

STEP BY STEP PROCEDURE

Step 1:

- Load the Dataset train.parquet file.

Step 2:

- Get the statistical information from the dataset.
- Preprocess the Data with checking null values, info().

Step3:

- Split the data into TARGET DRUG as eligible, other than Target Drug is not eligible.

Step 4:

- Necessary feature Engineering Has been done, like Frequency and Time Based procedure.

Step5:

- Calculate the No of Days between 1st prescription and target Drug used.

Step 6:

- Calculate the Total No of Prescription has been taken by the single patient.
- Encode the Incident Columns for a better Modelling.

Step 7:

- Split the dataset into Positive Set and Negative Set.
- Combine the new set of data to, which ready for model Training.
- Scale the data in order to get the even Distribution.

Step 8:

- I have done Different types of Classification model.
- Choose the Best model from that after performing different classification process.
- I have chosen Extragradient Boosting Classification for the better performance compared to others.

Step 9:

- Import the Test.parquet dataset.
- Do the same preprocessing steps for the test dataset.
- With the best model predict the Test data to the model.

Step10:

- Get the results in the CSV format , which predict the Target Drug 1 or 0.

NOTE:

I Have added the necessary Comment in that particular step, to get a proper understanding.