Performance Improvement of Decision Trees for Diagnosis of Coronary Artery Disease Using Multi Filtering Approach

Moloud Abdar
Département d'informatique
Université du Québec à Montréal
Montréal, Canada
m.abdar1987@gmail.com

Elham Nasarian
Department of Industrial Engineering
Islamic Azad University, Najafabad Branch, Najafabad, Iran
elham.ansarian88@gmail.com

Xujuan Zhou
School of Management and Enterprise
University of Southern Queensland
Springfield, Australia
xujuan.zhou@usq.edu.au

Ghazal Bargshady
School of Management and Enterprise
University of Southern Queensland
Springfield, Australia
ghazal.bargshady@usq.edu.au

Vivi Nur Wijayaningrum
Faculty of Computer Science
Brawijaya University
Malang, Indonesia
vivinurw@gmail.com

Sadiq Hussain
System Administrator
Dibrugarh University
Dibrugarh, India
sadiq@dibru.ac.in

***Abstract*—The heart is one of the strongest muscular organs in the human body. Every year, this disease can kill many people in the world. Coronary artery disease (CAD) is named as the most common type of heart disease. Four well-known decision trees (DTs) are applied on the Z-Alizadeh Sani CAD dataset, which consists of J48, BF tree, REP tree, and NB tree. A multi filtering approach, named MFA, was used to modify the weight of attributes to improve the performance of DTs in this study. The model was applied on three main coronary arteries including the Left Anterior Descending (LAD), Left Circumflex (LCX), and Right Coronary Artery (RCA). The obtained results show that data balancing has a valuable impact on the performance of DTs. The comparison results show that this study provides the best results applied on the Z-Alizadeh Sani dataset compared to previous studies. The proposed MFA could improve the performance of the classic DTs algorithms significantly, with the highest accuracies obtained by NB tree for LAD, LCX, and RCA are 94.90%, 92.97% and 93.43%, respectively.**

Keywords-*heart disease; coronary artery disease; data mining; machine learning, classification*

I. Introduction

Automatic computer-based Diagnosis has been developed in order to decrease patient mortality and has been applied in various healthcare issues (e.g. different diseases and cancers). [1]. The ADDI systems consist of different data mining (DM) and machine learning (ML) algorithms which are faster, more accurate, and sometimes cheaper than classical diagnosis methods. The effectiveness and authority of ML/DM methods is derived from a capability to use proportionate techniques to find hidden patterns and as a result construct different models from raw data [2].

Nowadays, there are huge amount of raw data in different fields. ML/DM methods investigate data to find patterns. The extracted patterns reveal useful information from raw data that cannot be seen from such raw data. Generally speaking, ML techniques can be grouped into three major categories: supervised learning, unsupervised learning and reinforcement learning [3]. Each of these categories includes different algorithms which were applied on different subjects such as diseases and cancers [4, 5, 6, 7, 8, 9], security and networks [10, 11], transportation [12, 13], economics [14], energy [15,16,17], etc. Some of most well-known ML/DM algorithms are: Artificial neural network (ANNs), Support vector machines (SVMs), Decision trees (DTs), k-nearest neighbours (KNN), Naïve Bayes, K-means and C-means. As discussed, ML/DM techniques have been successfully used for diagnosing of different diseases and cancers. Coronary artery disease (CAD), a common type of heart disease, is one of the most important diseases. Heart disease is one of leading causes of death worldwide [18]. According to a report published by the World Health Organization (WHO), CAD was the cause of death for an estimated 17.7 million people in 2015. This shows that CAD has a grave impact on public health which means to decrease mortality rates of CAD by using new methods are clearly in the public interest.

There are three main coronary arteries are as follows: the Left Anterior Descending (LAD), Left Circumflex (LCX), and Right Coronary Artery (RCA) [19]. It should be noted that CAD occurs when one of these coronary arteries is blocked. Therefore, this paper investigates each of these coronary arteries using ML/DM techniques. This research tries to introduce an accurate and efficient technique for diagnosing CAD by using these coronary arteries (LAD, LCX, and RCA).

We investigate three main coronary arteries (LAD, LCX, and RCA) using classical DTs. The DTs algorithms are: J48, BF tree, REP tree, and NB tree. In this paper, a new multi filtering approach (MFA) is proposed as our balancing technique that includes both supervised and unsupervised techniques. The methodology includes four main filtering techniques and the MFA can modify the weights of attributes. The proposed methodology can increase the capability of prediction of DTs algorithms. To the best of our knowledge, the obtained outcomes are best results on the Z-Alizadeh Sani coronary artery disease dataset.

The rest of study is structured as follows. In Section II, some related works are discussed. Section III discusses about the material and method used. The proposed methodology is described in Section IV. The obtained outcomes are presented in Section V. Then, we conclude the research in Section VI.

II. RELATED WORK

In this section, we reviewed some related studies that have used ML/DM techniques to detect CAD. In the past few years, various ML/DM algorithms have been applied on CAD datasets successfully; some will be presented here briefly.

Abdar [5] compared the performances of four different decision trees on a heart disease dataset. The results demonstrated that C5.0 showed better performance than other algorithms with 85.33%. In [3], the author reviewed and provided a workflow for the diagnosis of coronary artery disease. Alizadehsani et al. [19] selected most important features among the Z-Alizadeh Sani coronary artery disease (CAD) dataset using Information Gain approach. Then, SMO classifier was applied on the dataset and reached the highest accuracy with 94.08%. Alizadehsani et al. [20] applied a bagging ensemble technique with C4.5 algorithm. The study investigated the performance of methods on the dataset for each LAD, LCX, and RCA coronary arteries separately. The highest accuracies obtained by bagging C4.5 was 79.54% (LAD), 61.46% (LCX), and 68.96 (RCA). Alizadehsani et al. [21] improved the accuracy of the methods applied on the Z-Alizadeh Sani dataset. To this end, SVM algorithm was chosen as the base classifier. They combined information gain for all arteries. The proposed approach indicated significant performance with high accuracy for each LAD, LCX, and RCA coronary arteries which are 86.14%, 83.17%, and 83.50% for LAD, LCX, and RCA, respectively. Stuckey et al. [22] analyzed the signals obtained from 606 persons at rest just frontward to angiography to investigate CAD. Their report indicated that the machine-learned algorithm showed a high sensitivity of 92% (95% CI: 74%-100%). Electrocardiogram (ECG) is an important method that can be used for detection of CAD [23, 24, 25]. Giri et al. [23] used Discrete Wavelet Transform (DWT) approach in order to decompose the heart rate signals into frequency sub-bands. The best accuracy using the proposed methodology and Gaussian Mixture Model (GMM) classifier was 96.8%. In addition, Patidar et al. [24] introduced tunable-Q wavelet transform (TQWT) using features extraction from heart rate signals for CAD diagnosis. The least squares support vector machine (LS-SVM) and TQWT showed the highest accuracy (99.72%).

Convolutional neural network (CNN) algorithm was applied in [25], which investigated two and five seconds durations of ECG signal segments. The accuracy of applied CNN for normal and abnormal ECG data of two seconds was 94.95% while the accuracy of five seconds was 95.11%. Alkeshuosh et al. [26] applied Particle swarm optimization (PSO) to create the best rules among heart disease dataset. The performance of proposed PSO was then compared with C4.5 algorithm. Their findings indicated that the average accuracy of PSO was 87% while related accuracy of C4.5 was 63%.

III. PROPOSED METHODOLOGY

In this study, four decision trees were selected to investigate CAD dataset. Four well-known algorithms are J48, BF tree, REP tree, and NB tree methods. Moreover, both supervised and unsupervised filtering approaches were chosen to change the weight of attributes. It should be noted that resample (both supervised and unsupervised) and class balancer (supervised) techniques were applied. This study applies multi filtering approach which several simple techniques can be applied at the same time. The proposed methodology is illustrated in Fig. 1.

As Fig. 1 shows, a Multi Filtering Approach (MFA) is used to change the weights of features. The proposed approach includes three supervised filtering techniques (2 resample and one class balancer techniques) and one unsupervised technique (resample). The proposed MFA methodology applies three supervised filtering techniques and then one unsupervised filtering technique. In other words, MFA allows the filtering techniques to change the weights of attributes. A 10-fold cross validation was used. This means that in each training-testing set, 90% of data will be used for training and the rest of data (10%) will be used for testing. The WEKA 3.9.1 was used to apply proposed methodology. The classical algorithms were implemented on a system with 3.40 GHz Intel Core i7 CPU and 8 GB RAM, with Windows 7 operating system.

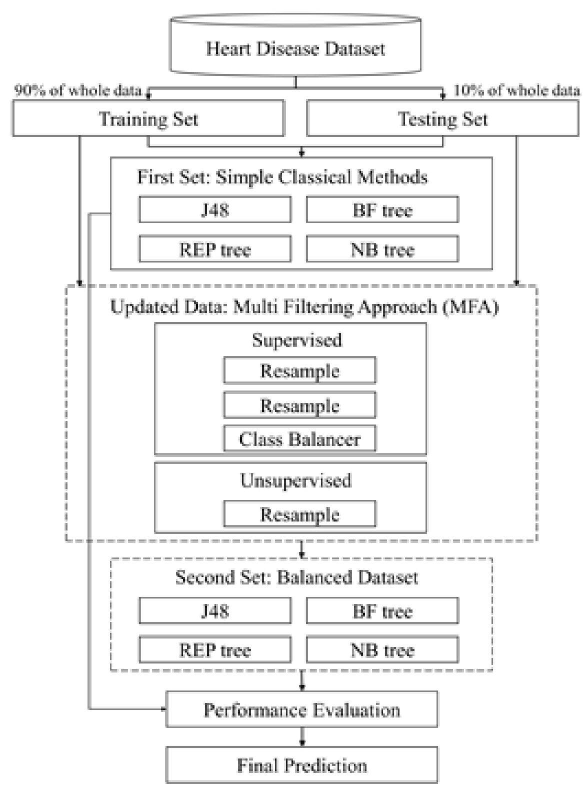

Figure 1. The proposed multi filtering approach for prediction of CAD

IV. EXPERIMENTS AND RESULTS

Current research uses the extension of the Z-Alizadeh Sani coronary artery disease dataset which can be downloaded from UCI Machine Learning Repository [27]. The dataset has three main target features: LAD, LCX, and RCA. The dataset includes 303 records with 55 features as input and also three features as output (LAD, LCX, and RCA: attributes for prediction). The dataset is categorized into four major types as follows: demographical features, symptoms and examination features, ECG features, and laboratory and echocardiography features which more information can be found about the dataset in [20][21][28]. It should be noted that the dataset has three prediction features including LAD, LCX, and RCA. Hence, we performed the application of ML/DM techniques on all of these features.

A. *First Experiment*

In the first experiment, simple classical DTs techniques were applied. The initial weights of all coronary arteries (LAD, LCX, and RCA) are shown in TABLE I.

As TABLE I shows in LAD category, "Stenotic" class in the major class (including more records) whereas in both LCX and RCA categories "Normal" class is the major class. Thus, the DTs are applied on the original dataset without any modification. The results are presented in TABLES II, III, and IV.

TABLE I. INETIAL WEIGHTS

Label		Count	Weight
LAD	Normal	126	126
	Stenotic	177	177
LCX	Normal	184	184
	Stenotic	119	119
RCA	Normal	189	189
	Stenotic	114	114

TABLE II. COMPARISON BETWEEN THE ACCURACY OF THE DTs FOR LAD

Method	Precision	Recall	F-measure	Accuracy
J48	65.00	72.20	72.40	72.27
BF tree	70.90	65.90	74.40	74.58
REP tree	**74.70**	**74.90**	**74.70**	**74.91**
NB tree	70.50	70.60	70.60	70.62

TABLE III. COMPARISON BETWEEN THE ACCURACY OF THE DTs FOR LCX

Method	Precision	Recall	F-measure	Accuracy
J48	55.30	55.80	55.50	55.77
BF tree	**61.90**	**62.00**	**62.00**	**62.04**
REP tree	58.10	60.10	58.10	60.06
NB tree	59.70	60.10	59.80	60.06

TABLE IV. COMPARISON BETWEEN THE ACCURACY OF THE DTs FOR RCA

Method	Precision	Recall	F-measure	Accuracy
J48	60.80	62.40	61.00	62.37
BF tree	63.00	64.40	63.10	64.35
REP tree	61.00	63.40	60.10	63.36
NB tree	**65.50**	**66.00**	**65.70**	**66.00**

The obtained results demonstrate that REP tree had the best performance for LAD with 74.91% whereas BF tree and NB tree had better performance for LCX and RCA with 62.04% and 66%, respectively.

B. *Second Experiment*

In the second experiment, we first apply multi filtering approach (MFA) technique in order to reweigh the features for all three coronary arteries (LAD, LCX, and RCA). The reweighted results are displayed in TABLE V. Then, the classical DTs methods which are used in the first experiment are applied and results are shown in TABLES VI, VII, and VIII.

As TABLE V shows the MFA modified the weights of attributes. These new weights are different to previous ones presented in TABLES I. As TABLE V presents, the weights are different with the real count of records in the dataset. This means that the proposed technique can give weights to attributes to decrease the impact of unbalancing of the original data. The results of proposed approach are presented in TABLES VI, VII, and VIII.

As TABLES VI, VII, and VIII show the proposed MFA could improve the performance of classic DTs algorithms significantly. According to TABLES VI, VII, and VIII, the highest accuracies for LAD, LCX, and RCA are 95.36%, 96.29%, and 96.14%, respectively. To compare the performance of DTs optimized by MFA is presented in the following section.

TABLE V. THE NEW WEIGHTS OBTAINED USING MFA.

Label		Count	Weight
LAD	Normal	107	128.655
	Stenotic	196	167.763
LCX	Normal	180	148.207
	Stenotic	123	156.592
RCA	Normal	187	149.897
	Stenotic	116	154.158

TABLE VI. COMPARISON BETWEEN THE PERFORMANCE OF DTs FOR LAD OPTIMIZED BY USING MFA

Method	Precision	Recall	F-measure	Accuracy
J48	92.60	92.40	92.40	92.36
BF tree	90.20	90.20	90.20	90.22
REP tree	86.00	85.80	85.90	85.83
NB tree	**95.00**	**94.90**	**94.90**	**94.90**

TABLE VII. COMPARISON BETWEEN THE PERFORMANCE OF DTs FOR LCX OPTIMIZED BY USING MFA

Method	Precision	Recall	F-measure	Accuracy
J48	88.90	88.90	88.90	88.89
BF tree	83.70	83.70	83.70	83.66
REP tree	83.60	83.60	83.60	83.56
NB tree	**93.10**	**93.00**	**93.00**	**92.97**

TABLE VIII. COMPARISON BETWEEN THE PERFORMANCE OF DTs FOR RCA OPTIMIZED BY USING MFA

Method	Precision	Recall	F-measure	Accuracy
J48	92.10	92.00	92.00	92.02
BF tree	89.30	89.20	89.20	89.22
REP tree	83.70	83.40	83.40	83.44
NB tree	**93.70**	**93.40**	**93.40**	**93.43**

V. DISCUSSION

In previous section, we showed the results of our proposed method. However, this section discusses the outcomes of proposed methodology and compares with other results in the literature and also some well-known algorithms to show the effectiveness of it. The improvement of DTs using MFA for LAD, LCX, and RCA are presented in Fig. 2, Fig. 3, and Fig. 4.

According to Fig. 2, Fig. 3, and Fig. 4, it can be observed that presentation of data plays a significant role in order to have better prediction results. For more clarity, the results should be compared with previous studies those applied their methods on the same dataset. Hence, TABLE IX indicates the comparison of our outcomes with prior results.

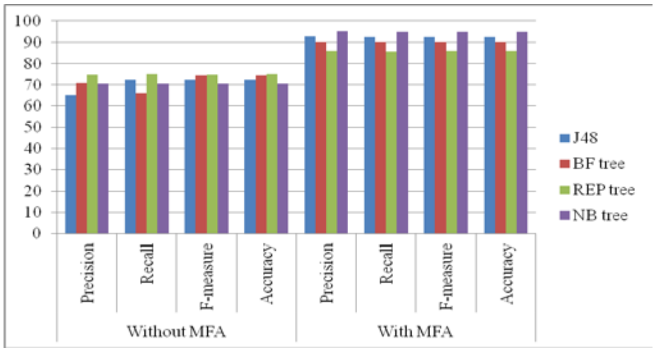

Figure 2. The improvement of DTs using MFA for LAD

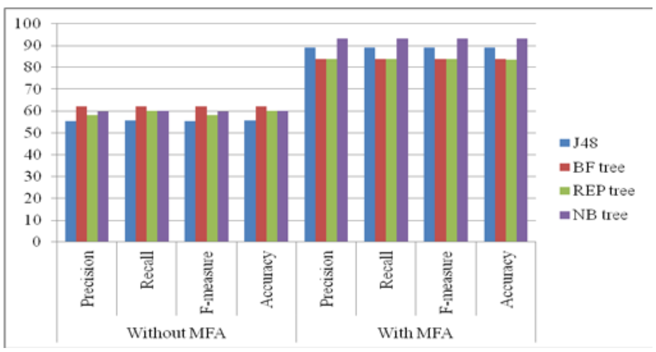

Figure 3. The improvement of DTs using MFA for LCX

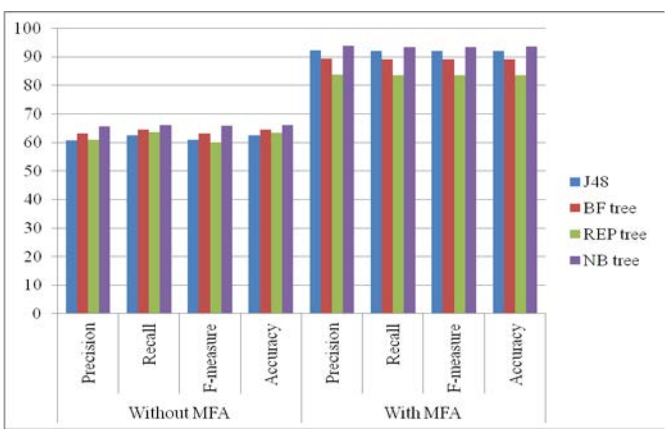

Figure 4.The improvement of DTs using MFA for RCA

TABLE IX. COMPARISON BETWEEN OUR RESULTS AND OTHER STUDIES

Study	Method	Accuracy
Alizadehsani et al. (2016) [20]	Combined information gain for all arteries-SVM	86.14 (LAD) 83.17 (LCX) 83.50 (RCA)
Alizadehsani et al. (2013) [21]	Bagging-C4.5	79.54 (LAD) 61.46 (LCX) 68.96 (RCA)
Alizadehsani et al. (2012) [28]	C4.5	74.20 ± 5.51% (LAD) 63.76 ± 9.73% (LCX) 68.33 ± 6.90% (RCA)
WEKA 3.9.1	C-SVC	58.41 (LAD) 60.72 (LCX) 62.37 (RCA)
WEKA 3.9.1	Naïve Bayes	71.94 (LAD) 64.35 (LCX) 65.34 (RCA)
WEKA 3.9.1	MLP	69.63 (LAD) 62.70 (LCX) 64.02 (RCA)
This study	MFA+ NB tree	94.90 (LAD) 92.97 (LCX) 93.43 (RCA)

As TABLE IX shows, the obtained outcomes in this research are much better than previous results. Moreover, we compared our outcomes with other three methods (C-SVC, Naïve Bayes, and MLP algorithms) using WEKA open source data mining tool. It should be noted that Alizadehsani et al. [20, 21, 28] applied feature selection in order to increase the prediction results; however, in this study we did not eliminate any feature. This means that the study used more features and such data is commonly available in the real world. Thus, we would argue that our results are more reliable and practicable. The number of techniques used for MFA in this study was chosen manually; while this number can be found using some evolutionary algorithms (EAs) such as genetic algorithm [29, 30, 31] and PSO [26]. Moreover, deep learning algorithm can be used with our proposed methodology [32].

VI. CONCLUSION

This research investigated the performance of four classical DTs algorithms including J48, BF tree, REP tree, and NB tree on the CAD dataset. Three major coronary arteries (LAD, LCX, and RCA) were selected. In other words, we applied the methods on each artery separately. Due to low performance of those methods, a new multi filtering approach (MFA) was proposed in order to improve the quality of the prediction. Our findings showed that the MFA can improve the prediction outcomes of DTs. The highest accuracies were obtained using NB tree and MFA approach for LAD, LCX, and RCA which were 94.90%, 92.97%, and 93.43%, respectively. As a conclusion, this study introduced a new approach that can outperform previous algorithms having much better results on CAD dataset.

REFERENCES

[1] T. Y., Satyanarayana, D., Prasad, M. G., & K. D., Dhruve, " Melanoma is Skin Deep" A 3D reconstruction technique for

computerized dermoscopic skin lesion classification, 5, IEEE Journal of Translational Engineering in Health and Medicine, 5, 1-17, 2017.

[2] I., Kavakiotis, O., Tsave, A., Salifoglou, N., Maglaveras, I., Vlahavas, & I., Chouvarda, " Machine learning and data mining methods in diabetes research", Computational and structural biotechnology journal, 15, 104-116, 2017.

[3] F. A., Cárdenas-López, L., Lamata, J. C., Retamal, & E., Solano, " Multiqubit and multilevel quantum reinforcement learning with quantum technologies", PloS one, 13(7), e0200455, 2018.

[4] M, Abdar, " Using Decision Trees in Data Mining for Predicting Factors Influencing of Heart Disease", Carpathian Journal of Electronic & Computer Engineering, 8(2), 2015.

[5] M., Hassoon, M. S., Kouhi, M., Zomorodi-Moghadam, & M, Abdar, " Rule Optimization of Boosted C5. 0 Classification Using Genetic Algorithm for Liver disease Prediction. In Computer and Applications (ICCA)", International Conference on (pp. 299-305), IEEE, 2017.

[6] M., Abdar, Zomorodi-Moghadam, M., X., Zhou, R., Gururajan, X., Tao, P. D., Barua, & R., Gururajan, " A new nested ensemble technique for automated diagnosis of breast cancer", Pattern Recognition Letters, 2018.

[7] P., Pławiak, "Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system", Expert Systems with Applications, 92, 334-349, 2018.

[8] P., Pławiak, "Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals", Swarm and evolutionary computation, 39, 192-208, 2018.

[9] Ö., Yıldırım, P., Pławiak, R. S., Tan, & U. R., Acharya, " Arrhythmia detection using deep convolutional neural network with long duration ECG signals", Computers in biology and medicine, 102, 411-420, 2018.

[10] A. L., Buczak, & E., Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection", IEEE Communications Surveys & Tutorials, 18(2), 1153-1176, 2016.

[11] R., Sommer, & V., Paxson, " Outside the closed world: On using machine learning for network intrusion detection", IEEE Symposium on (pp. 305-316), 2010.

[12] H. I., Ashqar, M. H.,Almannaa, M., Elhenawy, H. A., Rakha, & L., House, " Smartphone Transportation Mode Recognition Using a Hierarchical Machine Learning Classifier and Pooled Features From Time and Frequency Domains", IEEE Transactions on Intelligent Transportation Systems, 2018.

[13] X., Ma, H., Yu, Y., Wang, & Y., Wang, " Large-scale transportation network congestion evolution prediction using deep learning theory", PloS one, 10(3), e0119044, 2015.

[14] M., Abdar, & N. Y., Yen, " Understanding regional characteristics through crowd preference and confidence mining in P2P accommodation rental service", Library Hi Tech, 35(4), 521-541, 2017.

[15] M., Fesanghary, S., Asadi, & Z. W., Geem, " Design of low-emission and energy-efficient residential buildings using a multi-objective optimization algorithm", Building and environment, 49, 245-250, 2012.

[16] M., Nazari-Heris, A. F., Babaei, B., Mohammadi-Ivatloo, & S., Asadi, "Improved harmony search algorithm for the solution of non-linear non-convex short-term hydrothermal scheduling". Energy, 151, 226-237, 2018.

[17] S., Asadi, " A multiobjective harmony-search algorithm for building life-cycle energy optimization". In Construction Research Congress 2014: Construction in a Global Network (pp. 484-493), 2014.

[18] T. Y., Wah, R., Gopal Raj, & U., Iqbal, " A Review and Workflow. Cardiology research and practice", Automated Diagnosis of Coronary Artery Disease, 2018.

[19] R., Alizadehsani, J., Habibi, M. J., Hosseini, H., Mashayekhi, R., Boghrati, A., Ghandeharioun,... & Z. A., Sani, " A data mining approach for diagnosis of coronary artery disease", Computer methods and programs in biomedicine, 111(1), 52-61, 2013.

[20] R., Alizadehsani, M. H., Zangooei, M. J., Hosseini, J., Habibi, A., Khosravi, M.,Roshanzamir, ... & S., Nahavandi, "Coronary artery disease detection using computational intelligence methods". Knowledge-Based Systems, 109, 187-197, 2016.

[21] R., Alizadehsani, J., Habibi, Z. A., Sani, H., Mashayekhi, R., Boghrati, A.,Ghandeharioun, ... & F., Alizadeh-Sani, " Diagnosing coronary artery disease via data mining algorithms by considering laboratory and echocardiography features". Research in cardiovascular medicine, 2(3), 133, 2013.

[22] T. D., Stuckey, R. S., Gammon, R., Goswami, J. P., Depta, J. A., Steuter, F. J., Meine III, ... & P., Grouchy, "Cardiac Phase Space Tomography: A novel method of assessing coronary artery disease utilizing machine learning". PloS one, 13(8), e0198603, 2018.

[23] D., Giri, U. R., Acharya, R. J.,Martis, S. V., Sree, T. C., Lim, T. A., VI, & J. S., Suri, "Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform", Knowledge-Based Systems, 37, 274-282, 2013.

[24] S., Patidar, R. B., Pachori, & U. R., Acharya, "Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals". Knowledge-Based Systems, 82, 1-10, 2015.

[25] U. R.,Acharya, H., Fujita, O. S., Lih, M., Adam, J. H., Tan, & C. K., Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network". Knowledge-Based Systems, 132, 62-71, 2017.

[26] A. H., Alkeshuosh, M. Z., Moghadam, I., Al Mansoori, & M., Abdar, "Using PSO Algorithm for Producing Best Rules in Diagnosis of Heart Disease". In Computer and Applications (ICCA), International Conference on (pp. 306-311). IEEE, 2017.

[27] Extention of Z-Alizadeh sani dataset Data Set, https://archive.ics.uci.edu/ml/datasets/extention+of+Z-Alizadeh+sani+dataset. [accessed on November 5, 2018].

[28] Alizadehsani, R., Habibi, J., Bahadorian, B., Mashayekhi, H., Ghandeharioun, A., Boghrati, R., & Sani, Z. A. "Diagnosis of coronary arteries stenosis using data mining". Journal of medical signals and sensors, 2(3), 153, 2012.

[29] V. N., Wijayaningrum, & W. F., Mahmudy, " Optimization of Ship's Route Scheduling Using Genetic Algorithm". Indonesian Journal of Electrical Engineering and Computer Science, 2(1), 180-186, 2016.

[30] Q., Kotimah, W. F., Mahmudy, & V. N., Wijayaningrum, "Optimization of Fuzzy Tsukamoto Membership Function using Genetic Algorithm to Determine the River Water". International Journal of Electrical and Computer Engineering (IJECE), 7(5), 2838-2846, 2017.

[31] V. N., Wijayaningrum, W. F., Mahmudy, & M. H., Natsir, " Optimization of Poultry Feed Composition Using Hybrid Adaptive Genetic Algorithm and Simulated Annealing". Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9(2-8), 183-187, 2017.

[32] J., Soar, G., Bargshady, X., Zhou, & F., Whittaker, (2018, July). Deep Learning Model for Detection of Pain Intensity from Facial Expression. In International Conference on Smart Homes and Health Telematics (pp. 249-254). Springer, Cham.

- **The below form will not be published, but it will help us to understand your paper better**

Authors' Background

Name	Email	Position (Prof , Assoc. Prof. etc.)	Research Field	Homepage URL
Moloud Abdar	m.abdar1987@gmail.com	PhD. student	Data Mining, Machine Learning, Evolutionary Computation, User Modeling & Understanding	
Elham Nasarian	elham.ansarian88@gmail.com	MSc	Data Mining, Machine Learning	
Dr. Xujuan Zhou	Xujuan.Zhou@usq.edu.au	Senior Lecturer	data mining and machine learning	
Ghazal Bargshady	ghazal.bargshady@usq.edu.au	PhD. student	Computer vision	
Vivi Nur Wijayaningrum	vivinurw@gmail.com	MSc	Data Mining, Evolutionary Computation,	
Dr. Sadiq Hussain	sadiq@dibru.ac.in	System Administrator, Dibrugarh University, India	Data Mining, Machine Learning	