

FORECAST OF VEHICLES ACCIDENT SEVERITY

Md Islam

OUTLINE

- Introduction
- Goals
- Methodology
- Discussion
- Conclusion

INTRODUCTION

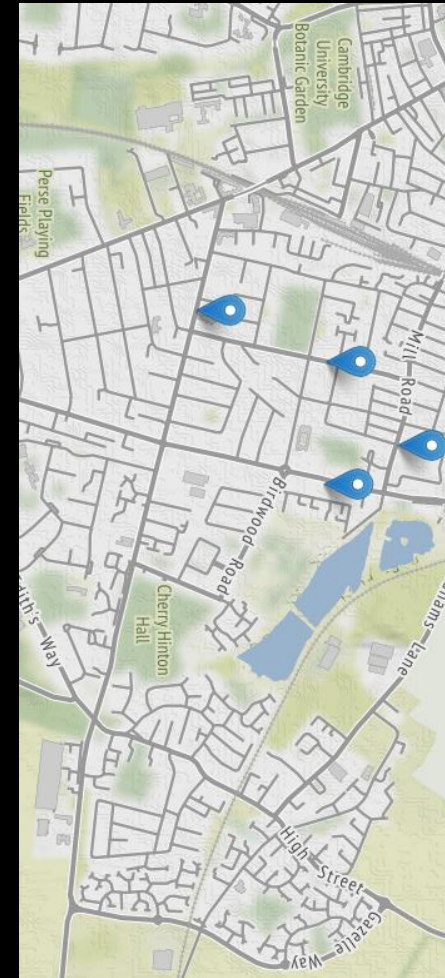
- Road traffic accidents of England
 - Growth of many activities involve cities for accidents
 - Geographic positions of those cities
 - Weather preferred place to go to avoid accidents
 - Effective factors reduce accidents

GOALS

- The goals is to
 - Identify the cause of accidents in UK and how it will be predicted before any accident
 - It will predict two side of the country which is east and northeast part.
 - Highlight the potentiality and needs to develops vehicle crashing.
- The goal is of interest for pedestrians, drivers and all kinds of commuters.

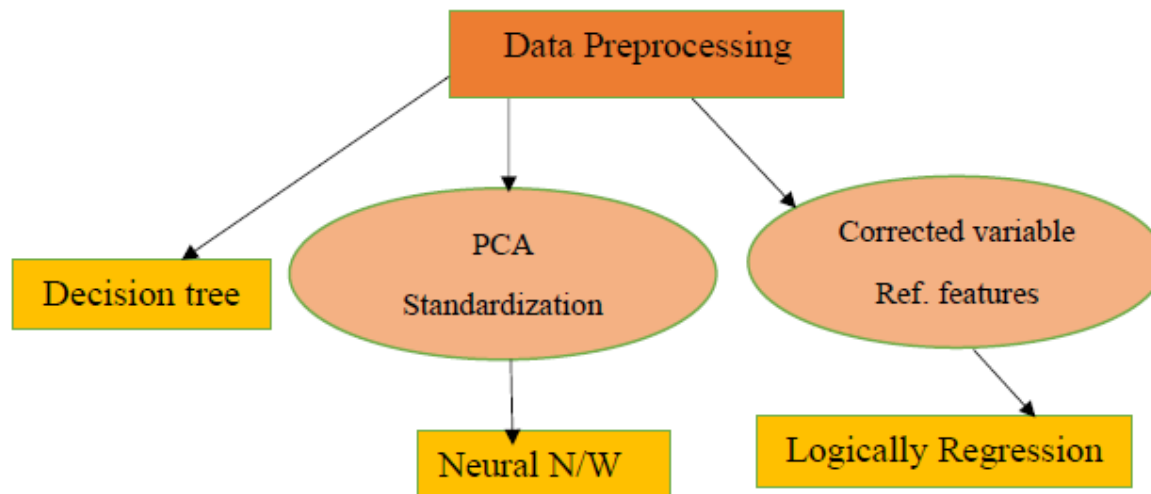
GOALS

On the map we could see the geographic position of Traffic and commuters each with their respective neighborhood. We obtained the information about the neighborhoods cause some accident severity for some season where traffic is the most one



METHODOLOGY

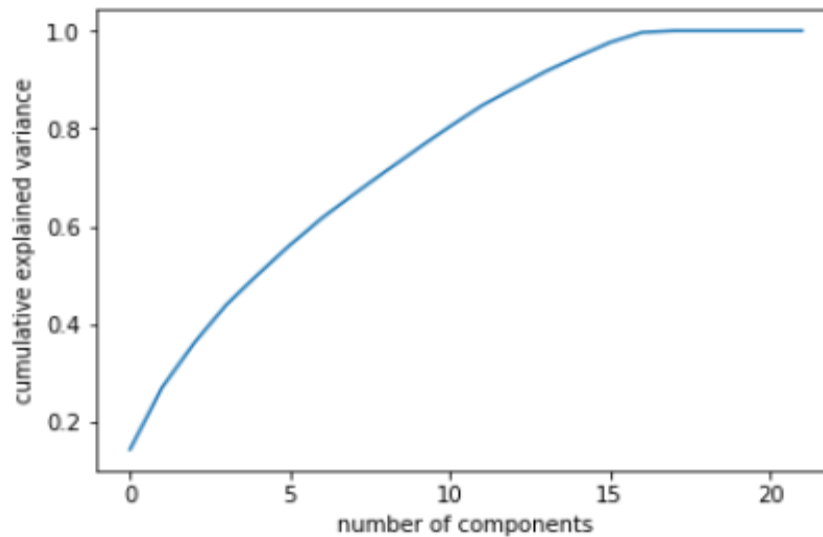
A. Data preprocessing



- The three datasets were merged
- Deleting the missing value data

METHODOLOGY

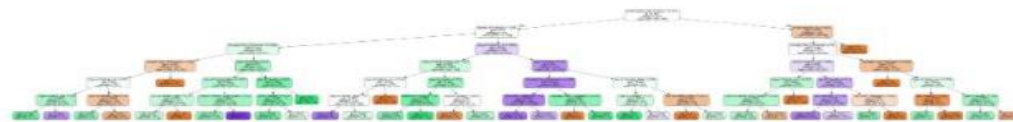
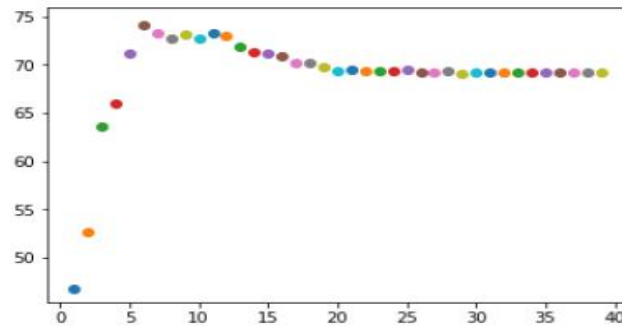
B. PCA



It contains first 12 components which covers mostly variance

METHODOLOGY

C. Decision tree



On the above the graph depicts that the best accuracy based on the customized datasets and the tree also shows the importance of road surface, casualty of sex etc.

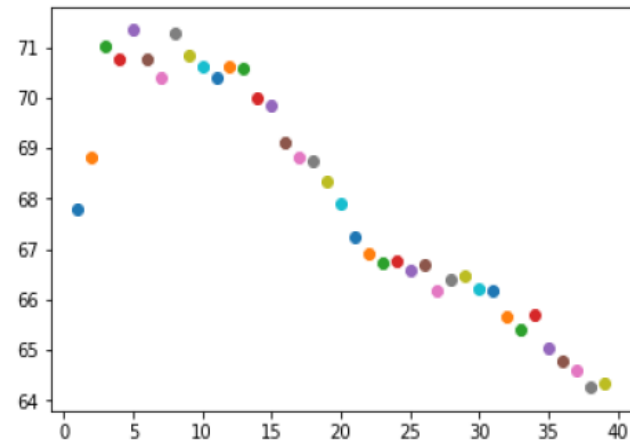
Selected forest: The standard deviation is 3.98%

METHODOLOGY

D. Neural N/W

- It contains 24 nodes
- Maintaining Standardized data via PCA
- The mean is 72.66%

E. K-Nearest Neighbors



DISCUSSION

- Quality of our dataset used for the geocoding of department, borough and neighborhood. sample dataset meaning we were limited in our search. With full list gear this study toward a comparison between the two cities of Country
- Foursquare API to obtain the venues of each region. Too big radius. Query within a radius of 500 for concise and accurate result. Nevertheless necessary to geocode venues and places in the whole cities of England.
- The mean of K-fold is 71.37%
- The mean of PCA is 53.12%

CONCLUSION

G. Conclusion

Algorithm	Mean Accuracy	Standard Deviation
Decision Tree	73.95	2.64
Random Forest	67.15	3.98
Neural Network	72.66	2.95
KNN	71.37	2.77
Logistic Regression with PCA	53.12	2.32
Logistic Regression without PCA	54.09	3.03

On the above, the final model decision tree where mean is 73.95%, therefore we can say that the most 3 features involve the severity of vehicle accident are – Road Surface dry, road surface wet, and Casualty of Sex.