Rule Induction and Prediction of Chronic Kidney Disease Using Boosting Classifiers, Ant-Miner and J48 Decision Tree

Arif-Ul-Islam
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
a.rifislamm.cse@gmail.com

Shamim H Ripon
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
dshr@ewubd.edu

***Abstract*—Chronic Kidney Disease (CKD) is one of the deadliest diseases that slowly damages human kidney. The disease remains undetected in its early stage and the patients can only realize the severity of the disease when it gets advanced. Hence, detecting such disease at earlier stage is a key challenge now. Data mining is a branch of Artificial Intelligence that is widely used to derive interesting patterns from a large volume of medical data. While various data mining techniques used by Experts, boosting and rule extraction techniques have rarely been applied in analyzing Kidney diseases. Boosting is a method of ensemble technique that enhances the prediction power of a data mining model. AdaBoost and LogitBoost are used here for comparing the performance of classification. Ant-Miner is also a data mining algorithm that applies Ant Colony Optimization technique. Ant-Miner along with Decision tree have been used in the paper to derive rules. The aim of this paper is two-fold: analyzing the performance of boosting algorithms for detecting CKD and deriving rules illustrating relationship among the attributes of CKD. The best information retrieved by both classification and rule generation techniques are promising and can be adopted by the Medical Scientists for their research purpose.**

***Keywords*—Chronic Kidney Disease, Boosting Algorithms, J48, Ant-Miner**

I. Introduction

Two bean-shaped organs, named kidney, are two important parts in human body. Kidney removes waste from blood by filtering. If this filtering system is hampered, protein can seep to urine and waste elements can remain in blood. And gradually, kidney lose its ability to filter. This failure of kidney is called Chronic Kidney Disease (CKD), also known as Chronic Renal Disease. Whole body is affected by kidney failure. Risk factors of CKD are diabetes, smoking, lack of sleeping, hypertension , improper diet etc. Among them diabetes is the more dangerous factor [1]. Generally, medical experts determine this kidney disease by a value called Glomerular Filtration Rate (GFR) [2]. It is calculated by age, blood test, sex etc. CKD ranks number 27 and 18 in 1990 and 2010 respectively as world's prime reason of death [3]. 956,000 people died in 2013 because of CKD [4]. At the last stage, the patient must take dialysis or do kidney transplantation. One of the best ways to reduce this death rate is early treatment [5]. But in developing countries, patients take treatment when they reached in serious state. An automated system can be built in order to detect CKD affected patients before reaching in last stage. Clinical data such as age, sex, blood pressure etc. of patient can be used to achieve this. Many researches have been done in order to construct artificial systems, which serve solutions for disease detection.

To predict diseases, data mining or machine learning models are playing a vital role. By making some mathematical approaches, data mining models extract patterns from data and later these patterns are used for the survival of patients. Multilayer Perceptron (MP), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF) etc. are some renowned machine learning methods which were successfully implemented to examine and classify the kidney disease in [6][7][8].

The reason we choose Boosting algorithms, J48 and Ant-Miner because they are easy to implement and proved to be better in previous researches. Boosting algorithm is an ensemble type machine learning algorithm which converts weak classifiers to strong model to achieve a better accuracy. AdaBoost is a popular boosting algorithm. It was also applied to chronic kidney disease for classification purpose in [26]. These researches achieved better accuracy score. LogitBoost is another boosting approach which was proved to be better than AdaBoost in [25]. On the other hand, some rule induction methods namely J48, RIPPER, Ant-Miner etc. have a huge impact in the field of medical science. These classification rules are formed of "IF..THEN". IF part contains rule antecedent and THEN part contains rule consequent. Consequent part is the prediction of the class. Results of a comparison, described in [27], between Ant-Miner, J48 and other rule induction methods showed that Ant-Miner was superior to others. LogitBoost and Ant-Miner were never used on CKD dataset and they were justified to be comparatively better performer than AdaBoost and J48 respectively by researchers. Thus the goal of this research is set to provide 1) a comparative study between two boosting algorithms AdaBoost and LogitBoost for diagnosis of CKD 2) analysis of decision rules inducted by J48 decision tree and Ant-Miner over CKD dataset. Comparatively best

978-1-5386-9111-3/19/$31.00 ©2019 IEEE

output from our research may contribute to the medical field to support the patient detection who have chronic renal failure and to identify the people who can be at risk of having CKD by extracting decision rules.

Remaining of this paper is structured as follows- next section is about past data mining works in the research field of chronic kidney disease, section III talks about dataset, methods and implementation details, section IV deals with experimental results and in the end the paper has been concluded.

II. Related Works

This section describes the past related experiments, applied on CKD dataset. A number of works have been done for extracting useful informations from Chronic Kidney Disease dataset using data mining techniques. Authors of [9] proposed a study which involved six classifiers: KNN, NB, SVM, decision tables, RF, J48 and three ensemble methods: bagging, random subspace, AdaBoost in 2016. Their result showed that J48 basis algorithm and random tree basis algorithm provided 100% accuracy. Anu Chaudhary and Puneet Garg [10] examined chronic kidney disease using k-means algorithm and a-priori. Result was evaluated using Receiver Operating Characteristic (ROC). A study for recognizing the CKD using SVM, DT, NB and KNN algorithm was presented in [11]. Experimental results proved that DT was better than other algorithms. Ani R et al. 2016 [12] performed various classification algorithms such as NB, Back Propagation Network (BPN), DT, Random Subspace, LDA classifier and KNN. Among them Random subspace achieved better accuracy. DT and NB classification techniques were applied to predict CKD for prevention of death rate caused by CKD [13]. Authors implemented these data mining methods using rapid miner tool. The performance accuracies of DT and NB are 96% and 81% respectively. In 2017, an experiment [26] conducted by M. S. Wibawa, I. M. D. Maysanjaya and I. M. A. W. Putra showed that combination of KNN, CFS and AdaBoost was 98.1% success. Another study [29] by Engin Avci, Songul Karakus, Ozlem Ozmen and Derya Avci showed that J48 had highest prediction accuracy of 99% comparing with K-Star and SVM. Ruey Kei Chiu et al. [30] built a system which can predict Chronic Kidney Disease at early stage. They used various neural network algorithms to find the best model. Our proposed methodology is going to analyse the performance of AdaBoost and LogitBoost classifiers and find the best decision rules based on the comparative research among J48 and Ant-Miner methods.

III. Proposed Methodology

The experiment is implemented on a system having 8GB RAM and 3.5 GHz Intel Core i-7 processor. In this paper, analysis of the chronic kidney disease dataset is divided into two parts (classification and rule generation). Classification and rule induction are done using WEKA 3.8 tool [28] and Myra tool [14]. Weka is a tool for data mining tasks. It contains procedures for preprocessing, classifications, clustering, association rules etc. On the other hand, Myra tool is a combination of many Ant Colony Optimization algorithms such as AntMiner, cAntMiner, cAntMinerPB, Ant-Tree Miner etc.

- CKD - Patient having Chronic Kidney Disease
- NOTCKD - Patient not having Chronic Kidney Disease

TABLE I
Types of the attributes

Parameters	Type
Age (age)	Discrete Integer Values
Blood pressure (bp)	Discrete Integer Values
Specific gravity (sg)	Nominal Values
Albumin (al)	Nominal Values
Sugar (su)	Nominal Values
Red blood cells (rbc)	Nominal Values
Pus cell (pc)	Nominal Values
Pus cell clumps (pcc)	Nominal Values
Bacteria (ba)	Nominal Values
Blood glucose random (bgr)	Discrete Integer Values
Blood urea (bu)	Discrete Integer Values
Serum creatinine (sc)	Numeric Values
Sodium (sod)	Discrete Integer Values
Potassium (pot)	Numeric Values
Hemoglobin (hemo)	Numeric Values
Packed cell volume (pcv)	Discrete Integer Values
WBC count (wc)	Discrete Integer Values
RBC count (rc)	Numeric Values
Hypertension (htn)	Nominal Values
Diabetes mellitus (dm)	Nominal Values
Coronary artery disease (cad)	Nominal Values
Appetite (appet)	Nominal Values
Pedal edema (pe)	Nominal Values
Anemia (ane)	Nominal Values
Class (class)	Nominal Values [Target or Class Label]

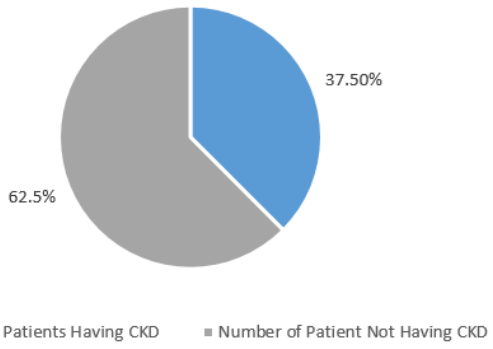

Fig. 1. Percentage of Healthy and Unhealthy Patients

A. Dataset

The dataset contains information of 2800 patients where number of healthy patients is 1050 and 1750 adults have Chronic Kidney Disease. 24 records such as age, blood pressure, sugar, blood urea, appetite, red blood cells, sodium,

diabetes mellitus etc. have been collected from each patient. Among them, 13 are categorical and 11 are numerical. Table I describes the types of the attributes. Table II shows the descriptive analysis of numerical attributes of CKD dataset. There are two class labels (Figure 1). Among 2800, most people are from age group 40 to 60 (Figure 2).

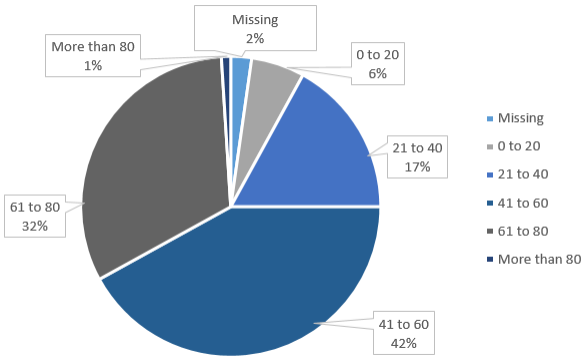

Fig. 2. Age Analysis of People in Dataset

TABLE II
DETAILS OF NUMERICAL ATTRIBUTES

Parameters	Minimum	Maximum	Mean	Std.Dev
Age	2	90	51.483	17.17
bp	50	180	76.469	13.684
bgr	22	490	148.037	79.282
bu	1.5	391	57.426	50.503
sc	0.4	76	3.072	5.741
sod	4.5	163	137.529	10.409
pot	2.5	47	4.627	3.194
hemo	3.1	17.8	12.526	2.913
pcv	9	54	38.884	8.99
wbcc	2200	26400	8406.122	2944.474
rbcc	2.1	8	4.707	1.025

B. Methodology

Proposed methodology consists of below parts.

- Applying classification methods AdaBoost and LogitBoost on CKD dataset using decision stamps as base learner
- Constructing decision rules generated from CKD dataset using J48 decision tree and Ant-Miner algorithm
- a. Analyzing prediction performances based on Root Mean Squared Error, Kappa and F-measure b. Comparing rules extracted by J48 and Ant-Miner

Workflow of our research is shown in Figure 3.

C. Preprocessing and Implemented Methods

Dataset has columns of discrete and nominal values. As column values are repeated, therefore, missing values replaced by mode values of that particular column. Later, two types of algorithms are implemented in this proposed methodology. They are a. Boosting Algorithm b. Rule Induction Algorithm. They are discussed below.

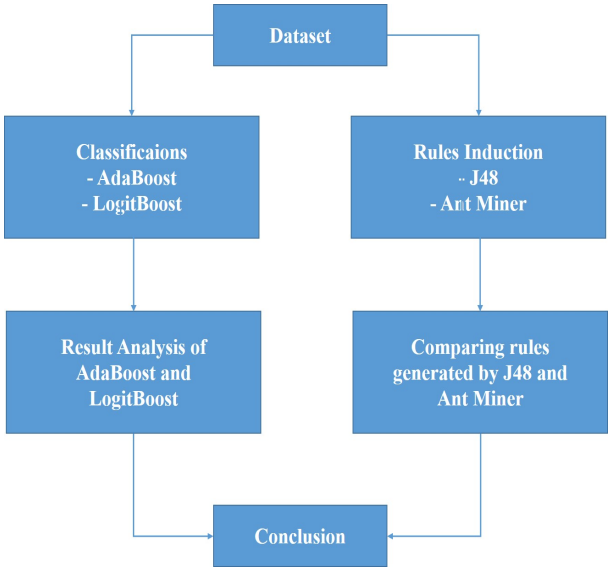

Fig. 3. Research Structure

Boosting Algorithms: Boosting algorithms combine several weak classifiers to form a strong classifier to improve the classification accuracy. A more practical algorithm Adaptive Boosting (AdaBoost) was proposed by Freund and Schapire in 1997. Though AdaBoost reduces training error, it suffers from the over-fit problem [15]. Later, in 2000 Friedman et al. showed that LogitBoost had overcome this situation with a better generalization. Boosting algorithms solved various healthcare problems such as prediction of protein structure class in [15], cancer classification in [16] and recognition of breast cancer in [17].

Rule Induction Methods: Generally, decision rules are generated from a dataset to discover relations within that dataset. Rules can be inducted from medical dataset, which are easily read by most humans. J48 decision tree and Ant-Miner are two famous rules induction methods. They are used in medical fields such as detecting lung cancer [18], diabetes [19], cancer [20]. J48 is known as C4.5 decision tree algorithm [21]. C4.5 was proposed in 1993 [22]. J48 is a Java implementation of C4.5 in Weka [23]. J48 can handle both categorical and numerical values [21]. It builds tree from labeled data and constructs a set of decision rules. Information gain is used to find the best branches. Node and leaf represents attributes and class respectively. Ant-Miner is an Ant Colony Optimization (ACO) algorithm which was proposed in [24]. In real life, ants follow a path from their home to food. At the time of moving, they leave pheromone in the road. When an ant has to choose a path among two or more paths, it used to take the road which has more pheromone [24]. In Ant-Miner algorithm, each time a solution is built for target. And sequentially it finds a list of decision rules from training cases. Rule construction process of Ant-Miner is shown below-

- Every time one classification rule is discovered and it is added to the list of discovered rules.
- The training instances which are covered by that rule will be removed from the training set
- Above process is iteratively done while number of uncovered training cases is more than some threshold value.

IV. Result Analysis

In this section, performance of the proposed method is dealt. To evaluate the performance of the classification methods, f-measure, root mean squared error (RMSE) and kappa are used. K-fold cross validation method is popularly applied in case of analyzing machine leaning approaches as it is essential for avoiding overfitting problem. In this research, 10-fold is used.

A. Performance of AdaBoost and LogitBoost

Accuracy is the common measurement for comparison among machine learning algorithms. Accuracy is calculated by summation of true positive and true negative divided by total instances.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

In equation 1, TP is true positive, TN is true negative, FP is false positive and FN is false negative. Both boosting algorithms AdaBoost and LogitBoost performed very well and achieved good accuracy against the CKD dataset. Because boosting algorithms iteratively do the classification until it achieves the better accuracy. Performance comparison between AdaBoost and LogitBoost is shown in Table III and Figure 4.

TABLE III
Performance Analysis of AdaBoost and LogitBoost

	AdaBoost	LogitBoost
Accuracy	99%	99.75%
Correctly classified instance	2772	2793
Incorrectly classified instance	28	7

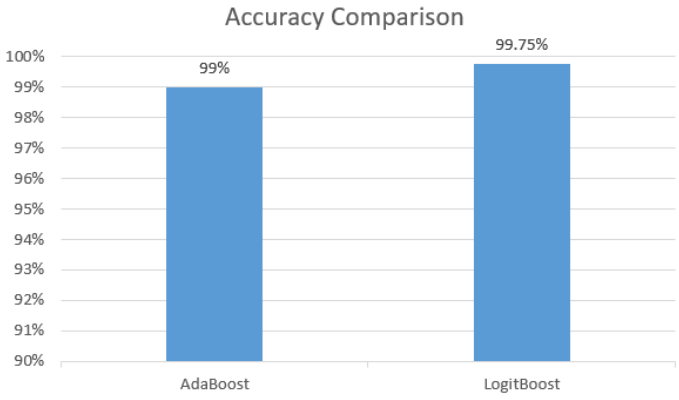

Fig. 4. Accuracy Comparison of AdaBoost and LogitBoost

TABLE IV
RMSE, F-Measure and Kappa Values of AdaBoost and LogitBoost

	AdaBoost	LogitBoost
RMSE	0.0902	0.0641
F-Measure	0.990	0.998
Kappa	0.9788	0.9947

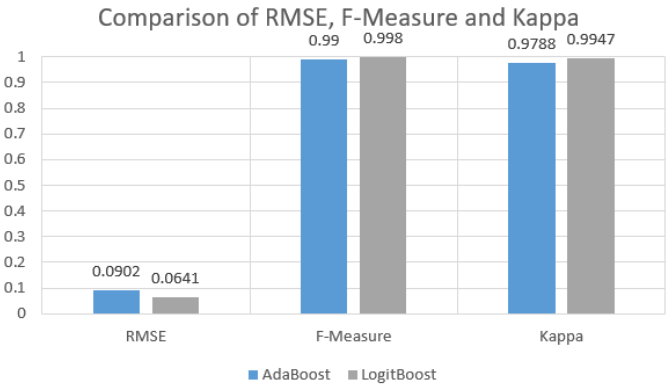

Fig. 5. RMSE, F-Measure and Kappa Comparison Between AdaBoost and LogitBoost

Root mean squared error interpretation: It tells the error between actual result and predicted result. Equation 2 shows the formula of RMSE. It ranges from 0 to 1. As low value of RMSE means better fit model, performance of AdaBoost was lower than LogitBoost's performance (Figure 5).

$$RMSE = \sqrt{\overline{(R - P)^2}} \quad (2)$$

In equation 2, R is actual results and P is the predicted outcome.

F-Measure interpretation: F-measure is a combined metric of precision and recall. Precision is defined as how many patients actually have CKD among all people who are thought to have CKD. Equation 3 describes the formula of calculating f-measure. Recall refers to how much model can identify among the patients who truly have CKD. Higher value of precision and recall is good. Table V contains precision and recall values which were resulted from applying AdaBoost and LogitBoost. It can be interpreted that LogitBoost performed well, as precision, recall and harmonic mean f-measure values of LogitBoost is higher (Figure 5).

$$Fmeasure = \frac{(2 * precision * recall)}{(precision + recall)} \quad (3)$$

TABLE V
Precision and Recall of AdaBoost and LogitBoost

	AdaBoost	LogitBoost
Precision	0.990	0.998
Recall	0.990	0.998

Kappa statistics interpretation: Value of Kappa ranges between 0 to 1 inclusive. When Kappa value is above 0.75, that means there is a strong relation between actual values and predicted values. Here, Kappa value of LogitBoost is higher than the AdaBoost, that means predictions of LogitBoost is more accurate than the predictions of AdaBoost (Figure 5).

Both Boosting algorithms had accuracy close to 100% because they constructed a strong classifier based on several weak classifiers and thus they improved their performances. But as Logitboost is a updated version of Adaboost, it came up with satisfactory results.

B. Rules Construction

The decision rules have been collected from tree built by J48 and Ant-Miner. Each rule is formed by the path from root to leaf. Rule antecedent may have several "AND" signs. Leaf is the prediction label. Both of them produced total 14 nodes and 9 leaf nodes. Figure 6 shows the J48 decision tree for classification of CKD patients.

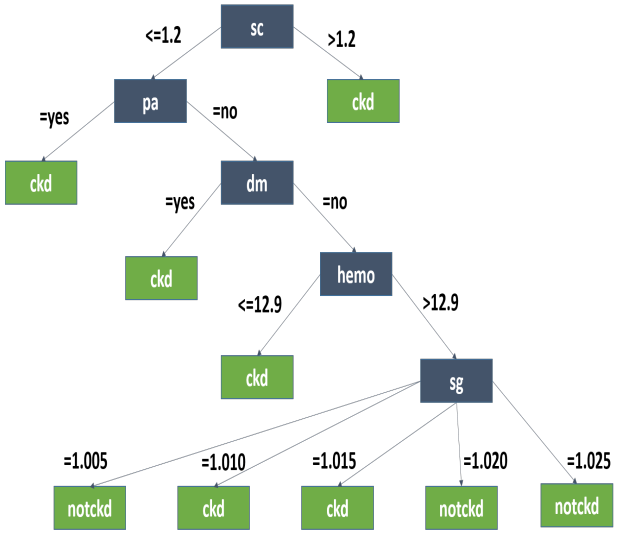

Fig. 6. J48 Decision Tree

Rules generated by J48 are-

R1: If (sc>1.2) Then CKD
R2: If (sc<=1.2 && pe==yes) Then CKD
R3: If (sc<=1.2 && pe==no && dm==yes) Then CKD
R4: If (sc<=1.2 && pe==no && dm==no && hemo<=12.9) Then CKD
R5: If (sc<=1.2 && pe==no && dm==no && hemo>12.9 && sg== 1.005) Then NOTCKD
R6: If (sc<=1.2 && pe==no && dm==no && hemo>12.9 && sg== 1.010) Then CKD
R7: If (sc<=1.2 && pe==no && dm==no && hemo>12.9 && sg== 1.015) Then CKD
R8: If (sc<=1.2 && pe==no && dm==no && hemo>12.9 && sg== 1.020) Then NOTCKD
R9: If (sc<=1.2 && pe==no && dm==no && hemo>12.9 && sg== 1.025) Then NOTCKD

There are nine (9) rules output by J48. They classified 2772 instances successfully with 99% accuracy. Figure 7 shows a decision tree built by Ant-Miner. There are also 9 rules, derived from Ant-Miner.

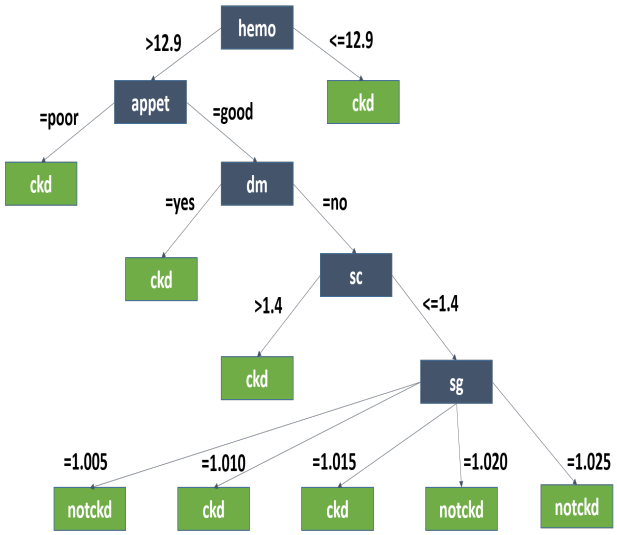

Fig. 7. Decision Tree by Ant-Miner

R1: If (hemo<=12.9) Then CKD
R2: If (hemo>12.9 && appet==poor) Then CKD
R3: If (hemo>12.9 && appet==good && dm==yes) Then CKD
R4: If (hemo>12.9 && appet==good && dm==no && sc>1.4) Then CKD
R5: If (hemo>12.9 && appet==good && dm==no && sc<=1.4 && sg==1.005) Then NOTCKD
R6: If (hemo>12.9 && appet==good && dm==no && sc<=1.4 && sg==1.010) Then CKD
R7: If (hemo>12.9 && appet==good && dm==no && sc<=1.4 && sg==1.015) Then CKD
R8: If (hemo>12.9 && appet==good && dm==no && sc<=1.4 && sg==1.020) Then NOTCKD
R9: If (hemo>12.9 && appet==good && dm==no && sc<=1.4 && sg==1.025) Then NOTCKD

Above rules constructed by Ant-Miner identified 2786 healthy and unhealthy patients successfully.

TABLE VI
ANALYSIS OF CLASSIFICATION RULES

	J48	**Ant-Miner**
Number of rules	9	9
Accuracy	99%	99.5%
Correctly classified instance	2772	2786
Incorrectly classified instance	28	14

One observation can be made is that though both rule extractors gave better decision rules, Ant-Miner had a comparative edge over J48 becuase it iteratively tried untill it achieved the effective accuracy.

V. CONCLUSION

People of all over the world are suffering from CKD. Treatment at the early stage of the disease can cure the patient.

Researches showed that many data mining techniques had been applied for CKD classification. Among those algorithms, AdaBoost classifier and J48 rule induction method performed well. While comparing the classification performance of boosting algorithms, our experimental results revealed that the performance of AdaBoost was less than that of LogitBoost by a fraction. Various performance indicators have been applied for the comparative analysis. Whereas, for rule generation, Ant-Miner performed better than J48 decision tree. The combination of these two comparatively better approaches can be used as an important and valuable tool to help experts to treat CKD patients. In future, our target is to research on analysis of other diseases based on various boosting algorithms with different base learners and other variations of Ant Colony Optimization algorithms.

REFERENCES

[1] K. Chimwayi, N. Haris, R. Caytiles and N. Iyengar, "Risk Level Prediction of Chronic Kidney Disease Using Neuro- Fuzzy and Hierarchical Clustering Algorithm (s)", International Journal of Multimedia and Ubiquitous Engineering, vol. 12, no. 8, pp. 23-36, 2017.

[2] "Glomerular Filtration Rate (GFR)", National Kidney Foundation. [Online]. Available: https://www.kidney.org/atoz/content/gfr.

[3] J. V, L. Nayak and K. Dharmarajan, "A Survey o n Chronic Kidney Disease Detection Using Novel Methods", International Journal of Pure and Applied Mathematics, vol. 119, 2018

[4] N. Tazin, S. A. Sabab and M. T. Chowdhury, "Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique," 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec), Dhaka, 2016, pp. 1-6. doi: 10.1109/MEDITEC.2016.7835365

[5] C. Y. Ho et al., "Ultrasonography Image Analysis for Detection and Classification of Chronic Kidney Disease," 2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems, Palermo, 2012, pp. 624-629. doi: 10.1109/CISIS.2012.180

[6] D. S. Sisodia and A. Verma, "Prediction performance of individual and ensemble learners for chronic kidney disease," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 1027-1031. doi: 10.1109/ICICI.2017.8365295

[7] A. Al-Hyari, A. Al-Taee and M. Al-Taee, "Diagnosis and Classification of Chronic Renal Failure Utilising Intelligent Data Mining Classifiers", International Journal of Information Technology and Web Engineering, vol. 9, no. 4, pp. 1-12, 2014.

[8] Parul Sinha and Poonam Sinha, "Comparative Study of Chronic Kidney Disease Prediction using KNN and SVM", International Journal of Engineering Research and, vol. 4, no. 12, 2015.

[9] K. Eroğlu and T. Palabaş, "The impact on the classification performance of the combined use of different classification methods and different ensemble algorithms in chronic kidney disease detection," 2016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), Bursa, 2016, pp. 512-516.

[10] N. Tazin, S. A. Sabab and M. T. Chowdhury, "Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique," 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec), Dhaka, 2016, pp. 1-6. doi: 10.1109/MEDITEC.2016.7835365

[11] N. Tazin, S. A. Sabab and M. T. Chowdhury, "Diagnosis of Chronic Kidney Disease using effective classification and feature selection technique," 2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec), Dhaka, 2016, pp. 1-6. doi: 10.1109/MEDITEC.2016.7835365

[12] R. Ani, G. Sasi, U. R. Sankar and O. S. Deepa, "Decision support system for diagnosis and prediction of chronic renal failure using random subspace classification," 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, 2016, pp. 1287-1292. doi: 10.1109/ICACCI.2016.7732224

[13] K. Anantha Padmanaban and G. Parthiban, "Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease", Indian Journal of Science and Technology, vol. 9, no. 29, 2016.

[14] F. Barril Otero, "MYRA: a Java ant colony optimization framework for classification algorithms", 2017.

[15] Y. Cai, K. Feng, W. Lu and K. Chou, "Using LogitBoost classifier to predict protein structural classes", Journal of Theoretical Biology, vol. 238, no. 1, pp. 172-176, 2006.

[16] M. Dettling and P. Buhlmann, "Boosting for tumor classification with gene expression data", Bioinformatics, vol. 19, no. 9, pp. 1061-1069, 2003.

[17] W. Zhang, F. Zeng, X. Wu, X. Zhang and R. Jiang, "A Comparative Study of Ensemble Learning Approaches in the Classification of Breast Cancer Metastasis," 2009 International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, Shanghai, 2009, pp. 242-245. doi: 10.1109/IJCBS.2009.23

[18] M. V. Dass, M. A. Rasheed and M. M. Ali, "Classification of lung cancer subtypes by data mining technique," Proceedings of The 2014 International Conference on Control, Instrumentation, Energy and Communication (CIEC), Calcutta, 2014, pp. 558-562. doi: 10.1109/CIEC.2014.6959151

[19] K. R. Pradeep and N. C. Naveen, "Predictive analysis of diabetes using J48 algorithm of classification techniques," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, 2016, pp. 347-352. doi: 10.1109/IC3I.2016.7917987

[20] S. Sharma, S. Ghosh, N. Anantharaman and V. Jayaraman, Simultaneous Informative Gene Extraction and Cancer Classification Using ACO-AntMiner and ACO-Random Forests. Springer, 2012.

[21] O.O. Adeyemo, T. O Adeyeye, D. Ogunbiyi, "Comparative Study of ID3/C4.5 Decision tree and Multilayer Perceptron Algorithms for the Prediction of Typhoid Fever", African Joumal of Computing & ICT IEEE, 2015.

[22] J.R. Quinlan, "C4. 5: Programs for Machine Learning.", 1993.

[23] S. Perveen, M. Shahbaz, A. Guergachi, and K. Keshavjee, "Performance Analysis of Data Mining Classification Techniques to Predict Diabetes,"Procedia Comput. Sci., vol. 82, no. Supplement C, pp. 115–121, Jan. 2016.

[24] R. S. Parpinelli, H. S. Lopes and A. A. Freitas, "Data mining with an ant colony optimization algorithm," in IEEE Transactions on Evolutionary Computation, vol. 6, no. 4, pp. 321-332, Aug 2002. doi: 10.1109/TEVC.2002.802452

[25] J. Wu, W. Zhang and R. Jiang, "Comparative study of ensemble learning approaches in the identification of disease mutations," 2010 3rd International Conference on Biomedical Engineering and Informatics, Yantai, 2010, pp. 2306-2310. doi: 10.1109/BMEI.2010.5639753

[26] M. S. Wibawa, I. M. D. Maysanjaya and I. M. A. W. Putra, "Boosted classifier and features selection for enhancing chronic kidney disease diagnose," 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, 2017, pp. 1-6.

[27] Uthayakumar, D. Nivetha, D. Vinotha and M. Vasanthi, "Classification Rule Discovery Using Ant - Miner Algorithm: An Application Of Network Intrusion Detection", nternational Journal Of Modern E ngineering Research (IJMER), vol. 4, no. 8, 2014.

[28] G. Holmes, A. Donkin and I. H. Witten, "WEKA: a machine learning workbench," Intelligent Information Systems,1994. Proceedings of the 1994 Second Australian and New Zealand Conference on, Brisbane, Qld., 1994, pp. 357-361. doi: 10.1109/ANZIIS.1994.396988

[29] E. Avci, S. Karakus, O. Ozmen and D. Avci, "Performance comparison of some classifiers on Chronic Kidney Disease data," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, 2018, pp. 1-4. doi: 10.1109/ISDFS.2018.8355392

[30] R. K. Chiu, R. Y. Chen, Shin-An Wang and Sheng-Jen Jian, "Intelligent systems on the cloud for the early detection of chronic kidney disease," 2012 International Conference on Machine Learning and Cybernetics, Xian, 2012, pp. 1737-1742. doi: 10.1109/ICMLC.2012.6359637