COMORBIDITY PROGNOSTICATION USING MACHINE LEARNING

Bhavana G[1], Prof Rajeshwari N [2]

Student, MCA Department, Bangalore Institute of Technology, Bangalore, Karnataka, India.[1]

Assistant professor, Department of MCA, Bangalore Institute of Technology, Bangalore, Karnataka, India[2]

Abstract: The inspiration is gathered for this project to relieve the burden of multiple diseases and shift the negatives into proactive healthcare strategies. Leaving in the edge of evolved technologies utilizing the advantages of health system that may help many people to reduce cost and increase their health profiles.

The project's scope is extensive, focusing on predicting the likelihood of various diseases in individuals based on patients unique health profiles. At the heart of this initiative are machine learning models, specifically assemblage technique such as Random Forests and Gradient Boosting, which utilize diverse data sources to provide precise and personalized risk assessments.

I. INTRODUCTION

The "Comorbidity Prognostication using machine learning" project work is an innovative and comprehensive platform that utilizes advanced machine learning and data analysis technologies to aid in the analysis and diagnosis of various diseases. This system is trained on data from multiple sources, such as medical records, patient histories, and clinical records, to enhance the quality and accuracy of diagnoses. Our diagnostic platform ensures patient diagnosis while maintaining data privacy and security through advanced technology.

Title "comorbidity" means one or more than one disease at a time and "prognostication refers to as prediction , hence detection one or more than one disease at a time using various ML algorithms which provides with the major accuracy. The method employed in this project involves the collection and processing of data from images, electronic medical records, and health histories using algorithms like Random Forest, XGBoost, and other training models. This approach integrates learning models into an easy-to-use platform.

The comorbidity prognostication system is an inventive and inclusive platform designed to facilitate the screening and to undergo diagnosis of various diseases, using advanced machine learning and data analytics techniques. The prognostication system trains to improve the efficiency and accuracy to perform medical diagnostics by using data sets from diverse sources such as medical profile, patient histories and their laboratory record.

The objective of comorbidity prognostication system is to develop an integrated diagnostic platform for patients' diagnosis. By using advanced machine learning , and to provide safety and security of data and privacy for the data sources collected from.

The projects scope is to focus on diseases such as diabetes, cardiovascular heart disease, breast cancer, brain tumor and pneumonia and to achieve high end accuracy level to get the required result prediction. In this project the methodology used is to collect and process the data from images, EHRs and lab records with the help of algorithms like CNN algorithm, random forest, XGBoost and other learning models.

Keywords: Machine Learning, Decision Tree, Adaboost, XGBoost, CatBoost, CNN, ML Technique.

II. PROBLEM STATEMENT

The objective is to create a machine learning model capable of accurately detecting various diseases, like diabetes, chronic heart disease, brain disease, breast cancer and pneumonia. This model will be trained using pertinent medical data, such as patient demographics, medical histories, symptoms, and diagnostic test results. The aim is to develop a system that facilitates early disease detection and diagnosis, which could enhance patient outcomes and lower healthcare costs. Additionally, the model must be robust and generalizable, maintaining high accuracy when applied to new patient data from diverse populations and geographic regions.

III.   SCOPE

This project aims to develop a comprehensive healthcare system designed to detect multiple diseases, including diabetes, chronic heart disease, brain disease, breast cancer. Utilizing advanced techniques in machine learning, data analytics, and medical imaging analysis, this system will support early disease identification, risk assessment, and personalized healthcare recommendations. The project will integrate various medical data sources and create user-friendly interfaces for both patients and healthcare professionals. Ultimately, the goal is to enhance public health, improve disease management, and facilitate early intervention and better healthcare outcomes for individuals worldwide.

IV.   LITERATURE SURVEY

The section examines the previously proposed models aimed at predicting diseases related to our research. Numerous studies have focused on detecting various diseases, employing different data mining techniques to predict them accurately.

One study, titled "Prediction of Diabetes Using Machine Learning Classification Techniques," utilized several classification algorithms, such as Random Forest and Convolutional Neural Networks (CNN), on datasets containing features like the number of pregnancies times, blood pressure, and body mass index. The study found that the Random Forest algorithm achieved the highest accuracy at 74%, while the lowest accuracy reported was 71.3%.

His work "Understanding the lifestyle of people to identify.." is been discussed for reducing the risk like diabetes using deep data mining techniques and discussed sub-classifications of diabetes.

The authors used Bayes algo and SVM classification algorithms for the collected dataset by a survey using google surveys and reported the accuracy and for SVM Naïve Bayes respectively 64.92 and 60.44.

Doctors find it hard to diagnose, especially in its recent life stages when patients have only mild cognitive problems. However, those patients are most likely to get benefit from treatment at this phase; that is why it makes sense to enhance the diagnostic process. In paper that we introduced and validates an approach of the machine-learning type that allows personalized and cost-effective diagnosis. Specifically, it employs locally weighted learning to adjust the classifying model for every individual and calculate, involve. The primary purpose of this paper is to establish a scientific base for anticipating the occurrence and spreading of hand, foot and mouth disease based on the affect of different meteorological conditions on the disease . Here the data collecting agents gathered the data on hand, foot and mouth disease and the weather from 2008-2010, then carried out correlation analysis, multiple linear regression, exponential curve fitting methods.

The paper emphasizes the crucial role of the heart in living organisms and the importance of accurate diagnosis and prediction of heart-related diseases to prevent fatalities. Machine learning and Artificial Intelligence were employed to predict heart disease using k-nearest neighbour decision tree, linear regression, and SVM algorithms with a UCI repository dataset for training and testing. The accuracies of these algorithms were compared: SVM (83%), Decision Tree (79%), Linear Regression (78%), and k-nearest neighbor (87%).

3. The system identifies liver disease as a major health concern in India, also considered a life-threatening disease globally. Due to the difficulty in detecting liver disease at an early stage, machine learning algorithms were utilized to analyze medical data and improve early detection and treatment outcomes.

V.   PROPOSED SYSTEM

The proposed system for multiple disease detection leverages machine learning to create a model that accurately identifies five diseases: diabetes, chronic heart disease, brain cancer, pneumonia, and other conditions based on symptoms, using the Naive Bayes algorithm. By analyzing patient data, including medical history, symptoms, and lab test results, the system will apply various machine learning techniques to identify patterns and predict the likelihood of diseases.

The objective is to provide clinicians with accurate and timely information, improving patient outcomes, reducing costs, and enhancing healthcare delivery. This will be achieved through integration with electronic medical records and a user-friendly interface, facilitating the overall process. The project will implement a section focused on detecting diseases based on symptoms.

Accuracy achieved for different diseases using different algorithms.

SN	DISEASE NAME	Algorithm	Accuracy
1	Prediction using symptoms	Navies bayes	70.5%
2	Diabetes	Random forest	66.8%
3	Heart disease	XGBoost	86.96%
4	Brain tumor	CNN	100%
5	Breast caner	Random forest	91.81%
6	Pneumonia	CNN	83.17%

Advantages:

1.   Early Detection and Timely Intervention:. Early detection often results for better outcomes, reduced healthcare costs, and improved quality of life for individuals with these diseases.

2.   Optimized Healthcare Delivery: By providing clinicians with accurate and timely information, the system can optimize healthcare delivery.. This can lead to more effective and personalized healthcare services

CNN (CONVOLUTION NEURAL NETWORK)
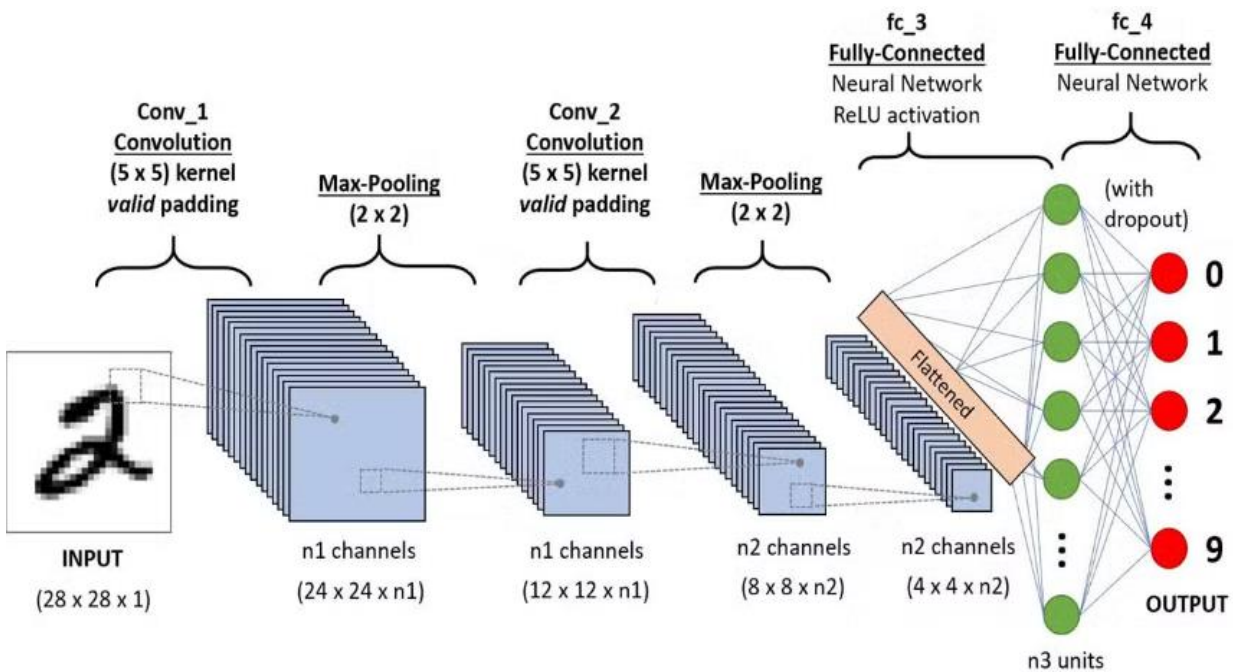
•   A Convolutional Neural Network (CNN) is a deep learning model designed to process input images, assigning significance through learnable weights and biases to different elements within the image, enabling it to distinguish between various objects. Unlike traditional classification algorithms, ConvNets require minimal pre-processing, as they can automatically learn and develop filters and features through training, rather than relying on manually engineered filters.
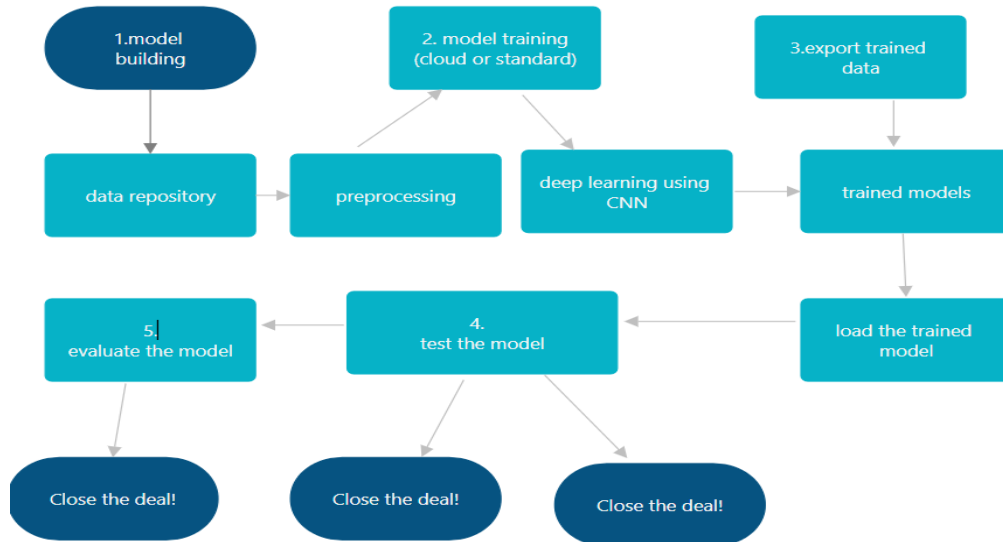
VI.   SYSTEM ARCHITECTURE

Fig : System Architecture

Use case diagram:
A use-case diagram is a representation that elaborates how various participants, such as users or external systems, interact with a system to achieve prediction accuracy or perform prediction tasks.

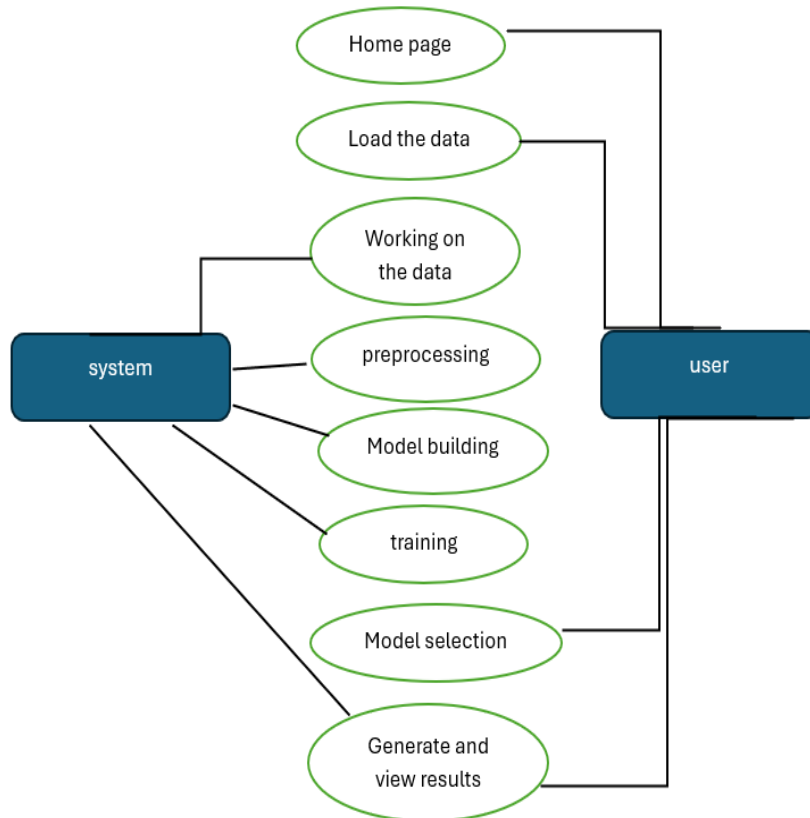Fig :use case diagram

S/W Configuration:

Software's   : Python 3.9 or high version

IDE   : PyCharm.

Framework   : Flask, pandas, NumPy and Scikit-Learn

Learning Outcome:

1. Regarding Machine Learning
2. Supervised Machine Learning
3. Classification Technique
4. About PyCharm
5. Flask Frame work

VII.   CONCLUSION

Incorporating Machine Learning into disease prediction could revolutionize healthcare by enhancing prediction accuracy, allowing for early interventions, supporting personalized medicine, improving resource allocation, and producing data-driven insights. To maximize the potential of Machine Learning in this field, continuous research, cooperation between healthcare professionals and data scientists, and adherence to ethical standards are essential.

This field comprised of multiple research in the both existing and proposed system in order decrease the bridging distance between patients and doctor and their health condition.

This build is used to pre-detect the disease that is suffering either through the symptom's or parameters of required features

VIII.   FUTURE ENHANCEMENTS

An overview of this comorbidity prognostication using machine learning has showed up with fast promising results and trusted efficiency and accuracy through predicting multiple diseases in several areas hence the futuristic enhancement includes

Developing the build model with the outstanding trusted result ,developing the chatbot system where the patients can have an interactions with the model in a user friendly manner.

Gathering the perfect data and need to reduce the ambiguity between choosing the model and data simultaneously.

Realtime monitoring and feedback of this application can include doctors monitoring system hence the patients also can trust it with full sincerity.

Integrating the drug and diet recommendation to any problem that is been detected in the same model so that diet and exercise plays a crucial role to control the early disease development in the patients lifestyle routine. Easy and user-friendly. We will also be storing the detection records.

Deploying it with IOT [internet of things] devices can help the economic benefits for the developed organization and gives the users authenticate prediction results

REFERENCES

[1]. J. L. Scully, "What is a disease?" EMBO Rep., vol. 5, no. 7, pp. 650–653, 2004.
[2]. R. Leaman, R. Islamaj Dogan, and Z. Lu, "DNorm: Disease name normalization with pairwise learning to rank," Bioinformatics, vol. 29, no. 22, pp. 2909–2917, Nov. 2013.
    https://www.slideshare.net/slideshow/multiple-disease-prediction-system/253060965
[3]. N. Armstrong and P. Hilton, "Doing diagnosis: Whether and how clinicians use a diagnostic tool of uncertain clinical utility," Social Sci. Med., vol. 120, pp. 208–214, Nov. 2014. From scholar paper.

[4]. P. Croft, D. G. Altman, and J. J. Deeks, ''The science of clinical practice: Disease diagnosis or patient prognosis? Evidence about 'what is likely to happen' should shape clinical practice,'' BMC Med., vol. 13, no. 1, p. 20, 2015. Google scholar.

[5]. E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. Sun, ''Doctor aiPredicting clinical events via recurrent neural networks,'' in Proc. Mach. LearnHealthcare Conf., 2016, pp. 301–318. From google scholar

[6]. C. C. Lee, ''Fuzzy logic in control systemsFuzzy logic controller. I,''IEEE Trans. Syst., Man, Cybern., vol. 20, no. 2, pp. 404–418, Mar./Apr. 1990. From IEEE scholar.

[7]. J. Yen and R. Langari,Fuzzy Logic: Intelligence, Control, and Information, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1999. Google scholar

[8]. H. D. Beale, H. B. Demuth, and M. Hagan, Neural Network Design. Boston, MA, USA: PWS, 1996.

[9]. PARSHANT1, DR. ANU RATHEE2:'' Multiple Diseasing Prediction Using ml'' Volume 6 , IRE Journals-2023 from google journals

[10]. ''An Approach to detect multiple diseases using machine learning algorithm'': Indukuri Mohit et al 2021

[11]. E3S Web of Conferences ''Feasible Prediction of Multiple Diseases using Machine Learning'': -2023

[12]. ''MULTIPLE DISEASE PREDICTION USING MACHINE LEARNING, DEEP LEARNING AND STREAM-LIT'', Mallula Venkatesh, e-ISSN: 2582-5208 . from google scholar.

[13]. ''Multi Disease Prediction System''. Divya Mandem1, B. Prajna- 2021, Volume 8

[14]. Naveen Kishore G,V .Rajesh ,A.Vamsi Akki Reddy, K.Sumedh,T.rajesh Sai Reddy, ''Prediction Of Diabetes Using Machine Learning Classification Algorithms'' .from google scholar.