

# Multi Feature DCR based Drug Compound Selection and Recommendation System for Efficient Decision-Making using Genetic Algorithm

ST. Aarthy<sup>1</sup>, Dr. J. L. Mazher Iqbal<sup>2</sup>

Research Scholar, Dept. of Electronics and Communication Engineering<sup>1</sup>

Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India<sup>1</sup>

Professor, Dept. of Electronics and Communication Engineering<sup>2</sup>

Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India<sup>2</sup>

**Abstract**—The performance of treating the cardiac diseases is dependent on the kind of drug being selected. There exist numerous decisive support systems which work according to certain characteristics and factors like drug availability, and popularity. Still, they struggle to achieve expected performance in supporting the medical practitioner. To handle this issue, a multi feature drug curing rate based drug compound selection and recommendation system (MDCRSR) is presented. The method utilizes medical histories and data set of various medical organization around the disease considered. Using the traces, the method identifies the drug compounds and features to perform preprocessing which eliminates the noisy data points. Further, the features of the traces are extracted to perform training with genetic algorithm. At the test phase, the method estimates the fitness measure for different drug combination and compounds by measuring their Drug Curing Rate (DCR). The method performs cross over and mutation to produce various populations of drug compounds. According to the curing rate, the drug compound pattern or population is selected and ranked. The ranked results are populated to the medical practitioner. The method improves the performance of recommendation system as well as drug compound selection.

**Keywords**—Decisive support Systems; GA; DCR; drug selection; compound selection; fitness; recommendation system; cardiac disease; RMDCRSR

## I. INTRODUCTION

The human society faces variety of diseases but not all of them are harmful and claim the human life, but there are few diseases which would claim the human life without giving any time. The cardiac disease is one among them which occurs in various ways and they can be classified as Dextrocardia, Tachycardia, Bradycardia, Hyperkalcaemia, Sinoatrial block, and Myocardialischaemia. Among these diseases, few of them instantly block the heart and leads to cardiac arrest and lead to death. However, by identifying the disease at the early stage, they can be treated with set of drugs. For other diseases, there are varieties of drugs available in various compounds, each of them differs with their efficiency in curing the disease. Here it is about the selection of exact drug for any disease with the available drug compounds.

The medical practitioners are capable of identifying the disease and even they require set of automated decisive

support system in the classification and detection of any disease. Also, they have lot of drugs and compounds in front of them, and they can choose any of them to serve with patients. But, the efficiency of curing the disease cannot be justified and the performance of curing the disease is highly dependent on the combination of drugs being selected. As, there are numerous drugs, the medical practitioner would confuse in the selection of exact drug and compounds. To handle this issue, a real time drug and compound selection approach is discussed.

The recommendation systems are over the decisive support systems which consider variety of medical logs and features. According to the features and traces, the recommendation system identifies set of features in terms of drugs and would measure their performance according to their success and failure. By measuring the value of performance, they can be ranked and populated to the medical practitioner to serve the patient. For example, towards cardiac disease there are many drugs available like Aspirin, Lidocaine, Disopyramide, Procainamide which are used towards arrhythmic diseases. However, these drugs are available, the performance of the drug must be considered while selecting a drug.

The optimal selection of drug is dependent on how efficient the drug compound is in curing the disease. It is more important to identify the exact drug and to perform this, the Genetic algorithm has been adapted in this paper. Genetic algorithm is a scientific approach of searching the drug more effective in curing the disease. The genetic algorithm is applied in several scientific problems of medical issues and the same can be used in the selection of drugs in this model. The drugs of any disease would come on different milligrams and applied to several diseases. The GA algorithm would search on the optimal combination of drug towards any disease by measuring Drug Curing Rate (DCR). The RMDCRSR algorithm estimates the fitness.

## II. RELATED WORKS

There are number of approaches available towards recommendation generation and drug selection for heart diseases. This section details set of approaches related to the problem.

A sentiment analysis-based drug recommendation scheme is presented in [1], which take the reviews of sentiment analysis and support the decision-making problem. The sentiment measurement scheme uses the drug rating and selects according to the patient conditions. Accordingly, recommendations are generated. Similarly, in [2], an implicit feedback based approach cross recommendations (IFCR) which uses epileptics' medical history in identifying the relation of syndrome among drugs. A context aware approach towards hypertensive drugs is presented to support recommendations on personalized scope [3]. The method uses Semantic Web Rule Language (SWRL) to generate recommendations. A detailed review on several recommendation system is presented in [4], to support healthcare professionals.

A user preference-based recommendation scheme is developed in [5], where the recommendations are generated with two approaches and the user can select accordingly. An android based application is developed to support the selection of required medication to manage diabetic in [6]. The recommendations are generated according to the input made.

A group-based approach is discussed in [7], with a consensus reaching process to stimulate the recommendations related to the group users. An evidence-based recommendation scheme is presented in [8], which uses physical activity, exercise in diabetes patients. A naïve bayes classification approach towards disease prediction and drug recommendation is presented in [9], which uses the profile of patients like blood pressure, heart rate, cold and fever in classification where the drug selection is performed according to the symptoms. Similarly, for the hypertensive patients a physical activity recommendation system is presented in [10], which uses the user profile in recommendation generation named HyperModel2PAR.

A genetic algorithm based multi expert scheme (GA) MES-GRS-GA is presented in [11], which discuss consensus scheme towards group recommendation system. A predictive system is designed towards supporting cardiac disease in [12], which analyzes various machine learning algorithms and their results. A Fourier transformation-based heart disease prediction system (FTHDPS) is presented towards predicting chronic heart diseases with time series dat. ANN has been used in recommendation generation [13].

An intelligent HRS using Restricted Boltzmann Machine (RBM)-Convolutional Neural Network (CNN) is presented in [14], which uses the big data in recommendation generation. A hybrid diagnosis scheme for coronary artery disease is presented in [15], with neural network. Similarly, a Fuzzy analytic hierarchy process (Fuzzy\_AHP) technique is presented towards cardiac disease which estimates global weights for various features of individual and classification is performed with ANN [16]. A Congenital Heart Disease diagnosis scheme is presented in [17], which uses BPNN towards recommendation.

Towards providing security for software components a Mamdani fuzzy inference system is presented in [18], which evaluates different security measures and produces recommendations. A multiple kernel learning adaptive neuro fuzzy inference (MKL with ANFIS) is presented towards diagnosis of heart disease which support the classification of individuals [19]. A detailed review on drug recommendation is presented in [20], which used several articles. A relational connection based heterogeneous model is presented in [21], which uses drug, proteins and drugs with side effects. The DTIPred is discussed according to random walk and convolutional neural network. The model integrates various interactions and representations.

A cognitive intervention-based approach is designed in [22], which clubs the result of three different approaches in drug recommendations. A meta-analysis of randomized controlled trial based anti-inflammatory drugs towards cardiac disease is presented in [23]. At last, a cross sectional long term follow up scheme with transanal endorectal pull-through (TERPT) is presented in [24] to support diagnosis and surgery.

#### A. Multi Feature Drug Curing Rate Based Recommendation System with GA

The proposed multi feature drug curing rate based recommendation system reads the medical data set. According to the traces available, the method first preprocesses the data set to remove the noisy records. Further, the method finds the drugs list and their compounds to extract different features. Using the features extracted, the method applies genetic algorithm to estimate the fitness value of different drug compound to select optimal and efficient drug compound towards specific disease. According to the fitness function, the method computes the value of multi feature drug curing rate based on which the drugs are ranked to support medical practitioner. The detailed approach is presented in this section.

The architecture of proposed RMDCRSR model is presented in Fig. 1, and shows various functional components of the mode. Each of the functional stages is discussed in detail in this section.

#### B. Preprocessing

The medical data set given has been fetched here and the preprocessing algorithm finds the list of features available in the entire trace from the data set. According to the features identified, the traces are traversed to verify the features presence with value. If a trace without the feature and value is identified, then it has been removed from the data set. Such traces removed with noisy records are used to perform feature extraction.

Consider the data set given is Meds, which contains  $\emptyset$  records and each has k number of features, then the noise removal is performed as follows:

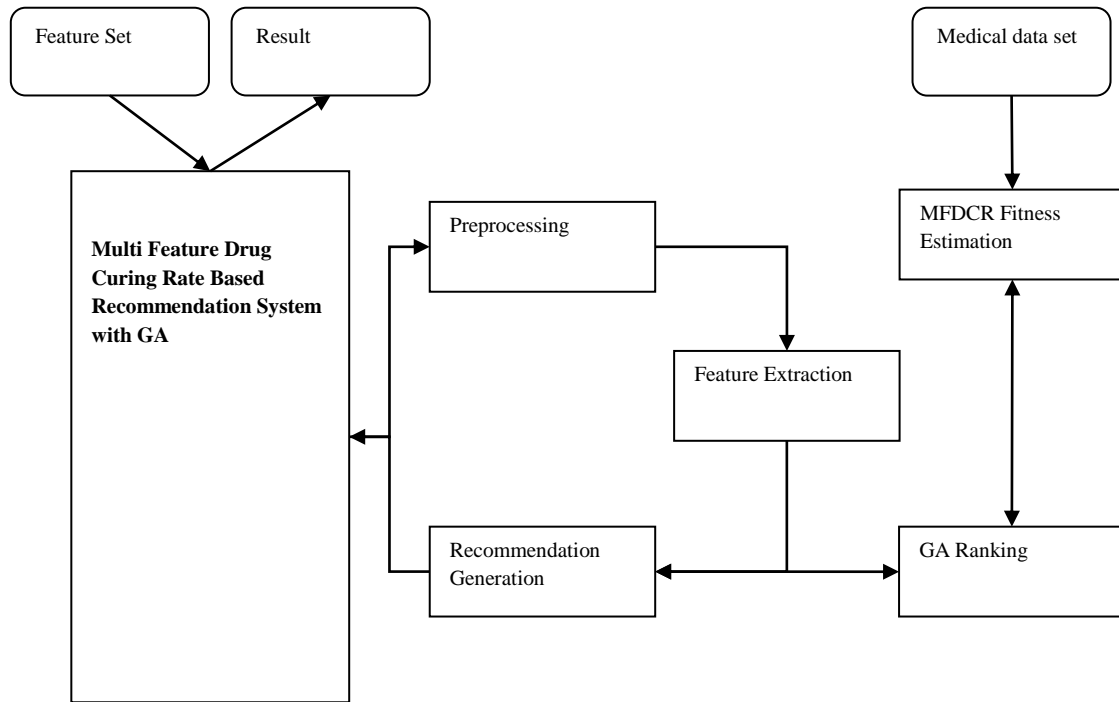


Fig. 1. Architecture of proposed MDCRSR model.

First the list of features are identified using equation (1). Features list  $F_{e\ list} =$

$$F_{e\ list} = \bigcup_{i=1}^{size(Meds)} (Meds(i).features \setminus F_{e\ list}) \cup F_{e\ list} \quad (1)$$

Now, the noisy records are identified and eliminated as follows:

$$Medsp = \bigcup_{i=1}^{size(Meds)} (if\ Meds(i) \in (\forall\ features\ (F_{e\ list}))\ Medsp \cup Meds(i): Meds \cap Meds(i)) \quad (2)$$

Such noise removed data set has been used to extract the features and to support decision making.

### C. Feature Extraction

The data set preprocessed contains number of features and this phase extract several features from the data set. First, the traces belong to specific disease are identified and separated. With the traces separated, the set of drugs given for the disease has been identified and their subsequent compounds are identified. For example, for arrhythmic disease number of drugs would be used and at each drug there would be different manufacturer and volume of drug available, the feature extraction module, finds such drugs and volumes in distinct manner and collects the drugs and compounds. Identified drugs list and compounds are used to perform drug analysis to support recommendation generation. Medical Data set is considered as  $P_{mds}$  and drug vector as  $DV$ . The proposed

system will read the medical dataset and initialize the drug list  $DI$ .

For each record  $r$  Identify set of drugs  $D_s =$

$$D_s = \sum_{i=1}^{size(P_{mds})} Drugs \in p_{mds}(i) \quad (3)$$

For each drug  $d$

$$DI = \bigcup_{i=1}^{size(D_s)} (D_s(i) \in DI? D_s \cap D_s(i): DI \cup D_s(i)) \quad (4)$$

Drug vector  $Dv = \{DI\}$

The feature extraction algorithm identifies the list of drugs according to varying compounds and volumes. According to the drugs identified, the method performs recommendation generation.

### D. MFDCR Fitness Estimation

The fitness function of the proposed model estimates the fitness value by computing the multi factor drug curing rate (MFDCR). Each drug has specific curing rate on specific disease and according to that the method computes the DCR value for each disease. Based on the value of DCR, the method computes the value of MFDCR. It has been measured according to the difference on different logs status. To measure the value of fitness, the method split the traces of disease under success and failure classes. Once the traces are split, then the fitness value is measured as follows:

Consider the population given is  $p = \{D1, D4, D7, D9, D11\}$  where  $Dx$  represent the drug name and according to that

the fitness value is measured in terms of multi feature drug curing rate (MFDCR). Similarly, consider the trace set Sdt contains the traces of drugs with success treatment and Fdt represent the traces of drugs with failure treatment, then the method computes the Drug Curing Rate as follows:

First the drug trace has been identified as follows:

$$\text{Success Drug Trace ST} = \left( \sum_{i=1}^{\text{size}(Sdt)} Sdt(i) \in Dk \right) \&\& Sdt(i).state == Success \quad (5)$$

$$\text{Failure Drug Trace FT} = \left( \sum_{i=1}^{\text{size}(Sdt)} Sdt(i) \in Dk \right) \&\& Sdt(i).state == Failure \quad (6)$$

According to the values of equation (5, 6), the value of drug curing rate is measured as follows:

$$DCR = \frac{\text{Dist}(\text{Size}(ST) - \text{Size}(FT))}{\text{size}(ST) + \text{size}(FT)} \quad (7)$$

According to the drug curing rate of different drugs, the value of fitness function as MFDCR has been measured as follows:

$$MFDCR = \frac{\sum_{i=1}^{\text{size}(DI)} DI(i).DCR}{\text{size}(DI)} \quad (8)$$

Where, DI-represent the drug list and size(DI) represent the total number of drugs. According to the value of MFDCR, the fitness of the drugs pattern or population has been verified.

#### E. GA-Ranking

The proposed modified genetic algorithm finds the drugs list and their compounds. With the set of drugs, the method initially generates the population according to the size of drug list. For the populated drug compound, the method computes the MFDCR fitness value. If the fitness value greater than the threshold, then the population is selected as the result as recommendation. Otherwise, the process is iterated by generating different combinations by crossover operation and mutation operations. In this algorithm, the crossover and mutation operations are generated by computing the combinations with the number of drugs available in the earlier population. At each population, the method computes the value of MFDCR value. Finally, the method is iterated till an optimal drug composition is selected according to the value of MFDCR and threshold. Also, the method tries to reduce the number of compositions or drugs by choosing the drug population according to the MFDCR value. Finally, top few drug compositions or populations are populated as result to the user.

The proposed method Takes the Medical data set Meds, Disease D and returns Recommendations Rc.

When MFDCR < Th then

$$\text{Population set Ps} = \underset{i=1}{\text{GenerateCombinations}(i, p)}^{\text{size}(p)} \quad (9)$$

For each population pi

$$MFDCR = \text{Fitness}(pi)$$

If MFDCR > Th then

Population list PI = Rank Populations by MFDCR.

Recommendations Rc = PI

The genetic algorithm based ranking approach reads the data set and generates the initial population and computes the fitness value. According to the threshold the recommendation are generated and if the value of MFDCR is less than threshold, the process is iterated by performing crossover and mutation functions on the population of drug given. According to the value of MFDCR, the combinations of drug are ranked to produce recommendations to the user.

### III. RESULTS AND DISCUSSION

The proposed real time multi feature DCR based drug recommendation scheme with Genetic algorithm has been implemented and evaluated for its performance under various parameters. The evaluation has been performed according to the data set maintained by different medical organizations and the efficiency of the methods are measured on various factors. The performance of the proposed algorithm has been evaluated with MIMIC-III data set. MIMIC-III is an open source database provided by American critical care unit located at Beth Israel Deaconess Medical Center, Boston in the period of 2001 and 2012, which has been obtained from MIMIC-III. It covers variety of information from general patient details, intensive care unit features obtained, treatment details, diagnosis reports from lab, clinical and various other details at different stage with the improvement with the treatment and Data includes vital signs, medications, laboratory measurements, observations and notes charted by care providers, fluid balance, procedure codes, diagnostic codes, imaging reports, hospital length of stay, survival data, and more. In total it covers thousands of features in 26 tables which can be used to perform analysis on heart diseases.

TABLE I. ANALYSIS ON RECOMMENDATION GENERATION

Recommendation Generation Performance			
	50 Drugs	100 Drugs	200 Drugs
SWRL	73	79	83
HyperRecSysPA	77	82	85
MES-GRS-GA	81	86	89
MFDCR_GA	87	92	97

#### A. Performance

In the Table I performance of recommendation generation has been measured at different number of drugs cases. In each test case, the proposed MFDCR\_GA approach has produced higher recommendation performance than other approaches.

The performance of different methods in recommendation generation has been measured and presented in Fig. 2. The proposed MFDCR\_GA has produced higher performance at different number of drugs cases.

In Table II false ratio of recommendation generation has been measured for different approaches and the proposed MFDCR\_GA based approach has produced less false ratio than other approaches.

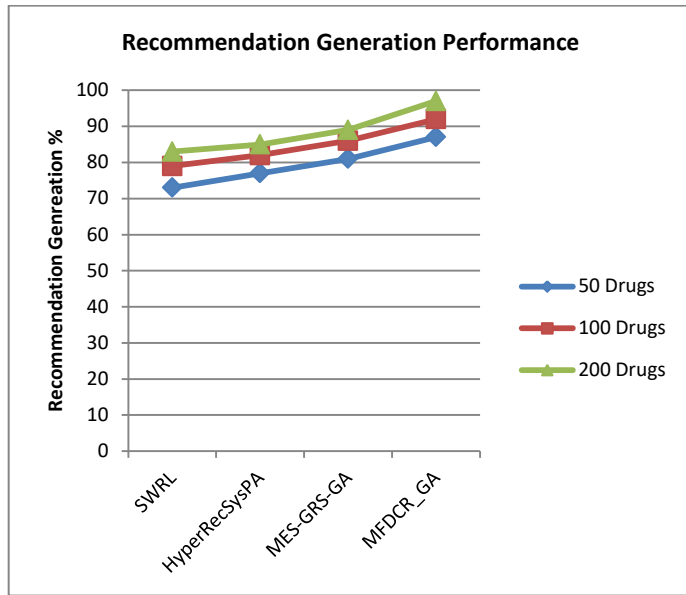


Fig. 2. Performance on recommendation generation.

TABLE II. ANALYSIS ON FALSE RATIO

False Ratio in Recommendation Generation			
	50 Drugs	100 Drugs	200 Drugs
SWRL	27	21	17
HyperRecSysPA	23	18	15
MES-GRS-GA	19	14	11
MFDCR_GA	13	8	3

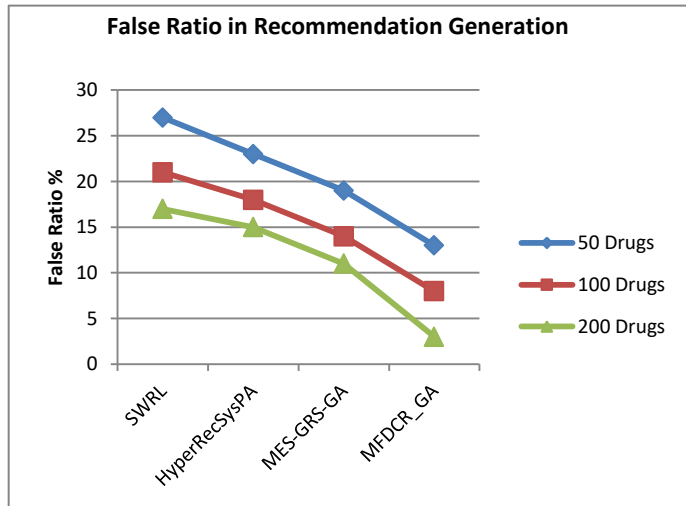


Fig. 3. False ratio in recommendation generation.

The ratio of false recommendation produced by the methods are measured and presented in Figure 3. In each test case, the proposed MFDCR\_GA approach has produced less false ratio than other approaches.

In the table 3 time complexity in recommendation generation has been measured for different approaches and the proposed MFDCR\_GA based approach has produced less time complexity than other approaches.

TABLE III. TIME COMPLEXITY

Time Complexity in Recommendation Generation			
	50 Drugs	100 Drugs	200 Drugs
SWRL	57	71	96
HyperRecSysPA	53	68	87
MES-GRS-GA	46	62	81
MFDCR_GA	23	28	39

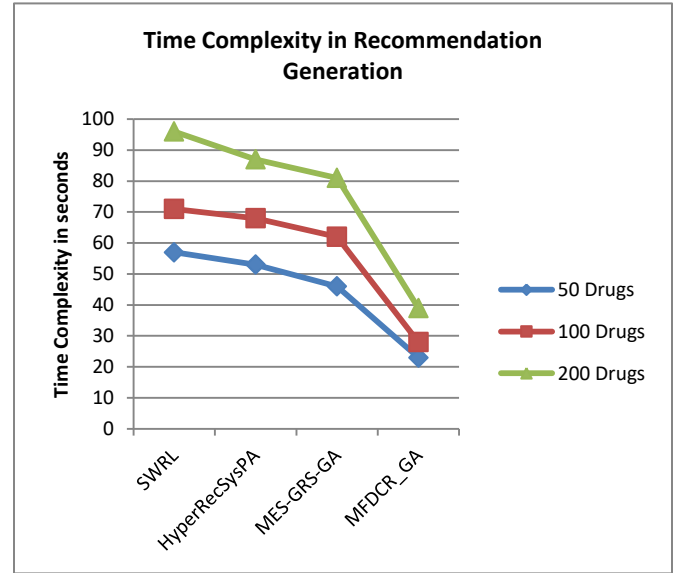


Fig. 4. Analysis on time complexity.

The time complexity in recommendation generation has been measured and presented in Fig. 4, where the proposed MFDCR\_GA has produced less time complexity than other approaches.

#### IV. CONCLUSION

This paper presented a detailed implementation of multi feature DCR recommendation system with Genetic algorithm. The method reads the data set and finds set of drugs for different disease class. Further the value of drug curing rate is measured on each drug towards various disease classes. Also, the method computes the multi feature drug curing rate (MFDCR) towards various disease classes. According to the value of MFDCR the drugs and their compounds are ranked to generate recommendation to the medical practitioners. The proposed approach has been evaluated with MIMIC-III database and the method improves the performance of recommendation generation with least false ratio and time complexity.

#### REFERENCES

- [1] Md. Deloar Hossain; Md. Shafiu Azam; Md Jahan Ali; Hakilo Sabit, Drugs Rating Generation and Recommendation from Sentiment Analysis of Drug Reviews using Machine Learning, IEEE, Emerging Technology in Computing, Communication and Electronics (ETCCE), 2020.
- [2] C. Chen, L. Zhang, X. Fan, Y. Wang, C. Xu and R. Liu, "A epilepsy drug recommendation system by implicit feedback and crossing recommendation", Proceedings – 2018 IEEE Smart World Ubiquitous

- Intelligence and Computing Advanced and Trusted Computing Scalable Computing and Communications Cloud and Big Data Computing Internet of People and Smart City Innovations Smart World IUICIATCIScaIComICBDCo, 2018.
- [3] D. Chen, D. Jin, T. T. Goh, N. Li and L. Wei, "Context-Awareness Based Personalized Recommendation of Anti-Hypertension Drugs", *J. Med. Syst.*, vol. 40, no. 9, pp. 1-10, Sep. 2016.
- [4] AThi Ngoc Trang Tran, *Recommender systems in the healthcare domain: state-of-the-art and research issues*, Springer, Journal of Ambient Intelligent Information Systems, 2020.
- [5] Atas, M., Tran, T.N.T., Felfernig, A., Polat-Erdeniz, S., Samer, R., & Stettinger, M. (2019). Towards similarity-aware constraint-based recommendation. In *Advances and trends in artificial intelligence, lecture notes in computer science*, (pp. 287–299). Springer.
- [6] Bankhele, S., Mhaske, A., & Bhat, S. (2017). V., s.: a diabetic healthcare recommendation system. *International Journal of Computer Applications*, 167, 14–18.
- [7] Castro, J., Quesada, F.J., Palomares, I., & Martínez-López, L. (2015). A consensus-driven group recommender system. *International Journal of Intelligent Systems*, 30(8), 887–906.
- [8] Colberg, S., Sigal, R., Yardley, J., Riddell, M., Dunstan, D., Dempsey, P., Horton, E., Castorino, K., & Tate, D. (2016). Physical activity/exercise and diabetes: a position statement of the american diabetes association. *Diabetes Care*, 39, 2065–2079.
- [9] Gujar, D., Biyani, R., Bramhane, T., Bhosale, S., & Vaidya, T.P. (2018). Disease prediction and doctor recommendation system. *International Research Journal of Engineering and Technology (IRJET)*, 5, 3207–3209.
- [10] Luciano Rodrigo Ferretto; Ericles Andrei Bellei, *A Physical Activity Recommender System for Patients With Arterial Hypertension*, *IEEE Access*, Vol.8, 2020.
- [11] Ritu Meena and Sonajharia Minz, *Group Recommender Systems – An Evolutionary Approach Based on Multi-expert System for Consensus*, De Gruyter, 2018.
- [12] Yar Muhammad, Muhammad Tahir, *Early and accurate detection and diagnosis of heart disease using intelligent computational model*, *scientific reports*, 2020.
- [13] Subhashini narayanan, *A novel recommender system based on FFT with machine learning for predicting and identifying heart diseases*, *Neural Computing and Applications*, 2019.
- [14] Abhaya Kumar Sahoo, *DeepReco: Deep Learning Based Health Recommender System Using Collaborative Filtering*, *MDPI, Computation*, 7(2), 2019.
- [15] Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., Moosaei, H. & Yarifard, A. A. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Comput. Methods Programs Biomed.* 141, 19–26 (2017).
- [16] Samuel, O. W., Asogbon, G. M., Sangaiah, A. K., Fang, P. & Li, G. An integrated decision support system based on ANN and Fuzzy\_AHP for heart failure risk prediction. *Expert Syst. Appl.* 68, 163–172 (2017).
- [17] Vanisree, K. & Singaraju, J. Decision support system for congenital heart disease diagnosis based on signs and symptoms using neural networks. *Int. J. Comput. Appl.* 19, 6–12 (2015).
- [18] Nazir, S., Shahzad, S., Mahfooz, S. & Nazir, M. Fuzzy logic based decision support system for component security evaluation. *Int. Arab J. Inf. Technol.* 15, 224–231 (2018).
- [19] Manogaran, G., Varatharajan, R. & Priyan, M. Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system. *Multimedia Tools Appl.* 77, 4379–4399 (2018).
- [20] Idris Rabi, *Recommender System Based on Temporal Models: A Systematic Review*, *MDPI, Applied Sciences*, 10(7), 2020.
- [21] Xiaoqiang Xu Ping Xuan, *Infering drug target interaction based randomwalk and convolution neural network*, *IEEE/ACM Trans Comput Biol Bioinform* 2021.
- [22] Zana Bućinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. *To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making*. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 2021.
- [23] Ivan Wudexi, Elica Shokri, *Comparative Effectiveness of Anti-Inflammatory Drug Treatments in Coronary Heart Disease Patients: A Systematic Review and Network Meta-Analysis*, *Hindawi, Mediators of Inflammation*, 2021.
- [24] Johan Hasseri, Josefine Hedbys, *Treatment and Patient Reported Outcome in Children with Hirschsprung Disease and Concomitant Congenital Heart Disease*, *Hindawi, Biomedical research international*, 2017.