CS550/DSL501: Machine Learning (2024–25–M)
Project Report Phase 2

Team name: MT67
1) R. C. Yajour Kichenamourty - M24MT006
2) Sagar Sudhir Pathak - M24MT007

Explainable AI Based Disease Prediction System

1 Problem Statement

People nowadays suffer from a variety of diseases because of environmental factors and their lifestyle choices. As a result, disease prediction at an earlier stage becomes a critical task. In the current healthcare landscape, disease prediction and diagnosis often rely on traditional methods, which may not always yield accurate or timely results, especially for complex diseases. With the growing availability of healthcare data from Electronic Health Records (EHRs), medical imaging, and wearable devices, the application of Machine Learning (ML) in healthcare has gained prominence. However, while ML models have shown great potential in disease prediction, they often operate as "black boxes," making it difficult for medical professionals and patients to trust and understand the predictions. This lack of transparency limits the adoption of ML system in critical healthcare environments, where explainability and trust are significant.

The problem lies in developing an ML-based healthcare application that can not only predict diseases with high accuracy but also provide clear explanations for its predictions. This will allow healthcare professionals to make informed decisions and gain trust in AI-powered medical tools.

2 GitHub

GitHub Link

3 Individual Contributions

3.1 R. C. Yajour Kichenamourty - M24MT006

Yajour's primary responsibility was the initial data preprocessing and feature engineering for the dataset. This involved:

- **Data Collection:** Sourced data from publicly available health datasets such as Electronic Health Records (EHRs), wearable device data, and medical imaging repositories.
- **Data Cleaning:** Managed missing values, normalized data, and handled outliers using Python libraries such as `pandas` and `numpy`.

- **Feature Engineering:** Selected key features for disease prediction by using domain knowledge and statistical methods (e.g., correlation analysis) to filter out unimportant features.
- **Model Training:** Conducted experiments with several machine learning models like Random Forest, Gradient Boosting, and XGBoost for initial disease prediction.
- **Performance Tuning:** Performed hyperparameter tuning using Grid Search and Random Search techniques to improve model performance.
- **Documentation:** Handled documentation and reporting for the explainability aspect of the project, detailing how the system's predictions could be trusted based on the generated explanations.

3.2 Sagar Sudhir Pathak - M24MT007

Sagar was responsible for implementing the Explainable AI (XAI) components of the system, which included:

- **Exploration of XAI Techniques:** Investigated several XAI methods, such as LIME (Local Interpretable Model-agnostic Explanations), SHAP (SHapley Additive exPlanations), Decision Trees, Logistic Regression, Explainable Boosting and integrated Grad-CAM for explainability in image-based models.
- **Model Integration:** Integrated SHAP values with the disease prediction model to provide feature-level explanations for predictions. This ensured that healthcare professionals could understand the influence of different factors on predictions.
- **System Architecture:** Designed the overall architecture for the healthcare prediction system, ensuring that both the prediction and explanation components worked seamlessly.
- **Model Evaluation:** Evaluated the explainability of the system by conducting tests with sample predictions and explanations, ensuring transparency and accuracy in predictions.

  In addition, Sagar contributed to developing an interface for integrating machine learning models into a healthcare application, including assisting with testing the system in a controlled environment.

4 Tasks and Milestones Achieved

The project invloved training of 3 different datasets viz., Heart Disease, Alzeimers' and Diabetes over 5 Explainable AI models viz., SHAP, LIME, Decision Tree, Logistic Regression and Explainable Boosting Algorithm.

- **Problem Understanding and Data Collection (Week 1-2):** Successfully identified a healthcare problem where AI-based disease prediction could be useful. Collected relevant healthcare datasets for training the machine learning model.

- **Data Preprocessing and Feature Engineering (Week 3-4):** Cleaned and preprocessed the dataset. Selected relevant features based on domain knowledge and statistical techniques.
- **Model Training and Evaluation (Week 5-6):** Trained multiple machine learning models for disease prediction and evaluated them on accuracy and precision. Finalized Random Forest as the best-performing model.
- **Integration of XAI Techniques (Week 7-8):** Integrated SHAP and LIME into the model to provide explainability for predictions. Successfully generated human-readable explanations for individual predictions.
- **Deployment and Testing (Week 9):** Deployed the system in a test environment and validated predictions with explanations. Worked on refining the user interface to display model explanations effectively.

The RECALL values showed Explainable Boosting Algorithm performing best for Heart Disease and Diabetes dataset while Logistic Regression was efficient for Alzeimers' dataset. The Front end application was also made using the above said XAIs for respective datasets.

5 Further Works

- **Optimization for Real-Time Predictions:** The current system is capable of providing predictions and explanations but may not be optimized for real-time performance. Further work can involve optimizing the model's inference time to make it more suitable for real-time use in clinical settings.
- **User Interface Improvements:** Currently, the user interface for displaying explanations is functional but basic. Future work will focus on improving the visual representation of SHAP values and other XAI metrics to make them more intuitive for medical professionals.
- **Generalization for Other Diseases:** While the model is focused on a specific set of diseases, further work will focus on extending the model to cover a wider range of diseases by using more diverse datasets.
- **Model Calibration:** Explore model calibration techniques to improve the trustworthiness of the predicted probabilities, ensuring that predictions are not only accurate but also well-calibrated.

THANK YOU