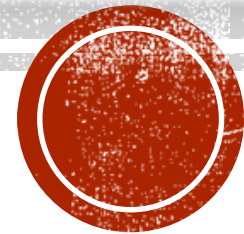# CHICAGO CRIME FORECAST

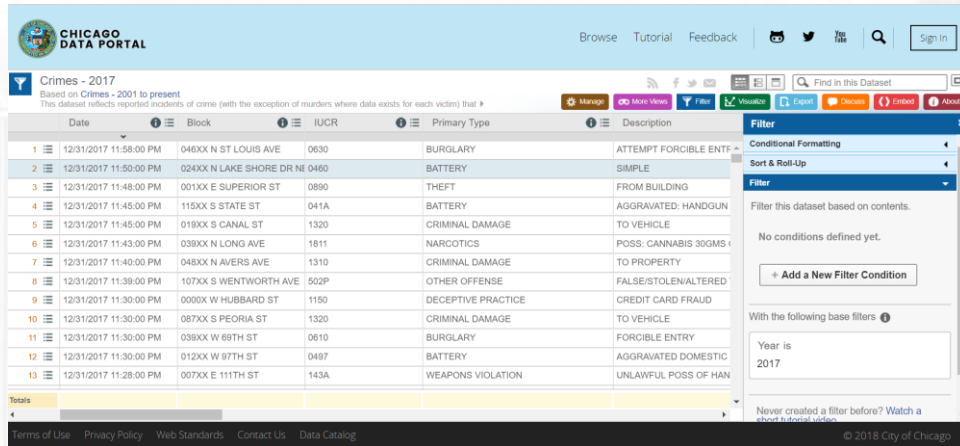It's really not that bad

Varun Gupta – MSiA423

MOTIVATION

1. Provide means to compare crime rates in different neighborhoods of Chicago, especially for visitors who want to look up zip codes.

2. Explore the curious relationship, even studied by the Chicago Tribune, between crime and the weather.

3. Opportunity to explore Chicago's public data portal.

4. **Clear MSiA423**

http://ec2-18-188-134-206.us-east-2.compute.amazonaws.com:5000/

DATA STUFF

City of Chicago Data Portal

Dark Sky API

- Historical Crime Data with Ward Information dating from Jan 2011- Feb2018. Data included date, ward, crime type, category, etc.
- Official Ward to Zipcode mapping data
- Historical Weather data for Chicago on each day across these years, courtesy Dark Sky API.
- Forecast data from Dark Sky, pulled multiple times a day using Cron to store it in RDS.
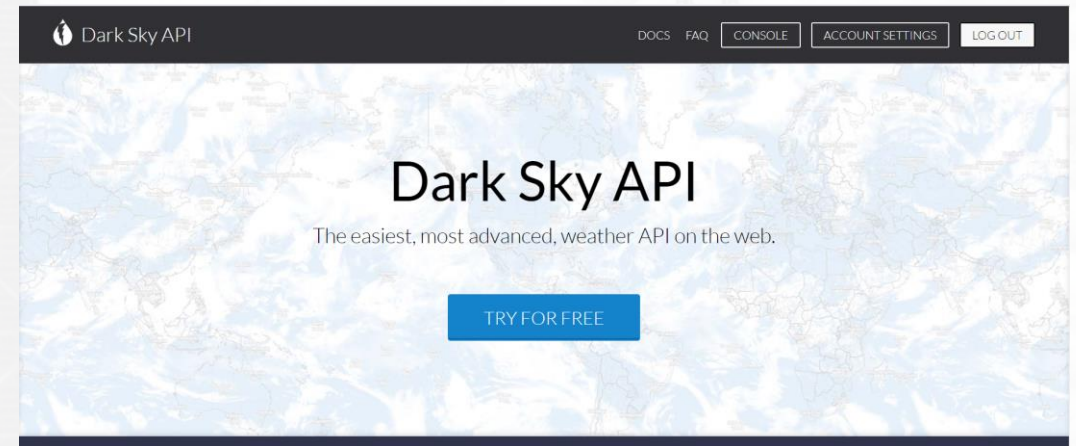
MODEL/ALGORITHM

▪ Data aggregated by day and zip-code, joined with weather data. Response variable for regression was total crime count.
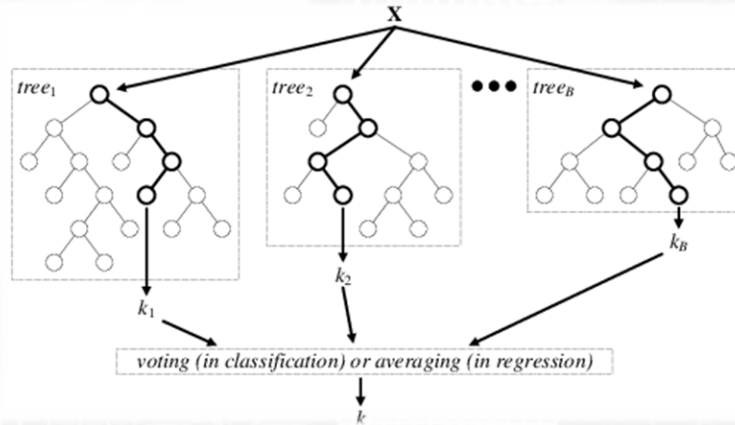
▪ Finally used Random Forest model (Looking to upgrade this).

  ▪ Why? Trees handle categorical variables with many levels very well.

  ▪ Ensembles strongly improve variance without much increase in bias.

  ▪ Success Criterion - Tuned for cross-validated SSE in crime count prediction.

  ▪ Tuned parameters – 950 trees, max depth of 10

INSIGHTS AND RESULTS

- Chicago Data Portal -> Amazing Initiative

- ***Apparent*** temperature more significant than real temperature -> Used apparent temp. for model and predictions.

- Zipcode, temperature and precipitation intensity were all very significant predictors.

Other Insights:

- Insight – APIs can go sideways and leave you depressed.

- Scikit-learn doesn't do trees right.

- *Insight* – I can myself make a working web application that I can see through my browser using an actual URL: Mind=blown

QUESTIONS?

THANK YOU!

- Contact:
  - Email: varungupta2018@u.northwestern.edu
  - Phone: +1-312-964-0262
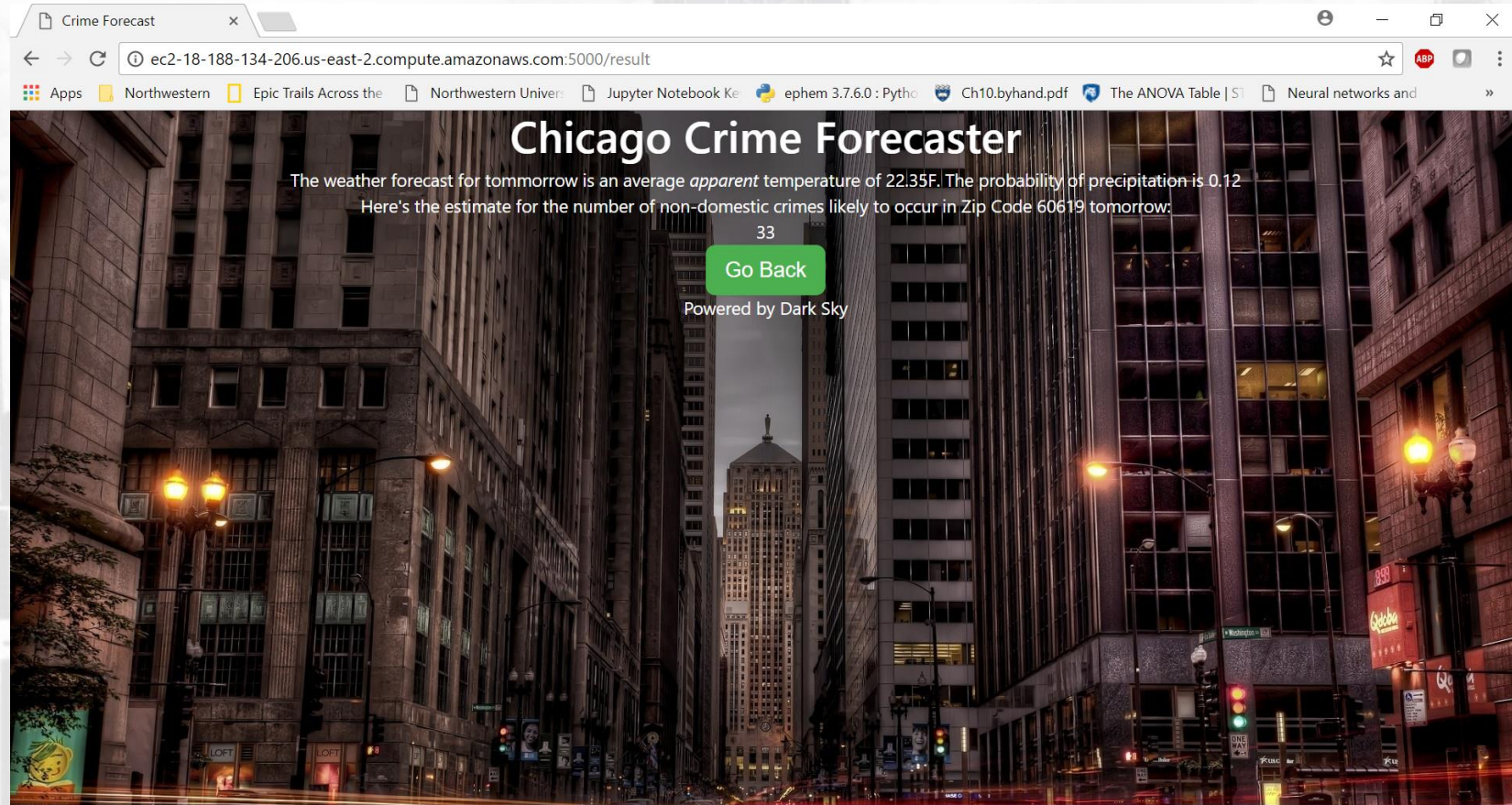 - Github: varungupta93

SCREENSHOTS

Minimalist Home Page

SCREENSHOTS

Results page (Maybe don't go to this Zipcode)