NAME - ADARSHA KUMAR GUPTA

SECTION – KM046

REGISTRATION NUMBER - 11811996

EMAIL – adgta007@gmail.com

Abstract

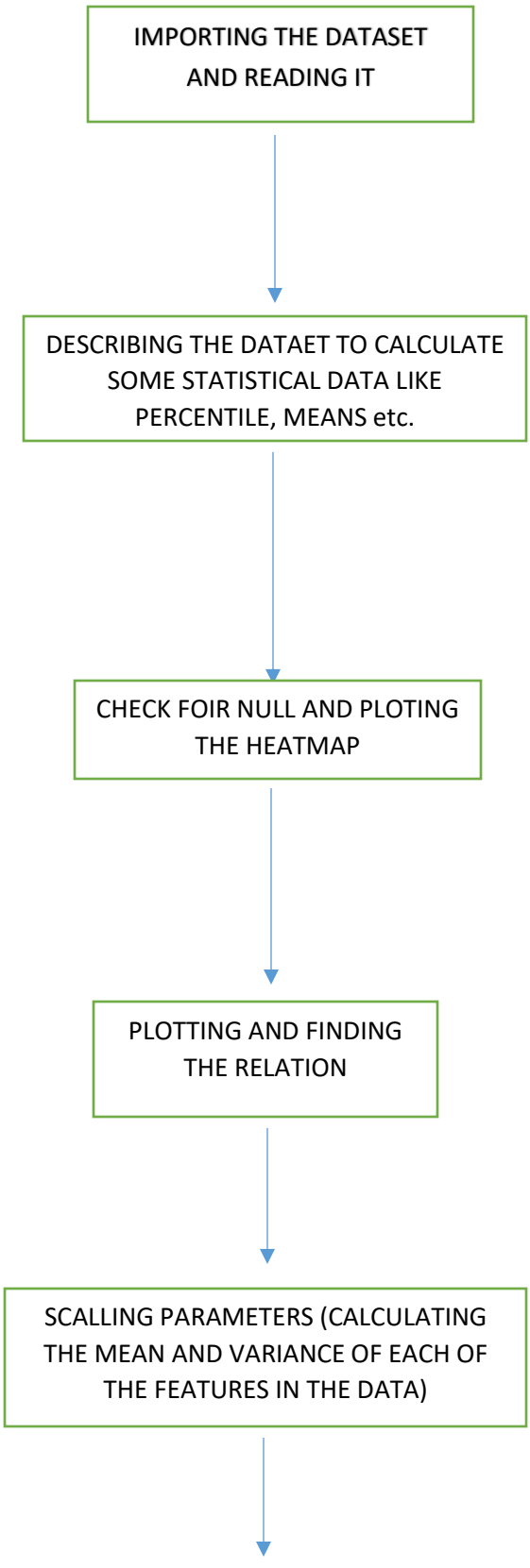
The project is related to Heart Disease Prediction and the data is taken from Kaggle. There are various columns like chest pain type, resting blood pressure (in mm Hg on admission to the hospital), cholesterol level, fasting blood sugar, resting ECG, etc. These columns helped to get an understanding of which factors are responsible for the disease. Finding the relation among the columns using various visual plotting, made it possible to understand the dataset.

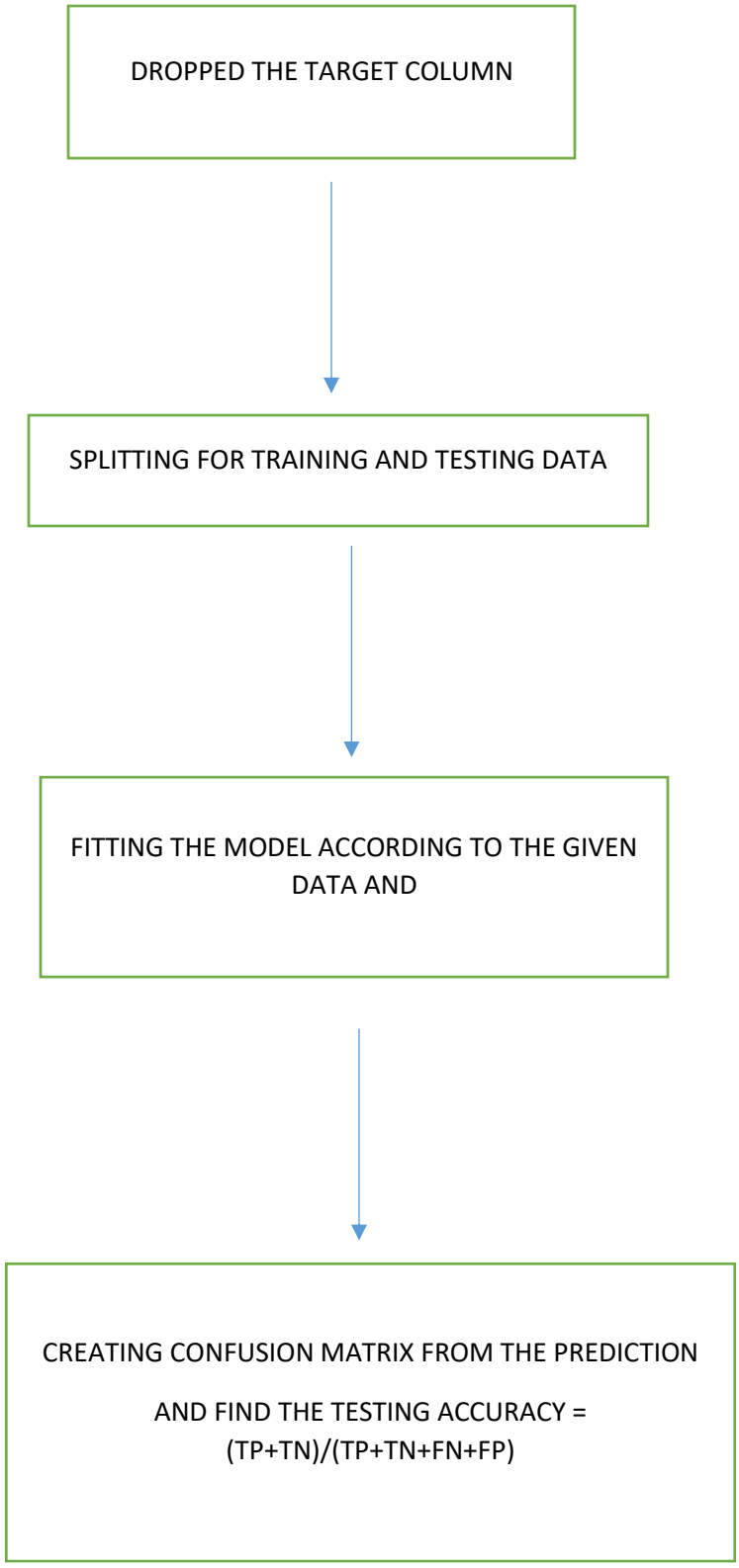
Dataset

Link of the repository - https://www.kaggle.com/ronitf/heart-disease-uci

Github - https://github.com/travellerR/Heart-Disease-Prediction-Using-Logistic-Regression

FLOWCHART


IMPORTING THE DATASET AND READING IT
DESCRIBING THE DATAET TO CALCULATE SOME STATISTICAL DATA LIKE PERCENTILE, MEANS etc.
CHECK FOIR NULL AND PLOTING THE HEATMAP
PLOTTING AND FINDING THE RELATION
SCALLING PARAMETERS (CALCULATING THE MEAN AND VARIANCE OF EACH OF THE FEATURES IN THE DATA)

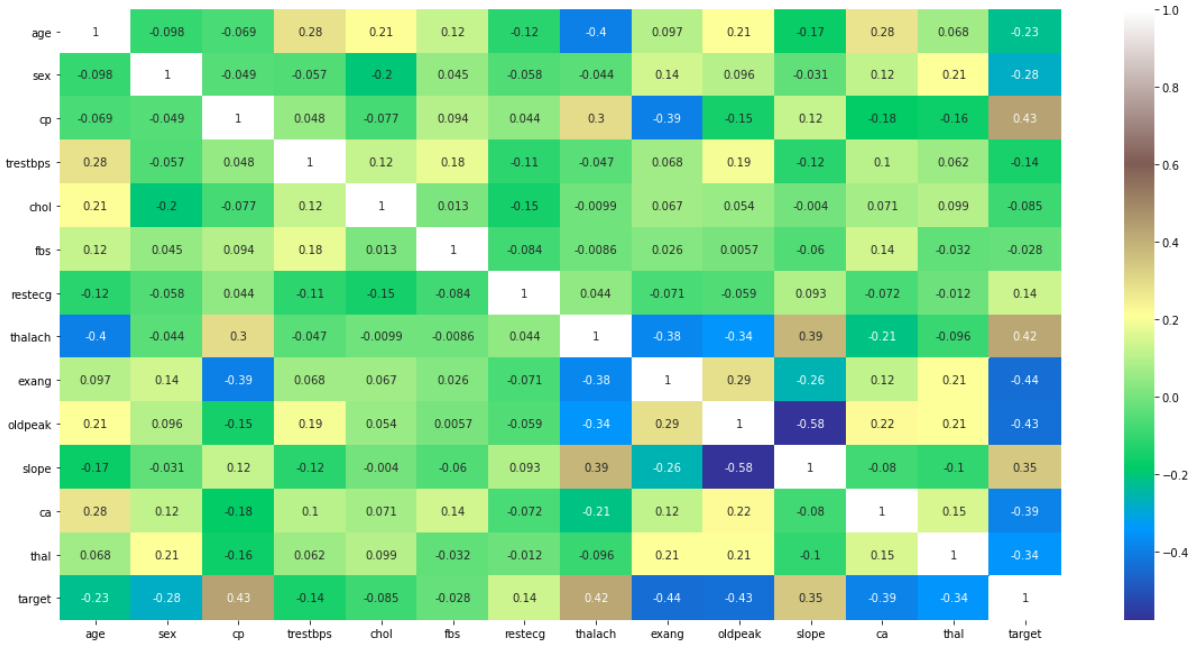



```
Testing Accuracy: 0.9230769230769231
```

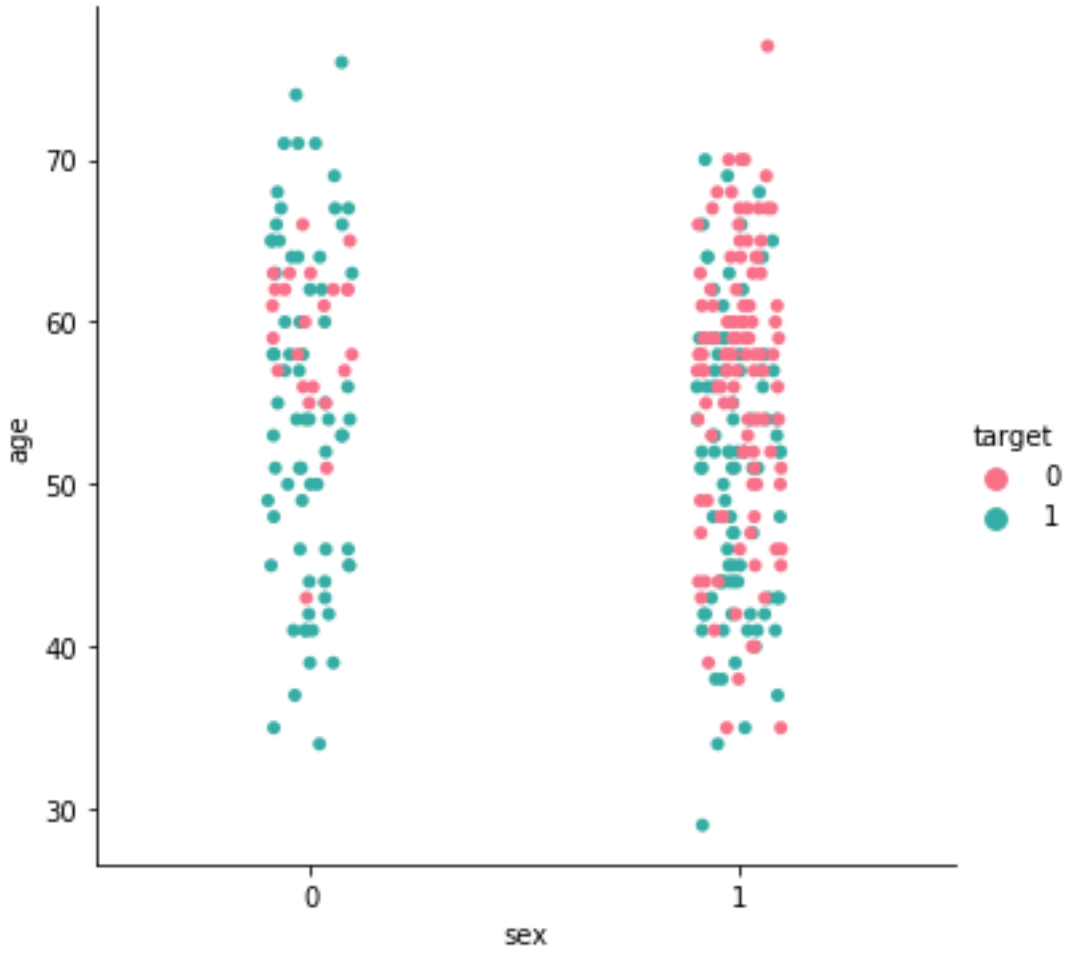
EXPERIMENTAL SETUP

JUPYTER NOTEBOOK

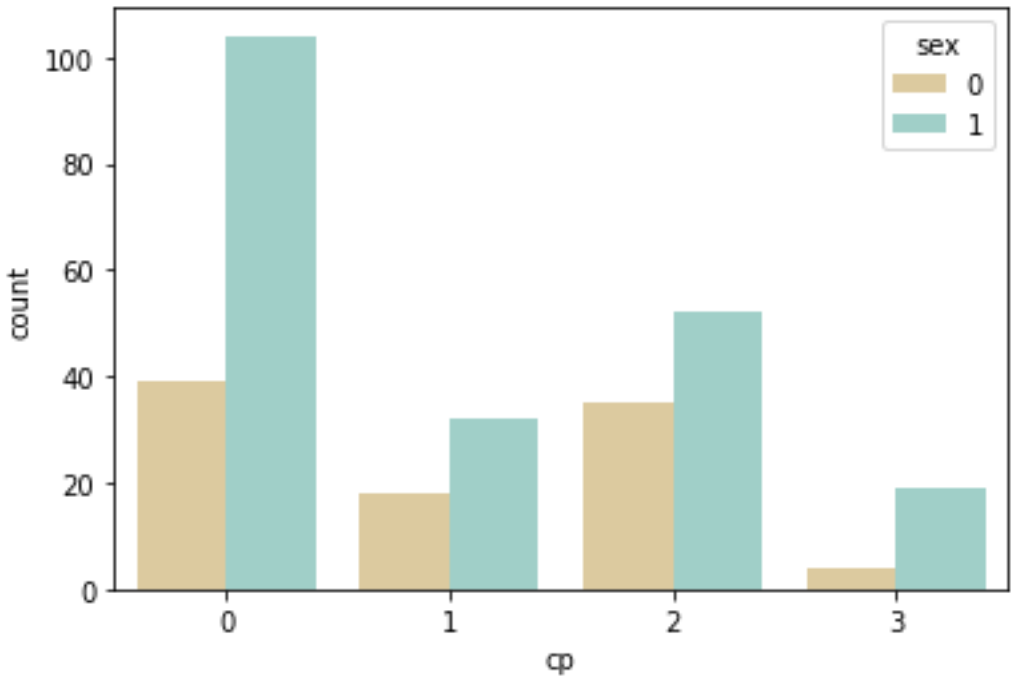
RESULTS



HEATMAP



CATPLOT



COUNTPLOT

CONCLUSION

After plotting the various relations, it was found the that the various factors that were actually responsible for the heart disease were -

- AGE
- RESTING BLOOD PRESSURE
- CHOLESTROL
- MAXIMUM HEART RATE ACHIEVED
- DEPRESSION INDUCED BY EXERCISE RELATIVE TO REST

USING LOGISTIC REGRESSION THE RESULTS WERE 92% ACCURATE.