Yifan Yao
ECE 532
Madison car crash prediction project
Email: yyao66@wisc,edu

Object:

The purpose for this project is giving a route choice for drivers for their safety.

The first part is figuring out for this person which area needs to be avoided because of high risk, which means make a prediction after input some person information like age, sex, time, currrent location, then using histrical crah data to give a some risk area.

The second part is more focus on route choosing. The person from A goes to B, based on the risk area prediction from the first part, giving a safety route for this person.

Project Dataset:

The dataset will be used is crash data from Wisconsin Traffic Operations and Safety Laboratory (TOPS Lab), which includes all crashes in Madison in 2019. The total number of crashes is 5414, which means the row of the dataset is 5415. The dataset gives 126 features of these crashes, which include driver information as well as crash information. For this project may focus on some general personal information like sex, age, current location and time, also for some feature which is signifcently connect with crash like using alcohol, drug, high speed.

Algorithms that will be investigated:

Because all the data has been recorded after crashes and it is hard to get data of traffic flow to normalize it,it does not have labels. Thus algorithms to be used need to be unsupervised learning.

For the first part, which is using historical crash-data to predict the risk area, the algorithm that might be used for clustering is k-means.

For the second part, which is route choice, the algorithm that might be used is PageRank and networks.

For the third algorithm, I am still thinking about, it might be Singular value decomposition or Principal component analysis.

Project Github:

https://github.com/Mieggy/car-crash-prediction-project.git

Project timeline:

Initial proposal -- Oct 22nd

First part about prediction risk area -- Nov 17th

Second part about route choice -- Dec 1st

Final project report -- Dec 12th