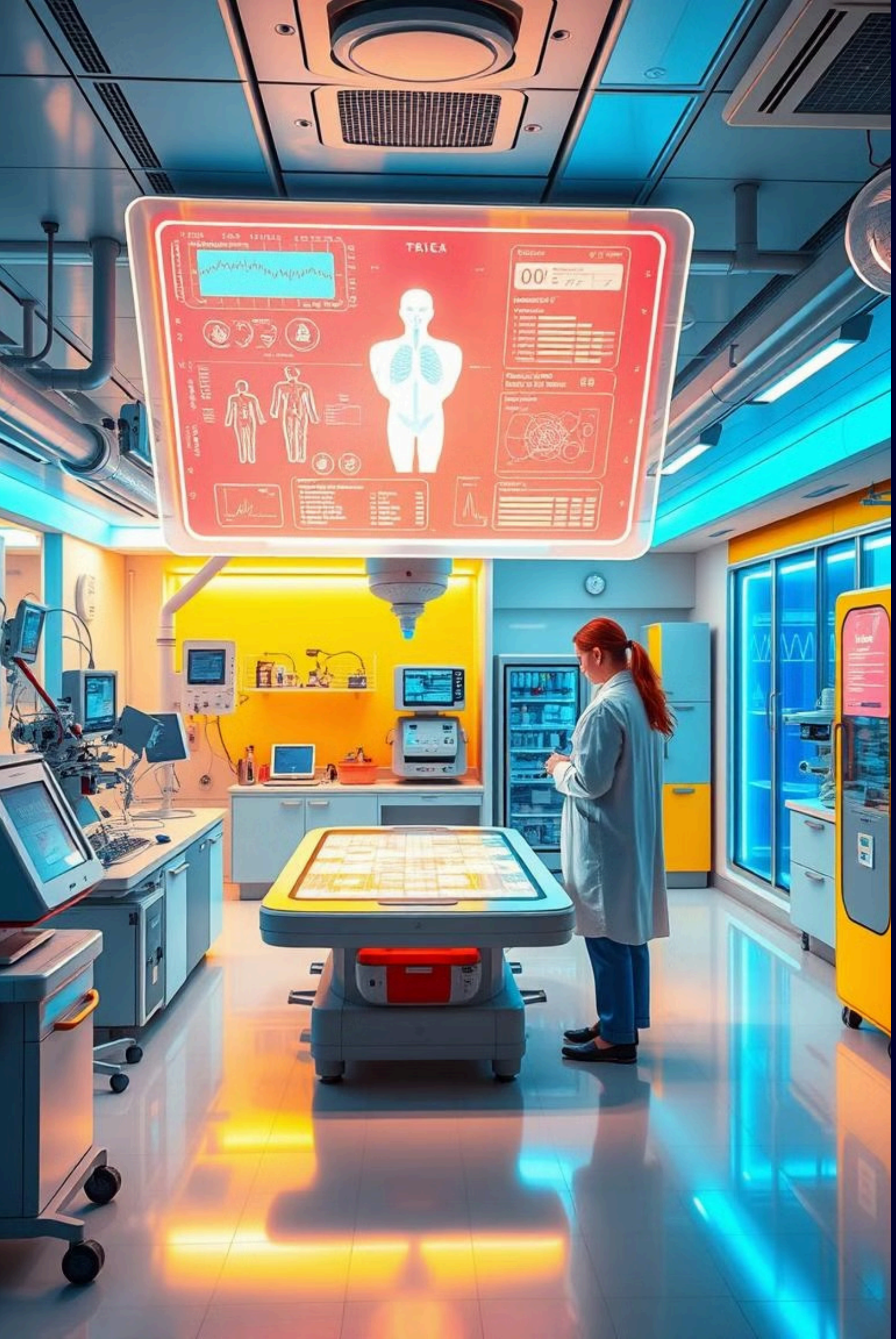

Predicting Heart Failure: An Approach using AI models

This presentation outlines an AI project that focuses on predicting healthcare outcomes. We explored key aspects of model development, including data preprocessing, model selection, and the critical considerations for achieving fairness, explainability, and effectiveness.

Created By: Team 4

Our Data Source and Data Transformation Process

Data Source

We leveraged a large dataset of electronic health records (EHRs) from Kaggle: https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease/data

Transformation Process

The raw data of over 400K records was transformed into a structured format suitable for machine learning. This involved data cleaning, feature engineering, and handling missing values.

Heart Failure was the target or 'Y' variable, and the prediction of heart failure consisted of one of two columns being true:

data['HeartFailure'] = (data['HadHeartAttack'] 1.0) | (data['HadAngina'] 1.0)

Model Development Process and Key Considerations

1

Model Selection

We experimented with various machine learning algorithms, including logistic regression, support vector machines, and random forests, to identify the most suitable model for our specific prediction task.

2

Performance Evaluation

Model performance was rigorously evaluated using metrics such as accuracy, precision, recall, and F1-score to ensure the model's effectiveness in predicting healthcare outcomes.

3

Ethical Considerations

We emphasized fairness and explainability to ensure that the model's predictions were not biased and could be readily interpreted by healthcare providers.

Challenges with an Unbalanced Model

1. ## Data Imbalance

   The dataset exhibited class imbalance, where one outcome (e.g., heart failure) was significantly less prevalent than the other (e.g., no heart failure). This posed a challenge for model training and evaluation.

2. ## Overfitting Risk

   An imbalanced dataset could lead to overfitting, where the model learned to predict the majority class very well, but struggled with the minority class.

3. ## Addressing Imbalance

   We employed various techniques to address class imbalance, including resampling, Class weighting, and applying SMOTE to oversample the minority class (Heart Failure).

   We will finally performed threshold tuning to bring recall to an acceptable level.

Results Improvements

91% -> 77%

Accuracy

The model iterations ultimately achieved an accuracy of 77% in predicting heart failure, demonstrating its effectiveness in making accurate predictions. However, from earlier models, reduced accuracy was a tradeoff to increase recall and to reduce the chance of reporting false negatives.

- High accuracy largely reflects the model's ability to predict the majority class (0), as most examples belong to this class.
- Accuracy is not the most meaningful metric for this problem since false negatives (failing to identify heart failure) are critical.

52% ->74%

ROC AUC

ROC AUC measures the model's ability to distinguish between positive cases (heart failure) and negative cases across all possible thresholds, making it a comprehensive evaluation metric.

It's ideal for heart failure prediction because it works well with imbalanced datasets and prioritizes reducing false negatives, which is critical in healthcare.

ROC AUC stands for **Receiver Operating Characteristic - Area Under the Curve**.

• **ROC (Receiver Operating Characteristic):** A curve that plots the True Positive Rate (Recall) against the False Positive Rate at various thresholds.

• **AUC (Area Under the Curve):** The total area under the ROC curve, quantifying the model's ability to distinguish between classes.