Problem Statement:

The development of drugs is critical in providing therapeutic options for patients suffering from chronic and terminal illnesses. “Target Drug”, in particular, is designed to enhance the patient's health and well-being without causing dependence on other medications that could potentially lead to severe and Life threatening side effects. These drugs are specifically tailored to treat a particular disease or condition, offering a more focused and effective approach to treatment, while minimising the risk of harmful reactions. The objective in this assignment is to develop a predictive model which will predict whether a patient will be eligible*** for “Target Drug” or not in the next 30 days. Knowing if the patient is eligible or not will help physicians treating the patient make informed decisions on which treatments to give.

Potential Application of the problem statement:

The problem statement, which involves developing a predictive model to determine a patient's eligibility for the "Target Drug" in the next 30 days, has several potential applications that can significantly benefit the healthcare industry and patient care. Some of these potential applications include:

1. **Personalized Medicine and Treatment Planning:** The predictive model can enable healthcare providers to personalize treatment plans based on the predicted eligibility of patients for the "Target Drug." This can lead to more targeted and effective interventions, improving overall treatment outcomes and patient satisfaction.

2. **Optimized Resource Allocation:** By accurately predicting a patient's eligibility for the "Target Drug," healthcare facilities can optimize resource allocation, ensuring that the drug is allocated to the most suitable patients. This can help in managing drug supplies more efficiently and reducing unnecessary costs.

3. **Improved Patient Care and Outcomes:** The application of the predictive model can lead to improved patient care and better health outcomes. Identifying eligible patients for the "Target Drug" can result in timely and appropriate interventions, leading to better disease management and improved overall patient health.

   Overall, the application of the predictive model can lead to a more efficient and targeted approach to patient care, contributing to improved treatment outcomes, reduced healthcare costs, and enhanced overall patient well-being.

Type of ML problem:

1. The problem described, which involves developing a predictive model to determine a patient's eligibility for the "Target Drug" in the next 30 days, can be categorised as a **binary classification problem** in the field of machine learning.
2. The model is trained using labelled data, where each instance has a set of features (e.g., patient characteristics) and a corresponding label indicating whether the patient is eligible for the "Target Drug" or not.
3. The predictive model is then trained to learn the patterns and relationships between the input features and the target label, enabling it to make accurate predictions on new, unseen data.
4. By utilising various machine learning algorithms such as random forests, Extra Gradient Boosting, the binary classification problem can be effectively addressed.
5. The goal of the model is to accurately classify patients into two categories, i.e., those who are eligible and those who are not eligible for the "Target Drug," based on the relevant input features.
6. The performance of the model is typically evaluated using metrics such as F1-score, auc_roc score, cross_val_score and area under the receiver operating characteristic curve (AUC-ROC), among others, to assess its predictive capabilities and generalisation to unseen data.

Data:

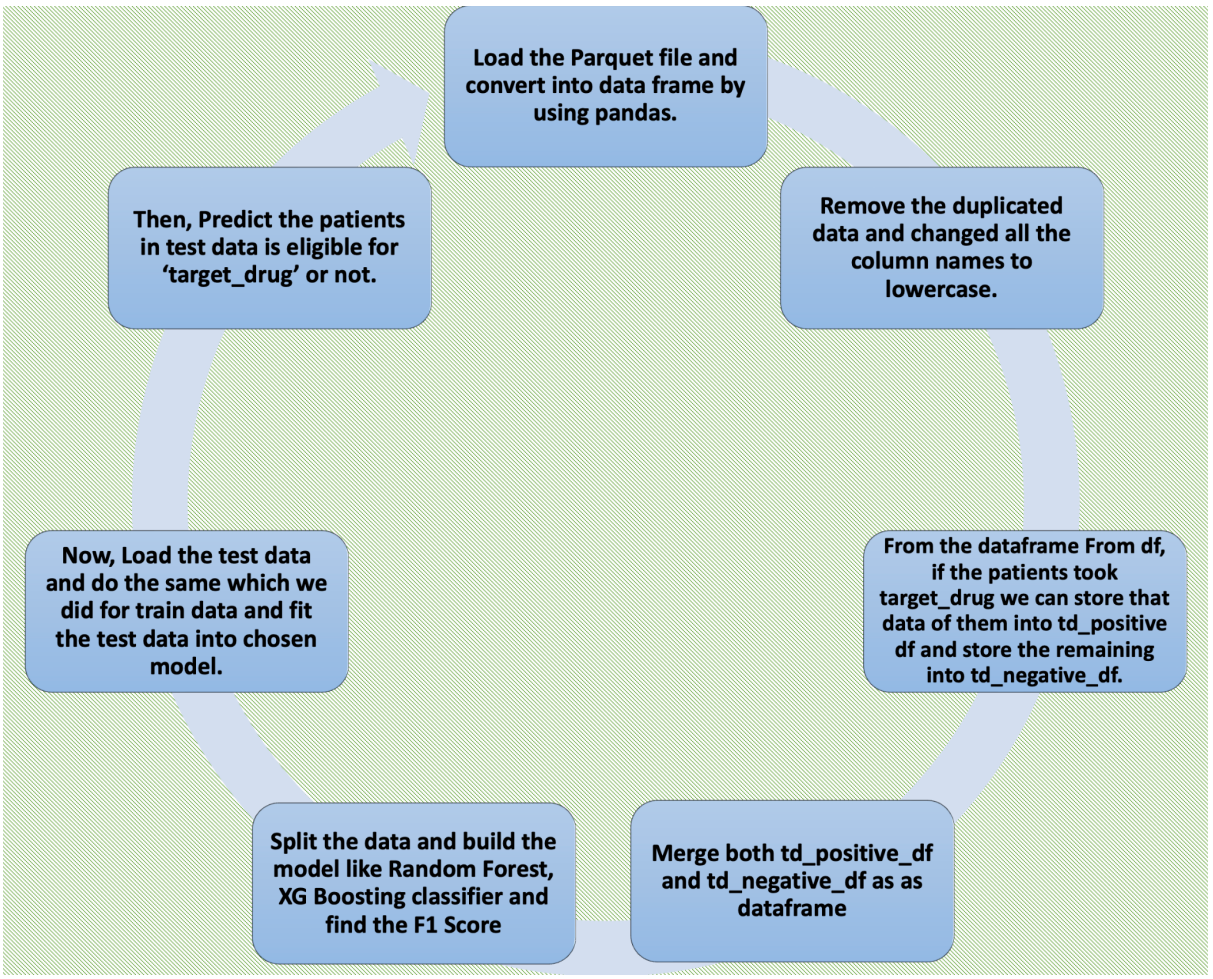
https://drive.google.com/file/d/1oHnw-M9jOshB3WkbKrMBWepjlEHAdwA1/view

By ensuring the availability of comprehensive, accurate, and updated data encompassing these key aspects, the development of an effective predictive model becomes more robust and reliable, enabling healthcare professionals to make informed decisions regarding the eligibility of patients for the "Target Drug.

Problems with data and Solution:

During the analysis of the train dataframe, it was discovered that there were 35,571 duplicate entries present in the dataset. Duplicates in a dataset can skew the analysis results and potentially lead to inaccurate conclusions. Therefore, to ensure the integrity and reliability of the analysis, the decision was made to remove these duplicate entries from the dataframe. By dropping the duplicates, the dataset was streamlined and prepared for more accurate and meaningful analysis, enabling a more precise understanding of the underlying patterns and insights within the data.

Solution:



Final Solution:

The development and implementation of the predictive model signify a pivotal advancement in the realm of patient care, providing physicians with a robust tool to accurately pinpoint individuals eligible for the administration of the specialised "Target Drug.By leveraging the model's precise predictions, physicians can make well-informed decisions on treatment protocols, ensuring that patients receive the most appropriate and impactful interventions to address their individual health concerns.

https://github.com/JoeSnowA/Akaike-Technologies