

# Vehicle Safety Analyzer

Is Your Car on a Thief's Wishlist? See If You're at Risk!

Angie Jaimes, Negin Djalali, Kate Reva

Summer 2024

# The Problem

Car theft in Canada has surged dramatically, with a car stolen every **5 min**



**300%** increase in Toronto since 2015, affecting over a million drivers annually



# Andrew's Story

— — —



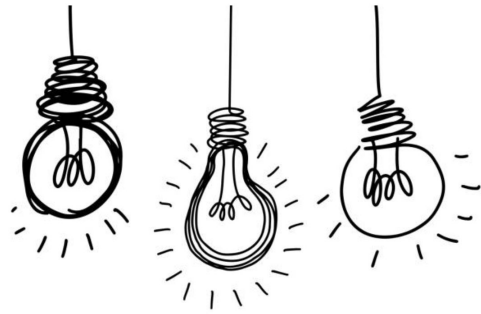






# Vehicle Safety Analyzer

---



Our **mission** is to empower vehicle owners with actionable insights to improve security measures and make informed decisions about vehicle safety.

# Target Audience

---



**10M**  
registered drivers



# Impact

---

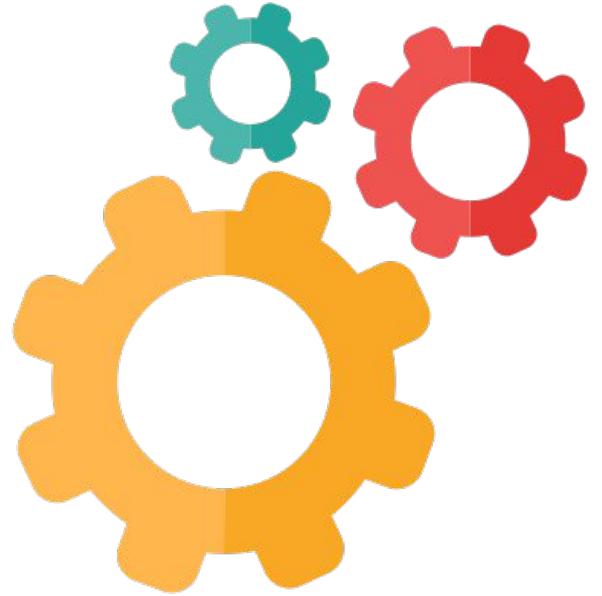


# Demo

---

# Technical Discussion

---



# Data



## Vehicle population data - 2022

**Dataset:** All registered vehicles

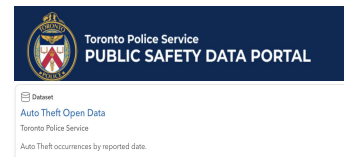
**Source:** Government of Ontario

### Data features:

- Vehicle class
- Make
- Model
- Model-Year
- Total

### Data characteristics:

- Registered cars in 2022
- ~ 18,000 rows for passenger cars
- Over 5.5 million cars



**Dataset:** Vehicles reported **stolen**

**Source:** Toronto Police

### Data features:

- Date and time (reported and actual)
- Incident id
- Neighborhood
- Geo information Long and Lat
- Premise Type

### Data characteristics:

- 2014-2024
- ~ 55,000 unique cases

# The Model Goal

---

**Is Your Car on a Thief's  
Wishlist? See If You're at Risk!**

**Outcome Variable:** Category that indicates risk of a car being stolen

# Data challenges

— — —



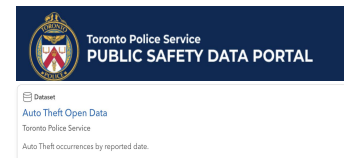
**Vehicle population data - 2022**

**Dataset:** All registered vehicles

**Source:** Government of Ontario

**Data missing for prediction:**

- Neighborhood information



**Dataset:** Vehicles reported **stolen**

**Source:** Toronto Police

**Data missing for prediction:**

- Vehicle make, model, year

# Data challenges



Vehicle population data - 2022

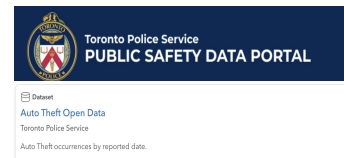
**Dataset:** All registered vehicles

**Source:** Government of Ontario

**Data missing for prediction:**

- Neighborhood information

**Solution:** Leverage neighborhood information from stolen vehicles to randomly generate neighborhoods



**Dataset:** Vehicles reported **stolen**

**Source:** Toronto Police

**Data missing for prediction:**

- Vehicle make, model, year

**Solution:** Leverage vehicles stolen in Ottawa to generate Make, Model and Model Year for Ontario



### Vehicle population data - 2022

~5.5 mln registered vehicles in Ontario

Add Neighborhood, Location type synthetic data from Toronto stolen vehicles. Randomly generate month, hour and Day of the Week

~55,000 stolen vehicles in Toronto

Add Make, Model, vehicle year from Ottawa Police Reports

MAKE, MODEL, VEHICLE YEAR (2005 and above)

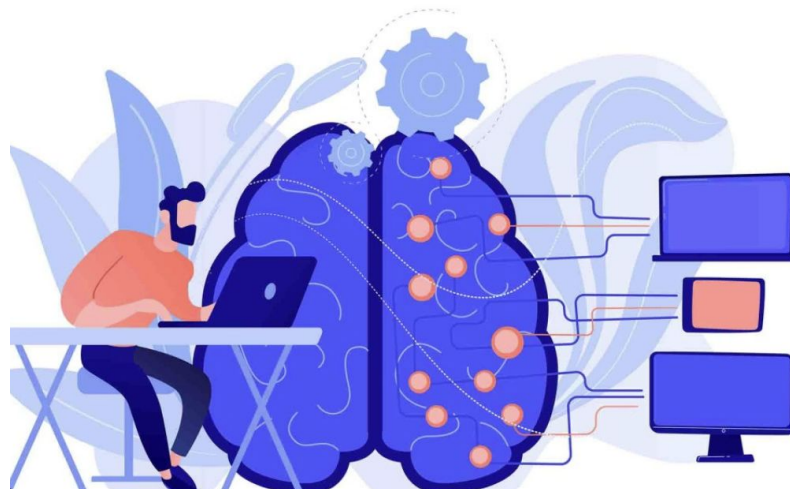
	MAKE	MODEL	MODEL-YEAR	MAKE_MODEL_YEAR	MAKE_MODEL	Stolen	MONTH	DOW	HOUR	LOCATION_TYPE	NEIGHBOURHOOD	Latitude	Longitude
0	ACUR	ARL	2012	ACUR--ARL--2012	ACUR-ARL	0	4	3	17	Other Commercial / Corporate Places (For Profi...	West Humber-Clairville	43.712560	-79.595638
1	ACUR	ARL	2012	ACUR--ARL--2012	ACUR-ARL	0	8	3	19	Single Home, House (Attach Garage, Cottage, Mo...	West Humber-Clairville	43.712560	-79.595638
2	ACUR	ARL	2012	ACUR--ARL--2012	ACUR-ARL	0	12	3	14	Gas Station (Self, Full, Attached Convenience)	Bedford Park-Nortown	43.728451	-79.421177
3	ACUR	ARL	2012	ACUR--ARL--2012	ACUR-ARL	0	4	2	22	Single Home, House (Attach Garage, Cottage, Mo...	Scarborough Village	43.741775	-79.216566



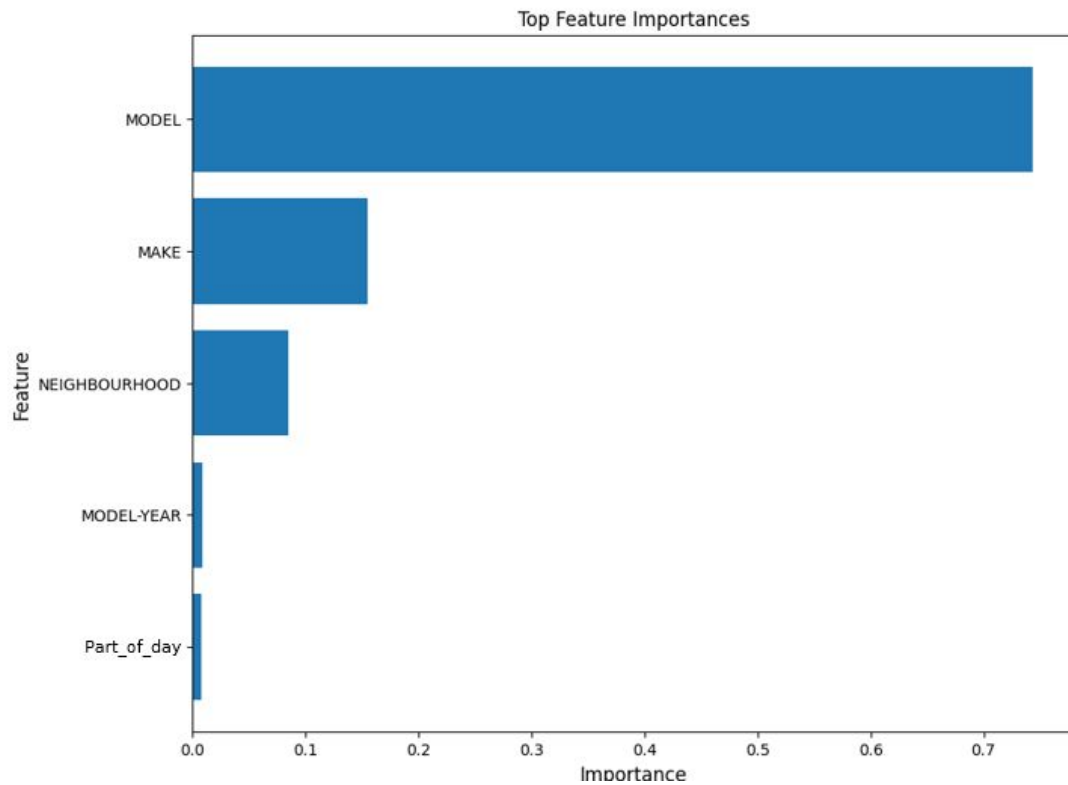
# Modeling

---

- Models that were built
  - Random Forest
  - XGBoost
- Split the Data to test and train (80/20)
- Preprocess the categorical columns



# Feature Selection



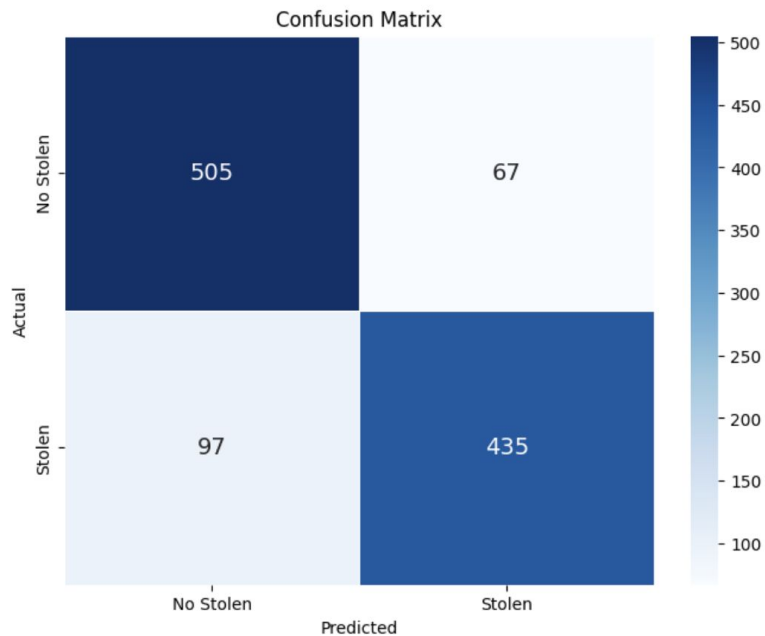
# XBoost

— — —

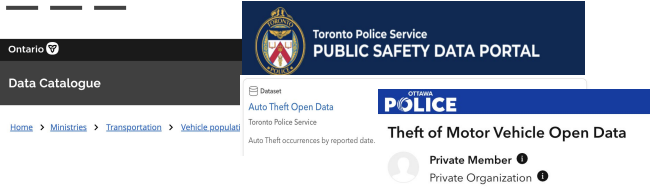
**Improved Performance**  
**Class Imbalance Handling**

**Opportunity:** Continue reducing false negatives

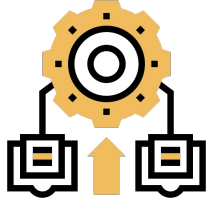
	precision	recall	f1-score	support
0	0.84	0.88	0.86	572
1	0.87	0.82	0.84	532
accuracy			0.85	1104
macro avg	0.85	0.85	0.85	1104
weighted avg	0.85	0.85	0.85	1104



# Data Pipeline and ML Architecture



Data collection, ETL, synthetic data generation, EDA, data pre-processing



Train ML Model

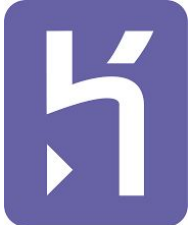


Save with **pickle**



Flask

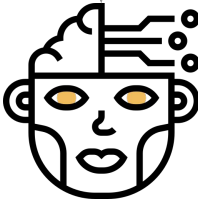
Create Prediction API and HTML



Deploy to Heroku



Setup Web Interface



Serve Predictions

# If we had more time...

---



# Conclusion

---

We hope our personalized recommendations will **improve** community safety.

By empowering drivers with actionable insights and effective safety tips, we contribute to creating a **safer** environment for all.



# Acknowledgements

---

We would like to thank our Capstone Instructors, Fred Nugen and Danielle Cummings for their actionable and thoughtful feedback throughout the process.

We are grateful for everyone who tested our product and provided feedback to make it better.