Prediction of Heart Attack using SMLT

E.Aloyses[1], C.Vidya Bhaskar[2]
[1, 2] Dept of Computer Engineering
[1, 2] DR.MGR Educational And Research Institute, Chennai, India

Abstract- *Heart is the main organ that pumps blood to the body and for proper functioning of body. Heart disease is a fatal human disease increasing globally in both developed and undeveloped countries and consequently, causes death. Normally, in this disease, the heart fails to supply a sufficient amount of blood to other parts of the body in order to accomplish their normal functionalities. It associates many risk factors in heart disease and a need of the time to get accurate, reliable, and sensible approaches to make an early diagnosis to achieve prompt management of the disease. Data mining is a commonly used technique for processing enormous data in the healthcare domain. Researchers apply several data mining and machine learning techniques to analyse huge complex medical data, helping healthcare professionals to predict heart disease. The proposed method is to build a machine learning model capable of classifying the person has heart disease or not. Different algorithms are compared and the best model is used for predicting the outcome*

I. INTRODUCTION

Data Science

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name for computer science. In 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic. However, the definition was still in flux.The term “data science” was first coined in 2008 by D.J. Patil, and Jeff Hammerbacher, the pioneer leads of data and analytics efforts at LinkedIn and Facebook. In less than a decade, it has become one of the hottest and most trending professions in the market.Data science is the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data.Data science can be defined as a blend of mathematics, business acumen, tools, algorithms and machine learning techniques, all of which help us in finding out the hidden insights or patterns from raw data which can be of major use in the formation of big business decisions.

Data Scientist:

Data scientists examine which questions need answering and where to find the related data. They have business acumen and analytical skills as well as the ability to mine, clean, and present data. Businesses use data scientists to source, manage, and analyze large amounts of unstructured data.

Required Skills for a Data Scientist:

Programming: Python, SQL, Scala, Java, R, MATLAB.
Machine Learning: Natural Language Processing, Classification, Clustering.
Data Visualization: Tableau, SAS, D3.js, Python, Java, R libraries.
Big data platforms: MongoDB, Oracle, Microsoft Azure, Cloudera.

II. ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.Artificial intelligence (AI) is intelligence demonstrated by machines, as opposed to the natural intelligence displayed by humans or animals. Leading AI textbooks define the field as the study of "intelligent agents" any system that perceives its environment and takes actions that maximize its chance of achieving its goals. Some popular accounts use the term "artificial intelligence" to describe machines that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving", however this definition is rejected by major AI researchers.Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing, and speech recognition and machine vision.AI applications include advanced web search engines, recommendation systems (used

by Youtube, Amazon and Netflix), Understanding human speech (such as Siri or Alexa), self-driving cars (e.g. Tesla), and competing at the highest level in strategic game systems (such as chess and Go), As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI effect. For instance, optical character recognition is frequently excluded from things considered to be AI, having become a routine technology.

Artificial intelligence was founded as an academic discipline in 1956, and in the years since has experienced several waves of optimism, followed by disappointment and the loss of funding (known as an "AI winter"), followed by new approaches, success and renewed funding. AI research has tried and discarded many different approaches during its lifetime, including simulating the brain, modeling human problem solving, formal logic, large databases of knowledge and imitating animal behavior. In the first decades of the 21st century, highly mathematical statistical machine learning has dominated the field, and this technique has proved highly successful, helping to solve many challenging problems throughout industry and academia.

The various sub-fields of AI research are centered on particular goals and the use of particular tools. The traditional goals of AI research include reasoning, knowledge representation, planning, learning, natural language processing, perception and the ability to move and manipulate objects. General intelligence (the ability to solve an arbitrary problem) is among the field's long-term goals. To solve these problems, AI researchers use versions of search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistics, probability and economics. AI also draws upon computer science, psychology, linguistics, philosophy, and many other fields.

The field was founded on the assumption that human intelligence "can be so precisely described that a machine can be made to simulate it". This raises philosophical arguments about the mind and the ethics of creating artificial beings endowed with human-like intelligence. These issues have been exploredby myth, fiction and philosophy since antiquity. Scien ce fiction and futurology have also suggested that, with its enormous potential and power, AI may become an existential risk to humanity.

As the hype around AI has accelerated, vendors have been scrambling to promote how their products and services use AI. Often what they refer to as AI is simply one component of AI, such as machine learning. AI requires a foundation of specialized hardware and software for writing and training machine learning algorithms. No one programming language is synonymous with AI, but a few, including Python, R and Java, are popular.

In general, AI systems work by ingesting large amounts of labeled training data, analyzing the data for correlations and patterns, and using these patterns to make predictions about future states. In this way, a chatbot that is fed examples of text chats can learn to produce life like exchanges with people, or an image recognition tool can learn to identify and describe objects in images by reviewing millions of examples.

AI programming focuses on three cognitive skills: learning, reasoning and self-correction.

Learning processes. This aspect of AI programming focuses on acquiring data and creating rules for how to turn the data into actionable information. The rules, which are called algorithms, provide computing devices with step-by-step instructions for how to complete a specific task.

Reasoning processes. This aspect of AI programming focuses on choosing the right algorithm to reach a desired outcome.

Self-correction processes. This aspect of AI programming is designed to continually fine-tune algorithms and ensure they provide the most accurate results possible.

AI is important because it can give enterprises insights into their operations that they may not have been aware of previously and because, in some cases, AI can perform tasks better than humans. Particularly when it comes to repetitive, detail-oriented tasks like analyzing large numbers of legal documents to ensure relevant fields are filled in properly, AI tools often complete jobs quickly and with relatively few errors.

Artificial neural networks and deep learning artificial intelligence technologies are quickly evolving, primarily because AI processes large amounts of data much faster and makes predictions more accurately than humanly possible.

Natural Language Processing (NLP):

Natural language processing (NLP) allows machines to read and understand human language. A sufficiently powerful natural language processing system would enable natural-language user interfaces and the acquisition of knowledge directly from human-written sources, such as newswire texts. Some straightforward applications of natural

language processing include information retrieval, text mining, question answering and machine translation. Many current approaches use word co-occurrence frequencies to construct syntactic representations of text. "Keyword spotting" strategies for search are popular and scalable but dumb; a search query for "dog" might only match documents with the literal word "dog" and miss a document with the word "poodle". "Lexical affinity" strategies use the occurrence of words such as "accident" to assess the sentiment of a document. Modern statistical NLP approaches can combine all these strategies as well as others, and often achieve acceptable accuracy at the page or paragraph level. Beyond semantic NLP, the ultimate goal of "narrative" NLP is to embody a full understanding of common sense reasoning. By 2019, transformer-based deep learning architectures could generate coherent text.

Existing System:

The Acoustic analysis has the potential to be a relatively low cost and non-invasive diagnostic tool for point-of-care assessment and remote monitoring of LVAD recipients. Prior work on acoustic analysis in LVAD recipients has focused on analysis of LVAD sounds to detect pump thrombosis. The analysis of heart sounds in LVAD recipients has been limited by the dominant LVAD sounds present within their precordial sounds. We have developed a novel signal processing pipeline to mitigate LVAD-generated sounds within precordial sounds recorded from LVAD recipients, potentially rendering automated heart sound analysis in this population feasible. Their analysis revealed that pump activity in LVAD recipients induces heart sounds with diverse acoustic signatures, such as S1 and S2 sounds with varying frequency ranges and relative amplitudes, and variabilities in S2 components based on changes in aortic valve mobility across different pump speeds.

System study:

Objectives

The goal is to develop a machine learning model for Heart Stroke Prediction, to potentially replace the updatable supervised machine learning classification models by predicting results in the form of best accuracy by comparing supervised algorithm.

Project Goals

Exploration data analysis of variable identification

- Loading the given dataset
- Import required libraries packages
- Analyze the general properties
- Find duplicate and missing values
- Checking unique and count values

Uni-variate data analysis

- Rename, add data and drop the data
- To specify data type

Exploration data analysis of bi-variate and multi-variate

- Plot diagram of pairplot, heatmap, bar chart and Histogram

Method of Outlier detection with feature engineering

- Pre-processing the given dataset
- Splitting the test and training dataset
- Comparing the Decision tree and Logistic regression model and random forest etc.

Comparing algorithm to predict the result

- Based on the best accuracy

Scope of the Project

Here the scope of the project is that integration of clinical decision support with computer-based patient records could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome. This suggestion is promising as data modeling and analysis tools, e.g., data mining, have the potential to generate a knowledge-rich environment which can help to significantly improve the quality of clinical decisions.

Literature survey:

1.REVIEW OF HEART DISEASE PREDICTION SYSTEM USING DATA MINING AND HYBRID INTELLIGENT TECHNIQUES
(R. Chitra1 and V. Seenivasagam2)
2.EFFECTIVE PREDICTION OF CARDIOVASCULAR DISEASE USING CLUSTER OF MACHINE LEARNING ALGORITHMS
(G. Jignesh Chowdary1 , Suganya. G 2 , Premalatha. M3)
3. A Survey on Prediction Techniques of Heart Disease using Machine Learning
(Mangesh Limbitote, Dnyaneshwari Mahajan, Kedar Damkondwar, Pushkar Patil)

4. Heart Disease Prediction Using Effective Machine Learning Techniques
(Avinash Golande, Pavan Kumar T)
5. **:**Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques
1Animesh Hazra, 2Subrata Kumar Mandal, 3Amit Gupta, 4Arkomita Mukherjee and 5Asmita Mukherjee

Proposed System:

The proposed method is to build a machine learning model for classification of heart disease. The process carries from data collection where the past data related to heart disease are collected. Data mining is a commonly used technique for processing enormous data in the healthcare domain. The heart disease if found before proper treatment can save lives. Machine learning is now applied and mostly used in health care where it reduces the manual effort and better model makes error less which leads in saving the life. The data analysis is done on the dataset proper variable identification done that is both the dependent variables and independent variables are found. Then proper machine learning algorithm are applied on the dataset where the pattern of data is learnt. After applying different algorithms a better algorithm is used for the prediction of outcome.

List of Modules:

- Data Pre-processing
- Data Analysis of Visualization
- Comparing Algorithm with prediction in the form of best accuracy result
- Deployment Using Flask

III. ALGORITHM AND TECHNIQUES

Algorithm Explanation

In machine learning and statistics, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation. This data set may simply be bi-class (like identifying whether the person is male or female or that the mail is spam or non-spam) or it may be multi-class too. Some examples of classification problems are: speech recognition, handwriting recognition, bio metric identification, document classification etc. In Supervised Learning, algorithms learn from labelled data. After understanding the data, the algorithm determines which label should be given to new data based on pattern and associating the patterns to the unlabelled new data.

Used Python Packages:

sklearn:

- In python, sklearn is a machine learning package which include a lot of ML algorithms.
- Here, we are using some of its modules like train_test_split, DecisionTreeClassifier or Logistic Regression and accuracy_score.

NumPy:

- It is a numeric python module which provides fast maths functions for calculations.
- It is used to read data in numpy arrays and for manipulation purpose.

Pandas:

- Used to read and write different files.
- Data manipulation can be done easily with data frames.

Matplotlib:

- Data visualization is a useful way to help with identify the patterns from given dataset.
- Data manipulation can be done easily with data frames.

IV. CONCLUSION

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set is higher accuracy score will be find out. This application can help to find the Prediction of Heart Attack.

V. FUTURE WORK

- Heart Attack prediction to connect with cloud.
- To optimize the work to implement in Artificial Intelligence environment.

REFFERENCES

[1] S. Agrawal et al., "Thirty-day readmissions after left ventricular assist device implantation in the United States," Circ. Heart Fail., vol. 11, no. 3, p. e004628, Mar. 2018.

[2] S. A. Akhter et al., "Hospital readmissions after continuous-flow left ventricular assist device implantation: incidence, causes, and cost analysis," Ann. Thorac. Surg., vol. 100, no. 3, pp. 884-9, Sep. 2015.
[3] M. R. Mehra et al., "Healthcare resource use and cost implications in the MOMENTUM 3 long-term outcome study," Circulation, vol. 138, no. 18, pp. 1923-1934, Jul. 2018.
[4] R. J. Gordon et al., "Prospective, multicenter study of ventricular assist device infections," Circulation, vol. 127, no. 6, pp. 691-702, Feb. 2013.
[5] P. Patel et al., "Acoustic signatures of left ventricular assist device thrombosis," ASME J. Med. Diagn., vol. 2, no. 2, May 2019, Art no. 024501.
[6] C. Feldmann et al., "An acoustic method for systematic ventricular assist device thrombus evaluation with a novel artificial thrombus model," J. Thorac. Dis., vol. 10, no. Suppl 15, pp. S1711-s1719, Jun. 2018.
[7] N. Reiss et al., "Telemonitoring and Medical Care of Heart Failure Patients Supported by Left Ventricular Assist Devices - The Medolution Project," Stud. Health. Technol. Inform., vol. 236, pp. 267-274, May 2017.
[8] D. Feldman et al., "The 2013 International Society for Heart and Lung Transplantation Guidelines for mechanical circulatory support: executive summary," J Heart Lung Transplant, vol. 32, no. 2, pp. 157-87, Feb. 2013.
[9] H. N. Sabbah and P. D. Stein, "Investigation of the theory and mechanism of the origin of the second heart sound," Circ. Res., vol. 39, no. 6, pp. 874- 882, Dec. 1976.
[10] G. E. Burch, "The Human Heart as a Pump," in Heart Perfusion, Energetics, and Ischemia, vol. 62, L. Dintenfass, D. G. Julian, and G. V. F. Seaman Eds., (NATO Advanced Science Institutes Series. Boston, MA, USA: Springer, 1983, pp. 81-95.
[11] M. R. Hodkiewicz and M. P. Norton, "The effect of change in flow rate on the vibration of double-suction centrifugal pumps," Proc. Inst. Mech. Eng., Part E: J. Process Mech. Eng., vol. 216, no. 1, pp. 47-58, Feb. 2002.
[12] F. Arvin et al., "Frequency shifting approach towards textual transcription of heartbeat sounds," Biol. Proced. Online, vol. 13, Oct. 2011, Art no. 7.
[13] A. Albraik et al., "Diagnosis of centrifugal pump faults using vibration methods," in J. Phys. Conf. Ser., Huddersfield, UK, Jun. 18-20 2012, vol. 364: IOP Publishing, p. 012139.
[14] J. K. Perloff, "Auscultation - The Audible Language of the Heart," in Physical Examination of the Heart and Circulation, 4th ed. Shelton, CT, USA: People's Medical Publishing House, 2009, ch. 6, pp. 157-223.
[15] Y. Chen, "Valvular Heart Disease," in Clinical Guide to Cardiology, C. F. Camm and A. J. Camm Eds. Hoboken, UK: John Wiley & Sons, Incorporated, 2016, ch. 14, pp. 194-196.
[16] A. K. Dwivedi et al., "Algorithms for automatic analysis and classification of heart sounds–a systematic review," IEEE Access, vol. 7, pp. 8316-8345, Dec. 2018.
[17] R. S. Gardner et al., "HeartLogic multisensor algorithm identifies patients during periods of significantly increased risk of heart failure events: results from the MultiSENSE study," Circ. Heart Fail., vol. 11, no. 7, p. e004669, Jul. 2018.
[18] C. N. Gupta et al., "Neural network classification of homomorphic segmented heart sounds," Appl. Soft Comput., vol. 7, no. 1, pp. 286-297, Jan. 2007.
[19] G. L. Yost et al., "Acoustic characterization of axial flow left ventricular assist device operation in vitro and in vivo," ASAIO J., vol. 62, no. 1, pp. 46- 55, Jan.-Feb. 2016. [20] F. Kaufmann et al., "Acoustic spectral analysis for determining pump thrombosis in rotary blood pumps," ASAIO J., vol. 60, no. 5, pp. 502-7, Sep.- Oct. 2014.