**Health Genie – Disease Prediction and Health Monitoring System**

**by-** Zuber Khan, Sakshi Ajabe, Shruti Kalekar

**Abstract:** “Health Genie is an innovative health platform leveraging machine learning to predict diseases and monitor health status based on patient symptoms and specific medical parameters. With dedicated sections for prevalent conditions like diabetes, heart diseases, and lung cancer, the platform offers personalized risk assessments and actionable insights, empowering individuals to proactively manage their health and facilitating early intervention for improved outcomes in healthcare delivery.

Furthermore, Health Genie incorporates advanced machine learning algorithms trained on extensive datasets to continuously refine disease prediction models, ensuring adaptability to evolving healthcare scenarios. Through its user-friendly interface, individuals can input symptoms, medical history, and lifestyle factors to receive real-time feedback and personalized guidance for disease management. By bridging traditional healthcare approaches with cutting-edge technology, Health Genie represents a significant advancement in proactive health management, enabling early detection, personalized interventions, and ultimately, fostering better health outcomes for individuals and reducing healthcare burdens on society.

**Keywords:** Machine Learning • Streamlit • Decision Tree • Random Forest Classifier

.

# 1 Introduction

In the era of digital health solutions, the development of predictive and monitoring platforms has become imperative for enhancing healthcare delivery and ensuring early intervention. This research introduces Health Genie, an innovative health platform designed to predict diseases and monitor health status utilizing machine learning techniques. Health Genie utilizes patient symptoms as input data to predict various diseases with high accuracy. Additionally, it features dedicated sections for prevalent conditions such as diabetes, heart diseases, and lung cancer, leveraging specific medical parameters for precise prediction and monitoring.

The platform's predictive capabilities are facilitated by advanced machine learning algorithms trained on extensive datasets, enabling it to analyze symptom patterns and identify potential health risks. Furthermore, Health Genie incorporates comprehensive databases and expert knowledge to refine disease prediction models continuously, ensuring adaptability to evolving healthcare scenarios and emerging medical insights.

Through its user-friendly interface, Health Genie empowers individuals to proactively manage their health by providing personalized risk assessments and actionable insights. Users can input symptoms, medical history, and lifestyle factors, receiving real-time feedback on potential health concerns and preventive measures. Moreover, the platform offers tailored guidance for individuals diagnosed with specific conditions, facilitating personalized disease management strategies and promoting better health outcomes.

By bridging the gap between traditional healthcare approaches and cutting-edge technology, Health Genie represents a significant advancement in disease prediction and monitoring. Its implementation has the potential to revolutionize healthcare delivery by enabling early detection of diseases, proactive health management, and personalized interventions, ultimately leading to improved patient outcomes and reduced healthcare burdens..

Moreover, after successfully developing the disease prediction models, Health Genie aims to enhance accessibility and usability by implementing its machine learning algorithms into a user-friendly web application using Streamlit. This transition to a web-based platform will enable seamless interaction for users, allowing them to access personalized health insights and disease predictions conveniently from any device with internet connectivity. By integrating machine learning capabilities with a streamlined interface, Health Genie strives to democratize proactive health management, empowering individuals to make informed decisions about their well-being and fostering a culture of preventive healthcare.

# 2 Literature Survey

Compressive Strength Prediction using Cross Validation.[1]

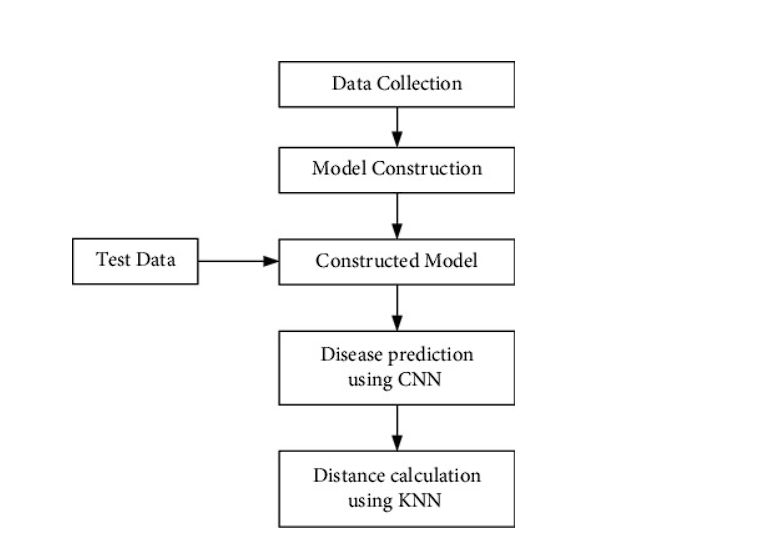
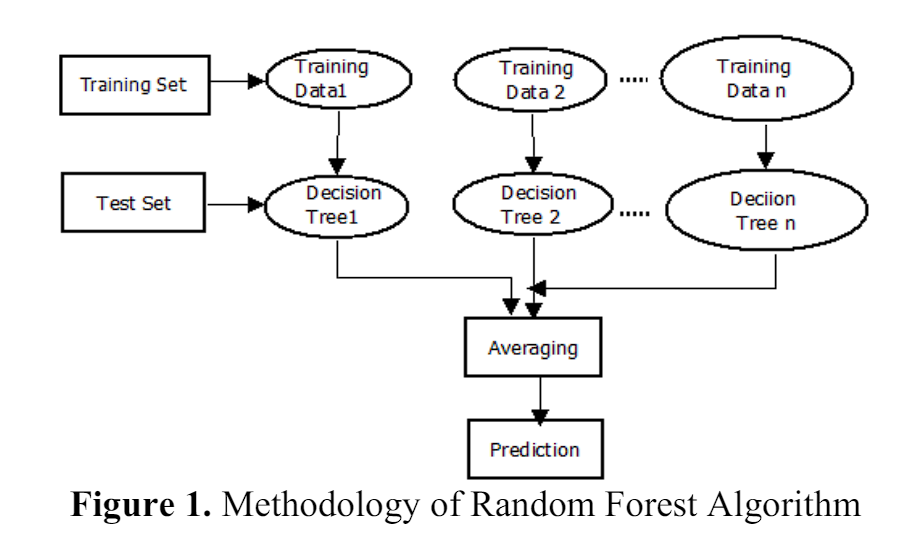
Nowadays, humans face various diseases due to the current environmental condition and their living habits. The identification and prediction of such diseases at their earlier stages are much important, so as to prevent the extremity of it. It is difficult for doctors to manually identify the diseases accurately most of the time. The goal of this paper is to identify and predict the patients with more common chronic illnesses. This could be achieved by using a cutting-edge machine learning technique to ensure that this categorization reliably identifies persons with chronic diseases. The prediction of diseases is also a challenging task. Hence, data mining plays a critical role in disease prediction. The proposed system offers a broad disease prognosis based on patient's symptoms by using the machine learning algorithms such as convolutional neural network (CNN) for automatic feature extraction and disease prediction and K-nearest neighbor (KNN) for distance calculation to find the exact match in the data set and the final disease prediction outcome. A collection of disease symptoms has been performed for the preparation of the data set along with the person's living habits, and details related to doctor consultations are taken into account in this general disease prediction. Finally, a comparative study of the proposed system with various algorithms such as Naïve Bayes, decision tree, and logistic regression has been d

Figure 1. Architecture of proposed disease and risk prediction system.

Human Disease Prediction using Machine Learning Techniques and Real-life Parameters.[3]

Human disease predication is a crucial part of human life. Early disease prediction of a human is an important step in the treatment of disease. Since the very beginning, a doctor has handled it almost exclusively. Thus, the healthcare industry thrives on innovation to make logistics efficient [1]. Innovation is the heart of the medical industry. It is what drives new treatments, cures and therapies [2]. Innovation is also what keeps the medical industry current and relevant. The scope of development in the medical industry is vast [3, 4]. There are many areas where innovation is needed to make progress. Some of these include developing new treatments for diseases, finding ways to improve patient care, and making medical procedures more efficient. In the current digital age, innovation in the medical industry can be achieved through the digitalization of medical \*Corresponding Author Email: mmkasar@bvucoep.edu.in (M. Kasar) processes [5]. One of the most pressing issues in the medical industry is the workload on the doctors [6] and the unaffordable consultation cost [7]. This issue is highlighted mainly in the disease prediction with the symptoms of the patients as input. The current methodology of the medical industry consists of the patient visiting a generalist doctor and explaining to the doctor the conditions, and symptoms faced by the patient upon which the doctor infers possible diseases and then channels them to a specialist doctor [8]. The logistics behind this methodology can be minimized with the help of a machine learning algorithm: Random Forest [9]. This algorithm is used for classifying multiple diseases based on symptoms and geographic locations. These locations help determine the results as the database assumes that for a particular location, there exist some symptoms that only occur at that location.

3. 1. Random Forest Algorithm The random forest produces decision trees from multiple data using their average for regression and most of the voting for categorization [21]. The research reported by Paul et al. [22] used the Random Forest Algorithm as the main algorithm. The random forest algorithm is used to train the model with the dataset which contains a combination of symptoms and the corresponding diseases [22]. The driving force behind using the random forest algorithm is that it has the capacity to handle data sets with continuous variables, as in regression, and categorical variables, as in classification [21, 23]. It produces superior results with regard to classification problems. The working method of the Random Forest is illustrated in Figure 1. Step 1: Select arbitrary samples from a given data set or training set. Step 2: This method will create a decision tree for every training data set. Step 3: Using the decision tree's average, voting will be done. Step 4: Lastly, select the predicted outcome that garnered the greatest support as the final prediction outcome. The Random Forest Algorithm analyses the symptoms and geographical region in the provided database to make judgments about a disease. Then it analyzes the outcome with the labels supplied before 

.

Disease Prediction From Various Symptoms Using Machine Learning..[4]

Accurate and on-time analysis of any healthrelated problem is important for the prevention and treatment of the illness. The traditional way of diagnosis may not be sufficient in the case of a serious ailment. Developing a medical diagnosis system based on machine learning (ML) algorithms for prediction of any disease can help in a more accurate diagnosis than the conventional method. We have designed a disease prediction system using multiple ML algorithms. The dataset used had more than 230 diseases for processing. Based on the symptoms, age, and gender of an individual, the diagnosis system gives the output as the disease that the individual might be suffering from. The weighted KNN algorithm gave the best results as compared to the other algorithms. The accuracy of the weighted KNN algorithm for the prediction was 93.5 %. Our diagnosis model can act as a doctor for the early diagnosis of a disease to ensure the treatment can take place on time and lives can be saved

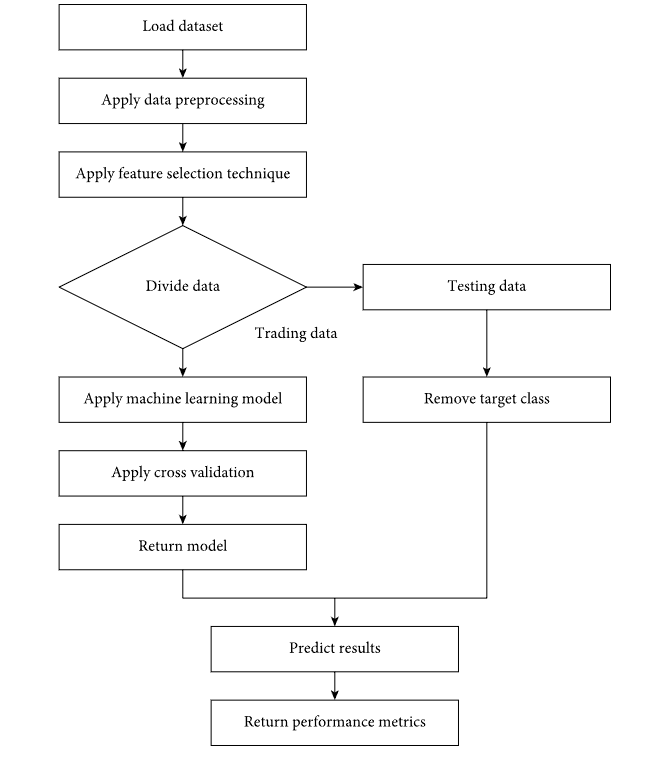


Figure 1. 3 Functioning of the ML models. The dataset was split into input consisting of age, gender, and symptoms and the output as the diseases based on the input factors. We randomly split the available data into train and test sets. These sets were then encoded and further trained using different algorithms. After which the algorithms test the training set and predict the values, resulting in the accuracy of different ML algorithms. The predicted values were then decoded to give the output as the disease

# 3. Findings and Discussions

According to the above literature we analyze numerous machine learning techniques for Health Genie- Disease Prediction and Health Monitoring System. Below are the findings we identified after this investigation:

# Significance of Disease Prediction: Literature emphasizes the critical need for early disease prediction to prevent the progression of illnesses and improve treatment outcomes, particularly in the context of chronic diseases.

# Machine Learning in Disease Prediction: Studies highlight the effectiveness of machine learning algorithms, such as decision trees, random forests, and weighted KNN, in accurately predicting diseases based on symptoms, age, gender, and other relevant factors.

# Health Genie Platform Overview: Health Genie is introduced as an innovative health platform that utilizes machine learning techniques to predict diseases and monitor health status. It leverages patient symptoms and specific medical parameters to provide personalized risk assessments and actionable insights.

# Dedicated Sections for Common Conditions: Health Genie features dedicated sections for prevalent conditions like diabetes, heart diseases, and lung cancer. This specialization enables precise prediction and monitoring tailored to specific medical conditions.

# Continuous Model Refinement: The platform continuously refines its disease prediction models using advanced machine learning algorithms trained on extensive datasets. This ensures adaptability to evolving healthcare scenarios and enhances prediction accuracy over time.

# User-Friendly Interface: Health Genie offers a user-friendly interface where individuals can input symptoms, medical history, and lifestyle factors to receive real-time feedback and personalized guidance for disease management.

# Transition to Web-Based Platform: Health Genie's transition to a web-based platform using Streamlit aims to enhance accessibility and usability for users. This shift enables seamless interaction and access to personalized health insights from any internet-enabled device.

# Democratizing Proactive Health Management: By bridging traditional healthcare approaches with cutting-edge technology, Health Genie democratizes proactive health management, empowering individuals to make informed decisions about their well-being and fostering a culture of preventive healthcare.

# Potential Impact on Healthcare Delivery: The findings suggest that platforms like Health Genie have the potential to revolutionize healthcare delivery by enabling early disease detection, personalized interventions, and ultimately, improving patient outcomes while reducing healthcare burdens on society.

# 4. Conclusion

Health Genie represents a groundbreaking advancement in healthcare, utilizing machine learning to predict diseases and monitor health status based on patient symptoms and medical parameters. With dedicated sections for prevalent conditions and a user-friendly web interface, it empowers individuals to proactively manage their health. By bridging traditional healthcare with innovative technology, Health Genie holds immense promise in revolutionizing healthcare delivery, fostering preventive healthcare practices, and ultimately improving patient outcomes. Further development and research in this area are crucial for maximizing its potential impact on healthcare systems and society.

**References:**

1. Rayan Alanazi.: " Identification and Prediction of Chronic Diseases Using Machine Learning Approach”

2. K. Gaurav, A. Kumar, P. Singh, A. Kumari, M. Kasar\*, T. Suryawanshi

Department of Computer Science and Engineering, Bharati Vidyapeeth Deemed to be University College of Engineering, Pune, India

3.

Rayan Alanazi.: "Predicting the Compressive Strength of Concrete using Neural Network and Kernel

Ridge Regression," (2016).

4.

Rinkal Keniya · Aman Khakharia · Vruddhi Shah · Vrushabh Gada ·

Ruchi Manjalkar · Tirth Thaker · Mahesh Warang · Ninad Mehendale ∗.: " Disease prediction from various symptoms using machine

learning," (2018).

