A Survey of Machine Learning on Heart Disease

Megala.R[1], Sandhiya M[2]
[1]*Assistant Professor, Department of Computer Science and Engineering, Karpagam College of Engineering, Coimbatore, India*
[2]*Assistant Professor, Department of Electronics and Communication Engineering, Karpagam College of Engineering, Coimbatore, India*

***Abstract*- Heart disease is the leading cause of death all over the world .The diagnostic system, designed for detecting and classifying heart diseases. The aim of this paper is how the various Machine Learning algorithms is used for detecting the heart disease. ML algorithms used in diagnosis of various diseases like diabetics, heart problems, cancer. Most frequently used machine learning algorithm is support vector machine and decision tree. Future work of this paper is to use neural network, ensemble algorithm for better accuracy and sensitivity etc.**

***Index terms*- Machine Learning, support vector machine, neural network, Supervised learning, Unsupervised learning, Reinforcement learning, Artificial neural networks**

I.INTRODUCTION

Machine learning allows software applications to become more accurate and predicting outcomes without being explicitly programmed. Machine Learning is used to identify the abnormalities at early stage. Machine learning use statistical analysis to predict an output. Machine learning is a subset of artificial intelligence which focuses on machine learning from their experience. Making predictions based on its experience.

- Representation
- Evaluation
- Optimization

Medical diagnosis is a critical task by different intelligent systems. Medical treatment starts with frequent monitoring, diagnosis, screening, and treatment. Machine learning used to produce a descriptive analysis of clinical features.

II. METHODS IN MACHINE LEARNING

1. SUPERVISED LEARNING

Train the machine using the input data which is labelled. It means some data is already tagged with the correct answer. Supervised learning (SL) algorithm analyses the labelled training data and produces a correct outcome. Supervised Learning algorithms are trained using a training set of data, based on which the results are predicted. To predict the output value for the given input vector. The final output can be a continuous value or a discrete value. Continuous value is for regression problem and the discrete value is for classification problems. The training data set contains the sample input and output values. Popular supervised learning algorithms are classification and regression. supervised learning is of two types,

1. Parametric models
2. Non-parametric models.

Parametric models, the predictive function is the combination of fixed number of parametric. First stage is the learning stage using the training data set. After this stage the training data can be discarded as the prediction for new input is depends upon the learned parameters. Linear regression and classification are some of the parametric models. Successful parametric model is the neural networks.

In non parametric models, the number of parameters is dependent on the training data set. The training data set is maintained for the prediction. The most commonly used non parametric models are support vector machines and nearest neighbour algorithm.

Supervised learning is categorised into two:

Classification: Categorize data into a desired and distinct number of classes.

Regression: Target prediction value based on independent variables.

2. UNSUPERVISED LEARNING

Information neither labelled nor classified Group the unsorted information according to similarities, patterns and differences without any prior training of data.
Unsupervised learning is categorised into two:
Clustering: Discover the inherent group of the data
Association: It describe large portions of data

3. SEMI-SUPERVISED
Semi supervised learning some of the data is labelled. Semi supervised Models can be categorized under four classes:

1. Generative model
2. Model where the decision boundary is in low-density region
3. Graph-based model
4. Two-step model

4. REINFORCEMENT LEARNING
Making decisions sequentially. Current input and the next input depends upon the previous input. Using evaluative feedback Reinforcement learning (RL) explore the given test data and find the correct output. The main features of Reinforcement learning are delayed reward and trial-and-error. Reinforcement learning algorithm follows the Markov Decision Process (MDP). It is intimated when the result is wrong, then the algorithm explores the possibilities to find the right result. This algorithm is mainly used in finance, robotics, inventory management

Types of Reinforcement:
1.Positive :
Event occurs due to a particular behaviour and increases the strength.
Advantages are:
i)Maximizes Performance
ii)Sustain Change for a long period of time
Disadvantages are:
i)Too much Reinforcement can lead to Semi-Supervised Machine Learning
ii)overload of states which can diminish the results

2.Negative :
Strengthening of a behaviour Advantages are:
i)Increases Behaviour
ii)Provide defiance to minimum standard of performance
Disadvantages are :
i)Provides enough to meet up the minimum behaviour

III. DIAGNOSIS OF HEART DISEASES BY USING DIFFERENT MACHINE LEARNING ALGORITHMS

1.BAYESIAN NETWORKS
It can be also called as Probabilistic directed or acyclic graphical model. Bayesian Network represents a set of variables and their conditional dependencies via a directed acyclic graph. Bayesian networks are predicting the likelihood that any one of several possible known causes was the contributing factor. Bayesian network represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases. They can be used for a wide range of tasks including prediction, anomaly detection, diagnostics etc .
Joint probability distribution of random variables A_0, A_1, .. A_n, denoted as P(A_0, A_1, …, A_n), is equal to P(A_1 | A_2, …, A_n) * P(A_2 | A_3, …, A_n) * … * P(A_n) by the chain rule of probability.

$$\mathrm{P}\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} \mathrm{P}\left(A_k \,\middle|\, \bigcap_{j=1}^{k-1} A_j\right)$$

conditional independence between two random variables, A and B, given another random variable, C, is equivalent to satisfying the following property: P(A,B|C) = P(A|C) * P(B|C). A and B are independent, as long as the value of C is known and fixed is that P(A|B,C) = P(A|C).

2.FUNCTIONAL TREE
A functional tree is a diagram showing the dependencies between the functions of a system. It breaks a problem into simpler parts. To study the use of functional nodes at different places and for different types of modelling, we introduce a simple unifying framework for multivariate tree learning. Framework combines a univariate decision tree with a linear function by constructive induction. When Multivariate decision nodes are built when growing the tree, while functional leaves are built when pruning the tree. Analysis tools are Bias-variance

decomposition of the error, cluster analysis, and learning curves.

3.NAVIE BAYES Naive Bayes is algorithm is based on predictive modelling. This classifier is a type of statistical classifier. Naive Bayes classifier has the minimum error rate, but may not be the case always. It assumes no dependency between attributes .Attempting to maximize the posterior probability in determining the class. This model is associated with two types of probabilities which can be calculated from the training data set directly:

a) The conditional probability of each class with each x value.

b) The probability of every class.

Bayesian classifier finds conditional probability of an instance of each class, based on the formula, and based on such conditional probability data, the instance is classified as the class with the highest conditional probability. If these probabilities are calculated, then the probabilistic model can be implemented to make predictions with new data using Naïve Bayes Theorem.

4.ARTIFICIAL NEURAL NETWORKS

Artificial neural networks also called Multilayer Perceptron. It is capable of modelling extremely complex non-linear functions. It consist of 3 layers of input, output and hidden layer. Hidden layer comprises units that transform the input to a pattern that the output layer manipulates. Artificial neural networks are excellent tools for the purpose of finding patterns that are so complex or ambiguous to a human programmer to extract and teach the machine how to recognize.

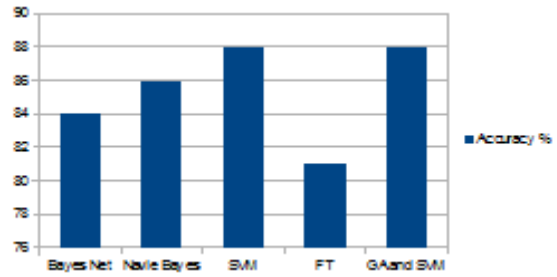
5.SUPPORT VECTOR MACHINE

SVM is used to analyse data and discover patters in classification and regression analysis. Support vector machine is used when your data has exactly two classes. An Support Vector Machine classifies data by finding the best hyper plane that separates all data points of one class from other class. A margin must have no points in its interior region. Support vector machine is based on mathematical functions used to model complex, and real world problems. It performs well on data sets that have many attributes, such as the CHDD Support Vector Machines map the training data into kernel space. There are many differently used kernel spaces – quadratic, polynomial, Radial Basis Function kernel, linear(uses dot product), Multilayer Perceptron kernel, etc.There are multiple methods of implementing SVM, such as quadratic programming, sequential minimal optimization, and least squares. SVM is kernel selection and method selection not over optimistic or pessimistic. Attributes and class labels relation ships are nonlinear, due to the large number of features, RBF kernel may not improve performance.

6.GENETIC ALGORITHM

Genetic Algorithm has the potential to recognize the optimal feature set. High resolution analysis of temporal, spatial, and spectral aspects of the data and granting for their interactions, leads to a very high dimensional feature space .Genetic algorithms are stochastic search algorithms. Genetic Algorithm loosely based on the mechanics of population genetics and selection. Pprobabilistic search methods that the states which they explore are not determined solely by the properties of the problems. GA are used in artificial intelligence like other search algorithms are used in artificial intelligence — to search a space of potential solutions to find one which solves the problem.

Machine Learning Algorithm	Accuracy %
Bayes Net	84
Navie Bayes	86
SVM	88
FT	81
GA and SVM	88



CONCULSION AND FUTURE WORK

This paper proposes to arrive at an accurate prediction of heart disease using combinations of SVM, Genetic Algorithm, Bayes Net etc. The classification of accuracy, sensitivity, and specificity of the SVM and GA have been found to be high thus making it a good option for the diagnosis. Future

work involves the development of a tool to predict the risk of heart disease . It can be extended for use on other models such as neural networks, ensemble algorithms, etc.

GLOSSARY

ANN	Artificial Neural Network
GA	Genetic Algorithm
ML	Machine Learning
SVM	Support Vector Machine
MDP	Markov Decision Process
RL	Reinforcement Learning
CHD	Coronary Heart Disease
RBF	Radial Basis Function
BN	Bayesian Network

REFERENCES

[1] Nashif, S., Raihan, Md.R., Islam, Md.R. and Imam, M.H. (2018) Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. World Journal of Engineering and Technology, 6, 854-873. https://doi.org/10.4236/wjet.2018.64057

[2] Elgendi, M., Liang, Y. and Ward, R. (2018) Toward Generating More Diagnostic Features from Photoplethysmogram Waveforms. Diseases, 6, https://doi.org/10.3390/diseases6010020[Citation Time(s):1]

[3] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," Neural and Evolutionary Computing, pp. 1-18, 2012

[4] M. C. Tu and D. Shin, "Effective diagnosis of heart disease through Bagging approach," The IEEE 2nd International Conference on Biomedical Engineering and Informatics,pp. 1-4, 2009.

[5] A Sandryhaila, JMF Moura, Big data analysis with signal processing on graphs: representation and processing of massive data sets with irregular structure. IEEE Signal Proc Mag 31(5), 80–90 (2014)

[6] K Slavakis, GB Giannakis, G Mateos, Modeling and optimization for big data analytics: (statistical) learning tools for our era of data deluge. IEEE Signal Proc Mag 31(5), 18–31 (2014)

[7] N. Cristianini and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, 2000.

[8] H. Drucker, D. Wu, and V. Vapnik. Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5), 1999.

[9] Animesh Hazra, Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee, "Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review", Research Gate Publications, July 2017, pp.2137-2

[10] Noura Ajam, "Heart Disease Diagnoses using Artificial Neural Network", The International Insitute of Science, Technology and Education, vol.5, No.4, 2015, pp.7-11.

[11] Prajakta Ghadge, Vrushali Girme, Kajal Kokane

[12] Dr.S.Seema Shedole, Kumari Deepika, "Predictive analytics to prevent and control chronic disease", https://www.researchgate.net/punlication/316530782, January 2016.

[13] Kanika Pahwa and Ravinder Kumar et al. "Prediction of Heart Disease Using Hybrid Technique For Selecting Features", 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON

[14] Saba Bashir, Usman Qamar, M.Younus Javed et al. "An Ensemble based Decision Support Framework for Intelligent Heart Disease Diagnosis" International Conference on Information Society (i-Society 2014).

[15] Shaikh Abdul Hannan, A.V. Mane, R. R. Manza,and R. J. Ramteke, Dec 2010,"Prediction of Heart Disease Medical Prescription using Radial Basis Function", IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), DOI: 10.1109/ICCIC.2010.5705900 ,28-29 .

[16] Sairabi H. Mujawar, and P. R. Devale, October 2015,"Prediction of Heart Disease using Modified k-means and by using Naive Bayes", International Journal of Innovative Research in Computer and Communication Engineering(An

ISO 3297: 2007 Certified Organization) Vol. 3, Issue 10, pp. 10265-10273.

[17] Kajaree Das1, Rabi Narayan Behera 2 "A Survey on Machine Learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 5, Issue 2, pp. 10265-10273.

[18] Francesco Musumeci, Member, IEEE, Cristina Rottondi, Member, IEEE, Avishek Nag, Member, IEEE, Irene Macaluso, Darko Zibar, Member, IEEE, Marco Ruffini, Senior Member, IEEE, and Massimo Tornatore, Senior Member, IEEE "An Overview on Application of Machine Learning Techniques in Optical Networks".