	Bangladesh Army University of Science and Technology (BAUST)		
	Department of Computer Science and Engineering		
	Assignment #1, Winter 2023	Level-4	Term-I
	Course Code: CSE 4131	Course Title: Artificial Neural Networks and Fuzzy Systems	
	Submission Date:	CO Number: CO2	Full Marks: 15
ID:200101055		Name: Monisha Bakshi	

Diabetes Prediction using Machine Learning.

1. Introduction

Diabetes is a metabolic disease affecting a multitude of people worldwide.Its incidence rates are increasing alarmingly every year. Diabetes is a disease whereby blood sugar (glucose) is not metabolized in the body. This increases the glucose in the blood to alarmingly high levels. This is known by the name hyper glycemia. In this condition, body is unable to produce sufficient insulin. The other possibility is that body cannot respond to the produced insulin. Diabetes is incurable; it has to be controlled. A diabetic person can develop severe complications like nerve damage, heart attack, kidney failure and stroke.

2. Literature Review

2.1 Diabetes detection using deep learning algorithms

For the purpose of detecting diabetes, the suggested method makes use of a deep neural network.Three convolutional layers are followed by three fully linked layers in the proposed deep learning architecture. The binary cross-entropy loss function and the Adam optimizer were used to train the model. The performance of the model was evaluated using various metrics such as accuracy, precision, recall, and F1 score.

Result: The results showed that the proposed method achieved an accuracy of 83.69%, a precision of 69.49%, a recall of 70.39%, and an F1 score of 69.93%. According to these findings, the suggested approach can be a useful tool for the early detection of diabetes, which is essential for prompt medical intervention and better patient outcomes.

Limitations: Only 768 samples were used in the study, which may limit the generalizability of the findings to other datasets or populations. The study's failure to compare the performance of the suggested strategy to other ways of detecting diabetes limits our ability to evaluate how successful it is compared to other strategies.

2.2 A Multilayer Hybrid Machine Learning Model for Diabetes Detection

The principles of machine learning and its algorithms is used in order to detect the possibility of a diabetic patient based on their level of glucose concentration , insulin levels and other medically point of view required test reports. The basic diabetes detection model uses Bayesian classification machine learning algorithm, but even though the model is able to detect diabetes, the efficiency is not acceptable at all times because of the drawbacks of the single algorithm of the model.

Result: In this research study, our working material was diabetes data set with 768 instances which we first of all divided into two further sets of 700 instances for training and testing of individual algorithms and remaining 68 instances to be used as the live input for the proposed model. The findings revealed that the suggested model had an F1 score of 90.98%, an accuracy of 92.18%, a sensitivity of 89.62%, and a specificity of 93.28%. These findings suggest that the multilayer hybrid machine learning model that has been suggested can be a useful tool for the early identification of diabetes.

Limitations: Only used the Pima Indians Diabetes dataset, which might not reflect the broader population and contains missing values that were reconstructed using the mean value.

2.3 Artificial Breath Classification Using XGBoost Algorithm for Diabetes Detection

This paper presents a system for analyzing exhaled air with the use of various sensors. Breath simulations with acetone as a diabetes biomarker were performed using the proposed e-nose system. The XGBoost algorithm for diabetes detection based on artificial breath analysis is presented.

Result: The XGBoost model, a gradient boosting approach that can handle both linear and non-linear connections between features, is then trained and evaluated using the pre-processed data. The findings revealed that the suggested method had an F1 score of 91.9%, an accuracy of 92.6%, a sensitivity of 90.0%, and a specificity of 94.0%.

Limitations: The proposed method's main drawback is that it only uses one breath sample to diagnose diabetes, which may not be sufficient. Therefore, additional research is required to evaluate the generalizability and robustness of the suggested technique for diabetes detection using larger datasets and more varied samples.

2.4 Diabetes Mellitus Disease Prediction Using Machine Learning Classifiers with Oversampling and Feature Augmentation

This paper proposes an ML based approach for diabetes mellitus disease prediction. For diabetes prediction, many ML algorithms are compared and used in the proposed work, and finally the three ML classifiers providing the highest accuracy are determined: RF, GBM, and LGBM.

Result: The LGBM algorithm produces the highest accuracy (89.5%) for the Pima dataset and the same accuracy is also obtained when the RF classifier is used without data preprocessing. When the data is preprocessed, the accuracy obtained for Pima dataset is highest (92.5%) when LGBM algorithm is used. For the DMS dataset, the accuracy obtained is highest (95.27%) when RF algorithm is used without preprocessing, and the accuracy obtained after preprocessing is highest (98.99%) for LGBM.

Limitations: Used only single dataset for experiment. Another drawback is the dataset's asymmetry, which might have had an impact on how well the oversampling and feature augmentation strategies worked.

3. Result and Analysis

Compare the results of all four methods :

i. The “Diabetes detection using deep learning algorithms” method use a collection of clinical measurements is used to identify diabetic patients from non-diabetic individuals using deep learning algorithms.

ii. The “A Multilayer Hybrid Machine Learning Model for Diabetes Detection” method that integrates various methods to correctly categorize people with diabetes.

iii. The “Artificial Breath Classification Using XGBoost Algorithm for Diabetes Detection” method , the results have shown that the designed system based on the XGBoost algorithm is highly selective for acetone, even at low concentrations. Moreover, in comparison with other commonly used algorithms, it was shown that XGBoost exhibits the highest performance and recall.

iv. The “Diabetes Mellitus Disease Prediction Using Machine Learning Classifiers with Oversampling and Feature Augmentation” method that enhances the effectiveness of conventional classifiers in predicting diabetes by combining oversampling and feature augmentation.

4. Conclusion

Diabetes is a chronic illness with no known cure, only treatment options. Considerable part of human population is under the grip of diabetes which is incurable. If not managed well, diabetes can Conflict of interest.

References

[1] Anna Paleczek * , Dominik Grochala and Artur Rydosz - Artificial Breath Classification Using XGBoost Algorithm for Diabetes Detection

[2] Sahil Parab,∗, Piyush Rathod,∗∗, Durgesh Patil,∗∗∗, and Vishwanath Chikkareddi1,∗∗∗∗, Ramrao Adik Institute Of Technology, Navi Mumbai, India - A Multilayer Hybrid Machine Learning Model for Diabetes Detection

[3] B. Shamreen Ahamed , Meenakshi S. Arya, and Auxilia Osvin V. Nancy, College of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani Campus, No. 1, Jawaharlal Nehru Road, Vadapalani, Chennai, Tamil Nadu, India, MIT World Peace University, Pune, India - Diabetes Mellitus Disease Prediction Using Machine Learning Classifiers with Oversampling and Feature Augmentation

[4] Swapna G.∗, Vinayakumar R., Soman K.P. - Diabetes detection using deep learning algorithms