Accident Severity in Seattle from 2004 to 2020

Background

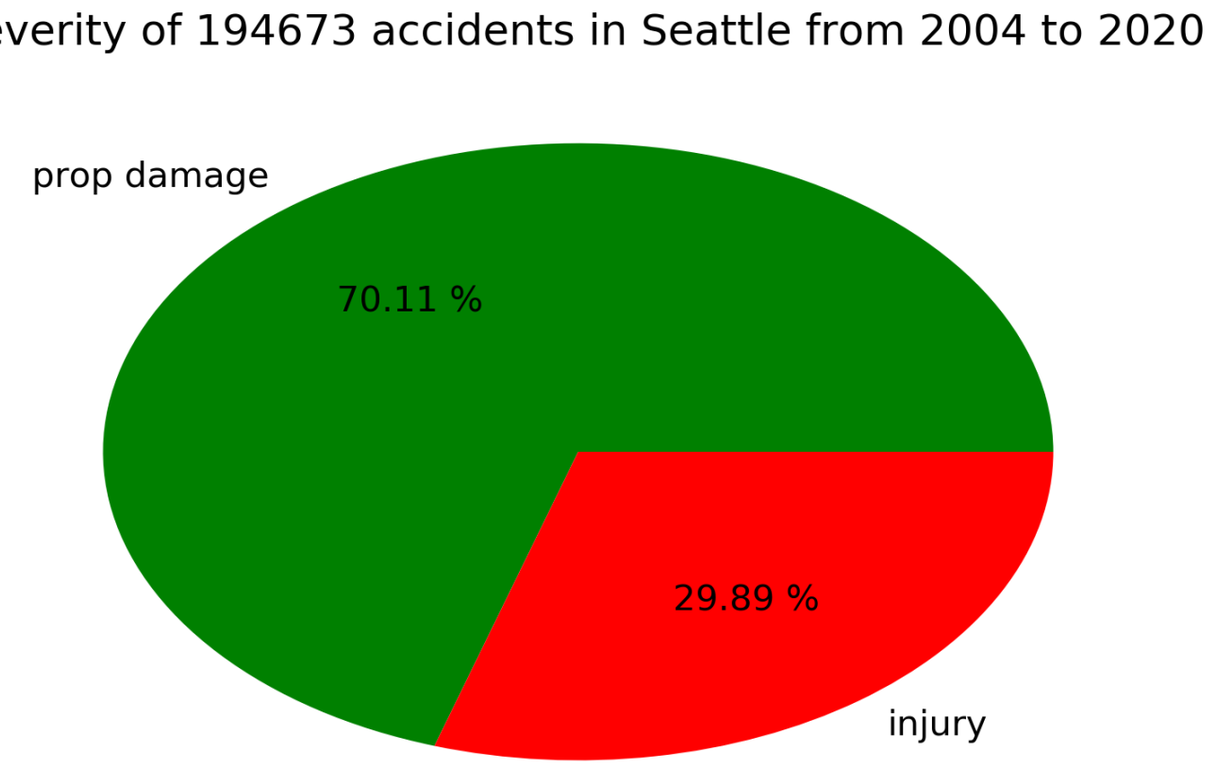
Growth rate of population in Seattle ranks No.2 in US. As a fast growing city, the traffic congestion problems have become a terrible problem. According to the 2019 Urban Mobility Report, Seattle rank 7th on the problem of time delay for auto commuters traveling during peak periods (6 a.m. to 10 a.m. and 3 p.m. to 7 p.m.). Increasing number of traffic accidents is one of the most contribution to traffic jams.

Problems

If prediction of severity of traffic accidents can be made, traffic congestion problem can be relieved.
With the development of machine learning, we want to use the data of traffic attributes including road condition, weather and other factors to predict accident severity.

Data

The data of severity of accident I used is from ArcGIS Metadata form. It concludes 194673 collisions happened in Seattle from 2004 to 2020. 37 attributes have been recorded on each collision. There are two kinds of severity have been recorded in this data set, injury and prop damage. 70% are prop damaged and 30% are injury.

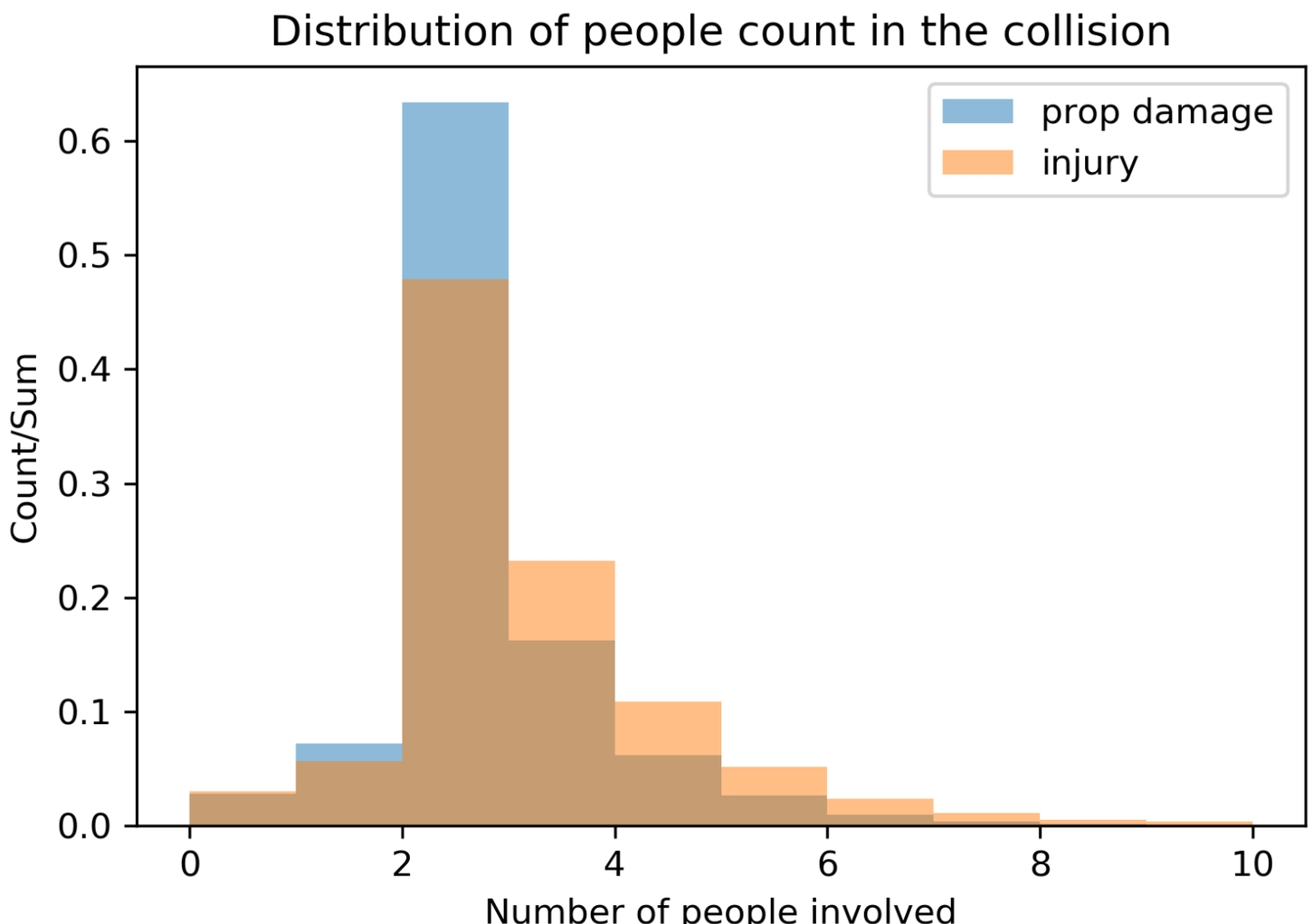


Features

1. COLLISIONTYPE:Collision type
2. PERSONCOUNT: The total number of people involved in the collisions.
3. INATTENTIONIND: Whether or not the collision was caused by the inattention
4. WEATHER: Weather condition
5. ROADCOND: Road condition during the collision.
6. LIGHTCOND: Light condition during the collision.
7. SPEEDING: Whether or not speeding was a factor in the collision.

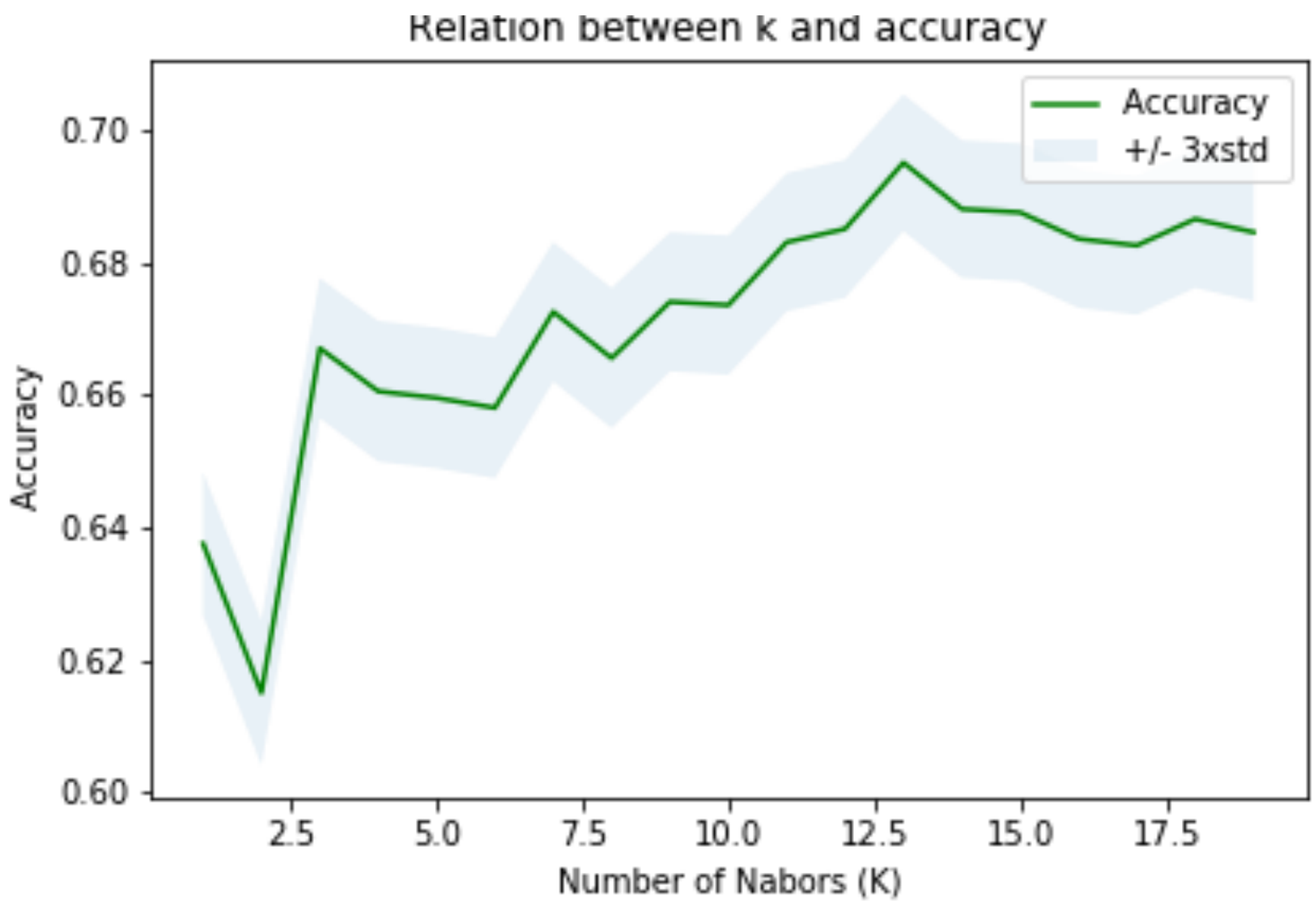
Example of Relations between features and target

- Person count: number of people is high more probability cause a more serious accident.



Models

- K nearest neighbor algorithms
- When k=12 is the best model to predict



Evaluation

- Jaccard score and F1-Score
- All are 0.685: stable and predictable

Algorithm	Jaccard	F1-score
KNN(k=12)	0.685	0.6849747983870969

Conclusion

We have developed an K nearest neighbor model to predict the severity of the collisions. We used 7 features which are highly related to the severity of the collisions, including: collision type, person involved, reason caused the collision, weather, road condition and driving condition. It appears that when the driving condition is good, it will cause a more serious accidents. The reason might be when the weather is good and road condition is good, people will be more careless. Here we recommend that the drivers should always be careful to save people from accidents.