Classifying Risk of Injury for Traffic Crashes in the City of Chicago

Lorela Blaka
Allison Gao
Raylin Soriano

Presentation Outline

- Purpose of Analysis
- Data & Methods
- Results
- Conclusions
- Future Steps

Purpose of Analysis

Stakeholder: The Chicago Department of Transportation (CDOT)

Target: Predicting levels of injury (mild, medium, and severe)

Question of interest: What variables can help us predict different levels of injury from car crashes?

Data & Methods

- **Data**: 2019 Chicago car crashes
  - N = 550,000
  - Source: Chicago Data Portal

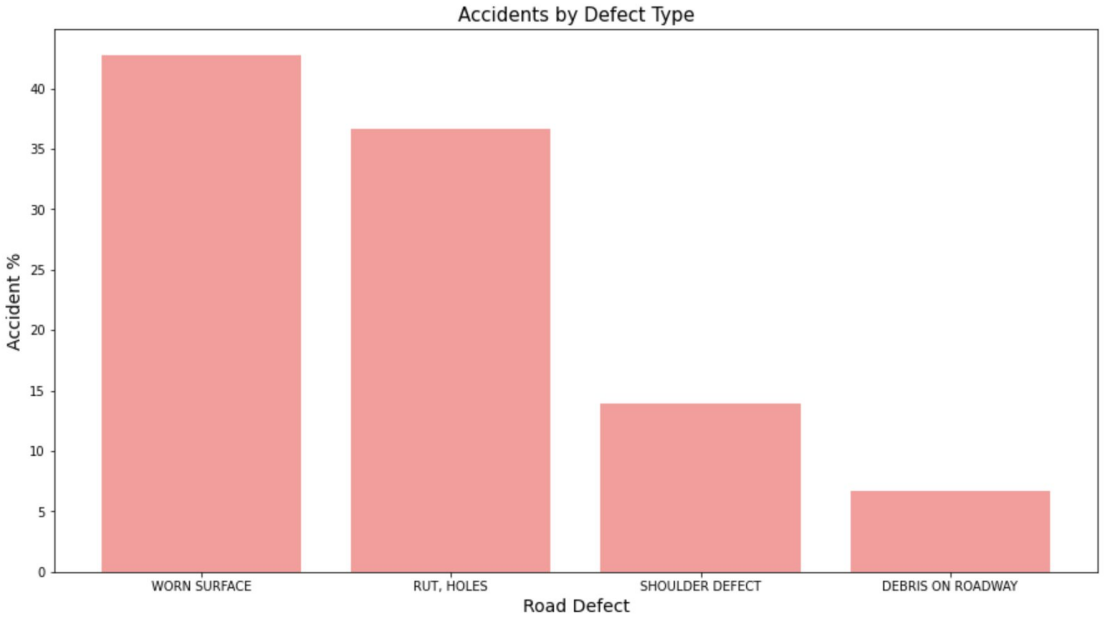
- **Method:**
  - Several classification models
  - 9 main features

Modeling Results

	Score Type	Dummy	Logistic	Random Forest	XGBoost
Training Data	Accuracy	0.8729	0.8720	0.9224	0.9715
	Macro Precision	0.2906	0.3290	0.7690	0.9275
Test Data	Accuracy				0.9488
	Macro Precision				0.8690

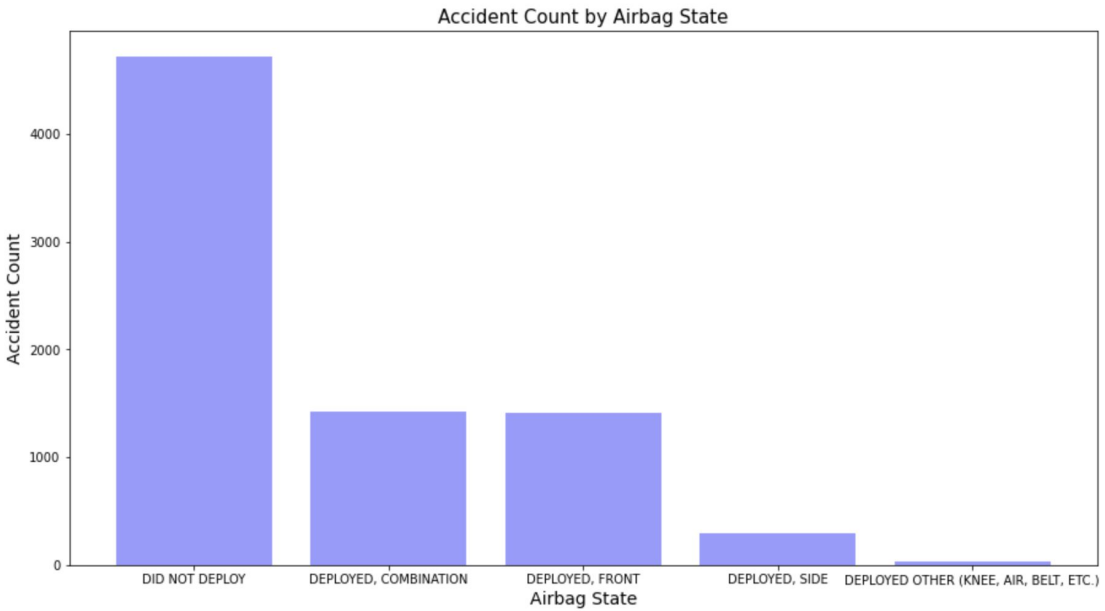
1

The percent of accidents for most severe injuries was highest when the road was defected



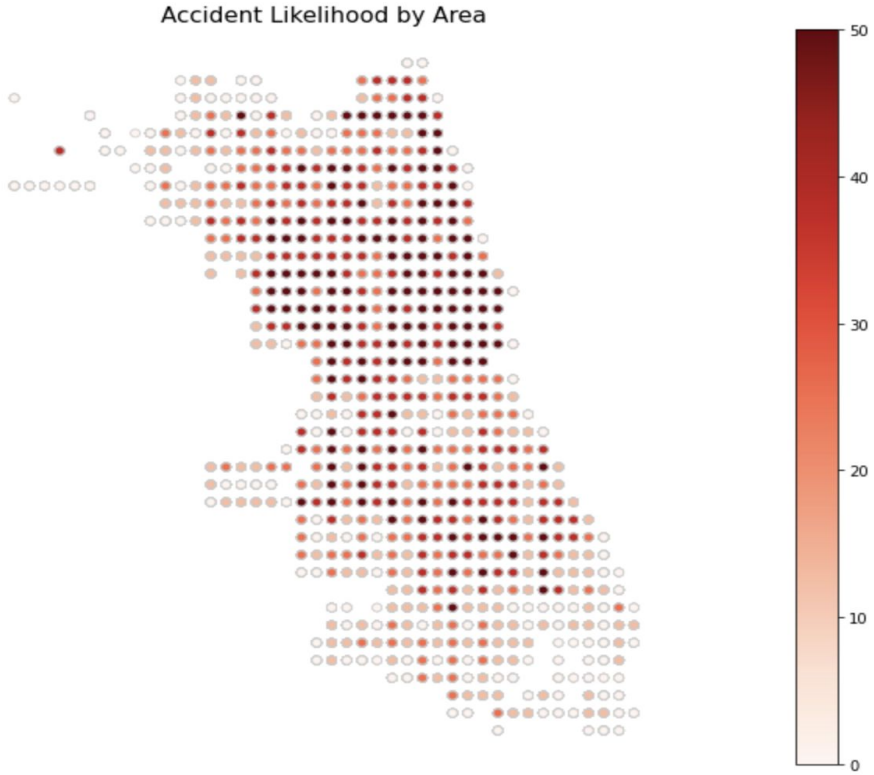
2

The accident count for most severe injuries was highest when the airbag was not deployed



3

The accident likelihood is highest in areas of Northern Chicago



Recommendations

1. Fix road infrastructure to reduce severe level of injury
2. Investigate airbag safety issues
3. Implement safety measures in high accident zones (i.e. address congestion issue, implement more safety signs)

Future Steps

- Increase dataset to include greater number of years

- Further modeling and hyperparameter tuning. For example, compare difference between rural and urban Chicago

Thank you!

Questions, Comments, Feedbacks