Predicting the Unpredictable: A General Disease Forecasting System

Ideation

Silicon Institute of Technology

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

May 9, 2023

Table of Contents

- Introduction
- How it Works
- Types
- Diabetes Prediction System
- Heart Disease Prediction System
- Breast Cancer Prediction System
- Lung Cancer Prediction System
- Benefits
- Challenges
- Future Developments
- Conclusion

Motivation

- Most hospitals today employ some sort of hospital information systems to manage their healthcare or patient data.
- These systems typically generate huge amounts of data which take the form of numbers,text, charts and images.
- Unfortunately, these data are rarely used to support clinical decision making. There is a wealth of hidden information in these data that is largely untapped. This raises an important question: "How can we turn data into useful information that can enable healthcare practitioners to make intelligent clinical decisions?" This is the main motivation for this research

Introduction

The General Disease Prediction System is an innovative technology that uses advanced artificial intelligence algorithms to predict the occurrence of diseases in individuals.

This system is designed to analyze various factors such as age, medical history, lifestyle habits, and environmental conditions to provide accurate predictions

How it Works

The General Disease Prediction System works by collecting and analyzing large amounts of data from various sources such as medical records, genetic testing, and environmental sensors. The system then uses machine learning algorithms to identify patterns and correlations between different variables to accurately predict the likelihood of developing certain diseases.
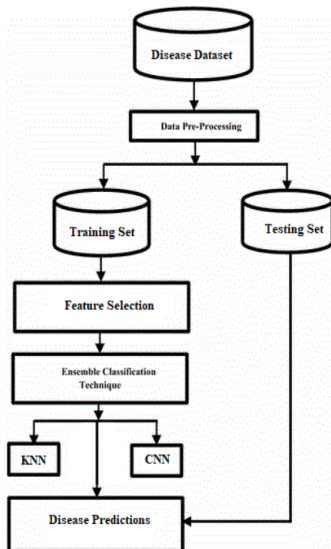
General Disease Prediction System

General Disease Prediction System

- Artificial Intelligence made computer more intelligent and can enable the computer to think.
- AI study consider machine learning as subfield in numerous research work. Different analysts feel that without learning, insight can't be created.
- There are numerous kinds of Machine Learning Techniques like Unsupervised, Semi Supervised, Supervised,Reinforcement, Evolutionary Learning and Deep Learning.
- These learnings are used to classify huge data very fastly. So we use K-Nearest Neighbor (KNN) and Convolutional neural network (CNN) machine learning algorithm for fast classification of big data and accurate prediction of disease.
- Because medical data is increasing day by day so usage of that for predicting correct disease is crucial task but processing big data is very crucial in general so data mining plays very important role and classification of large dataset using machine learning becomes so easy.

General Disease Prediction System Types Diabetes Prediction System Lung Cancer Prediction System Heart Disease Prediction System Breast

○○●   ○○   ○○○   ○○○   ○○○   ○○○○○

System Architecture

Types

Types of Prediction System

- Diabetes Prediction System
- Lung Cancer Prediction System
- Heart Disease Prediction System
- Breast Cancer Prediction System

Diabetes Prediction System

Diabetes Prediction System

- Diabetes is one of deadliest diseases in the world. It is not only a disease but also a creator of different kinds of diseases like heart attack, blindness, kidney diseases, etc.Diabetes Mellitus is defined as a group of metabolic disorders mainly caused by abnormal insulin secretion and/or action. Insulin deficiency results in elevated blood glucose levels (hyperglycemia) and impaired metabolism of carbohydrates, fat and proteins.

- There are two major clinical types :
  –Type 1 diabetes (T1D)
  –Type 2 diabetes (T2D)

- According to the etiopathology of the disorder. T2D appears to be the most common form of diabetes,mainly characterized by insulin resistance.

- The main causes of T2D include lifestyle, physical activity,dietary habits and heredity, whereas T1D is thought to be due to autoimmunological destruction of the Langerhans islets hosting pancreatic beta cells.

Working Principle

- Machine learning is the scientific field dealing with the ways in which machines learn from experience. The purpose of machine learning is the construction of computer systems that can adapt and learn from their experiences.
- We have developed a system using data mining which has the ability to predict whether the patient has diabetes or not.
- Furthermore, predicting the disease early leads to treating the patients before it becomes critical. Data mining has the ability to extract hidden knowledge from a huge amount of diabetes-related data.
- The aim of this research is to develop a system which can predict the diabetic risk level of a patient with a higher accuracy. This research has focused on developing a system based on three classification methods namely, Support Vector Machine, Logistic regression and Artificial Neural Network algorithms.

Lung Cancer Prediction System

Lung Cancer Prediction System

- Lung cancer is a harmful disease that causes a huge number of deaths globally.
- The primal encounter of lung cancer is necessary to decrease the mortality rate of patients. Thus it is a great challenge encountered by doctors and researchers to detect and diagnose lung cancer.
- Detection of lung cancer can be done by using medical images such as computed tomography, chest X-ray; MRI scans, etc., ML approaches recognize the main characteristics of complex lung cancer datasets.
- A CAD (Computer-Aided Diagnosis) was developed in the early 1980s to enhance the survival rate and efficiency that aid the doctors in interpreting medical images.
- Some of the machine learning algorithms that have a profound impact in health care are decision trees, linear regression, random forest,SVM, naive Bayes, K-nearest neighbors and so on.

Working Principle

- A prototype lung cancer disease prediction system is developed using data mining classification techniques.
- The system extracts hidden knowledge from a historical lung cancer disease database.
- The most effective model to predict patients with Lung cancer disease appears to be Naïve Bayes followed by IF-THEN rule, Decision Trees and Neural Network.
- Decision Trees results are easier to read and interpret.
- The drill through feature to access detailed patients' profiles is only available in Decision Trees

Heart Disease Prediction System

Heart Disease Prediction System

- Heart disease, alternatively known as cardiovascular disease, encases various conditions that impact the heart and is the primary basis of death worldwide over the span of the past few decades.
- 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.
- This project will be presenting various attributes related to heart disease, and the model on basis of supervised learning algorithms as Naïve Bayes, decision tree, K-nearest neighbor, and random forest algorithm.

Working Principle

- For predicting heart attack, significantly 15 attributes are listed and with basic data mining technique other approaches e.g. ANN,Time Series, Clustering and Association Rules, soft computing approaches etc. can also be incorporated.

- The outcome of predictive data mining technique on the same dataset reveals that Decision Tree outperforms and some time Bayesian classification is having similar accuracy as of decision tree but other predictive methods like KNN, Neural Networks, Classification based on clustering are not performing well.

- Decision Tree and Bayesian Classification further improves after applying genetic algorithm to reduce the actual data size to get the optimal subset of attribute sufficient for heart disease prediction.

Breast Cancer Prediction System

Breast Cancer Prediction System

- Most common type of medical hazard found in middle aged women is, breast cancer. Mortality rate of women due to breast cancer can be reduced if can be detected at a relatively early stage.
- With the help of latest, efficient and advanced screening methods, the majority of such cancers are diagnosed when the disease is still at a localized stage
- The utility of machine learning techniques in healthcare analysis is growing progressively. Certainly analysis of patient's clinical data are the most considerable features in diagnosis. Most of the possible medical flaws can be avoided by the using classification systems
- Accurate and timely prediction of breast cancer allows physicians and healthcare providers to make most favorable decision about the patient treatment

Proposed Framework

A classification model is proposed with boosted accuracy to predict the breast cancer patient. The framework is composed of the following important phases:

- Dataset Selection
- Data Preprocessing
- Learning by Classifier (Training) i.e. SVM, Linear Regression and KNN.
- Achieving trained model with highest accuracy
- Using trained model for prediction

Training and Classification

- Classifications of the data sets are done on the basis of specific properties posses by the sample variable that the capable to classify them, and each sample variable is assigned a malignant or benign class.
- Classification is principally done by making predictions based on known sample data that has been learned from training data.
- Designed algorithm is first trained on the known data labels and further uses this learning to predict the class labels for the new unknown set of data sample
- The classification objective set for this study is to achieve enhanced accuracy by using LR, SVM and KNN classifiers and determine which one suits the most for diabetes classification technique

K-Neighbors Classifier Pseudo code

KNeighborsClassifier(X,y, x) // X: training data, y: class labels of X, x: unknown sample
i=1
do
Calculate distance D(Xi, x)
while(i<=m)
Compute set Z having indices for the k smallest distances D(Xi, x).
Return label with {yi where $i \in$ Z}

Features

Features

- The general disease prediction system is equipped with various features that enable it to accurately predict the spread of diseases.
- One of the primary features of the system is its ability to analyze large amounts of data in real-time.
- This allows it to identify patterns and trends that can be used to predict future outbreaks of diseases.
- Another feature of the system is its ability to integrate with other healthcare systems, such as electronic health records and hospital information systems. This enables it to access a patient's medical history and genetic makeup, which can be used to predict their likelihood of developing certain diseases.
- In addition to these features, the general disease prediction system also has a user-friendly interface that makes it easy for healthcare professionals to use.
- The system provides visualizations and dashboards that allow users to quickly and easily interpret complex data.
- It also has built-in analytics tools that enable users to conduct further analysis and refine their predictions over time.

Applications

Applications

- The general disease prediction system has a wide range of applications in the healthcare industry. One of the primary applications is in disease surveillance, where it can be used to monitor and track the spread of infectious diseases. The system can also be used to predict outbreaks of diseases based on historical data and current trends.

- Another application of the system is in personalized medicine. By analyzing a patient's genetic makeup and medical history, the system can predict their likelihood of developing certain diseases and recommend personalized treatment plans. This can lead to better patient outcomes and more effective use of healthcare resources.

Future Developments

Future Developments

- As technology continues to advance, the General Disease Prediction System will likely become even more accurate and effective.
- New developments in wearable devices, genetic testing, and artificial intelligence will enable the system to provide even more personalized and precise prediction

Conclusion

Conclusion

- The General Disease Prediction System represents a major breakthrough in the field healthcare, offering numerous benefits for patients and healthcare providers alike.
- While there are still challenges to be addressed, the potential impact of this technology on disease prevention and treatment is immense

References

References

- General Disease Prediction System:
  https://sci-hub.se/10.1109/iccmc.2019.8819782
- Diabetes Prediction System : https://sci-hub.se/10.1007/978-3-319-41561-131
- Lung Cancer Prediction System :
  https://sci-hub.se/10.1109/icimia48430.2020.9074947
- Breast Cancer Prediction System : https://sci-hub.se/10.1016/j.procs.2018.05.197
- Heart Disease Prediction System :
  https://sci-hub.se/10.1109/aiccsa.2008.4493524