Comparison of Different Decision Tree Algorithms for Predicting the Heart Disease

Deepak Saraswat(✉) and Preetvanti Singh

Department of Physics and Computer Science, Dayalbagh Educational Institute, Agra, India
deepakcsdei@gmail.com, preetvantisingh@gmail.com

Abstract. Data mining procedures are utilized to extract meaningful information for effective knowledge discovery. Decision tree, a classification method, is an efficient method for prediction. Seeing its importance, this paper compares decision tree algorithms to predict heart disease. The heart disease data sets are taken from Cleveland database, Hungarian database and Switzerland database to evaluate the performance measures. 60 data records for training and 50 data records for testing were taken as input for comparison. In order to evaluate the performance, fourteen attributes are considered to generate confusion matrices. The results exhibited that the algorithm that highest accuracy rates for predicting heart disease is Random forest, and thus can be considered as the best procedure for prediction.

Keywords: Heart disease · Classification technique · Decision tree · Decision tree algorithms · Performance measures

1 Introduction

The knowledge discovery process, now-a-days, has become more complex because of increasing size and complexity of the data sets. Data mining procedures are utilized to extract meaningful information for effective knowledge discovery. These procedures can be classified as descriptive procedures and predictive procedures. Descriptive procedures of data mining provide latest information on past or recent events, and for validating results, necessitate post-processing methods. Predictive procedures, on the other hand, predict the patterns and properties of vague information. Commonly used data mining procedures are Clustering, Classification, Association, Outlier Detection, Prediction, and Regression.

Classification is utilized for discovering knowledge based on different classes. It determines a model to describe and distinguish data classes based on trained data set, and identifies to which of the categories a new observation belongs to. Decision tree fosters a classification model in a tree-like structure. This type of mining, where the data set is distributed into smaller subsets and the associated Decision Tree (DT) is incrementally built, belongs to supervised class learning. The benefits of decision trees are:

A. Bhattacharjee et al. (Eds.): MIND 2020, CCIS 1241, pp. 245–255, 2020.
https://doi.org/10.1007/978-981-15-6318-8_21

- Easy integration due to intuitively representing the data,
- Investigative discovery of knowledge
- High accuracy
- Easily interpretable, and
- Excludes unimportant features

Because of above mentioned benefits, decision tree classifier is utilized for knowledge extraction in areas like education [33, 40], tourism [18, 34], healthcare [30, 31] and others. The healthcare industry creates colossal information from which it becomes extremely difficult to extract useful information. Decision tree is an efficient method for extracting effective knowledge from this titanic of information and providing reliable healthcare decision. It has been utilized in making effective decisions in various medical science areas like cancer detection, heart disease diagnosis and others [9, 11, 32, 45]. Presenting a brief overview of the algorithms for developing decision trees, and then comparing these algorithms for predicting heart disease based on performance measures is the foremost goal of this paper.

Heart diseases are a major source of death worldwide. As of 2016 there have been more than 17.6 million deaths per year. The death toll is expected to exceed 23.6 million by 2030 [3]. India too is witnessing shocking rise in the occurrence of heart disease (HD) [12]. Researchers have developed various decision tree algorithms to effectively diagnosis and treat heart diseases. Decision trees and rough set approach was utilized by Son, Kim, Kim, Park and Kim [41] to develop a model for heart failure. Chaurasia and Pal [8], Sa [35], and Amin, Chiam, and Varathan [4] developed a prediction system for HD by utilizing decision tree in combination with other data mining algorithms. Mathan, Kumar, Panchatcharam, Manogaran, and Varadharajan [22] presented forecast frameworks for heart diseases using decision tree classifiers. Wu, Badshah and Bhagwat [45] developed prediction model for HD survivability. Saxena, Johri, Deep and Sharma [37] developed a HD prediction system using KNN and Decision tree algorithm. Shekar, Chandra and Rao [38] developed a classifier to provide optimized feature for envisaging the type of HD using decision tree and genetic algorithm. Vallée, Petruescu, Kretz, Safar and Blacher [43] evaluated the role of APWV index in predicting HD. Pathak and Valan [29] proposed a forecasting model for HD diagnosis by integrating rule-based approach with decision tree. Sturts and Slotman [42] predicted risks for the patients who are re-admitted within 30 days after hospital discharge for CHF by using decision trees analysis. Seeing the importance of decision tree in healthcare, this paper presents a brief overview and comparison of seven DT algorithms based on various evaluation measures to diagnosis the heart disease.

2 Material and Method

2.1 Overview of Decision Tree Algorithms

Decision tree algorithm is a supervised learning method which is implemented on the basis of the data volume, available memory space and scalability, in serial or parallel style. The DT algorithms considered in this study are: J48, Decision stump, LMT, Hoeffding tree, Random forest, Random tree and REPTree. These are the most used algorithms for predicting various diseases (Table 1).

a. The J48 algorithm develops decision tree by classifying the class attribute based on the input elements.
b. The Hoeffding tree algorithm learns from huge data streams.
c. A Random tree algorithm draws a random tree from a set of possible trees and the distribution of trees is considered uniform.
d. A Random forest algorithm draws multiple decision trees using a bagging approach.
e. Logistic model tree (LMT) interprets combination of tree induction and linear logistic regression.
f. Decision stump builds simple binary decision stumps for both nominal and numeric classification task.
g. REPTree algorithm generates a regression or decision tree using information gain or variance.

Table 1. Different algorithms are applied in many areas.

S. No.	Authors name	Year	Algorithms	Areas
1	Vijiyarani and Sudha [44]	2013	Decision stump, Random forest, and LMT	Heart disease
2	Pandey, Pandey, Jaiswal and Sen [27]	2013	J48	
3	Chaurasia and Pal [7]	2014	J48	
4	Masethe and Masethe [20]	2014	J48 and REPTree	
5	Karabulut and Ibrikci [15]	2014	LMT	
6	Lohita, Sree, Poojitha, Devi and Umamakeswari [19]	2015	Random forest, J48, and REPTree	
7	Pachauri and Sharma [26]	2015	J48, Decision stump, and Random forest	
8	Bahrami and Shirvani [6]	2015	J48	
9	Kasar and Joshi [17]	2016	J48 and CART	
10	Alickovic and Subasi [2]	2016	Random forest	
11	Masetic and Subasi [21]	2016	Random forest	
12	Shrivas and Yadu [39]	2017	Decision stump	

(*continued*)

Table 1. (*continued*)

S. No.	Authors name	Year	Algorithms	Areas
13	Karthikeyan and Thangaraju [16]	2013	J48 and Random forest	Liver disorder
14	Novakovic and Veljovic [24]	2014	J48 and Decision stump	
15	Nahar and Ara [23]	2018	Decision stump, J48, REPTree, LMT, Random tree, Hoeffding tree, and Random forest	
16	Parimala and Porkodi [28]	2018	J48, LMT, Random tree, and REPTree	
17	Hasan, Bakar, Siraj, Sainin and Hasan [11]	2015	LMT, Random forest, and Random tree	Cancer
18	Azar, Elshazlyb, Hassanien and Elkorany [5]	2014	Random forest	Lymph diseases
19	Iyer, Jeyalatha and Sumbaly [13]	2015	J48	Diabetes
20	Perveen, Shahbaza, Guergachi and Keshavjeec [31]	2016	J48	
21	Alehegn, Joshi and Mulay [1]	2018	Decision stump	
22	Olayinka and Chiemeke [25]	2019	LMT, REPTree, Hoeffding tree, and J48	Malaria
23	Jena and Kamila [14]	2015	J48	Kidney disease
24	Gomathi and Narayani [10]	2018	Random forest, J48, and Hoeffding tree	Systemic Lupus Erythematous
25	Salih and Abraham [36]	2015	J48, LMT, Random forest, and Random tree	HD, Asthma, and Diabetes
26	Fatima and Pasha [9]	2017	J48	HD, Diabetes, Liver disease, and Dengue
27	Yang, Guo and Jin [46]	2018	J48, Decision stump, and Random tree	Cancer and Heart disease

2.2 Data Set

In order to attain the second goal of the present paper, three data sets from Cleveland database, Hungarian database and Switzerland database are considered for evaluating the performance measures of the DT algorithms. 60 data records for training and 50 data records for testing were taken as input for comparison. As shown in Table 2, fourteen attributes are considered for evaluating the performance measures.

Table 2. Description of the input attributes.

S. No.	Attributes	Description	Values
1.	Age	Years	Continuous
2.	Sex	M/F	1 if Male, 0 if Female
3.	CPT	Type of Chest pain	1 if Typical, 2 if Atypical angina, 3 if Non-angina pain, 4 if Asymptomatic
4.	RBP	Resting BP	Cont. (mm Hg)
5.	C	Cholesterol	Cont. (mm/dL)
6.	REG	Resting electrographic	0 = Normal, 1 = Abnormal, 2 = Probable
7.	BS	Blood sugar	True, when greater than or equal to 120 mg/dL and False, otherwise
8.	MHR	Maximum heart rate	Continuous
9.	EIA	Exercise induced angina	0 if no, 1 if yes
10.	OdPeak	Depression by exercise relative to rest	Cont.
11.	Slp	Slope	1 if unsloping, 2 if flat, 3 if downsloping
12.	Ca	Number of major vessels	Value (0–3)
13.	Tha	Type of Defect	3 if normal, 6 if fixed, 7 if reversible
14.	HDNum	The predicted attribute	0 = Heart Disease No, $1 \leq$ value ≤ 4 = Heart Disease Yes

3 Decision Tree Analysis

The performance measures are generated by using the information mining instrument Weka 3.9.3. Data pre-processing is done by means of the Replace Missing Values channel to filter all records and replace missing qualities. Next confusion matrices are developed by applying considered DT algorithms with 2 classes as Class 1 = YES (heart disease is present), and Class 2 = NO (heart disease not present), and True Positive = correct positive predicted; False Positive = incorrect positive predicted; True Negative = correct negative predicted; False Negative = incorrect negative predicted; P are Positive samples; and N are Negative samples.

These matrices are then utilized to compute the accuracy measures using the equations:

$$TPrate = \frac{TP}{TP + FP} \tag{1}$$

$$FPrate = \frac{FP}{FP + TN} \tag{2}$$

$$Accuracy = \frac{TP + TN}{P + N} \tag{3}$$

$$Errorrate = \frac{FP + FN}{P + N} \tag{4}$$

4 Results: Comparison

In Table 3 discussed the comparison of considered algorithms.

Table 3. Working comparison of decision tree algorithms.

Algorithms	Measure	Procedure	Pruning	Data type
J48	Information gain and Entropy	Top-down construction	Pre-pruning Single Pass Pruning Process	Discrete, Continuous, and can handle incomplete data
Hoeffding tree	Information gain and Hoeffding bound	Top-down construction	Pre-pruning	Data streams
Random tree	Hold-out set (back fitting)	Stochastic procedure	No Pruning	Discrete, Continuous, takes the input feature vector
Random forest	Receiver operating characteristic area under the curve	Class for constructing a forest of random trees	No pruning	Discrete, Continuous, takes the input feature vector
Logistic model tree	Logistic regression functions	Top-down	CART-based pruning	Numeric, Nominal, can contain missing values
Decision stump	Mean square error and Entropy	One-level decision tree	Post pruning	Discrete, Continuous, Binary
REPTree	Information gain	Top-down	Reduce error pruning with back-fitting	Binary, Numeric, Unary

Table 4 shows the computed performance measures using Eq. (3) and Eq. (4) for the data.

Table 4. Values of correctly classifier instances (CCI) and incorrectly classifier instances (ICI).

	Cleveland		Switzerland		Hungarian	
	CCI	ICI	CCI	ICI	CCI	ICI
Decision stump	81.167	18.833	76.167	23.833	81.833	28.167
J48	79.500	20.500	76.167	23.833	81.833	28.167
Hoeffding tree	86.167	13.833	83.333	16.667	86.167	13.833
LMT	96.667	3.333	76.167	23.833	75.500	22.000
Random forest	100.000	0.000	87.000	13.000	86.333	13.667
Random tree	100.000	0.000	82.000	18.000	79.333	19.000
REPTree	68.667	31.333	65.667	34.333	67.667	29.000

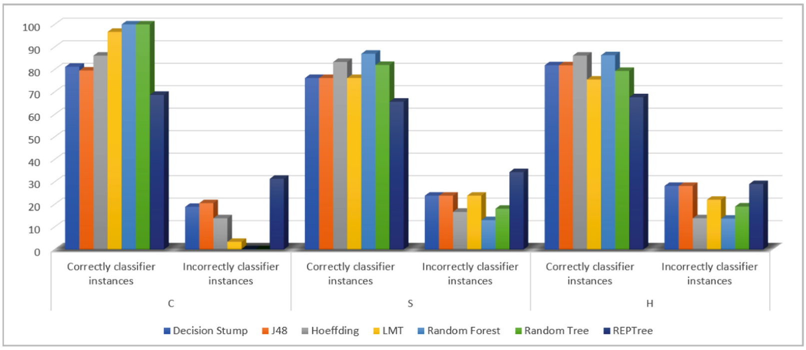

Fig. 1. Graph showing accuracy and error rate.

From the Fig. 1, it can be observed that for the considered data sets, Random Forest is showing max accuracy and least error rate.

Table 5. Class accuracy.

Algorithms	Class	Cleveland		Switzerland		Hungarian	
		TP rate	FP rate	TP rate	FP rate	TP rate	FP rate
J48	Yes	0.643	0.182	0.625	0.346	0.6	0.36
	No	0.818	0.357	0.654	0.375	0.64	0.4
Decision stump	Yes	0.786	0.182	0.625	0.396	0.6	0.36
	No	0.818	0.214	0.654	0.375	0.64	0.4
LMT	Yes	0.821	0.227	0.625	0.346	0.6	0.28
	No	0.773	0.179	0.654	0.375	0.72	0.4

(*continued*)

Table 5. (*continued*)

Algorithms	Class	Cleveland		Switzerland		Hungarian	
		TP rate	FP rate	TP rate	FP rate	TP rate	FP rate
Hoeffding tree	Yes	0.821	0.091	0.75	0.346	0.6	0.32
	No	0.909	0.179	0.654	0.25	0.68	0.4
REPTree	Yes	0	0	1	1	0.64	0.4
	No	1	1	0	0	0.6	0.36
Random tree	Yes	0.679	0.364	0.542	0.269	0.52	0.28
	No	0.636	0.321	0.731	0.458	0.72	0.48
Random forest	Yes	0.679	0.136	0.708	0.231	0.6	0.08
	No	0.864	0.321	0.769	0.292	0.92	0.4

From Table 5 it is clear that the TP Rate for the class = No is higher for Decision stump, Hoeffding tree, J48, Random forest, LMT and Random tree, which means the algorithms are successfully identifying the patients who do not have heart disease.

5 Conclusion and Future Scope

The primary goal of this paper was to compare most used decision tree algorithms and determine efficient method for predicting heart disease on the basis of computed performance measures Accuracy, True Positive Rate, Error rate and False Positive Rate. The algorithms considered in the study are Hoeffding tree, Decision stump, LMT, J48, Random tree, Random forest and REPTree were evaluated. From results it is clear that Random tree and Random forest are efficient method for generating decision tree. The reason for Radom forest being the best is, it splits on a sub set of a features and supports parallelism. Further the algorithm also supports high dimensionality, quick prediction, and outliners and non-linear data. However, the algorithm is less interpretable and can tend to over fit. In future, performance evaluation can be based on considering more attributes responsible for heart diseases. Other than healthcare, the framework can be utilized for evaluating performances in other domains also.

References

1. Alehegn, M., Joshi, R., Mulay, P.: Analysis and prediction of diabetes mellitus using machine learning algorithm. Int. J. Pure Appl. Math. **118**(9), 871–878 (2018)
2. Alickovic, E., Subasi, A.: Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. J. Med. Syst. **40**(4), 108 (2016). https://doi.org/10.1007/s10916-016-0467-8
3. American Heart Association. Heart disease and stroke statistics 2018 (2017). http://www.heart.org/idc/groups/ahamahpublic/@wcm/@sop/@smd/documents/downloadable/ucm_491265.Pdf

4. Amin, M.S., Chiam, Y.K., Varathan, K.D.: Identification of significant features and data mining techniques in predicting heart disease. Telematics Inform. **36**, 82–93 (2019). https://doi.org/10.1016/j.tele.2018.11.007
5. Azar, A.T., Elshazly, H.I., Hassanien, A.E., Elkorany, A.M.: A random forest classifier for lymph diseases. Comput. Methods Programs Biomed. **113**(2), 465–473 (2014). https://doi.org/10.1016/j.cmpb.2013.11.004
6. Bahrami, B., Shirvani, M.H.: Prediction and diagnosis of heart disease by data mining techniques. J. Multidisc. Eng. Sci. Technol. (JMEST). **2**(2), 164–168 (2015)
7. Chaurasia, V., Pal, S.: Data mining approach to detect heart diseases. Int. J. Adv. Comput. Sci. Inf. Technol. (IJACSIT). **2**, 56–66 (2014)
8. Chaurasia, V., Pal, S.: Early prediction of heart diseases using data mining techniques. Carib. J. Sci. Technol. **1**, 208–217 (2013)
9. Fatima, M., Pasha, M.: Survey of machine learning algorithms for disease diagnostic. J. Intell. Learn. Syst. Appl. **9**(1), 1 (2017). https://doi.org/10.4236/jilsa.2017.91001
10. Gomathi, S., Narayani, V.: Early prediction of systemic lupus erythematosus using hybrid K-Means J48 decision tree algorithm. Int. J. Eng. Technol. **7**(1), 28–32 (2018)
11. Hasan, M.R., Abu Bakar, N.A., Siraj, F., Sainin, M.S., Hasan, S.: Single decision tree classifiers' accuracy on medical data (2015)
12. https://timesofindia.indiatimes.com/india/heart-disease-deaths-rise-in-india-by-34-in-15-years/articleshow/64924601.cms
13. Iyer, A., Jeyalatha, S., Sumbaly, R.: Diagnosis of diabetes using classification mining techniques (2015). arXiv preprint arXiv:1502.03774, https://doi.org/10.5121/ijdkp.2015.5101
14. Jena, L., Kamila, N.K.: Distributed data mining classification algorithms for prediction of chronic-kidney-disease. Int. J. Emerg. Res. Manag. Technol. **4**(11), 110–118 (2015)
15. Karabulut, E.M., Ibrikci, T.: Effective automated prediction of vertebral column pathologies based on logistic model tree with SMOTE preprocessing. J. Med. Syst. **38**(5), 50 (2014). https://doi.org/10.1007/s10916-014-0050-0
16. Karthikeyan, T., Thangaraju, P.: Analysis of classification algorithms applied to hepatitis patients. Int. J. Comput. Appl. **62**(15), 25–30 (2013)
17. Kasar, S.L., Joshi, M.S.: Analysis of multi-lead ECG signals using decision tree algorithms. Int. J. Comput. Appl. **134**(16) (2016). https://doi.org/10.5120/ijca2016908206
18. Kuzey, C., Karaman, A.S., Akman, E.: Elucidating the impact of visa regimes: a decision tree analysis. Tourism Manag. Perspect. **29**, 148–156 (2019). https://doi.org/10.1016/j.tmp.2018.11.008
19. Lohita, K., Sree, A.A., Poojitha, D., Devi, T.R., Umamakeswari, A.: Performance analysis of various data mining techniques in the prediction of heart disease. Indian J. Sci. Technol. **8**(35), 1–7 (2015)
20. Masethe, H.D., Masethe, M.A.: Prediction of heart disease using classification algorithms. In: Proceedings of the World Congress on Engineering and Computer Science, vol. 2, pp. 22–24 (2014)
21. Masetic, Z., Subasi, A.: Congestive heart failure detection using random forest classifier. Comput. Methods Programs Biomed. **130**, 54–64 (2016). https://doi.org/10.1016/j.cmpb.2016.03.020
22. Mathan, K., Kumar, P.M., Panchatcharam, P., Manogaran, G., Varadharajan, R.: A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. Des. Autom. Embedded Syst. **22**(3), 225–242 (2018). https://doi.org/10.1007/s10617-018-9205-4

23. Nahar, N., Ara, F.: Liver disease prediction by using different decision tree techniques. Int. J. Data Min. Knowl. Manag. Process (IJDKP) **8**, 1–9 (2018). https://doi.org/10.5121/ijdkp.2018.8201
24. Novakovic, J.D., Veljovic, A.: Adaboost as classifier ensemble in classification problems. In: Proceedings Infoteh-Jahorina, pp. 616–620 (2014)
25. Olayinka, T.C., Chiemeke, S.C.: Predicting paediatric malaria occurrence using classification algorithm in data mining. J. Adv. Math. Comput. Sci. **31**(4), 1–10 (2019). https://doi.org/10.9734/jamcs/2019/v31i430118
26. Pachauri, G., Sharma, S.: Anomaly detection in medical wireless sensor networks using machine learning algorithms. Procedia Comput. Sci. **70**, 325–333 (2015). https://doi.org/10.1016/j.procs.2015.10.026
27. Pandey, A.K., Pandey, P., Jaiswal, K.L., Sen, A.K.: A heart disease prediction model using decision tree. IOSR J. Comput. Eng. (IOSR-JCE) **12**(6), 83–86 (2013)
28. Parimala, C., Porkodi, R.: Classification algorithms in data mining: a survey. Proc. Int. J. Sci. Res. Comput. Sci. **3**, 349–355 (2018)
29. Pathak, A.K., Arul Valan, J.: A predictive model for heart disease diagnosis using fuzzy logic and decision tree. In: Elçi, A., Sa, P.K., Modi, C.N., Olague, G., Sahoo, M.N., Bakshi, S. (eds.) Smart Computing Paradigms: New Progresses and Challenges. AISC, vol. 767, pp. 131–140. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9680-9_10
30. Paxton, R.J., et al.: An exploratory decision tree analysis to predict physical activity compliance rates in breast cancer survivors. Ethn. Health. **24**(7), 754–766 (2019). https://doi.org/10.1080/13557858.2017.1378805
31. Pei, D., Zhang, C., Quan, Y., Guo, Q.: Identification of potential type II diabetes in a Chinese population with a sensitive decision tree approach. J. Diabetes Res. (2019). https://doi.org/10.1155/2019/4248218
32. Perveen, S., Shahbaz, M., Guergachi, A., Keshavjee, K.: Performance analysis of data mining classification techniques to predict diabetes. Procedia Comput. Sci. **82**, 115–121 (2016). https://doi.org/10.1016/j.procs.2016.04.016
33. Rizvi, S., Rienties, B., Khoja, S.A.: The role of demographics in online learning; a decision tree based approach. Comput. Educ. **137**, 32–47 (2019). https://doi.org/10.1016/j.compedu.2019.04.001
34. Rondović, B., Djuričković, T., Kašćelan, L.: Drivers of E-business diffusion in tourism: a decision tree approach. J. Theor. Appl. Electron. Commer. Res. **14**(1), 30–50 (2019). https://doi.org/10.4067/S0718-18762019000100104
35. Sa, S.: Intelligent heart disease prediction system using data mining techniques. Int. J. Healthcare Biomed. Res. **1**, 94–101 (2013)
36. Salih, A.S.M., Abraham, A.: Intelligent decision support for real time health care monitoring system. In: Abraham, A., Krömer, P., Snasel, V. (eds.) Afro-European Conference for Industrial Advancement. AISC, vol. 334, pp. 183–192. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-13572-4_15
37. Saxena, R., Johri, A., Deep, V., Sharma, P.: Heart diseases prediction system using CHC-TSS evolutionary, KNN, and decision tree classification algorithm. In: Abraham, A., Dutta, P., Mandal, J., Bhattacharya, A., Dutta, S. (eds.) Emerging Technologies in Data Mining and Information Security, vol. 813, pp. 809–819. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-1498-8_71
38. Chandra Shekar, K., Chandra, P., Venugopala Rao, K.: An ensemble classifier characterized by genetic algorithm with decision tree for the prophecy of heart disease. In: Saini, H.S., Sayal, R., Govardhan, A., Buyya, R. (eds.) Innovations in Computer Science and Engineering. LNNS, vol. 74, pp. 9–15. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-7082-3_2

39. Shrivas, A.K., Yadu, R.K.: An effective prediction factors for coronary heart disease using data mining based classification technique. Int. J. Recent Innov. Trends Comput. Commun. **5** (5), 813–816 (2017)
40. Skrbinjek, V., Dermol, V.: Predicting students' satisfaction using a decision tree. Tert. Educ. Manag. **25**(2), 101–113 (2019). https://doi.org/10.1007/s11233-018-09018-5
41. Son, C.S., Kim, Y.N., Kim, H.S., Park, H.S., Kim, M.S.: Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. J. Biomed. Inform. **45**(5), 999–1008 (2012)
42. Sturts, A., Slotman, G.: Predischarge decision tree analysis predicts 30-day congestive heart failure readmission. Crit. Care Med. **48**(1), 116 (2020). https://doi.org/10.1097/01.ccm.0000619424.34362.bc
43. Vallée, A., Petruescu, L., Kretz, S., Safar, M.E., Blacher, J.: Added value of aortic pulse wave velocity index in a predictive diagnosis decision tree of coronary heart disease. Am. J. Hypertens. **32**(4), 375–383 (2019). https://doi.org/10.1093/ajh/hpz004
44. Vijiyarani, S., Sudha, S.: An efficient classification tree technique for heart disease prediction. In: International Conference on Research Trends in Computer Technologies (ICRTCT-2013) Proceedings published in International Journal of Computer Applications (IJCA), vol. 201, pp. 0975–8887 (2013)
45. Wu, C.S.M., Badshah, M., Bhagwat, V.: Heart disease prediction using data mining techniques. In: Proceedings of the 2019 2nd International Conference on Data Science and Information Technology, pp. 7–11 (2019). https://doi.org/10.1145/3352411.3352413
46. Yang, S., Guo, J.Z., Jin, J.W.: An improved Id3 algorithm for medical data classification. Comput. Electr. Eng. **65**, 474–487 (2018). https://doi.org/10.1016/j.compeleceng.2017.08.005