1. Background

1.1. Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications.

[Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications - PMC](#)

Over the last few years, the application of ML in the healthcare domain has greatly improved the early diagnosis and control of chronic illnesses, particularly diabetes mellitus. Chronic hyperglycemia disease or diabetes has its risks in development of cardiovascular disease, nephropathy etc. It is important therefore that early diagnosis especially be made and managed effectively in order to reduce these risks.

In the present research, Butt et al. (2021) discuss the role of ML algorithms for the classification as well as prediction of diabetes. Utilizing the PIMA Indian Diabetes dataset, the researchers implemented three classifiers: Among them there are two tree-based algorithms: Random Forest (RF) and Gradient Boosting Machine (GBM); two deep-learning-based algorithms: Multilayer Perceptron (MLP) and Recurrent Neural Network (RNN); and two statistical-based algorithms: Logistic Regression (LR) and Decision Tree (DT). Their results show that MLP was the best classifier, with the highest diabetes classification accuracy of 86.08%. For the predictive analysis, Long Short-Term Memory (LSTM) networks were used with an overall prediction accuracy of 87.26. These outcomes suggest that the application of the ML models has the ability to improve the diagnosis and prediction of diabetes.

Thus, in terms of algorithms you may work on Logistic Regression model as well as K-Nearest Neighbors (KNN), and Naive Bayes. These algorithms are quite appropriate for problems of binary classification such as the prediction of diabetes. Logistic Regression is a probabilistic method of classification, while KNN is a simple direct instance-based learning and Naive Bayes is a probabilistic classifier using Bayes' theorem with stringent assumptions of independence. They will give a good idea of how to go about implementing them and comparing the results back to your dataset.

Therefore, there are huge implications of applying ML algorithms in the prediction of diabetes. Refereeing and evaluating the results of Logistic Regression, KNN and Naive Bayes classifiers and following guidelines for data preparation and model assessment, your work can provide valuable insights to this important field of healthcare analytics (Umair Muneer Butt, 2021).