# Introduction to Old Models

- In the model proposed by [1] showed important ML approaches to predict the disease but this model which was proposed by [1] works on the K-Nearest Neighbour (KNN) and Convolution Neural Network (CNN) approach of the machine learning algorithm. Both the KNN and CNN approaches are used in this system which is different from the approach which is used in our project. The CNN uses both the structures as well as the unstructured data for the prediction of the disease which makes it more time consuming.

  The accuracy of the system proposed by [1] comes out to be very high i.e. above 95% for the KNN algorithm and 100% for the CNN algorithm that is very high for a ML model, In such cases the model is said to be overfitting.

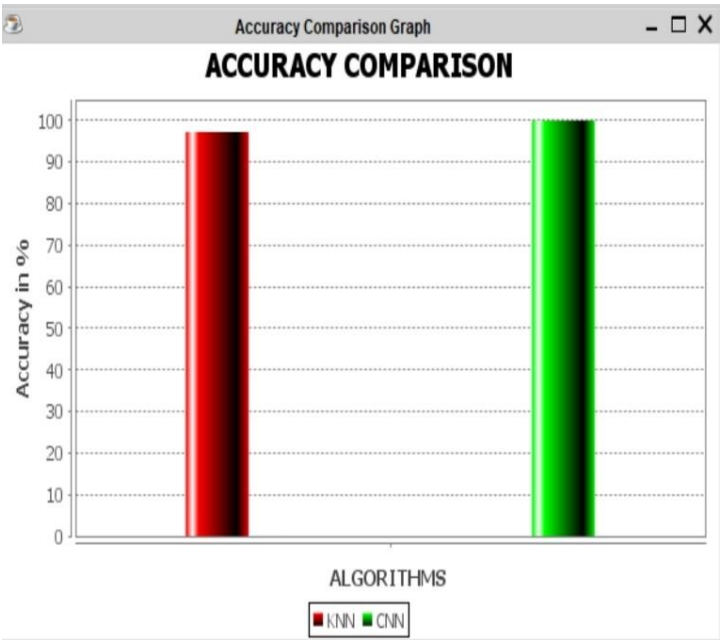


Figure 2.1: Accuracy Of The Model

- The model proposed by [2] is used for Disease Prediction and uses different ML algorithms like Iforest for correcting the dataset problems and SMOTET for balancing the dataset and then it uses the Ensemble learning technique. The Input the the ML model is

  taken only by the electronic reports which are produced by the blood examination of the patient or the user. Some of the input taken in this model are glucose level, cholesterol, lipoprotein, blood pressure and other inputs which are only be possible by the physical examination the user or the patient.

- The model proposed by [3] uses big data analytics and the deep learning models for the prediction of Disease The dataset is big so it uses the Big Data analytics like Map reduce is used in this model and on that the deep learning models are used for the prediction of the Disease which makes it a very big process and it becomes very time consuming. This model needs the full medical examination of the user or the patient foe the prediction

of the disease. Full medical history of the patient or the user is taken as an input to this model which is stored with the help of the big data tools and then used by the deep learning models to predict the disease. this model also needs all the medical record of the patient like all the medications which the patient or the user was taking and the list of doctors which he or she has visited which help in proper analysis of the patient's problem.

- In the model proposed by [4] uses different ML algorithms like Random Forest, Logistic Regression, Decision tree and others for the sake of prediction of the Disease and is used for the prediction of the Heart Disease, Breast Cancer and Diabetes. All the algorithms used in the system have their own way of predicting the Disease and are used accordingly. Different dataset are used in this model for the different disease like the heart disease has a different dataset and the Breast Disease has a different dataset. For the different dataset the algorithms have their different accuracy % accordingly and are used as per the accuracy.

- In the model proposed by [5] uses different data mining and the classification algorithms for the prediction of disease. This model is mainly used for the prediction of the Heart Disease and the algorithms which are used in this model are Decision Tree and the Naive Bayes algorithm which are used for the prediction of the Disease and various data mining techniques are also used in this model for correcting and balancing the dataset so that the system can work correctly and can predict the correct Disease. This model also needs the blood report of the patient or the user of the model. Some of the inputs which are used in this system are Cholesterol Level, Blood Pressure, Glucose Level in the body etc. This model has accuracy rate of 91% for the decision tree algorithm and 87% for the Naive Bayes algorithm but has a very limited scope in the prediction of the Diseases as it can only predict the Disease which are related to the Heart, Diabetes and Breast Cancer and cannot predict the general Diseases.

- In the model proposed by [6] makes use of Support Vector Machine (SVM) technique of the Machine Learning for the prediction of the Diseases. The dataset used in this model has dome general symptoms like eating habits, physical activity and they are rated in this model between 1-5 where 1 is for excellent and 5 is for very bad. This model helps to predict that weather a person's lifestyle is healthy or not and does he or she have any

disease or not, The model does not predict the name of the Disease or any problem whichthe patient is facing or not. The data from the user is collected by the means of a form and then is used by SVM for the prediction. This model is more focused on the lifestyleis the user that the user is active or not that how much physical work is he or she doing in day to day life and how much stress he or she have in life and on the basis of that thehealth of the user is predicted.

In the model proposed by [7] uses big data techniques for the disorders and helps inthe prediction of the disease like thyroid, chronic diseases. This model uses the MahoutHadoop technique of the big data analytics for the prediction of the disease Mahout has all the data mining techniques in it which makes the system efficient and powerful. In this model the Mahout part of the Hadoop system helps in the analysis of the data whichis stored in Hbase and on the basis of that the disease are predicted in this model. The size of the dataset is very large hence the overall system becomes very time consuming and the system requirements are also very high to run this system so it needs the very high and fast processing environment for its functioning and