Identification of Collision-prone Zones Based on Pedestrian Violation Behaviours

Haniyeh Ghomi, Mohamed Hussein

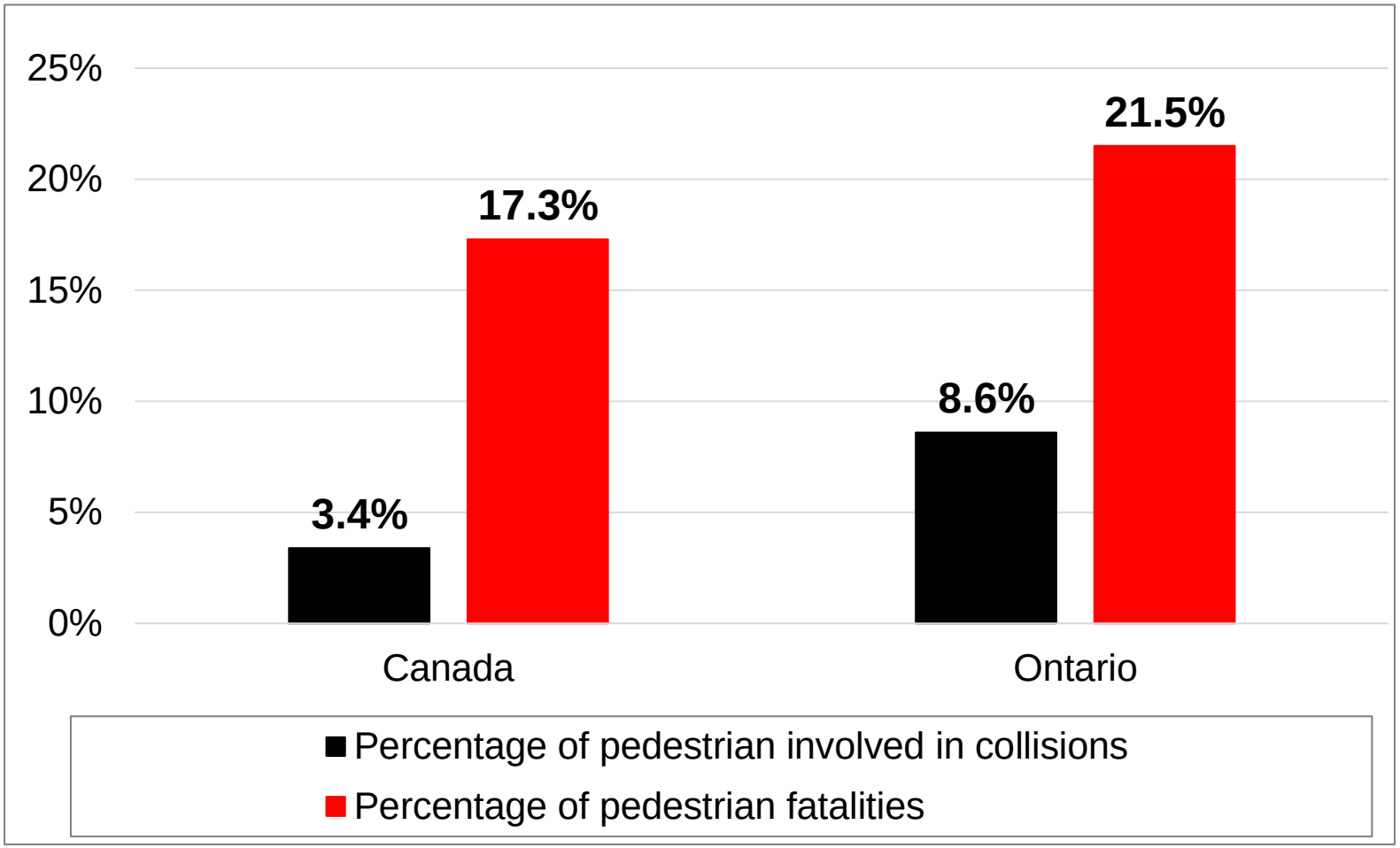
McMaster University

CARSP Conference 2022
Sudbury, Ontario
June 19-21, 2022
Collaborating on the United Nations' (UN) Decade of Action for Road Safety

CARSP ACPSER

McMaster University

Pedestrian Collision Statistics in 2019 in Canada



Pedestrian Unsafe Behaviours

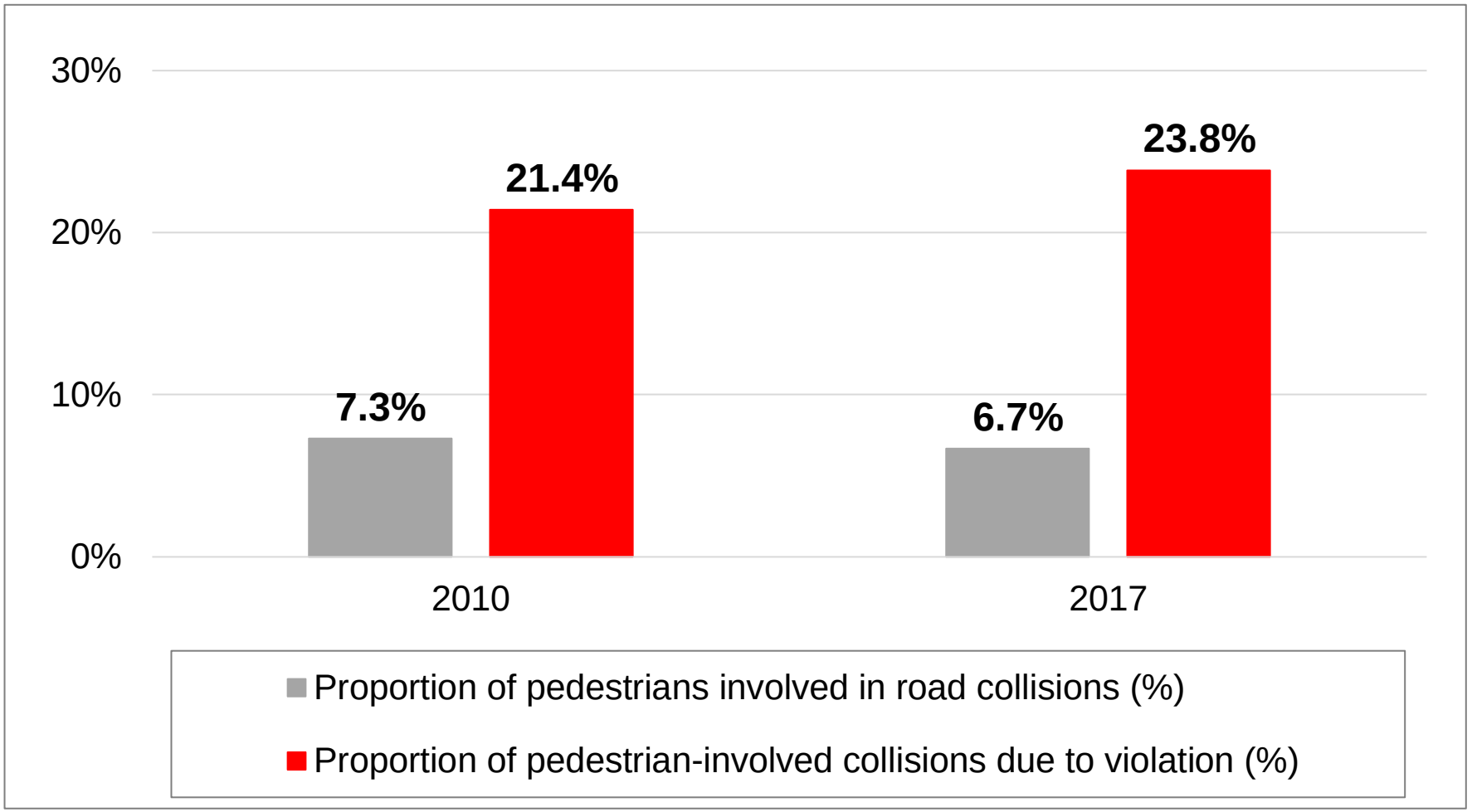
Spatial violation

Crossing a roadway or an intersection at undesignated spaces

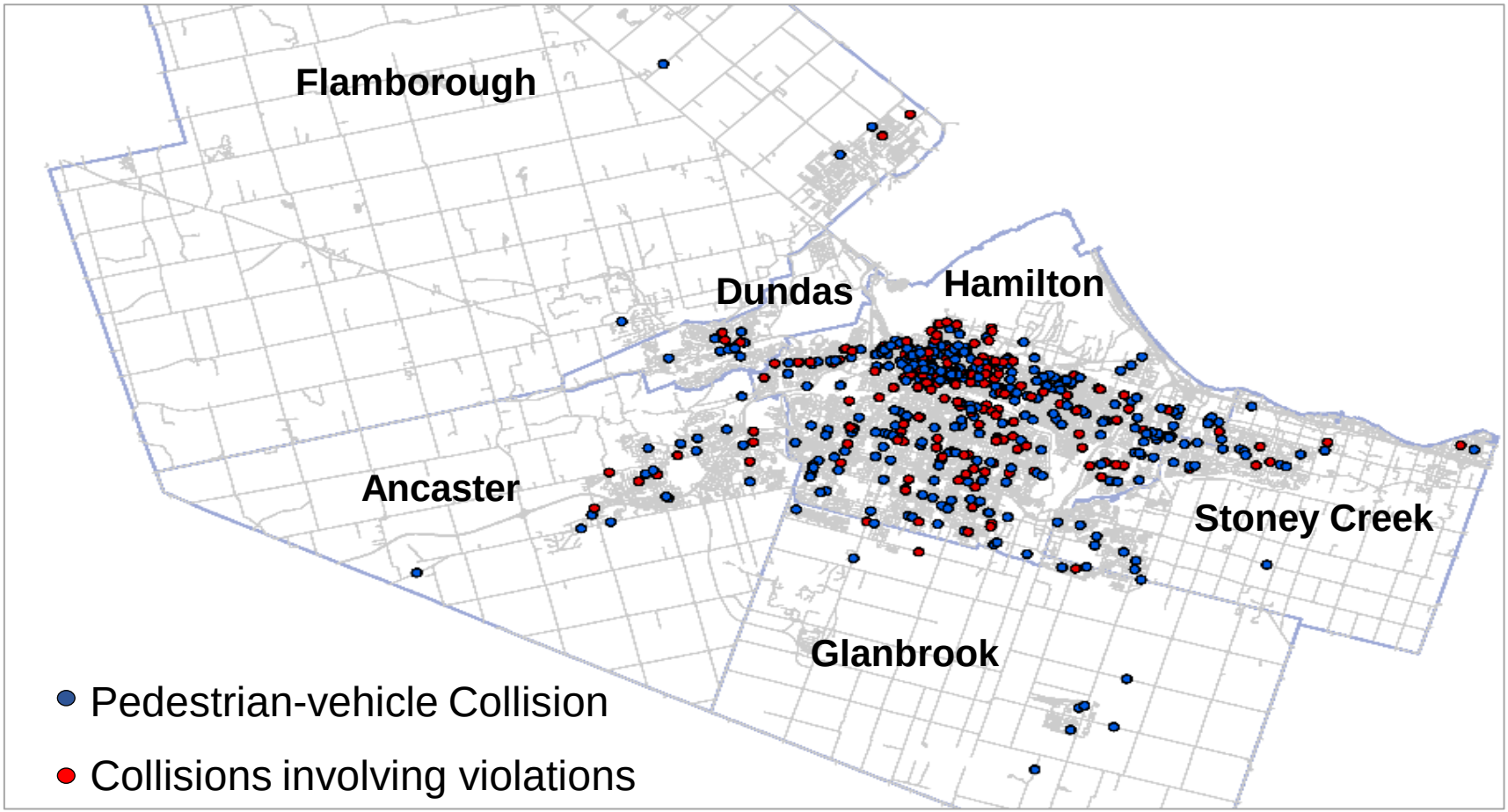
Temporal violation

Crossing a signalized crosswalk during undesignated signal phases

Pedestrian Collision Statistics in Hamilton



Spatial Distribution of the Pedestrian-related Collisions in the City of Hamilton



Objective

Full Bayesian Analysis

- Identify the main contributing factors to both total collisions and collisions that involve pedestrian violations

Hotspots Identification

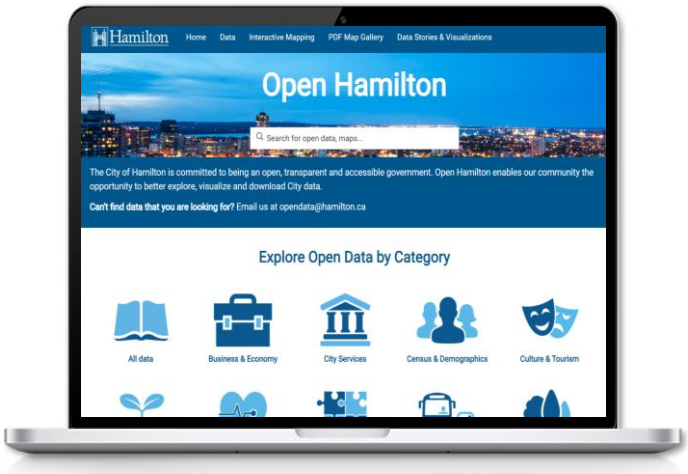
- Identify the collision-prone zones for both types of collisions
- Rank them using the Potential Safety Improvement (PSI) indicator

Support Vector Machine (SVM) model

- Investigate the main differences between hotspots and non-hotspots

Data

8 years collision dataset (2010-2017)



Collision Statistics

- 2089 pedestrian-involved collisions
- 45 fatal collisions (2.15%) and 1859 injury collisions (88.99%)

Pedestrian Violation Statistics

- 509 collisions (24.4% of total collisions)

Potential Contributing Factors

Pedestrian Connectivity

Intersection density
Degree of network coverage
Complexity

Pedestrian Route Directness

Average edge length
Linearity

Land use

Residential Density
Commercial Density
Institutional/Office Density

Exposure

Log (VKT)
Log (PKT)

Built Environment

Signal Density
Bus Stop Density

Socio-economic

Household
Job

Full Bayesian Macro-level Prediction Models

Variables	Total Collisions		Violation-related Collisions	
	Mean	Std. Dev.	Mean	Std. Dev.
Intersection Density	7.861	1.92	21.207	1.44
Degree of network coverage	-0.588	2.13	0.124	1.32
Complexity	3.967	3.52	7.899	2.26
Average edge length	-0.0385	4.40	0.756	1.57
Linearity	0.684	1.93	2.44	8.11
Signal Density	206.8	2.01	-113	1.4
Bus Stop Density	-35.92	1.87	44.05	1.37
Household	0.264	2.10	0.027	1.23
Job	0.113	2.81	0.056	7.22
Residential Density	0.020	2.45	0.02	1.05
Commercial Density	0.019	3.02	0.01	1.16
Institutional/Office Density	0.026	7.83	0.018	1.35
Log(VKT)	7.76	2.09	5.35	7.39
Log (PKT)	1.285	2.63	1.872	1.04

Key findings of Full Bayesian model

Same impact on both categories

- Road user exposure
- Intersection density
- Network complexity
- Sidewalk linearity
- Socio-economic factor
- Land use

Variables	Total Collisions	Violation-related Collisions
Intersection Density	7.861	21.207
Degree of network coverage	-0.588	0.124
Complexity	3.967	7.899
Average edge length	-0.0385	0.756
Linearity	0.684	2.44
Signal Density	206.8	-113
Bus Stop Density	-35.92	44.05
Household	0.264	0.027
Job	0.113	0.056
Residential Density	0.020	0.02
Commercial Density	0.019	0.01
Institutional/Office Density	0.026	0.018
Log(VKT)	7.76	5.35
Log (PKT)	1.285	1.872

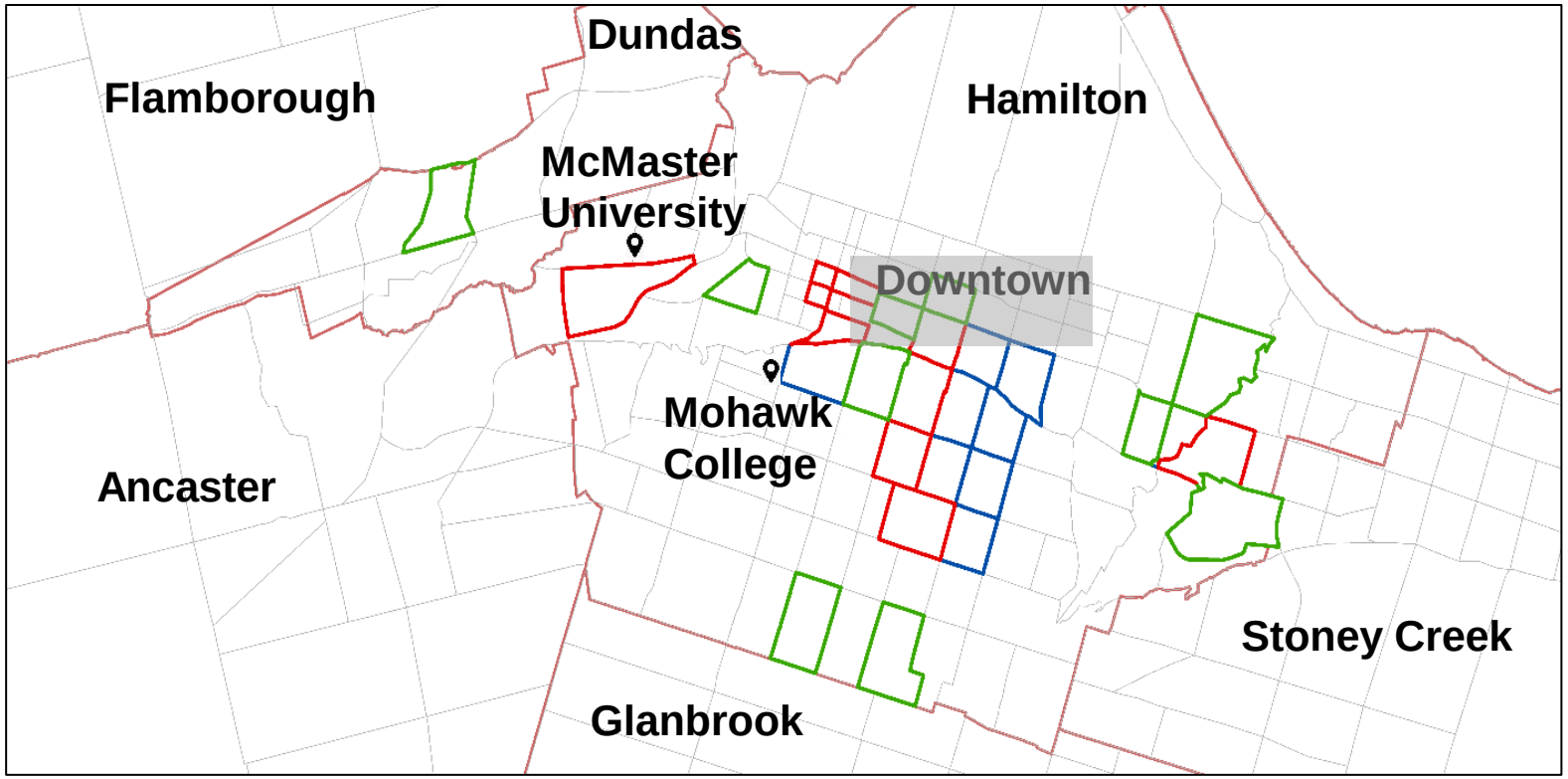
Key findings of Full Bayesian model

Different impact on both categories

- Degree of network coverage
- Average edge length
- Signal density
- Bus stop density

Variables	Total Collisions	Violation-related Collisions
Intersection Density	7.861	21.207
Degree of network coverage	-0.588	0.124
Complexity	3.967	7.899
Average edge length	-0.0385	0.756
Linearity	0.684	2.44
Signal Density	206.8	-113
Bus Stop Density	-35.92	44.05
Household	0.264	0.027
Job	0.113	0.056
Residential Density	0.020	0.02
Commercial Density	0.019	0.01
Institutional/Office Density	0.026	0.018
Log(VKT)	7.76	5.35
Log (PKT)	1.285	1.872

Identification of Collision-prone Zones



- Collision prone zones
- Violation prone zones
- Both collision and violation prone zones

Support Vector Machine Prediction

Criteria	Collision prone zones vs. non-collision prone zones	Hotspots in both scenarios vs. only collision prone zones
CCR	88.13%	87.44%
MAE	0.1186	0.1256
RMSE	0.1444	0.1706
Roc Area	0.869	0.826

Results of SVM Models

Variables	SVM Model 1	SVM Model 2
Intersection Density	0.9642	0.7595
Degree of network coverage	0.1887	0.0756
Complexity	0.5417	0.5085
Average edge length	0.593	0.0798
Linearity	0.4596	0.0693
Signal Density	0.8432	0.6264
Bus Stop Density	0.4783	0.5252
Household	0.6764	0.3212
Job	0.0578	0.3452
Residential Density	0.1972	0.2914
Commercial Density	0.3649	0.3856
Institutional/Office Density	0.4322	0.4488
Log(VKT)	0.6722	0.2370
Log (PKT)	0.5885	0.1131
Intercept	1.7308	1.3331

Key findings of SVM Model

- Contributing factors in distinguishing hotspots and non-hotspots zones:

Total pedestrian-vehicle collisions

Intersection density
Signal density
Household density
VKT
PKT

Violation involved collisions

Intersection density
Pedestrian network complexity
Signal density
Bus stop density
Institutional/office land use density

Conclusion

- Intersection density is the most influential factor in the frequency of violation-related collisions
- Pedestrian network features and land use areas are significant on hotspots identifications
- Locations with poor pedestrian network connectivity require countermeasures that mitigate pedestrian violations

Future Directions

- Explore more precise measures for pedestrian exposure, including collecting extra survey data or implementing activity-based algorithms to estimate the volume
- Conduct some other advanced techniques to identify the collision-prone zones, such as Deep Learning models
- Investigate the impact of other contributing factors on pedestrian violations including income, car ownership, and household characteristics

Thank You

ghomirah@mcmaster.ca