Retraction

Retraction: Machine Learning based Classification for Heart Disease Identification (*J. Phys.: Conf. Ser* 1916 012174)

Published 17 December 2021

This article has been retracted by IOP Publishing following an allegation that this article contains text overlap from multiple unreferenced sources [1, 2]. IOP Publishing has investigated and agree the article constitutes plagiarism. IOP Publishing also expresses concern regarding a number of nonsensical phrases used in the article, which suggests the article may have been created at least partly by artificial intelligence or translation software. IOP Publishing also notes sections of this article were published in multiple other journals at a similar time [3, 4, 5, 6], by different author groups. These issues all bring the legitimacy of this article into serious doubt.

The authors have not responded to confirm whether they agree or disagree to this retraction.

IOP Publishing wishes to credit Problematic Paper Screener [7] for bringing some of these issues to our attention.

1. "Machine learning" Wikipedia, Wikimedia Foundation, https://en.wikipedia.org/wiki/Machine_learning
2. "Cardiovascular disease" Wikipedia, Wikimedia Foundation, https://en.wikipedia.org/wiki/Cardiovascular_disease
3. Sukanth, N. et al., 2021. Heart Disease Classification using Machine Learning Algorithm. *International Journal of Innovative Research in Computer and Communication Engineering*, 9(3), pp.1108-1114.
4. Karthikeyan, N. et al., 2021. Machine learning based classification models for heart disease prediction. *Journal of Physics: Conference Series*, 1916.
5. Priyadharshini, K. et al., 2021. Coronary Infarction Prediction Using Correlation Analysis aspects based on Parallel Distributed Processing Network. *Annals of the Romanian Society for Cell Biology*, 25(4), pp.2864-2869.
6. Vennila, V. et al., 2021. Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction. *Annals of the Romanian Society for Cell Biology*, 25(3), pp.8467-8474.
7. Cabanac G, Labbe C, Magazinov A, 2021, arXiv:2107.06751v1

Retraction published: 17 December 2021
Machine Learning based Classification for Heart Disease Identification

Siamala Devi S 1, Harini Karthika G 1, Deepika M 1

1Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore, Tamilnadu
s.siamaladevi@skct.edu.in

Abstract. Cardiac illness is one of the unpredictable infections and around the world numerous individuals experienced this sickness. On schedule and effective recognizable proof of coronary illness assumes a critical part in medical care, especially in the arena of cardiology. A productive and precise framework is proposed to finding coronary illness and the framework depends on AI procedures. Supervised learning algorithms such as Multi-Layer Perceptron (MLP), Multinomial Logistic Regression (MLR), Fuzzy Unordered Rule Induction Algorithm (FURIA) and C4.5 are then used to model CAD cases. This approach is tested on medical data that has 26 features and 335. MLR accomplishes most noteworthy expectation precision of 88.4 %. This methodology is benchmarked on Cleveland heart coronary illness information also. For this situation additionally, MLR, beats different methods. Projected hybridized model increases the exactness of arrangement calculations from 8.3 % to 11.4 % for the Cleaveland information. The proposed technique is, along these lines, a promising tool for finding CAD patients with improved forecast precision.

Keywords: Multi-Layer Perceptron, Multinomial Logistic Regression, Fuzzy Unordered Rule Induction Algorithm

1. INTRODUCTION

Data Mining is an insightful interaction planned to inspect information viewing for predictable instances as well as orderly influences among factors, and afterward to support the discoveries by relating the distinguished instances to novel subdivisions of information. A decisive aim of information extraction is expectation, and perceptive information mining is the widely recognized kind of information extraction and one that has the maximum immediate commercial applications. The cycle of information mining has three phases: (1) the underlying investigation, (2) model structure or sample distinguishing proof with approval/check, and (3) organization Simulated intelligence counts build a model reliant on model facts, known as planning data to make assumptions or results deprived of being unequivocally tweaked to do in that capacity. Compute based intelligence computations were utilized in a wide grouping of usages, for instance, email filtering and PC vision, where it is inconvenient or infeasible to make ordinary figuring to play out the necessary endeavors. A subset of AI is immovably related to computational bits of knowledge, which bases on making assumptions by means of PCs; yet not all AI is real learning. The investigation of accurate improvement permits on methods, theory and application spaces to the domain of AI. Data mining is a notable domain in which the assessment of every individual data is done by means
of autonomous learning. Simulated intelligence incorporates PCs discovering how they can perform tasks without being explicitly redone to do thusly. It incorporates PCs acquiring from data gave with the objective that they complete certain endeavors. For essential tasks distributed to PCs, it is possible to program counts prompting the machine how to execute all methods expected to handle the recent concern; on the part, no learning is required. For additional created tasks, it will in general be going after for a human to genuinely make the necessary counts. Before long, it can wind up being all the more remarkable to empower the machine to develop its own computation, instead of having human designers show each necessary advance. The request for AI uses various approaches to manage teach PCs to accomplish tasks where no totally tasteful computation is available. In circumstances where enormous amounts of potential answers exist, one technique is to name a segment of the correct answers as considerable. This would then have the option to be used as getting ready data for the PC to improve the calculation.

Disorders in the categories of cardiac disorders causes more harm that includes vein illness and other issues related to heart and veins. The articulation &quot; coronary disease&quot; is consistently used equally with the articulation &quot; cardiovascular contamination. &quot; Cardiovascular infection generally insinuates conditions that incorporate restricted or blocked veins. CAD is an issue with the veins that convey blood to the heart muscle. In the event that these veins get minuscule, or in the event that they become impeded, blood can't move through them regularly. Since less blood is provided to the heart muscle, the muscle cannot work at typical limit. The heart muscle can get debilitated and frail. Heart muscle can even bite the dust if blood stream stops. Obstructed courses in the heart are frequently brought about by smoking, elevated cholesterol, hypertension, diabetes, and acquired qualities from guardians. These issues harm the covering of the heart's veins and cause them to become limited or hindered completely. Congestive Heart Failure (CHF) this is a condition that implies that the heart isn’t siphoning at ordinary levels. Two basic causes are a powerless or wiped out heart muscle and strange heart valves. The valves may not let enough blood through in light of the fact that they are excessively limited. Or then again the valve may have hole and let blood stream in reverse (a misguided course) inside the heart. At the point when the heart valves don’t work ordinarily, the heart muscle needs to accomplish additional work and it can get drained.

Highlight determination is the strategy towards minimizing the amount of information while developing a prescient model. It is highly preferable to reduce the number of factors to diminish the cost related to execution which improves the demonstration of the adopted strategy. Certain component determination procedures contain assessing the assembly among each info variable and the objective variable using measurements and selecting those information influences that have the most grounded association with the objective variable. These policies can be rapid and successful, in spite of the fact that the conclusion of accurate measures that depends upon the evidence of both the data and yield factors. All things considered, it very well may be demanding for an AI expert to pick a suitable factual measure for a dataset when executing channel based element choice. In this post, you will find how to pick factual measures for channel based element choice with mathematical and clear cut information. There are two primary kinds of highlight choice procedures: administered and una ided, and managed techniques might be partitioned into covering, channel and natural. Channel based element determination techniques utilize factual measures to score the connection or reliance between input factors that can be sifted to pick the most pertinent highlights. Factual measures for include choice must be painstakingly picked dependent on the information sort of the info variable and the yield or reaction variable.

Cardio Vascular Disease (CVD) is a category of contaminations which embeds the heart or veins. CVD passions coronary channel contaminations (CAD, for instance, angina and myocardial confined
putrefaction (by and large known as a respiratory disappointment). Other CVDs consolidate stroke, cardiovascular breakdown, hypertensive coronary ailment, rheumatic coronary ailment, cardiomyopathy, sporadic heart rhythms, inborn coronary ailment, valvular coronary ailment, carditis, aortic aneurysms, periphery course affliction, thromboembolic disease, and venous blood vessel breakage. The secret segments move dependent upon the disease. Coronary vein disease, stroke, and periphery supply course ailment incorporate atherosclerosis. Rheumatic coronary sickness may follow untreated throat. It is surveyed that up to 90% of CVD may be preventable. Neutralization of CVD incorporates improving peril factors through: shrewd counting calories, work out, evading of tobacco smoke and confining alcohol utilization.

![Architecture Diagram](Retracted)

**Figure 1. Architecture Diagram**

Treating risk factors, for instance, hypertension, blood lipids and diabetes is similarly favourable. Treating individuals who have strep throat with anti-infection agents can diminish the danger of rheumatic coronary illness. The utilization of anti-inflammatory medicine in individuals, who are generally sound, is of indistinct advantage. The architecture diagram is represented in figure 1.
2. LITERATURE REVIEW

Despite the new innovative progressions in the clinical area, the quantity of death coming about because of HF sickness is as yet on the increment, and this could be credited to insufficient analytic apparatuses. As part exertion to build up a clinical choice help stage for the precise forecast of HF chances in sick people, a mixture approach dependent on NN and Fuzzy_AHP[1] was suggested in this investigation. Not at all like the customary ANN-based techniques that expected an indistinguishable danger commitment for all the credits of HF, the recently proposed strategy considered every one of the ascribes as an alternate danger supporter dependent on the clinical information on an accomplished clinical master. Utilizing the UCI clinical dataset with an aggregate of 297 examples got from potential HF sick people, our outcomes exhibited that the recently projected strategy had a superior execution regarding expectation exactness, cross-entropy, ROC plot, and execution plot measures in contrast with the customary ANN approach.

Another methodology is introduced in this paper to distinguish cardiovascular breakdown infection by utilizing improved SVM[2] in light of contrast streamlining strategy. The projected strategy separates among ordinary and strange ECG signals dependent on pulse term of PQRST waves prompting anticipating and identifying the HF from an ECG signal as it were. The SVM and the double SVM are utilized to progress execution of the calculation. In this calculation, sufficiency and time span are utilized as one zone in typical arrangement. The SVM and the double SVM are powerful in light of the fact that duality gives an alternate yet integral perspectives on the ECG signal. The planned thought depends by utilizing the non-parametric way to deal with deal with coach ISVM and its double to obtain dual strategies of SVM.

A tale ReliefF and Rough Set-(RFRS) established characterization framework is projected for coronary illness conclusion. The principle curiosity is the projected method: the blend of the ReliefF and RS strategies towards characterizes coronary illness issues in an effective and quick way. The RFRS grouping framework comprises of two subsystems: the RFRS highlight determination subsystem and the order subsystem. The Statlog (Heart) dataset from the UCI AI data set [3] was chosen to test the framework. The exploratory outcomes show that the reduce R2 (C1, C3, C7, C8, C11, C12, C13) accomplishes the most noteworthy grouping precision (92.59%) utilizing an outfit classifier with the C4.5 choice tree as the feeble student. The outcomes likewise show that the RFRS technique has better execution thought about than three normal classifiers as far as ACC, affectability, and explicitness.

SBS include choice calculation was utilized to choose more proper highlights to expand the characterization exactness and decreased the computational season of prescient framework. Cleveland coronary illness dataset was in this investigation and 70% for preparing and 30 % for testing of the dataset. The exactness metric utilized for execution assessment of the framework. The test results show that the utilization of SBS calculation to pick the proper number of highlights that can be utilized for better order exactness utilizing K-Nearest Neighbor [4].

Seven notable classifiers like calculated relapse, K-NN, ANN, SVM, NB, DT, and irregular timberland remained utilized with three element determination calculations Relief, mRMR, and LASSO used to choose the significant highlights. 0c K-overlap cross-approval strategy stayed utilized in the framework aimed at approval [5].
CBMIR is a functional recovery worldview to help advantageous versatile picture recovery. Hashing can be applied as a powerful method to encourage huge scope CBMIR[6], because of its proficient transmission, low stockpiling cost and quick recovery reaction. In any case, most existing hashing strategies are planned dependent on unadulterated visual factual data without thinking about the useful content, which is generally connected with Web pictures. Despite the fact that CMH can conceivably use messages, it actually neglects to completely utilize the writings. This paper proposes a novel solo hashing to extraordinarily use helper writings to improve the viability of hashing in visual space. We pick up hashing codes and capacities inside a brought together THH system, which models high-request semantic relationships of pictures and jam them in the hashing codes through unaided learning.

Distinguishing the handling of crude medical care information of heart data will help in the drawn out saving of living souls and early location of irregularities in heart conditions. AI strategies were utilized in this work to deal with simple data and provide another and novel acumen towards coronary illness. Coronary illness estimate is testing and vital in the clinical field. Be that as it may, the demise percentage can be absolutely measured if the sickness is notable at the beginning phases and precaution measures are established quickly. Further expansion of this examination is exceptionally alluring to guide the examinations to genuine world datasets rather than simply hypothetical methodologies and reproductions. The proposed cross breed HRFLM [7] approach is utilized joining the attributes of Random Forest (RF) and Linear Method (LM). HRFLM end up being very precise in the expectation of coronary illness. The future course of this exploration can be performed with different combinations of AI methods to better expectation procedures. Besides, new component choice strategies can be created to get a more extensive view of the critical highlights to expand the exhibition of coronary illness expectation.

Certain investigations are performed to recognize the presentation of different calculations. FS calculation Relief utilized for significant highlights choice since highlights determination expanded classifier execution as far as precision and lessen preparing season of classifier. The mortal genome arrangement stayed an extraordinary accomplishment, and incredible advancement regarding infections examination was usual. The work is to partake in the conception of multifaceted illnesses. The last aim was to fabricate an arrangement equipped to inspect the SNP[8] evidence and diagnose solid and influenced tests.

As introduced already, it is apparent that the proposed strategy gives recognizable enhancements in the exactness over completely looked at FS calculations in the past investigations more than three datasets [9], for the BC dataset serious outcomes remain acquired, through precision like different structures. In any case, the proposed strategy has the upside of being additional time effective.

AI information mining methods assume a significant part in medical services benefits by conveying a framework to examine the clinical information for analysis of illnesses. The effective recognition of sugar levels is a basic clinical problem for clinical specialists and analysts. To handle this issue, an E- medical services framework is proposed for the location of diabetes utilizing ML information extracting strategies. In the proposed technique, utilized the DT (ID3) [10] calculation for highlights determination as highlights choice is fundamental for powerful preparing and trying of the classifier. Furthermore, gathering learning DT parameter choice calculations Ada Boost and Random Forest are additionally utilized for include choice. The DT AI classifier is utilized aimed in the process of identifying diabetes.
Parkinson’s infection is a risky human illness, and various individuals have been experienced this sickness around the globe. Along these lines, solid procedure is needed for a sufficient acknowledgment of the PD. In this article, we proposed a solid technique for Parkinson’s infection acknowledgment utilizing fitting AI draws near. In particular, SVM has been applied for the arrangement of Parkinson’s sickness and sound subjects. Help and ACO based incorporated technique has been received for the determination of related highlights. Furthermore, K-folds cross approval technique has been utilized for ideal worth determination of Hyper- boundaries for the best model. Besides, assessment measurements have been utilized to assess the exhibition of proposed model.

3. EXISTING SYSTEM

In Existing clinical frameworks, including medical clinic the executives frameworks and dynamic frameworks, center on gathering and mining the whole clinical information. The whole quiet records are stacked and all components are thought of. The clinical information can’t be effortlessly broke down, on the grounds that for creating a probabilistic rating, indications as well as elements prefer test results, current pandemics, clinical history, outer environment conditions, and different variables are required, which might be available in the report. Based on this new measure, an information mining calculation was created to mine the causal connection among drugs and their related with hazard for CAD coronary illness. The restrictive causal-influence was utilized to rank the possible causal relationship between every one of the chose drugs. Calculation could successfully spread the word about ADRs rank high among all the manifestations in the information base.

- Existing frameworks have neglected to use and comprehend the significance of misdiagnosis.
- A vital characteristic which interconnects and addresses every one of these issues. Mining the misdiagnosis property is the key on the grounds that the principal analysis by the clients would have effectively covered all the fundamental factors like patient’s clinical history, climatic conditions, neighborhood, and different components, permitting the client to simply focus on either missed factors like secret indications, winning conditions, complexities, and so forth, or heart Diseases that are like the one previously analyzed.

4. PROPOSED SYSTEM

In the proposed work user will look for the coronary illness analysis (coronary illness and treatment related data) by giving indications as an inquiry in the interne searcher. These indications are pre-handled to make the further interaction simpler to discover the side effects catchphrase which assists with recognizing the coronary illness rapidly. The manifestations which catchphrase is coordinated with the put away clinical info data set to recognize the different heart Diseases identified with that watchword. Different heart Diseases is distinguished; it will make the example coordinating about the various heart Diseases and furthermore discover the likelihood of heart Diseases. Then the heart Disease will make a differential diagnosis to find the heart Disease accuracy.
The keyword which is a preprocessed manifestation is coordinated with the heart Diseases put away in the nearby data set to distinguish the comparing coronary illness identified with those indications given by the client. This needs to look through a record data set of in excess of 20000 heart Diseases and significantly more side effects, which is very tedious, so CFS+PSO characterization was applied to arrange heart Diseases highlights into subgroups. On the off chance that a gathering of side effects coordinate higher inclination is given to that subgroup and looking in that new more modest subgroup along these lines decreases information base access. In design acknowledgment, CFS with PSO Feature Selection scheming is a technique for gathering objects reliant on adjacent preparing representations in the element space. CFS+PSO are a type of example based learning, or apathetic comprehending where the size is just approached nearby and all intention is conceded until order. This element has been recognized as the most reasonable for the current framework.

The Wisconsin Prognostic Cleave Land Train Dataset is taken from the UCI repository site and saved as a content record. This record is then brought into Excel bookkeeping page and the qualities are saved with the comparing credits as section headers. The missing qualities are supplanted with fitting qualities. The ID of the patient cases does not add to the classifier execution. Consequently it is eliminated and the result characteristic characterizes the objective or ward variable subsequently diminishing the list of capabilities size to 33 credits. The algorithmic procedures applied for include significance investigation and arrangement is intricately introduced in the accompanying areas.

4.1 DIMENTIONALITY REDUCTION (CFS+PSO)

The generic problem of supervised feature selection can be illustrated as follows. Given an informational index \( \{(x_i, y_i)\}_{i=1}^n \) where \( x_i \in \mathbb{R}^d \) and \( y_i \in \{1, 2, \ldots, c\} \), we plan to discover an element subset of size \( m \) which contains the most useful highlights. The two well-performing highlight choice calculations on the WPBC data set are momentarily laid out beneath.

Particle Swarm Optimization (PSO) is a worldwide enhancement system that reenacts the social conduct saw in a group (multitude) of birds looking for food. A basic inquiry system in PSO guides the calculation toward the best arrangement through consistent refreshing of the psychological information and social conduct of the particles in the multitude. Highlight choice is an interaction of choosing a subset of applicable highlights from countless unique highlights to accomplish comparable or better grouping execution and improve the calculation productivity. As a significant information pre-handling method, investigation into include choice has been completed in the course of recent many years. Deciding an ideal component subset is a confounded issue.

4.2 MODEL FOR CAD IDENTIFICATION

Multinomial Logistic Regression model (MLR): It is an augmentation of calculated relapse with edge estimator. MLR is a straightforward expansion of double strategic relapse that takes into consideration multiple classes of the ward or result variable. Like twofold calculated relapse, MLR utilizes Maximum probability assessment to assess the likelihood of downright enrollment. Multinomial Logistic Regression is the direct relapse examination to lead when the reliant variable is ostensible with multiple levels.
Default risk is the chance that associations or individuals will be not ready to make the essential portions on their commitment responsibilities. Toward the day’s end, credit default danger is the probability that if you advance money, perhaps they will not have the choice to give the money back on time. It gives reproducible and target conclusion, and subsequently can be an important extra apparatus in medical follows. Outcomes are similarly, encouraging then the planned strategy will be useful in infection diagnostics. To lighten the impact of default risk, banks every now and again power charges that identify with the record holder’s level of default peril. A huge degree of peril stimulates a advanced required return. At that point the information is grouped utilizing MLP, MLG, FURIA and C4.5 utilizing all the highlights of CAD information. Risk forecast instruments are created to recognize patients in danger and to encourage doctor dynamic.

4.3 CLASSIFICATION ALGORITHMS

By considering the preparation dataset to progress boundary conditions that might be exploited to choose every target class. By determining all these boundary values, the aim is to foresee the objective class. To order sex (target class) utilizing hair length as highlight boundary, we could prepare a model utilizing any grouping calculations to concoct some arrangement of limit conditions that can be utilized to separate the male and female sexes utilizing hair length as the preparation include. In sex characterization case the limit condition could the best possible hair length esteem in table 1. Then the data are clustered using MLP, MLG, FURIA and C4.5 using all the features of CAD In gathering, the thought is not to foresee the objective class in sequence, it’s additionally attempting to assemble the comparative sort of things by knowing the most fulfilled condition, all the things in a comparable assembly should be equated and ought no dualistic distinctive gathering things to not be comparative. An investigated outcome proves the dominance of the suggested scheme in terms prediction accuracy of CAD along with the parameters chosen by CFS & PSO. Very less number of inputs is required to relate this strategy. The accuracy can be additionally enlarged with extra data instance.

5. EXPERIMENTAL RESULTS

The clinical data is taken as input, dimensionality reduction is done by which data is clustered based on model construction which shown in Figure 2. Prediction accuracy of classification model is usually measured regarding accuracy and misclassification rate.

\[
\text{Accuracy} = \frac{(tp+tn)}{(tp+tn+fp+fn)}
\]

\[
\text{Misclassification rate} = \frac{(fp+fn)}{(tp+fn+fp+fn)}
\]

Where, tp- true positive rate, tn-true negative rate, fp- false positive rate, tn- true negative rate.

Prediction accuracy of classification model is usually measured in terms of accuracy and misclassification rate.

| TABLE 1: The overall Performance of MLP, MLG, FURIA and C4.5 using all the Features of CAD data |
|--------------------------------------------------|----------------|----------------|----------------|----------------|
| Kappa statistic                                  | MLP            | MLR            | FURIA          | C4.5           |
| 63.2                                             | 10.91          | 6.9            | 0.261          |                |
Mean absolute error & 50.59 & 47.66 & 45.4 & 44.77 \\
Root mean squared error & 53.08 & 48.83 & 624.9 & 73.1 \\
Relative absolute error & 101.18 & 95.32 & 90.7 & 89.5 \\
Root relative squared error & 106.39 & 97.2 & 124.98 & 94.63 \\
Coverage of cases (0.95 level) & 98.98 & 100 & 67.33 & 100 \\
Mean rel. region size (0.95 level) & 98.99 & 100 & 62.71 & 98.35 \\
Accuracy in % & 93.67 & 92.7 & 94.7 & 94.9 \\

Figure 2: Performance of MLP, MLG, FURIA and C4.5 using all the features of CAD data
6. CONCLUSION

Clinical finding is a significant region of exploration which assists with recognizing the event of a coronary illness. The framework, utilizing different methods referenced, will thus uncovered the root coronary illness alongside the arrangement of most plausible heart Diseases which have comparative side effects. The information base utilized is a portrayal data set so to decrease the dataset tokenization, separating and stemming is finished. The venture grants a novel mixture strategy to recognize then affirm CAD cases requiring little to no effort by utilizing clinical information that can be effectively gathered at clinics. Intricacy of the framework is diminished through decreasing the dimensionality of the informational collection with PSO. It gives regenerative and target findings, and subsequently can be a significant extra device in medical practices. Outcomes are equivalently, encouraging and along these line the proposed half and half technique will be useful in coronary illness diagnostics. Trial results exhibit the Predominance of the proposed half breed technique concerning forecast precision of CAD with the highlights chose by CFS and PSO, we need a couple of clinical information to apply this model. The precision can be additionally expanded with more information occasions.

REFERENCES

[1] Samuel, O. W., Asogbon, G. M., Sangainih, A. K., Fang, P., & Li, G. (2017). An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction. Expert Systems with Applications, 68, 163-172.

[2] Geweid, G. G., & Abdallah, M. A. (2019). A new automatic identification method of heart failure using improved support vector machine based on duality optimization technique. IEEE Access, 7, 149595-149611.

[3] A. Haldorai and A. Ramu, Security and channel noise management in cognitive radio networks, Computers & Electrical Engineering, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784

[4] A. Haldorai and A. Ramu, Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability, Neural Processing Letters, Aug. 2020. doi:10.1007/s11063-020-10327-3

[5] Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., & Sun, R. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine leaning algorithms. Mobile Information Systems, 2018.

[6] Zhu, L., Shen, J., Xie, L., & Cheng, Z. (2016). Unsupervised topic hypergraph hashing for efficient mobile image retrieval. IEEE transactions on cybernetics, 47(11), 3941-3954.

[7] Mohan, S., Thirumalai, C., & Sriavastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542-81554.

[8] Haq, A. U., LI, J., KHAN, J., MEMON, M. H., PARVEEN, S., RAJI, M. F., ... & MONDAY, H. N. (2019, December). Identifying the predictive capability of machine learning classifiers for designing heart disease detection system. In 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing (pp. 130-138). IEEE.

[9] Alzubi, R., Ramzan, N., Alzoubi, H., & Amira, A. (2017). A hybrid feature selection method for complex diseases SNPs. IEEE Access, 6, 1292-1301.

[10] Haq, A. U., Li, J. P., Khan, J., Memon, M. H., Nazir, S., Ahmad, S., & Ali, A. (2020). Intelligent Machine Learning Approach for Effective Recognition of Diabetes in E-Healthcare Using Clinical Data. Sensors, 20(9), 2649.