

# A Survey on Heart Disease prediction using Machine Learning Algorithms

Hasna V N<sup>1</sup>, Ms. Hrudya K P<sup>2</sup>

<sup>1</sup> M Tech, Computer science and Engineering, IES Engineering College, Kerala, India

<sup>2</sup> Assistant professor, Name science and Engineering, IES Engineering College, Kerala, India

## ABSTRACT

Heart diseases or cardio vessel Diseases (CVDs) unit the foremost reason for an enormous style of death among the globe over the previous vary of decades and has emerged as a result of the foremost dangerous condition, accurate and potential system to diagnose such diseases in time for proper treatment. Machine Learning algorithms and techniques have been applied to varied medical information sets to change the analysis of large and complex information. many researchers, in recent times, have been observe several machine learning techniques to help the health care business and along the professionals among the designation of heart connected diseases. During this paper presents a survey of assorted models supported such algorithms and techniques and analyse their performance. Models based on supervised learning algorithms like Support Vector Machines (SVM), K-Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and Logistic regression unit found very regarded among the researchers.

**Keyword:** - Supervised learning Algorithms, Support Vector Machines (SVM), K-Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and Logistic regression

## 1. INTRODUCTION

Heart is a vital organ of the organic structure. It pumps blood to a part of our anatomy. If it fails to perform properly, then the brain and varied different organs can shut down, and inside few minutes, the person can die. Modification in modus Vivendi, work connected stress and unhealthy food habits contribute to the rise in rate of several heart connected diseases.

Heart diseases have emerged jointly of the foremost distinguished cause of death all round the world. in step with World Health Organization, heart connected diseases square measure accountable for the taking 17.7 million lives once a year, thirty first of all international deaths. In India too, heart connected diseases became the leading reason behind mortality. Heart diseases have killed one.7 million Indians in 2016, in step with the 2016 international Burden of illness Report, released on Sept fifteen, 2017. Heart connected diseases increase the disbursal on health care Associate in nursing additionally scale back the productivity of an individual. Estimates created by the planet Health Organization (WHO), counsel that Asian country have lost up to \$237 billion, from 2005-2015, thanks to heart connected or vessel diseases. Thus, possible and correct prediction of heart connected diseases is very important.

Medical organizations, all round the world, collect knowledge on various health connected problems. This knowledge is exploited mistreatment various machine learning techniques to achieve helpful insights. But the data collected is incredibly large and, many another times, this knowledge will be terribly streaky. These datasets, that square measure too overwhelming for human minds to understand, is simply explored mistreatment varied machine learning techniques. Thus, these algorithms became very helpful, in recent times, to predict the presence or absence of heart connected diseases accurate.

Machine learning (ML) is that the sub domain of computing (AI). these days we have a tendency to are mistreatment millilitre in day-to-day life. Millilitre based mostly laptop programs will access knowledge and use it to find out themselves. It suggests that past expertise is employed for prediction in millilitre. Algorithms are of 4 types: supervised Learning during which direct supervising is concerned developer label the dataset restricts the

boundaries of formula, Unsupervised Learning supervising isn't needed, semi supervised machine learning each kind supervised and unsupervised utilized in combine format and Reinforcement Learning exploration of factor one by one menus event take as input for next event. During this paper the main target is on supervised machine learning algorithms. In supervised machine learning algorithms, because the name indicates there's presence of a supervisor World Health Organization train the machine for prediction.

In different words we have a tendency to train the machine with the assistance of a tagged dataset. tagged dataset is that the one that is labelled with the proper answer or category which might be tagged once foreseen by machine. In our prediction system, we have a tendency to are predicting the tagged knowledge into 2 categories: having a heart condition or not having a heart condition. heart condition dataset from UCI machine learning repository is employed. This dataset contains seventy-six numerous attributes and 303 instances. Attribute choice is a very important issue for the accuracy of result as a lot of relevant knowledge can predict results accurately. choice of attributes from the dataset desires correct domain information which will facilitate choose fewer attributes to predict results accurately. Out of the seventy-six attributes gift within the Cleveland heart condition dataset, fourteen attributes are chosen for prediction.

## 2. LITERATURE SURVEY

The literature survey consists of initial four papers surveyed from medical background. They are studied to determine the importance of various options and their typical values. Being from medical background, no machine learning algorithms is applied for the centre malady detection and therefore don't seem to be enclosed within the survey table.

The literature survey when these papers is from engineering background wherever totally different options are thought of and varied machine learning algorithms are applied for automatic detection of presence of heart condition. The numbers within the punctuation mark below corresponds to the numbers mentioned within the 'References'.

The table shows the different authors titled Paper and the accuracy.

SL NO	AUTHOR	METHOD	DESCRIPTION	ACCURACY
1	Purushottam, Kanak Saxena and Richa Sharma,	Efficient Heart Disease Prediction System	Classification rules generated by Decision tree algorithm	86.7%
2	Senthilkumar, Mohan	Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques	Combining the characteristics of Random Forest (RF) and Linear Method (LM)	88.7%
3	Ali, Liaqat	An Optimized Stacked SVM Based Expert system for the Effective Prediction of Heart Failure	In this method two models SVM model is linear and L1 regularized and SVM is L2 regularized	92.22%
4	Singh, Yeshvendra K	Heart Disease Prediction System Using Random Forest	Random Forest (RF) with crossvalidation	85.81%
5	Santhana Krishnan. J	Prediction of Heart Disease Using Machine Learning Algorithms	Two algorithms are used separately: Decision tree & Naive Bayes algorithm	Decision tree = 91% Naive Bayes =87%
6	Tang, X., Liu, Z., Li, T., Wu,	Expert Systems with Applications	The proposed model is thus useful as a decision support system for 32 heart disease	86%

	W. and Wei, Z		diagnosis	
7	Bigsby, K.G., Ohlmann, J.W. and Zhao, K	Design And Implementing Heart Disease Prediction Using Naives Bayesian.	The prevailing technique surpasses the Naive Bayes by yielding an accuracy of 89.77% in spite of reducing the attributes. AES yields in high security performance evaluation in comparison to PHEA (Parallel Homomorphic Encryption Algorithm).	89.77%
8	Bashir, Saba	Improving Heart Disease Prediction Using Feature Selection Approaches	The accuracy of Decision Tree is 82.22%, Logistic Regression 82.56%, Random Forest 84.17%, and Naïve Bayes 84.24% and Logistic Regression SVM is 84.85	DT : 82.22% LR : 82.56%
9	Chitra, R., and V.Seenivasagam.	Review of heart disease Prediction system using data mining and hybrid intelligent techniques	The computer aided heart disease prediction system helps the physician as a tool for heart disease diagnosis	87.56%
10	Rairikar, Abhishek	Heart disease prediction using data mining techniques	That KNN is best among all the classification techniques after we name prediction or classification of a nonlinear knowledge. KNN uses nearest neighbour to find the optimal solution	
11	Patel, Jaymin, Dr TejalUpadhyay, and Samir Patel	Heart Disease Prediction Using Machine learning and Data Mining Technique	Data Mining; Decision Support System Classification	55.77%
12	Anbarasi, M., E. Anupriya, and N. C. S. N. Iyengar	Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm	Decision Tree data mining technique outperforms other two data mining techniques after incorporating feature subset selection with relatively high model construction time	
13	Nishant, Gupta	Intelligent heart disease prediction in cloud environment through ensembling	The results have proved the minimum execution and maximum accuracy of the proposed model.	
14	Purushottam, Kanak Saxena and Richa Sharma,	Efficient Heart Disease Prediction System	Classification rules generated by Decision tree algorithm	86.7%

### 3. CONCLUSIONS

This paper aimed to analyse the appliance of Heart disease prediction in medical domain and a few of the algorithms went to predict diseases. its discovered that results might vary for various unwellness identification supported the tools and techniques used. Machine learning provides sensible leads to unwellness identification once applicable tools and techniques applied. Therefore knowledge in machine learning is that the promising field for care predictions.

### 4. ACKNOWLEDGEMENT

This paper and the research behind it would not have been possible without the exceptional support of my supervisor, Hrudya K P. I am also grateful for the insightful comments offered by my colleagues allowed me to continue my research with the book much longer than I could have hoped. Finally I would like to thank my college and university for the unconditional support for completing this work.

### 5. REFERENCES

- [1] Purushottam, Kanak Saxena and Richa Sharma, "Efficient heart disease prediction system." *Procedia Computer Science* 85 (2016): 962-969.
- [2] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective heart disease prediction using hybrid machine learning techniques." *IEEE Access* 7 (2019): 81542-81554.
- [3] Ali, Liaqat, et al, "An optimized stacked support vector machines based expert system for the effective prediction of heart failure." *IEEE Access* 7 (2019): 54007-54014.
- [4] Klyuchka, Y.A., Cherednichenko, O.Y., Vasylenko, A.V. and Yakovleva, O.V., 2017. Forecasting the results of football matches on the Internet based information. *Вісник Національного технічного університету ХПІ. Серія: Системний аналіз, управління та інформаційні технології*, (55), pp.51-59.
- [5] Tang, X., Liu, Z., Li, T., Wu, W. and Wei, Z., 2018, August. The Application of Decision Tree in the Prediction of Winning Team. In *2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)* (pp. 239-242). IEEE.
- [6] Bigsby, J.G., Ohlmann, J.W. and Zhao, K., 2019. The turf is always greener: Predicting decommitments in college football recruiting using Twitter data. *Decision Support Systems*, 116, pp.1-12.
- [7] Shin, J. and Gasparyan, R., 2014. A novel way to soccer match prediction. *Stanford University: Department of Computer Science*.
- [8] Dubitzky, W., Lopes, P., Davis, J. and Berrar, D., 2019. The open international soccer database for machine learning. *Machine Learning*, 108(1), pp.9-28.
- [9] Esumeh, E.O., 2015. Using machine learning to predict winners of football league for bookies. *Int. J. Artif. Intell.*, 5, p.22.
- [10] Hucaljuk, J. and Rakipović, A., 2011, May. Predicting football scores using machine learning techniques. In *2011 Proceedings of the 34th International Convention MIPRO* (pp. 1623-1627). IEEE.
- [11] Ljung, D., Carlsson, N. and Lambrix, P., 2018, September. Player pairs valuation in ice hockey. In *International Workshop on Machine Learning and Data Mining for Sports Analytics* (pp. 82-92). Springer, Cham.
- [12] Subbalakshmi, G., K. Ramesh, and M. Chinna Rao. "Decision support in heart disease prediction system using naive bayes." *Indian Journal of Computer Science and Engineering (IJCSE)* 2.2 (2011): 170-176.
- [13] Masethe, Hlaudi Daniel, and Mosima Anna Masethe. "Prediction of heart disease using classification algorithms." *Proceedings of the world Congress on Engineering and computer Science*. Vol. 2. 2014.
- [14] Saxena, Kanak, and Richa Sharma. "Efficient heart disease prediction system." *Procedia Computer Science* 85 (2016): 962-969.
- [15] Thomas, J., and R. Theresa Princy. "Human heart disease prediction system using data mining techniques." *2016 international conference on circuit, power and computing technologies (ICCPCT)*. IEEE, 2016.
- [16] Sultana, Marjia, Afrin Haider, and Mohammad Shorif Uddin. "Analysis of data mining techniques for heart disease prediction." *2016 3rd international conference on electrical engineering and information communication technology (ICEEICT)*. IEEE, 2016.