optimization towards measuring the attractiveness of firefly, this research work defines a new operator termed as rebound-expenditure. It is dependent on attractiveness of firefly from another one is computed depending on rebound-expenditure.

Let M_t indicate a firefly and M_s indicate an another firefly where the brightness of M_s is better than M_k. The rebound is received by M_t from M_s as $rebound(t,s)$ and it is describes as the difference related to two fireflies with objective space.

$$rebound(t,s) = \frac{obf(M_t) - obf(M_s)}{obf_{max} - obf(M_s)} \tag{9}$$

where obf_{max} is the maximum value present in objective function between all fireflies.
The expenditure paid by M_t while learning from M_s is represented as $expenditure(t,s)$ is described as the difference between two fireflies present in the search space.

$$expenditure(t,s) = \sum_{a=1}^{P} |m_{t,b} - m_{s,b}| \tag{10}$$

Commonly, firefly intends to observe from object that attempts gain more with minimum expenditure. From this inspiration, the attractiveness described in Eq.(4) is possible to modify as Eq.(11).

$$B(t,s) = \frac{1}{2} \times ((expenditure(t,s) + 1)^{-1} - rebound(t,s)) \tag{11}$$

Each firefly represents a binary vector in EFS-PO, the operator related to the movement are formulated using Eq.(3) won't support for FS related problems. Hence, this research work develops a new operator related to firefly movement depending on the rebound-expenditure and making an adaptive jump to update the firefly position based on Eq.(12) and Eq.(13).

$$m_{t,s} = \begin{cases} 1 - v_{t,s}, a > rand_1 \\ v_{t,s}, otherwise \end{cases} \tag{12}$$

$$m_{t,s} = \begin{cases} m_{s,b}, B(t,s) > rand_2 \\ m_{t,b}, otherwise \end{cases} \tag{13}$$

Dataset and Tool

To make an analysis about the proposed protocol against existing protocols, this research work has chosen benchmark datasets namely Cleveland dataset, Z-Alizadeh Sani dataset and Statlog dataset. Details regarding the count of attribute and record are provided in Table 1.

Table 1. Dataset Details

Dataset Name	Attributes	Records
Cleveland (Gupta et al., 2020)	75	303
Z-Alizadeh Sani (Arabasadi et al., 2017)	56	303
Statlog (Fitriyani et al., 2020)	13	270

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

Results and Discussion

Analysis of TP, TN, FP and FN

Figure 1, Figure 2 and Figure 3 shows the performance analysis of EFS-PO against existing algorithms namely MIFH (Gupta et al., 2020) and HNNGA (Arabasadi et al., 2017) for the performance metrics variables (i.e., TP, TN, FP, FN) towards the prediction of CAD. In Figure 1, Figure 2 and Figure 3, x-axis is marked with performance metrics variables and y-axis is marked with number of records in dataset. MIFH and HNNGA have lower performance than the proposed method EFS-PO. It makes a clear indication that EFS-PO has outperformed MIFH and HNNGA by selecting suitable features to predict CAD while evaluating with 3 different datasets namely Cleveland dataset, Z-Alizadeh Sani dataset and Statlog dataset. Average TP value of MIFH, HNNGA and EFS-PO are 81, 98.667 and 127.333 respectively. Average TN value of MIFH, HNNGA and EFS-PO are 77.667, 84.667 and 97.333 respectively. Average FP value of MIFH, HNNGA and EFS-PO are 67, 58.667 and 33.333 respectively. Average FN value of MIFH, HNNGA and EFS-PO are 66.333, 50 and 34 respectively. The reason for EFS-PO better performance is selecting the features in an optimized manner.

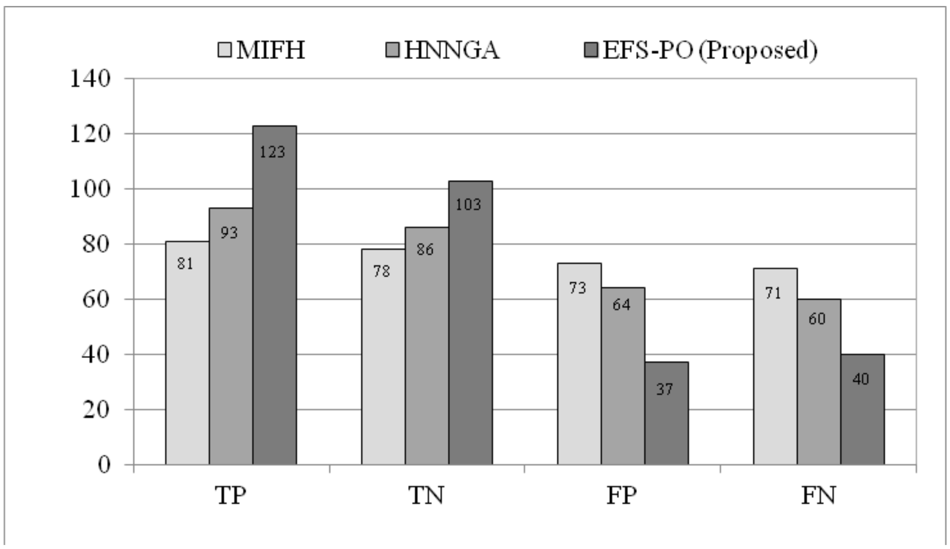

Figure 1. EFS-PO Vs TP,TN,FP,FN - Cleveland Dataset

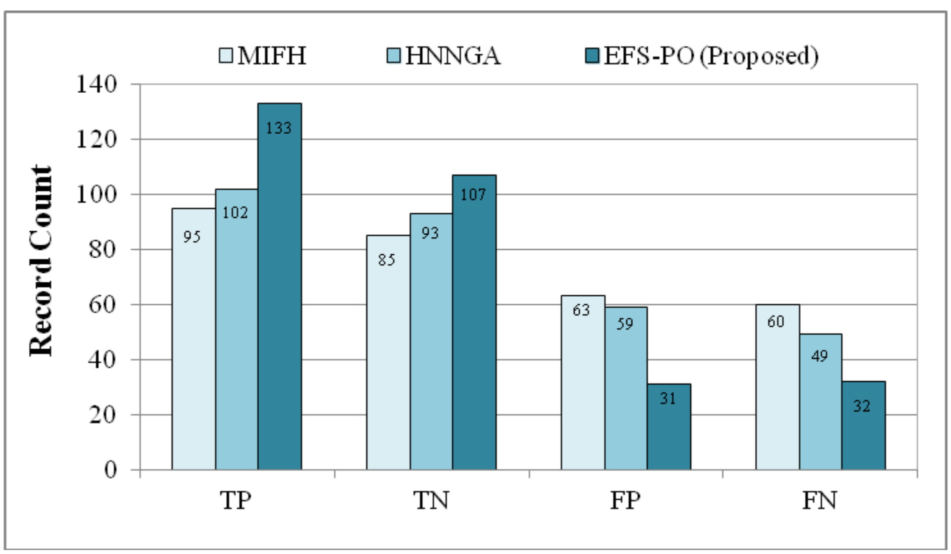

Figure 2. EFS-PO Vs TP,TN,FP,FN - Z-Alizadeh Sani Dataset

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

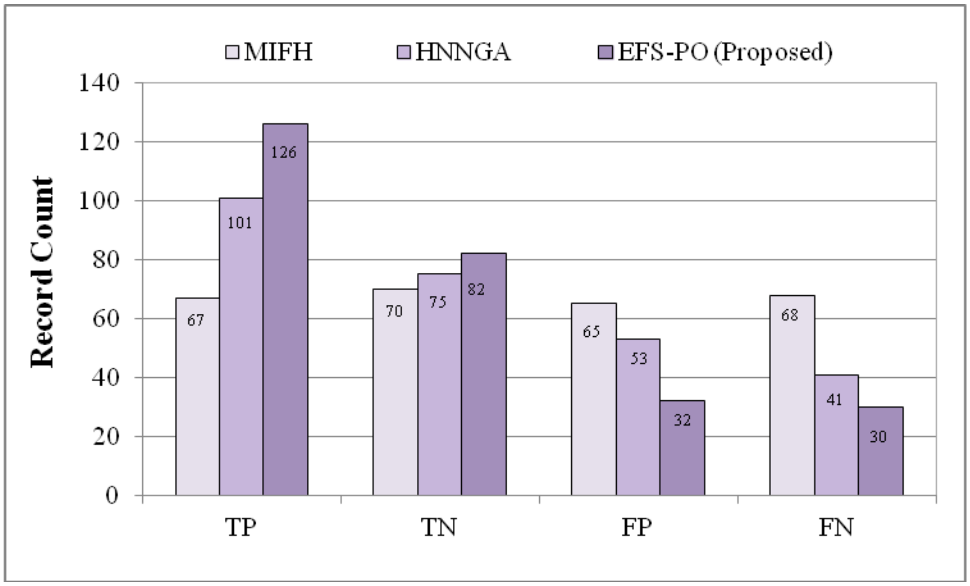

Figure 3. EFS-PO Vs TP,TN,FP,FN - Statlog Dataset

Analysis of Sensitivity, Specificity and Accuracy

Figure 4, Figure 5 and Figure 6 shows the performance analysis of EFS-PO against existing algorithms namely MIFH (Gupta et al., 2020) and HNNGA (Arabasadi et al., 2017) for the performance metrics sensitivity, specificity and accuracy towards the prediction of CAD. In Figure 4, Figure 5 and Figure 6, x-axis is marked with performance metrics and y-axis is marked with results in percentage. MIFH and HNNGA have lower performance than the proposed method EFS-PO. It makes a clear indication that EFS-PO has outperformed MIFH and HNNGA by selecting suitable features to predict CAD while evaluating with 3 different datasets namely Cleveland dataset, Z-Alizadeh Sani dataset and Statlog dataset. Average Sensitivity of MIFH, HNNGA and EFS-PO are 54.736%, 66.487% and 78.945%. Average Specificity of MIFH, HNNGA and EFS-PO are 53.647%, 59.037% and 74.346%. Average Accuracy of MIFH, HNNGA and EFS-PO are 54.207%, 62.873% and 76.944%. The significant reason for EFS-PO better results is because of maintaining rebound-expenditure way of performing the classification.

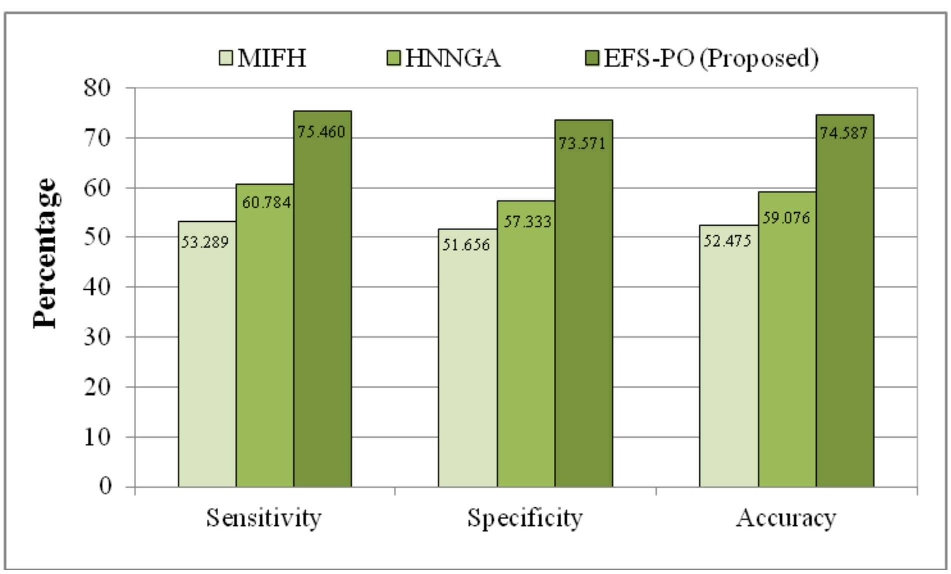

Figure 4. EFS-PO Vs Sensitivity, Specificity, Accuracy - Cleveland Dataset

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

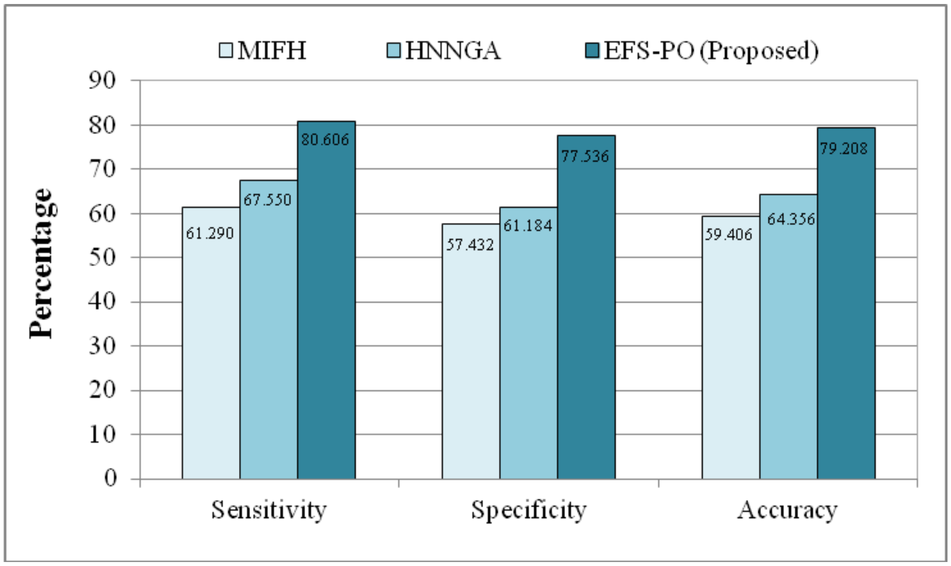

Figure 5. EFS-PO Vs Sensitivity, Specificity, Accuracy - Z-Alizadeh Sani Dataset

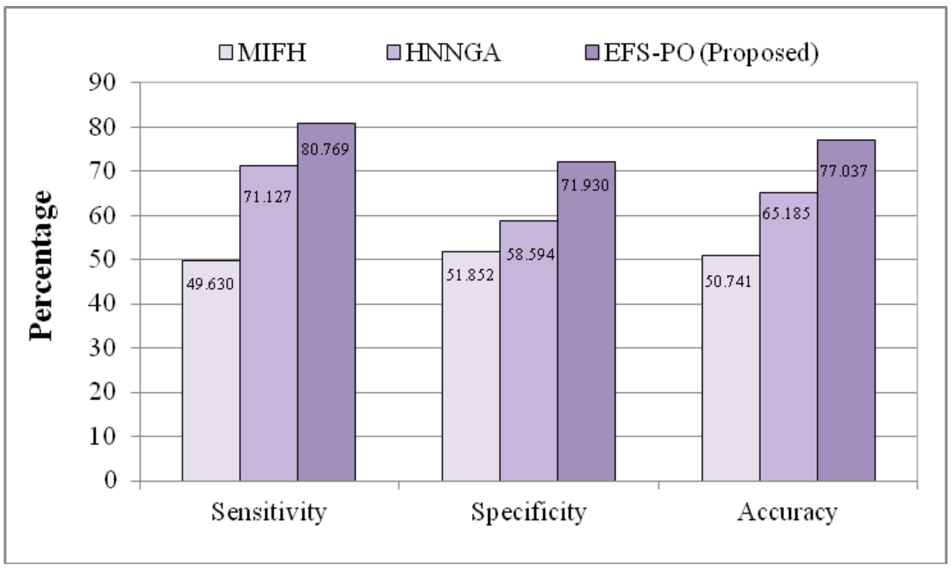

Figure 6. EFS-PO Vs Sensitivity, Specificity, Accuracy - Statlog Dataset

Analysis of Precision, Recall and F-Measure

Figure 7, Figure 8 and Figure 9 shows the performance analysis of EFS-PO against existing algorithms namely MIFH (Gupta et al., 2020) and HNNGA (Arabasadi et al., 2017) for the performance metrics precision, recall and f-measure towards the prediction of CAD. In Figure 7, Figure 8 and Figure 9, x-axis is marked with performance metrics and y-axis is marked with results in percentage. MIFH and HNNGA have lower performance than the proposed method EFS-PO. It makes a evident that EFS-PO has better performance than MIFH and HNNGA by selecting suitable features to predict CAD while evaluating with 3 different datasets namely Cleveland dataset, Z-Alizadeh Sani dataset and Statlog dataset. Average Precision of MIFH, HNNGA and EFS-PO are 54.494%, 62.725% and 79.240%. Average Recall of MIFH, HNNGA and EFS-PO are 54.736%, 66.487% and 78.945%. Average F-Measure of MIFH, HNNGA and EFS-PO are 54.610%, 64.543% and 79.089%. The significant reason for EFS-PO better results is selecting the better features in an optimized manner and then performing the classification.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

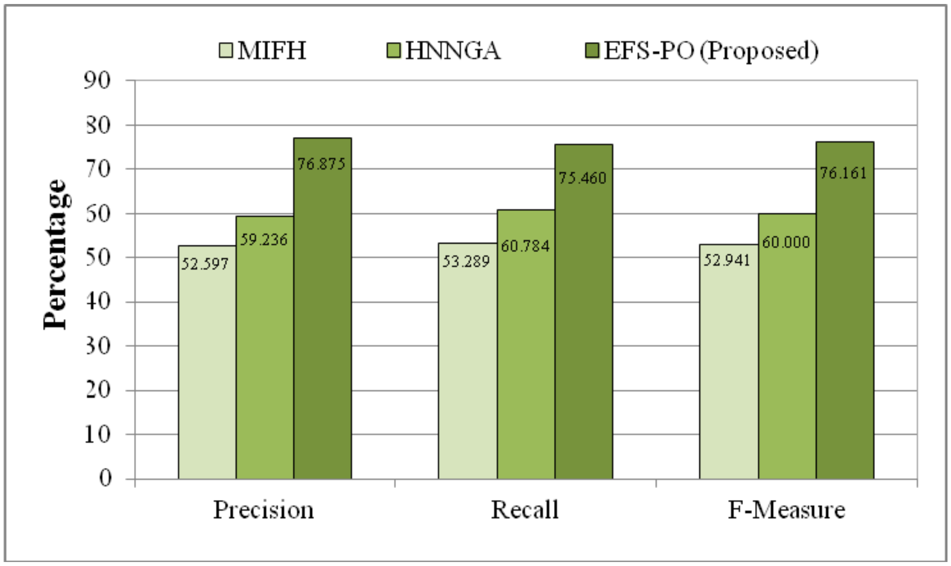

Figure 7. EFS-PO Vs Precision, Recall, F-Measure - Cleveland Dataset

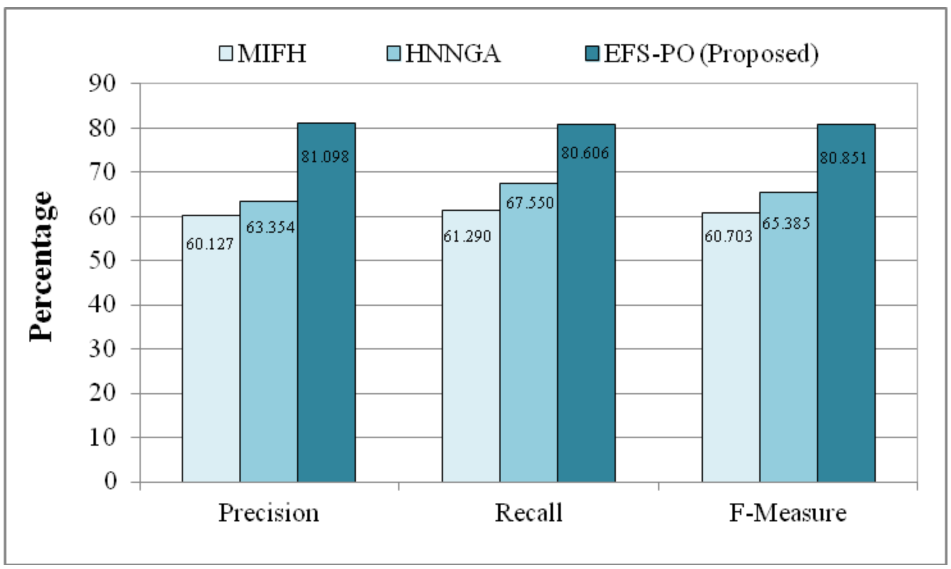

Figure 8. EFS-PO Vs Precision, Recall, F-Measure - Z-Alizadeh Sani Dataset

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4238 - 4248
Received 25 April 2021; Accepted 08 May 2021.

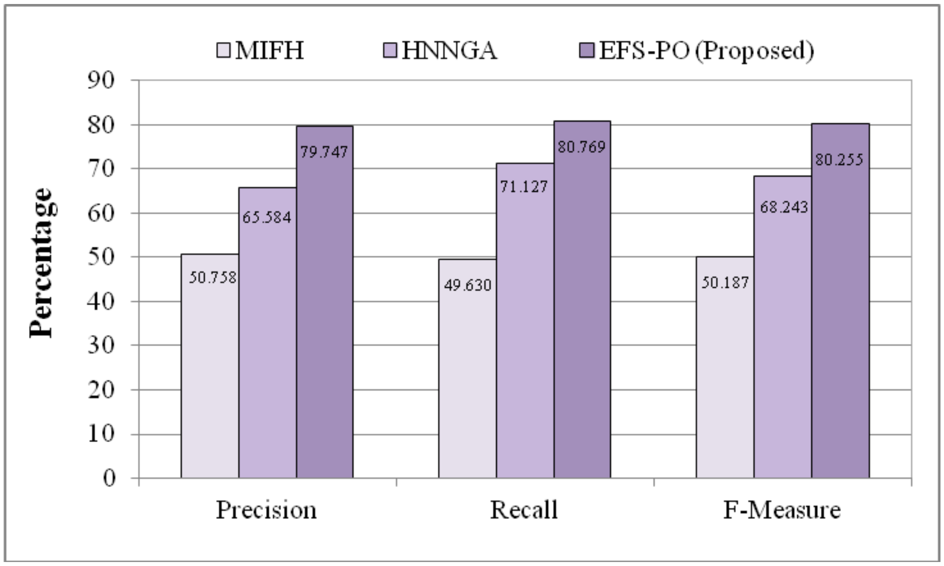

Figure 9 EFS-PO Vs Precision, Recall, F-Measure - Statlog Dataset

Conclusion

This paper has proposed an optimization based feature selection strategy towards the prediction of coronary artery disease (CAD). It adopts the natural characteristics of firefly to select the most appropriate feature for predicting CAD. Performance of the currently available algorithms degrades when the feature count increases. The proposed EFS-PO has better performance with the dataset having higher and lower number of features. Performance of EFS-PO has been tested with datasets namely Cleveland Dataset, Z-Alizadeh Sani Dataset and Statlog Dataset. EFS-PO has obtained better results because of selecting the appropriate features using bio-inspired concept.

References

[1] Abdar, M., Książek, W., Acharya, U. R., Tan, R. S., Makarenkov, V., & Pławiak, P. (2019). A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, *179*, 104992. https://doi.org/10.1016/j.cmpb.2019.104992

[2] Alizadehsani, R., Abdar, M., Roshanzamir, M., Khosravi, A., Kebria, P. M., Khozeimeh, F., Nahavandi, S., Sarrafzadegan, N., & Acharya, U. R. (2019). Machine learning-based coronary artery disease diagnosis: A comprehensive review. In *Computers in Biology and Medicine* (Vol. 111, p. 103346). Elsevier Ltd. https://doi.org/10.1016/j.compbiomed.2019.103346

[3] Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics*, *36*, 82–93. https://doi.org/10.1016/j.tele.2018.11.007

[4] Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., Moosaei, H., & Yarifard, A. A. (2017). Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Computer Methods and Programs in Biomedicine*, *141*, 19–26. https://doi.org/10.1016/j.cmpb.2017.01.004

[5] de Carvalho, H. H., Moreno, R. L., Pimenta, T. C., Crepaldi, P. C., & Cintra, E. (2013). A heart disease recognition embedded system with fuzzy cluster algorithm. *Computer Methods and Programs in Biomedicine*, *110*(3), 447–454. https://doi.org/10.1016/j.cmpb.2013.01.005

[6] de Menezes, F. S., Liska, G. R., Cirillo, M. A., & Vivanco, M. J. F. (2017). Data classification with binary response through the Boosting algorithm and logistic regression. *Expert Systems with Applications*, *69*, 62–73. https://doi.org/10.1016/j.eswa.2016.08.014

[7] Esmaelian, M., Shahmoradi, H., & Vali, M. (2016). A novel classification method: A hybrid approach based on extension of the UTADIS with polynomial and PSO-GA algorithm. *Applied Soft Computing Journal*, *49*, 56–70. https://doi.org/10.1016/j.asoc.2016.07.017

[8] Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2020). HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System. *IEEE Access*, *8*, 133034–133050.

https://doi.org/10.1109/ACCESS.2020.3010511

[9] Głowacki, J., Krysiński, M., Czaja-Ziółkowska, M., & Wasilewski, J. (2020). Machine Learning-based Algorithm Enables the Exclusion of Obstructive Coronary Artery Disease in the Patients Who Underwent Coronary Artery Calcium Scoring. *Academic Radiology*, *27*(10), 1416–1421. https://doi.org/10.1016/j.acra.2019.11.016

[10] Gupta, A., Kumar, R., Singh Arora, H., & Raman, B. (2020). MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis. *IEEE Access*, *8*, 14659–14674. https://doi.org/10.1109/ACCESS.2019.2962755

[11] Han, D., Liu, J., Sun, Z., Cui, Y., He, Y., & Yang, Z. (2020). Deep learning analysis in coronary computed tomographic angiography imaging for the assessment of patients with coronary artery stenosis. *Computer Methods and Programs in Biomedicine*, *196*, 105651. https://doi.org/10.1016/j.cmpb.2020.105651

[12] JayaSree, M., & Koteswara Rao, L. (2020). Survey on - Identification of Coronary Artery Disease using Deep Learning. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2020.09.526

[13] Kawasaki, T., Kidoh, M., Kido, T., Sueta, D., Fujimoto, S., Kumamaru, K. K., Uetani, T., Tanabe, Y., Ueda, T., Sakabe, D., Oda, S., Yamashiro, T., Tsujita, K., Kato, S., Yuki, H., & Utsunomiya, D. (2020). Evaluation of Significant Coronary Artery Disease Based on CT Fractional Flow Reserve and Plaque Characteristics Using Random Forest Analysis in Machine Learning. *Academic Radiology*, *27*(12), 1700–1708. https://doi.org/10.1016/j.acra.2019.12.013

[13] Lee, J. G., Ko, J., Hae, H., Kang, S. J., Kang, D. Y., Lee, P. H., Ahn, J. M., Park, D. W., Lee, S. W., Kim, Y. H., Lee, C. W., Park, S. W., & Park, S. J. (2020). Intravascular ultrasound-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. *Atherosclerosis*, *292*, 171–177. https://doi.org/10.1016/j.atherosclerosis.2019.10.022

[14] Li, H., Wang, X., Liu, C., Zeng, Q., Zheng, Y., Chu, X., Yao, L., Wang, J., Jiao, Y., & Karmakar, C. (2020). A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection. *Computers in Biology and Medicine*, *120*, 103733. https://doi.org/10.1016/j.compbiomed.2020.103733

[15] Long, N. C., Meesad, P., & Unger, H. (2015). A highly accurate firefly based algorithm for heart disease prediction. *Expert Systems with Applications*, *42*(21), 8221–8231. https://doi.org/10.1016/j.eswa.2015.06.024

[16] Mishra, R. K., Tison, G. H., Fang, Q., Scherzer, R., Whooley, M. A., & Schiller, N. B. (2020). Association of Machine Learning–Derived Phenogroupings of Echocardiographic Variables with Heart Failure in Stable Coronary Artery Disease: The Heart and Soul Study. *Journal of the American Society of Echocardiography*, *33*(3), 322-331.e1. https://doi.org/10.1016/j.echo.2019.09.010

[20] Nilashi, M., Ibrahim, O. bin, Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. *Computers and Chemical Engineering*, *106*, 212–223. https://doi.org/10.1016/j.compchemeng.2017.06.011

[21] Ramkumar, J., & Vadivel, R. (2020a). Improved Wolf prey inspired protocol for routing in cognitive radio Ad Hoc networks. *International Journal of Computer Networks and Applications*, *7*(5), 126–136. https://doi.org/10.22247/ijcna/2020/202977

[22] Ramkumar, J., & Vadivel, R. (2020b). Meticulous elephant herding optimization based protocol for detecting intrusions in cognitive radio ad hoc networks. *International Journal of Emerging Trends in Engineering Research*, *8*(8). https://doi.org/10.30534/ijeter/2020/82882020

[23] Shao, Y. E., Hou, C. D., & Chiu, C. C. (2014). Hybrid intelligent modeling schemes for heart disease classification. *Applied Soft Computing Journal*, *14*(PART A), 47–52. https://doi.org/10.1016/j.asoc.2013.09.020

[24] Tayefi, M., Tajfard, M., Saffar, S., Hanachi, P., Amirabadizadeh, A. R., Esmaeily, H., Taghipour, A., Ferns, G. A., Moohebati, M., & Ghayour-Mobarhan, M. (2017). hs-CRP is strongly associated with coronary heart disease (CHD): A data mining approach using decision tree algorithm. *Computer Methods and Programs in Biomedicine*, *141*, 105–109. https://doi.org/10.1016/j.cmpb.2017.02.001

[25] Vafaie, M. H., Ataei, M., & Koofigar, H. R. (2014). Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. *Biomedical Signal Processing and Control*, *14*, 291–296. https://doi.org/10.1016/j.bspc.2014.08.010

[26] Vivekanandan, T., & Sriman Narayana Iyengar, N. C. (2017). Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Computers in Biology and Medicine*, *90*, 125–136. https://doi.org/10.1016/j.compbiomed.2017.09.011

[27] Wang, L., Haug, P. J., & Del Fiol, G. (2017). Using classification models for the generation of disease-specific medications from biomedical literature and clinical data repository. *Journal of Biomedical Informatics*, *69*, 259–266. https://doi.org/10.1016/j.jbi.2017.04.014

[28] Yang, S., Koo, B.-K., Hoshino, M., Lee, J. M., Murai, T., Park, J., Zhang, J., Hwang, D., Shin, E.-S., Doh, J.-H., Nam, C.-W., Wang, J., Chen, S., Tanaka, N., Matsuo, H., Akasaka, T., Choi, G., Petersen, K., Chang, H.-J., … Narula, J. (2020). CT Angiographic and Plaque Predictors of Functionally Significant Coronary Disease and Outcome Using Machine Learning. *JACC: Cardiovascular Imaging*. https://doi.org/10.1016/j.jcmg.2020.08.025

[29] Zhang, Z., Dong, J., Luo, X., Choi, K. S., & Wu, X. (2014). Heartbeat classification using disease-specific feature selection. *Computers in Biology and Medicine*, *46*(1), 79–89. https://doi.org/10.1016/j.compbiomed.2013.11.019