Machine Learning Methods in Early Diagnosis of Coronary Artery Disease

Luxmi Verma[1]*, Sangeet Srivastava[2] and Varun Sapra[3]

[1]Uttaranchal University, India

[2]The NorthCap University, India

[3]University of Petroleum and Energy Studies, India

*Corresponding author: Luxmi Verma, Uttaranchal University, Dehradun, India, Email: luxmi.verma@gmail.com

Abstract

With the advancement in information technology in health care sector, the medical hospitals are generating a huge volume of data including patient's details, medications, ECG etc. It was required to understand this large collection of data and deduce the meaningful patterns out of it. In the recent past, Data mining methods proved best for finding the hidden patterns out of the large data. These methods have been applied by researchers to transmute data into knowledge. One of the most life threatening disease worldwide is Coronary artery disease. It is one of the lifestyle diseases where cost of detection and treatment is high. One of the gold standards for diagnosis of disease is coronary angiography, which is invasive method and requires high end tests with a lot of technical expertise. So many researchers are seeking intelligent computation methods for disease detection. This review gives the insight of how computational methods are used in detection of Coronary Artery Disease.

Keywords: Coronary artery disease; Coronary angiography; Data mining

Abbreviations: CVD: Cardiovascular Diseases; CAD: Coronary Artery Disease; DM: Data Mining

Introduction

Health care sector is one of the most concerned sector for any country. One of the major chunk of the government budget is allocated for the research and advancement in this sector. At the same time information technology is the other booming sector and playing a support system to all major sectors specially the healthcare. With this advent, the health sector is producing complex medical data about patients including inter alia, clinical parameters, hospital resources, medical devices, disease diagnosis and patients' records. If this voluminous data is processed and understood properly, then it can assist medical practitioners to take better and early decisions regarding the disease. CAD needs to be detected early to avoid the risk of being exaggerated further. Many machine learning methods have been applied successfully by researchers to discover new and interesting patterns and detection of CAD. One of the prime features of machine learning is its capacity to construct models that are capable of generating knowledge, predicting the new unseen cases and produce patterns based on the learning from historical data. Some of the major application areas where machine learning algorithms made a mark are insurance frauds detection [1,2], infection control surveillance [3-6], disease diagnostic [7,8], treatment effectiveness and healthcare management [9-11].

CVD are becoming common with changing life styles across the world. They majorly happen due to various disorders of blood and heart vessels, and are foremost causes of disability and death [12,13]. With the availability of the clinical data of probable and potential patients, it becomes imperative to find suitable models to detect these diseases both economically and accurately. One of the most common CVD, that leads to cardiac arrest or myocardial infraction is due to accumulation of plaques in coronary arteries [14]. CAD patients can be diagnosed accurately using angiography which is an invasive, highly technical, painful and costly procedure.

Angiography can be risky and may lead to further aggravation of the disease and is not suitable for screening of larger population. Due to limitations of Coronary Artery Disease diagnostic methods, researchers are seeking other non-invasive methods that are less expensive, less risky, less complex, fast and easily reproducible. Various computation methods such as signal processing and image processing [15-18], Fuzzy logic [19-28], DM techniques [29] namely Decision Tree, Support Vector Machine, Neural Network, Logistic Regression techniques are used by researcher for diagnosis of CAD by using non invasive clinical parameters of subjects (Table 1). Various computational methods used by researchers for early identification of Coronary Artery Disease.

Table 1: Various computational methods used by researchers for early identification of Coronary Artery Disease.

References	Methodology	Application Summary
Karaolis et al. [18]	Decision tree	Identification of risk factors for event related to heart disease such as Percutaneous Coronary Intervention, Coronary Bypass Graft Surgery and Myocardial Infarction. Data collected from Paphos district hospital data
Nahar et al. [19]	Naïve Bayes, Sequential Minimal optimization, Ada boost, J48 and part, classifiers	Combined knowledge based features and correlation based feature subset selection to predict heart disease using UCI Machine repository data
Nahar et al. [20]	Apriori, Predicted Apriori and Tertius	Investigation of CAD risk factors in Males and Females using UCI Machine repository data
Aalizadehsani et al. [17]	Bagging, Sequential Minimal Optimization, Neural Network and Naïve Bayes	Investigation and construction of three features such as Left Anterior Descending (LAD), Left Circumflex (LCX), or Right Coronary Artery (RCA) form the existing features to determine the higher probability of CAD. The data set used for study is Z-Alizadeh Sani data with 54 attributes and 303 instances
Wiharito et al. [23]	Decision Tree	Dimensionality reduction with information gain, synthetic minority over sampling technique to solve class imbalance problem using UCI Machine repository data
Arabasadi et al. [21]	Genetic Algorithm and Neural Network	Analysis of dimensionality reduction techniques such as Gini Index, Information Gain, Principal Component Analysis and Weighted by SVM. Improvement in prediction accuracy with hybridization. Used Z-alizadesh Sani data for investigation
Tsipouras et al. [22]	Decision tree and fuzzy rules	Developed fuzzy rule based system for heart disease diagnosis
Anooj et al. [24]	Mamdani Fuzzy Inference System	Developed fuzzy rule based decision support for heart disease diagnosis using UCI Machine repository data.
Labsaba et al. [25]	Multi objective genetic algorithm, ensemble method	Fuzzy logic based system using UCI machine repository data
Pal et al. [26]	Rule organization method	Prototype of expert system for heart disease detection, Rules were guided by medical experts
Marateb et al. [27]	Neuro Fuzzy Classifier	Sequential feature selection and multiple logistic regression to reduce the feature space using UCI Machine repository data

These intelligent methods proved really effective in the diagnosis and treatment of diseases like diabetes, Parkinson's, cancer, thyroid, hepatitis, kidney disease and heart disease. The effectiveness of medical treatment is evaluated by comparing causes, symptoms and course of treatments and suggests the best practice procedure based on the produced knowledge and helps the medical practitioners to provide better and affordable healthcare services.
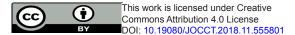
References

1. Bayerstadler A, van Dijk L, Winter F (2016) Bayesian multinomial latent variable modeling for fraud and abuse detection in health insurance. Insurance: Mathematics and Economics 71: 244-252.

2. Joudaki H, Rashidian A, Minaei-Bidgoli B, Mahmoodi M, Geraili B, et al. (2015) Improving fraud and abuse detection in general physician claims: a data mining study. International Journal of Health Policy and Management 5(3): 165-172.

3. Brossette S, Sprague A, Jones W, Moser S (2000) A data mining system for infection control surveillance. Methods of Information in Medicine 39(4-5): 303-310.

4. Obenshain M (2004) Application of data mining techniques to healthcare data. Infect Control Hosp Epidemiol 25(8): 690-695.

5. Ma L, Tsui F, Hogan WR, Wagner MM, Ma H (2003) A framework for infection control surveillance using association rules. In AMIA Annual Symposium proceedings, pp. 410-414.

6. Zikos D, Diomidous M (2012) Integration of data analysis methods in syndromic surveillance systems. Studies in health technology and informatics 180: 1114-1116.

7. Ordonez C (2006) Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine 10(2): 334-343.

8. Melillo P, Izzo R, Orrico A, Scala P, Attanasio M, et al. (2015) Automatic Prediction of Cardiovascular and Cerebrovascular Events Using Heart Rate Variability Analysis. Plos One 10(3): e0118504.

9. Dai W, Brisimi T, Adams W, Mela T, Saligrama V, et al. (2015) Prediction of hospitalization due to heart diseases by supervised learning methods. International Journal of Medical Informatics 84(3): 189-197.

10. Chih-Wen Cheng, Chanani N, Venugopalan J, Maher K, Wang M (2013) ICU ARM-an ICU clinical decision support system using association rule mining. IEEE Journal of Translational Engineering in Health and Medicine 1(1): 122-131.

11. Wong N (2014) Epidemiological studies of CHD and the evolution of preventive cardiology. Nature Reviews Cardiology 11(5): 276-289.

12. Berezin AE (2015) Stable coronary artery disease patients: different practice patterns in everyday clinical situations. EBioMedicine 2(11): 1576.

13. Acharya UR, Bhat PS, Iyengar SS, Rao A, Dua S (2003) Classification of heart rate data using artificial neural network and fuzzy equivalence relation. Pattern Recognition 36(1): 61-68.

14. Acharya R, Kannathal N, Krishnan SM (2004). Comprehensive analysis of cardiac health using heart rate signals. Physiological measurement 25(5): 1139-1151.

15. Acharya RU, Faust O, Alvin APC, Sree SV, Molinari F, et al. (2012) Symptomatic vs. asymptomatic plaque classification in carotid ultrasound. J Med Syst 36(3): 1861-1871.

16. Zhu KY, Ang A, Acharya R, Lim CM (2011) Human cardiovascular model and applications. J Med Syst 35(5): 885-894.

17. Alizadehsani R, Habibi J, Hosseini MJ, Mashayekhi H, Boghrati R, et al. (2013) A data mining approach for diagnosis of coronary artery disease. Comput Methods Prog Biomed 111(1): 52-61.

18. Karaolis MA, Moutiris JA, Hadjipanayi D, Pattichis CS (2010) Assessment of the risk factors of coronary heart events based on data mining with decision trees. IEEE Trans Inf Technol Biomed 14(3): 559-566.

19. Nahar J, Imam T, Tickle K, Chen Y (2013) Association rule mining to detect factors which contribute to heart disease in males and females. Expert Systems with Applications 40(4): 1086-1093.

20. Nahar J, Imam T, Tickle K, Chen Y (2013) Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. Expert Systems with Applications 40(1): 96-104.

21. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard A (2017) Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. Computer Methods and Programs in Biomedicine 141: 19-26.

22. Tsipouras M, Exarchos T, Fotiadis D, Kotsia A, Vakalis K, et al. (2008) Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling. IEEE Trans Inf Technol Biomed 12(4): 447-458.

23. Wiharto W, Kusnanto H, Herianto H (2016) Interpretation of clinical data based on c4.5 algorithm for the diagnosis of coronary heart disease. Healthc Inform Res 22(3): 186-195.

24. Anooj PK (2012) Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. Journal of King Saud University-Computer and Information Sciences 24(1): 27-40.

25. Lahsasna A, Ainon RN, Zainuddin R, Bulgiba A (2012) Design of a fuzzy-based decision support system for coronary heart disease diagnosis. J Med Syst 36(5): 3293-3306.

26. Pal D, Mandana K, Pal S, Sarkar D, Chakraborty C (2012) Fuzzy expert system approach for coronary artery disease screening using clinical parameters. Knowledge-Based Systems 36: 162-174.

27. Marateb HR, Goudarzi S (2015) A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system. J Res Med Sci 20(3): 214-223.

28. Ordonez C (2006) Association rule discovery with the train and test approach for heart disease prediction. IEEE Trans Inf Technol Biomed 10(2): 334-343.

29. Verma L, Srivastava S, Negi PC (2016) A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data. J Med Syst 40(7): 178.

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
  (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
https://juniperpublishers.com/online-submission.php