2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) | 979-8-3503-8592-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/CONECCT62155.2024.10677165

Explainable Artificial Intelligence for Prediction of Diabetes using Stacking Classifier

Aruna Devi B[1] and Karthik N[2]
Computer Science and Engineering
National Institute of Technology Puducherry
Karaikal, Puducherry, India
[1]arunaakshal@gmail.com
[2]nkarthikapce@gmail.com

***Abstract*— The escalating incidence of diabetes in globe has prompted the medical sector to explore innovative approaches aimed at enhancing their medical technologies. Integrating machine learning (ML) algorithms into clinical care can play a pivotal role in early diabetes detection, thus helping to mitigate the potential health complications associated with the condition. Moreover, the latest Explainable Artificial Intelligence (XAI) techniques have the potential to facilitate user understanding and trust in AI-driven decisions. This work proposes a method for the precise detection of diabetes through meticulous data preprocessing, the construction of an ensemble ML algorithm and the interpretation of the model's outcomes using XAI. Early detection of diabetes enables timely intervention through medication, dietary adjustments, and lifestyle modifications, leading to improved blood sugar regulation and reduced risk of diabetes-related complications. The proposed work uses preprocessing techniques like K Nearest Neighbors (KNN) imputation, One-Class Support Vector Machine (OCSVM) anomaly detection, Synthetic Minority Over-Sampling Technique and Edited Nearest Neighbour (SMOTE + ENN) data balancing technique, and ensemble model has KNN, Support Vector Machine (SVM), and eXtreme gradient boosting (XGB) as baseline models and Random Forest (RF) as meta classifier. This research underscores the importance of building a reliable model for diabetes prediction and interpreting the results using the Local Interpretable Model-Agnostic Explanation (LIME) technique. This work addresses challenges such as missing data, anomalies, data imbalance, and appropriate model selection, while highlighting the significance of comprehending the model's outcomes. The proposed ensemble model achieved an accuracy of 97%.**

Keywords—diabetes, missing values, imbalanced dataset, anomalies, ensemble model, XAI, LIME

I. INTRODUCTION

Glucose is a sugar that moves through the bloodstream and provides energy to the body's cells. Human body has the ability to produce glucose, and at the same time, glucose is derived from the food that is consumed. Insulin is a hormone secreted by the pancreas that helps your cells absorb glucose and use it as fuel. Glucose doesn't reach the cells for proper usage and stays in blood, when the pancreas doesn't secrete enough or any insulin or doesn't use insulin properly. This increase in blood glucose level in blood is called diabetes. Elevated blood glucose levels. Gradually can cause health problems related to heart, kidneys, feet and eyes. In 2021, an estimated 536.6 million people between the ages of 20 to 79 worldwide have diabetes, accounting for 10.5% of the global population. By 2045, it is projected that the number of individuals with diabetes in this age group will increase to 783.2 million, which would represent 12.2% of the global population [1]. Detecting diabetes early allows the patient to start treatment and reduce the health hazards. Healthcare practitioners can identify people who may be at risk of developing diabetes with the use of machine learning algorithms and patient data, and also assist in monitoring and managing blood sugar levels more effectively [2].

There has been an enormous amount of medical data recently due to the growing patient population, social, medical data, and sensed data. Such data may not always be complete and correct; there may be missing values, data imbalance and anomalies in one or more features due to human or sensor errors [3]. Building a trustworthy ML model becomes challenging if these issues are not appropriately addressed. Transparency and interpretability issues typically impede the adoption and dissemination of the best machine-learning models in clinical settings [3]. Improved human comprehension and trust in prediction and decisions driven by AI is possible with Explainable AI.

The different ML and XAI methods that have been used In the near past to identify and diagnose diabetes are covered in this section. Extreme Gradient Boosting (XGB) and six additional conventional classifiers were used by the authors in [4] to train the model, with an emphasis on imputing missing values. They used the Soft Voting Ensemble (SVE) technique, which weights each classifier's output equally in order to combine its results. The Pima Indian Diabetes Dataset, which contains information on people with and without diabetes, was used in the study. In [5] By examining datasets from several nations, this study advances the understanding of diabetes prediction. The findings reveal the extent to which ML algorithms in particular, boosting algorithms work for precisely predicting the onset of diabetes, the boosting algorithm includes Gradient Boost, CatBoost, AdaBoost and XGBoost. Basic models such as Random Forests and Decision Trees demonstrated satisfactory performance. The results of the study in [6] show that models built with random forest algorithms for classification have an accuracy that is 81.1% higher than models built with other algorithms. Since the model's accuracy was not up to par, the authors used ensemble learning techniques like bagging, boosting, and averaging. They discovered that the averaging method had the least amount of error compared to the other techniques. With the use of two machine learning techniques, Random Forest and Logistic Regression, the work in [7] aims to identify prospective cases of diabetes. The accuracy of the prediction is calculated for each algorithm in which random forest outperformed logistic regression. With the help of XAI, this study [8] has developed a diabetes indicator model that improves diabetes classification by taking into account both common and external diabetes-causing factors, such as age, diabetes, BMI, glucose, and insulin. The Random Forest Classifier, Decision Tree, Xg Boost, and SVM are all part of the ML model. In [9] the study suggests a methodical approach to enhance the comprehensibility of AI applications in healthcare. The proposed methodology's applicability is demonstrated through the use of several AI applications for type 2 diabetes diagnosis. An artificial neural network is used for classification and regression while employing a local interpretable model-agnostic explanation, which provides a more direct and comprehensible approach. AI can be applied to the diagnosis of diabetes in essentially two ways. Treating it as a binary classification problem is one approach. Users are categorized as either yes (diabetic) or no. The other creates a forecasting model to estimate the likelihood that a user has diabetes [9]. Using a dataset collected from Sylhet Hospital, the authors of [10] compared several machine learning algorithms to find the most important feature for diabetes prediction. These algorithms included RF, Decision Tree, Artificial Neural Networks (ANN), KNN, Support Vector Machine, and XGB. They also employed feature attribution using SHAP.

A. *Research Motivation*

Availability of good quality data is the main challenge of ML models for making accurate predictions. The majority of real-world datasets often necessitate more information and may encompass anonymized data. These characteristics can have a deleterious effect on the reliability of machine learning models constructed using them. These model predictions may lead to inaccurate conclusions and put the safety of medical decisions at risk. ML models often function as black box models which make medical professionals difficult to understand the results of ML models. Hence, interpreting the outcomes using XAI could make the ML model trustworthy.

B. *Research Question*

In this work, we concentrate on the following research questions.

RQ1: How does ML-based preprocessing like imputation, anomaly detection, and balancing dataset impact the performance of diabetes detection model?

RQ2: How does XAI help to build a trustworthy diabetes prediction model?

C. *Problem Statement*

To develop a trustworthy diabetes detection model which involves precise preprocessing, building an ensemble model, and interpreting the results for better comprehension.

The structure of the work is as follows. In Section II, explains the proposed work. The preprocessing strategy used is described in Section III, and the methodology and evaluation metrics are described in Section IV. XAI is explained in Section V. In Section VI, the dataset is described and the results are reported. A thorough understanding of the measured parameters and their importance is provided. Finally, Section VII lays out the main conclusions and directions for further research.

II. PROPOSED WORK

The proposed work has three main objectives. The proposed workflow is displayed in Fig. 1. The first objective is to preprocess the medical dataset. The preprocessing includes ML-based imputation followed by anomaly detection and removal of anomalies, and dataset balancing. This work utilizes the KNN imputation, OCSVM anomaly detection and SMOTE+ENN dataset balancing technique. After properly preprocessing the dataset, the work chooses the best-performing ML models to make ensemble model for prediction. To ensure the reliability of the diabetes prediction model, this work uses (XAI) to analyse and explain the outcomes.

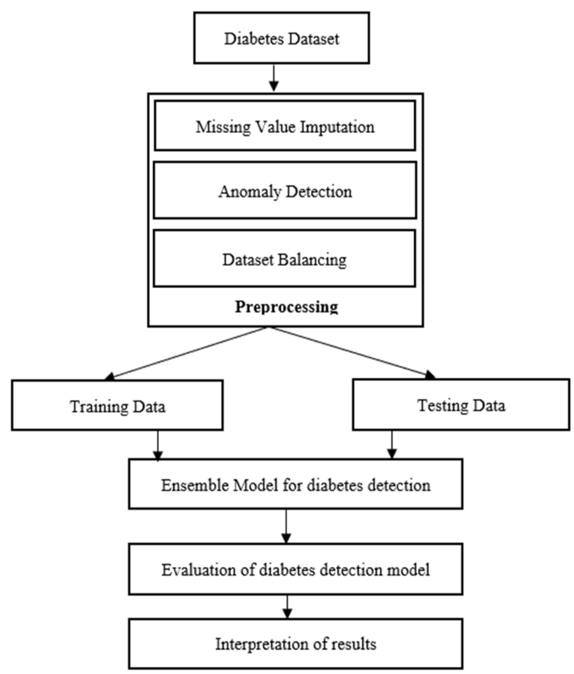

Fig. 1. Workflow of the proposed sycstem

III. PREPROCESSING

A. *Data Imputation*

Lack of certain data values in a given dataset has impeded the attainment of accurate outcomes. NaNs, empty spaces, undefined, null, or any other placeholders are used to represent the missing values traditionally [11]. Data may go missing in datasets due to various reasons, such as equipment failures, unavailable data, non-respondence to surveys, inaccurate measurements, improper data entry, or errors in laboratory procedures [12]. Two peculiar problems can arise from improper handling of missing values: a high false rate and an extended classification time [13]. This work employs KNN based imputation to replace the zero values present in the dataset [14].

B. Anomaly Detection

Outliers are unusual datapoints that deviates from the typical trend of the data. To identify outliers, statistical techniques such as Z-scores and Interquartile Range (IQR) can be applied, enabling the detection of these anomalous data points. There are other algorithms that can also be used for anomaly detections, such as: Local Outlier Factor (LOF) Isolation Forest (iForest), One-Class SVM (OCSVM); Elliptical Envelope (eEnvelope) Lightweight On-line Detector of Anomalies (LODA). After performing these anomaly detection algorithms over the chosen dataset, OCSVM outperformed other algorithms hence for this work OCSVM was used to detect the anomalies [15].

C. Balancing Dataset

The dataset is called a class imbalance if the ratio between the classes of the majority and minority of it is not equal. Data-level, algorithm-level, cost-sensitive are three ways of solving the problem of data imbalance. This work uses the hybrid sampling methods Synthetic Minority Over-Sampling Technique and Edited Nearest Neighbour, abbreviated as SMOTE + ENN. SMOTE is more robust in noisy environments than random oversampling. The ENN random under sampling method is one in which the examples whose class label differs from the majority class label of the examples closest to them are removed [15].

IV. DIABETES PREDICTION AND EVALUATION

In healthcare, machine learning plays a significant role in the identification and diagnosis of various diseases and medical conditions. To accurately classify patients based on a specific set of features, machine learning techniques are necessary. The term "ensemble" generally denotes a collection of items. In the realm of machine learning, ensemble learning combines multiple machine learning models to yield heightened accuracy in predictions. In this study, we utilized a stacking approach for the ensemble model, which combines several machine learning models through a meta-model. The complete dataset was used to train the base models, with KNN, XGB, and SVM utilized as base learners, and RF employed as the meta-model. The results of the base models are fed into meta-model to find the features from the base models to reach the highest accuracy. The models used in ensemble learning was taken from the work [15]. Fig. 2. depicts the ensemble model of the proposed work.

The proposed ensemble model for diabetes prediction is evaluated using metrics like accuracy, precision, recall, and F1 score. When assessing the reliability and validity of data, conclusions, or forecasts, accuracy is a crucial concept. Typically, it's stated as the percentage of precise predictions or the extent to which a computed or measured value resembles the intended or actual value. The exactness or precision of a system is assessed by measuring the degree to which repeated measurements or predictions yield consistent results. The ability of a classification model to correctly identify each pertinent instance of a particular class in a dataset is measured by recall metrics. They are also referred to as sensitivity or true positive rate [15].

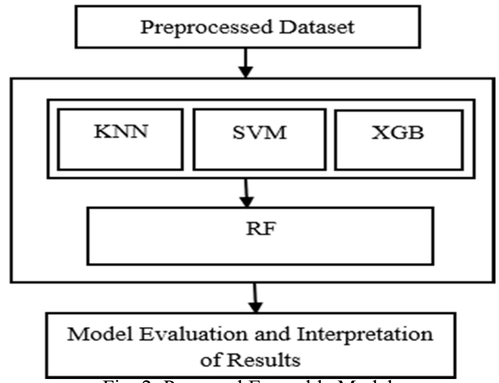

Fig. 2. Proposed Ensemble Model

V. EXPLAINABLE ARTIFICIAL INTELLIGENCE

The reason for the effectiveness of machine learning is its ability to build predictive models from data automatically even for complicated or high-dimensional data types without assuming a priori the data distribution. However, a significant drawback of machine learning that attracts criticism is that a large proportion of the algorithms it employs are entirely opaque: these models are “black box” models, in which practitioners find it difficult to discern why the models make the predictions they do [3]. The use of artificial intelligence in biomedicine is reshaping decisions and care in a more meaningful way. Explainable AI applications enable medical professionals to make more precise diagnoses, treatment decisions, risk assessments, and suggestions. The XAI models are transparent and interpretable, which allows clinicians to have confidence in and confirm the results, leading to better patient care and outcomes. Local interpretable model-agnostic explanations and Shapley Additive Explanations (SHAP), or correlation analysis, were employed to forecast the optimal set of feature vectors for classification [16].

VI. DATASET DESCRIPTION AND RESULTS

A. Dataset

We utilize the UCI Machine Learning Archive's PIMA Indian Diabetes dataset to illustrate our proposed work. The dataset consists entirely of subjects who are Pima-Indian women over the age of 21. The dependent variable in the dataset has two possible values: '0' represents a negative diabetes test, and '1' represents a positive test. Of the total cases, 268 (34.9%) are in class '1' and 500 (65.1%) are in class '0'. Because of this distribution, the PIMA Indians diabetes dataset is considered to be of poor quality, unbalanced, and contains missing values. These issues can lead to errors in classification methods. [16].

B. Results

This section discusses the outcomes of diabetes prediction using an ensemble of KNN, SVM, XGB as base learners, and RF as meta learners, as well as KNN imputation, OCSVM anomaly detection, and Smote ENN data balancing technique. The outcomes of the proposed work are shown in Fig 3. It demonstrates that the suggested ensemble model has an accuracy of 0.98, precision of 0.99, recall of 0.98, and F1 score of 0.99. Precise preprocessing methods such as anomaly detection, data balancing, and missing value imputation contributed to our ensemble model's enhanced performance.

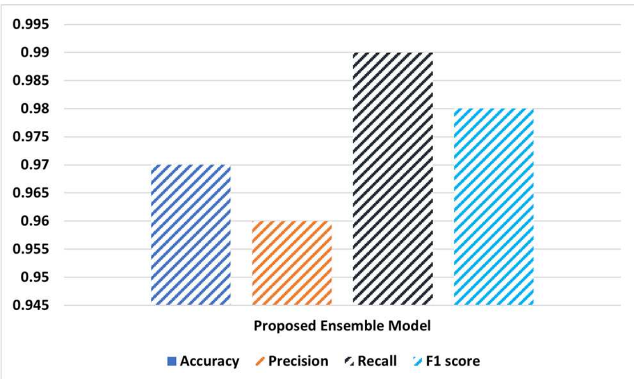

Fig. 3. Results of Proposed Ensemble Model

The proposed ensemble model is trained and tested against test data. In Fig 4a, the actual feature values of 15th observation in the test dataset is shown and it indicates the presence of diabetes with a value of 1. Figure 4b illustrates this with a yellow colour bar on the right indicating diabetes-positive and a blue colour bar on the left indicating non-diabetes. The variables Glucose value is greater than 158, Insulin is higher than 205, Pregnancies is greater than 6, Age is greater than 31, BMI is greater than 36, Skin Thickness is greater than 30, and Diabetes Pedigree Function is greater than 0.40, all these variables support the presence of diabetes for this particular observation.

In Fig 5a, the actual feature values of 20th observation in the test dataset is shown and it indicates the presence of diabetes with a value of 1. In Figure 5b, the yellow colour bar on the right represents diabetes-positive, while the blue colour bar on the left indicates support for non-diabetes. The variables Glucose is higher than 123 and less than 159, BMI is greater than 27, Insulin is greater than 92, Skin Thickness is greater than 24, Blood Pressure higher than 66, and Diabetes Pedigree Function is greater than 26. These features strongly support the positive presence of diabetes for this observation.

Fig 6a shows the feature values of the 3rd observation of the test dataset and the observation has the value of 0 indicating the absence of diabetes. In fig 6b the yellow colour bar on the right side denotes the diabetes positive, blue colour bar on the left side indicate the absence of diabetes. The variables Pregnancies is 2, Age is 25, BMI is less than 27, Glucose is less than 99, Insulin is less than 88 and Diabetes Pedigree Function is less than 0.26, all these variables support the absence of diabetes for the taken observation. Only variables Skin Thickness and Blood Pressure support the observation as diabetic.

Feature	Value
Glu	196.00
Insu	278.00
Pregs	7.00
BP	90.00
Age	41.00
BMI	39.80
SkTh	31.00
DPF	0.45

Fig. 4a: Feature Values of 15th observation

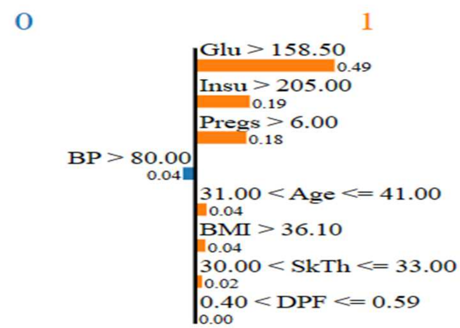

Fig. 4b: Explanation for diabetes prediction of 15th observation

Feature	Value
Glu	154.00
Pregs	1.00
BMI	29.45
Insu	126.00
SkTh	26.00
BP	69.00
DPF	0.33
Age	39.00

Fig. 5a: Feature Values of 20th observation

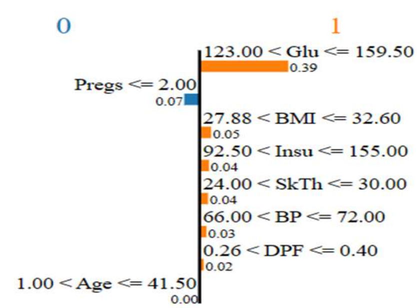

Fig. 5b: Explanation for diabetes prediction of 20th observation

Feature	Value
Glu	68.00
Insu	66.00
Pregs	2.00
BMI	25.00
SkTh	32.00
DPF	0.19
BP	70.00
Age	25.00

Fig. 6a: Feature Values of 3rd observation

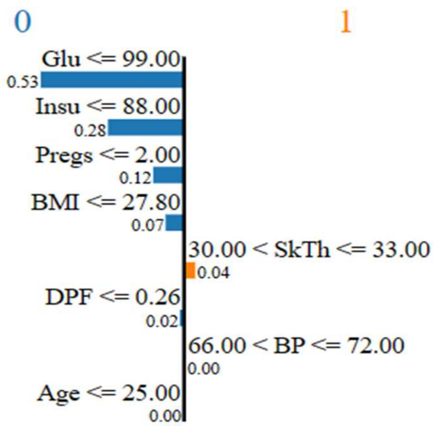

Fig. 6b: Explanation for diabetes prediction of 15th observation

VII. CONCLLUSION

Diabetes is a long-term condition that might impact multiple areas of the body. Patients with high blood glucose levels cannot produce enough insulin. Diabetes prognosis can help patients and healthcare professionals get the right treatment. In this work we proposed an ensemble model to predict diabetes by precisely preprocessing the well-known online available dataset. KNN imputation, OCSVM anomaly detection, SMOTE ENN data balancing, and an ensemble of machine learning models were all used in this work. Recall, accuracy, precision, and F1 score were employed. The proposed ensemble models achieved the accuracy of 0.97, precision, recall and F1 Score of 0.96,0.99 and 0.98. Since ML models are black box models, one of the disadvantages of using them is the interpretation of results. To overcome the disadvantage and for better comprehension of results of the proposed ensemble model this work employed LIME XAI.

REFERENCES

[1] H. Sun, P. Saeedi, S. Karuranga, M. Pinkepank, K. Ogurtsova, B. B. Duncan, C. Stein, A. Basit, J. C. N. Chan, J. C. Mbanya, M. E. Pavkov, A. Ramachandaran, S. H. Wild, S. James, W. H. Herman, P. Zhang, C. Bommer, S. Kuo, E. J. Boyko, and D. J. Magliano, ‘‘IDF diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045,’’ Diabetes Res. Clin. Pract.,vol. 183, Jan. 2022, Art. no. 109119.

[2] A. Triantafyllidis, H. Kondylakis, K. Votis, D. Tzovaras, N. Maglaveras, and K. Rahimi, ‘‘Features, outcomes, and challenges in mobile health interventions for patients living with chronic diseases: A review of systematic reviews,’’ Int. J. Med. Informat., vol. 132, Dec. 2019, Art. no. 103984.

[3] Wee, Boon Feng, Saaveethya Sivakumar, King Hann Lim, W. K. Wong, and Filbert H. Juwono. "Diabetes detection based on machine learning and deep learning approaches." Multimedia Tools and Applications (2023): 1-33.

[4] Behera, Dayal Kumar, et al. "Extreme gradient boosting and soft voting ensemble classifier for diabetes prediction." 2021 19th OITS International Conference on Information Technology (OCIT). IEEE, 2021.

[5] Shampa, Sadia Afrin, Md Saiful Islam, and Ayatun Nesa. "Machine Learning-based Diabetes Prediction: A Cross-Country Perspective." 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM). IEEE, 2023.

[6] Prajapati, Yash R., Darshan G. Hihoriya, and Shanti Verma. "Early Detection and Prediction of Diabetes Using Ensemble Classifier." 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2023.

[7] Mahajan, Srishti, et al. "Diabetes Mellitus Prediction using Supervised Machine Learning Techniques." 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT). IEEE, 2023.

[8] Nagaraj, P., et al. "A prediction and recommendation system for diabetes mellitus using XAI-based lime explainer." 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, 2022.

[9] Wang, Yu-Cheng, Tin-Chih Toly Chen, and Min-Chi Chiu. "A systematic approach to enhance the explainability of artificial intelligence in healthcare with application to diagnosis of diabetes." Healthcare Analytics 3 (2023): 100183.

[10] Vakil, Vishal, et al. "Explainable predictions of different machine learning algorithms used to predict Early Stage diabetes." arXiv preprint arXiv:2111.09939 (2021).

[11] K. Z. Hasan et al., ‘‘Performance evaluation of ensemble-based machine learning techniques for prediction of chronic kidney disease,’’ in Emerging Research in Computing, Information, Communication and Applications, vol. 1. India: Springer, 2019, pp. 415–426.

[12] Dzulkalnine, Mohamad Faiz, and Roselina Sallehuddin. "Missing data imputation with fuzzy feature selection for diabetes dataset." SN Applied Sciences 1.4 (2019): 362

[13] Emmanuel, Tlamelo, et al. "A survey on missing data in machine learning." Journal of Big Data 8.1 (2021): 1-37.

[14] Aruna Devi B, and Karthik N. " Data Imputation using Correlation-based Ma-chine Learning Algorithms." Intelligent Systems Design and Applications: 23rd International Conference on Intelligent Systems Design and Applications (ISDA 2023), Volume 2. Springer International Publishing, 2024. In press.

[15] Aruna Devi B and Karthik, N. "The Effect of Anomaly Detection and Data Balancing in Prediction of Diabetes." International Journal of Computer Information Systems and Industrial Management Applications 16.2 (2024): 13-13.

[16] Alnowaiser, Khaled. "Improving Healthcare Prediction of Diabetic Patients Using KNN Imputed Features and Tri-Ensemble Model." IEEE Access (2024)