Predicting Traffic Accident Severity

JIALIANG ZHANG

2020-10-07

TRAFFIC ACCIDENT SEVERITY INSIGHT

❖**Scenario Description:**

Traffic accidents are occurring everyday.

❖**Data insight benefits:**

The corresponding government could allocate the resources feasibly.

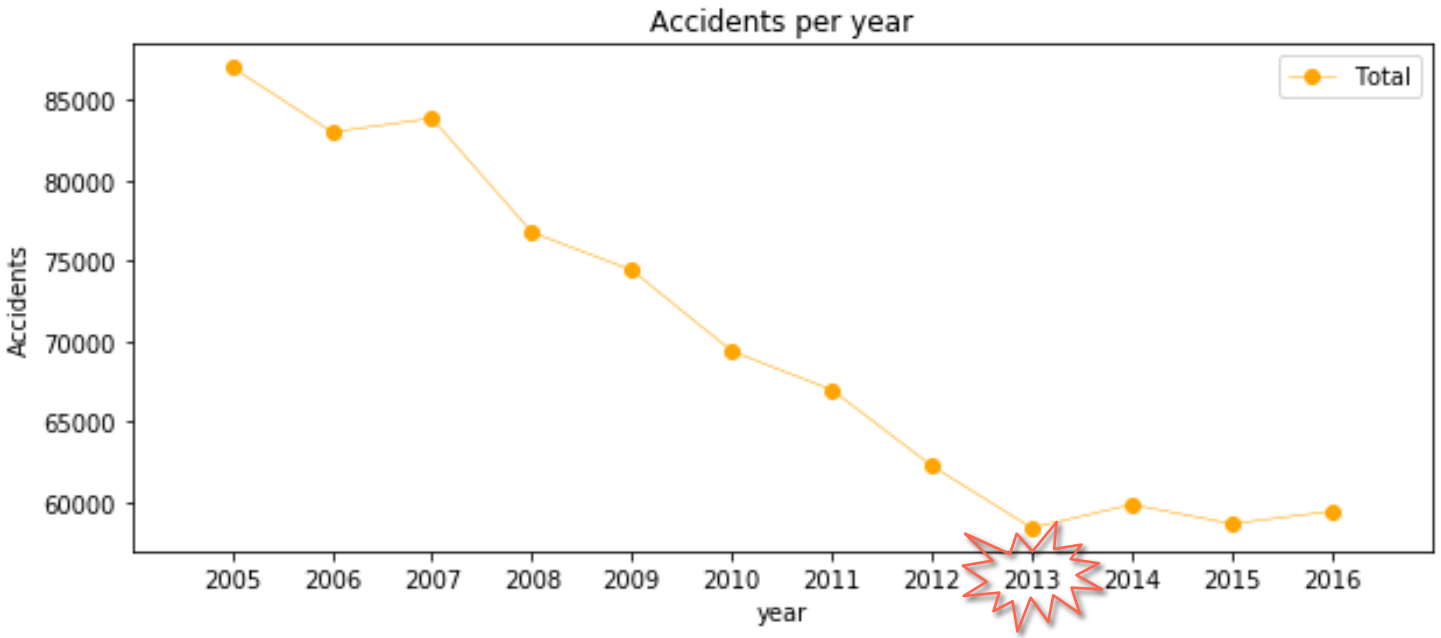
Drivers could be notified to change the planned route for avoiding the traffic accident.

DATA ACQUISITION AND CLEANING

- The original data comes from the following Kaggle data set .
- 29 data labels are selected as feature.
- Duplicated and redundant data were dropped.
- Missing and abnormal data were replaced with appropriated values.

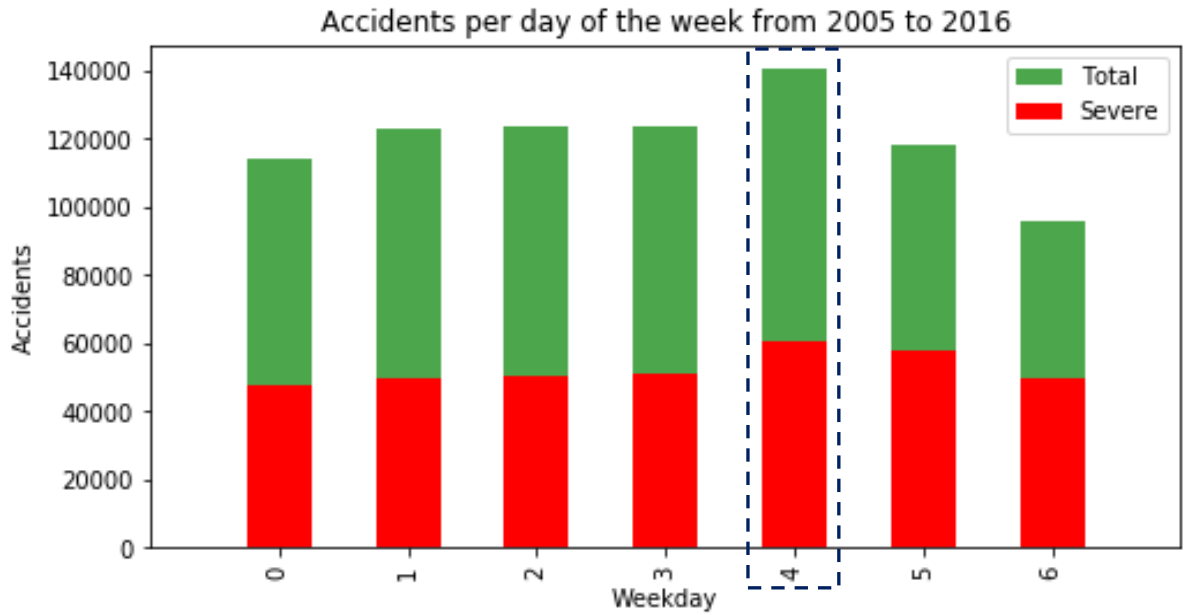
ACCIDENT PER YEAR 2005-2016

The accident rate (per year) decrease to stable after 2013.



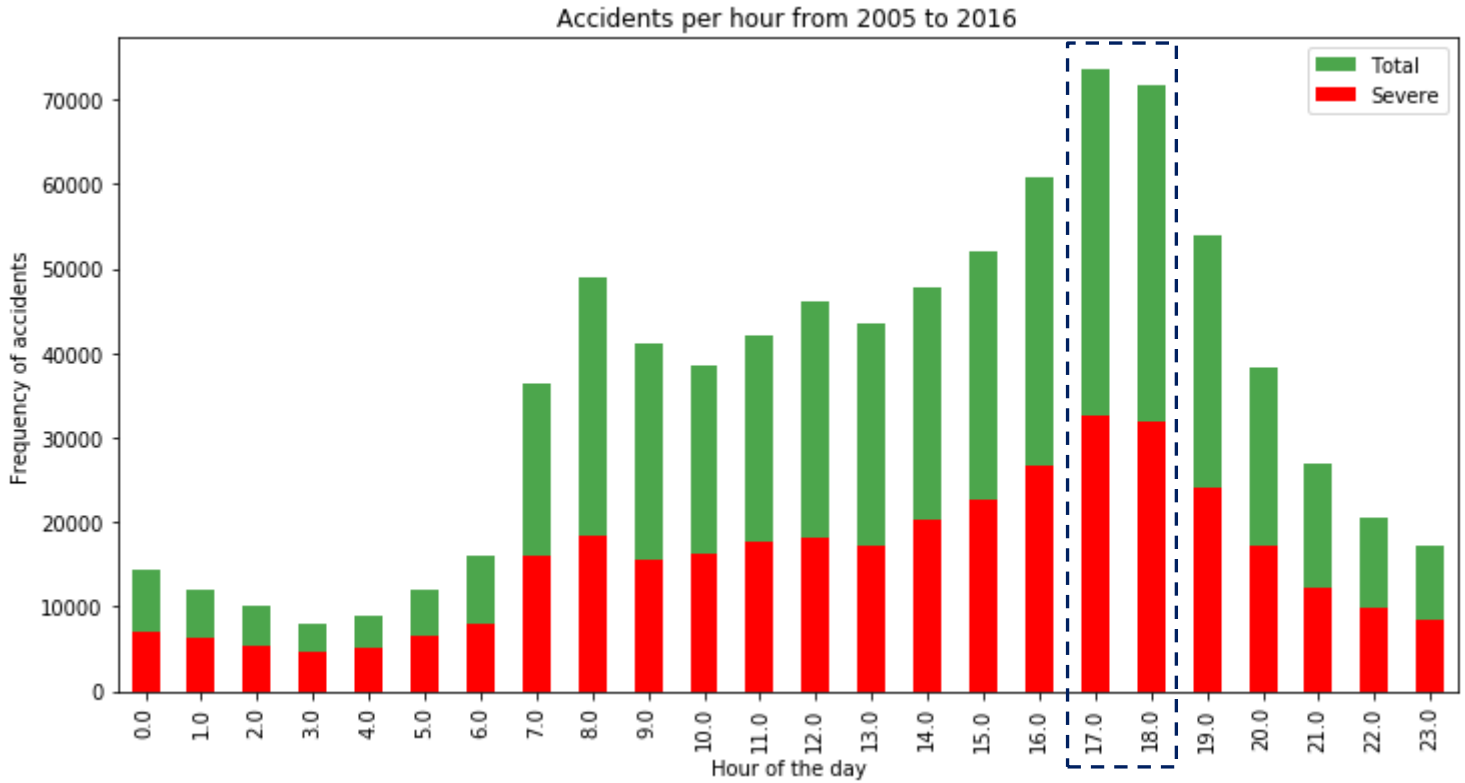
ACCIDENT PER DAY OF THE WEEK

The accident rate (per year) will reach to PEAK on Friday.



ACCIDENT PER HOUR OF A DAY

The accident rate (per year) will reach to PEAK at 17:00—18:00.



MODEL DEVELOPMENT

Algorithm in the Model development:

- **Random Forest,** 10 decision trees
- **Logistic Regression**, C=0.001
- **K-Nearest Neighbor**, K=16
- **Support Vector Machine**, 7500 sample data

MODEL ACCURACY COMPARISON

Best Performance Model: *Random Forest*.

Algorithm	Jaccard	f1-score	Precision	Recall	Time(s)
Random Forest	0.722	0.72	0.724	0.591	6.588
Logistic Regression	0.661	0.65	0.667	0.456	6.530
KNN	0.664	0.66	0.652	0.506	200.58
SVM	0.659	0.65	0.630	0.528	403.92

CONCLUSION AND FUTURE DIRECTIONS

- Built useful models to predict traffic accident severity.
- Accuracy of the models has room for improvement.
- Take more features into consideration.

Thanks!

2020-10-07