Crime Predictor

Team 30

OBJECTIVES

- Build a visual analytics tool to assist law enforcement by predicting the likelihood of different types of crime in a given geographical area.

- Identify a specific set of features that lead to increased crime instances (like a street with many abandoned houses, graffiti, etc.) so that steps can be taken to rectify them.

CURRENT PRACTICES & LIMITATIONS

Current practices:

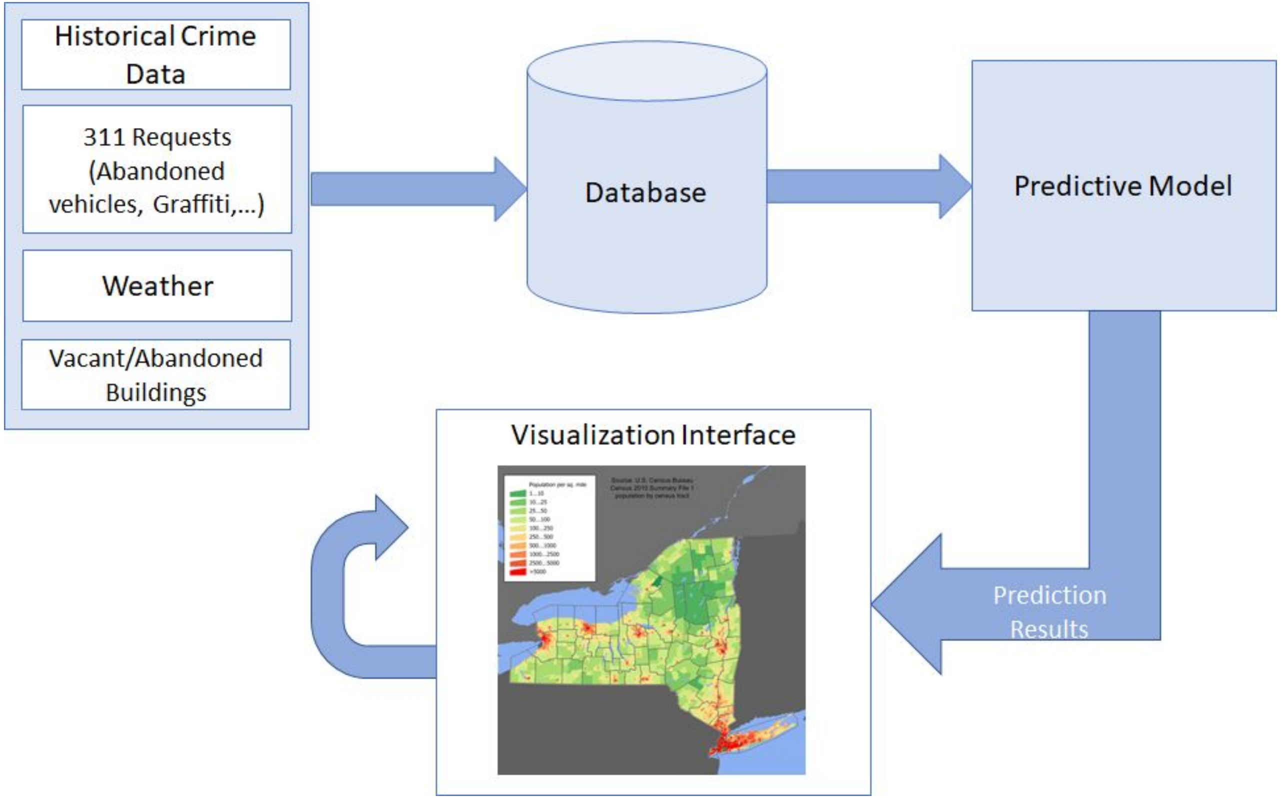
- Decision making based on experience, intuition and hunches.
- Use of historical crime records and individual criminal histories.
- Use of surveillance cameras

Limitations:

- Other local level metrics not taken into consideration.
- Bias.

Our Approach

Will these additional variables improve prediction accuracy?



Literature Survey

[4] proposes a visualization-based crime prediction tool that considers the type of crime to identify the crime hotspots and apply nearest neighbor approach to predict crime locations and show them on google maps.

[5] describes a study that identified the seasonal pattern between property crime and major festive seasons in Malaysia and suggested that a seasonal index can be a good indicator to be utilized for fighting crime.

[6] proposed a crime prediction method by merging multi-modal data across multiple domains with environmental context data. They integrated past criminal activity records for certain region and modeled them based on deep learning to predict the occurrence of crimes.

Risks & Payoffs

Payoffs:

- Improved prediction accuracy.
- Improved resource allocation based on predictions.
- Evidence-based policy making.
- Reduced crime incidents with better policing and patrolling.

Risks:

- Biased predictions based on demographic factors, historical data.
- Additional variables may not improve prediction accuracy.

Measuring Impact

- Can law enforcement use this tool to reduce crime incidents?
- Can this approach lead to better planning and resource allocation?
- Ease in data exploration and better decision making.

TIME & COST

- Developer Time - 200 hours
  - 40-50 hours for data exploration, cleaning and merging.
  - 80-100 hours for modelling and refining.
  - 40-50 hours for getting the visualizations and web-app ready.

- Server costs
  - Costs associated with remote web-app hosting and testing.

PLAN OF ACTIVITIES

Mid-term:

- Collect data from different sources.
- Clean and merge data.
- Train a base model to predict the likelihood of each type of crime in a given geographical area based on historical data.

Final:

- Build a web app that takes in selected constraints and shows the likelihood of types of crimes on a map.
- Design and build interactive visualizations for result exploration.
- Iteratively add other factors to the base model to improve prediction accuracy.
- Find the most critical local factors that contribute to crime.

REFERENCES

1. Pexels.com. (2018). *Crime Scene Do Not Cross Signage · Free Stock Photo*. [online] Available at: https://www.pexels.com/photo/crime-scene-do-not-cross-signage-923681/ [Accessed 16 Oct. 2018].
2. Commons.wikimedia.org. (2018). *File:New York Population Map.png - Wikimedia Commons*. [online] Available at: https://commons.wikimedia.org/wiki/File:New_York_Population_Map.png [Accessed 16 Oct. 2018].
3. Anon, (2018). [online] Available at: https://www.independent.co.uk/news/uk/home-news/police-big-data-technology-predict-crime-hotspot-mapping-rusi-report-research-minority-report-a7963706.htm [Accessed 16 Oct. 2018].
4. ToppiReddy, H., Saini, B. and Mahajan, G. (2018). Crime Prediction & Monitoring Framework Based on Spatial Analysis. *Procedia Computer Science*, 132, pp.696-705.
5. Ismail, S. and Ramli, N. (2013). Short-term Crime Forecasting in Kedah. *Procedia - Social and Behavioral Sciences*, 91, pp.654-660.
6. Kang H-W, Kang H-B (2017) Prediction of crime occurrence from multi-modal data using deep learning. PLoS ONE 12(4): e0176244. https://doi.org/10.1371/journal.pone.0176244
7. Hongjian Wang, Daniel Kifer, Corina Graif, and Zhenhui Li. 2016. Crime Rate Inference with Big Data. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). ACM, New York, NY, USA, 635-644.