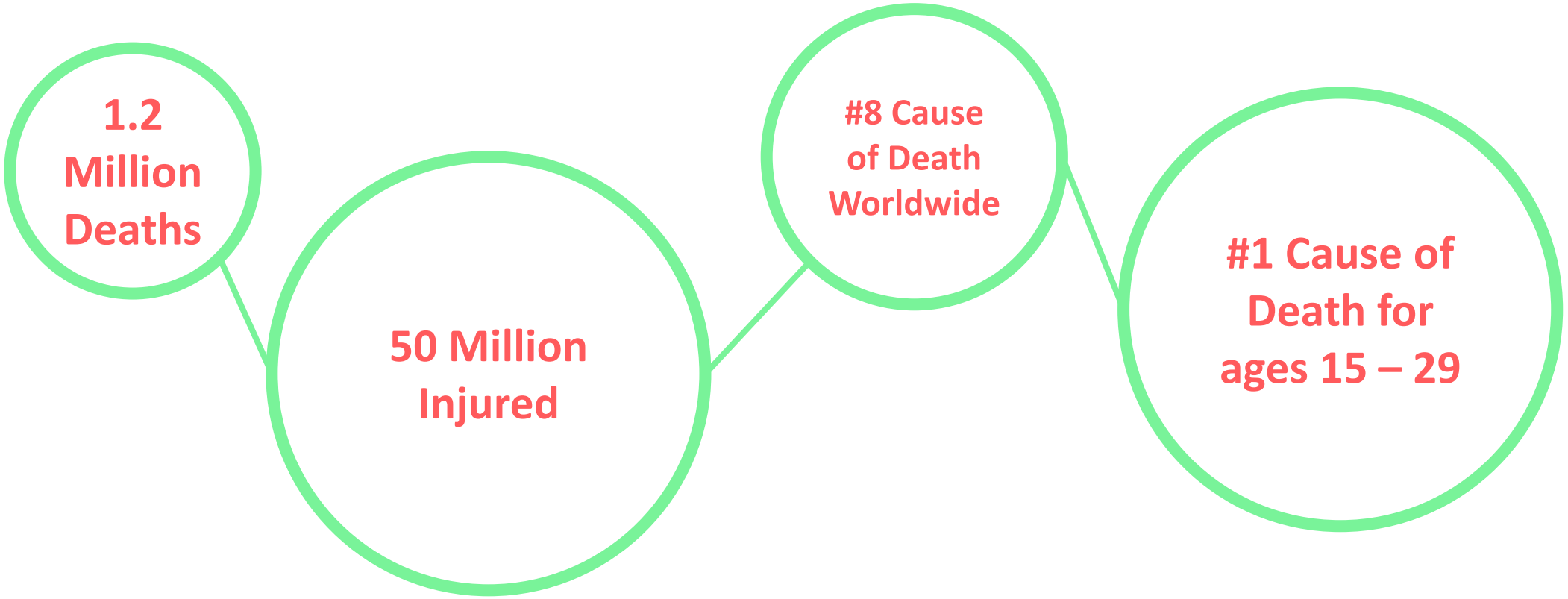
DriveAlive

Toward Smart Cities: Data-driven Road Accident Prevention

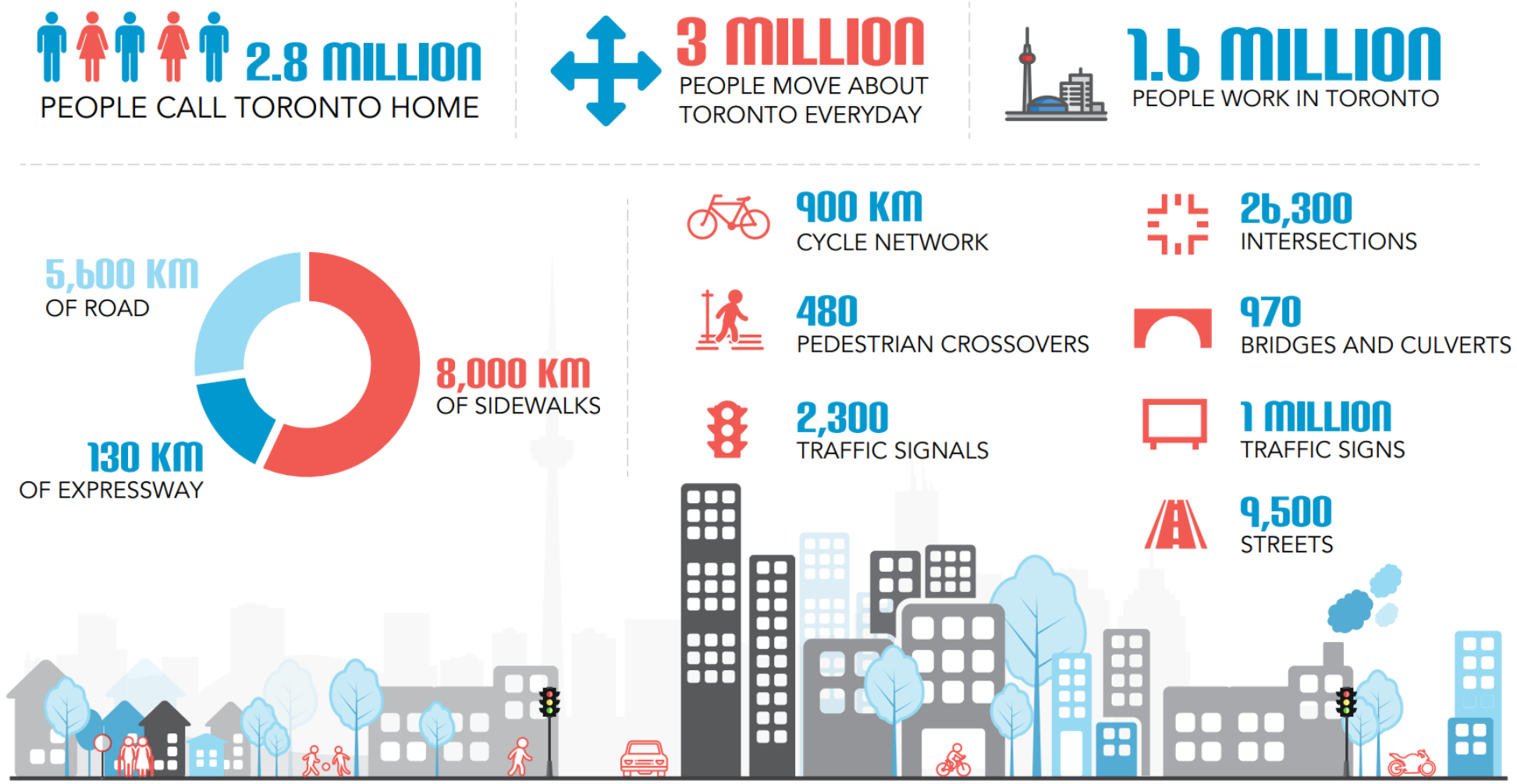
- Amanda Aleong
- Francis Cruz
- Youngmok Ko
- Shi Yu
- Han Hu
- Hao Xing

Smart Cities Need Smarter Roads




Toronto: Canada’s Largest Metropolitan City

Case Study: City of Toronto

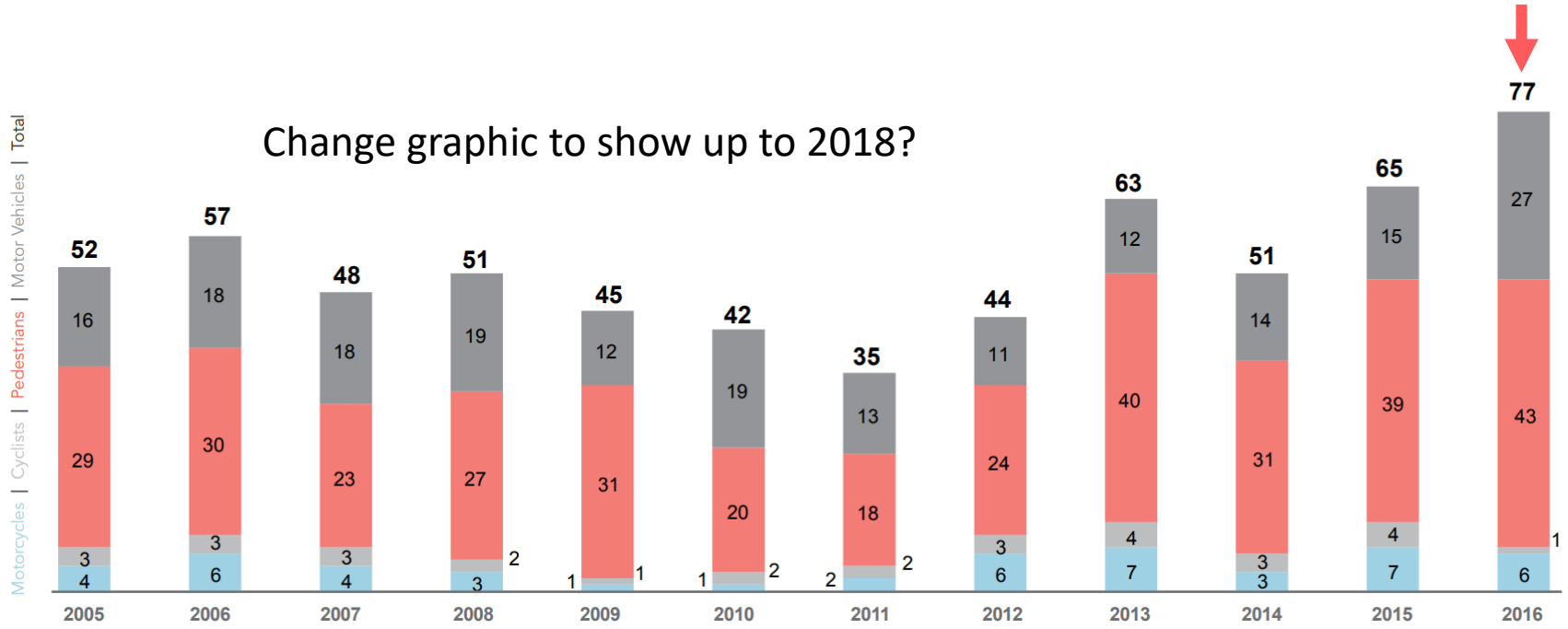


Source: City of Toronto Vision Zero Report

Toronto Collisions on the Rise

Despite efforts introduced by Toronto's VisionZero plan, fatalities have reached a 10 year high and stagnated around this value

Total Traffic Fatalities in Toronto, Jan. 1, 2005 - Dec. 31, 2016



Source: City of Toronto Vision Zero Report

Problem Definition

Current top-down approach: Slow acting policies and infrastructure changes with only marginal results

Need for bottom-up approach: Give people the tools they need to make safer decisions on the road

Target Stakeholders

PRIMARY USER: DRIVERS

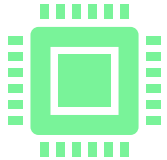
PRIMARY STAKEHOLDER:
MAPPING APPS

OTHER STAKEHOLDERS:
GOVERNMENT, TRANSPORTATION AUTHORITIES, POLICE, HOSPITALS

Value Proposition in Toronto

CAD 3 Billion in Social Costs Due to Road Accidents

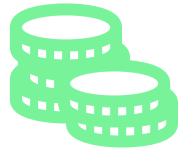
3.1 Million use Mapping Apps Daily



CAD 5.2k in Operating Costs

CAD 220 Million in Annual Healthcare Costs due to Road Accidents

Estimated CAD100k in Sales Revenue as Mapping Add-on



Over 13,600% Return on Investment

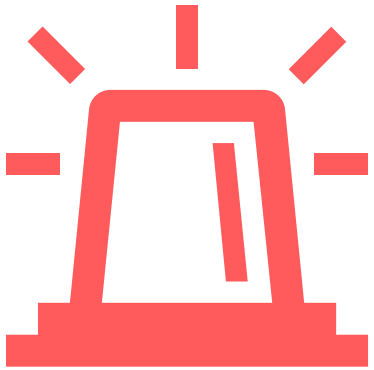
“ Safe roads for all by 2030” - WHO

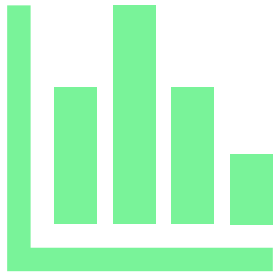
Can we wait for 2030?

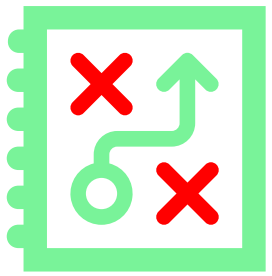
The time to drive is now…

So…

DriveAlive

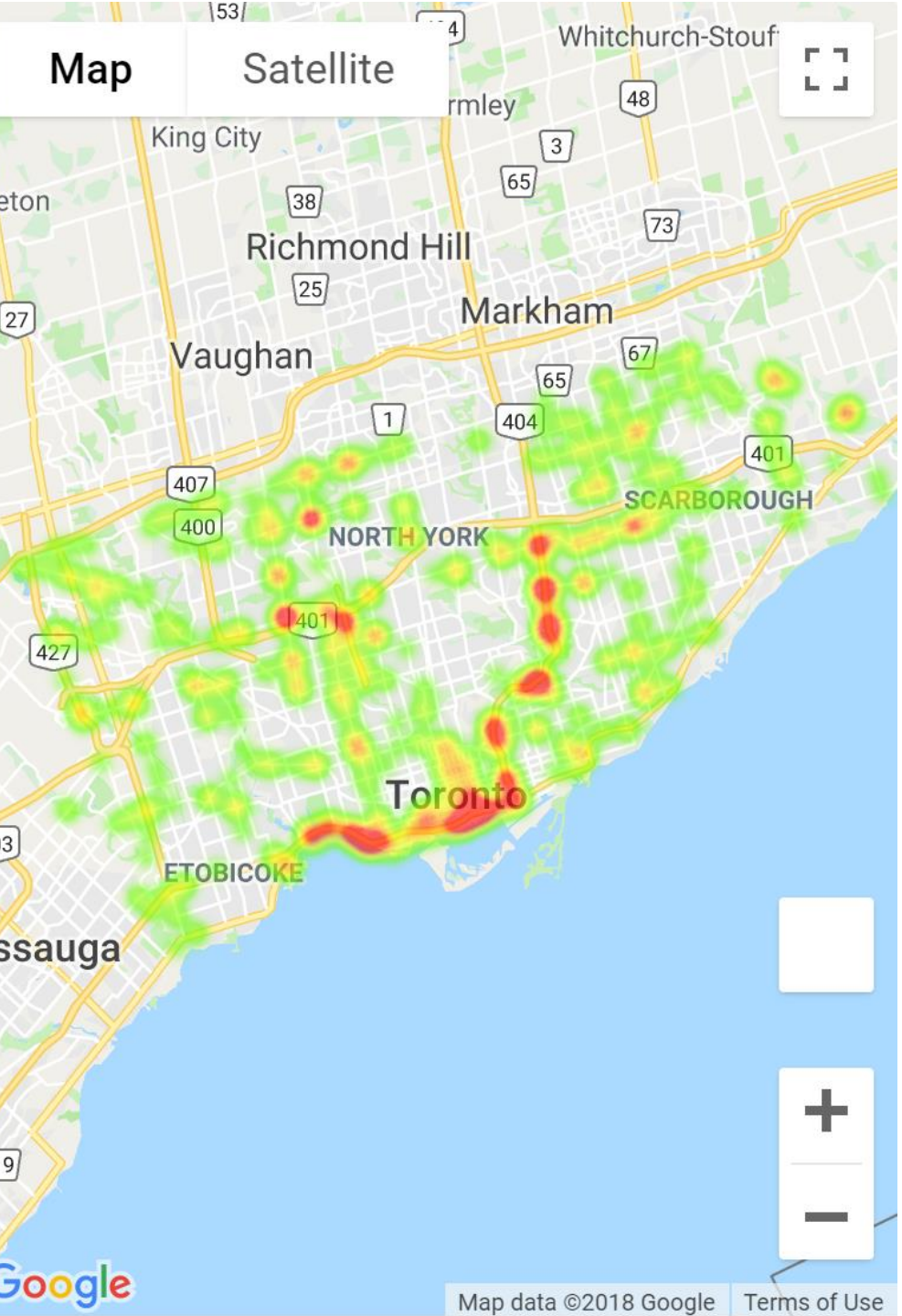






How Does it Work?

- Historical collision, traffic and road data
- Machine Learning algorithm for collision prediction
- Collision report suggests the safest (and fastest) route to your destination



Collision Mapping

Collision avoidance is a temporal problem

Collision avoidance is a spatial problem

Binary Classification

Historical Collision Data

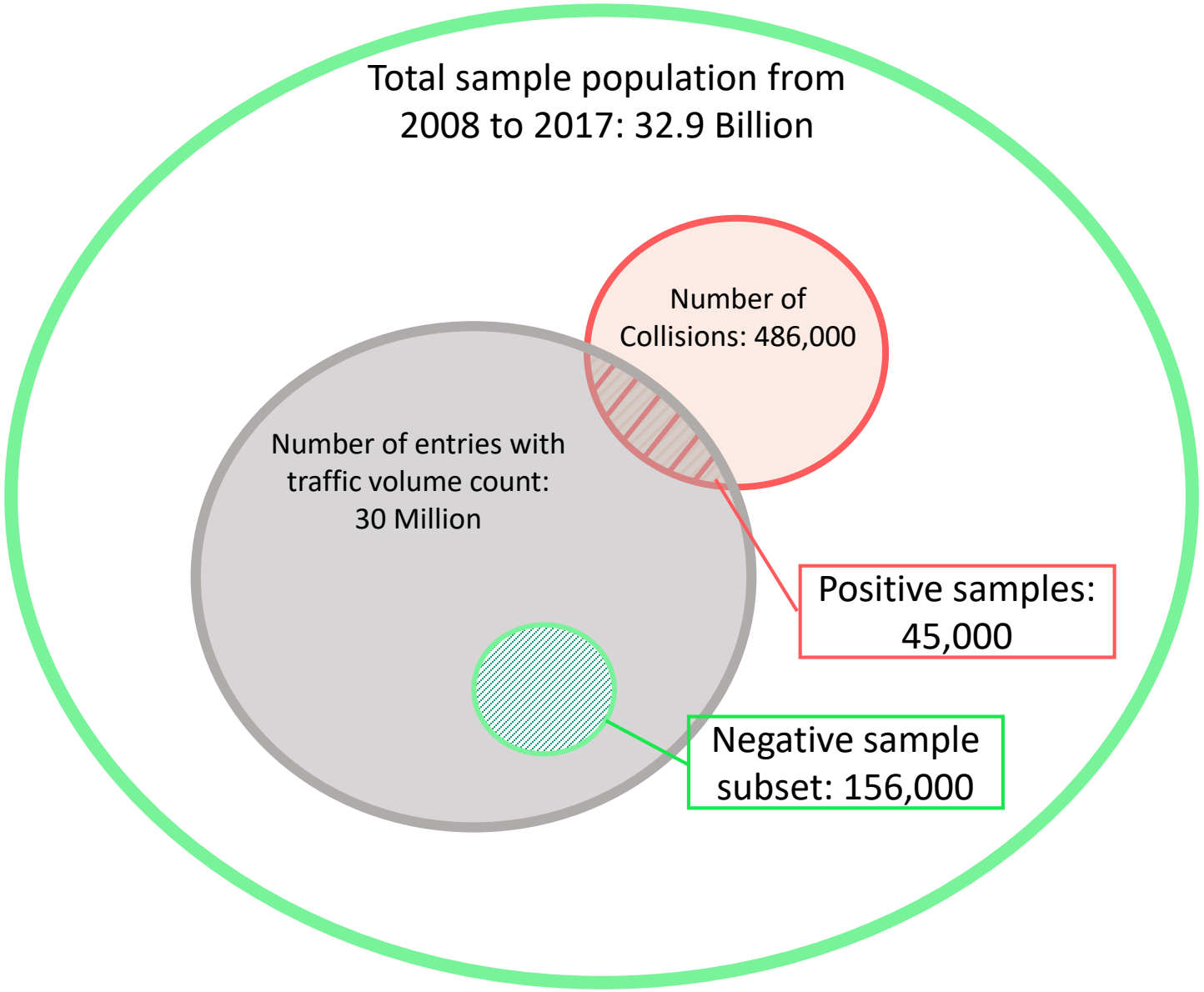
- Compile data from 5 data sets to fill feature columns: Collision Events, Volume, Road Data
- Discard cases with insufficient features available

Obtaining Negative Samples

- Random sample a road segment, day and time
- Check if the combination is in the collision dataset
- If not, fill in features and add to negative sample list

Clustering

- Toronto is huge!
- Cluster areas based on density of sample reports



Negative Sampling

Sample population:

- All possible combination of road segment and time over the last 9 years.

Positive samples:

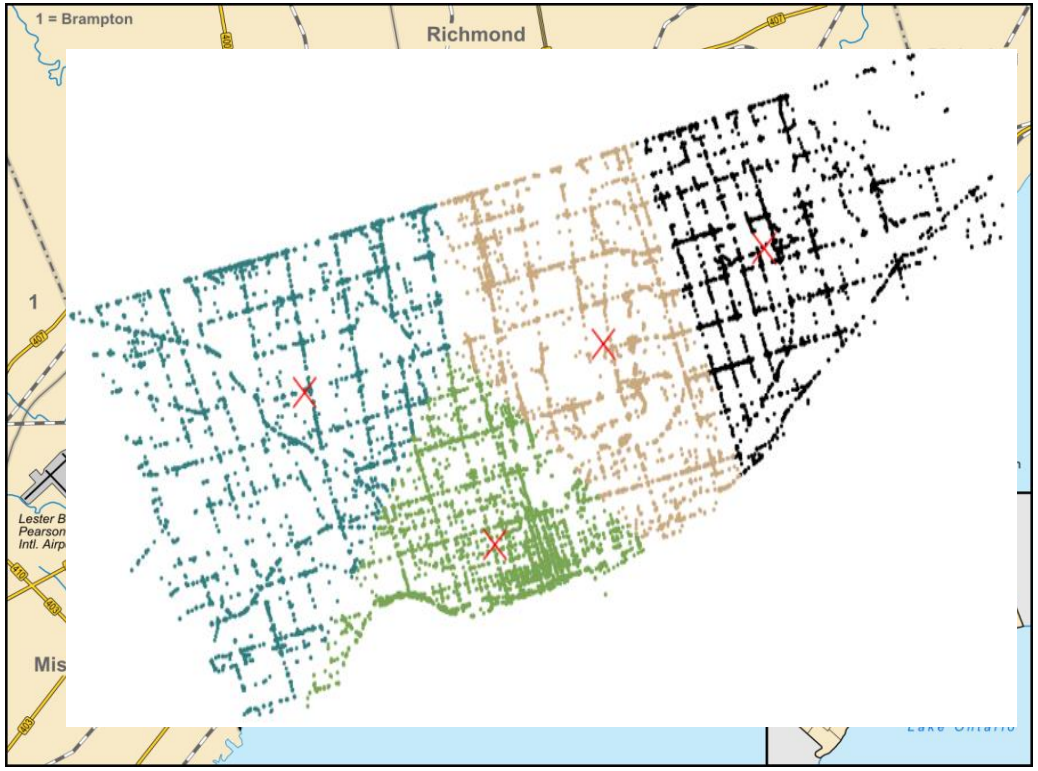
- All collision events with corresponding traffic volume count.

Negative samples:

- Combination of road segment and time that did not result in a collision, but has traffic volume count.

Feature Engineering: K-Means Clustering

- Sample points were clustered by Euclidean distance
- Cluster tags were used as an additional feature to the model
- Represents the regional variation within Toronto

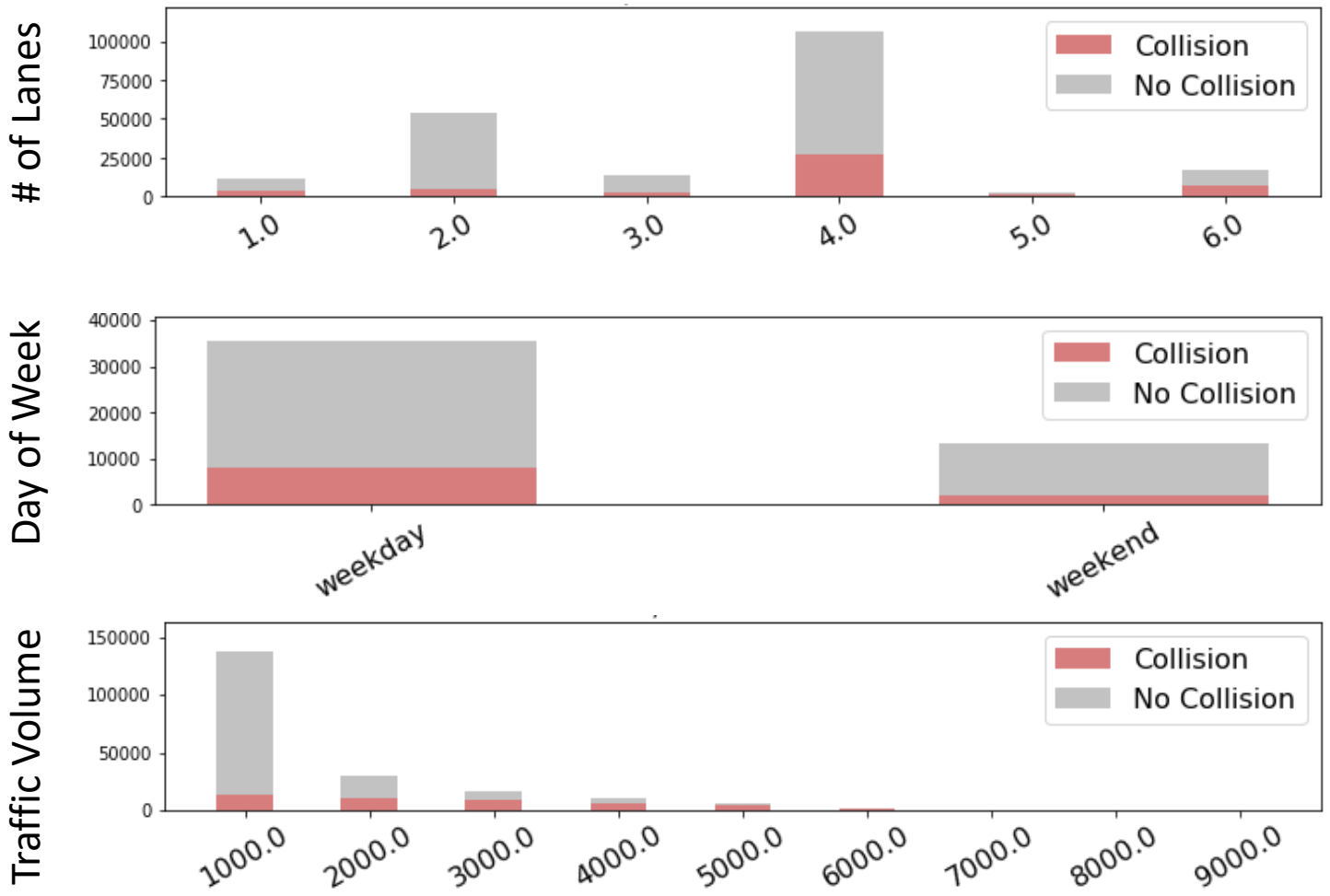


The Waypoints: Key Features

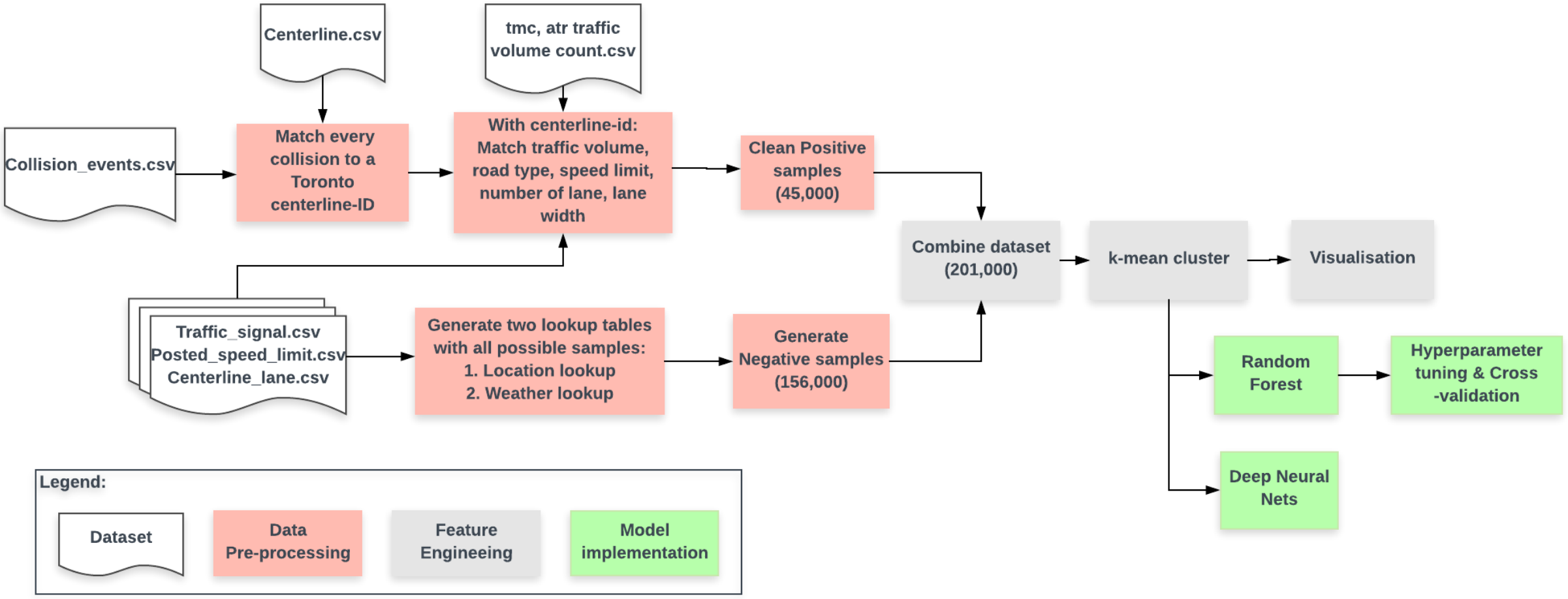
Spatial Features

Temporal Features

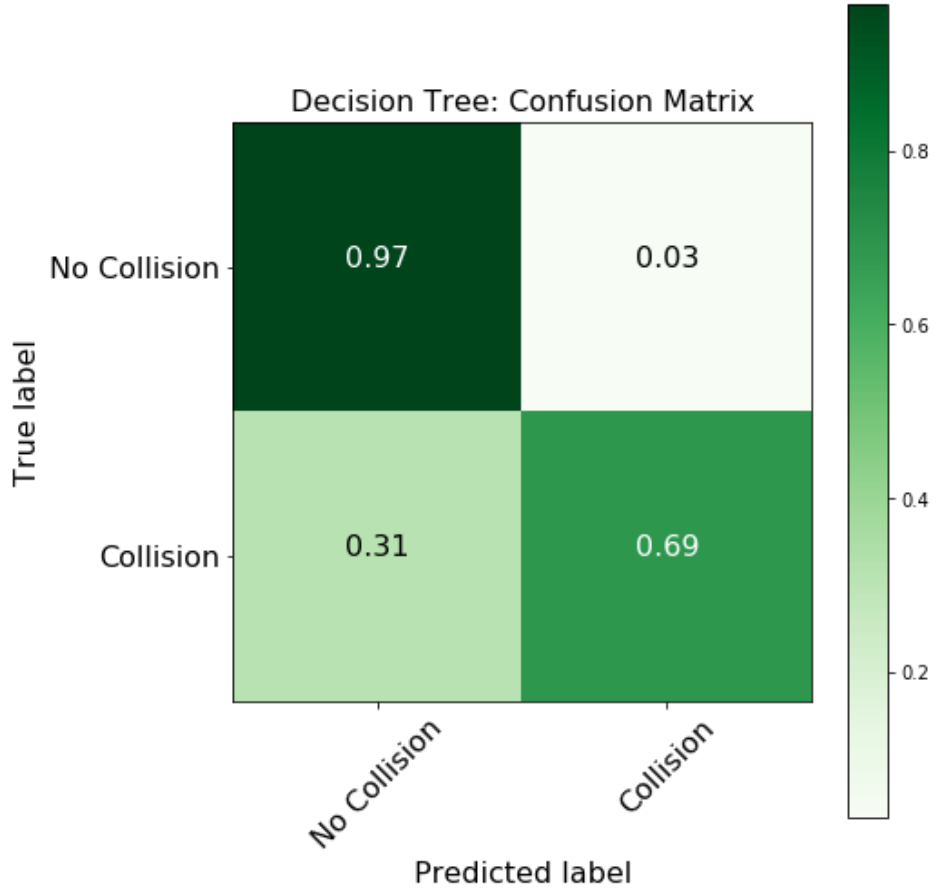
Spatio-temporal Features



Behind the Scenes



A System You Can Trust

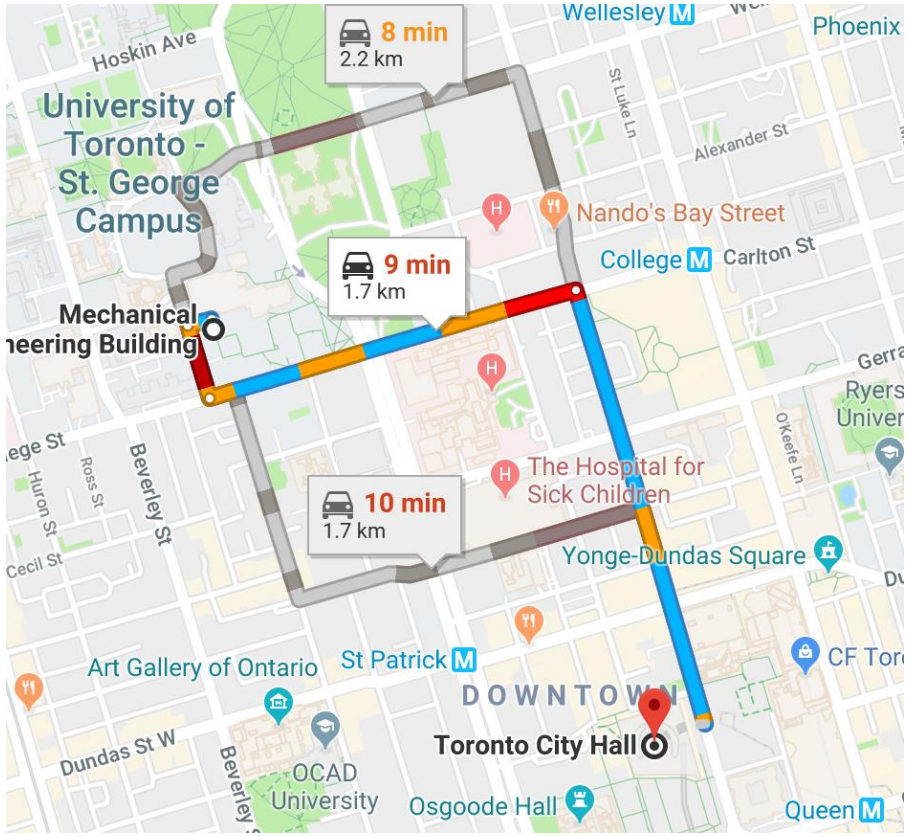


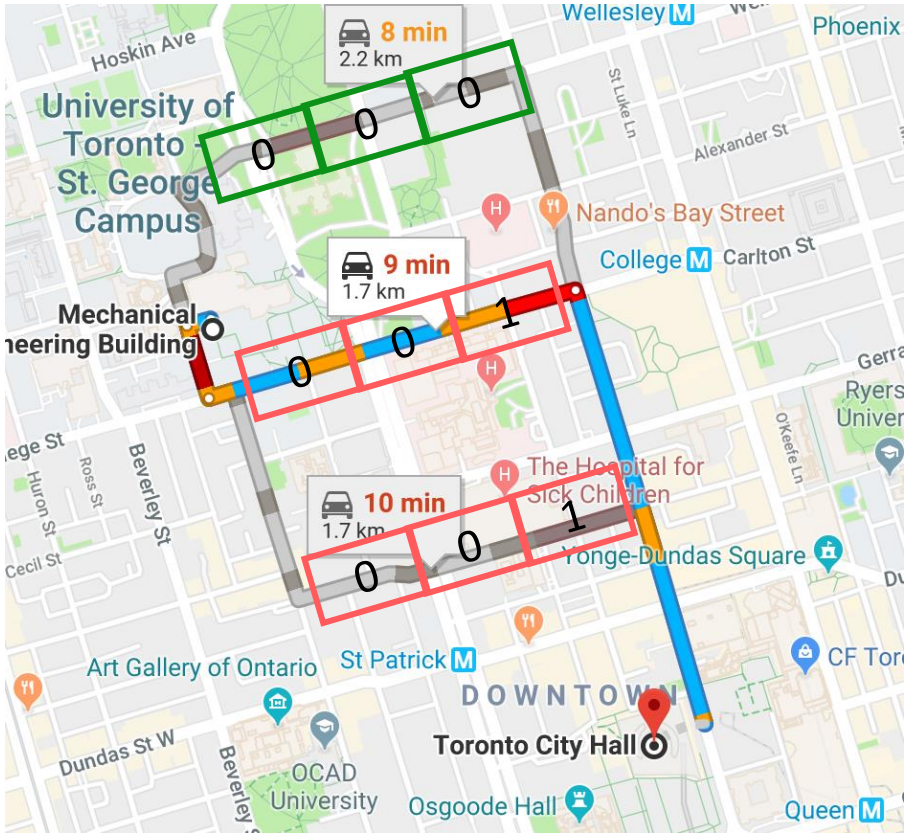
90.6% Cross-Validated Accuracy

97% Recall for No Collisions

67% Recall for Collisions

With DriveAlive , drivers can now find not only the **fastest** but also the **safest** route to their destination with just the touch of a button.





DriveAlive