Crime reduction with machine learning: a comparative analysis

Credits

Significance to the discipline

- Machine learning for crime prediction has been explored academically, but does not have widespread adoption
- Potentially due to academic focus on model performance over practicalities of implementation
- Challenges include:
  - Insights are often not actionable
  - Lack of interpretability
- Study seeks to combine performance, interpretability and computational efficiency metrics to evaluate true "usefulness"
- The "usefulness" metric would be applicable to domains outside of crime prediction

Research questions

- Which metrics should be used to measure machine learning model performance, interpretability, and computational efficiency, to provide an overall blended score.
- Which machine learning models provide the best overall performance, interpretability and computational efficiency when predicting locations of crime.

Aims and objectives

Aims

Comparing machine learning performance for crime prediction is difficult due to different datasets and metrics. This study will define a methodology and use a single dataset to compare multiple solutions.

Objectives

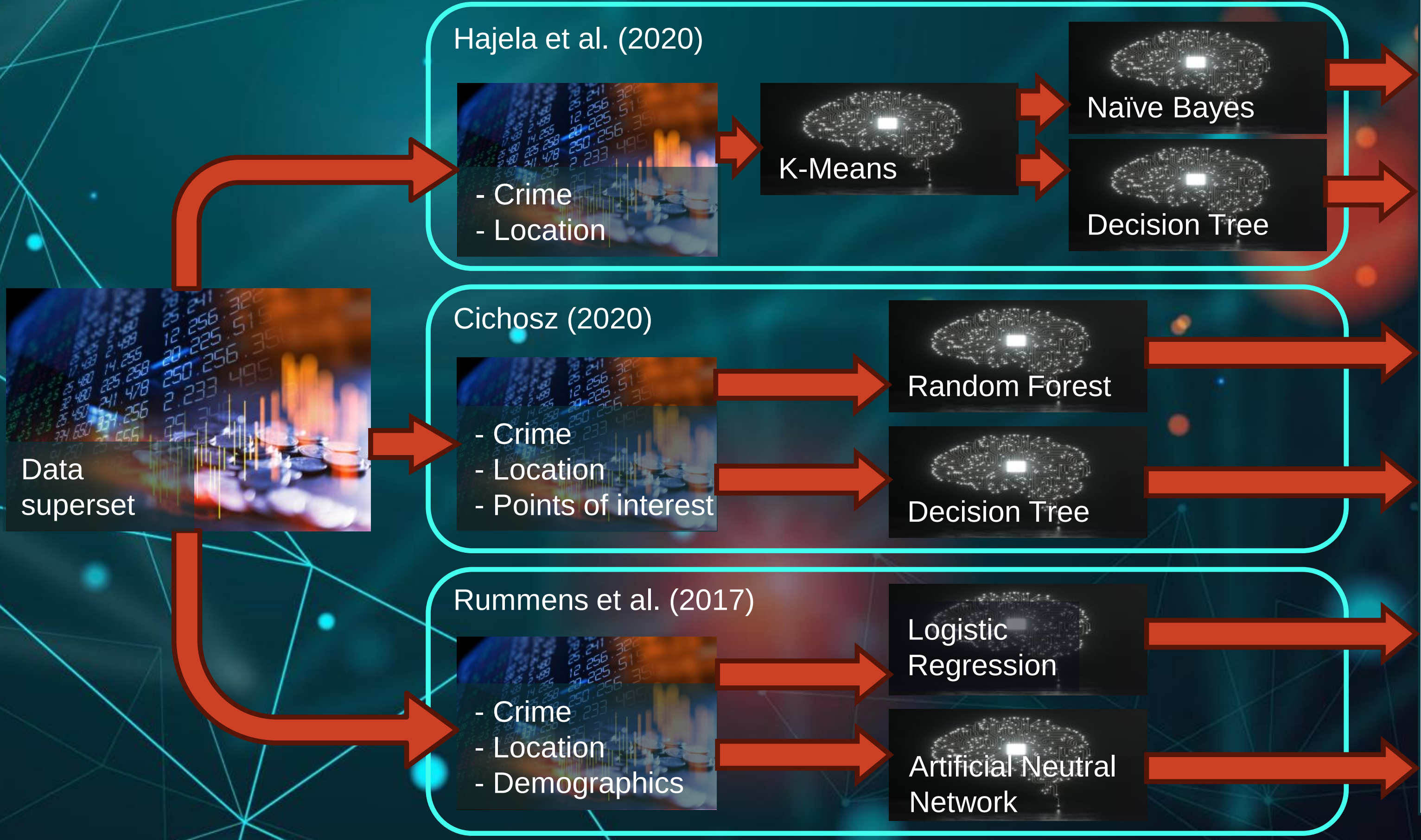
1. Review best practices for measuring machine learning model interpretability and computational efficiency
2. Propose a methodology to measure model usefulness incorporating those metrics
3. Compare three published solutions using the same dataset to prove the methodology and recommend the best model for predicting crime

Key literature

- Literature review (Feaviour, 2024) identified various solutions
- Challenges (varied by paper) included:
  - Lack of real-world applicability
  - Lack of attention to interpretability
  - Lack of attention to computational efficiency
- Three solutions selected with real-world applicability to test against each other
- Additional literature search to identify best practice for assessing interpretability and computational efficiency

Reference	Machine learning model(s)	Solution overview	Granularity	Relevence to day-to-day policing	Limitations
Jain et al. (2017)	K-Means	Cluster by crime type and location	Variable; typically a few streets	Low	Not a predictor. More appropriate for strategic resource planning
Iqbal et al., (2013)	Decision Tree Naïve Bayes	Demographic data to predict if a state has high, medium or low violent crime	State	Low	Predictions at state level so limited use in day-to-day policing
Zaidi et al., (2020)	Random Forest Support Vector Machine	Demographic data to predict if a state has high, medium or low violent crime	State	Low	Predictions at state level so limited use in day-to-day policing
Ahishakiye et al. (2017)	Decision Tree	Demographic data to predict if a county has high, medium or low violent crime	County	Low	Predictions at county level so limited use in day-to-day policing
Safat et al. (2021)	Logistic regression Decision Tree Random Forest Multilayer Perceptron Naïve Bayes Suport Vector Machine GXBoost K Nearest Neigbour	Compare classifier algorithms using two datasets then test time series prediction using ARIMA	District	Potential to inform short-term resource planning	Longer-term accuracy is questionable
Sivaranjani et al. (2016)	K Nearest Neighbour K-Means Agglomerative Clustering DBSCAN	Classify crimes with KNN then cluster crimes of the same category to give each city a high/medium/low preopensity per crime type	City	Potentially useful for city-level planning	Lacks granularity to inform day-to-day policing
Hajela et al. (2020)	K-Means Naïve Bayes Decision Tree	Cluster crime hot spots then classify crime type at each hot spot	Variable based on cluster size; typicaly district but centre point is visible	Hot spots are based upon coordinates, so potential for targetted interventions	Each cluster only has one class, so intervensions would be based upon most prevelent crime only
Cichosz (2020)	Logistic Regression Support Vector Machine Decision Tree Random Forest	Points of interest (POI) aggregated into cells in a geographic grid to predict hotpots of different crime types based upon POI atributes. Performed well when trained on one area and tested on another	300m²	Predicting crime types based upon POI could enable targetted policing	The grid doesn't give a precise prediction of where crime will occur
Rummens et al. (2017)	Logistic Regression Artificial Neural Network	Hotspots by crime type plotted in a grid, split by day and night. Predictions using rolling time window	200m²	Predictive patterns by day and night could help with proactive policing	Only three crime types, but more could be added
Stalidis et al. (2021)	CNN LSTM	Time series maps with crime plots passed through CNN to create feature vectors then LTSM to extract feature vectors to predict emerging crime hotspots	500m²	Predicted emerging crime hotspots could be used to target police activity	Relatively high compute required in pre-processing the maps and running the algorithms

Methodology



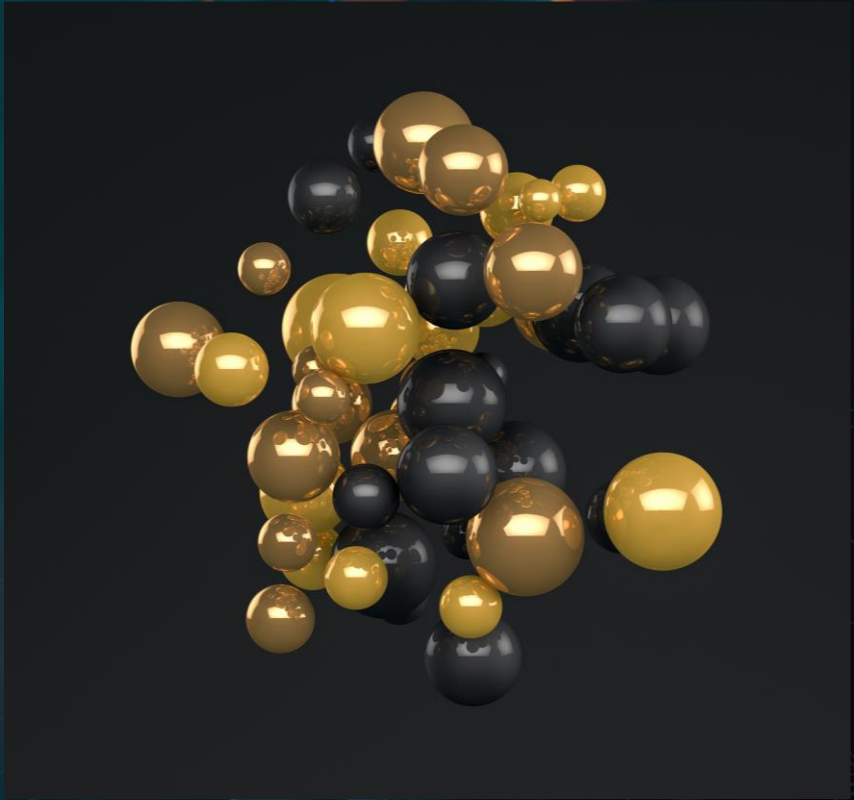
Accuracy	Precision	Recall	F1-Score	AUC	Interpretability	Computational efficiency

Ethical considerations

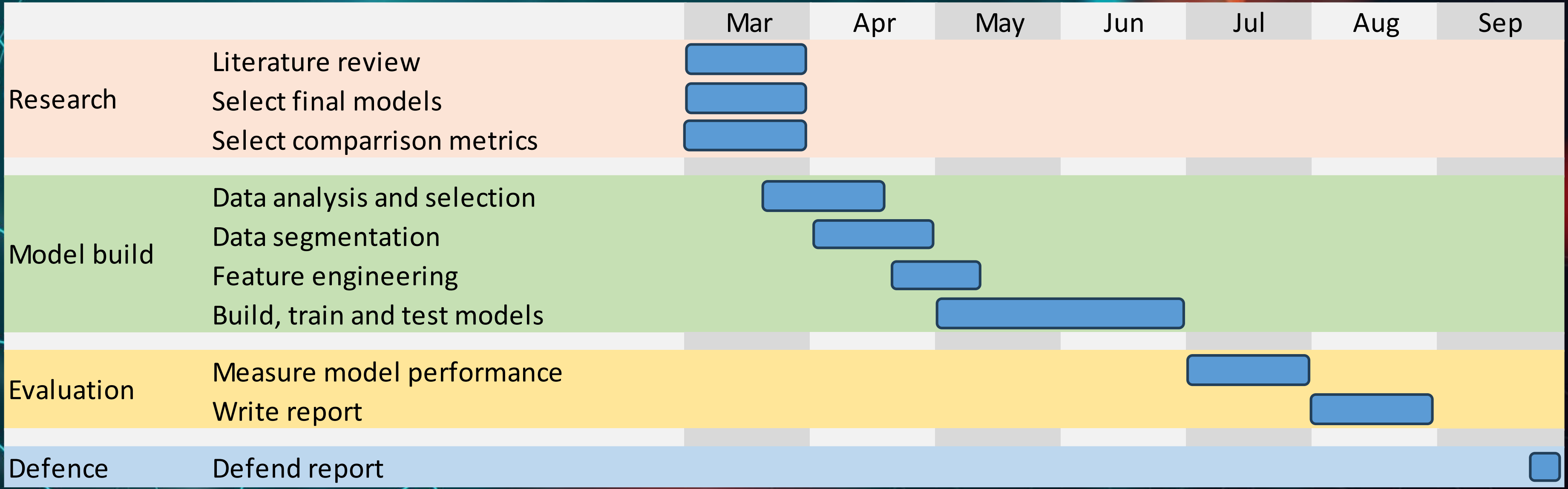
- Data used should not contain personal information
- Outcome will provide a more ethical model for machine learning model assessment due to inclusion of interpretability and computational efficiency metrics

Artefacts to be created

- Three datasets; one for each solution
- Six trained and tested machine learning models, all available on GitHub
- One report comparing the results of the six models with a recommendation



Schedule


Mar
Apr
May
Jun
Jul
Aug
Sep
Research
Literature review
Select final models
Select comparrison metrics
Model build
Data analysis and selection
Data segmentation
Feature engineering
Build, train and test models
Evaluation
Measure model performance
Write report
Defence
Defend report

References

- To be added
- TBC