ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EVALUATING TRAVEL MODE DECISIONS AND TRANSPORT MODELS IN UNDERSTANDING TRANSIT EQUITY: THE CASE OF GREATER TORONTO AND HAMILTON AREA

Ph.D. THESIS

Elnaz YOUSEFZADEH BARRI

Department of Urban and Regional Planning

Urban and Regional Programme

AUGUST 2022

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EVALUATING TRAVEL MODE DECISIONS AND TRANSPORT MODELS IN UNDERSTANDING TRANSIT EQUITY: THE CASE OF GREATER TORONTO AND HAMILTON AREA

Ph.D. THESIS

Elnaz YOUSEFZADEH BARRI
(502162804)

Department of Urban and Regional Planning

Urban and Regional Programme

Thesis Advisor: Assoc Prof. Eda BEYAZIT İNCE
Thesis Co-Advisor: Assoc Prof. Steven FARBER

AUGUST 2022

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

TOPLU TAŞIMADA EŞİTLİĞİ ANLAMAYA YÖNELİK OLARAK YOLCULUK TÜRÜ KARARLARININ VE ULAŞIM MODELLERININ DEĞERLENDİRİLMESİ: BÜYÜK TORONTO ALANI VE HAMİLTON BÖLGESİ VAKA ÇALIŞMASI

DOKTORA TEZİ

Elnaz YOUSEFZADEH BARRI
(502162804)

Şehir ve Bölge Planlaması Anabilim Dalı

Şehir ve Bölge Planlama Programı

Tez Danışmanı: Assoc Prof. Eda BEYAZIT İNCE
(Varsa) Eş Danışman: Assoc Prof. Steven FARBER

AĞUSTOS 2022

Elnaz YOUSEFZADEH BARRI, a Ph.D. student of ITU Graduate School student ID 502162804, successfully defended the dissertation entitled “EVALUATING TRAVEL MODE DECISIONS AND TRANSPORT MODELS IN UNDERSTANDING TRANSIT EQUITY: THE CASE OF GREATER TORONTO AND HAMILTON AREA”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc Prof. Eda BEYAZIT İNCE** ....
Istanbul Technical University

Co-advisor : **Assoc Prof. Steven FARBER** ....
University of Toronto

Jury Members : **Prof. Dr. Handan TÜRKOĞLU** ....
Istanbul Technical University

Assoc Prof. Kevser İSMET ÜSTÜNDAĞ ....
Mimar Sinan Fine Art University

Assoc Prof. Hüseyin ONUR TEZCAN ....
Istanbul Technical University

Prof. Dr. Turgay Kerem KORAMAZ ....
Istanbul Technical University

Assist Prof. Nicholas J. KLEIN ....
Cornell University

Date of Submission : 1 August 2022
Date of Defense : 16 August 2022

To my family,

FOREWORD

The accomplishment of this dissertation has made tremendous contributions to my professional and personal life. This work has broadened my academic experience in international scientific communities, enhanced my research and analysis skills, and taught me to have enough perseverance and adhere to the goal to the end of the journey. Accordingly, I would like to thank all those who support me to make this study possible and memorable.

Firstly, I want to give the biggest thanks to my advisors, Professor Eda Beyazit and Professor Steven Farber, for their endless help and contribution to this work. They have a valuable contribution to my education, research, and professional progress. Their mentorship and encouragement throughout these years propelled me further forward. Moreover, the support and guidance of my committee members, Professors Handan Türkoğlu, Kevser Üstündağ, and Hüseyin Onur Tezcan, helped me a lot understand how my research could best advance.

This study would not have been possible without the support of the Istanbul Technical University and the University of Toronto. I would like to acknowledge the University of Toronto's Transportation Research Institute and helpful staff for providing the necessary data.

For me, the last two years were a great and enjoyable step in my academic and professional life. I would like to thank Professor Farber, who trusted my abilities and let me join his wonderful research lab at the Department of Geography & Planning, University of Toronto. I want to thank my dear friends Jeff Allen and Ignacio Tiznado-Aitken, who were always one step ahead of me, for their support and for sharing their experiences. I would also like to thank all SAUSy lab members who made my last year better than ever.

Many thanks also go to my parents. Without their patience throughout my life, I would have not pursued a Ph.D. Last but not the least, I want to thank my spouse, Hadi Jahanshahi, for his unconditional support emotionally and professionally. Without his boundless help and motivation, this success was impossible. He is the true hero behind this accomplishment, and I owe him a great deal.

AUGUST 2022

Elnaz YOUSEFZADEH BARRI
(Urban & Regional Planner)

TABLE OF CONTENTS

Page

FOREWORD ... **ix**
TABLE OF CONTENTS ... **xi**
ABBREVIATIONS ... **xv**
SYMBOLS ... **xvii**
LIST OF TABLES ... **xix**
LIST OF FIGURES ... **xxi**
SUMMARY ... **xxiii**
ÖZET ... **xxvii**
1. INTRODUCTION ... **1**
1.1 Chapter Overview ... 1
1.2 Background ... 1
1.3 Research Objectives ... 3
1.4 Research Questions ... 5
1.5 Thesis Structure ... 6
2. LITERATURE REVIEW ... **7**
2.1 Chapter Overview ... 7
2.2 Transportation Planning and Equity ... 7
2.3 Travel Behaviour Analysis ... 8
2.4 Transit Investments and Mode Shift ... 10
2.5 Travel Behaviour Analysis Models ... 13
2.5.1 Trip chain analysis ... 13
2.5.2 Travel mode use analysis ... 15
2.5.2.1 Classification models ... 16
2.5.2.2 Count models ... 17
2.5.2.3 Machine learning applications in transit equity studies ... 19
2.5.2.4 Interpretable machine learning models ... 20
2.6 Summary and Research Gaps ... 21
3. STUDY AREA AND DATA ... **25**
3.1 Chapter Overview ... 25
3.2 Study Context ... 25
3.2.1 The population density ... 26
3.2.2 Transit ridership ... 27
3.2.3 Housing price ... 29
3.3 Transportation Tomorrow Survey (TTS) ... 29
3.4 Descriptive Summary ... 31
3.4.1 General overview ... 31
3.4.2 Transit accessibility measurement ... 32

3.4.3 Low-income vs. high-income ... 35
3.4.3.1 Socioeconomic and built environment variables ... 35
3.4.3.2 Trip information ... 36
3.4.4 Transit use, income, car ownership, and accessibility levels ... 38
4. METHODS ... 41
4.1 Chapter Overview ... 41
4.2 Trip Chain Analysis ... 41
4.2.1 Trip sequence generation ... 41
4.2.2 Hierarchical clustering method ... 45
4.3 Statistical and Machine Learning Methods ... 46
4.3.1 Statistical models ... 46
4.3.2 Supervised learning models ... 47
4.3.3 Cross-validation ... 49
4.4 Performance Metrics ... 49
4.4.1 Classifier's performance metrics ... 50
4.4.2 Regressor's performance metrics ... 51
4.5 Non-parametric Statistical Tests for Comparing Multiple Groups ... 52
4.6 Spatial Efficiency Measure (SPAEF) ... 53
4.7 Model Interpretability Tools ... 53
4.7.1 Global interpretability ... 54
4.7.2 Local interpretability ... 55
4.8 Conclusion ... 56
5. TRAVEL BEHAVIOUR CLUSTERING ... 59
5.1 Chapter Overview ... 59
5.2 Cluster Analysis of Low-income Travellers ... 60
5.2.1 Activity patterns of low-income clusters ... 60
5.2.2 Socioeconomic characteristics of low-income clusters ... 63
5.2.3 Built environment characteristics of low-income clusters ... 64
5.3 Comparing Low-income and High-income Household's Travel Behaviour ... 65
5.3.1 Carless clusters ... 65
5.3.2 Car-owning clusters ... 68
5.4 Conclusion ... 70
6. IMPACTS OF TRANSIT INVESTMENTS ON SHIFTING MODE ... 73
6.1 Chapter Overview ... 73
6.2 ZINB Model Results ... 74
6.3 Sensitivity Analysis ... 77
6.4 Conclusion ... 83
7. COMPARING STATISTICAL AND MACHINE LEARNING MODELS ... 85
7.1 Chapter Overview ... 85
7.2 Comparing Models on Predicting the Probability of Taking Transit ... 86
7.3 Comparing Models on Estimating the Number of Transit Trips ... 88
7.4 Sensitivity Analysis ... 90
7.5 Model Interpretation ... 93
7.5.1 Feature importance ... 94
7.5.2 Partial dependency plot (PDP) ... 95
7.5.3 Individual conditional expectation (ICE) ... 95
7.5.4 Shapley value (SHAP) ... 96
7.5.5 Local interpretable model-agnostic explanations (LIME) ... 97

7.6 Conclusion 98
8. CONCLUSIONS 101
8.1 Chapter Overview 101
8.2 Thesis Summary 101
8.3 Thesis Contributions 106
8.4 Policy Recommendations 108
8.4.1 Policies for equity outcomes 108
8.4.2 Policies for sustainable outcomes 109
8.4.3 Policies for model selection 111
8.5 Study Limitations and Future Work 112
8.5.1 Reproducibility 114
REFERENCES 115
APPENDICES 129
APPENDIX A1: Feature importance algorithm 130
APPENDIX A2: Partial Dependence Plot (PDP) algorithm..130
APPENDIX A3: Individual Conditional Expectation (ICE) algorithm..131
APPENDIX A4: Local interpretable model-agnostic explanations (LIME) algorithm 131
APPENDIX B1: Detailed comparison of classification models' performances ...132
APPENDIX B2: Detailed comparison of regression models' performances 133
CURRICULUM VITAE 135

ABBREVIATIONS

AUC-ROC	:	Area Under the Curve of Receiver Operating Characteristic
CV	:	Cross-Validation
DT	:	Decision Tree
ICE	:	Individual Conditional Expectation
IRR	:	Incidence Rate Ratio
LIME	:	Local Interpretable Model-agnostic Explanation
LinR	:	Linear Regression
LogR	:	Logistic Regression
MedAE	:	Median Absolute Error
ML	:	Machine Learning
MNL	:	Multinomial Logit
NB	:	Naive Bayes
NL	:	Nested Logit
NN	:	Neural Networks
OR	:	Odds Ratio
PDP	:	Partial Dependency Plot
R^2	:	R-squared
RF	:	Random Forest
RMSE	:	Root Mean Squared Error
RRSE	:	Root Relative Squared Error
RSS	:	Residual Sum of Squares
SHAP	:	Shapley value
SVM	:	Support Vector Machine
TSS	:	Total Sum of Squares
XGB	:	Extreme Gradient Boosting
ZINB	:	Zero-Inflated Negative Binomial Regression

SYMBOLS

A_i	**:** Accessibility measure in zone i
O_j	**:** Number of jobs in zone j
t_{ij}	**:** Travel time between zone i and zone j
i, j	**:** Zones
y_i	**:** The dependent variable of individual i
$\bar{y}$	**:** Mean of the dependent variable
$\hat{y}_i$	**:** The estimated dependent variable of individual i
p_i	**:** The probability of taking transit for the individual i
n_i	**:** Number of the expected transit trips for the individual i
x_i	**:** The transit accessibility value of the individual i
E_{x_i}	**:** The elasticity of transit accessibility for the individual i
β_i	**:** The coefficient of transit accessibility for individual i
w_i	**:** The expansion value of the individual i
$\bar{E}_x$	**:** The weighted elasticity of transit accessibility variable
y_i^c	**:** The current number of transit trips for the individual i in class c
$\hat{y}_i^c$	**:** The estimated number of transit trips for the individual i in class c
$\Delta\hat{y}^c$	**:** The predicted change in the number of transit trips in class c
n^c	**:** The total number of individuals in class c
I	**:** Individual
$\mathcal{M}$	**:** Travel modes
$\mathcal{S}$	**:** Sequence of trip legs
$\mathcal{T}$	**:** Trip chain
a,b	**:** Numeric vectors
s_{ij}	**:** Average cosine similarity between individuals i and j
c, c'	**:** Classes of the classification task
T_c	**:** True prediction of the class c
F_c	**:** False prediction of the class c

LIST OF TABLES

Table 2.1	Contribution of the current study compared to the literature. ...	**22**
Table 3.1	Descriptive statistics of explanatory variables for respondents ($n = 249,632$; expandable to $\mathcal{N} = 5,387,081$).	**33**
Table 3.2	The descriptive summary of explanatory variables for carless and car-owner households with income less than $\$40k$ and greater than $\$125k$ ($n = 65,458; \mathcal{N} = 1,419,640$).	**36**
Table 3.3	The expanded frequency of trip purposes for low- and high-income households in their daily trips.§	**37**
Table 3.4	The expanded frequency of the travel modes for low- and high-income households in their daily trips.§	**37**
Table 4.1	Frequency table of trip legs for each individual ($\mathcal{S}$).............	**43**
Table 5.1	Cluster description of low-income households and the probability of belonging to the given cluster.....................	**61**
Table 5.2	Cluster description of high-income households and the probability of belonging to the given cluster.....................	**66**
Table 6.1	ZINB model results (N= 3,279,979; n = 149,177).	**76**
Table 6.2	The expanded number of new transit trips generated for each class after transit accessibility improvement.....................	**82**
Table 7.1	Performance comparison of each classifier for predicting the probability of taking transit in individuals' daily trips based on 10-fold cross-validation..	**88**
Table 7.2	Performance comparison of each regressor for predicting the number of transit trips in individuals' daily trips based on 10-fold cross-validation..	**89**
Table B.1	Comparing the performance of the best classifiers and the baseline classifier using the Friedman Aligned Ranks test and its *post hoc* analysis.	**132**
Table B.2	Comparing the performance of the best regressors and the baseline regressors using the Friedman Aligned Ranks test and its *post hoc* analysis.	**133**

LIST OF FIGURES

Figure 3.1	Study Area (The Greater Toronto and Hamilton Area).	**26**
Figure 3.2	The population density of different household income and car-ownership classes (missing data are excluded). The number of people in each class is normalized by the total population of the same class and the total number of people in each census tract to eliminate the size effect. The color indicates the normalized persons per square kilometer. VA = Vehicles per adult in the household.	**28**
Figure 3.3	Percentage of individuals taking at least one transit trip in their daily trip chain in the GTHA.	**29**
Figure 3.4	Average house price for all types of dwellings in the GTHA.	**30**
Figure 3.5	Average of gravity-based accessibility to jobs by transit in the GTHA.	**34**
Figure 3.6	The percentage of individuals using transit in each class (people with missing income data and no trips are excluded).	**38**
Figure 3.7	Distribution of accessibility, income and car ownership for all households.	**39**
Figure 4.1	Sample trip chains with their prime mode.	**42**
Figure 4.2	An example of a trigram with a gap of 2 to generate a set of the complete trip chain.	**42**
Figure 4.3	Motivating example.	**44**
Figure 5.1	The travel pattern of four clusters for low-income carless and car-owners (The transparency shows the frequency of each trip segment, destination/travel mode: the lighter it is, the less frequent it will be). (a) Low-income, carless individuals. (b) Low-income, car-owner individuals.	**62**
Figure 5.2	The travel pattern of four clusters for high-income carless and car-owners (The transparency shows the frequency of each trip segment, destination/travel mode: the lighter it is, the less frequent it will be). (a)High-income, carless individuals. (b)High-income, car-owner individuals.	**67**
Figure 6.1	Elasticity estimates of transit accessibility for 25 LogR models (income and car ownership levels).§*Elasticity estimates were insignificant at the 0.05 level.*	**79**
Figure 6.2	The expanded changes in transit trips per person by accessibility improvements. (The expanded population of each stratum is shown on the graphs).	**80**

Figure 6.3 The ratio of newly generated transit trips given existing and new users after transit improvement (200k). **81**

Figure 7.1 The experimental design of the study. **86**

Figure 7.2 Friedman test result for classification "accuracy" after Bergmann-Hommel *post hoc* procedure. (a) Corrected pairwise p-values using Bergmann-Hommel *post hoc* procedure. (b) Average rank of classifiers ($\alpha = 0.05$). Edges between algorithms indicate an insignificant difference. **89**

Figure 7.3 Friedman test result of regression "RMSE" values after Bergmann-Hommel *post hoc* procedure. (a)Corrected pairwise p-values using Bergmann-Hommel *post hoc* procedure. (b)Average rank of regressors ($\alpha = 0.05$). Edges between algorithms indicate an insignificant difference. **90**

Figure 7.4 The sensitivity of all models to the accessibility improvement for low-income carless households (The baseline model is shown in red, the statistical models in black, and the ML models in blue. **91**

Figure 7.5 The spatial prediction of newly generated transit trips by low-income carless group after improving job accessibility (200k jobs) using different algorithms. **93**

Figure 7.6 The heatmap of SPAEF scores. **94**

Figure 7.7 The importance of each feature for predicting the number of transit trips using RF model. **95**

Figure 7.8 ICE plot and PDP for predicting the number of transit trips after increasing transit accessibility by 200k jobs (Blue lines indicate ICE plots). **96**

Figure 7.9 SHAP values' distribution and mean. Features are sorted by their mean SHAP values. **97**

Figure 7.10 Two sampled individuals, one not using transit and one a frequent transit user. (a) A non-transit user. (b) A transit user. **98**

EVALUATING TRAVEL MODE DECISIONS AND TRANSPORT MODELS IN UNDERSTANDING TRANSIT EQUITY: THE CASE OF GREATER TORONTO AND HAMILTON AREA

SUMMARY

In recent decades, the incorporation of equity considerations in the transportation domain and the equity analysis of transport projects and policies are rapidly increasing. These approaches mainly include travel behaviour analysis with equity indicators and the socioeconomic impacts of transport investments on individuals. Accordingly, the cost and benefits of transport investments for residents are evaluated. Moreover, travellers' travel behaviour, daily activity patterns, and travel mode decisions are estimated through their trip chains analysis. These assessments can offer a broad perspective on individuals' travel needs and constraints. They also offer valuable insight for transportation planners and policymakers in understanding how different transport investments impact society. Therefore, they enable authorities and planners to develop equitable transport policies and travel demand management to address various environmental problems.

This dissertation focuses on understanding how different socioeconomic groups plan their daily trips and reports important findings on their responses to transport investments, aiming to improve individuals' activity participation and alleviate travel barriers. The study also evaluates travel behaviour and mode use models and investigates the potential of machine learning algorithms for travel behaviour prediction in the Greater Toronto and Hamilton Area (GTHA), one of the largest and fastest-growing regions in Canada.

The primary data source used for this study is the 2016 Transportation Tomorrow Survey (TTS) dataset, a large-sample household travel survey including a one-day household travel diary conducted in the Greater Golden Horseshoe Area. The TTS data is a part of an ongoing data collection program started in 1986 and is collected every five years. This regional survey is conducted to travel demand management, and it can use for transportation planning programs and models.

In the first step, this study explores how income and car-ownership levels determine activity patterns and travel decisions of travellers using an aggregated form of activity type and travel mode as a unit of trip chain analysis. A presumption-free clustering framework is leveraged to mitigate the subjectivity of rule-based approaches for trip chain analysis. This approach extracts the homogeneous clusters of activity patterns. Second, the impacts of transit improvements in low-income communities are explored based on the assumption that transit investments could result in changing travel mode use and generating more transit and fewer car trips. Such analysis is performed by exploring the association between transit use and transit accessibility improvements

using stratified regression models. Lastly, the effects of travel behaviour models are evaluated in terms of their predictive performance in policy-making and transportation planning. This study investigates how the model selection affects the prediction of transit use and compares the predictive performance of traditional and Machine Learning (ML) algorithms. Then, it evaluates a transit investment policy by contrasting the predicted activities and the spatial distribution of transit trips generated by the vulnerable households after improving accessibility.

The findings of this study reveal that income and car-ownership levels influence a traveller's travel decisions and change their mobility patterns. The findings show that females, regardless of income or car ownership, frequently take transit in their daily trip chains. Among low-income carless individuals, most of their daily trips include the mobility of care, where women more often than men play this traditional role in a household by either public transit or a car as a passenger. In the low-income car-owner subsample, females still use public transit for their work trips, whereas males regularly use the household's car to commute to work. It confirms that women benefit less from having access to a car in families with a shared private vehicle. Males of wealthy carless households integrate public transit and active transportation for their daily trips when they live in high-density and more accessible neighbourhoods.

Furthermore, evaluating transit improvements in low-income communities shows that low-income households with one or more cars per adult have the most elastic relationship between transit accessibility and transit use; they are more likely to be transit riders if transit improves. However, in auto-centric areas with poor transit, the transit use of low-income households drops off sharply as car ownership increases. It implies that low-income car-owning households might become too reliant on their vehicles as soon as they own them. Moreover, the sensitivity analysis exploring how changes to accessibility affect transit trip generation highlights that the accessibility gains in the region provide more opportunity for increasing transit ridership among car-deficit households when transit is improved. Therefore, the analysis suggests some insight into engaging individuals in taking transit and resulting in overall transit ridership in the region.

Given the model selection, the results show that ML algorithms outperform all other statistical models and have great potential for enhancing travel behaviour predictions without sacrificing interpretability. Random Forest (RF), XGBoost (XGB), and Neural Networks (NN) classifiers and regressors significantly outperform other algorithms. Among them, RF is the most accurate approach for predicting low-income families' transit demand according to its predictive performance. However, statistical models perform poorly when forecasting transit users' behaviours. Further, the spatial distribution of newly generated transit trips after transit improvements is not identical; thus, traditional models may arrive at a different, probably inaccurate, policy recommendation in addressing social, spatial, and environmental problems. Moreover, applying model-agnostic interpretation tools to ML models shows that these techniques can uncover each model's underlying process, which was supposed to be a "black box". All in all, ML models demonstrate significant improvement in accuracy and interpretability.

The findings point out that understanding and estimating individuals' travel decisions and preferences with a reliable model enables policymakers to establish an appropriate transit framework that benefits low-income people and alleviates transit inequality in society. This study suggests that evaluating individuals' travel behaviour in terms of their income and car-ownership levels may give a new and different outlook on transport planning in metropolitan cities. Overall, a fair transportation investment that meets environmental, economic, and social goals necessitates a thorough understanding of different socioeconomic groups' travel requirements and responses. The findings help planners rethink transport policies and strategies that increase activity participation and reduce environmental impacts.

TOPLU TAŞIMADA EŞİTLİĞİ ANLAMAYA YÖNELİK OLARAK YOLCULUK TÜRÜ KARARLARININ VE ULAŞIM MODELLERININ DEĞERLENDİRİLMESİ: BÜYÜK TORONTO ALANI VE HAMİLTON BÖLGESİ VAKA ÇALIŞMASI

ÖZET

Son yıllarda, ulaşım alanına eşitlik konularının dahil edildiği ve ulaşım projelerinin ve politikalarının değerlendirilmesinde eşitlik temelli analizlerin yaygın bir şekilde yer aldığı görülmektedir. Bu yaklaşımlar temel olarak eşitlik göstergeleri ile yolculuk davranışı analizini ve ulaşım yatırımlarının bireyler üzerindeki sosyoekonomik etkilerini içermektedir. Buna göre ulaşım yatırımlarının maliyet ve faydaları değerlendirilmektedir. Ayrıca, yolcuların davranışları, günlük aktivite kalıpları ve seyahat türü kararları, yolculuk zincirleri analizi yoluyla tahmin edilmektedir. Bu değerlendirmeler, bireylerin yolculuk ihtiyaçları ve kısıtlamaları hakkında geniş bir perspektif sunabilir. Ayrıca, farklı ulaşım yatırımlarının toplumu nasıl etkilediğini anlama konusunda ulaşım planlayıcıları ve politika yapıcılar için değerli bilgiler sunarlar. Bu nedenle, yetkililerin ve plancıların çeşitli çevresel sorunları ele almak için adil ulaşım politikaları ve seyahat talebi yönetimi geliştirmelerini sağlar.

Bu tez, farklı sosyoekonomik grupların günlük gezilerini nasıl planladıklarını anlamaya odaklanmakta ve bireylerin aktivite katılımını artırmayı ve seyahat engellerini hafifletmeyi amaçlayan ulaşım yatırımlarına nasıl yanıt verdiklerine ilişkin önemli bulguları rapor etmektedir. Çalışma ayrıca seyahat davranışı ve ulaşım türü kullanım modellerini değerlendirmekte ve Kanada'nın en büyük ve en hızlı büyüyen bölgelerinden biri olan Greater Toronto ve Hamilton Bölgesi'nde (GTHA) seyahat davranışı tahmini için makine öğrenmesi (machine learning) algoritmalarının potansiyelini araştırmaktadır.

Bu çalışma için kullanılan birincil veri kaynağı, Greater Golden Horseshoe Bölgesi'nde yürütülen bir günlük yolculuk günlüğünü içeren büyük örneklemli bir hane halkı seyahat anketi olan 2016 Ulaştırma Yarını Anketi (Transportation Tomorrow Survey - TTS) veri setidir. TTS verileri, 1986'da başlatılan ve her beş yılda bir toplanan devam eden bir veri toplama programının bir parçasıdır. Bu bölgesel anket, seyahat talep yönetimi için yapılmıştır ve ulaşım planlama programları ve modelleri için kullanılabilir.

İlk adımda, bu çalışma, bir yolculuk zinciri analizi birimi olarak toplu bir etkinlik türü ve seyahat türü biçimini kullanarak, gelir ve araç sahipliği düzeylerinin yolcuların etkinlik modellerini ve seyahat kararlarını nasıl belirlediğini araştırmaktadır. Yolculuk zinciri analizi için kural tabanlı yaklaşımların öznelliğini azaltmak için varsayımsız bir kümeleme çerçevesi kullanılır. Bu yaklaşım, homojen aktivite kalıpları kümelerini çıkarır. İkincisi, düşük gelirli topluluklardaki toplu taşıma iyileştirmelerinin etkileri, toplu taşıma yatırımlarının yolculuk türü kullanımını değiştirebileceği ve

daha fazla toplu taşıma ve daha az otomobil yolculuğu üretebileceği varsayımına dayalı olarak araştırılmaktadır. Bu tür bir analiz, katmanlı regresyon modelleri kullanılarak toplu taşıma kullanımı ile toplu taşıma erişilebilirliği iyileştirmeleri arasındaki ilişkiyi araştırmaktadır. Son olarak, seyahat davranışı modellerinin etkileri, politika oluşturma ve ulaşım planlamasındaki öngörülü performansları açısından değerlendirilmektedir. Bu çalışma, model seçiminin geçiş kullanımının tahminini nasıl etkilediğini araştırmakta ve geleneksel ve Makine Öğrenmesi (ML) algoritmalarının tahmin performansını karşılaştırmaktadır. Ardından, erişilebilirliği iyileştirdikten sonra hassas haneler tarafından oluşturulan toplu taşıma yolculukların mekansal dağılımını ve öngörülen faaliyetleri karşılaştırarak toplu taşıma yatırım politikasını değerlendirmektedir.

Bu çalışmanın bulguları, gelir ve araç sahipliği düzeylerinin yolculuk kararlarını etkilediğini ve hareketlilik modellerini değiştirdiğini ortaya koymaktadır. Bulgular, kadınların gelir durumlarından veya araba sahibi olmalarından bağımsız olarak, günlük yolculuk zincirlerinde sıklıkla topplu taşıma kullandıklarını göstermektedir. Düşük gelirli arabasız bireyler arasında, günlük yolculukların çoğu, kadınların erkeklerden daha sık olarak toplu taşıma veya araba ile yolcu olarak gerçekleştirilen bakım hareketliliğini içerir. Düşük gelirli otomobil sahibi gruplar alt örneğinde, kadınlar iş yolculukları için hala toplu taşıma araçlarını kullanırken, erkekler işe gidip gelmek için düzenli olarak haneye ait olan otomobili kullanmaktadır. Kadınların ortak özel aracı olan ailelerde otomobile erişimden daha az yararlandığını doğrulamaktadır. Otomobilsiz yüksek gelirli hanelerdeki erkek bireyler, yüksek yoğunluklu ve daha erişilebilir mahallelerde yaşadıklarında, günlük yolculukları için toplu taşıma ve aktif ulaşımı entegre ederek kullanmaktadır.

Ayrıca, düşük gelirli topluluklardaki toplu taşıma iyileştirmelerinin değerlendirilmesi, yetişkin başına bir veya daha fazla otomobil bulunan düşük gelirli hanelerin toplu taşıma erişilebilirliği ile toplu taşıma kullanımı arasında en esnek ilişkiye sahip olduğunu göstermektedir; toplu taşıma iyileşirse, toplu taşımayı kullanmaları olasıdır. Bununla birlikte, ulaşımın zayıf olduğu araç odaklı bölgelerde, düşük gelirli hanelerin toplu taşıma kullanımı, araba sahipliği arttıkça keskin bir şekilde düşmektedir. Bu durum, düşük gelirli otomobil sahibi hanelerin, araçlarına sahip olur olmaz çok fazla bağımlı hale gelebileceğini ima etmektedir. Ayrıca, erişilebilirlikteki değişikliklerin toplu taşıma yolculuk üretimini nasıl etkilediğini araştıran duyarlılık analizi, bölgedeki erişilebilirlik kazanımlarının, ulaşım iyileştirildiğinde, otomobil eksikliği olan haneler arasında toplu taşıma yolcu sayısını artırmak için daha fazla fırsat sağladığını vurgulamaktadır. Bu nedenle, analiz, bireyleri toplu taşımaya katılmaya ve bölgede genel toplu taşıma yolculuğunu artırmaya yönelik öneriler sunmaktadır.

Model seçimi göz önüne alındığında, sonuçlar, ML algoritmalarının diğer tüm istatistiksel modellerden daha iyi performans gösterdiğini ve yorumlanabilirlikten ödün vermeden seyahat davranışı tahminlerini geliştirmek için büyük potansiyele sahip olduğunu göstermektedir. Rastgele Orman algoritması (RF), XGBoost (XGB) ve Sinir Ağları (NN) sınıflandırıcıları ve regresörleri, diğer algoritmalardan önemli ölçüde daha iyi performans göstermektedir. Bunlar arasında RF, tahmin performansına göre düşük gelirli ailelerin transit talebini tahmin etmek için en doğru yaklaşımdır. Ancak, toplu taşıma kullanıcılarının davranışlarını tahmin ederken istatistiksel modeller zayıf performans gösterir. Ayrıca, toplu taşıma iyileştirmelerinden sonra yeni oluşturulan toplu taşıma yolculuklarının mekansal dağılımı aynı değildir; bu nedenle, geleneksel modeller sosyal, mekansal ve çevresel sorunları ele alırken farklı, muhtemelen yanlış bir politika önerisine ulaşabilir. Ayrıca, ML modellerine modelden bağımsız yorumlama araçlarının uygulanması, bu tekniklerin her modelin bir "kara kutu" olması gereken temel sürecini ortaya çıkarabileceğini göstermektedir. Sonuç olarak, ML modelleri doğruluk ve yorumlanabilirlikte önemli bir gelişme göstermektedir.

Bulgular, bireylerin yolculuk kararlarını ve tercihlerini güvenilir bir modelle anlamanın ve tahmin etmenin, politika yapıcıların düşük gelirli insanlara fayda sağlayan ve toplumdaki toplu taşıma eşitsizliğini azaltan uygun bir toplu taşıma çerçevesi oluşturmasına olanak tanıdığına işaret etmektedir. Bu çalışma, bireylerin yolculuk davranışlarını gelir ve araç sahibi olma düzeyleri açısından değerlendirmenin büyük şehirlerde ulaşım planlamasına yeni ve farklı bir bakış açısı kazandırabileceğini düşündürmektedir. Genel olarak, çevresel, ekonomik ve sosyal hedefleri karşılayan adil bir ulaşım yatırımı, farklı sosyoekonomik grupların seyahat gereksinimlerinin ve davranışlarının kapsamlı bir şekilde anlaşılmasını gerektirir. Bulguların, plancılara yol göstererek, bireylerin aktivitelere katılımı artıran ve çevresel etkileri azaltan ulaşım politikalarını ve stratejilerini yeniden düşünmelerine yardımcı olacağı ümit edilmektedir.

1. INTRODUCTION

1.1 Chapter Overview

This chapter provides an overview of the equity-based transportation planning and travel behaviour models. Section 1.2 discusses the background of the study, and Section 1.3 lists the research objectives and focus. Section 1.4 provides the main research questions of the thesis. Lastly, the structure of the thesis is summarized in Section 1.5.

1.2 Background

Travellers aim to reach their desired destinations or engage in activities at different locations. To forecast this travel demand, planners and researchers often require a comprehensive understanding of the travellers' trip decisions and activity patterns. Despite several definitions, a trip chain in this study is considered as a composition of consecutive activities scheduled over a period of time, started and terminated at home (Primerano, Taylor, Pitaksringkarn, & Tisato, 2008). The analysis of these trip chains (also known as tours (Bowman & Ben-Akiva, 2001; Krizek, 2003)) allows travel demand planners and policymakers to examine how individuals plan their daily trips and which factors may influence their travel decisions (Strathman, Dueker, & Davis, 1994; Currie & Delbosc, 2011a). It also uncovers its consequent impacts on the number of stops, trip sequences, distance traveled, time allocation, and travel mode use.

Identifying the individual daily activity patterns regarding their socioeconomic characteristics and land use variables is a key dimension in transport planning. In recent years, growing attempts have been made about the mobility needs of low-income communities, who are at the risk of transport disadvantage (Lucas, 2012; Martens, 2016). These groups usually experience disproportionate accessibility barriers (Tiznado-Aitken, Lucas, Muñoz, & Hurtubia, 2020) and have experienced an extensive relocation into the suburbs in many cities due to increasing decentralization or suburbanization (Hulchanski, 2010; Hochstenbach & Musterd, 2018; Allen & Farber, 2021). In Canada, several studies show that the lack of transportation services in inner suburbs hinders low-income individuals' ability to participate in activities (Foth, Manaugh, & El-Geneidy, 2013; Allen & Farber, 2020b). Accordingly, a broader understanding of the daily trip chains and travel patterns of low-income populations would give a better overview of their travel behaviour, decisions, and needs. Stratifying low-income people with homogeneous travel decisions can help develop inclusive and targeted transport projects. Furthermore, improving mobility equity and unlocking suppressed activity of low-income groups have the potential to help them overcome social exclusion difficulty and benefit from transport policy decisions.

By the same token, transportation equity advocates recommend improving public transit in low-income neighbourhoods to alleviate socio-spatial inequalities and increase the quality of life. However, transit infrastructure investments historically have largely focussed on attracting choice riders in an effort to take cars off the road, and reap congestion and environmental benefits (Bhattacharjee & Goetz, 2012; Carey, 2002; Pucher, 2002). As a result, many socioeconomically disadvantaged communities, home to transit-dependent populations, were largely overlooked during the transit planning processes of the post-war era. The rationale is that investing in low-income neighbourhoods, where transit ridership is already very high, would be less likely to result in mode-shifting, congestion relief, and environmental benefits. More recently, justice and equity objectives for transportation investments are receiving growing attention in both research and planning practice. The focus is shifting towards the alleviation of transport disadvantages to encourage fairness in the opportunity for people to reach daily activity destinations.

With the justice turn in transportation planning, much more information is now available about the social benefits of achieving more equity in the distribution of transit benefits among population groups, including rationales grounded in theoretic (Lucas, 2012; Martens, 2016) and empirical work (Allen & Farber, 2020b; Stanley et al., 2011). But in striving for equity, must planners put aside their desires to similarly achieve the conventional benefits of congestion relief and environmental benefits? How true is the received wisdom that investments in "transit-dependent" communities will not result in sizable benefits associated with growth in transit mode share? This study argues that too little is known about the degree of transit demand in low-income communities, and how sensitive low-income populations are to transit accessibility improvements. More research is needed to better understand whether the social goals associated with low-income transit investments can align with the congestion and environmental goals associated with growing transit mode shares.

A transport planning model plays a key role in evaluating, estimating, and managing changes in behavioural patterns. Accurate modeling of travel behaviour is an important component in transportation planning and travel demand management. For decades, extensive efforts have been devoted to identifying and improving methods in travel behaviour research, including such watershed moments as the derivation of discrete choice models, the shift from trip-based to activity-based models, and, more recently, experimentation with "big data" and Machine Learning (ML) methodologies. These technical and theoretical advancements go hand in hand with developing more nuanced understandings of the travel needs and revealed travel outcomes of different members of the population. From a social justice perspective, improving model accuracy is therefore vitally important to understand how different people respond to different types of changes in their transportation and land use environment and, accordingly, how to better plan for the needs of historically marginalized communities. Within justice-based transportation planning, travel behaviour models help researchers predict activity and travel behaviour outcomes associated with transit investments, which help planners evaluate the equity implications of different planning scenarios. Moving beyond the typical buffering exercises involved in US-based Title VI and Environmental Justice analyses, travel behaviour-based assessments examine how transit investments unlock potential for higher life quality. It can be done by forecasting, by population segment, behavioural responses such as changes in auto-ownership, transit mode share, and out-of-home activity participation rates.

1.3 Research Objectives

Building on the discussion provided in Section 1.2, this thesis follows three main objectives in association with the transportation and equity concept.

- First, this study investigates the travel pattern of different households within the Greater Toronto Area and Hamilton and examines how income and car-ownership levels affect their travel decisions and behaviour. Particularly, it explores whether there are travel pattern differences between carless and car-owners of low- and high-income households. To do this, a new approach for detecting and clustering travel patterns of people is used. The aggregation of activity types and transport mode used as the unit of analysis is considered. This analysis of trip sequences as a whole gives precise insights into individuals' travel preferences and constraints. Moreover, it offers a more comprehensive framework to evaluate which transport policies can yield benefits or impose burdens on different population groups.

- Second, the importance of considering traditional goals in conjunction with those of social equity through travel behaviour change in low-income neighbourhoods is discussed. Accordingly, this thesis investigates how different income and car-ownership groups respond to transit accessibility improvements. Further, it explores the extent to which transit investments in low-income neighbourhoods are likely to increase transit use, therefore reducing vehicle kilometers traveled (VKT), traffic congestion, air pollution, and other externalities. Consequently, the findings can guide planners and policy-makers to take account of transit investments in low-income neighbourhoods for alleviating both transport and financial burdens while reaping the societal benefits of positive environmental and congestion outcomes.

- Finally, this thesis aims at making a methodological contribution to transportation planning and policy-making. The study assesses the potential for using ML-based travel behaviour models to accurately predict travel behaviour responses to transit investments among marginalized populations. Moreover, it investigates the interpretability of ML models compared to the traditional approaches. Since the 1980s, most mode-choice problems have been addressed and modeled by traditional discrete choice models – e.g., multinomial logit. However, ML's flexibility in dealing with non-linearity and capturing complex and previously undiscovered relationships between input and output variables makes them promising for modeling heterogeneous travel behaviour patterns. To the best of our knowledge, there have been few efforts that examine the pros and cons of using ML models within equity-focused research and planning. The end of this study is to discover how the potential forecasting accuracy benefits of ML approaches stack up against their potential drawbacks, namely, that they are not derived from behavioural theory, and do not provide easily interpretable relationships between input and output variables.

1.4 Research Questions

The research questions are classified into three main parts to expand the discussion with respect to each research objective. This thesis answers each research question in its dedicated chapter as follows.

(1) Investigating how income and car-ownership levels determine activity patterns and mode choice decisions (see Chapter 5)

(RQ1-1) How does car-ownership affect the trip chaining behaviours of low-income communities?

(RQ1-2) How do the trip chaining decisions of low-income households differ from those of high-income households?

(2) Exploring how transit investments affect mode choice decisions of households with different income and car-ownership levels (see Chapter 6)

(RQ2-1) To what extent can transit investments in lower socio-economic neighbourhoods enhance transit mode share?

(RQ2-2) To what extent are low-income car-owners sensitive to transit improvements and shift their travel mode use?

(3) Analyzing how the model selection (e.g., statistical and ML algorithms) influences travel behaviour prediction, transportation planning, and policy-making (see Chapter 7)

(RQ3-1) How accurate are ML models compared to traditional models in predicting travel behaviour in response to transit investments?

(RQ3-2) To what extent are ML models interpretable?

1.5 Thesis Structure

This first chapter provides an introduction to the thesis. The contents of the following chapters are summarized below.

- Chapter 2 discusses the transportation and equity concept, reviews travel behaviour literature, debates transit investments and mode shift decisions, and provides an overview of several models used for travel behaviour analysis in transportation planning.
- Chapter 3 introduces the study area for undertaking the analysis. It also provides a description of the data source used and a descriptive summary of the dataset.
- Chapter 4 describes the algorithms used for the analyses within the scope of this thesis. The chapter sheds light on the structure of the models estimated in the next chapters of the thesis.
- Chapter 5 explores how travel patterns of households in the Greater Toronto and Hamilton Area differ by different income and car ownership levels using a cluster-based framework.
- Chapter 6 extends the scope of the study undertaken in Chapter 5 and investigates whether transit investments can affect transit trip generations of residents, particularly low-income car-owners. This work is covered in a paper titled "Can transit investments in low-income neighbourhoods increase transit use? Exploring the nexus of income, car-ownership, and transit accessibility in Toronto", published in the journal of *Transportation Research Part D: Transport and Environment* (Yousefzadeh Barri et al., 2021).
- Chapter 7 builds on Chapter 6 and analyzes how the model selection affects travel behaviour estimations, taking transit trips, while comparing statistical and ML models.
- Chapter 8 concludes the thesis and summarizes the findings of the studies. It provides policy recommendations according to the results obtained from the experiments, discusses the limitations on conducting the study, and suggests directions for future works.

2. LITERATURE REVIEW

2.1 Chapter Overview

This chapter provides an overview of the main research areas related to the studies undertaken within the scope of this thesis. Section 2.2 opens a brief discussion on the transportation network and its social impacts, Section 2.3 reviews studies about the travel behaviour and patterns of travellers. Section 2.4 debates whether transit investments in low-income communities may shift travel modes, and Section 2.5 focuses on methods used in travel behaviour analysis. Section 2.6 points out the gaps in the literature and the contribution of this thesis in filling these gaps.

2.2 Transportation Planning and Equity

Mobility and the ability to travel are the essential requirements of individuals in their everyday lives. The primary advantage of transportation infrastructure is (should be) to enhance this ability and allow all people to travel without any difficulties. In this way, any difficulties, defects, or deprivation in transport systems will create an improper situation for mobility that leads to the transport inequality. Therefore, one of the main objectives of transportation planning is to increase activity participation by empowering people to fulfill their mobility needs (Martens, 2016). It is achievable only by understanding how various socioeconomic strata schedule their daily trips and what factors influence their choices. In recent years, improving the available transport resources of low-income communities, enhancing their access to opportunities, and reducing the risk of social exclusion have been the subject of considerable academic debates (Lucas, 2012; Manaugh, Badami, & El-Geneidy, 2015; Martens, 2016). Several studies in Canadian cities show that a large number of low-income households are living in neighbourhoods with inadequate levels of transit services, moving from the city center to suburbs (Manaugh et al., 2015; Allen & Farber, 2020b). Due to several transport barriers, their activity participation is lower, and they encounter more suppressed trips. Some low-income households may prefer to own a private vehicle due to the lack of transport alternatives, generating forced car ownership and car-dependence (Mattioli, Anable, & Vrotsou, 2016; Mattioli, 2017), which usually translates into considerable car ownership costs and risk of indebtedness (Currie & Delbosc, 2011b; Walks, 2018).

From a political economy perspective, transportation projects and investments are theorized to be formed by the interest of powerful actors or power associations. These influential groups intervene in policy-making decisions to support their values and interests (Glaeser & Ponzetto, 2018). There may be evidence of this in the Toronto case, with much of the transit expansion in the region occurring in the form of commuter rail lines that mainly link middle- and upper-income suburban areas to the Central Business District. By focussing investments in suburban rail expansion for long-distance commuters, the needs of inner-suburban residents risk going unmet (Giuliano, 2005; Brown & Thompson, 2009). Indeed, the City of Toronto has the majority of North America's highest ridership bus routes, almost all operating in mixed traffic, and under crowded or crushed conditions during both peaks. These riders remain disempowered due to systemic marginalization along with income, race, and immigration lines (Hertel, Keil, & Collens, 2016; Lo, Shalaby, & Alshalalfah, 2011; Palm, Shalaby, & Farber, 2020). Moreover, they have largely been unsuccessful in attracting transit improvements for their daily, local travel needs. As a result, transportation planners and policymakers need to gain a more comprehensive insight into the differences between the travel behaviour of low-income and high-income individuals if they want to address the daily travel needs of disadvantaged groups.

2.3 Travel Behaviour Analysis

Over recent decades, researchers have focused on traveller's trip chains as an indicator of travel behaviour to investigate people's travel patterns and predict future travel demand (Ma, Mitchell, & Heppenstall, 2014; Y. Huang, Gao, Ni, & Liu, 2021). Although there is no unanimous definition of a trip chain in travel behaviour literature, it is primarily defined as a sequence of activities with single or multiple stops that begins from and ends at home (McGuckin & Murakami, 1995; Shiftan, 1998; Primerano et al., 2008). Under this definition, a movement between a pair of activities or stops is called a trip segment or trip leg. The number of activities, travel mode choice, duration of travel, the complexity of the trip chain, and distance traveled are examined to understand trip chain mechanisms and users' activity patterns (Currie & Delbosc, 2011a; Goulet-Langlois, Koutsopoulos, & Zhao, 2016; Schneider et al., 2021). Researchers usually emphasize socioeconomic and spatial factors such as individual characteristics, household structure and built environment attributes that may affect traveller's decisions and behaviour (Cervero & Kockelman, 1997; Currie & Delbosc, 2011a; Ma et al., 2014). Hence, individuals' travel schedule preferences and decisions, along with other socioeconomic and spatial factors, may result in heterogeneous travel behaviour outcomes. Studying the travel behaviour of travellers may allow us to evaluate where and to what extent transport investments alleviate travel barriers and improve individuals' activity participation.

Several travel surveys and studies have demonstrated that economically and socially disadvantaged groups, particularly low-income households, use public transit more frequently than other socioeconomic categories (Giuliano, 2005; Pucher & Renne, 2003; Rosenbloom, 1998). Furthermore, due to structural racism and sexism, racialized people, women, and non-binary people are more likely to have lower incomes and are more likely to face safety issues when traveling from harassment and threat of violence (Oswin, 2014; Scholten & Joelsson, 2019). This transit dependency becomes more evident while exploring the relationship between gender and mobility (Ravensbergen, Fournier, & A, 2022). Often, low-income women have less access to a private car and drive fewer times than men (Naess, 2008; Madariaga, 2016). Given the lack of transit services, active transport infrastructure, and access to a personal vehicle, they face numerous mobility challenges. Furthermore, most women make most non-work lengthy trips due to carrying a disproportionate burden of household responsibilities and caring tasks (Madariaga, 2016; J. Lee, Vojnovic, & Grady, 2018; Craig & van Tienoven, 2019). Therefore, although women undertake more non-work trips due to uneven division of household tasks, they also have less propensity to get a car as a driver and rely on other modes (Vance & Iovanna, 2007; Scheiner & Holz-Rau, 2012).

Travel behaviour analysis in some car-dependent cities shows that low-income households make fewer and shorter trips. They are more likely to walk than high-income households since they are living in neighbourhoods with proper access to destinations, greater street connectivity, and a better land-use mix (Turrell, Haynes, Wilson, & Giles-Corti, 2013; Foth et al., 2013; Sagaris & Tiznado-Aitken, 2020; Allen & Farber, 2020b). In contrast, wealthier families have more tendency to chain trips and have complex tours compared to low-income populations (Ye, Pendyala, & Gottardi, 2007; Cheng, Bi, Chen, & Li, 2013). Most of these multi-stop tours are taken by car in that a flexible travel mode is required to chain multiple trip legs in a single journey. This difference could be due to the high costs of trips and activity participation – whether in time or money, or due to other time-geographical or accessibility limitations. This suppressed demand offers an opportunity to improve mobility equity by removing barriers and equalizing the number of trips regardless of income, ceteris paribus. Achieving this goal is made difficult by the relatively recent reversal of the income-distance gradient observed across many global cities, including the GTHA (Kneebone & Garr, 2010; Glaeser, Kahn, & Rappaport, 2008). Poverty is increasing in the suburbs partly due to inner-city gentrification and the changing geography of affordable housing (Ding, Hwang, & Divringi, 2016; Ellen & O'Regan, 2011; Pucher & Renne, 2003). The combination of the auto centric design of cities with the suburbanization of poverty has resulted in a large group of financially constrained drivers who are driving because of a lack of alternatives, as well as transit users living in poorly served neighbourhoods far from social and economic activities. Consequently, these conditions serve to suppress activity participation, or shift travel burdens unduly on already structurally marginalized groups, further worsening the risks of social exclusion (Allen & Farber, 2021; Lucas, 2012; Martens, 2016).

2.4 Transit Investments and Mode Shift

Planners have traditionally focused on the value of transit investments and their efficiency in terms of environmental sustainability and value-of-time savings. This measurement regime ultimately supports the goal of reducing car-based trips via attracting choice riders to transit (Richmond, 2001). From this perspective, numerous transport agencies and planners are evaluating the performance of rail projects with a congestion-relief target or conducting air quality analyses. For instance, Bhattacharjee and Goetz (2012) have analysed how successful the newly opened light rail system in

Denver was in increasing transit ridership, taking cars off the road, and thus relieving congestion. They measured the temporal and spatial changes in the levels of highway traffic in terms of VMT changes. Their main purpose was investigating the spatial distribution of riders who switch from car to transit reflected in average VMT changes over 16 years. Research by Baum-Snow, Kahn, and Voith (2005) explored the effects of new or extended transit rail between 1970 and 2000 on the transit mode share in sixteen major U.S. cities. Their study illustrated that rail transit projects do not necessarily lead to an overall increase in transit ridership; instead, increases in rail transit ridership stemmed from those switching from bus to rail. Nevertheless, these conventional approaches overlooked socioeconomically disadvantaged communities, their needs, and behaviour during the transit investment process.

Blumenberg and Thomas (2014) indicated that car-dominant countries have witnessed the rapid increase in car ownership among low-income households. Focusing on the impact of private vehicles on trips, Blumenberg and Pierce (2012) have identified the profound impact of car ownership on the increase in travelled miles of low-income adults. Furthermore, some researchers have found that auto ownership plays a crucial role in accessing employment opportunities and higher earnings for vulnerable groups (Gurley & Bruce, 2005; Raphael & Rice, 2002). Baum (2009), for instance, measured the effect of car ownership on the probability of employment using a longitudinal survey, concluding that owning a car significantly produces positive employment outcomes and promotes welfare receipt exits. Curl, Clark, and Kearns (2018) voiced a similar concern when evaluating the trend of car ownership regarding financial difficulties for households living in disadvantaged communities between 2006 and 2011 in Glasgow. They found that a large number of car owners keep their cars despite experiencing economic stresses because they consider it as a necessity for reaching their life opportunities. Their findings showed that having children and searching for a job deter the majority of low-income households in deprived neighbourhoods to relinquish their private vehicles. Moreover, the necessity of having a private vehicle to meet mobility needs forces them to buy a car, even if they are unwilling or cannot afford it (Curl et al., 2018; Potoglou & Kanaroglou, 2008; Pucher & Renne, 2003). Therefore, it may be true that transit investments in those low-income neighbourhoods will not help with mode shift because car-ownership brings a variety of opportunities for the poor, and they are committed to car-use after the heavy investment made in the car.

Other studies have shown unstable car ownership trends in low-income communities. This population segment more frequently changes its car ownership rate across time compared to other groups in society (Klein & Smart, 2017). Using a mobility biography approach, Klein and Smart (2019) have examined how a variety of life events affects the car ownership decisions of households over ten years. Their findings revealed that losing a job or worsening health has the most significant effect on giving up a car for the poor. Besides, looking at the financial burden of having a car, Currie and Senbergs (2007) explored the relationship between car-related expenditures (e.g., car purchase, insurance, and charges) and its financial difficulties for low-income car owners living on the fringe of Melbourne. They found that low-income, car-owning households living in outer Melbourne make fewer and shorter distance trips compared to other car owners in the same region. Interestingly, these fewer trips are highly reliant on their private vehicles and less frequently done by transit. They prefer car trips to transit trips, probably because using their vehicles provides a reduced cost of travel. Therefore, they do not opt to pay for transit when they can drive comfortably and less costly to their destinations. Consequently, they suggest transit investments could address transport disadvantage and mitigate financial burdens on vulnerable households as they believe that people in poverty will give up their private vehicles.

2.5 Travel Behaviour Analysis Models

Travel behaviour analysis often aims to forecast travel demand in a city based on future transportation investments. When individual travel information is available, a disaggregated model easily captures traveller's behavioural patterns and predicts their decisions following changes in the transportation system. For transit equity analysis, the impacts of changes in transportation plans on different communities and equity outcomes of those policies can be evaluated by predicting travellers' participation in activities (Hodgson & Turner, 2003; Martens, 2016; Fransen, Farber, Deruyter, & De Maeyer, 2018; Allen & Farber, 2020b). Accordingly, planners understand to what extent the outcome of policy options might be inequitable across different segments of the population. For instance, evaluating transit mode share or accessibility among people who have been historically marginalized enables transport planners and decision-makers to explore whether inequities are being redressed by proposed transit projects. Reviewing studies reveals that spatial attributes of the built environment (Cervero & Kockelman, 1997; Ewing & Cervero, 2010), socioeconomic characteristics of travellers (Turner & Niemeier, 1997; Dieleman, Dijst, & Burghouwt, 2002), and the level of transit service (Taylor, Miller, Iseki, & Fink, 2003; Moniruzzaman & Páez, 2012) affect an individual's travel mode choice and behavioural pattern. Thus, it is desirable to have a disaggregated, comprehensive and accurate travel behaviour model that deals with complex behavioural attributes and predicts travellers' responses to the variations in the transportation system.

2.5.1 Trip chain analysis

Transportation planners and policy-makers gain greater information for travel demand management by predicting residents' travel behaviour and clustering daily activity patterns through trip chain analysis. A considerable body of trip chain analyses have focused on understanding the activity types and travel mode decisions as researchers believe that the order of these decisions dramatically shapes travel patterns (Ye et al., 2007; Yang, Shen, & Li, 2016). Some studies have also explored the relationship between trip chaining complexity and individuals' travel mode choices (Currie & Delbosc, 2011a; Schneider et al., 2021). In trip chain assessments according to trip purposes, researchers usually divide trip destinations into work or non-work trips under different assumptions. Afterward, they explore the pattern of these predefined sub-divisions (Frank, Bradley, Kavage, Chapman, & Lawton, 2008; Ma et al., 2014; Chowdhury & Scott, 2020). For example, Ye et al. (2007) classify tours into work trips, if a trip chain has at least one working stop, or into non-work trips if no working stop exists. Frank et al. (2008) develop a hierarchy of the trip purposes and travel modes to determine if a trip sequence belongs to work trips or non-work ones. On the other hand, Ma et al. (2014) define work and non-work trips in multipurpose tours according to the order of work and non-work stops. Thus, there is no unanimous guideline even for dividing the trip activities into two simple categories. While some researchers group trip chains into work and non-work trips based on their predefined rules, others classify them into car or transit trips and then explore pattern differences of each sub-division.

Besides trip purpose evaluation in trip chain analyses, travel mode decisions are also considered a primary predictor of travel behaviour patterns. Researchers rely on smart-card data (e.g., transit or bike-sharing data) or mode-based travel surveys to investigate the travel patterns of residents and their specific journeys in terms of their mode decisions (Chu & Chapleau, 2010; J. Zhao, Wang, & Deng, 2015; Goulet-Langlois et al., 2016; Y. Zhang, Brussel, Thomas, & van Maarseveen, 2018). Other studies search for bidirectional causality of mode choice decision and trip chain characteristics using several choice models such as the multinomial logit model (Wan et al., 2019), probit model (Ye et al., 2007), and nested logit model (Y. Huang et al., 2021). Using the national household travel survey, Rafiq and McNally (2021) explore the activity patterns of transit users and cluster them according to their similar travel behaviour. They set an arbitrary rule and identify individuals who have at least one transit trip segment in their trip chain as transit users. In another study, Ho and Mulley (2013) group trip chains with multiple modes into a travel mode class based on the longest distance traveled by the given travel mode. To explore the causality of bicycle choice and activity pattern, Z. Li, Wang, Yang, and Jiang (2013) conduct a study to capture the order of decision. They define a tour as a bicycle tour if a bicycle is the dominant trip mode. These studies provide insights into individuals' activity patterns based on their mode decisions, although they do not reflect the interconnection of trip purposes and the transition from one mode to another in the daily tours.

Another thread of activity pattern analysis in travel behaviour literature is comparing an episodic sequence of states. The spatio-temporal structure of activities in trip chain schedules is explored and compared to understand how individuals allocate their time to specific activities. Using GPS-based data stream or space-time activity surveys, researchers track and capture the location and time duration of activities and trips of each individual (Song et al., 2021; Hafezi, Liu, & Millward, 2019). Then, they generate sequences of activities according to the visited locations and time spent. Accordingly, the unit of analysis is a trip sequence representing the status of an individual with temporally equal intervals (e.g., 5 min intervals). For instance, a 2-hour stop in a workplace from an individual's trip chain is transformed into a sequence of 24 workplace stops at 5-min intervals (Song et al., 2021; Saneinejad & Roorda, 2009). As a result, each trip segment of a trip chain would be a converted trip sequence with several adding dummy characters to reflect the duration of the activity.

Most of these activity sequence studies use developed sequence alignment methods, applied first by Wilson (1998) in social science and derived from DNA sequence studies. They compare the activity episodes of individuals to measure the distance of states and then cluster them based on their similar activity patterns (Saneinejad & Roorda, 2009; Kwan, Xiao, & Ding, 2014; F. Liu, Janssens, Cui, Wets, & Cools, 2015; Hafezi et al., 2019). Using smart card data, Goulet-Langlois et al. (2016) analyze the heterogeneity among transit users through clustering their activity sequences in 4 weeks. This study represents the structure of activity sequences using transit and their variability over a month. A similar analysis by Song et al. (2021) investigates individuals' activity patterns using GPS-based tracking data. They applied the sequence alignment approach to sequences of users' daily travel diaries to cluster their activity patterns. This activity sequence analysis focuses on individuals' duration of the activity and cannot determine traveller's transition decisions thoroughly. It also requires a detailed and large amount of data. Although such analyses enrich travel behaviour studies, they are unable to show the transition between destinations and individuals' travel decisions.

2.5.2 Travel mode use analysis

Among travel behaviour studies, mode choice modelling and different ML algorithms have received particular attention for exploring and predicting individuals' travel mode decisions (Koushik, Manoj, & Nezamuddin, 2020). Several ML algorithms including Decision Tree (DT), Random Forest (RF) (Cheng, Chen, Yang, Wu, & Yang, 2019; Yan, Liu, & Zhao, 2020), Extreme Gradient Boosting (XGB) (Wang & Ross, 2018; Shao, Zhang, Cao, Yang, & Yin, 2020), Neural Networks (NN) (Xie, Lu, & Parkany, 2003), and Support Vector Machine (SVM) (Y. Zhang & Xie, 2008; Zhou, Wang, & Li, 2019) for travel demand and mode choice predictions. Traditional discrete choice or statistical models have been adopted to explore travel mode decisions of marginalized groups (e.g., (Mercado, Paez, Farber, Roorda, & Morency, 2012; Jiao & Wang, 2021)). Socially disadvantaged and low-income individuals, who are unable to afford a car, are often marginalized and are greatly faced with transportation barriers (Lucas, 2012; Lucas, Philips, Mulley, & Ma, 2018). However, there are still limited studies of the travel behaviour of marginalized populations, most at risk of transportation disadvantage, using ML models.

In evaluating transportation policies, both classification and regression algorithms are used to predict future behavioural outcomes. The next section first summarizes how conventional and ML models have been applied to two main types of transportation questions: classification and count predictions. Then, it explores studies comparing statistical and ML models in travel behaviour. Afterward, it discusses the trade-off between predictive performance and interpretability of ML models and our contribution to the literature.

2.5.2.1 Classification models

Classification models are commonly used in different areas of transportation research, including examining the crash injury severity (F. Hu, Lv, Zhu, & Fang, 2014; Rezapour, Moomen, & Ksaibati, 2019), predicting household car ownership (Curl et al., 2018), and uncovering activity participation probability (Allen & Farber, 2018). More specifically, discrete choice models are extensively applied in mode-choice analysis for predicting a commuter's travel mode use (Mercado et al., 2012; Jun, Kim, Kwon, & Jeong, 2013), exploring gender equity in bicycle mode choice (Abasahl, Kelarestaghi, & Ermagun, 2018), or investigating the role of shared mobility in serving the transit-dependent populations (Jiao & Wang, 2021). Notably, all these studies focus on conventional statistical models such as Binary and Multinomial Logistic Regression, Multinomial

Logit model (MNL), Nested Logit model (NL), and Mixed logit model. Although these models facilitate straightforward interpretation, they come with their own limitations. They require predefined assumptions about the data for modelling (Cheng, Chen, De Vos, Lai, & Witlox, 2019), and any violation of these assumptions gives rise to a biased and unreliable prediction (J. Lee et al., 2018; Wang & Ross, 2018). For instance, one of the restrictions of MNL is the IIA assumption that implies the choice probability of each pair of options does not change if the other alternatives are absent (McFadden, 1973). However, NL and Mixed logit, other variants of the standard logit, have addressed this shortcoming by having more sophisticated model specifications, thus improving modeling performance compared to MNL. In contrast to statistical models, ML algorithms with their data-driven nature have become a promising alternative for predicting individual's behavioural responses (Koushik et al., 2020). Unlike statistical models, they are assumption-free algorithms that detect and learn patterns from the existing (observed) data and then apply them to predict unobserved data (Murphy, 2012). Furthermore, discrete choice models rely on the principle of utility maximization theory -i.e., individuals' preferences and behaviour are explained in terms of gaining the most benefit and highest satisfaction from their decisions. Moreover, in a discrete choice model, a modeler requires to manually input the relationship between features and labels, whereas ML models can learn those complex relationships from the dataset (Hillel, Bierlaire, Elshafie, & Jin, 2021), and they become a suitable alternative for traditional approaches. Nonetheless, the interpretability of ML models, considered to be black-box models, is still a big concern in the domain (Rudin, 2019). There are also heated debates regarding the bias and ethics in ML algorithms, especially against females, underrepresented groups, and low-income people (Mehrabi, Morstatter, Saxena, Lerman, & Galstyan, 2021; I. Y. Chen et al., 2021; N. T. Lee, 2018). Therefore, while using them, a researcher needs to consider different aspects to have a fair and interpretable ML model.

2.5.2.2 Count models

Travel behaviour researchers use statistical count models for predictive modeling. A wide range of regression models, including Ordinary Least Square regression, Zero-Inflated count models, Hurdle, Negative Binomial, and Poisson, are popular in travel behaviour studies. These methods are utilized for estimating commuters' transit trip frequency (Yousefzadeh Barri et al., 2021; Böcker, van Amen, & Helbich, 2017; Legrain, Buliung, & El-Geneidy, 2016), variations in bicycle ridership counts (Roy, Nelson, Fotheringham, & Winters, 2019), the relationship between residential self-selection and non-work trips (Chatman, 2009), built environment and trip generation association (Q. Zhang, Clifton, Moeckel, & Orrego-Oñate, 2019), bus fare evasion rates (Guarda, Galilea, Paget-Seekins, & de Dios Ortúzar, 2016), and peak-car phenomenon (leveling off car travel) (Kamruzzaman, Shatu, & Habib, 2020) as they are easy to apply and interpret. The implementation and interpretation of statistical models are straightforward, although they fail to handle the nonlinear relationship between variables.

Typically, the advantage of ML models is to learn and model the intricate interactions between a dependent and a set of independent features (Xie et al., 2003; Cheng, Chen, Yang, et al., 2019), making them suitable for modelling travel behaviour. Accordingly, many studies have recently used ML algorithms in transportation studies. For instance, Shao et al. (2020) investigate the nonlinear relationship between land use and metro ridership at the station level using the XGB model. In modelling travel demand for ride-sourcing, Yan et al. (2020) estimate the number of ride-sourcing trips using the RF algorithm and the traditional multiplicative model and then compare their predictive performance in forecasting the future number of ride-sourcing trips.

Several studies have compared the predictive performance of ML models with that of statistical models, and most of them have indicated that ML models are promising tools compared to statistical models (e.g., (Xie et al., 2003; Hagenauer & Helbich, 2017; Cheng, Chen, Yang, et al., 2019; Zhou et al., 2019; Yan et al., 2020; E. Chen, Ye, & Wu, 2021)). Focusing on travel mode choice modelling, Hagenauer and Helbich (2017), Cheng, Chen, Yang, et al. (2019), Zhou et al. (2019), and X. Zhao, Yan, Yu, and Van Hentenryck (2020) use multiple ML and statistical classifiers to compare their predictive capability and behavioural analysis. Most of these studies focus on the classification problem rather than regression. However, there are a few studies constructing a regression problem using both statistical and ML models. For instance, T. Kim, Sharda, Zhou, and Pendyala (2020) have adopted mixed models to achieve more accurate travel demand prediction models. They consider the travel demand for on-demand ride-hailing services and develop a new framework integrating the

Linear Regression model (LinR) and Long-term Short Memory model to predict the potential travel demand. In a similar approach, E. Chen et al. (2021) propose a hybrid model incorporating Geographically Weighted Regression (GWR) in the structure of RF to consider spatial heterogeneity and explore the complex association between built environment variables and bus-metro transfers (trip generations from bus to metro). Then, they compare the predictive performance of Multiple Linear Regression, SVM, RF, GWR, and their hybrid model to examine their suggested model's advantages over other models. Therefore, ML models allow us to capture complex and nonlinear variable relationships and overcome the limitations of statistical modelling methodologies.

2.5.2.3 Machine learning applications in transit equity studies

Over the decades, investigating transit services with an equity lens has become a crucial concern. Recently, some researchers have begun to adopt different ML techniques to address regression, classification, or clustering problems for their transportation equity studies. For instance, Tran, Draeger, Wang, and Nikbakht (2022) explore the travel experience and behavioural responses of transit riders before and during the COVID-19 pandemic. They utilize Twitter data for sentiment analysis as an ML algorithm to understand how a significant transit disruption may aggravate the vulnerabilities of transit-dependent users. Jiao, Degen, and Azimian (2022) use the RF model to analyze e-scooter ridership in poorly transit-served neighbourhoods and address the inequality in transportation supply. In measuring the vulnerability of households to transport energy burdens, S. Liu and Kontou (2022) propose a new framework to quantify transport-related energy poverty. They apply linear regression and several non-linear methods (e.g., DT, XGB, RF, and NN) to estimate households' average fuel consumption. They find that low-income households are more at risk of higher transport fuel costs compared to their counterparts. Among transportation studies, substantial studies in travel demand prediction (e.g.,(Yan et al., 2020; T. Kim et al., 2020)), mode choice modelling (e.g.,(Hagenauer & Helbich, 2017; X. Zhao et al., 2020)), and traffic predictions (e.g.,(Cai et al., 2016; Y. Liu, Liu, & Jia, 2019; Cui, Ke, Pu, & Wang, 2020)) have adopted different ML algorithms. However, the potential of ML regressors in formulating the travel behaviour and transit use of low-income people and comparing their predictive performance with that of statistical models are not yet fully explored. Therefore, there is still a need to better investigate ML applications in the transit equity context in the literature. Moreover, interpretation of models is likely more important within an equity context, compared to more basic travel demand forecasting, as researchers and planners are interested in the specific policy levers they can use to improve the well-being of marginalized populations.

2.5.2.4 Interpretable machine learning models

With advances in ML techniques, their applications were applied in the transportation field. Despite their high predictive power, there is still a lack of research using ML techniques due to interpretability concerns (Rudin, 2019; Koushik et al., 2020). Even more important than the model performance is the interpretability of algorithms. After knowing the probability of an event's occurrence, it is required to discover how the prediction is made. Although an accurate predictive model can enhance the final decision, understanding the rationale behind the suggestions leads to a better insight into the problem. Researchers generally use inherently interpretable models rather than ML models that are assumed to be black boxes. Notably, the output of most ML models is not directly interpretable (X. Zhao et al., 2020). However, to decipher the way an ML model arrives at its conclusion, several interpretation techniques have been recently introduced. Unlike intrinsic interpretability, model-agnostic interpretability tools are employed following estimation of an ML model in a *post hoc* analysis. The flexibility of these *post hoc* interpretation methods lets researchers use any ML model in different fields. Applying interpretation tools to an algorithm is a way to understand the rationale behind the decision and justify the predicted outcome. *Post hoc* interpretability approaches uncover the effects of independent features on the response variable and interpret their influence to propose appropriate policy implications. Therefore, they summarize the behaviour of a model and explain how important a predictor is for the final decision enabling planners to identify key variables for effective decision making. More specifically, such explanations can provide enough evidence to implement policies for a given scenario.

When it comes to the granularity level of interpretability, global and local interpretation techniques describe the aggregated behaviour of the model and each individual or group, respectively (Du, Liu, & Hu, 2019; Molnar, 2020). Feature importance (Breiman, 2001; Hagenauer & Helbich, 2017), Partial Dependence Plot (PDP) (Friedman, 2001), and Accumulated Local Effects (ALE) (Molnar, 2020) are examples of global interpretability tools. They explain the effect of each feature on the average prediction of a model. Conversely, Individual Conditional Expectation (ICE) (Goldstein, Kapelner, Bleich, & Pitkin, 2015), Shapley value (SHAP) (Lundberg & Lee, 2017), and Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro, Singh, & Guestrin, 2016) explain how a model predicts an output for an individual (Molnar, 2020). Hence, contrary to the popular belief of black-box ML models, there are a plethora of *post hoc* tools that have the potential to address this limitation.

2.6 Summary and Research Gaps

Building on the above literature review, this thesis forms on the travel behaviour analysis and transit equity themes with the focus on activity patterns and mode choice decisions. According to the relevant studies from the literature provided in the previous sections, some research gaps, the key aspects of this thesis in terms of travel behaviour studies, and its methodological novelty are listed below.

As discussed in Section 2.3, travel behaviour of travellers have been investigated extensively using individual trip chains analysis. Previous studies mainly examined the travel pattern of travellers in terms of different socioeconomic factors, built environment variables, and trip characteristics. However, few studies have focused on clustering travel patterns of residents based on their income and car-ownership concerning their trip sequences. Given the importance of trip analysis, it is essential to study the association between trip purpose and its relevant mode for various income and car-ownership levels. This discussion is listed as the first research gap. Furthermore, as discussed in Section 2.5.1, trip chain studies mostly investigate activity patterns using predefined rules and assumptions. They also observe the activity types and mode choices separately which can not reflect their interconnections. This identifies as the second research gap which is filled by this thesis.

Moreover, if transportation planners and policymakers want to address the daily travel needs of disadvantaged groups in the context of transport equity, they need to understand the travel behaviour of low-income households. Therefore, this thesis further evaluate how different transport investments may alleviate travel barriers and improve individuals' activity participation, particularly those with higher constraints and needs. The studies on travel behaviour and travel mode choice of disadvantaged groups, particularly low-income households, show that their greater reliance on transit as debated in Section 2.4, partly implies why policymakers overlook them in the traditional transit planning process. On the other hand, several studies show that the rate of car ownership has increased among low-income households despite the huge financial burdens on their living expenses. Therefore, the need for evaluating transit projects and investments in low-income communities to reduce transit inequity and increase activity participation is highlighted. Exploring how much sensitive are low-income households to transit investments is the main concern of this discussion.

As emphasized in Section 2.5.2, there is still limited travel behaviour and mode choice studies exploring the application of ML models in predicting transit use and addressing low-income households travel needs. Table 2.1 lists travel behaviour studies that use ML algorithms and conduct a comparative analysis with statistical models. Most of the studies focus on mode choice modelling as a classification task. In comparing various ML algorithms, different predictive performance metrics are calculated, including accuracy, precision, recall, F1-Score, and AUC-ROC (Murphy, 2012). However, most of these studies have compared the predictive performance of models through accuracy alone. These metrics are thoroughly explained in Section 4.4 and the difference between these metrics and their reliance on the predefined threshold is discussed in Section 4.4.1. Therefore, a lack of understanding about other performance measures is remarkable in the literature. Moreover, only one of them uses statistical tests to compare the significance of the differences among models (Hagenauer & Helbich, 2017).

Table 2.1 : Contribution of the current study compared to the literature.

	Unit of analysis	Supervised Learning		Statistical Test	CV	Interpretability	
		classification	regression			global	local
Xie et al. (2003)	Mode choice analysis	✓				✓	
Y. Zhang and Xie (2008)	Mode choice analysis	✓					
Hagenauer and Helbich (2017) *	Mode choice analysis	✓		✓	✓	✓	
Wang and Ross (2018)	Mode choice analysis	✓			✓	✓	
Cheng, Chen, Yang, et al. (2019) *	Mode choice analysis	✓				✓	
Zhou et al. (2019) *	Mode choice analysis	✓			✓		
X. Zhao et al. (2020) *	Mode choice analysis	✓			✓	✓	
Yan et al. (2020)	Ride-sourcing trips		✓		✓	✓	
E. Chen et al. (2021)	Bus-metro transfers		✓		✓	✓	
Our study *	Transit trip generation	✓	✓	✓	✓	✓	✓

* These studies have compared more than two different ML algorithms with statistical models.

Not all travel behaviour studies have used a validation technique for the unbiased train-test split – e.g., cross-validation (CV) – to compare the algorithms' performances.

More importantly, the local interpretability of ML models is mainly disregarded. These works mostly applied global interpretability tools and ignored the importance of local interpretation. This study leverages statistical tests and interpretability tools to shed light on the differences among models and their possible explanations. The novelty and difference of this study compared to the existing literature are summarized in Table 2.1. Besides the technical difference, this study takes into account the transit ridership of vulnerable groups, making it a different unexplored domain for ML vs. statistical model comparison. Therefore, this study aims to fill these gaps methodologically.

3. STUDY AREA AND DATA

3.1 Chapter Overview

This chapter provides an overview of the study context and the data used for our analyses. Section 3.2 provides a brief introduction to the study area in which our research undertake. Section 3.3 describes the dataset used for this study. Finally, Section 3.4 describes the statistic summary of the subsample used for this study.

3.2 Study Context

This study takes place in a contemporary Canadian context, the Greater Toronto and Hamilton Area (GTHA). It is the largest urban agglomeration in Canada, with a population of more than 7 million people based on 2021 population estimates (Statistics Canada, 2021a) and one of the fastest-growing regions in North America. The GTHA contains the cities of Toronto and Hamilton and four regional municipalities including, Durham, York, Peel, and Halton. The city of Toronto, with approximately 3 million residents is the most populated urban region. Both York and Peel regions include more than 1 million people, while the remaining municipal regions have less than 1 million residents. Metrolinx, the Greater Toronto Transportation Authority, is responsible for developing Regional Transportation Plans (RTP), operating commuter rail, and more recently, planning for all new heavy rail development (e.g., LRT, BRT, and subways) in the GTHA (see Figure 3.1). The overall public transportation system in the GTHA comprises nine local transit agencies, operating subway lines, surface buses, and streetcar routes, together with regional bus and rail lines. Toronto's transit system operated by the Toronto Transit Commission (TTC) includes four subway lines, surface bus, and streetcar routes. A regional rail and commuter bus system, GO Transit (operated by Metrolinx), connects suburban regions to themselves and the city center, and is used primarily for long-distance commutes. Other municipalities in the region primarily offer local bus services, with some lite BRT functionality along select corridors.

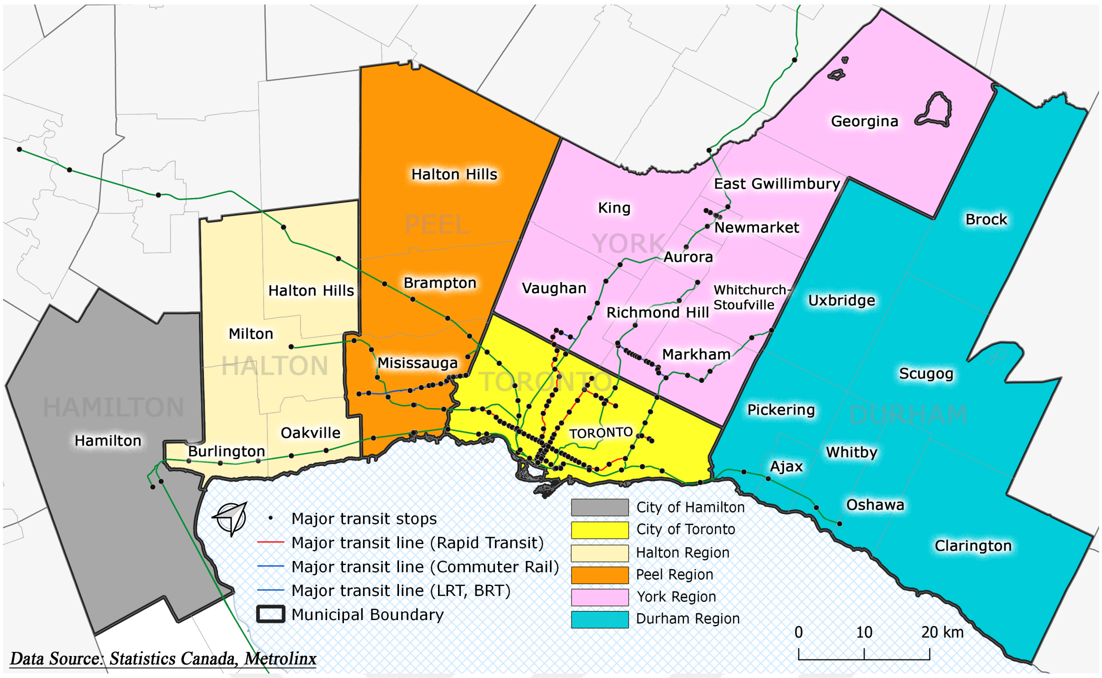


Figure 3.1 : Study Area (The Greater Toronto and Hamilton Area).

3.2.1 The population density

Despite overall economic growth in Canada, its major cities have encountered growing socio-spatial inequalities, with increasing polarization as neighbourhoods change over time (Ades, Apparicio, & Séguin, 2012; Hulchanski, 2010). Toronto is now the most unevenly distributed metro area in Canada, according to the Gini coefficient for income [1], and within the Toronto area, inequalities between neighbourhoods are very high (Dinca-Panaitescu et al., 2017). Figure 3.2 displays the population density patterns of our study area for different income and car-ownership levels together with the population of each stratum. In the bottom row of Figure 3.2, the 'U' shape pattern illustrates the higher concentrations of low-income households (<$40k) in downtown Toronto and its inner suburbs. Regarding households with zero vehicles per adult (VA=0), as income increases, they become more concentrated in downtown Toronto, whereas the low-income carless cohort is more dispersed in the region.

[1]Gini coefficient or Gini index is a measure of inequality to illustrate the wealth distribution within society. It ranges from 0 to 1. The high Gini coefficient for a country means that the gap between the income levels of the poor and the affluent is high.

Previous studies in the GTHA have shown that vulnerable groups have, on average, shorter transit travel times for their work commutes, and higher levels of accessibility than their counterparts (El-Geneidy et al., 2016; Foth et al., 2013). This can largely be explained by a) the vestiges of a sizable inner-city low-income population, and b) the proliferation of middle and upper-income households throughout the vast and poorly served outer suburbs of the region. Despite this overall distribution, recent work shows that there are hundreds of thousands of low-income households located in low-accessibility parts of the GTHA (Allen & Farber, 2019).

3.2.2 Transit ridership

Given the concentration of rapid transit within the City of Toronto, Figure 3.3 illustrates that most transit riders are living within the Toronto municipal boundaries, with concentrations closely mirroring both transit levels of service as well as the "U" shaped pattern of low-income brackets. In between the dense urban core and the poorly served outer suburbs, lies a transition zone characterized by people still having access to moderate levels of public transit, largely aligned with the service area of the TTC within the City of Toronto; consequently, 20-40% of residents keep transit in their daily trip basket in this zone. Furthermore, there is a decline in transit use in the city centre, mainly due to the availability of biking and walking options for reaching destinations despite the high level of transit accessibility.

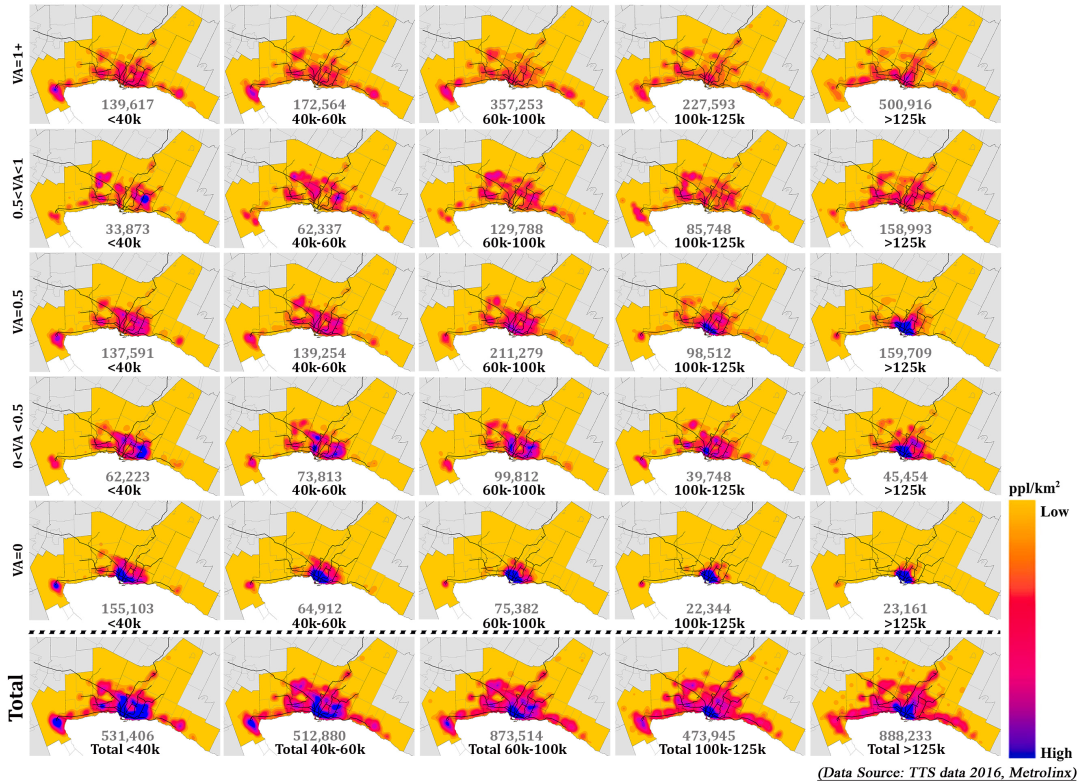


Figure 3.2 : The population density of different household income and car-ownership classes (missing data are excluded). The number of people in each class is normalized by the total population of the same class and the total number of people in each census tract to eliminate the size effect. The color indicates the normalized persons per square kilometer. VA = Vehicles per adult in the household.

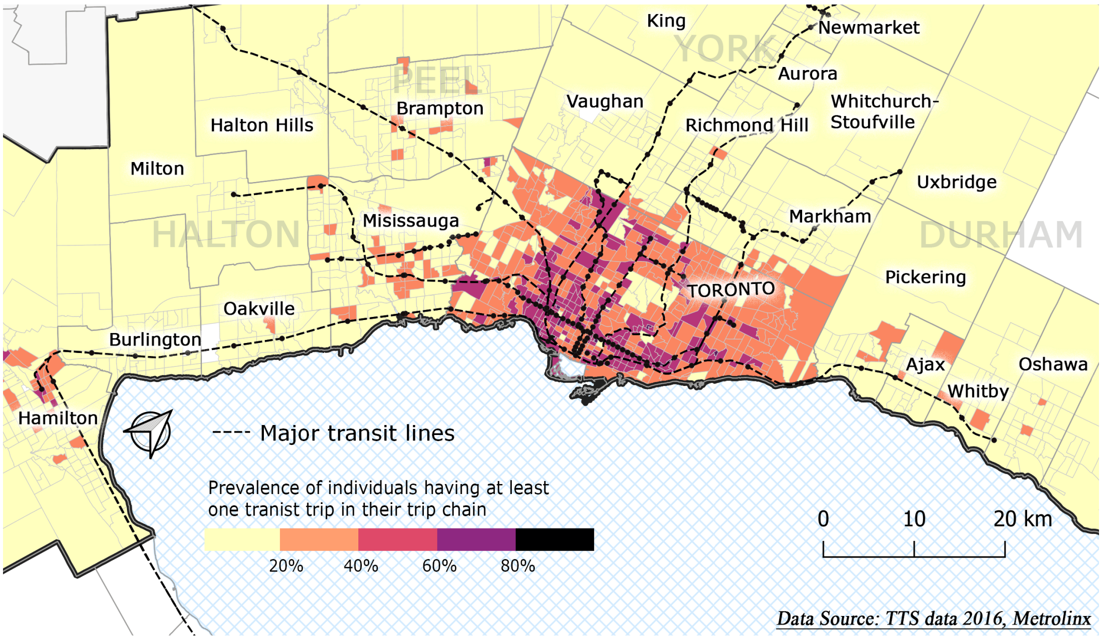


Figure 3.3 : Percentage of individuals taking at least one transit trip in their daily trip chain in the GTHA.

3.2.3 Housing price

Figure 3.4 displays the whole region and the average house price of all properties in 2016 using the data from RPS Real Property Solutions Inc. According to this data, the Toronto central regions along the transit network have the highest average home price. Following the north-south line of the transit network, several parts of the York Region, including Richmond Hill and Vaughan, observe a higher mean of home prices. Therefore, homes adjacent to the main transit networks are not affordable in GTHA. It forces low-income households to relocate to the neighbourhoods with affordable houses but less transit accessibility.

3.3 Transportation Tomorrow Survey (TTS)

The data source used for this study comes from the 2016 Transportation Tomorrow Survey (TTS). This travel survey is an ongoing data collection program started in 1986 and conducted every five years[2]. It is a large-sample, one-day household travel diary of people in the Greater Toronto Area and Hamilton (GTHA) plus adjacent regional municipalities. One available household member reports all trips taken by the entire

[2]Due to the pandemic, the survey in 2021 is postponed to 2023. By the time the dissertation is submitted, 2016 TTS data is the latest available survey in the region.

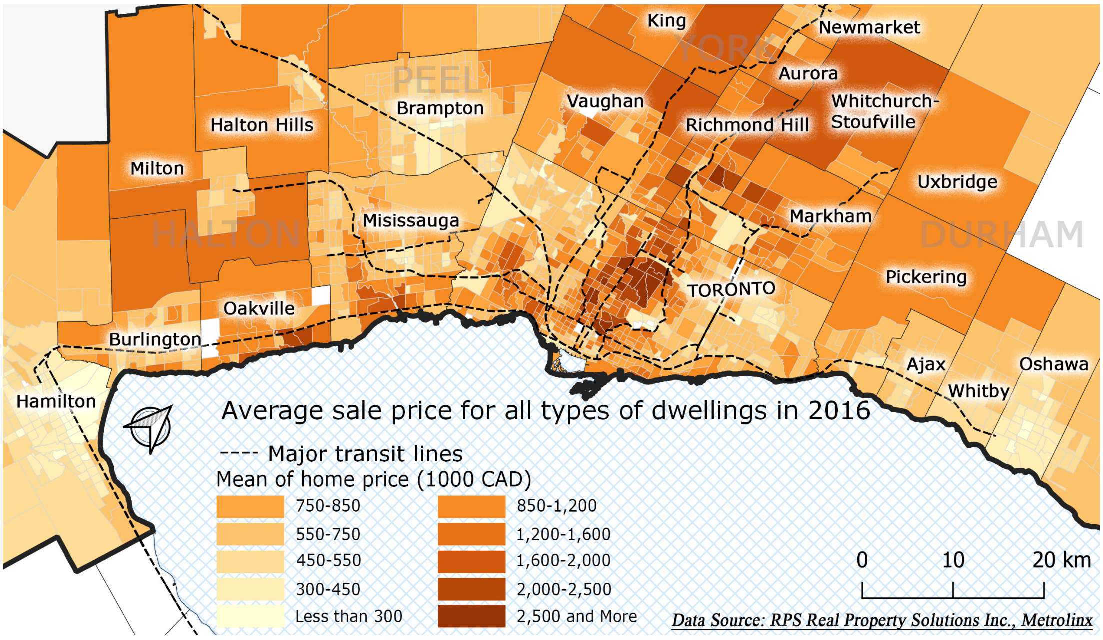


Figure 3.4 : Average house price for all types of dwellings in the GTHA.

household. This survey collects the personal (e.g., age, gender, having driving license, employment status, etc.), household (e.g., location, income level, number of vehicles, etc.), and trip (e.g., trip origin, trip destination, prime mode, etc.) information of each person 11 years and older in the household. For the 2016 survey cycle, the targeted sampling rate was 5 percent of households, except Hamilton, with only a 3 percent sampling rate target due to insufficient local funding. Notably, there might be sampling bias in the survey itself, which is partially addressed by weighting adjustments. By correcting for representation by dwelling type, family size, age, and gender, the data expansion procedure may better reflect additional variables (car ownership or employment status). The data is utilized with the caveat that there might still be additional aspects that cannot be found or corrected using the expansion factor. Further information on the households' total incomes was first recorded in this survey in 2016. All the survey data currently is under the supervision of the Data Management Group.

To provide a reliable population estimation, a set of expansion factors is included in the TTS 2016 data. In the 2016 survey, the data expansion process consists of data weighting method, considering dwelling type, household size, and the distribution of the population by age and gender matching the population distributions of the 2016 Canadian Census (Data Management Group, 2017). The anonymized, individual, and trip-level data is used for this study. In this research, only a subset of individual trips, those starting from and terminating in GTHA regions considers. Furthermore, the dataset is limited to adults aged 18 years or older –i.e., working-aged people assumed to be making autonomous residential location, vehicle ownership, and daily travel decisions.

3.4 Descriptive Summary

To undertake this study, the dataset is prepared by removing incomplete data and filtering out adults living outside the GTHA. According to the 2016 TTS data, there are six categories for household's total income level: "$0 to $14,999, $15,000 to $39,999, $40,000 to $59,999, $60,000 to $99,999, $100,000 to $124,999, $125,000 and above". This detailed information allows having an equity and poverty-related study. The low-income cut-off (LICO) is the income-related measure defined by Statistic Canada. It is calculated based on family size and community size and represents the poverty line for households that spend more than 20% of their income on basic needs such as food, shelter, etc. If the household's income level is below the LICO measure, the household considers a low-income household. According to this measure, a household in urban regions with more than 500k residents is a low-income household if its income level is less than 40k in 2016 (Statistics Canada, 2021b). Therefore, the two first income categories merge in this study to define low-income households with income levels less than $39,999. Moreover, high-income households are determined as those with income levels greater than $125k. Notably, this amount is the highest category in the TTS, although it may not represent the high-income level for the whole GTHA. The remaining categories are identified as middle-income families.

3.4.1 General overview

Table 3.1provides a descriptive summary of the dataset, including a total of 122,724 households (249,632 individuals) aged 18 years and older and a total of 538,364 trips. These figures are expandable to 5,387,081 people, 2,532,632 households, and 11,610,043 trips, respectively. A uniform age and gender distribution in the study area are observed. More than 50 percent of responders indicated a household salary of greater than $60k. In this work, families with a household income below $40k are considered the low-income group. Moreover, it can be seen that the region is car-dominant since more than 40% of households own their private vehicles. Interestingly, Table 3.1 shows that one-person households include only 11.5% of the dataset,different from other developed countries showing a high trend toward one-person households[3]. Even compared to the report made by the United Nations Economic Commission for Europe for Canada[4], which shows 28.2% of people are living alone, the GTHA has less than half of the country's average. Regarding employment, a low rate of people working from home is observed; however, this employment status should have changed after the pandemic and increasing teleworking. As the driver's licence is mainly considered the formal ID, 82% of people own one. Finally, all the aforementioned independent attributes are checked for possible multicollinearity, and no significant correlation among them is observed.

3.4.2 Transit accessibility measurement

Researchers and policymakers use various accessibility assessment approach for evaluating equity in the transportation domain. The well-known definition of accessibility is "a measurement of the spatial distribution of an activity" (Hansen, 1959, p. 04). The widely used accessibility measurement is location-based accessibility that demonstrates the level of access to spatially distributed activities by different travel modes (car, public transport, etc.). It defines the number of reachable activities such as jobs, schools, shopping centers, and health services for different groups of people within a certain travel time threshold. Access to jobs is a commonly used measure of transit benefits and can be a crucial predictor of travel behaviour (Allen & Farber, 2020b; Foth et al., 2013; Sanchez, Shen, & Peng, 2004; Tyndall, 2017). In this study, job accessibility is used as a proxy for overall transit benefits, which is defendable given the high degree of correlation between access to jobs via transit, and access to other daily destination types.

Gravity-based accessibility to jobs by transit, calculated in a recent analysis in the GTHA (Allen & Farber, 2019), is used in this study. This measure estimates the total number of reachable jobs from each Dissemination Area (origin), a census geographical unit with a population of 400 to 700 persons. The gravity-based accessibility is

[3]https://ourworldindata.org/grapher/one-person-households

[4]https://w3.unece.org/PXWeb/en/Table?IndicatorCode=318

computed as

$$A_i = \sum_{j=1}^{J} O_j f(t_{ij}) \tag{3.1}$$

where A_i is the accessibility measure in zone i, O_j is the count of jobs found in census tract j, $f(t_{ij})$ is the impedance function used to operationalize the diminishing

Table 3.1 : Descriptive statistics of explanatory variables for respondents ($n = 249,632$; expandable to $\mathcal{N} = 5,387,081$).

Variables	Individuals in the GTHA	
	Expanded frequency	**Expanded Proportion**
Age group		
18-25	702,598	13.04%
26-35	964,695	17.91%
36-45	955,252	17.73%
46-55	1,057,659	19.63%
56-65	841,574	15.62%
65+	853,731	15.86%
Missing	11,572	0.21%
Gender		
Female	2,806,404	52.10%
Male	2,580,677	47.90%
Household's total income per year		
\$0 to \$39,999	838,021	15.56%
\$40,000 to \$59,999	713,772	13.25%
\$60,000 to \$99,999	1,148,963	21.33%
\$100,000 to \$124,999	598,691	11.11%
\$125,000 and above	1,089,156	20.22%
Missing	998,478	18.53%
Vehicles per adult		
(VA=0)	597,833	11.10%
(0<VA<0.5)	610,964	11.34%
(VA=0.5)	1,243,951	23.10%
(0.5<VA<1)	835,877	15.51%
(VA=1+)	2,097,184	38.93%
Missing	1,272	0.02%
Household size		
One-person	622,417	11.55%
Two-people	1,419,702	26.35%
Three-people	1,088,185	20.20%
Four-people	1,199,803	22.27%
Five or more people	1,056,974	19.62%
Employment status		
Full time employment	2,694,603	50.02%
Part time employment	545,006	10.12%
Work at home (full time or part time)	255,900	4.75%
Not employed (including students)	1,888,816	35.06%
Missing	2,756	0.05%
Possession of a driver's license		
Having driver's license	4,423,009	82.10%
Not having driver's license	858,401	15.93%
Missing	105,671	1.96%
	Mean	**SD**
Population Density (per person)	6,864	7,924
Business Density (per person)	700	1,422
Intersection Density (per person)	54	36

attraction of jobs with travel time, and t_{ij} is the travel time between i and j estimated with OpenTripPlanner using GTFS and OpenStreetMap data as inputs. This travel time includes walking time to and from stops, waiting time for the transit vehicle, in-vehicle travel time by transit, and transferring time. The impedance function is defined as

$$f(t_{ij}) = \begin{cases} 180(90+t_{ij})^{-1} - 1 & t_{ij} < 90 \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

giving a weight of 0.5 to a 30-minute trip, roughly equal to the median duration commute trip in the GTHA across all modes. In this function, the maximum travel time value is limited to 90 minutes since very few people travel to jobs more than 90-minutes away (Allen & Farber, 2020a).

Figure 3.5 shows average gravity-based accessibility to jobs by transit in the GTHA at the Census Tract (CT) level. For this map, the measure is normalized by 100k reachable jobs and it illustrates the number of reachable jobs in the region. High levels of transit accessibility belong to places downtown Toronto and around transit lines, while suburban neighbourhoods have poor transit networks and low levels of transit accessibility.

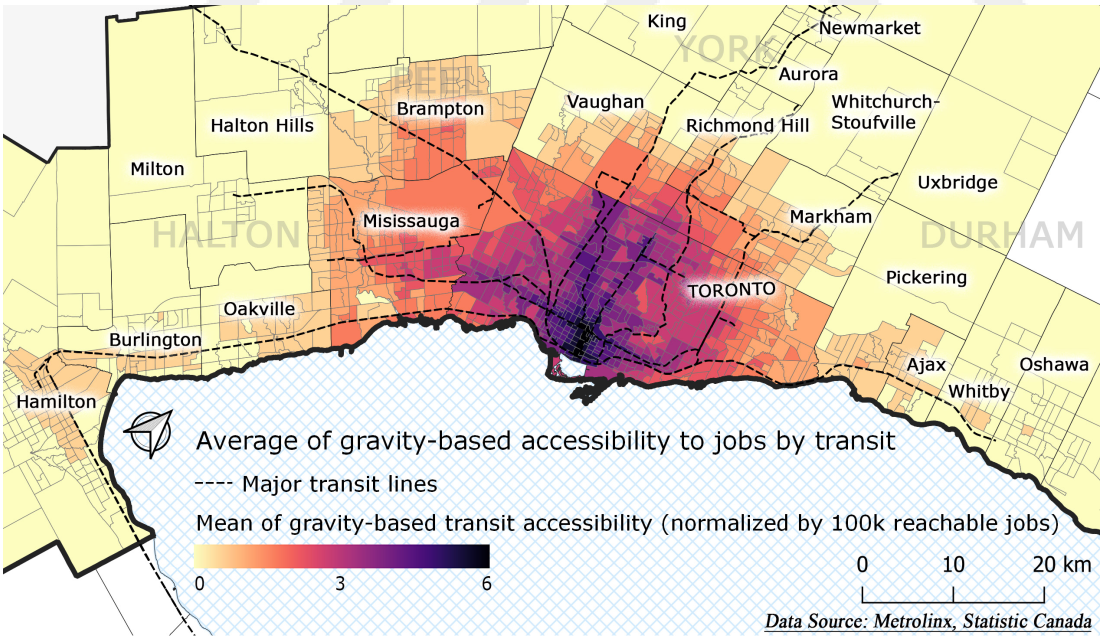


Figure 3.5 : Average of gravity-based accessibility to jobs by transit in the GTHA.

3.4.3 Low-income vs. high-income

3.4.3.1 Socioeconomic and built environment variables

A subset of the whole dataset is used to explore the differences in the travel behaviour of low-income and high-income households. Table 3.2 provides summary statistics for the socioeconomic characteristics of low-income and high-income carless and car-owners as well as their built environment information. Notably, all figures provided in the table are expanded. This dataset includes 65,458 individuals, expandable to 1,419,640 people. People with missing income or vehicle data, no trips, or an extraordinary number of trips (i.e., greater than 25 trips) are excluded. Among all subsamples, the high-income carless dataset is the smallest group having 23,161 individuals, whereas high-income car-owners are the most populated subgroup. The table shows that high-income carless households are, on average, younger than high-income car-owners. Expectedly, this younger stratum has the lowest unemployment rate (i.e., 93% of them are full-time or part-time employees) since they are of working age.

Most poor carless individuals are female, whereas there is a relative balance between males and females in other subsamples. Moreover, the larger share of low-income carless families includes women in one-person households. Comparing low-income and high-income carless households, the results show that 68.8% of high-income households have transit passes, the highest rate among all strata. In contrast, low-income carless households that are more transit-dependent less likely to possess transit pass. The findings reveal that high-income carless families live in neighbourhoods with higher transit accessibility and business density (e.g., downtown and around transit lines) than their counterparts. Therefore, their public transit use is justifiable as it is easily accessible in these regions. Notably, houses located in more accessible areas are not affordable for disadvantaged households (see Figure 3.4). The descriptive analysis shows that low-income and high-income car-owners mainly live in more remote neighbourhoods where population and business density are low, and transit accessibility is not appropriate (e.g., suburbs). In other words, low-income households prefer or are forced to locate in remote areas, which have lower housing prices, at the cost of owning private cars to compensate for reduced transit accessibility (see Figure 3.5).

Table 3.2 : The descriptive summary of explanatory variables for carless and car-owner households with income less than $\$40k$ and greater than $\$125k$ ($n = 65,458; \mathcal{N} = 1,419,640$).

Samples	Low-income Carless	Low-income Car-owner	High-income Carless	High-income Car-owner
Normal (n)	5,354	16,859	1,267	41,978
Expanded ($\mathcal{N}$)	155,103	376,303	23,161	865,073
Variables	**Proportion**		**Proportion**	
Individual attributes				
Age group				
18-25	21.4%	14.3%	5.8%	10.6%
26-35	17.0%	12.8%	43.0%	16.9%
36-45	12.1%	16.8%	28.1%	24.1%
46-55	13.6%	17.3%	11.1%	26.2%
56-65	13.9%	14.9%	7.9%	16.3%
65+	21.9%	24.0%	4.2%	5.8%
Gender				
Female	60.5%	52.1%	44.6%	47.9%
Male	39.5%	47.9%	55.4%	52.1%
Household size				
One-person	44.1%	19.6%	19.6%	2.4%
Two-people	29.1%	26.8%	54.5%	22.5%
Three-people	14.1%	18.1%	17.7%	23.2%
Four-people	8.0%	17.6%	5.6%	31.7%
Five or more people	4.6%	17.9%	2.7%	20.3%
Employment status				
Full-time & part-time employee	48.6%	53.3%	93.0%	86.8%
Unemployment	51.4%	46.7%	7.0%	13.2%
Having transit pass				
Yes	57.1%	18.5%	68.8%	23.8%
No	42.9%	81.5%	31.2%	76.2%
Having driving license				
Yes	44.7%	87.5%	79.4%	96.0%
No	55.3%	12.5%	20.6%	4.0%
	Sample $\mu \pm \sigma$		**Sample** $\mu \pm \sigma$	
Built environment attributes				
Measure of accessibility to jobs using a gravity function§	3.0±1.5	1.8±1.3	4.5±0.9	1.8±1.6
Population density (per person)‡	1.9±1.8	1.0±1.1	2.6±1.8	0.8±1.0
Business density (per person)‡	1.6±2.6	0.7±1.2	4.3±4.4	0.9±2.0
Intersection density (per person)‡	1.2±0.8	0.9±0.6	1.8±0.9	1.0±0.7

§ Gravity-based accessibility to jobs by transit estimates the total number of reachable jobs found in census tract from each dissemination area.

‡ Local built environment characteristics of travelers come from the weighted sum of values normalized by area in each dissemination area. These values are further divided by the mean density of all individuals.

3.4.3.2 Trip information

All trips of individuals are composed of various destinations and relevant travel modes. Table 3.3 presents trip purpose frequencies of low-income and high-income households. High-income carless households, mostly of work age, have work trips as the most frequent destination (53%) and shopping and school trips as the least frequent trips (10% and 1%, respectively) among all other subsamples. Thus, their higher employment rate confirms that they generally commute to work and are probably inclined to online shopping.

Table 3.3 : The expanded frequency of trip purposes for low- and high-income households in their daily trips.§

Symbol	Description	Low-income				High-income			
		Carless		Car-owner		Carless		Car-owner	
H	Home*	17,797	(8%)	80,813	(12%)	3,249	(8%)	248,324	(14%)
W	Work	63,153	(28%)	171,764	(25%)	20,354	(53%)	705,664	(40%)
M	Marketing/Shopping	47,618	(21%)	137,481	(20%)	3,841	(10%)	186,870	(11%)
S	School	25,761	(11%)	41,704	(6%)	463	(1%)	38,383	(2%)
F	Picking up or dropping off someone (Facilitator passenger)	3,294	(1%)	91,497	(13%)	475	(1%)	212,879	(12%)
D	Taking kids to daycare	3,435	(2%)	12,009	(2%)	1,464	(4%)	62,426	(4%)
O	Other discretionary activities	64,854	(29%)	146,167	(21%)	8,705	(23%)	302,263	(17%)

§ Trip chains which are not started and ended at home are removed.

* Home is excluded if it is at the start or end of the trip chain.

Table 3.4 summarizes mode share of each trip segment in our dataset. It illustrates that low-income carless more frequently use public transit for their daily trips than other categories. The result is consistent with the literature that the most socially disadvantaged households are more transit-dependent in their daily trips than others (Pucher & Renne, 2003). Low-income car-owners extensively use a private car as their prime travel mode compared to low-income carless households, probably to justify their car-ownership costs. Interestingly, high-income carless households make 48% of their daily trips by public transit and 42% of those by walking or cycling as they afford to live in regions with much higher transit accessibility (see Table 3.2 and Figure 3.5). This scenario confirms that the higher levels of walking and cycling access provide more opportunities for local active trips. Considering both high- and low-income car-owners, the results show that although driving a car and taking transit are their most frequent travel mode, low-income households also use a car as a passenger more often than their counterparts (13% vs. 8%). It may be an indication of more car-sharing among low-income families. Other than this observation, car owners of different income levels have almost identical travel mode decisions.

Table 3.4 : The expanded frequency of the travel modes for low- and high-income households in their daily trips.§

Symbol	Description	Low-income				High-income			
		Carless		Car-owner		Carless		Car-owner	
C	Car as a driver	2,962	(1%)	729,314	(70%)	1,028	(2%)	1,950,778	(75%)
K	Car as a passenger	29,545	(8%)	135,020	(13%)	1,411	(2%)	196,364	(8%)
P	Public transit and Go rail	241,054	(64%)	119,938	(11%)	28,933	(48%)	289,711	(11%)
A	Cycling and Walking (active transport)	86,729	(23%)	50,372	(5%)	25,652	(42%)	129,904	(5%)
T	Taxi and paid rideshare	11,882	(3%)	4,946	(0%)	3,610	(6%)	15,565	(1%)

§ Trip chains which are not started and ended at home are removed.

3.4.4 Transit use, income, car ownership, and accessibility levels

Figure 3.6 demonstrates a cross-tabulation of income, vehicle ownership, and the percentage of individuals with at least one transit trip per day. While carless households overall have very high rates of transit use, the rates are highest among low-income households (<$40k) at 73.3% vs. 60.9% for the wealthier households ($125k). Moreover, while there are more than 155,000 people living in carless low-income households, there are only 23,000 in carless wealthier households[5]. Notwithstanding the potential for wealthier carless households to travel by taxi and ridehailing more easily than low-income carless counterparts, the transit-use gap between incomes is also informed by the maps in Figure 3.2, showing that high-income carless households are extremely concentrated in the core of the city, with added ability to either walk or bike. Conversely, low-income carless households are dispersed into the inner suburbs, where there are more barriers to active travel.

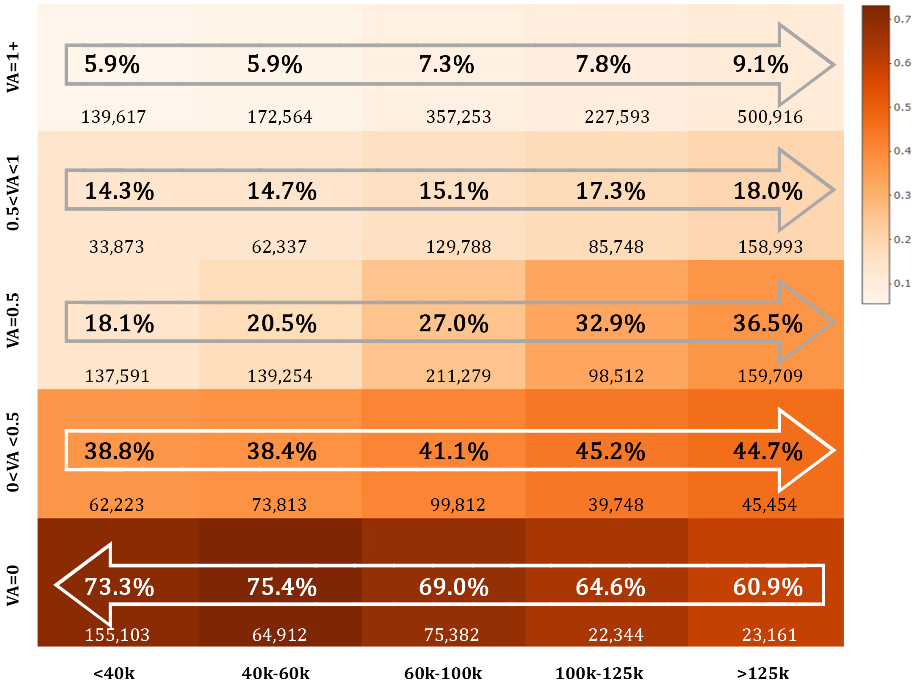


Figure 3.6 : The percentage of individuals using transit in each class (people with missing income data and no trips are excluded).

Interestingly, for car-owners, Figure 3.6 shows that the percentage of people that use transit tends to increase with income. Conversely, among carless households, transit use tends to decline with income, likely related to residential concentration of the

[5] It is important to note that the TTS income categories are designed to give granularity at the lower-end of the income scale, with nearly a full quarter of households earning higher than $125,000.

carless wealthy, and the relative affordability of taxis. Moreover, it shows a much steeper drop-off in transit use when moving from 0 cars to 0.5 cars per adult among the low-income households compared to the wealthy. At face value, these statistics support the hypothesis that when low-income households purchase a car, given its large expense relative to income, they do so with intentions to use it fully. A study by Giuliano (2005) declares that the poor own a car since it is their only solution for household maintenance and income earning. Also, for many trips in the GTHA, the marginal costs by car are far cheaper than for transit, making the use of a car a cost-saving decision, except for trips heading to downtown Toronto along rapid transit corridor.

Figure 3.7 shows the relationship between accessibility, income, and car ownership levels in the study area. It illustrates that carless households, regardless of income level, tend to reside in neighbourhoods with higher levels of accessibility. Moreover, increasing income corresponds with an increase in transit accessibility for zero-car households. It suggests that high-income carless (and car-deficit, i.e., households with less than a car per driver) residents afford to locate in places with higher levels of transit accessibility.

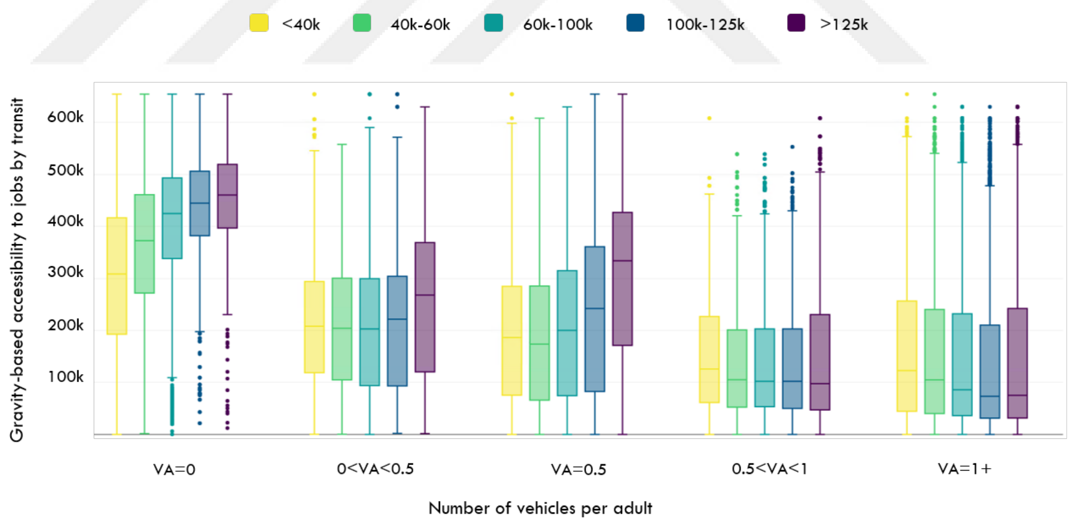


Figure 3.7 : Distribution of accessibility, income and car ownership for all households.

Furthermore, it indicates that as the number of private vehicles increases, households tend to locate further from the core, in car-dependent neighbourhoods where transit accessibility levels are far lower, and the differences in accessibility are far less pronounced across income groups. The findings are consistent with the maps shown in Figure 3.2.

4. METHODS

4.1 Chapter Overview

This chapter describes the method and algorithms used for the analysis in detail. Section 4.2 introduces generating the sequential trip segments for exploring the behavioural differences amongst residents according to their trip purpose and mode used in GTHA. Afterward, it discusses the hierarchical clustering method and distance matrix. Section 4.3 provides an overview of statistical and ML algorithms to compare them, and their predictive performance and interpretability are investigated in Section 4.4.

4.2 Trip Chain Analysis

This section explains the preparing process of the dataset and the structure of the model to cluster people based on their daily trip chain.

4.2.1 Trip sequence generation

To evaluate residents' daily activity patterns, first, a set of sequential trip segments, including trip destination and travel mode is created. The raw dataset has a series of end-to-end trips with their prime travel mode for each individual in a household. Therefore, all trip segments of an individual with their relevant travel mode is combined to create their complete daily trip chain. The notation used for each trip purpose and transport mode is based on the symbols shown in Table 3.3 and Table 3.4. Figure 4.1 shows an example of trip chains of individuals, where the travel mode is in red, and the origin/destination is in black. The trip chains of the first and the third individuals include two travel modes; however, the trip chains of the second and fourth individuals have a single mode in their daily trips. For instance, the trip chain of the first person includes two travel modes, i.e., car as a passenger (`K`) and public transit (`P`).

1) HKOPWPH
2) HKWKH
3) HTWAMAWPH
4) HCDCWCH

Figure 4.1 : Sample trip chains with their prime mode.

After constructing individuals' trip chains based on their travel diary, the trip chains are broken into three particular subchains, namely, mode chains, destination chains, and destination-mode chains. First, each individual's travel modes are broken into unigrams (i.e., a contiguous set of modes). For instance, for the first person in Figure 4.1, the set of modes is $\mathcal{M}_1 = \{\texttt{K},\texttt{P},\texttt{P}\}$. Second, the trip legs of an individual based on their activity chains are broken into bigrams (e.g., $\mathcal{S}_1 = \{\texttt{HO},\texttt{OW},\texttt{WH}\}$ for the first person). This way, the users' transition among different destinations is captured. Finally, the trip chains split into trigram of the origin, travel mode, and destination. With a gap of 2 and the window size of 3, the complete trip chain of $\mathcal{T}_1 = \{\texttt{HKO},\texttt{OPW},\texttt{WPH}\}$ is obtained for the first individual. A gap of 2 for the moving trigram is required since the origin and destination should be on both sides while the travel mode remains in the middle (see Figure 4.2).

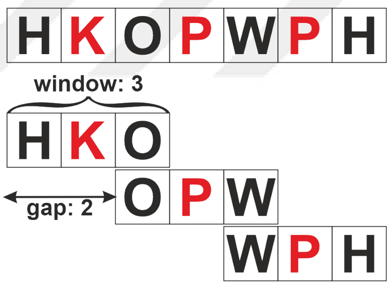


Figure 4.2 : An example of a trigram with a gap of 2 to generate a set of the complete trip chain.

How is the similarity of individuals determined using trip purposes, transport modes, and sequences of activities-modes? The three sets of sequences is generated for the first two individuals as follows.

$$\mathcal{S}_1 = \{\texttt{HO},\texttt{OW},\texttt{WH}\} \qquad \mathcal{S}_2 = \{\texttt{HW},\texttt{WH}\}$$

$$\mathcal{M}_1 = \{\texttt{K},\texttt{P},\texttt{P}\} \qquad \mathcal{M}_2 = \{\texttt{K},\texttt{K}\}$$

$$\mathcal{T}_1 = \{\texttt{HKO},\texttt{OPW},\texttt{WPH}\} \qquad \mathcal{T}_2 = \{\texttt{HKW},\texttt{WKH}\}$$

The first one starts the trip from home (H), makes a discretionary trip (O) using a car as a passenger (K), then goes to a workplace (W) by public transit (P), and finally returns home by transit (`P`). On the other hand, the second individual leaves home (`H`) to work (`W`) as a car passenger (`K`) and then returns home (`H`) with the same travel mode. Afterward, these textual representations are converted to a numeric vector of term frequencies. Counting the number of behaviour repetitions for each individual displays their travel characteristic compared to others. Table 4.1 shows the frequency table for the trip legs of each individual in Figure 4.3. The frequency tables is used to specify the similarity of individuals. Similarly, the frequency tables are generated for the travel modes ($\mathcal{M}$) and the complete trip chains ($\mathcal{T}$) of individuals. These three tabular frequency datasets are employed to determine the similarity of individuals' travel behaviours.

Table 4.1 : Frequency table of trip legs for each individual ($\mathcal{S}$).

Individual ID	Trip chain	HO	OW	WH	HW
1	**HOWH**	1	1	1	0
2	**HWH**	0	0	1	1

Measuring the similarity/distance is fundamental in pattern recognition and clustering task. To cluster people according to their travel behaviour, the pairwise dissimilarity between individuals' trip sequences is computed. Previous studies emphasized the particular use of the cosine similarity in the case of frequency-based matrices or sparse matrices mostly comprised of zero values (B. Li & Han, 2013; Sidorov, Gelbukh, Gómez-Adorno, & Pinto, 2014; M. Li, 2019; Jahanshahi & Baydogan, 2022). Cosine similarity, a popular metric in sequence mining, is the cosine of the angle of two vectors in n dimensions. It is a recommended method for normalizing the length of vectors (A. Huang, 2008; Murthy, 2012). The cosine distance of two numeric vectors a and b associated with individuals $\mathcal{I}_1$ and $\mathcal{I}_2$ is computed as

$$\mathrm{Dis}_{\cos}(a,b) = 1 - \frac{a \cdot b}{\|a\|\|b\|} = 1 - \frac{\sum_{i=1}^{n} \mathrm{a}_i \mathrm{b}_i}{\sqrt{\sum_{i=1}^{n} \mathrm{a}_i^2}\sqrt{\sum_{i=1}^{n} \mathrm{b}_i^2}} \tag{4.1}$$

where $a \cdot b$ is the dot product of vectors a and b, $\|.\|$ is the magnitude of a vector, and n is the length of each vector. The output of the cosine similarity metric is within the range of 0 and 1, where 0 indicates a considerable dissimilarity, and 1 represents the highest similarity between two vectors. The cosine similarity values is subtracted from 1 to generate the cosine distance for clustering purposes. Accordingly, the pairwise distance matrix is constructed for all individuals, then input this distance matrix into any clustering algorithm to generate distinct clusters.

The cosine similarity between two individuals should consider all sets of sequences shown in Figure 4.3. For instance, if only their complete trip chain consider, the similarity becomes zero. However, if the mode choice considers, their similarity becomes 0.45, and if activities/purposes consider, their similarity becomes 0.41 (see Figure 4.3).

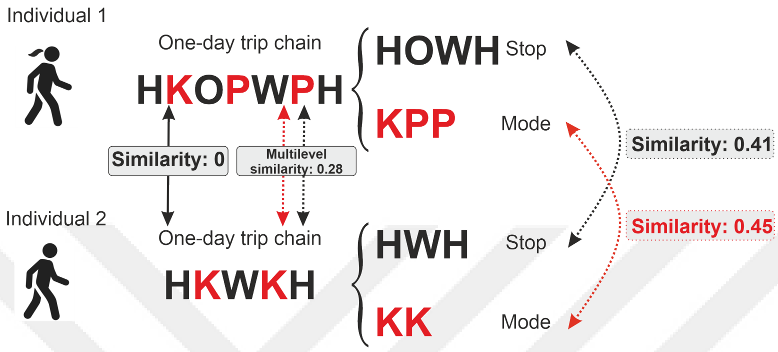


Figure 4.3 : Motivating example.

Therefore, in order to avoid simplistic rules in such a multifaceted problem, the average cosine similarity of all three vectors is used to decide the overall trip similarity of two individuals (s_{ij}) as follows.

$$s_{ij} = \frac{\mathrm{Sim}(\mathcal{S}_1,\mathcal{S}_2)+\mathrm{Sim}(\mathcal{M}_1,\mathcal{M}_2)+\mathrm{Sim}(\mathcal{T}_1,\mathcal{T}_2)}{3} \tag{4.2}$$

Equation 4.2 is a multilevel similarity of individuals and is assumption-free (i.e., no presumption is required to determine the prime mode or cluster). If the individuals' travel behaviour is clustered based on a simplified rule (e.g., transit vs. non-transit users), the importance of their destinations (i.e., trip purposes) probably is overlooked. In such rule-based approaches, the frequency of taking each travel mode or even the transition between modes, especially if a person is multi-modal is also disregarded. This study uses the cosine similarity accounting for the number of, diversity of, and transition between each trip leg/mode, better representing a person's travel behaviour.

4.2.2 Hierarchical clustering method

The clustering method is an unsupervised learning technique that investigates the dataset's structure by discovering and extracting its inherent patterns. Through clustering, the pairwise distances between a set of observations are measured, and then the dataset is partitioned into various subgroups (Sasirekha & Baby, 2013; Saxena et al., 2017). Observations with the highest similarities are assigned to a subgroup called a "cluster". As a result, each cluster will be composed of homogeneous observations but dissimilar from others. In other words, clusters seek to have the lowest intra-cluster distances and the highest inter-cluster distances (Sasirekha & Baby, 2013; Murtagh & Legendre, 2014).

In this study, an agglomerative hierarchical clustering algorithm, one of the most widely used methods in data analytics is utilized. It divides our high-dimensional dataset into homogenous clusters with similar travel patterns. The agglomerative hierarchical clustering approach is a bottom-up method in which cluster hierarchies are built by partitioning the individual data points into subclusters from the bottom and then merging subclusters with similar patterns into a high-level cluster at the top. Since the individuals' behaviour rather than global behaviour considers, the agglomerative approach is preferable to the divisive one. As recommended by Murtagh and Legendre (2014), the Ward's method is utilized in this study to minimize the total within-group variance. The cosine distance of individuals' trip chains, explained in Section 4.2.1, is used as an input for the agglomerative hierarchical clustering. After applying the clustering algorithm, the dataset is partitioned into four distinct clusters based on average silhouette width metric. Average Silhouette Width (ASW) is a metric to estimate the optimal number of clusters by computing the inter and intra clusters' sum of squares.

Although the agglomerative hierarchical clustering algorithm is employed to cluster people, any other unsupervised algorithm (e.g., k-means) could have being used for the same purpose. However, k-means is not recommended for this task as it suffers from random initialization and requires preknowledge of the parameter k (Pena, Lozano, & Larranaga, 1999; Celebi & Kingravi, 2012). When the estimation of cluster numbers is not logical at the beginning of the procedure, the hierarchical algorithm is recommended (León, Mkrtchyan, Depaire, Ruan, & Vanhoof, 2014; Pizzol, Strambi, Giannotti, Arbex, & Alves, 2021).

4.3 Statistical and Machine Learning Methods

First, a brief discussion about the specifications of statistical and ML models is provided in this section. Then, various performance metrics used to compare these models are discussed.

4.3.1 Statistical models

Regression analysis, one of the widely used statistical modeling approaches, explores the dependency of a predictor and responses. It helps to understand a causality relationship between two or more continuous variables. LinR is one of the simplest regression models with a linearity assumption. It is applied when the dependent variable is continuous. In contrast, the Logistic Regression (LogR) applies for binary outcomes, and it investigates the association between a categorical dependent variable and other independent covariates.

One of the well-known count regression models is negative binomial regression, in which the dependent variable follows the negative binomial distribution. It can be utilized if the dataset has overdispersed count outcome — that is, the variance of the data is equal to or greater than its mean. Zero-Inflated Negative Binomial Regression (ZINB) and Hurdle model are examples of zero-inflated count models, particularly used in dealing with excessive zeros in the data. In the zero-inflated model, excessive zeros are divided into "structural" and "sampling" zeros. Sampling zeros come from the unusual Poisson or negative binomial distribution, assumed to be generated by chance. On the other hand, structural zeros are observed by non-risk groups who structurally are a source of zero (M.-C. Hu, Pavlicova, & Nunes, 2011; Hua, Wan, Wenjuan, & Paul, 2014). Either a mixture or a two-part modelling type, both ZINB and Hurdle models consist of two processes and deal with two types of distributions: zeros and counts. In the first process, a binomial model is utilized to estimate the probability of zeros versus non-zeros (Zuur, Ieno, Walker, Saveliev, & Smith, 2009; Cameron, Trivedi, Jackson, & Chesher, 1998). In this step, the zero-inflated model assumes zeros both as structural and sampling zeros, while the hurdle model assumes all zeros as structural ones then formulate a pure mixture of zero and positive (non-zero) models (M.-C. Hu et al., 2011; Hua et al., 2014). Therefore, their difference lies in the way they treat different types of zeros.

Another distinction between ZINB and the Hurdle model is in the second step, the count model portion. A truncated-at-zero count model is implemented for the count portion of the Hurdle model, while a negative binomial is used for the count portion of the zero-inflated model (Zuur et al., 2009). Naive Bayes (NB) one of the simplest learning algorithm uses Bayes' rules. All of the statistical models are derived based on assumptions about the data. Violation of these assumptions leads to inefficient and/or biased estimations. Conversely, there are no hypotheses or restrictive considerations in ML methods.

4.3.2 Supervised learning models

An ML algorithm as a computational process adjusts and improves its architecture through learning from the environment. This learning process stemmed from using and experiencing input data to achieve a required output. Since this training process constructs a fundamental part of this technique, most ML methods are classified based on their learning into two broad categories, *supervised* and *unsupervised learning* (El Naqa, Li, & Murphy, 2015). In *supervised learning* methods, the dataset labels are known, helping the learning process to predict the outcome of new, unseen data. Both classification and regression problems are classified under supervised learning as their labels are known and are categories or numeric values, respectively.

A DT consisting of branches, decision nodes, and terminal leaves comes from the recursively partitioned feature space of the training set. The purpose of these tree-like structures, such as CART (Breiman, Friedman, Stone, & Olshen, 1984), is to construct disjoint subnodes through a set of decision rules according to features. In a fully developed tree, this splitting process of the dataset iterates until all possible decision boundaries are tested and finally arrive at a terminal leaf, i.e., a homogeneous subnode. The impurity level of each decision node and the expected entropy reduction is computed to quantify the best split (Wang & Suen, 1984; El Naqa et al., 2015). The terminal node in DT classifiers is the probability of a class, whereas the numeric estimated value for the dependent feature is the terminal leaf of DT regressors. To avoid overfitting problems and improve the predictive performance of models, ensemble techniques such as bagging (Breiman, 1996) and boosting have been proposed.

RF is an ensemble ML algorithm that aggregates a collection of DTs with a random selection of features independent of previous attributes in each split (Breiman, 2001). Similar to all bagging models, the ultimate prediction result of RF is taken based on a majority vote of successive trees. Another tree-based ensemble model is Extreme Gradient Boosting (XGB), a scalable gradient tree boosting system (T. Chen & Guestrin, 2016). It constructs consecutive weak trees by incrementally adding a new DT to prevent overfitting issues and improve predictive performance for hard-to-predict instances. In a set of sequential trees, each tree is fitted on the residuals of the previous tree to minimize the loss of the last iteration (Friedman, 2001).

While tree-based algorithms are formed of branches and leaves, NN, considered a "black box" algorithm, consist of layers and neurons (nodes). NN models are triggered by feeding input data and going through the activation functions to estimate the output values (nodes) using the sum of weighted connections in hidden layers. The most widely used way of optimizing weights is the backpropagation method, in which the weights are iteratively updated to minimize the total loss.

SVM is a supervised learning model which can be divided into linear and non-linear models. To classify a linear dataset, a hyperplane (straight line) is determined to define a boundary with a maximum margin between two classes and separate the data points (Suthaharan, 2016). In a case that the training dataset cannot be separated into two-dimension, the SVM transforms the given data into a high-dimension feature space and searches for an optimal hyperplane using *kernel* functions. This hyperplane in the transformed space is the line close to support vectors —i.e., data points on the margin— with a maximum margin between those vectors (Cortes & Vapnik, 1995; Vapnik, 2013).

One of the simplest supervised learning classifiers is the Naive Bayes (NB) model that has a naive assumption about the training data. They assume that there is conditional independence between all features given the class. However, NB algorithm is implemented based on Bayes' rule and computes the probability estimations of each class with an acceptable accuracy (Lewis, 1998; McCallum & Nigam, 1998). This probabilistic approach can also be categorized in statistical models.

4.3.3 Cross-validation

An unbiased model evaluation process has two folds: training the model on a training set and evaluating it on an unseen dataset, called a test set. Accordingly, the dataset is split arbitrarily into two parts, namely the training and test sets. It mitigates the overfitting problem. Nevertheless, this random split can produce bias (El Naqa et al., 2015). The stratified k-fold cross-validation (CV) technique is used in the study to alleviate the bias issue. It is the most common approach in which the entire available dataset is divided into exclusive k subsets of almost equal size. The "stratified" CV is adopted in case imbalanced classes exist. Hence, the same ratio between non-transit and transit riders are maintained in the folds of our study; otherwise, the majority class, i.e., non-transit users, might be overrepresented in some folds. In this technique, the model iterates the training and validation sets k times where a subset is selected as a test set and the remaining ones as the training set. Each algorithm's performance is estimated using 10-fold cross-validation. In each iteration, a model is fitted on nine folds and test it on the remaining one. After ten iterations, ten independent performance scores for each model are recorded. Then, the average performance estimation during the CV is reported.

4.4 Performance Metrics

One of the major steps in model selection is evaluating the algorithm's performance. In predictive models, the estimation of predictive performance reflects how well the algorithm performs on unseen data. Therefore, selecting the best-performing model requires an approach to compare and rank the model's performance. In this study, multiple performance metrics are adopted to evaluate each technique. For classification problems, accuracy, precision, recall, F1-score, and the area under ROC (Receiver Operating Characteristics) curve are used. On the other hand, the performance of regressors is evaluated through R-Squared, Root Mean Squared Error (RMSE), Median Absolute Error (MedAE), and Root Relative Squared Error (RRSE) metrics.

4.4.1 Classifier's performance metrics

The most commonly used performance measure of classification models is accuracy. However, there are other performance metrics for classifiers such as precision, recall, F1-score, and the area under the ROC curve. In a binary classification task, labels are divided into two classes, $c \in \{0,1\}$. The prediction outcome is represented as true prediction of class c (T_c) and its false prediction (F_c). Accordingly, accuracy, precision, recall, and F1-score are computed as follows.

1. *Accuracy* of a classifier is the ratio of correctly classified observations of each class to the total observations as follows. It ranges from 0 to 1, where 1 is the highest accuracy.

$$\frac{\sum_{c\in\{0,1\}} T_c}{\sum_{c\in\{0,1\}} T_c + F_c} \tag{4.3}$$

2. *Precision* of class c is defined as the number of instances predicted as class c that belong to the same class, i.e., it measures how accurate the model is in predicting class c as follows. It ranges from 0 to 1, where 1 is the highest precision.

$$\frac{T_c}{T_c + F_c} \tag{4.4}$$

3. *Recall* of the class c represents how many data points that belong to class c are retrieved correctly, i.e., how much the model recalls the instances of the class c. It ranges from 0 to 1, where 1 indicates the highest recall.

$$\frac{T_c}{T_c + F_{c'}}, c \neq c' \tag{4.5}$$

4. *F1-Score* is the harmonic mean of precision and recall.

$$\frac{2 \times Precision \times Recall}{Precision + Recall} \tag{4.6}$$

5. *AUC-ROC* is a threshold-independent performance metric, whereas the previous ones assume that T_c or F_c is defined based on a threshold, typically 0.5. For instance, if the probability of belonging to class c is greater than 0.5, it is predicted as class c. However, the ROC curve is formed by the T_c rate against the F_c rate for all possible thresholds. Hence, the area under the ROC curve shows the

predictive model's performance comprehensively, thus alleviating the problem with the threshold settings (Hernández-Orallo, Flach, & Ferri, 2012). It ranges from 0 to 1, where 0.5 is the random guess and 1 shows the best performance.

4.4.2 Regressor's performance metrics

For evaluating the predictive quality of different regression models, the goodness-of-fit of each regressor is compared by measuring R-squared. Other performance metrics are calculated based on the loss functions of the predicted errors, such as root mean squared error, median absolute error, and root relative squared error. Unlike R-squared, the lower values for them are desirable.

1. *R-Squared (R^2)*, a widely used measure of goodness-of-fit for count-data models, is defined based on the proportion of variance in the dependent variable (y) predicted by independent variables (X). It ranges from 0 to 1, where $R^2 = 1$ associates with the best goodness-of-fit. It is computed as
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{RSS}{TSS} \tag{4.7}$$
where RSS is the residual sum of squares which means the sum of the squared difference between the observed and predicted values, while TSS is the total sum of squares.

2. *Root Mean Squared Error (RMSE)* is a frequently used metric for measuring the performance of regression models. This measure that shows the standard deviation of residuals is defined as
$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}}. \tag{4.8}$$

3. *Median Absolute Error (MedAE)* is the median of the absolute error of the predicted values. It is defined as
$$MedAE = \text{median}_{\forall i}(|y_i - \hat{y}_i|). \tag{4.9}$$

4. *Root Relative Squared Error (RRSE)* is the total squared error normalized by the total squared error of the predictor, shown as
$$RRSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \sqrt{\frac{RSS}{TSS}}. \tag{4.10}$$

4.5 Non-parametric Statistical Tests for Comparing Multiple Groups

After estimating the predictive performance of statistical and ML algorithms using various evaluation measures, the Friedman Aligned ranks test is used to statistically examine the significant difference in each performance among all algorithms. On top of that, Bergmann-Hommel *post hoc* analysis is employed to make a pairwise comparison between models (Derrac, García, Molina, & Herrera, 2011). The Friedman test is a non-parametric statistical analysis with the block design that uses the ranked values to perform a comparison between more than two models. This test is used to determine if there is a statistically significant difference between the prediction performance of at least two algorithms. In this study, the Friedman Aligned Ranks test, including the advanced ranking approach is employed. The method of aligned ranks is suggested when the size of our dataset is small (i.e., the data does not follow the normal distribution) (García, Fernández, Luengo, & Herrera, 2010; Derrac et al., 2011). Since for comparison study 10-fold cross-validation is used and only 10 values are compared, then a non-parametric test can be used. On the other hand, as the folds (datasets) remain the same for all iterations on different algorithms, a block design (i.e. a paired test) is used to compare the performance of the algorithms. The Friedman test facilitates such a pairwise comparison of the algorithms given each fold.

In this calculation, the average performance obtained from all algorithms in each fold for each metric is computed. Then, the difference between the performance score of each algorithm in each fold and the mean value of the same fold is calculated. Finally, these aligned scores, obtained through repeating this step for all folds and algorithms, are ranked from 1 to kn, associated with the best result and the worst one, respectively (García et al., 2010). These new ranks are called “aligned ranks”. The Friedman Aligned Ranks test statistic can be defined as

$$T = \frac{(k-1)\left[\sum_{j=1}^{k} \hat{R}_{.j}^{2} - (kn^2/4)(kn+1)^2\right]}{\{[kn(kn+1)(2kn+1)]/6\} - (1/k)\sum_{i=1}^{n} \hat{R}_{i.}^{2}} \tag{4.11}$$

where $\hat{R}_{.j}$ is the aligned rank total of the jth algorithms, k is the number of algorithms, $\hat{R}_{i.}$ is the aligned rank total of the ith fold, and n is the number of folds.

If the null hypothesis is rejected (i.e., the algorithms do not behave similarly and there is a significant difference between their performance), applying a *post hoc* analysis is neede to make a pairwise comparison. This pairwise comparison procedure detects which model performs better/worse than the others (Garcia & Herrera, 2008; García et al., 2010; Derrac et al., 2011). Accordingly, the Bergmann-Hommel *post hoc* applies as it is the best-performing procedure recommended by Derrac et al. (2011). In this study, the `scmamp` package is utilized for performing both the Friedman Aligned Ranks test and Bergmann-Hommel *post hoc* analysis in the `R` environment.

4.6 Spatial Efficiency Measure (SPAEF)

After sensitivity analysis of all algorithms in response to the transit improvement policy, the spatial distribution of newly generated transit trips estimated are mapped by all regressors. This study aims to compare the spatial patterns obtained from the estimated transit trips by each algorithm. Hence, a multi-component metric suggested by Demirel et al. (2018) is utilized

$$SPAEF = 1 - \sqrt{(\alpha-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{4.12}$$

where α is the Pearson correlation coefficient between map A and B (i.e., $\alpha = \rho(A,B)$), β is the ratio of coefficient of variations illustrating spatial variability (i.e., $\beta = \left(\frac{\sigma_A}{\mu_A}\right) / \left(\frac{\sigma_B}{\mu_B}\right)$), γ is the percentage of histogram intersection, and n is the number of bins (i.e., $\gamma = \frac{\sum_{j=1}^{n} min(K_j, L_j)}{\sum_{j=1}^{n} K_j}$). For γ, the histogram K of map A and the histogram L of map B are computed. For this study, the `Python` implementation of the SPAEF by the authors (github.com/cuneyd/spaef) is followed to implement it (Koch, Demirel, & Stisen, 2018).

4.7 Model Interpretability Tools

In addition to maximizing the prediction performance of a model, exploring which features affect the prediction outcome is essential for creating new knowledge about travel behaviour. Statistical models are categorized as intrinsically interpretable algorithms in which the coefficients of models readily reveal the significance and direction of each feature's impact on output. Conversely, it is generally assumed that ML models are "black-box" since their prediction results cannot be interpreted directly by the model. Interpretability is one of the main concerns when it comes to adopting ML algorithms (T. Kim et al., 2020; E.-J. Kim, 2021; Koushik et al., 2020). However, several *post hoc* interpretability techniques have been recently developed but seldom used in ML applications within travel behaviour research. To understand how a model predicts and which variables and to what extent contribute to its prediction, global and local interpretability tools are applied. It is necessary to mention that employing an intrinsically interpretable model is always preferred if the difference in its performance and that of black-box models is insignificant.

4.7.1 Global interpretability

Global interpretability clarifies how a model predicts in general and what is the entire behaviour of the model (Molnar, 2020). This approach quantifies the relationship and contribution of each feature to the model's prediction. Below are the descriptions for some of the interpretability tools.

Feature importance is a widely used global interpretation technique calculated as the total effect of each feature on the final prediction. It reflects how important a feature is for the predicted outcome of a model. The most used approach is permutation-based feature importance, in which the mean decrease in the performance of the out-of-bag sample is computed after permuting the values of a feature (Casalicchio, Molnar, & Bischl, 2018; Breiman, 2001). If the model's prediction error is increased after such permutation, it shows the feature is important, and the model's performance is sensitive to the change. On the other hand, shuffling the values of an unimportant feature does not significantly affect the model's performance. Ultimately, each feature is ranked by its variation in the model's prediction error after shuffling its values. Refer to Appendix A.1 for more details.

Partial Dependence Plot (PDP) is a popular method to compute the partial relationship between one or a set of features and the targeted response (Friedman, 2001). This plot represents how changes in the distribution of one or two features affect the average expected outcome of the model while fixing the values of the remaining features. When there is no correlation between a feature and other predictors, the PDP accurately shows how the features of interest affect the average prediction. The computation of a PDP is straightforward to interpret; however, its main disadvantage is the independence assumption between attributes (Molnar, 2020). The algorithmic way to obtain PDP is explained in Appendix A.2.

4.7.2 Local interpretability

Another category of model-agnostic interpretation tools is the local interpretability technique emphasizing individual instances and examining the features' effects on the outcome per instance (Molnar, 2020). ICE, SHAP, and LIME are examples of local interpretability tools.

Individual Conditional Expectation (ICE) plot is an extension of PDP that displays the effect of each attribute on the final prediction for individual observations (Goldstein et al., 2015). Instead of calculating the average partial relationship, ICE plots represent how much a change in the value of a set of features affects the prediction of a single instance (See Appendix A.3 for more details). To generate both PDP and ICE plots, `scikit-learn` package (in `Python`) is used.

Shapley value (SHAP) is a recently developed model-agnostic technique whose values are defined as the unified measure of feature importance. SHAP values, computed based on cooperative game theory, represent the contribution of each feature to the final prediction of a specific instance (Lundberg & Lee, 2017; Molnar, 2020). This local contribution is measured by the difference between the prediction value of a specific observation per feature (independent variable) and the average prediction of a model according to various possible coalitions. When the feature i joins a coalition S, its marginal contribution to the prediction f_x is computed as

$$\Delta_i(x) = f_x(S) - f_x(S \setminus i). \tag{4.13}$$

Therefore, the SHAP value evaluates the model's prediction and features' impact within every combination of features for each observation. The SHAP value (ϕ) for feature i at an observation x using model f can be calculated as

$$\phi_i(f,x) = \sum_{S \subseteq (\{1,\dots,p\} \setminus \{i\})} \frac{|S|!\,(p-|S|-1)!}{p!} \Big(f_x(S) - f_x(S \setminus i)\Big) \tag{4.14}$$

where p is the number of features, S is a subset of features, $f_x(S)$ is the prediction value including all features in a subset of S and $f_x(S \setminus i)$ is the prediction value of a subset S without feature i. For this study, the `shap` package in `Python` is used to compute SHAP values.

Local Interpretable Model-agnostic Explanations (LIME). Like SHAP values, this method also provides local explanations of any model and shows the heterogeneity of individual observations (Ribeiro et al., 2016). These explanations are given by approximating the underlying black-box model to a simple interpretable model, e.g., linear models or decision trees, around a single input point (Molnar, 2020). LIME perturbs numeric data features using standard normal distribution and categorical features according to the training distribution. It then learns locally weighted linear models on the attribute-space neighbor data points of a specific observation. Accordingly, it locally interprets the predicted values of an observation through an interpretable model. Thus, LIME models with the interpretability constraint explain the instance x using the following notation

$$\mathcal{L}(x) = \arg\min_{g \in G} L(f, g, \pi_x) + \Omega(g) \tag{4.15}$$

where g is an interpretable model (e.g., the linear regression model), which minimizes the loss L (e.g., RMSE), and Ω is the model complexity (e.g., the number of features). Hence, the aim is to minimize the difference error between the original model f and the explanation. On the other hand, π_x determines the proximity range around instance x, which is considered for its explanation (Please refer to Appendix A.4). The `lime` package in `Python` is used to implement the LIME interpretation.

4.8 Conclusion

The first aim of this study is to investigate the variations in travelers' trip chaining decisions and travel behaviour to understand how their activity-travel patterns differ. To this end, a clustering-based framework is employed to group individuals' trip preferences and travel patterns according to their similarities. For this analysis, a hierarchical clustering technique using an agglomerative algorithm is preferable to other clustering methods since it does not require defining the number of clusters in advance, unlike other algorithms such as k-means. The number of activity pattern clusters in this study is also initially undetermined. The agglomerative hierarchical clustering algorithm merges clusters to build homogeneous classes. This widely used bottom-up approach is appropriate for this study since it begins by grouping individuals' trip decisions as a single object, builds sub-clusters by merging similar clusters, and stops when it reaches homogeneous subsets. Accordingly, this clustering approach is proposed to reveal hidden information about activity patterns.

To examine the dependency of transit use on other variables, regression models are utilized. A ZINB regression model is used to explore the impact of the socio-economic, built environment, and trip factors on the number of transit trips. The ZINB method is a suitable model due to the high number of zero transit trips reported by individuals in their daily travel diaries. With extra zeros in the dataset, the zero-inflated regression model performs better in handling numerous zeros.

Since statistical and econometric models are easy to understand and have straightforward interpretations, they are commonly used to estimate changes in travel behaviour or traveler's mode choice decisions. However, they fail to represent the complex and non-linear relationship in the input. They also require a predefined assumption about the data on the decision process. ML algorithms, in contrast to statistical models, make no assumptions and identify the underlying pattern of the data by experiencing and learning from the dataset. The advanced computational power of ML algorithms to predict a traveller's complex behavioural responses with high prediction performance makes them a promising alternative in transportation planning modelling and travel demand management. Nevertheless, there is still limited research exploring the application of ML models in predicting transit trips. Selecting a predictive model with high predictive performance can affect transportation policies and travel demand management. In this study, the predicted performance of statistical models and ML algorithms is compared based on various performance metrics. The comparisons are made for classification and regression tasks to determine which model in travel behaviour analysis performs the best. Additionally, the interpretability of ML algorithms and features' contributions to the prediction is evaluated since identifying the factors that influence the outcome of the prediction is crucial for creating new insight into travel behaviour analysis.

5. TRAVEL BEHAVIOUR CLUSTERING

5.1 Chapter Overview

This chapter builds on the analyses conducted in Section 4.2 to identify residents' activity-travel patterns in the GTHA, and investigate how income and car ownership levels affect their travel decisions and behaviour according to their trip destinations and travel mode choices. Further, the heterogeneity in travel patterns of low- and high-income carless and car-owning households, and the effect that sociodemographic and built environment factors may have on this heterogeneity are investigated. Section 5.2 discusses the clustering result of low-income households. Then, the travel patterns of low-income clusters are compared with high-income ones in Section 5.3 to have a clear picture of the mobility decisions of different income levels. Particularly, this chapter answers the following research questions:

(RQ1-1) How does car-ownership affect the trip chaining behaviours of low-income communities?

(RQ1-2) How do the trip chaining decisions of low-income households differ from those of high-income households?

As explained in Section 3.4.3 (see Table 3.2), only individuals aged 18 years or older who are living in GTHA with either income levels less than $40k or more than $125k are examined. Two extreme income levels consider and they are categorized as low-income and high-income households. It helps explore the differences in travel behaviour according to income levels. The dataset is also limited to adults as they are assumed to be independent in their residential location, daily trips, and travel decisions. The subset includes trips that start from and end in GTHA and excludes people with an extraordinary number of trips (25 or more) or missing trip, income, and vehicle ownership information. Since car ownership affects the travel behaviour of a person, the dataset is divided into carless and car-owner people (Scheiner & Holz-Rau, 2012).

Notably, carless people do not use the car as their travel mode unless they rent a car or become a passenger in someone else's car. On the other hand, car-owners have the liberty to choose their travel mode.

5.2 Cluster Analysis of Low-income Travellers

To answer the first research question, the characteristics of low-income clusters in terms of activity patterns, socioeconomic, and built environment features is studied. To this end, car-owner and carless individuals are separately clustered into four clusters. Then, the descriptive attributes of each cluster are explored, and finally, the logistic regression (LogR) model is used to investigate the impact of socioeconomic and built environment features on assigning an individual to a particular cluster.

5.2.1 Activity patterns of low-income clusters

The activity pattern section of Table 5.1 shows the descriptive analysis of each cluster for low-income carless and car-owner households. Two clusters with their prime destination as work in the carless category are defined: Clusters 0 and 3. Cluster 0, including 46% of the low-income carless subsample, almost corresponds to work tour transit riders having nearly half of their trip chains as simple work tours and 39% of them as school and shopping trips. However, individuals of cluster 3, who commute by foot or bicycle, incorporate 17% of the poor carless subsample. One-fourth of their trip chains belong to simple tours with work destinations, and 35% of them are part of simple tours for school and shopping purposes. It illustrates that work and school trips have more similarities in individuals' daily activity patterns. The remaining two clusters (clusters 1 and 2), containing 37% of the low-income carless subsample, mainly consist of non-essential trips. Figure 5.1a represents the trip legs observed in the daily trip chain of at least 10 percent of the low-income carless strata. The majority in cluster 1 go to discretionary destinations by car as a passenger, whereas individuals in cluster 2 take public transit for similar trip purposes.

Table 5.1 : Cluster description of low-income households and the probability of belonging to the given cluster.

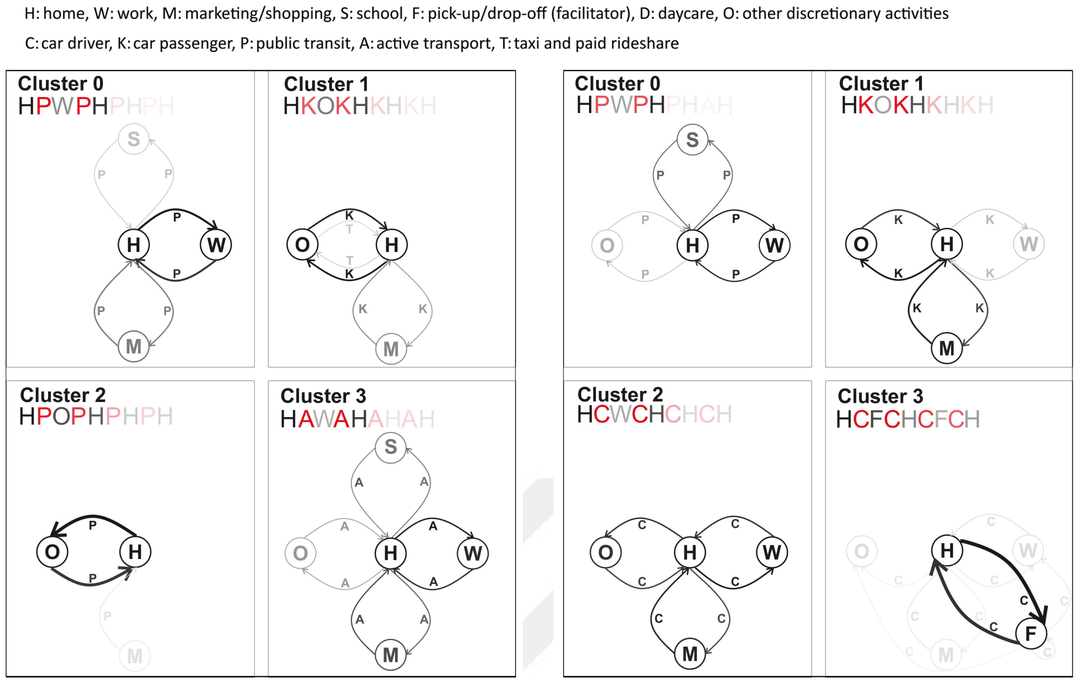
		Carless								Car-owner							
	Clusters	**#0**		**#1**		**#2**		**#3**		**#0**		**#1**		**#2**		**#3**	
	The number of individuals (cluster size %)	68,868 (46%)		14,302 (10%)		40,421 (27%)		25,635 (17%)		68,676 (19%)		49,755 (14%)		202,736 (56%)		41,587 (11%)	
Activity Patterns	**The most frequent destinations**	W[a]		O		O		W		W		O		W		F	
	Prime mode	P[b]		K		P		A		P		K		C		C	
	Top-3 most frequent trip sequences (%)	(41%) HPWPH		(28%) HKOKH		(44%) HPOPH		(26%) HAWAH		(27%) HPWPH		(25%) HKOKH		(28%) HCWCH		(16%) HCFCH	
		(25%) HPMPH		(14%) HKMKH		(4%) HPOPMPH		(18%) HAMAH		(20%) HPSPH		(23%) HKMKH		(17%) HCMCH		(9%) HCFCHCFCH	
		(14%) HPSPH		(10%) HTOTH		(3%) HPOPOPH		(17%) HASAH		(10%) HPOPH		(11%) HKWKH		(15%) HCOCH		(6%) HCFCWCH	
	Multi-modality (%)	10%		8%		34%		6%		9%		12%		5%		9%	
	Number of trips (μ)	2.3		2.4		2.9		2.6		2.2		2.6		2.8		4.6	
	Logistic Regression Independent Variables	(values)	OR[c]	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR
Socioeconomic Attributes	*Intercept*		0.080 ***		0.022 ***		0.013 ***		0.577 **		1.439 ***		0.710 *		0.000		0.000
	Age (μ)	(47)	1.002	(68)	**1.025** ***	(57)	**1.015** ***	(43)	0.974 ***	(40)	0.966 ***	(59)	0.997 **	(58)	**1.028** ***	(52)	0.983 ***
	Gender (male %) *(ref.: Female)*	(38%)	0.991	(24%)	0.727 ***	(35%)	1.047	(42%)	1.195 *	(40%)	0.829 ***	(21%)	0.404 ***	(53%)	**1.583** ***	(51%)	1.116 *
	Having transit pass *(ref.: No)*																
	Yes (%)	(70%)	**6.523** ***	(19%)	0.748 **	(54%)	3.845 ***	(32%)	0.719 ***	(51%)	1.765 ***	(13%)	0.462 ***	(7%)	0.194 ***	(8%)	0.308 ***
	Having driving license *(ref.: No)*																
	Yes (%)	(42%)	0.287 ***	(28%)	0.158 ***	(40%)	0.310 ***	(54%)	0.491 ***	(66%)	0.678 ***	(60%)	0.359 ***	(100%)	$1.03e8$	(100%)	$3.5e7$
	Distance of Mandatory trips (km) (μ)	(17.6)	1.000	(9.0)	0.997	(10.3)	0.976 ***	(3.4)	0.811 ***	(22.1)	1.004 ***	(12.7)	0.991 ***	(19.2)	1.014 ***	(17.7)	0.987 ***
	Distance of Discretionary trips (km) (μ)	(2.7)	0.923 ***	(8.8)	1.012 ***	(9.8)	1.033 ***	(1.2)	0.785 ***	(3.0)	0.947 ***	(10.1)	1.010 ***	(8.4)	0.986 ***	(17.3)	1.039 ***
Built environment Attributes	**Free parking at the workplace** *(ref.: No)*																
	Yes (%)	(23%)	0.848 *	(6%)	0.804	(7%)	0.595 ***	(12%)	0.628 ***	(20%)	0.479 ***	(18%)	1.225 *	(34%)	**2.650** ***	(30%)	**1.844** ***
	NA (%)	(55%)	0.786 ***	(91%)	1.733 **	(83%)	1.529 ***	(68%)	0.674 ***	(57%)	0.850 **	(78%)	1.482 ***	(62%)	1.303 ***	(66%)	1.813 ***
	Measure of accessibility to jobs using a gravity function (μ)	(3.0)	**1.457** ***	(2.0)	0.836 ***	(3.0)	**1.449** ***	(3.8)	**1.451** ***	(2.5)	1.056 **	(1.7)	0.771 ***	(1.6)	0.759 ***	(1.7)	1.000
	Population density (μ)	(1.9)	1.113 ***	(1.2)	1.040	(1.9)	1.026	(2.6)	1.016	(1.5)	0.932 ***	(1.0)	0.972	(1.0)	0.950 **	(1.0)	1.040
	Business density (μ)	(1.4)	0.840 ***	(0.9)	1.066 ·	(1.5)	0.939 ***	(3.0)	1.042 *	(1.1)	0.965 *	(0.6)	0.988	(0.6)	1.032 ·	(0.6)	0.791 ***
	Intersection density (μ)	(1.2)	1.014	(1.1)	1.014	(1.3)	1.206 ***	(1.6)	**1.45**0 ***	(1.1)	0.862 ***	(0.9)	0.804 ***	(0.9)	0.915 **	(1.0)	1.026

[a] H:home, W:work, M:marketing/shopping, S:school, F:pick-up/drop-off (facilitator), D:daycare, O:other discretionary activities

[b] C:car driver, K:car passenger, P:public transit, A:active transport, T:taxi and paid rideshare

[c] OR (odds ratio) values denote the probability of being in a given cluster. Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, · $p < 0.1$.

Notably, cluster 2 has the highest multi-modality rate, i.e., using different modes to reach their destination. They mainly chain their public transit trips with other travel modes.



(a) (b)

Figure 5.1 : The travel pattern of four clusters for low-income carless and car-owners (The transparency shows the frequency of each trip segment, destination/travel mode: the lighter it is, the less frequent it will be). (a) Low-income, carless individuals. (b) Low-income, car-owner individuals.

Regarding car-owner households, cluster 0 represents people taking public transit for work trips. Despite having a car, they still rely upon public transportation to reach their destination, probably due to car deficiency in their family. Individuals of clusters 1 and 2 frequently go to work, shopping, or other discretionary destinations as car passengers or car drivers, respectively. On the other hand, individuals in cluster 3 have pick-up/drop-off trips, accounting for the smallest category (11% of low-income car-owners). Expectedly, they have more trip segments, more trip numbers, and the longest trip chains compared to other clusters (see Figure 5.1b). This may indicate that using a car offers them more flexibility to visit different destinations while picking up other people. Interestingly, all low-income clusters have a relatively low multi-modality rate. It was expected that they use frequently multiple travel modes to complete their travel needs because they live in less accessible neighbourhoods. Comparing carless and car owners illustrates that low-income carless households are more likely to be multi-modal travellers because they live in high-accessible and high-density areas compared to car-owners. They use multiple modes to reach their destination even if they are poor and have no access to a private car. Therefore, the descriptive analysis suggests a relatively low multi-modality rate for low-income households, i.e., mainly relying on a single mode to complete their trip chain.

5.2.2 Socioeconomic characteristics of low-income clusters

The socioeconomic attributes section of Table 5.1 includes the independent variables of each activity pattern, their odds ratios, and the significance levels of coefficients obtained from the LogR models. A LogR model is fitted on each cluster to estimate the effects of independent variables on the membership of individuals in a specific cluster.

Regarding low-income carless households, the results show two distinct age groups: middle-aged individuals often taking work trips (cluster 0, 3) and elderly ones often making non-work trips (cluster 1, 2). Although cluster 0 of this group is a female-dominated category with the lengthiest mandatory trips, the odds ratios (ORs) of these factors are insignificant in determining the cluster. However, the ORs of having transit pass, accessibility, and population density of the same cluster positively contribute to the membership. 70% of individuals in this cluster have transit passes and use public transit as their prime mode, meaning that in the lack of private vehicles, middle-aged low-income women are forced to take lengthy work trips by transit. Cluster 3 of the poor zero-car group is the youngest among other clusters. They use active transportation for their short-distance mandatory and discretionary trips. The odds ratio of gender in this cluster indicates that being male leads to an increase of 19.5% in the likelihood of having this type of activity pattern. Clusters 1 and 2, mainly older females, are more responsible for making non-work trips. The mobility patterns of these groups underlines the traditional role of women in carrying more household labor than men (Madariaga, 2016; J. Lee et al., 2018; Craig & van Tienoven, 2019). They frequently make these trips, known as mobility of care trips, by public transit and by car as a passenger.

Regarding car-owners, although the car use is more often than public transit (see Table 3.4), cluster 0 reveals that females are the ones who mostly take public transit while sharing a private vehicle in their household. On the other hand, within cluster 2, the probability of driving a car for daily activities increases by 58.3% for men compared to women. Also, cluster 3 is a male-dominated class, using the car for pick-up/drop-off, with the highest number of trips (4.6). This is consistent with previous studies denoting males have the first right of using a car in a household due to gender inequality in access to the private vehicle between household members (Rosenbloom, 2004; Vance & Iovanna, 2007; Anggraini, Arentze, & Timmermans, 2008; Scheiner & Holz-Rau, 2012). Females of cluster 1 in the low-income car-owner group often make their non-work trips by car as a passenger. The finding aligns with the literature suggesting low-income women are more responsible for non-work trips, noting there is a considerable disparity in access to a car among men and women in their household (Simma & Axhausen, 2001; Blumenberg, 2016).

5.2.3 Built environment characteristics of low-income clusters

In the low-income carless groups, accessibility to transit is a significant predictor of having different travel mode decisions and activity patterns. Although accessibility has a positive relationship with clusters not taking a car, cluster 1 with people mainly using a car as a passenger and living in low accessible neighbourhoods demonstrates the opposite pattern. Cluster 3 has more variety in trip purposes (see Figure 5.1a) as they live in the most accessible regions, i.e., downtown Toronto and Hamilton, where walking and cycling are feasible. Unsurprisingly, an increase in the intersection density, i.e., more walkable streets, increases the probability of walking or cycling behaviour. It highlights the influence of built form and grid-like street networks on increasing the number of pedestrians and cyclists.

Regarding car-owner households, females of cluster 0 are probably car-deficit individuals living in inner suburbs or around transit stops. The positive and significant coefficient of transit accessibility for this cluster shows that an increase in accessibility enhances the likelihood of having this activity pattern by 5.7%. They choose or are forced to use public transit even with moderate transit accessibility (acc. score = 2.5). Comparing the transit accessibility of all low-income clusters demonstrates that accessibility is a more significant factor for low-income zero-cars than car-owners. The principal factor in driving a car for commuting among low-income car-owners is the free-parking availability at the workplace.

5.3 Comparing Low-income and High-income Household's Travel Behaviour

To better understand the effect of income levels on travel behaviour, a cross-group comparison between low- and high-income clusters is made. Also, car ownership leads to different travel patterns among clusters. Accordingly, in the second research question, the disparity in travellers' trip decisions considering their income and car ownership is explored. Table 5.2 describes the characteristics of high-income clusters in terms of activity patterns, socioeconomic attributes, and built environment features.

5.3.1 Carless clusters

Comparing low- and high-income carless households reveals low-income households have fewer and shorter distance trips compared to their high-income counterparts. Also, it is consistent with the literature (Blumenberg & Thomas, 2014) that the low-income clusters have a lower multi-modality rate than high-income ones. Among all wealthy carless strata, cluster 2 has the highest multi-modality rate, and its population mostly relies on public transport and active transportation in their daily trip sequences. For instance, they use public transport to commute to work and walk back to their home (see Figure 5.2a). The positive relationship between accessibility and intersection density for such an activity pattern underscores the importance of accessibility and connected street networks factors in encouraging multi-modal travel patterns. Such a multi-modal cluster cannot be captured by a simple rule-based approach, defined only based on the prime mode. It emphasizes the importance of using clustering algorithms in finding behavioural patterns in people's trip chains.

Table 5.2 : Cluster description of high-income households and the probability of belonging to the given cluster.

		Carless								Car-owner							
	Clusters	**#0**		**#1**		**#2**		**#3**		**#0**		**#1**		**#2**		**#3**	
	The number of individuals (cluster size %)	11,285 (51%)		1,271 (6%)		1,986 (9%)		7,641 (34%)		199,078 (24%)		129,439 (16%)		407,759 (49%)		93,550 (11%)	
Activity Patterns	**The most frequent destinations**	W[a]		W		W		W		W		O		W		F	
	Prime mode	P[b]		T		A/P		A		P		C		C		C	
	Top-3 most frequent trip sequences (%)	(53%) HPWPH		(9%) HTWTH		(13%) HPWAH		(58%) HAWAH		(39%) HPWPH		(19%) HCOCH		(60%) HCWCH		(11%) HCFCWCH	
		(6%) HPOPH		(7%) HKOKH		(7%) HAWPH		(3%) HAOAH		(9%) HAWAH		(14%) HCWCHCOCH		(5%) HCMCH		(8%) HCFCWCFCH	
		(5%) HPWPOPH		(7%) HTOTH		(4%) HPWPHAOAH		(3%) HAWAOAH		(6%) HKWKH		(9%) HCWCOCH		(2%) HCWCMCH		(7%) HCFCH	
	Multi-modality (%)	24%		53%		100%		9%		17%		15%		11%		11%	
	Number of trips (μ)	2.6		3.2		3.5		2.7		2.4		4.2		2.7		4.9	
	Logistic Regression Independent Variables	(values)	OR[c]	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR	(values)	OR
Socioeconomic Attributes	*Intercept*		0.001 ***		0.001 ***		0.000 ***		0.013 ***		18.649 ***		0.000		0.000		0.000
	Age (μ)	(41)	1.000	(44)	0.995	(41)	0.990	(39)	0.974 ***	(43)	0.984 ***	(49)	1.011 ***	(46)	1.004 ***	(47)	1.007 ***
	Gender (male %) *(ref.: Female)*	(52%)	1.045	(47%)	1.173	(61%)	1.344	(60%)	**1.800** ***	(45%)	0.703 ***	(56%)	**1.265** ***	(54%)	1.018	(51%)	1.024
	Having transit pass *(ref.: No)*																
	Yes (%)	(86%)	8.901 ***	(40%)	0.980	(74%)	4.007 ***	(46%)	1.130	(58%)	4.336 ***	(17%)	0.724 ***	(14%)	0.309 ***	(15%)	0.621 ***
	Having driving license *(ref.: No)*																
	Yes (%)	(76%)	0.261 ***	(72%)	0.224 ***	(85%)	0.699	(83%)	0.329 ***	(86%)	0.040 ***	(100%)	7.3e5	(100%)	2.8e6	(100%)	4.7e5
	Distance of Mandatory trips (km) (μ)	(17.9)	0.998	(15.3)	0.998	(9.3)	0.954 ***	(5.1)	0.859 ***	(32.1)	0.998 ***	(33.0)	0.995 ***	(40.8)	1.009 ***	(29.5)	0.987 ***
	Distance of Discretionary trips (km) (μ)	(3.2)	0.988 ·	(5.8)	1.000	(4.0)	1.018	(1.5)	0.959 *	(3.6)	0.962 ***	(18.7)	1.038 ***	(4.8)	0.950 ***	(19.1)	1.033 ***
Built environment Attributes	**Free parking at the workplace** *(ref.: No)*																
	Yes (%)	(13%)	0.512 ***	(11%)	0.464 ·	(4%)	0.273 *	(8%)	0.370 ***	(18%)	0.230 ***	(46%)	1.199 ***	(64%)	2.641 ***	(51%)	1.154 **
	NA (%)	(19%)	0.759 *	(38%)	1.430	(29%)	1.276	(23%)	0.666 **	(24%)	0.620 ***	(35%)	1.627 ***	(18%)	1.329 ***	(30%)	1.189 **
	Measure of accessibility to jobs using a gravity function (μ)	(4.3)	2.200 ***	(4.2)	1.806 ***	(4.7)	**3.377** ***	(4.8)	2.378 ***	(2.6)	1.219 ***	(1.6)	0.991	(1.5)	0.884 ***	(1.5)	0.956 **
	Population density (μ)	(2.5)	1.237 ***	(2.4)	1.179 *	(2.7)	1.121	(2.8)	1.050	(1.2)	0.969 ·	(0.7)	0.901 ***	(0.7)	0.971	(0.6)	0.901 **
	Business density (μ)	(3.5)	0.900 ***	(4.1)	1.006	(4.5)	0.897 **	(5.2)	0.957 **	(1.6)	0.990	(0.8)	0.999	(0.7)	0.980 *	(0.6)	0.859 ***
	Intersection density (μ)	(1.7)	1.183 **	(1.7)	1.022	(2.0)	1.276 *	(2.0)	**1.388** ***	(1.2)	0.953 ·	(0.9)	0.981	(0.9)	0.952 *	(0.9)	1.129 ***

[a] H:home, W:work, M:marketing/shopping, S:school, F:pick-up/drop-off (facilitator), D:daycare, O:other discretionary activities

[b] C:car driver, K:car passenger, P:public transit, A:active transport, T:taxi and paid rideshare

[c] OR (odds ratio) values denote the probability of being in a given cluster. Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, · $p < 0.1$.

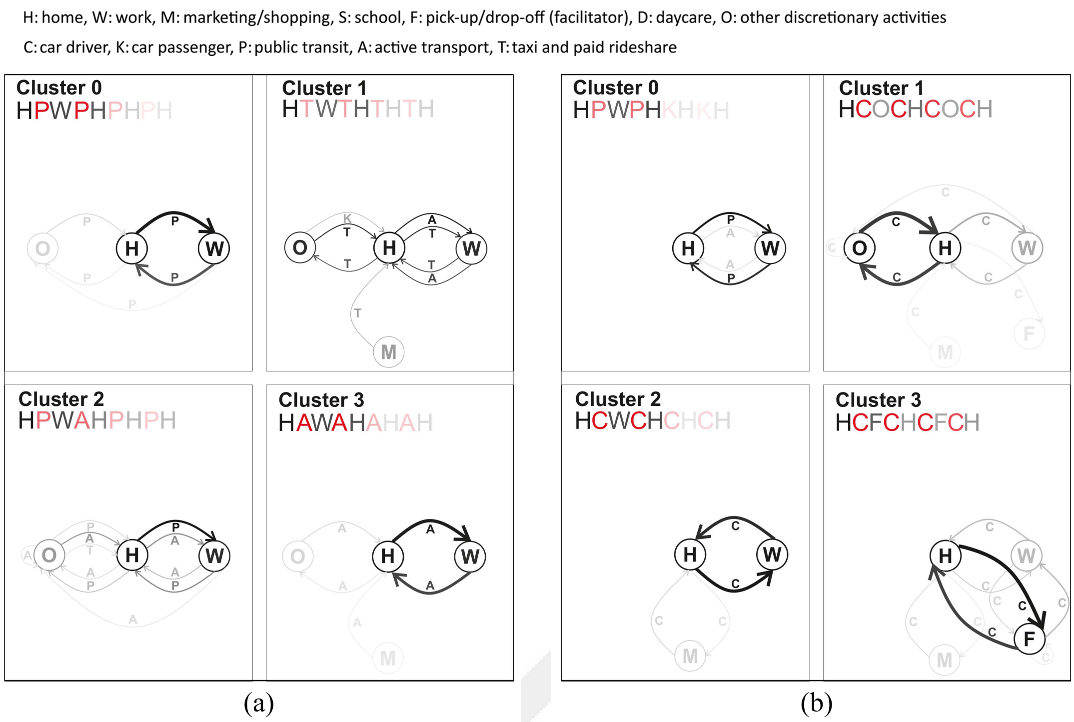


Figure 5.2 : The travel pattern of four clusters for high-income carless and car-owners (The transparency shows the frequency of each trip segment, destination/travel mode: the lighter it is, the less frequent it will be). (a)High-income, carless individuals. (b)High-income, car-owner individuals.

Expectedly, transit accessibility is a significant predictor for all carless clusters. Comparing the odds ratio of transit accessibility exhibits a significantly positive relationship for all high-income carless clusters compared to low-income zero-car ones. Cluster 0 of low-income carless households live in poorer accessible neighbourhoods compared to their counterparts (3.0 vs. 4.3). This may be due to higher home prices around the main transit stops and network. The lack of affordable houses near major transit hubs (see Figure 3.4) makes lower incomes more isolated in society, leading to shorter trip chains and distance traveled. They are forced to settle down in neighbourhoods with affordable houses but poor transit services while minimizing their number of trips due to time or money costs (Allen & Farber, 2020b; Paez, Ruben, Faber, Morency, & Roorda, 2009).

Cluster 0 of both income levels shares several identical characteristics. They mainly go to work by public transit and take the longest traveled distance mandatory trips. The main difference comes from the ratio of females to males in low-income households. The intersectionality among low-income, living in lower accessible regions, not having a private car, and being female makes them a vulnerable group in society.

The travel pattern of low- and high-income households seems different when they take other discretionary trips. Clusters 1 and 2 of the low-income carless group, who often have non-work trips, live in relatively low accessible, population, and business density neighbourhoods (see Figure 3.5). Accordingly, they mostly rely on others (car as a passenger) to take their trips or on public transit. On the other hand, if high-income carless households want to take discretionary trips, they mainly chain them with their work trips. They also take long traveled distance mandatory and discretionary trips by taxi or paid rideshare (15.3 km & 5.8 km, respectively).

Cluster 3 of both carless subcategories make their work trips mostly with active transportation. Among all other clusters of their subsample, they live in neighbourhoods with the highest levels of accessibility and density (e.g., downtown) and are the youngest cluster. They also have the shortest traveled distance mandatory trips compared to other groups of their subsamples as they mainly walk or cycle to their destination. On the other hand, high-income households of cluster 3 live in more accessible areas compared to their low-income counterparts (4.8 vs. 3.8). It indicates that the rich can afford to live in neighbourhoods with higher accessibility scores (e.g., downtown) in which walking and cycling are readily available. Surprisingly, despite the better accessibility of high-income households in this cluster, they still have lengthier trips than their counterparts in low-income groups. Besides active transport and transit, wealthy families take taxis or paid rideshare (e.g., Uber and Lyft) for their daily trips. Such a prime mode among low-income carless households is not observed since the cost can hinder their freedom of choice.

5.3.2 Car-owning clusters

Unlike zero-car families, low- and high-income car-owning households have significantly lengthier trips while living in more distant areas (e.g., remote areas in the suburbs). This indicates that driving a car gives more flexibility to residents to visit remote regions. In general, low-income car-owners are older and travel to closer destinations than high-income ones. Furthermore, males are the main drivers of the car in both car-owning families, and females choose or are forced to be transit riders or car passengers in completing their daily trip chains. It aligns with the literature that women have limited access to cars in car-owning households (Rosenbloom, 2004; Vance &

Iovanna, 2007; Anggraini et al., 2008; Scheiner & Holz-Rau, 2012). For car-owners, the free-parking availability at the workplace is an important factor. It has a positive association with using a car as a daily travel mode among car-owner households.

Both income levels still have a cluster of public transit users, whereas the poor take a car as a passenger as their prime mode, which is not the case in the high-income category. These transit riders have the least number of trips and the least complexity in their trip sequences. It may indicate that growing a variety of visited stops in a day requires a flexible mode. This finding aligns with the literature that more complex trip chains usually require a private car (Hensher & Reyes, 2000). The much better transit accessibility is among transit users of both income levels (i.e., cluster 0) compared to other groups of their subsample. Public transit does not provide the best alternative for a car when it comes to the convenience of multiple trips with different purposes. Therefore, people in car-deficit households (where they do not have one car per adult) are constrained to take transit for their daily trips, leading to less flexible trip decisions.

On the other hand, cluster 2 of car-owners have the same most frequent trip destination as work, whereas they often drive a car rather than taking public transit. Although their destinations are the same, car drivers may chain their trips to shopping or other activities. Hence, their discretionary traveled distance is slightly higher than transit users. Their low rate of accessibility score, population, and business density values may be the reason why they often use the car for their trips. While cluster 2 of low- and high-income subsamples make lengthy mandatory trips, high-income households take longer mandatory traveled distances than their counterparts (40.8km vs. 19.2km). In contrast, the poor make lengthier discretionary trips (8.4 vs. 4.8) since their cluster is more mixed between working destinations and discretionary trips.

Considering the discretionary trips, cluster 1 of high-income car-owners use their own car, whereas low-income ones take a car as a passenger for shopping and visiting discretionary destinations. It may be due to the lack of vehicles per adult in low-income households. Similarly, the same cluster of the high-income category has a much higher total distance traveled than low-income ones (51.7 vs. 22.8 km). It indicates that driving a car increases the flexibility of going to different destinations. However, when you are a passenger and depend on other drivers, you may visit limited destinations. Besides, non-work trips in low-income car-owner households are made mainly by females (see the significant OR of smaller than 1 on Table 5.1). These women also are more dependent on others by using the car as a passenger to make those trips. This observation indicates the traditional role of females in carrying out household tasks and the difficulty of depending on others to make their trips (J. Lee et al., 2018; Craig & van Tienoven, 2019).

The pick-up/drop-off cluster is observed only for car-owners since having a car allows them to pick up or drop off someone during their daily trip (cluster 3). Expectedly, their trip sequence and traveled distance are higher than those of any other cluster. It confirms that having a car allows travellers to have flexible and complex trip chains (Hensher & Reyes, 2000). After they pick up or drop off someone, they continue their trips to other destinations – e.g., commuting to work with their colleague who lives in their neighbourhood. They are mainly located in suburbs where transit accessibility is relatively low.

5.4 Conclusion

This chapter presents a thorough evaluation of the travel behaviour of low- and high-income carless and car-owner individuals in the Greater Toronto and Hamilton Area (GTHA). Further, it investigates the role of socioeconomic characteristics and built-environment attributes in shaping the different patterns.

The whole dataset is divided into four subsamples according to income and car-ownership levels. Then, each subsample is clustered into four homogeneous clusters using hierarchical clustering on individuals' daily activity patterns through their trip destinations and the mode used. Several conclusions can be drawn by investigating the results. The key findings relevant to research questions are summarized as follows.

(RQ1-1) How does car-ownership affect the trip chaining behaviours of low-income communities

> The mobility pattern of two clusters of low-income carless households underlines the traditional role of women in carrying more household labor than men. The findings show a cluster of young, carless households who walk or cycle in the neighbourhoods where accessibility, land use density, and street design prioritize active transportation over motorized cars. Furthermore, two key

female-dominated low-income car-owner clusters are observed: one for those forced to take transit with moderate accessibility and one relying on others for their non-work trips though having a shared car in their family. It is consistent with previous studies emphasizing inequity in access to a shared private vehicle among men and women in their households. Moreover, low-income car-owners are less likely to use multiple travel modes than carless households for their daily trips because they live in low-density and low-accessible neighbourhoods.

(RQ1-2) How do the trip chaining decisions of low-income households differ from those of high-income households?

The findings show that the higher housing prices around main transit lines lead to the isolation of low-income households in less accessible regions, resulting in their fewer trips, less multi-modality, shorter trip chains and distance traveled. It highlights that low-income carless households face mobility and activity participation barriers in the lack of multiple modes in low accessible and density areas. Low-income carless families live in lower accessible regions, do not have a private car, and are mostly female, making them a vulnerable group in society. The wealthier carless households afford taxis or paid ridesharing (e.g., Uber and Lyft) for their daily trips, whereas the cost hinders such freedom of choice for lower incomes. Despite the higher accessibility of wealthy carless households, they still have lengthier trips compared to low-income ones. The results show that both low- and high-income carless groups, mainly including younger males, use active transportation as their prime mode.

Low-income car-owning households are older and drive to nearby destinations compared to those with higher incomes. Low-income car-owning households have a cluster of car-deficit families relying on others for their discretionary trips (cluster 1), whereas such a cluster does not exist for high-income car-owners. This cluster of low-income car-owning families mainly includes females fulfilling the traditional role of carrying out household tasks. Moreover, a cluster of transit users is seen in both income levels who have the lowest flexibility in their trips, indicating the critical role of cars for households living in lower accessible regions. The pick-up/drop-off cluster of both income levels demonstrates relatively identical characteristics.

6. IMPACTS OF TRANSIT INVESTMENTS ON SHIFTING MODE

6.1 Chapter Overview

This chapter extends the aim of Chapter 5 in understanding the trip decisions and preferences of residents in GTHA. According to the arguments started in Section 2.4, this study empirically explores how transit investments, leading to accessibility improvements, may change travel mode and transit use of different income and car ownership strata. Overall, this chapter answers the questions below regarding the consequence of increasing transit accessibility on changing mode decisions among different income and car-ownership groups.

(RQ2-1) To what extent can transit investments in lower socio-economic neighbourhoods enhance transit mode share?

(RQ2-2) To what extent are low-income car-owners sensitive to transit improvements and shift their travel mode use?

A zero-inflated negative binomial model is employed to predict the number of transit trips that a person has per day. Section 6.2 investigates the first research question by exploring the models' estimations in two phases: 1) using the whole sample to fit a comprehensive model on the population and 2) exploring 25 stratified models for different income and car-ownership levels (5 car ownership levels $\times$ 5 income levels). This work allows us to explore the variability in response to accessibility. A sensitivity analysis is performed in Section 6.3 to explore how changes to accessibility will differentially affect transit trip generation throughout the region, particularly for low-income car-owners. It provides answers for the second research question and forms an additional layer of policy-relevant analysis, enabling us to directly evaluate the potential for transit investments in low-income communities to unlock suppressed demand for transit travel.

6.2 ZINB Model Results

The dataset used for this analysis is limited to all adults who are aged 18 years or older and living in GTHA (see Table 3.1). A zero-inflated negative binomial (ZINB) model is estimated to investigate the influence and significance of socio-demographic characteristics, local environments, and trip factors on the number of transit trips per individual. In the analysis, the dependant variable is the count of daily transit trips of an individual. A pool of candidate attributes that may affect taking transit are defined after removing highly correlated covariates. The ZINB was selected because a large number of individuals in the data were without any transit trips in their travel day (n=116,451).

ZINB models, unlike negative binomial, can deal with excess zeros and over-dispersion of the data (Sultana, Mishra, Cherry, Golias, & Tabrizizadeh Jeffers, 2018). In particular, the model comprises two distinct processes: one for generating zero values with the probability p_i, and the other for generating counts from negative binomial with the probability $1-p_i$. The zero-inflation portion of the model consists of a binary logit model predicting non-occurrence, i.e., not taking transit, whereas the count portion of the model predicts the frequency of occurrence, i.e., the number of public transit trips. The results of the zero-inflation portion and count portion of ZINB are presented in the form of odds ratios and incidence rate ratios, respectively. They are obtained by exponentiating the coefficients of each of the model portions. Therefore, the expected number of transit trips is computed as

$$E(y_i) = p_i \times 0 + (1-p_i) \times n_i \tag{6.1}$$

where n_i is the expected transit trip count given it is not zero.

For this study, the weighted ZINB in which the weights are normalized TTS expansion factors of each individual are utilized. Instead of using the direct expansion factors as the weight, the weights are rescaled in a way that it sums to the stratum sample size. They are normalized by the mean of expansion factors per stratum. These weights correct biases that may occur due to non-representative sampling in the study region.

Table 6.1 contains the odds ratios (ORs), incidence rate ratios (IRRs), and significance levels of model coefficients. OR values, obtained from the zero-inflation portion, demonstrate the probability of having zero transit trips. In other words, the values greater than one shows the increase in the probability of not taking transit, and vice versa. For instance, the coefficient for transit accessibility shows a negative relationship for the zero-inflation portion and a positive relationship for the count portion of the model, indicating an increase in accessibility leads to an increase in the probability of taking transit trips[1]. For car ownership, owning more vehicles per adult in a household reduces the likelihood of using transit. Unsurprisingly, households with one or more cars per adult have a significantly negative coefficient, indicating a lower probability of taking transit than the other car owners. This interpretation may also represent the association between poor transit service and a resident's propensity to own a car. This assumption is explored in Section 6.3. Contrasting the effect size of car-ownership and income levels, the findings show that the number of vehicles per adult has a much higher coefficient than income in the zero-inflation portion. Moreover, individuals with a driver's licence, after controlling for car ownership, are more reluctant to take transit for their daily trips than those without a license. Likewise, free parking spots at the workplace reduce the likelihood of using transit. On the other hand, the coefficient of holding a transit pass is a significant predictor for taking transit. Males show 37.6 percent less inclination to take transit compared to females. Younger individuals have more propensity to take transit (becoming one-year older reduces the probability of taking transit by 2.8 percent).

Since travel mode choice is a function of the built environment (Cervero & Kockelman, 1997), the residential neighbourhood characteristics for each individual is appended to the dataset. Accordingly, intersection density as a design metric comes from the total number of 3-way or more intersections per square kilometer. The population density in each Dissemination Area is from the 2016 Canadian Census, and business density comes from the Canadian business registry. These variables are measured as the sum of individuals and businesses per square kilometer, respectively. After controlling for transit accessibility, the coefficients for population and business density show a negative association with using transit. Similarly, the intersection density is negatively associated with using transit. These results are somewhat puzzling, but assume that

[1]The alternative definitions of the transit accessibility variable to account for nonlinearities (quadratic, cubic and sigmoid transformations was examined), but they do not result in improved model fits or any changes in interpretation. Therefore, only the linear effect of accessibility in the models is considered to reduce their complexity.

Table 6.1 : ZINB model results (N= 3,279,979; n = 149,177).

Dep Var.= The number of Transit Trips		
Description of Independent Variables	**ZINB Model**	
	Probability of no transit trip (OR)[a,b]	**Incidence rate ratio (IRR) for transit use[a,b]**
(intercept)	0.007 ***	1.070 *
Distance of mandatory trips (km)	0.989 ***	1.004 ***
Distance of discretionary trips (km)	1.017 ***	1.005 ***
Age	1.029 ***	1.002 ***
Household's total income per year		
(ref. category: <$40k)		
$40k-$60k	1.112 *	0.987
$60k-$100k	1.027	0.932 ***
$100k-$125k	1.036	0.920 ***
$125k+	1.138 **	0.919 ***
Number of vehicles per adult		
(ref. category: VA=0)		
0 <VA <0.5	11.414 ***	0.948 ***
VA=0.5	24.769 ***	0.921 ***
0.5 <VA <1	31.107 ***	0.901 ***
VA=1+	78.198 ***	0.877 ***
Gender		
(ref. category: Female)	1.371 ***	0.986
Free parking at workplace		
(ref. category: No)		
Yes	11.691 ***	0.980
NA	4.839 ***	1.014
Having driving license		
(ref. category: No)	7.477 ***	0.970 **
Having transit pass		
(ref. category: No)	0.058 ***	1.414 ***
Measure of accessibility to jobs using a gravity function (transit commute)	0.572 ***	1.069 ***
Population density [c]	1.057 **	1.004
Business density [c]	1.056 ***	0.953 ***
Intersection density [c]	1.002 ***	0.999 ***

[a] Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

[b] People with no trips or an extraordinary number of trips greater than 25 are removed.

[c] Local built environment characteristics of travelers come from the weighted sum of values normalized by area in each Dissemination Area.

they indicate that higher densities are associated with high levels of active travel, something that is not discernable within a single-mode model like the ZINB. Having long discretionary trips is a deterrent to using transit; however, longer mandatory trips have a positive association with the probability of taking transit.

On the other hand, the count-model IRR values greater than one show a positive impact on taking more transit trips, and those less than one have a negative impact on the number of transit trips. Considering the count model's IRR, free parking spot at workplace, population density and gender are not a significant predictor of the overall number of transit trips. Moving from the reference low-income group to households with total income greater than $125k per year corresponds to a %12.3 decline in having more transit trips. Moreover, having more vehicles per adult in a household decreases the number of transit trips. Similarly, individuals with a driver's license have a 2.3% lower transit trip rate. Conversely, holding a transit pass increase the number of transit trips by 41.4%. Both the longer discretionary and mandatory trips have a positive association with using more transit trips in their daily trips.

6.3 Sensitivity Analysis

In this section, 25 stratified logistic regression (LogR) models are generated, one for each combination of income and car-ownership strata (5×5). The objective of these models is to contrast the effect size of accessibility across different groups. To see the probability of taking transit, the LogR model is utilized only for this section. According to previous studies, the coefficient of the zero portion of the ZINB may be difficult to interpret by having structural and sampling zeros (Staub & Winkelmann, 2013; Hua et al., 2014). Therefore, the LogR model is selected for this task. To compare the effect size, an elasticity metric is used. Elasticity, as a unit-free measurement, is the ratio of the percentage change in an independent variable associated with the percentage change in a dependent variable. The elasticity of accessibility for each observation is defined as

$$E_{x_i} = \beta_i \times x_i(1 - P_i) \qquad \forall x_i \in \mathbb{R} \tag{6.2}$$

where E_{x_i} is the elasticity of individual i, β_i is the LogR coefficient of transit accessibility, x_i is the transit accessibility value of individual i, and P_i is the estimated probability of taking transit (Train, 2009). Then, these elasticities are averaged over the population in each stratum (Ewing & Cervero, 2010).

$$\bar{E}_x = \frac{\sum_{i=1}^{n} w_i E_i}{\sum_{i=1}^{n} w_i} \tag{6.3}$$

$\bar{E}_x$ is the weighted elasticity where w_i is the expansion value of individual i.

In Figure 6.1, each grid defines the elasticity of accessibility for the corresponding stratified LogR model. The weights of the model are the expansion factor of each stratum normalized by the mean expansion factor of the same group. Interpretation of elasticities are straightforward. For instance, the elasticity of 0.87 for the low-income households with one or more vehicles per adult indicates that a 1 percent increase in accessibility will result in a 0.87 percent increase in the probability of taking transit. Incidentally, low-income and high car-owning households appear to have the highest overall sensitivities to transit accessibility, indicating a latent demand for mode-switching if only transit were better provided to them. There are about 140,000 individuals in this strata, representing a large proportion (26%) of the low-income population. Overall, the top row shows that all households with one or more vehicles per adult, take transit as transit accessibility improves. On the other hand, carless households in the high and medium-income category prove to be insensitive to the change in their accessibility, having a p-value less than 0.05. It can be the upshot of the fact that most carless wealthy households are already transit users, or live in places that allow for active travel lifestyles. Therefore, their transit accessibility cannot be further improved to increase the probability that they will use transit. Of course, improving accessibility may have other personal benefits for those carless households, such as less travel and waiting times, greater reliability, and less crowding. Interestingly, the effect of accessibility on transit ridership increases as households own more personal vehicles. People owning cars are optional transit riders, thus, enhancing their accessibility probably provides impetus to use public transport. The elasticities tell us which individuals are more or less sensitive to accessibility improvements. Next, these elasticities are applied to the GTHA's population to ascertain how much opportunity there is to generate additional transit trips by focussing investments at different strata.

To estimate the number of new transit trips induced by a hypothetical transit investment, i.e., accessibility increase, the stratified ZINB models are utilized to determine the current number of transit trips (y_i^c) for individual i in class c as a baseline. Then, while all other independent variables remain constant, the level of accessibility is incrementally increased from 0 to 200k new jobs and estimated the new number of transit trips ($\hat{y}_i^c$) for each person. The smaller gains in accessibility (less than 50k jobs) would roughly be achievable by moderate investments in the existing transit system (e.g.,

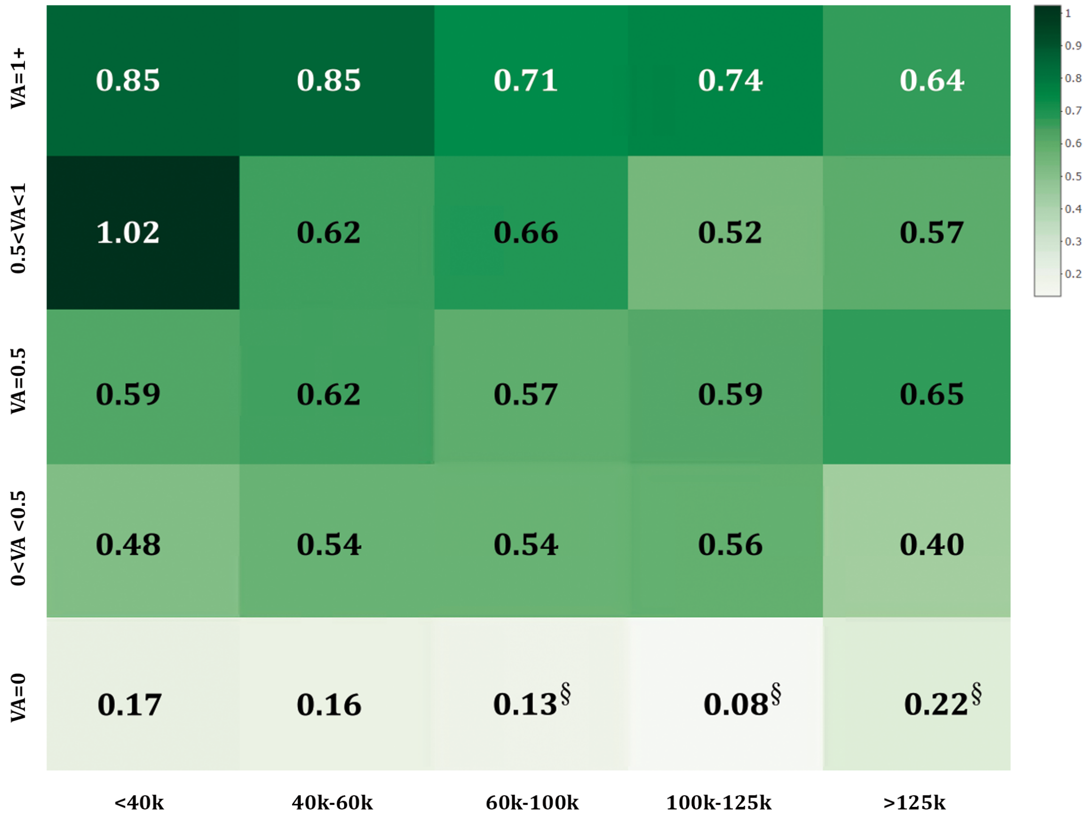


Figure 6.1 : Elasticity estimates of transit accessibility for 25 LogR models (income and car ownership levels).§*Elasticity estimates were insignificant at the 0.05 level.*

more frequent service), whereas the largest accessibility gain (i.e., 200k jobs) requires significant improvements in transit infrastructure (e.g., new rapid transit) (Allen & Farber, 2020b; Farber & Marino, 2017). Having the baseline values, the change in the number of transit trips per individual is computed in the weighted sample as follows:

$$\Delta\hat{y}^c = \frac{\sum_{i=1}^{n^c} \hat{y}_i^c - \sum_{i=1}^{n^c} y_i^c}{n^c} \qquad \forall c \in \{1 \ldots 25\} \tag{6.4}$$

where n^c is the total number of individuals in class c, and $\Delta\hat{y}^c$ is the predicted change in the number of transit trips due to a change in accessibility within population class c. The result of this analysis is demonstrated in Figure 6.2. The y-axis shows the expanded number of newly generated daily transit trips per person in the GTHA by increasing transit accessibility across the GTHA. These numbers include both transit trips shifted from other modes and entirely new transit trips, and don't differentiate between the two. Noting a large number of carless households reside in places with a high level of accessibility, a significant discrepancy in sensitivities of various income groups to accessibility improvements is still observed. Figure 6.2 shows that among zero-car groups, more transit trips are induced among low-income groups. Notably, these

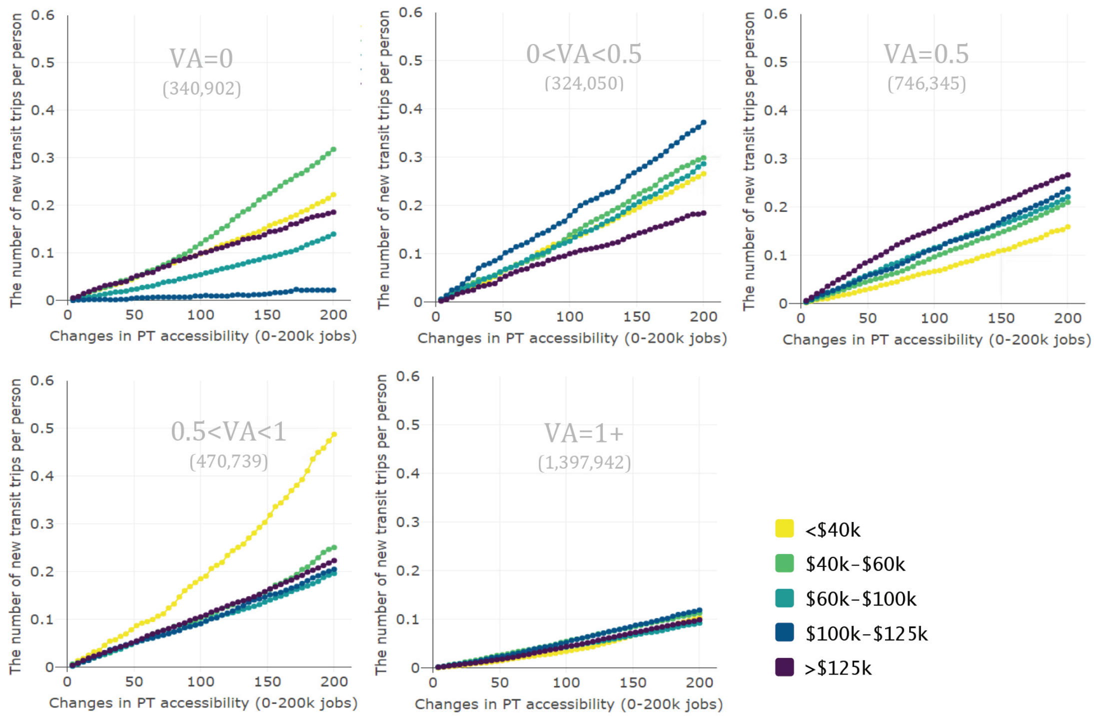


Figure 6.2 : The expanded changes in transit trips per person by accessibility improvements. (The expanded population of each stratum is shown on the graphs).

low-income carless households already take more transit trips than other income and car-ownership brackets (Figure 3.6). Moreover, Figure 6.1 showed the elasticity of non-transit riders of each class given accessibility improvement (converting from 0 to 1), while Figure 6.2 depicts the total number of newly generated transit trips per person —i.e., both new transit riders and increased transit trips among existing transit riders— after accessibility improvement. Overall, households with one or more cars per person, regardless of their income level, are less responsive to accessibility increase than other car-ownership brackets. The results show that when households own more cars, they are more willing to use them even if transit accessibility was improved. Notably, the accessibility coefficients of high and medium-income groups were insignificant for the carless strata in the models, meaning that these curves should only be used illustratively. Unsurprisingly, car-deficit households show more tendency toward taking transit after accessibility improvements because it opens a new door for them to select another travel mode. Notably, these households with less than a car per adult are strongly inclined to choose public transit since they have to share a car in a household.

Figure 6.3 shows how much transit ridership growth is associated with each source. It shows what percentage of new transit trips in each stratum originates from existing transit users and what percentage from new transit riders. It indicates improving transit accessibility has a significant impact on non-riders of households owning one or more cars per adult — i.e., VA=1+. These individuals have the highest sensitivity to transit accessibility improvement. On the other hand, most zero-car families with the least elasticity (Figure 6.1) are already transit riders. Therefore, accessibility improvements entice more existing riders of carless households to take more transit trips than non-riders.

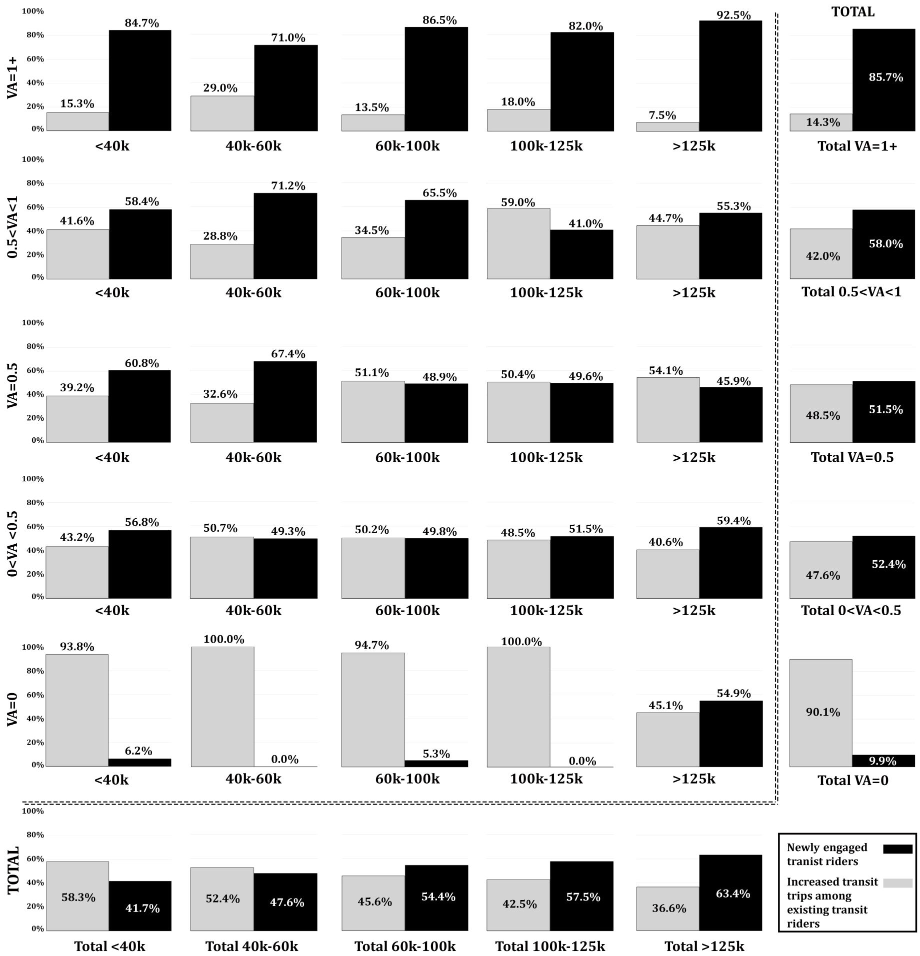


Figure 6.3 : The ratio of newly generated transit trips given existing and new users after transit improvement (200k).

Among carless households, non-riders of wealthy families are easily absorbed after transit accessibility improvement. However, the majority of low-income individuals are already transit riders. Therefore, they tend to increase their existing transit trips after accessibility gains. It illustrates that this stratum is still an unsaturated market, and its individuals need to be provided with transit investment. Interestingly, three car-deficit groups have almost equal potential for whether generating new transit trips or expanding their current transit trip.

Table 6.2 shows the response of each income and car-ownership level to a hypothetical level of transit improvement. It indicates the estimated number of newly generated trips after improving accessibility equally, by 50,000 or 200,000 jobs, or relatively, by 10 or 25 percent above existing levels for each respondent. The changes in accessibility have the most disproportional impact on the three classes of households with less than one car per adult (0<VA<1). The results illustrate that in total, 64% of new transit trips belong to these three groups although they are only 47% of the whole population. They are mostly living in inner-suburbs with a medium level of accessibility (Figure 3.7) and sharing one car in a household. Therefore, accessibility improvements help these family members to have another mode option for reaching their destinations. This finding is consistent with a recent study by Blumenberg, Brown, and Schouten (2020). They also found that car-deficit households (i.e., 0<VA<1) are more likely to use public transit. On the contrary, transit riders with zero cars, on average, make only 13.2% of the newly generated trips while comprising 10% of the population. As a result, transit investments will contribute more to increasing transit trips of car-deficit households. Moreover, families who own more than one vehicle per adult and live in suburban car-dependent neighbourhoods are less likely to shift their travel mode. It is also notable that there is no significant difference in the relative increase in new transit trips for different income groups. The percentage of new transit trips for different income strata has the same distribution as their population. It strongly suggests that increases in transit ridership resulting from improvements in transit accessibility come from car deficit households.

Table 6.2 : The expanded number of new transit trips generated for each class after transit accessibility improvement.

Strata	**Population**	**Accessibility Improvement**			
		50k	**200k**	**10%**	**25%**
VA=0	340,902 (10%)	12,611 (10%)	70,417 (12%)	9,344 (15%)	25,855 (16%)
0<VA<0.5	324,050 (10%)	20,532 (16%)	91,225 (16%)	9,134 (15%)	24,154 (15%)
VA=0.5	746,345 (23%)	41,103 (33%)	163,889 (28%)	21,713 (35%)	54,365 (33%)
0.5<VA<1	470,739 (14%)	24,265 (19%)	110,860(19%)	8,452 (13%)	21,002 (13%)
VA=1+	1,397,942 (43%)	26,633 (21%)	142,837(25%)	13,998 (22%)	36,963 (23%)
<$40k	531,406 (16%)	19,044 (15%)	104,732 (18%)	10,155 (16%)	28,538 (18%)
$40k-$60k	512,880 (16%)	21,229 (17%)	107,115 (18%)	10,490 (17%)	27,720 (17%)
$60k-$100k	873,514 (27%)	33,181 (27%)	144,396 (25%)	16,031 (26%)	39,979 (25%)
$100k-$125k	473,945 (14%)	18,519 (15%)	82,548 (14%)	7,764 (12%)	20,128 (12%)
>$125k	888,233 (27%)	33,171 (27%)	140,437 (24%)	18,201 (29%)	45,974 (28%)

6.4 Conclusion

In this chapter, the effects of accessibility improvement on transit use in the GTHA are evaluated. The key assumption is if transit accessibility is improved in low-income neighbourhoods, the low-income car-owning households will be more sensitive to these improvements due to getting rid of the financial burden of car ownership. They might be encouraged to use transit, leading to achieving simultaneously environmental and social goals.

Therefore, ZINB model is employed to estimate the number of transit trips made by individuals concerning their socio-demographic characteristics, local environments, and trip factors. Further, sensitivity analysis is conducted to explore the effect size of hypothetical transit improvements on individuals by different income and car ownership levels. The key findings of the analyses are summarized below.

(RQ2-1) To what extent can transit investments in lower socio-economic neighbourhoods enhance transit mode share?

Car-deficit households who have less than one car per adult ($0<VA<1$) are more inclined to use transit and generate newly transit trips if transit accessibility improves. As they share a car in a household and live in inner-suburbs where the level of transit accessibility is mediocre, transit investments provide them another mode option to meet their mobility needs. On the other hand, households in all income levels with more than one car per adult have less tendency to switch their travel mode. When they own a car, they want to use it even if transit infrastructures are improved. Among carless households, low-income individuals are still willing to take more transit trips and increase their existing transit trips after accessibility improvements, although they are already transit riders. On the contrary, a large amount of transit trips among high-income belongs to non-riders. It indicates that the low-income carless group still requires more transit improvements.

(RQ2-2) To what extent are low-income car-owners sensitive to transit improvements and shift their travel mode use?

Non-rides of households with one or more cars per adult are more sensitive to transit accessibility improvements and transit trip generations. They are more likely to use transit if transit services improve. However, in auto-centric areas with poor transit, the transit use of low-income households drops off sharply as car ownership increases. On the other hand, a sensitivity analysis suggests more opportunities for increasing transit ridership among car-deficit households when transit is improved. These findings indicate that improving transit in low-income inner suburbs, where most low-income car-owning households are living, would align social with environmental planning goals.

7. COMPARING STATISTICAL AND MACHINE LEARNING MODELS

7.1 Chapter Overview

This chapter presents the comparative analysis of two travel behaviour modeling approaches: statistical and ML models. It investigates how different ML techniques can improve prediction performance in transportation analysis projects. To this end, six of the most commonly used ML algorithms in travel behaviour studies, including DT, RF, XGB, NB, SVM, NN, and statistical methods such as LogR, LinR, ZINB, and Hurdle models as baselines are applied. Then, models are compared and the results of the analyses are reported in the following sections.

The travel behaviour of low-income households with a total income of less than $40k per year is evaluated in this study. As explained in Section 4.3.3, the dataset used for this analysis is divided into 10-folds using a stratified k-fold cross-validation approach to have an unbiased dataset. The performance of each model, fitted on nine folds and tested on one fold, is estimated and then averaged after ten iterations. Afterward, the recorded predictive performances of models are compared. This comparison is done in two parts: classification and regression tasks. The questions explored in this chapter are summarized below.

(RQ3-1) How accurate are ML models compared to traditional models in predicting travel behaviour in response to transit investments?

(RQ3-2) To what extent are ML models interpretable?

This chapter is organized into four subsections to share the results. Section 7.2 reports the findings of evaluating and comparing algorithms' performance on binary classification, i.e., predicting the probability of taking transit by individuals. The results of regression, i.e., estimating the number of transit trips taken per person are provided in Section 7.3. Afterward, the sensitivity of the classical and ML models to transit improvements is explored, and findings are explained in Section 7.4. Finally, the upshot of evaluating and comparing algorithms' performance is discussed in Section 7.5.

Figure 7.1 illustrates a detailed diagram outlining the steps undertaken and the dataset used in this section. A three-step approach is adopted. First, the predictive performance of each algorithm using various evaluation measures is estimated. To statistically examine the significant difference in each performance among all algorithms, a Friedman Aligned ranks test is used. On top of that, Bergmann-Hommel *post hoc* analysis is employed to make a pairwise comparison between models (Derrac et al., 2011). In the second step, a sensitivity analysis is utilized to explore how a model selection may influence different predictions, spatial distribution, and planning policies. To check the difference in the spatial pattern of the predicted new trips, the SPAtial EFficiency metric is applied for each map (Demirel et al., 2018). Finally, the feasibility of the interpretability of ML models is measured by applying global and local model-agnostic interpretation techniques.

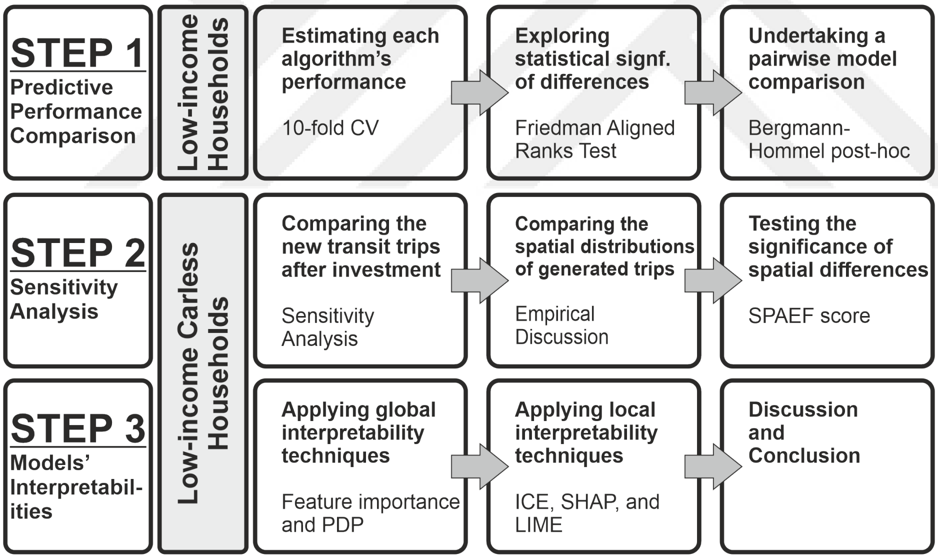


Figure 7.1 : The experimental design of the study.

7.2 Comparing Models on Predicting the Probability of Taking Transit

In this analysis, individuals having at least one transit trip in their trip chain are classified in the C_1 class, whereas C_0 belongs to non-transit users. To predict the probability of taking transit, six ML models as well as LogR as the baseline are employed. The dataset used for this step consists of travel behaviour of low-income households. Each model's predictive performance is estimated using 10-fold cross-validation technique discussed in subsection 4.3.3. Comparing the average performance of each algorithm, Table 7.1 shows that RF achieves the highest predictive accuracy of 89.56%. The models perform almost identically in forecasting the majority class, i.e., C_0 (non-transit users), by having an acceptable score for precision, recall, and F1-Score. However, there is a significant difference in predicting the probability of taking transit, i.e., C_1 (transit users). For instance, LogR performs poorly in recalling transit riders (70.57%) by overfitting non-transit users. This underestimation also applies to NB, considering its low performance in terms of the F1-score of the minority class. Overall, RF outperforms others across various performance metrics. It has a balanced performance for both the majority and minority classes. Moreover, it does not sacrifice the recall for having high precision. In terms of the AUC-ROC curve, still, RF has the highest performance, meaning that considering different thresholds for T_0 and F_0, it captures the behaviour of non-transit users efficiently.

To investigate the significance of performance differences, statistical tests are employed. Accordingly, multiple statistical comparison tests for the folds are selected (García et al., 2010). Friedman Aligned Ranks as a non-parametric test was chosen since the sample size is small. The null hypothesis of the test is that there is no significant difference in the algorithms' performance. The p-value of the test for the obtained accuracy measure is equal to $6.659e-10$ given ten folds. Hence, the results show that differences in algorithms' accuracies are statistically significant (for $\alpha = 0.05$). The same test is applied for all metrics, and similar findings are observed.

To explore the source of the difference, a *post hoc* analysis is applied. The *post hoc* procedure assesses the difference between all algorithms in terms of the absolute difference of the average ranking. It enables us to have a pairwise comparison among models. Moreover, adjusted p-values are computed using the *post hoc* procedure. Figure 7.2a is the matrix of corrected p-values for accuracy measure after applying the Bergmann-Hommel *post hoc* analysis. Dark colors show higher adjusted p-values, representing an insignificant difference between pairs of algorithms in terms of the accuracy metric. Accordingly, RF is ranked 1.4 on average, and based on the p-values, there is not enough evidence to confirm its outperformance compared to XGB and NN (see. Figure 7.2b). However, they are statistically better than the remaining algorithms, including the LogR, which ranked six on average. A detailed discussion

Table 7.1 : Performance comparison of each classifier for predicting the probability of taking transit in individuals' daily trips based on 10-fold cross-validation.

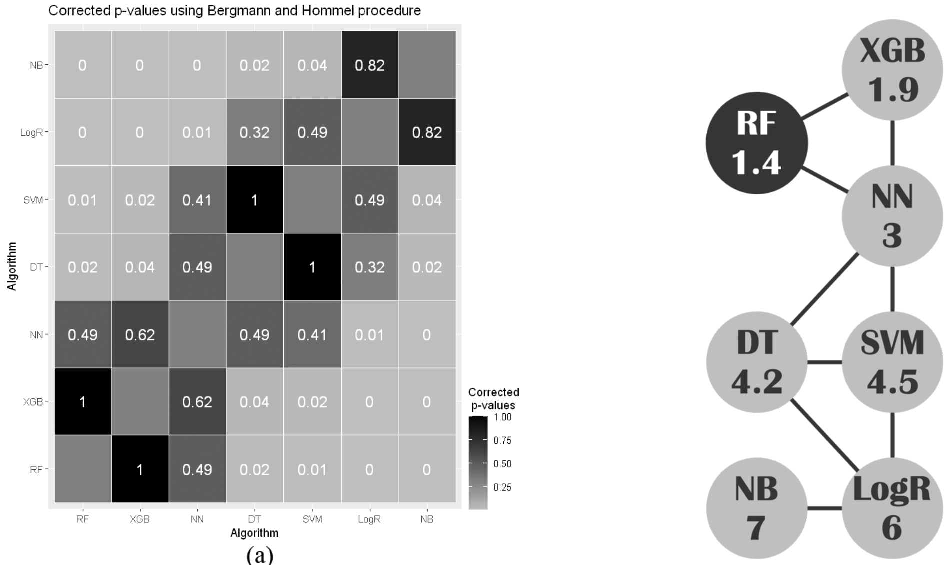
		DT	RF	XGB	NN	SVM	NB	LogR
Accuracy (%)	$\bar{X}$	88.12	**89.56**	89.33	88.75	88.02	83.93	86.07
	s	0.54	0.47	0.82	0.63	0.66	0.59	0.69
Precision$_{C_0}$ (%)	$\bar{X}$	91.27	92.00	91.81	**92.16**	90.59	89.82	89.31
	s	0.61	0.44	0.66	0.96	0.64	0.67	0.61
Recall$_{C_0}$ (%)	$\bar{X}$	92.54	**93.82**	93.70	92.44	93.23	87.89	91.86
	s	0.85	0.76	0.93	1.68	0.60	0.60	0.84
F1-Score$_{C_0}$ (%)	$\bar{X}$	91.90	**92.90**	92.74	92.28	91.89	88.84	90.56
	s	0.39	0.34	0.57	0.51	0.44	0.40	0.48
Precision$_{C_1}$ (%)	$\bar{X}$	79.31	**82.58**	82.21	79.77	80.37	69.36	76.54
	s	1.68	1.64	2.13	2.89	1.42	1.03	1.92
Recall$_{C_1}$ (%)	$\bar{X}$	76.28	78.15	77.61	**78.90**	74.06	73.32	70.57
	s	1.95	1.40	1.99	3.11	1.98	2.01	1.95
F1-Score$_{C_1}$ (%)	$\bar{X}$	77.75	**80.29**	79.82	79.24	77.07	71.27	73.37
	s	1.10	0.81	1.50	0.91	1.36	1.22	1.33
AUC-ROC (%)	$\bar{X}$	90.68	**94.37**	94.23	94.26	93.49	88.55	91.75
	s	0.70	0.42	0.55	0.46	0.49	0.58	0.59

*C_0 is the class of non-transit users (majority class), and C_1 is the class of transit users (minority class).

of the differences between ML classifiers and the traditional algorithms is provided in Appendix B.1.

7.3 Comparing Models on Estimating the Number of Transit Trips

To predict the number of transit trips, five ML regressors and three statistical models as the baseline are utilized. In this step, Hurdle and ZINB as count models are employed for the comparison. The reason for this decision is an excessive number of non-transit users in the dataset. LinR, another statistical method, is used as the baseline. Table 7.2 illustrates the average performance of each regressor across ten folds. Tree-based algorithms, i.e., RF, DT demonstrate a smaller error in predicting the number of transit users. The results show that RF alongside NN and XGB has a higher R-squared value on average. Considering RMSE and RRSE, the conclusion is the same; however, DT is the one that outperforms all other algorithms in terms of Median Absolute Error.



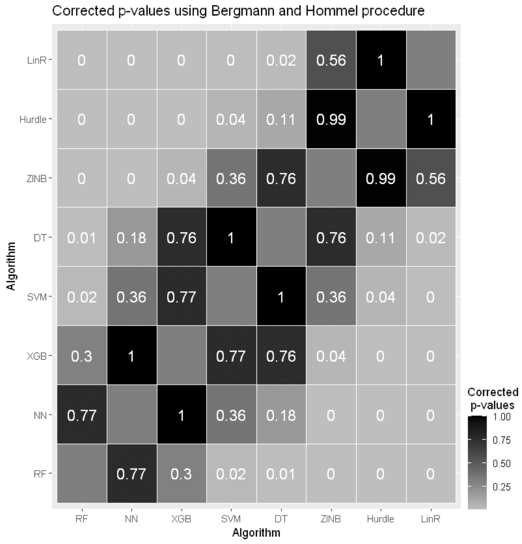
(a) (b)

Figure 7.2 : Friedman test result for classification "accuracy" after Bergmann-Hommel *post hoc* procedure. (a) Corrected pairwise p-values using Bergmann-Hommel *post hoc* procedure. (b) Average rank of classifiers ($\alpha = 0.05$). Edges between algorithms indicate an insignificant difference.

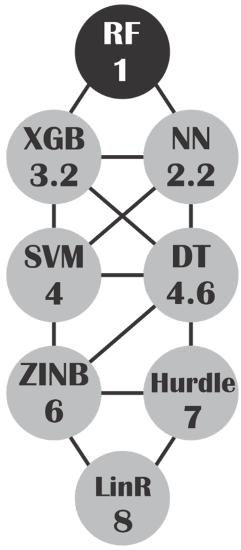
Table 7.2 : Performance comparison of each regressor for predicting the number of transit trips in individuals' daily trips based on 10-fold cross-validation.

		DT	RF	XGB	NN	SVM	LinR	ZINB	Hurdle
R-Squared (%)	$\bar{X}$	52.33	**58.22**	53.90	55.37	53.06	46.74	49.89	48.11
	s	1.84	1.10	1.42	1.50	1.55	1.61	2.02	2.16
RMSE loss (%)	$\bar{X}$	69.46	**65.04**	68.31	67.22	68.93	73.43	71.21	72.46
	s	1.30	1.43	1.46	1.64	1.67	1.63	1.48	1.62
MDAE loss(%)	$\bar{X}$	**6.82**	11.60	12.14	9.86	8.09	20.33	10.97	14.68
	s	0.26	0.90	0.68	1.46	0.21	0.84	0.72	0.55
RRSE loss(%)	$\bar{X}$	69.03	**64.63**	67.89	66.80	68.50	72.97	70.77	72.02
	s	1.33	0.86	1.05	1.13	1.13	1.11	1.43	1.50

Figure 7.3 depicts the pairwise comparison of regressors. The values greater than 0.05 in Figure 7.3a indicate an insignificant difference between algorithms. For instance, according to the Bergmann-Hommel *post hoc* test, there is no evidence for a statistical difference between RF and XGB in terms of RMSE values. On the other hand, Figure 7.3b shows only the insignificant pairwise difference among algorithms. The number on each node indicates the average rank of the algorithm given ten folds. Accordingly, RF has the lowest average rank of 1 and is on a par with XGB and NN in terms of RMSE. More details on the differences between ML models and the traditional baseline for other metrics are provided in Appendix B.2.



(a)



(b)

Figure 7.3 : Friedman test result of regression "RMSE" values after Bergmann-Hommel *post hoc* procedure. (a)Corrected pairwise p-values using Bergmann-Hommel *post hoc* procedure. (b)Average rank of regressors ($\alpha = 0.05$). Edges between algorithms indicate an insignificant difference.

7.4 Sensitivity Analysis

To evaluate transit investment policy in low-income communities, the effect of transit accessibility improvements on taking transit using sensitivity analysis of all regressors is investigated. For this work, low-income carless households as a subset of the dataset are selected to compare the predictive outcomes of all models for the disadvantaged group. Figure 7.4 shows the average number of new transit trips after incrementally increasing accessibility for low-income zero-car individuals. The ZINB model is selected as the baseline model, coloured red in Figure 7.4. Since it is an interpretable, less complex, and widely used model in the literature, it is considered a reference model. Accordingly, the ZINB's predictions with those of other models are compared. Among all regressors, the ZINB model predicts the average number of new transit trips after increasing accessibility to 200k jobs by more than 0.2 per person. Comparing the variations in the predicted trip numbers using statistical models represents that all ZINB, LinR, and Hurdle models follow a roughly identical trend. However, the average number of predicted transit trips after increasing accessibility to 200k jobs using LinR and Hurdle models is half of the ZINB ones. Comparing the results of the SVM algorithm with those of the reference model, it can be seen that the curve of newly generated transit trips using the SVM model increases up to 100k jobs. Then, it starts decreasing until it reaches 0 in 200k jobs. Such behaviour is also not observed by other models. Moreover, the predictive behaviour of ensemble models like RF and XGB is similar up until 60k jobs, but their prediction lines diverge beyond that. It concludes that there is a difference among all models in the average number of new transit trips after the largest accessibility gain (i.e., 200k jobs). Therefore, model selection may impact policy evaluation. To choose the optimal model, a researcher may consider the predictive performance and the interpretation of features contributing to that prediction. Before discussing possible interpretations of the best-performing model, this study explores how these prediction differences are spatially distributed in the region.

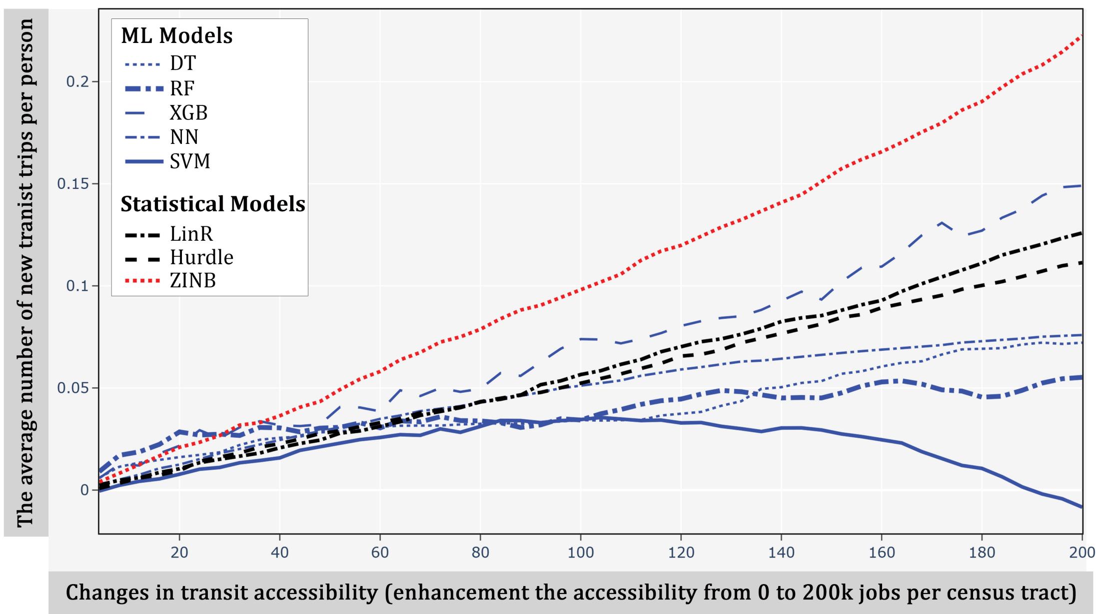


Figure 7.4 : The sensitivity of all models to the accessibility improvement for low-income carless households (The baseline model is shown in red, the statistical models in black, and the ML models in blue.

The spatial distribution of newly generated transit trips by low-income carless households after increasing the transit accessibility level by 200,000 jobs is mapped (See Figure 7.5). This accessibility gain explained by previous work (Farber & Marino, 2017) can be achieved by investments in higher-order transit services. I created 1000x1000 m hexagonal maps to investigate whether there is a spatial similarity between models' predictions of new transit ridership. The dark blue hexagons display the highest number of individuals with increased transit trips, while the light orange ones define that no individual inclines to increase their transit trips. The maps present a clear visual distinction in the spatial distribution of new transit trips predicted by each model. For instance, the statistical models have small numbers of new transit trips in the Hamilton, Brampton, and Newmarket regions, whereas ML models suggest new transit trips after accessibility improvement in the same regions. In a special case of SVM, all the new transit trips belong to the inner suburb of Toronto, and a negligible number of transit trips in Downtown Toronto is observed. This observation is consistent with Figure 7.4 in which SVM shows a decline in the number of new transit trips if the accessibility is significantly improved. Based on both Figures 7.4 and 7.5, only XGB have a similar number of transit trips and spatial patterns to those of the statistical methods.

Besides the visual evaluation, the study aims to statistically compare whether there is a significant difference between the spatial patterns of all maps. For this reason, the SPAtial EFficiency metric (SPAEF) (Demirel et al., 2018) is applied. This metric considers three statistical measures, including Pearson correlation, coefficient of variation, and histogram overlap, and their outputs are integrated into one measure. SPAEF values calculated for each map are reported in Figure 7.6. The high scores of SPAEF in the right bottom of this heatmap illustrate that all three statistical models have high spatial similarities. On the other hand, they show a different spatial distribution than that of to NN, DT, and SVM. Discarding the similarity of RF and XGB to traditional approaches, you can see a dark 5×5 cluster of ML models at the top left and another dark 3×3 cluster of traditional models at the bottom right. Thus, it concludes that utilizing ML algorithms instead of traditional models may suggest different spatial patterns of transit use after accessibility improvements (see Figure 7.4) and may result in a different spatial policy recommendation at the end (see Figure 7.5).

In terms of the equity implications of selecting a proper model, the results show that planners and policymakers may overlook some low-income carless households, living in suburbans and having the tendency to take transit if it is improved. The spatial distribution of transit trips predicted by traditional models shows a lower number of transit trips in some regions. This may signify the least return on investments in transit projects in those areas for planners. Accordingly, transportation planning authorities will probably give less priority to expanding transit networks in those regions and intervene in fulfilling the transit needs of groups at risk of transport disadvantage. Also, the gap between the activity participation rate of car-owners and carless families will

Figure 7.5 : The spatial prediction of newly generated transit trips by low-income carless group after improving job accessibility (200k jobs) using different algorithms.

remain. Although some policymakers may suggest facilitating the access and ownership of a private car for low-income carless households, it is neither an environmentally nor financially efficient solution.

7.5 Model Interpretation

In selecting a predictive algorithm, the interpretability of a model can be as important as the model's predictive performance. The feasibility of the interpretability of ML models is discussed to investigate whether there is indeed a trade-off between predictive

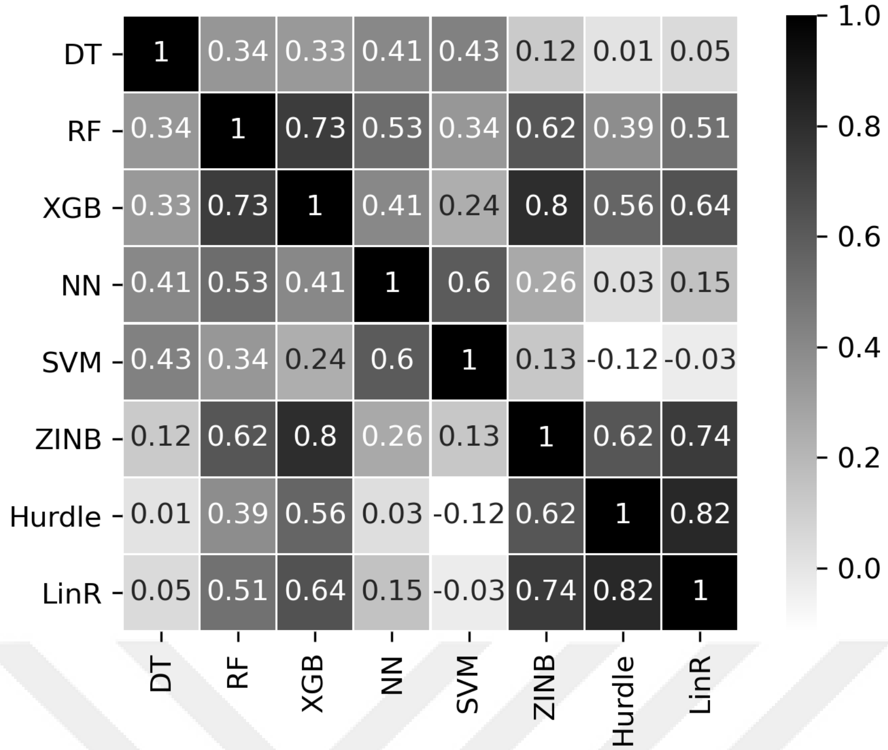


Figure 7.6 : The heatmap of SPAEF scores.

performance and interpretability of models. Accordingly, the model-agnostic is used as an interpretability approach; therefore, this process can be generalized to any other ML or even traditional model. In this section, global and local interpretability tools are applied for the best performing ML model, i.e., RF according to its predictive performance to understand how a model predicts and which factors and to what degree contribute to its prediction. Also, the number of transit trips is taken as an independent feature. This experiment can be replicated for the classification task, i.e., predicting the possibility of taking transit.

7.5.1 Feature importance

Given the Feature importance interpretation technique, the total effect of each variable on the final outcome is computed. Accordingly, Figure 7.7 shows the influence of each independent variable on having transit trips for the low-income carless stratum. It denotes that the most significant variable in predicting the number of transit trips is the mandatory trip length. Also, transit accessibility is among the most important variables confirming that this measure is strongly associated with activity participation. However, it does not show whether this feature affects the output positively or negatively. The results show that driving licence possession, the free parking spot at the workplace,

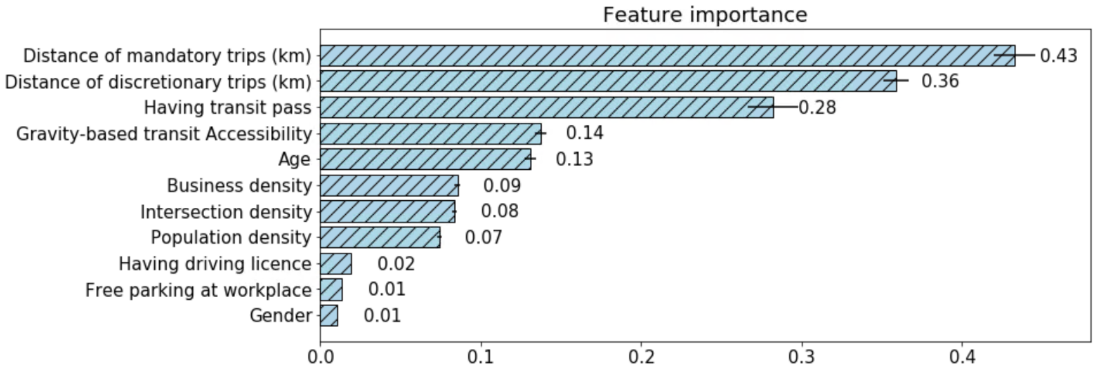


Figure 7.7 : The importance of each feature for predicting the number of transit trips using RF model.

and gender have the lowest score, i.e., the least importance in transit trip prediction. It is aligned with our expectations of the variables.

7.5.2 Partial dependency plot (PDP)

To determine the partial link between each attribute and the targeted variable, PD plot is generated. Furthermore, ICE plot are draw to understand the impact of each feature on the final prediction for individual instance. PD plot (red lines) in Figure 7.8 illustrates how the variations of each independent variable affect the expected average number of transit trips for vulnerable individuals. It also shows the direction of the effect between independent features and the dependent variable. The disadvantage of this technique is that it assumes there is no correlation between independent variables, whereas this assumption is often inaccurate in the real world. For instance, there might be some correlation between the number of people living in a region and its business density. However, PDP fails to show the mutual impact of these two variables. This plot indicates that the number of transit trips increases as either mandatory or discretionary trip lengths increase. However, it shows that gender, free parking at the destination, and driving licence do not influence one's number of transit trips.

7.5.3 Individual conditional expectation (ICE)

The ICE as a local interpretation technique is also applied to see the effect of each variable on the outcome of each observation. Each ICE line (blue lines) in Figure 7.8 represents how the dependent variable changes when an independent feature changes for observation, while the PDP line defines the average of the line of an ICE plot (Molnar, 2020). This change in the dependent variable, e.g., the number of trips, is estimated by keeping other attributes intact and incrementally increasing the specific feature for

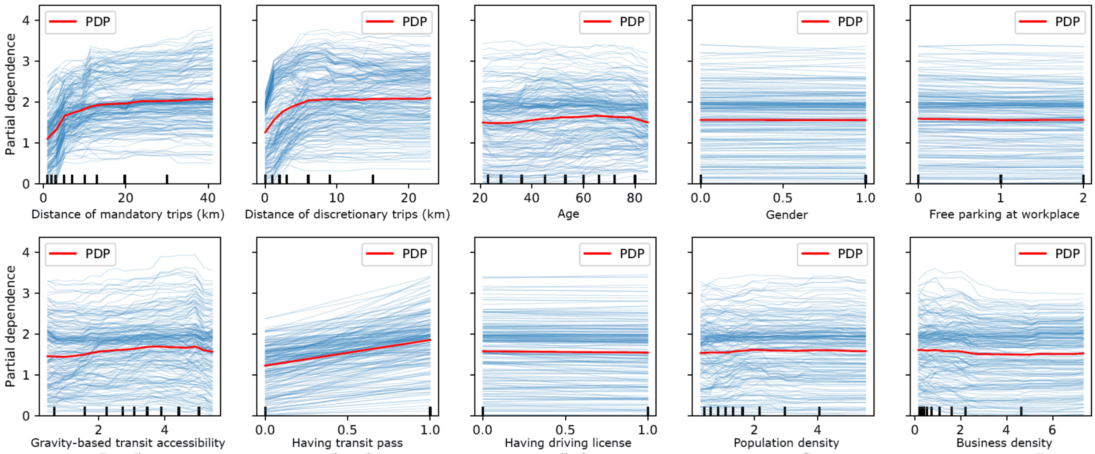


Figure 7.8 : ICE plot and PDP for predicting the number of transit trips after increasing transit accessibility by 200k jobs (Blue lines indicate ICE plots).

a single observation, e.g., a single person. The plot shows 200 random individual observations within the dataset and depicts how the prediction of the number of transit trips changes as the independent variables change (blue lines). Unlike the PD plot showing the average effect of an independent variable on the output, the possible anomalies in the ICE plot are seen. For instance, although the longer mandatory trips are, the more transit trips are taken, the results show some individuals for whom increasing the length of trips decreases their transit use. It is expected for suburbs with lower access to transit to use their own vehicle when their trip length is high. These individual-level findings cannot be obtained by merely checking the coefficients of a statistical model. Notably, these local interpretation tools are generalizable to any other statistical or ML algorithm, facilitating the interpretability of any model with higher granularity.

7.5.4 Shapley value (SHAP)

SHAP is applied to see how each feature affects the prediction of a single observation in different coalitions (see Figure 7.9). In this study, Shapley values show the average contribution of each feature to the predicted number of transit trips across all possible coalitions of features, including and excluding this feature value. Therefore, it is useful when the contributions of features are unequal, but they may affect each other. The sum of Shapley values for all attributes of an individual is equal to the predicted number of trips for oneself subtracted from the mean predicted number of trips for everyone.

Figure 7.9 illustrates that the length of the trips, whether mandatory or discretionary, together with having a transit pass contribute the most to the number of transit trips

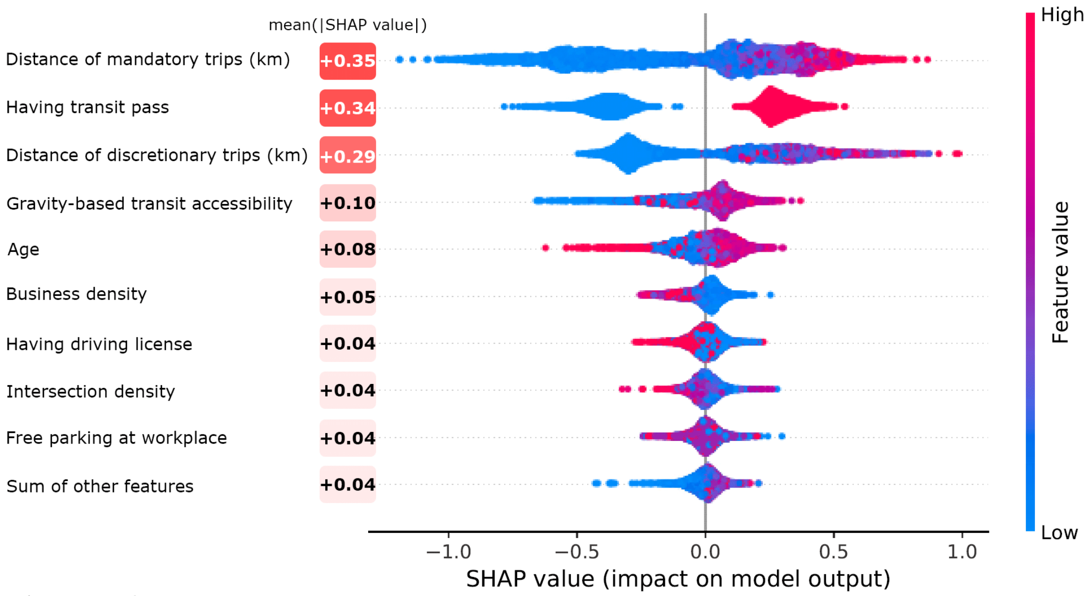
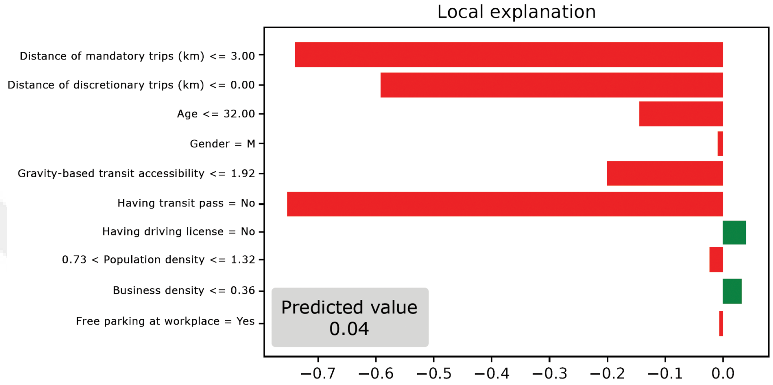


Figure 7.9 : SHAP values' distribution and mean. Features are sorted by their mean SHAP values.

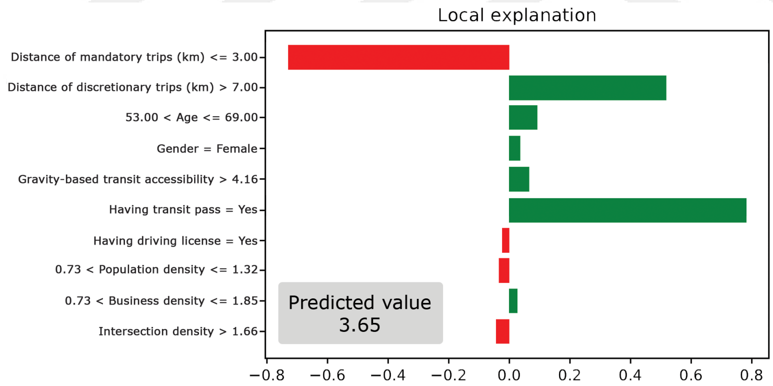
for low-income carless people – i.e., they have the highest mean SHAP value. The density of the length of the mandatory and discretionary trips shows how common different trip lengths are in the dataset, and the coloring indicates a smooth increase in the log odds ratio of the transit use as the trip length increases. For the transit pass possession, unsurprisingly, two clear clusters are observed: people owning a transit pass have a higher number of trips and vice versa. A longer tail to the left for transit accessibility means that living in low transit-accessible regions, e.g., suburban, can significantly reduce the number of transit use, but high accessibility does not necessarily significantly raise the number of transit trips either. For instance, in downtown, where biking and walking to destinations are convenient and at the same time transit is accessible, low-income individuals may prefer active transportation.

7.5.5 Local interpretable model-agnostic explanations (LIME)

To better understand the predicted values of a specific individual, the LIME tool is utilized. Figure 7.10 shows two randomly selected individuals, one who does not use transit and one who has five transit trips a day. LIME can be used to explore the notion behind the predicted values for a specific user. As the author of the original paper mentioned, it is also a way to check a model's trustability (Ribeiro et al., 2016). In this study, two extreme cases considers to see how each feature contributes to the final prediction. The y-axis shows the condition that holds for the feature value, and the x-axis shows the feature effect, i.e., its weight times its actual value. Figure 7.10a for the non-transit user shows that not having a transit pass, having short trips, being middle age, and having low transit access lead to his preference for other travel modes. Figure 7.10b for the frequent transit user indicates that her high number of discretionary trips, better transit accessibility, having a transit pass, and being older have an impact on her frequent transit trips. These detailed observations per individual are only possible when local interpretability was used. Therefore, besides the global interpretability of the models, a higher granularity of the interpretation sheds light on the model decisions and the soundness of its predictions.



(a)



(b)

Figure 7.10 : Two sampled individuals, one not using transit and one a frequent transit user. (a) A non-transit user. (b) A transit user.

7.6 Conclusion

The model selection which can accurately predict the travel behaviour responses to transport infrastructure changes is studied in this chapter. Since accurate travel behaviour models can affect travel demand management and transport policy-making, the potential of using ML algorithms to understand complex relationships between variables is examined.

Six of the most commonly used ML methods in travel behaviour studies and five statistical algorithms are selected to explore the differences in their predictive performances. The performance of each algorithm in estimating the number of transit trips after accessibility improvements for marginalized populations are analyzed. The main results and conclusion in response to the research question discussed before are summarized as follows.

(RQ3-1) How accurate are ML models compared to traditional models in predicting travel behaviour in response to transit investments?

The performance of different traditional and ML models in predicting the number of transit trips and the possibility of being a transit user is explored. These two tasks can be called regression and classification problems, respectively. Using Friedman's test with the Bergmann-Hommel *post hoc* procedure, the results show that Random Forest, XGBoost, and Neural Networks significantly outperform others in both regression and classification tasks. Random Forest as an ensemble method, with its ability to capture non-linear rules in the datasets, achieves the best rank among all other algorithms. Statistical models, mainly used in the literature, have lower performance in predicting the behaviour of transit users.

On the other hand, the sensitivity of the individuals to transit improvements may be interpreted differently if an inaccurate model is used. Therefore, the model choice has a significant impact on the suggested policy. Even the spatial distribution of the newly generated transit trips is different when traditional models are employed. Accordingly, utilizing the best-fitted model based on different performance metrics is recommended when proposing a policy.

(RQ3-2) To what extent are ML models interpretable?

Comparing the predictive performance of models through learning from data is of importance. On top of that, interpretability provides insights helping researchers realize how a model arrives at its accurate conclusion. Given the growing importance of ML interpretability, several model-agnostic interpretation methods are discussed in the study. These tools are flexible and can be generalized to any model type. Using both local and global interpretability of the ML models, this study shows that, despite the fallacy of calling them a black box, the model's predictive decisions can be demonstrated using numbers and figures. The features rank according to their importance, the direction and the significance of each feature in predicting transit use are shown, and the numeric detailed analyses per individual are provided. Thus, these model-agnostic interpretation tools can capture the notion behind each model's decision and have a balance between their performance and interpretability.

8. CONCLUSIONS

8.1 Chapter Overview

This chapter concludes the thesis and draws together arguments and results developed through the analyses. Section 8.2 provides a summary of the thesis and the main implications of the findings. The contribution of the dissertation to travel behaviour and mode choice decision domain are discussed in Section 8.3. Section 8.4 discusses policies to develop an equitable transportation system and reduce transport poverty. Finally, Section 8.5 highlights several limitations in conducting this study and suggests some directions for future works.

8.2 Thesis Summary

The primary role of transportation planning is to enhance individuals' activity participation to meet their mobility needs and reach their destinations. Consequently, understanding how different socioeconomic groups plan their daily trips and what variables impact their choices and preference are useful for travel demand management and transportation planning.

Accordingly, this thesis investigates the activity patterns, trip schedules, and travel mode decisions of individuals in the Greater Toronto and Hamilton Area (GTHA) using multiple models and tools for obtaining a robust and reliable estimation. The research aims of the study, background information on the research, and research questions formulated into three main parts are discussed in Chapter 1. Chapter 2 presents an overview of the literature with a focus on main research themes: transport equity, travel behaviour of travellers, trip chain analysis, traditional and equity-based transit investments approaches, the use of statistical and ML models in travel behaviour and mode choice studies. In Chapter 3, the study area and dataset used for the analyses are introduced. Chapter 4 explains the structure of models, algorithms, and tools utilized for conducting the research in the next three chapters.

In Chapter 5, travel patterns of low- and high-income carless and car owner households are investigated using a clustering framework for their trip sequences. Given their activity type and mode used, four distinct clusters are extracted for each group. Further, the impact of socioeconomic characteristics and built-environment attributes in structuring different travel patterns are comprehensively examined. In response to transit use rate, the results show that females, regardless of income and car-ownership levels, are the main transit riders whenever transit accessibility is appropriate. Among carless households, low-income women more often take public transit than their counterparts, although they live in neighbourhoods with low density and poor transit accessibility. Notably, low-income senior women are at risk of transport poverty because they largely depend on others to make their daily trips. Their activity patterns illustrate that they are a passenger of either their relatives or taxis/paid rideshare travel modes for their daily trips. Unlike the low-income people, wealthier carless individuals locate in high-density regions with acceptable accessibility to transit services, and land uses. Their mode frequency shows that they make 48% of their daily trips by transit and 42% of them by walking and cycling (see Table 3.4). Accordingly, low-income neighbourhoods of inner suburbs require to get the highest priority in transit planning or investments. The integration of transportation and land use planning in low-accessible places allows more trips to be made and increases activity participation rates. Density and mixes of land uses can minimize the number of car trips and support transit and active transportation use. Also, the housing affordability crisis can be addressed through planning or introducing changes in zoning regulations to force developers to build more affordable houses.

Car-owners who are populated households with four or more people tend to live in remote neighbourhoods with low levels of accessibility (e.g., suburbs) and drive a car to complete their daily trips. In low-income households, women are still passengers rather than drivers even after owning a private vehicle. They commute to work or school by transit when they are middle-aged and are located in places with appropriate transit accessibility. It supports the previous findings that females less benefit from accessing a car in households with a shared private vehicle (car-deficit households) (Naess, 2008; Madariaga, 2016). Furthermore, most women make most non-work lengthy trips due to carrying a disproportionate burden of household responsibilities and caring tasks (Madariaga, 2016; J. Lee et al., 2018; Craig & van Tienoven, 2019). Contrary to previous studies, the longest trip chains are, on average, related to car-based trips, and the least number of trips are made by public transit. A private vehicle or taxi user may chain more destinations per their daily tour. It suggests that using a car provides higher flexibility in trip-making behaviour. Interestingly, the most multi-modal travel pattern has a high number of trips and a lengthy trip chain. This pattern is particularly observed when public transit and active transportation are combined in places where transit accessibility levels are high. This is evidence of the successful integration of transit infrastructure and land use planning in downtown regions. On average, high-income households use multiple modes and more trips compared to their counterparts when they are zero-car families. This finding implies that in the lack of access to a private car, providing multi-modal mobility options may improve the mobility of low-income households.

Chapter 6 focuses on understanding how transit investments in low-income neighbourhoods might affect transit use of households and change their mode use behaviour. The hypothesis was low-income carless households may be largely insensitive to transit improvements, the so-called “captive transit” users for whom transit investments may not result in large environment or congestion co-benefits. The case for low-income car-owning households is less predictable, with competing arguments suggesting either:

(a) these households will be more sensitive to accessibility improvements since the costs of car ownership and use are high, and many could benefit from using transit rather than car if service levels were improved; or

(b) these households will be insensitive to accessibility improvements since once owning a car, it usually provides a reduced marginal cost of travel, and households with limited financial resources will not opt to pay for transit if they can drive places for “free” (or rather at a low marginal cost).

In response to this dichotomy, the finding show that transit use is more sensitive to transit investment in households that own cars, and most sensitive in low income, car-owning households (with an elasticity almost equal to 1). It is strong evidence in support of (a) that for non-transit rider low-income households, owning a car is a financial burden, and increased transit provides increased opportunity for mobility and transit use. At the same time, however, the results show that the tendency for individuals in low-income households to take transit reduces dramatically as soon as their household owns a private vehicle, whereas it drops more gradually for high-income households. This is evidence in support of (b), indicating that low-income car-owning households can become extremely car reliant, and even less multi-modal than their wealthier car-owning comparators.

In both cases, the findings are supported by spatial analysis of where different population strata live vis a vis existing levels of transit supply or relative environmental dominance of the automobile. Wealthier carless households tend to concentrate in neighbourhoods where existing transit accessibility levels are high. This pattern becomes complicated as income levels decrease. Low-income, carless households, are more dispersed in lower-accessibility areas compared to higher-income, carless households, who are more concentrated in the very core of the city.

Finally, since there is evidence to support that both (a) and (b) are true, the simulation analysis can provide some insight into how these forces combine to result in an overall transit ridership response in the region. The simulations, however crude, apply accessibility gains to individuals across the region to determine among which population groups the largest increase in transit use is seen. Here, the evidence is quite clear; more new trips are predicted to be made by households with less than one car per adult. There may indeed be more opportunity for increasing transit use overall by targeting car-deficit households. Conversely, accessibility improvements in areas where the accessibility gap between transit and car is large, and a significant number of car-deficit low-income households are residing, would be an effective way of both increasing transit ridership and improving equity. Nonetheless, this does not preclude the possibility of first investing in targeted areas where low-income, car-owning, and transit sensitive populations are presently residing. As seen in Figure 3.2, this population is mostly living in Toronto's inner suburbs, indicating that both social and mode-shifting goals can be achieved if investments were made there.

The finding of the thesis provides evidence that housing policies in coordination with transportation policies are essential to facilitate the transit accessibility of poor households (Pucher & Renne, 2003; Kramer, 2018). Housing policies that provide affordable housing in areas with higher levels of accessibility will be a strategy to prevent low-income households from incurring the financial burdens of car ownership. On the other hand, built environmental variables such as mixed land uses, walkable street networks, dense neighbourhoods, safety in neighbourhoods, employment, retail densities, and so on influence transit ridership. Accordingly, it is suggested that planners considering neighbourhood characteristics in evaluating local scale projects.

Chapter 7 builds on the argument of Chapter 6 and uses the same dataset for evaluating the prediction performance of different transportation models. This Chapter compares the application of traditional and ML methodologies in exploring how different people respond to different types of changes, transit accessibility improvements, in their land-use environment. Comparing the predictive performance of all models showed that Random Forest (RF) classifier and regressor are the most accurate methods for modeling the transit demand of low-income households. The higher value of F1-score for both non-transit and transit user classes indicated that the RF model has high robustness and precision. Additionally, its R-squared value on average was significantly higher compared to all other models. Conversely, traditional models, e.g., the ZINB model or LogR, showed a statistically lower performance based on both threshold-dependent and -independent metrics. From the equity-based perspective, the impact of transit investment, in terms of accessibility improvement, on the potential transit trips by low-income individuals was examined. The sensitivity of the models to the accessibility gains was significantly different. It showed a 17% difference in the number of predicted new transit trips across the models tested. Undoubtedly, any transit plans or policies framed by each model will have different equity impacts on low-income communities. Afterward, the newly generated transit trips across all models are mapped to examine the spatial distribution of transit trips in the region. The maps showed a heterogeneous spatial distribution of new transit trips by vulnerable groups among traditional and ML models. For instance, ML models proposed a potential for the new transit trips in Hamilton city – i.e., potential investment in that region. However, statistical models did not demonstrate a significant number of transit trips for the same region.

Further, the global and local interpretability of the best-performing model, i.e., RF was explored. Five different model-agnostic tools were utilized to investigate the effect of each feature on the number of predicted transit trips in two levels of granularities – e.g., group and individual interpretability. The length of the trips, having transit pass, and accessibility to public transit have the most impact on transit use. Throughout the analysis, findings suggest ML models are both accurate and interpretable via external interpretability tools. Based on the experiment, utilizing a model that is both accurate and rational is recommended– that is, its global and local interpretation can be supported by pre-existing knowledge or theory.

8.3 Thesis Contributions

As summarized in Section 8.2, this thesis thoroughly evaluates the travel behaviour of low- and high-income carless and car-owner individuals with a focus on their trip chains and their sensitivity to transit investments. Furthermore, it evaluates ML applications in transportation mode choice analysis to develop equitable transport policies and reduce transport poverty. The major contributions of this thesis to travel behaviour studies are listed below.

- In Chapter 5, this thesis contributes to the literature in at least four ways. First, previous studies have utilized predefined rules and arbitrary assumptions for trip chains. This study alleviates the subjectivity issue of rule-based approaches by leveraging presumption-free sequence clustering. Second, previous works looked at the trip purpose and mode choice as two separate variables. However, this study aggregates each trip's activity types and mode choices to construct trip sequences and understand travellers' behaviour. Third, this work considers all non-work activities separately, noting other studies tend to unify all non-work trip purposes into a single group. This decision provides deeper insights into different non-work activity types. Fourth, to the best of our knowledge, this is the first study to comprehensively analyze all possible trip chains, classify travellers' mobility patterns in terms of their trip destinations and mode use simultaneously, and compare travel patterns of populations in the GTHA.

  In sum, this study investigates travel patterns of residents by a cluster-based framework considering activity type and travel mode simultaneously in chaining trips in the context of transit equity. This approach provides insights into trip chain sequences, interdependencies, and the activity types related to mode choice behaviour.

- Concerning the arguments provided in Section 2.4, this thesis explores whether transit investments in low-income neighbourhoods are likely to result in increased transit use, thus congestion and environmental co-benefits, as well as reducing socio-spatial inequalities for disadvantaged communities. To this end, Chapter 6 of this thesis empirically evaluates two contradictory arguments mentioned in the literature. First, if owning a private vehicle bears a substantial burden on low-income households, they are expected to display more sensitivity to transit accessibility improvements and to be more likely to switch their mode from car to transit. Second, if low-income households are either already transit users or reluctant to shed their car after their sizeable investment, improving transit accessibility in low-income neighbourhoods will not necessarily be associated with mode shifting. Notably, auto ownership cost for a family includes expenses for car purchase, lease, loans, fuel, insurance, maintenance, parking, and so forth. However, our dataset does not include this information. This study reveals that car-deficit non-rider households are more likely to take transit trips after increasing transit accessibility. Among carless groups, existing transit riders of low-income carless households are encouraged to take more transit trips.

- Chapter 7 presents the methodological contribution of this thesis. The feasibility of using ML algorithms, whether classifiers or regressors, is explored compared to statistical models in predicting the number of transit use. In the comprehensive comparison, different predictive performance metrics are computed, and the interpretability of ML models using both global and local interpretability techniques is evaluated. To investigate whether model selection would affect the justice-based interpretation of the scenario, a subset of the dataset, i.e., low-income carless individuals, is chosen. To this end, a scenario with enhanced transit accessibility throughout the region is tested to see how people living in low-income households respond to transit investment policy. Further, the spatial distribution of the forecasted transit trips after transit improvements in the region is compared.

8.4 Policy Recommendations

The main purpose of this thesis is to improve individuals' activity participation by identifying their trip decisions to reduce inequalities in access to key destinations and limit travel barriers for all residents, particularly those who are at the risk of transport poverty. The findings in the previous chapters reveal that there is still a special need to develop transport policies in Toronto for improving public transit to facilitate the activity participation of residents in low-access neighbourhoods. These policies could be possible by enhancing transit infrastructure projects across the region or any fare integration programs to reduce the cost of using transit for low-income households. An equitable fare system working with existing and future transit developments could encourage residents to drive less and help in increasing regional transit ridership. Additionally, land use regulation enabling high-density, mixed-use developments close to transit stops and better integration of urban developments and transportation decisions could provide the environment facilitating walking, cycling, and transit use. The policies to be developed to encourage good travel behaviour and prevent negative externalities are as follows.

8.4.1 Policies for equity outcomes

Given how individuals travel in terms of their trip destinations and travel mode choices, this study provides evidence to focus on improving transit services in low-access neighbourhoods with more low-income families who are more dependent on transit for their daily trips.

The results indicate that public transit is frequently used by low-income females regardless of their car-ownership levels. They use public transit for completing their lengthy work trips while they are living in households with a shared private car. Their reliance on public transit, which is less convenient than private cars, highlights that women have less access to resources for their daily trips. Therefore, authorities and policymakers must consider female needs in transport planning and transit investments. Moreover, the findings reveal that females, who are primarily responsible for non-work trips in low-income carless households, take public transit or car as a passenger.

It may reflect a need to reconsider transit stops in locations where non-work trips are occurred to facilitate their travel needs. Further, public transit operators and companies should prioritize the quality, convenience, and safety of services for their main riders. Alternatively, the government may consider designing programs for low-income females to help access car-sharing services at low costs.

The results of this study confirm that low-income households have fewer trips than high-income families in their daily trip chains. Possible policies for enhancing their activity participation may be designed to improve transit access in areas where there is a high concentration of low-income residents. They should also aim to reduce the cost of travel in certain regions through transit fare reductions. Due to the high cost of designing a new transit line or extending the existing routes, the possible solution would be to increase the frequency of existing lines by reducing the travel time and adding multiple-car vehicles. Notably, it may be a more feasible strategy for implementation in urban areas where a fast and high-capacity transit network is required. Transit authorities may prioritize investing in Bus Rapid Transit (BRT) services in routes with a large demand. Bus rapid transit, including dedicated bus roadways, reduces the travel time by owning its right of way and avoids delays in the mixed, congested roads. BRT system also would be a cost-effective transit service because its operation, establishment, and maintenance cost are less than rail network services and infrastructures.

Furthermore, this study shows that improving transit accessibility in disadvantaged areas can increase the number of existing transit trips and unlock the suppressed demand for non-riders. The findings reveal that the potential benefits can be reaped in inner suburbs with a high concentration of car-deficit households. The cost of owning a car can be a financial burden for them. Therefore, improving transit accessibility and providing transit supply create opportunities for residents with a shared car in a household to engage in daily activities, increase the overall transit ridership, and alleviate equity concerns.

8.4.2 Policies for sustainable outcomes

In the course of achieving sustainable development objectives and reduced pollution emission targets, policies supporting active transportation modes and reducing the use of cars receive more attention from policymakers, authorities, and the government. According to the results of this study, although most low-income families live in Toronto municipal boundaries where public transit services are almost accessible, there is still a large number of car-owning low-income communities living in remote and low-accessibility neighbourhoods where the housing price is affordable for them. They own private vehicles and mostly make their trips by car to reach their destinations. Due to their significant investment in owning a private vehicle, they are often unwilling to give up using their car for daily trips.

With an aim of reducing their car trips, the result of this study shows that car-owning households can be persuaded to take transit trips, whether newly generated trips or switched from other modes, if transit accessibility is improved. Accordingly, transit investments in areas with a high concentration of car-owners may result in mode-shifting from private cars to public transit, thus resulting in environmental benefits. As a result, integrating land-use policies and transit planning is recommended to encourage more sustainable urban developments and travel behaviours. Urban planners and policymakers also should focus on enhancing new employment centers and mixed-use growth within the surroundings of existing transit services. These dense and transit-oriented developments contribute to access opportunities easily and reduce car-reliant trips, traffic congestion, and environmental problems. Therefore, this study demonstrates that the modal shift policies and interventions through improving transit services could be a viable solution to contribute to sustainable outcomes.

Further, the findings of this study show that individuals living in neighbourhoods, where the land use and intersection densities are high are more willing to use active transportation. Therefore, more connected street networks and intersections ease taking short trips by walk or bicycle. Accordingly, making long-term investments in neighbourhoods within walking distance to essential destinations could be an effective solution for sustainable outcomes. The key impact of neighbourhoods and street designs encouraging walking or biking may confirm that policymakers should prioritize built environment consideration in their policy plans to promote active transportation modes. They should review the design and planning of streets to provide a safe and attractive environment for pedestrians and cyclists. Implementing dedicated bike lanes and cycle tracks separated from mixed traffic could be another design solution to entice cyclists and improve their safety. Urban planners and designers may consider other streetscape design policies (e.g., providing adequate street lighting and pedestrian crossings) to enhance pedestrians' safety during the day.

8.4.3 Policies for model selection

Investigating the capability of statistical and ML models in predicting travel mode use and travel behaviour responses of vulnerable people shows that ML algorithms outperform the traditional models in terms of predictive performance. Moreover, there is a heterogeneous pattern among traditional and ML models in predicting transit trips and their spatial distribution maps. Therefore, model selection may have a crucial impact on a suggested policy; thus, choosing the proper algorithm is a vital step in equity-related studies.

The results of this study have significant implications for planning, policy, and travel demand modelling. As decision-makers are increasingly looking for ways to alleviate inequalities in access to transit and improve the activity participation of households living in low-income communities, this study can help agencies examine transit investment projects and transit-related policies. This framework demonstrates the possibility of using ML methods to enhance travel demand predictions. Still, the big question is which model should be used in practice. There is a trade-off between accuracy and interpretability. In any case, the following pipeline is recommended: first, to train different ML and statistical models; second, to statistically compare the result of each algorithm; third, to select an intrinsic interpretable model if the performance difference is negligible or to choose a more complex model when the difference is significant; fourth, to explain the best model using different interpretability tools and discuss its interpretation with an expert; and fifth, to rely on the model if its result is justified by the literature and empirical interpretation, and otherwise, to use an intrinsically interpretable model.

Regarding interpretability, there are different interpretation techniques, enabling researchers to investigate the effect of each feature globally or locally on the final prediction. To better understand the model's decision and the variables' impact on the final prediction, using model-agnostic interpretation tools is suggested. Accordingly, ML algorithms provide an opportunity for travel behaviour studies and policy-making plans without much compromising interpretability.

8.5 Study Limitations and Future Work

In this section, several limitations of conducting this study are underlined, and recommendations are offered for future research to improve data collection, modelling, and policy-making processes.

Given the nature of the Transportation Tomorrow Survey (TTS) data used for this analysis, a number of limitations can be eliminated in future research. Since the focus of data collection for this survey is on weekday activities, some discretionary activities mainly occur on the weekends are overlooked. Therefore, comprehensive data on non-work-related trips taken during the weekdays and weekends may collect for future studies. Recording various types of non-work trips can improve the activity pattern analysis of workers and non-workers. Additionally, each respondent of this household travel survey reports a one-day travel diary. Future works can survey for more than one day to better understand travellers' behaviour and concerns over a week.

In addition to extensive non-work-related trip data, there is a need to gather data on travellers' travel costs. This information would be used to assess if low-income households could afford and would be willing to use other transportation modes in their daily trips. Accordingly, it would make it easier to place its results within the current transport equity literature. According to its outcomes, it is possible to recommend Transit Fare Equity Programs as effective solution for enhancing low-income households' access to public transportation. The survey also does not include the attitudinal questions about transportation modes, residential selection, and other preferences. Consequently, this study is unable to measure or control these variables in terms of mode choice or residential selection. Several built environment and land use variables can also be recorded to improve the investigation of built environment characteristics on travellers' travel patterns.

Further, the latest travel survey in the GTHA related to 2016 is used for this study. However, the COVID-19 pandemic has massively changed the lives of people and their travel behaviour, particularly their transit use, due to the health risks. This drastic change in the social life of individuals may significantly affect their trip schedules and habits in the long run. It is highly recommended that activity patterns of households during pandemics and post-pandemic have been evaluated to gain a better overview of travellers' decisions (Shamshiripour, Rahimi, Shabanpour, & Mohammadian, 2020; Currie, Jain, & Aston, 2021; N. Zhang et al., 2021).

This study is a cross-sectional analysis using only 2016 TTS data, so it lacks measuring the travel behaviour change over a period of time and may exposure to a biased conclusion. Therefore, a longitudinal study evaluating the periodic changes in travel behaviours and decisions is recommended for future works since it would strengthen the analysis. Longitudinal research provides a considerable opportunity to measure travel behaviour and changes in mobility patterns over time. It allows decision-makers and planners to get a clear understanding of mode choice decisions and trip-making behaviours. The cross-sectional design of this study limits our ability to validate the findings over time. The snapshot of car-ownership and income level does not demonstrate whether low-income car-owning households are likely to give up their cars going forward. Longitudinal analysis gives a broader picture of the household's changing decisions over time, making causal analysis more feasible and enabling us to better estimate the long-run behavioural responses to accessibility improvements.

Another noteworthy caveat is that long-term residential selection for the neighbourhoods with an improved transit system should be explored. There is a possibility that low-income households will obtain the ability to live car-free in those neighbourhoods after new transit investments. The benefits may be reaped via long-term shifts in residential preference and car-ownership decisions and not in the "momentary" shifts in people's behaviour in their current accessibility and car ownership levels. Alternatively, it may be expected to observe that some gains made by low-income residents get lost due to gentrification and displacement processes over time. Again, this shortcoming of the present study points toward the need for longitudinal analysis.

Lastly, future studies should investigate other models and data sources to validate the result of the research in different urban contexts. It is worth mentioning that there is no single remedy to model all datasets. Future studies may replicate the experiment in new regions. Nevertheless, this study does not aim to propose the best model to predict travel behaviour but rather to shed light on the significance of model selection based on predictive performance and interpretability.

8.5.1 Reproducibility

To make the work reproducible, some codes are added to the dedicate GitHub page (https://github.com/ElnazYousefzadeh/PhDDissertation).

REFERENCES

Abasahl, F., Kelarestaghi, K. B., & Ermagun, A. (2018). Gender gap generators for bicycle mode choice in Baltimore college campuses. *Travel Behaviour and Society*, *11*, 78-85. doi: https://doi.org/10.1016/j.tbs.2018.01.002

Ades, J., Apparicio, P., & Séguin, A.-M. (2012). Are new patterns of low-income distribution emerging in Canadian metropolitan areas? *The Canadian Geographer/le géographe canadien*, *56*(3), 339–361.

Allen, J., & Farber, S. (2018). How time-use and transportation barriers limit on-campus participation of university students. *Travel Behaviour and Society*, *13*, 174-182. doi: https://doi.org/10.1016/j.tbs.2018.08.003

Allen, J., & Farber, S. (2019). Sizing up transport poverty: A national scale accounting of low-income households suffering from inaccessibility in Canada, and what to do about it. *Transport Policy*, *74*, 214-223. doi: https://doi.org/10.1016/j.tranpol.2018.11.018

Allen, J., & Farber, S. (2020a). A measure of competitive access to destinations for comparing across multiple study regions. *Geographical analysis*, *52*(1), 69–86.

Allen, J., & Farber, S. (2020b). Planning transport for social inclusion: An accessibility-activity participation approach. *Transportation Research Part D: Transport and Environment*, *78*, 102212. doi: https://doi.org/10.1016/j.trd.2019.102212

Allen, J., & Farber, S. (2021). Suburbanization of transport poverty. *Annals of the American Association of Geographers*, *111*(6), 1833-1850. doi: 10.1080/24694452.2020.1859981

Anggraini, R., Arentze, T. A., & Timmermans, H. J. (2008). Car allocation between household heads in car deficient households: a decision model. *European Journal of Transport and Infrastructure Research*, *8*(4).

Baum, C. L. (2009). The effects of vehicle ownership on employment. *Journal of Urban Economics*, *66*(3), 151-163. doi: https://doi.org/10.1016/j.jue.2009.06.003

Baum-Snow, N., Kahn, M. E., & Voith, R. (2005). Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000. *Brookings-Wharton Papers on Urban Affairs*, 147–206.

Bhattacharjee, S., & Goetz, A. R. (2012). Impact of light rail on traffic congestion in denver. *Journal of Transport Geography*, *22*, 262-270. (Special Section on Rail Transit Systems and High Speed Rail)

Blumenberg, E. (2016). Why low-income women in the us still need automobiles. *TPR: Town Planning Review*, *87*(5).

Blumenberg, E., Brown, A., & Schouten, A. (2020). Car-deficit households: determinants and implications for household travel in the US. *Transportation*, *47*(3), 1103–1125.

Blumenberg, E., & Pierce, G. (2012). Automobile ownership and travel by the poor: Evidence from the 2009 National Household Travel Survey. *Transportation Research Record*, *2320*(1), 28-36. doi: 10.3141/2320-04

Blumenberg, E., & Thomas, T. (2014). Travel behavior of the poor after welfare reform. *Transportation Research Record*, *2452*(1), 53-61. doi: 10.3141/2452-07

Böcker, L., van Amen, P., & Helbich, M. (2017). Elderly travel frequencies and transport mode choices in Greater Rotterdam, the Netherlands. *Transportation*, *44*(4), 831–852.

Bowman, J., & Ben-Akiva, M. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, *35*(1), 1-28. doi: https://doi.org/10.1016/S0965-8564(99)00043-9

Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123–140.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Brown, J. R., & Thompson, G. L. (2009). *The influence of service planning decisions on rail transit success or failure, MTI report 08-04* (Tech. Rep.). California: Mineta Transportation Institute Publications.

Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., & Sun, J. (2016). A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. *Transportation Research Part C: Emerging Technologies*, *62*, 21-34. doi: https://doi.org/10.1016/j.trc.2015.11.002

Cameron, A., Trivedi, P., Jackson, M., & Chesher, A. (1998). *Regression analysis of count data*. Cambridge University Press.

Carey, G. N. (2002). Applicability of bus rapid transit to corridors with intermediate levels of transit demand. *Journal of public Transportation*, *5*(2), 5.

Casalicchio, G., Molnar, C., & Bischl, B. (2018). Visualizing the feature importance for black box models. In *Joint european conference on machine learning and knowledge discovery in databases* (pp. 655–670).

Celebi, M. E., & Kingravi, H. A. (2012). Deterministic initialization of the k-means algorithm using hierarchical clustering. *International Journal of Pattern Recognition and Artificial Intelligence*, *26*(07), 1250018.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199-219. doi: https://doi.org/10.1016/S1361-9209(97)00009-6

Chatman, D. G. (2009). Residential choice, the built environment, and nonwork travel: Evidence using new data and methods. *Environment and Planning A: Economy and Space*, *41*(5), 1072-1089. doi: 10.1068/a4114

Chen, E., Ye, Z., & Wu, H. (2021). Nonlinear effects of built environment on intermodal transit trips considering spatial heterogeneity. *Transportation Research Part D: Transport and Environment*, *90*, 102677.

Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K., & Ghassemi, M. (2021). Ethical machine learning in healthcare. *Annual Review of Biomedical Data Science*, *4*(1), 123-144. doi: 10.1146/annurev-biodatasci-092820-114757

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/2939672.2939785

Cheng, L., Bi, X., Chen, X., & Li, L. (2013). Travel behavior of the urban low-income in China: Case study of Huzhou city. *Procedia - Social and Behavioral Sciences*, *96*, 231-242. (Intelligent and Integrated Sustainable Multimodal Transportation Systems Proceedings from the 13th COTA International Conference of Transportation Professionals (CICTP2013)) doi: https://doi.org/10.1016/j.sbspro.2013.08.030

Cheng, L., Chen, X., De Vos, J., Lai, X., & Witlox, F. (2019). Applying a random forest method approach to model travel mode choice behavior. *Travel Behaviour and Society*, *14*, 1 - 10. doi: https://doi.org/10.1016/j.tbs.2018.09.002

Cheng, L., Chen, X., Yang, S., Wu, J., & Yang, M. (2019). Structural equation models to analyze activity participation, trip generation, and mode choice of low-income commuters. *Transportation Letters*, *11*(6), 341-349. doi: 10.1080/19427867.2017.1364460

Chowdhury, T., & Scott, D. M. (2020). Role of the built environment on trip-chaining behavior: an investigation of workers and non-workers in Halifax, Nova Scotia. *Transportation*, *47*(2), 737–761.

Chu, K. K. A., & Chapleau, R. (2010). Augmenting transit trip characterization and travel behavior comprehension: Multiday location-stamped smart card transactions. *Transportation Research Record*, *2183*(1), 29-40. doi: 10.3141/2183-04

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273–297.

Craig, L., & van Tienoven, T. P. (2019). Gender, mobility and parental shares of daily travel with and for children: a cross-national time use comparison. *Journal of Transport Geography*, *76*, 93-102.

Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2020). Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values. *Transportation Research Part C: Emerging Technologies*, *118*, 102674. doi: https://doi.org/10.1016/j.trc.2020.102674

Curl, A., Clark, J., & Kearns, A. (2018). Household car adoption and financial distress in deprived urban communities: A case of forced car ownership? *Transport Policy*, *65*, 61-71. (Household transport costs, economic stress and energy vulnerability) doi: https://doi.org/10.1016/j.tranpol.2017.01.002

Currie, G., & Delbosc, A. (2011a). Exploring the trip chaining behaviour of public transport users in Melbourne. *Transport Policy*, *18*(1), 204-210. doi: https://doi.org/10.1016/j.tranpol.2010.08.003

Currie, G., & Delbosc, A. (2011b). Mobility vs. affordability as motivations for car-ownership choice in urban fringe, low-income Australia. In *Auto motives* (p. 193-208). Emerald Group Publishing Limited.

Currie, G., Jain, T., & Aston, L. (2021). Evidence of a post-COVID change in travel behaviour – self-reported expectations of commuting in melbourne. *Transportation Research Part A: Policy and Practice*, *153*, 218-234. doi: https://doi.org/10.1016/j.tra.2021.09.009

Currie, G., & Senbergs, Z. (2007). Exploring forced car ownership in metropolitan Melbourne. *Australasian Transport Research Forum: ATRF*.

Data Management Group. (2017). *TTS introduction.* http://dmg.utoronto.ca/transportation-tomorrow-survey/tts-introduction. (accessed: 21.06.2020)

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., & Stisen, S. (2018). Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. *Hydrology and Earth System Sciences*, *22*(2), 1299–1315.

Derrac, J., García, S., Molina, D., & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, *1*(1), 3–18.

Dieleman, F. M., Dijst, M., & Burghouwt, G. (2002). Urban form and travel behaviour: Micro-level household attributes and residential context. *Urban Studies*, *39*(3), 507-527. doi: 10.1080/00420980220112801

Dinca-Panaitescu, M., Hulchanski, D., Laflèche, M., McDonough, L., Maaranen, R., & Procyk, S. (2017). *The opportunity equation in the Greater Toronto Area: An update on neighbourhood income inequality and polarization*. United Way Toronto and York Region.

Ding, L., Hwang, J., & Divringi, E. (2016). Gentrification and residential mobility in Philadelphia. *Regional Science and Urban Economics*, *61*, 38-51. doi: https://doi.org/10.1016/j.regsciurbeco.2016.09.004

Du, M., Liu, N., & Hu, X. (2019). Techniques for interpretable machine learning. *Communications of the ACM*, *63*(1), 68–77.

El-Geneidy, A., Buliung, R., Diab, E., van Lierop, D., Langlois, M., & Legrain, A. (2016). Non-stop equity: Assessing daily intersections between transit accessibility and social disparity across the Greater Toronto and Hamilton Area (GTHA). *Environment and Planning B: Planning and Design*, *43*(3), 540-560.

Ellen, I. G., & O'Regan, K. M. (2011). How low income neighborhoods change: Entry, exit, and enhancement. *Regional Science and Urban Economics*, *41*(2), 89-97. doi: https://doi.org/10.1016/j.regsciurbeco.2010.12.005

El Naqa, I., Li, R., & Murphy, M. J. (2015). *Machine learning in radiation oncology: Theory and applications*. Springer.

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, *76*(3), 265-294.

Farber, S., & Marino, M. G. (2017). Transit accessibility, land development and socioeconomic priority: A typology of planned station catchment areas in the Greater Toronto and Hamilton Area. *Journal of Transport and Land Use*, *10*(1), 879–902.

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *J. Mach. Learn. Res.*, *20*(177), 1–81.

Foth, N., Manaugh, K., & El-Geneidy, A. M. (2013). Towards equitable transit: Examining transit accessibility and social need in Toronto, Canada, 1996–2006. *Journal of Transport Geography*, *29*, 1-10.

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation*, *35*(1), 37–54.

Fransen, K., Farber, S., Deruyter, G., & De Maeyer, P. (2018). A spatio-temporal accessibility measure for modelling activity participation in discretionary activities. *Travel Behaviour and Society*, *10*, 10-20. doi: https://doi.org/10.1016/j.tbs.2017.09.002

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189 – 1232. doi: 10.1214/aos/1013203451

Garcia, S., & Herrera, F. (2008). An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of machine learning research*, *9*(12).

García, S., Fernández, A., Luengo, J., & Herrera, F. (2010). Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power. *Information Sciences*, *180*(10), 2044-2064. (Special Issue on Intelligent Distributed Information Systems) doi: https://doi.org/10.1016/j.ins.2009.12.010

Giuliano, G. (2005). Low income, public transit, and mobility. *Transportation Research Record*, *1927*(1), 63-70. doi: 10.1177/0361198105192700108

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? the role of public transportation. *Journal of Urban Economics*, *63*(1), 1-24. doi: https://doi.org/10.1016/j.jue.2006.12.004

Glaeser, E. L., & Ponzetto, G. A. (2018). The political economy of transportation investment. *Economics of Transportation*, *13*, 4-26. (The political economy of transport decisions) doi: https://doi.org/10.1016/j.ecotra.2017.08.001

Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, *24*(1), 44-65. doi: 10.1080/10618600.2014.907095

Goulet-Langlois, G., Koutsopoulos, H. N., & Zhao, J. (2016). Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C: Emerging Technologies*, *64*, 1-16. doi: https://doi.org/10.1016/j.trc.2015.12.012

Guarda, P., Galilea, P., Paget-Seekins, L., & de Dios Ortúzar, J. (2016). What is behind fare evasion in urban bus systems? an econometric approach. *Transportation Research Part A: Policy and Practice*, *84*, 55-71. (Practical Applications of Novel Methodologies to Real Cases: Selected Papers from the XIII Pan-American Conference on Traffic and Transportation Engineering)

Gurley, T., & Bruce, D. (2005). The effects of car access on employment outcomes for welfare recipients. *Journal of Urban Economics*, *58*(2), 250-272. doi: https://doi.org/10.1016/j.jue.2005.05.002

Hafezi, M. H., Liu, L., & Millward, H. (2019). A time-use activity-pattern recognition model for activity-based travel demand modeling. *Transportation*, *46*(4), 1369–1394.

Hagenauer, J., & Helbich, M. (2017). A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*, *78*, 273 - 282. doi: https://doi.org/10.1016/j.eswa.2017.01.057

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, *25*(2), 73-76. doi: 10.1080/01944365908978307

Hensher, D. A., & Reyes, A. J. (2000). Trip chaining as a barrier to the propensity to use public transport. *Transportation*, *27*(4), 341–361.

Hernández-Orallo, J., Flach, P., & Ferri, C. (2012). A unified view of performance metrics: Translating threshold choice into expected classification loss. *Journal of Machine Learning Research*, *13*(91), 2813-2869.

Hertel, S., Keil, R., & Collens, M. (2016). Next stop equity: Routes to fairer transit access in the greater toronto and hamilton area. *Toronto, ON*.

Hillel, T., Bierlaire, M., Elshafie, M. Z., & Jin, Y. (2021). A systematic review of machine learning classification methodologies for modelling passenger mode choice. *Journal of Choice Modelling*, *38*, 100221. doi: https://doi.org/10.1016/j.jocm.2020.100221

Ho, C., & Mulley, C. (2013). Tour-based mode choice of joint household travel patterns on weekend and weekday. *Transportation*, *40*(4), 789–811.

Hochstenbach, C., & Musterd, S. (2018). Gentrification and the suburbanization of poverty: changing urban geographies through boom and bust periods. *Urban Geography*, *39*(1), 26-53. doi: 10.1080/02723638.2016.1276718

Hodgson, F., & Turner, J. (2003). Participation not consumption: the need for new participatory practices to address transport and social exclusion. *Transport Policy*, *10*(4), 265-272. (Transport and Social Exclusion) doi: https://doi.org/10.1016/j.tranpol.2003.08.001

Hu, F., Lv, D., Zhu, J., & Fang, J. (2014). Related risk factors for injury severity of e-bike and bicycle crashes in Hefei. *Traffic Injury Prevention*, *15*(3), 319-323. doi: 10.1080/15389588.2013.817669

Hu, M.-C., Pavlicova, M., & Nunes, E. V. (2011). Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse*, *37*(5), 367-375. doi: 10.3109/00952990.2011.597280

Hua, H., Wan, T., Wenjuan, W., & Paul, C.-C. (2014). Structural zeroes and zero-inflated models. *Shanghai archives of psychiatry*, *26*(4), 236.

Huang, A. (2008). Similarity measures for text document clustering. In *Proceedings of the sixth New Zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand* (Vol. 4, pp. 9–56).

Huang, Y., Gao, L., Ni, A., & Liu, X. (2021). Analysis of travel mode choice and trip chain pattern relationships based on multi-day GPS data: A case study in Shanghai, China. *Journal of Transport Geography*, *93*, 103070.

Hulchanski, J. D. (2010). *The three cities within Toronto: Income polarization among Toronto's neighbourhoods, 1970–2005. University of Toronto.* Toronto: Cities Centre Press.

Jahanshahi, H., & Baydogan, M. G. (2022). nTreeClus: A tree-based sequence encoder for clustering categorical series. *Neurocomputing*, *494*, 224-241. doi: https://doi.org/10.1016/j.neucom.2022.04.076

Jiao, J., Degen, N., & Azimian, A. (2022). Understanding the relationships among E-scooter ridership, transit desert index, and health-related factors. *Transportation Research Record*, 03611981221097094.

Jiao, J., & Wang, F. (2021). Shared mobility and transit-dependent population: A new equity opportunity or issue? *International Journal of Sustainable Transportation*, *15*(4), 294-305. doi: 10.1080/15568318.2020.1747578

Jun, M.-J., Kim, J. I., Kwon, J. H., & Jeong, J.-E. (2013). The effects of high-density suburban development on commuter mode choices in Seoul, Korea. *Cities*, *31*, 230-238. doi: 10.1016/j.cities.2012.06.016

Kamruzzaman, M., Shatu, F., & Habib, K. N. (2020). Travel behaviour in Brisbane: Trends, saturation, patterns and changes. *Transportation Research Part A: Policy and Practice*, *140*, 231-250. doi: https://doi.org/10.1016/j.tra.2020.08.019

Kim, E.-J. (2021). Analysis of travel mode choice in Seoul using an interpretable machine learning approach. *Journal of Advanced Transportation*, *2021*.

Kim, T., Sharda, S., Zhou, X., & Pendyala, R. M. (2020). A stepwise interpretable machine learning framework using linear regression (LR) and long short-term memory (LSTM): City-wide demand-side prediction of yellow taxi and for-hire vehicle (FHV) service. *Transportation Research Part C: Emerging Technologies*, *120*, 102786. doi: https://doi.org/10.1016/j.trc.2020.102786

Klein, N. J., & Smart, M. J. (2017). Car today, gone tomorrow: The ephemeral car in low-income, immigrant and minority families. *Transportation*, *44*(3), 495–510.

Klein, N. J., & Smart, M. J. (2019). Life events, poverty, and car ownership in the United States. *Journal of Transport and Land Use*, *12*(1), 395–418.

Kneebone, E., & Garr, E. (2010). The suburbanization of poverty: Trends in metropolitan America, 2000 to 2008. *Metropolitan Policy Program at Brookings*.

Koch, J., Demirel, M. C., & Stisen, S. (2018). The SPAtial Efficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models. *Geoscientific Model Development*, *11*(5), 1873–1886.

Koushik, A. N., Manoj, M., & Nezamuddin, N. (2020). Machine learning applications in activity-travel behaviour research: A review. *Transport Reviews*, *40*(3), 288-311. doi: 10.1080/01441647.2019.1704307

Kramer, A. (2018). The unaffordable city: Housing and transit in North American cities. *Cities*, *83*, 1-10. doi: https://doi.org/10.1016/j.cities.2018.05.013

Krizek, K. J. (2003). Neighborhood services, trip purpose, and tour-based travel. *Transportation*, *30*(4), 387–410.

Kwan, M.-P., Xiao, N., & Ding, G. (2014). Assessing activity pattern similarity with multidimensional sequence alignment based on a multiobjective optimization evolutionary algorithm. *Geographical Analysis*, *46*(3), 297–320.

Lee, J., Vojnovic, I., & Grady, S. C. (2018). The 'transportation disadvantaged': Urban form, gender and automobile versus non-automobile travel in the Detroit region. *Urban Studies*, *55*(11), 2470-2498. doi: 10.1177/0042098017730521

Lee, N. T. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication and Ethics in Society*.

Legrain, A., Buliung, R., & El-Geneidy, A. M. (2016). Travelling fair: Targeting equitable transit by understanding job location, sectorial concentration, and transit use among low-wage workers. *Journal of Transport Geography*, *53*, 1-11. doi: https://doi.org/10.1016/j.jtrangeo.2016.04.001

León, M., Mkrtchyan, L., Depaire, B., Ruan, D., & Vanhoof, K. (2014). Learning and clustering of fuzzy cognitive maps for travel behaviour analysis. *Knowledge and information systems*, *39*(2), 435–462.

Lewis, D. D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. In C. Nédellec & C. Rouveirol (Eds.), *Machine learning: Ecml-98* (pp. 4–15). Berlin, Heidelberg: Springer Berlin Heidelberg.

Li, B., & Han, L. (2013). Distance weighted cosine similarity measure for text classification. In H. Yin et al. (Eds.), *Intelligent data engineering and automated learning – ideal 2013* (pp. 611–618). Berlin, Heidelberg: Springer Berlin Heidelberg.

Li, M. (2019). An improved FCM clustering algorithm based on cosine similarity. In *Proceedings of the 2019 international conference on data mining and machine learning* (pp. 103–109).

Li, Z., Wang, W., Yang, C., & Jiang, G. (2013). Exploring the causal relationship between bicycle choice and trip chain pattern. *Transport Policy*, *29*, 170-177. doi: https://doi.org/10.1016/j.tranpol.2013.06.001

Liu, F., Janssens, D., Cui, J., Wets, G., & Cools, M. (2015). Characterizing activity sequences using profile hidden markov models. *Expert Systems with Applications*, *42*(13), 5705-5722. doi: https://doi.org/10.1016/j.eswa.2015.02.057

Liu, S., & Kontou, E. (2022). Quantifying transportation energy vulnerability and its spatial patterns in the United States. *Sustainable Cities and Society*, *82*, 103805. doi: https://doi.org/10.1016/j.scs.2022.103805

Liu, Y., Liu, Z., & Jia, R. (2019). DeepPF: A deep learning based architecture for metro passenger flow prediction. *Transportation Research Part C: Emerging Technologies*, *101*, 18-34. doi: https://doi.org/10.1016/j.trc.2019.01.027

Lo, L., Shalaby, A., & Alshalalfah, B. (2011). Relationship between immigrant settlement patterns and transit use in the greater toronto area. *Journal of Urban Planning and Development*, *137*(4), 470–476.

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy*, *20*, 105-113. doi: https://doi.org/10.1016/j.tranpol.2012.01.013

Lucas, K., Philips, I., Mulley, C., & Ma, L. (2018). Is transport poverty socially or environmentally driven? comparing the travel behaviours of two low-income populations living in central and peripheral locations in the same city. *Transportation Research Part A: Policy and Practice*, *116*, 622-634. doi: https://doi.org/10.1016/j.tra.2018.07.007

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777).

Ma, J., Mitchell, G., & Heppenstall, A. (2014). Daily travel behaviour in Beijing, China: An analysis of workers' trip chains, and the role of socio-demographics and urban form. *Habitat International*, *43*, 263-273. doi: https://doi.org/10.1016/j.habitatint.2014.04.008

Madariaga, I. S. d. (2016). Mobility of care: introducing new concepts in urban transport. In *Fair shared cities* (pp. 51–66). Routledge.

Manaugh, K., Badami, M. G., & El-Geneidy, A. M. (2015). Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in North America. *Transport Policy*, *37*, 167-176. doi: https://doi.org/10.1016/j.tranpol.2014.09.013

Martens, K. (2016). *Transport justice: Designing fair transportation systems*. Routledge.

Mattioli, G. (2017). "Forced car ownership" in the UK and Germany: socio-spatial patterns and potential economic stress impacts. *Social Inclusion*, *5*(4), 147–160.

Mattioli, G., Anable, J., & Vrotsou, K. (2016). Car dependent practices: Findings from a sequence pattern mining study of UK time use data. *Transportation Research Part A: Policy and Practice*, *89*, 56-72. doi: https://doi.org/10.1016/j.tra.2016.04.010

McCallum, A., & Nigam, K. (1998). A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (Vol. 752, pp. 41–48).

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. Frontiers in econometrics, Academic Press, New York.

McGuckin, N., & Murakami, E. (1995). Examining trip-chaining behaviour: a comparison of travel by men and women, Federal Highway Administration. *Washington, DC, FHWA*.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021, jul). A survey on bias and fairness in machine learning. *ACM Comput. Surv.*, *54*(6). doi: 10.1145/3457607

Mercado, R. G., Paez, A., Farber, S., Roorda, M. J., & Morency, C. (2012). Explaining transport mode use of low-income persons for journey to work in urban areas: a case study of Ontario and Quebec. *Transportmetrica*, *8*(3), 157-179. doi: 10.1080/18128602.2010.539413

Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

Moniruzzaman, M., & Páez, A. (2012). Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters. *Journal of Transport Geography*, *24*, 198-205. (Special Section on Theoretical Perspectives on Climate Change Mitigation in Transport) doi: https://doi.org/10.1016/j.jtrangeo.2012.02.006

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *Journal of classification*, *31*(3), 274–295.

Murthy, J. (2012). Clustering based on cosine similarity measure. *(IJESAT) International Journal Of Engineering Science & Advanced Technology*, *2*, 508-512.

Naess, P. (2008). Gender differences in the influences of urban structure on daily travel. *Gendered mobilities*, 173–192.

Oswin, N. (2014). Queer theory. In *The routledge handbook of mobilities* (pp. 105–113). Routledge.

Paez, A., Ruben, M., Faber, S., Morency, C., & Roorda, M. (2009). *Mobility and social exclusion in canadian communities: An empirical investigation of canadian communities.* Policy Research Directorate, Strategic Policy and Research, Human Resources and Social Development Canada.

Palm, M., Shalaby, A., & Farber, S. (2020). Social equity and bus on-time performance in Canada's largest city. *Transportation Research Record*, *2674*(11), 329-342. doi: 10.1177/0361198120944923

Pena, J. M., Lozano, J. A., & Larranaga, P. (1999). An empirical comparison of four initialization methods for the k-means algorithm. *Pattern recognition letters*, *20*(10), 1027–1040.

Pizzol, B., Strambi, O., Giannotti, M., Arbex, R. O., & Alves, B. B. (2021). Activity behavior of residents of Paraisópolis slum: Analysis of multiday activity patterns using data collected with smartphones. *Journal of Choice Modelling*, *39*, 100287. doi: https://doi.org/10.1016/j.jocm.2021.100287

Potoglou, D., & Kanaroglou, P. S. (2008). Modelling car ownership in urban areas: A case study of Hamilton, Canada. *Journal of Transport Geography*, *16*(1), 42-54. doi: https://doi.org/10.1016/j.jtrangeo.2007.01.006

Primerano, F., Taylor, M. A., Pitaksringkarn, L., & Tisato, P. (2008). Defining and understanding trip chaining behaviour. *Transportation*, *35*(1), 55–72.

Pucher, J. (2002). Renaissance of public transport in the united states? *Transportation Quarterly*, *56*(1), 33–49.

Pucher, J., & Renne, J. L. (2003). Socioeconomics of urban travel. evidence from the 2001 NHTS. *Transportation Quarterly*, *57*, 49–77.

Rafiq, R., & McNally, M. G. (2021). Heterogeneity in activity-travel patterns of public transit users: An application of latent class analysis. *Transportation Research Part A: Policy and Practice*, *152*, 1-18. doi: https://doi.org/10.1016/j.tra.2021.07.011

Raphael, S., & Rice, L. (2002). Car ownership, employment, and earnings. *Journal of Urban Economics*, *52*(1), 109-130. doi: https://doi.org/10.1016/S0094-1190(02)00017-7

Ravensbergen, L., Fournier, J., & A, E.-G. (2022). Mobility of care: An exploratory analysis in Montréal, Canada. In *Paper to be presented at the transportation research board 101st annual meeting* (pp. 1–16).

Rezapour, M., Moomen, M., & Ksaibati, K. (2019). Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: A case study in Wyoming. *Journal of Safety Research*, *68*, 107-118. doi: https://doi.org/10.1016/j.jsr.2018.12.006

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (p. 1135–1144). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/2939672.2939778

Richmond, J. (2001). A whole-system approach to evaluating urban transit investments. *Transport Reviews*, *21*(2), 141-179.

Rosenbloom, S. (1998). *Transit markets of the future: the challenge of change* (Vol. 28). Transportation Research Board.

Rosenbloom, S. (2004). Understanding women's and men's travel patterns. In *Research on women's issues in transportation: Report of a conference* (pp. 7–28).

Roy, A., Nelson, T. A., Fotheringham, A. S., & Winters, M. (2019). Correcting bias in crowdsourced data to map bicycle ridership of all bicyclists. *Urban Science*, *3*(2). doi: 10.3390/urbansci3020062

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215.

Sagaris, L., & Tiznado-Aitken, I. (2020). Sustainable transport and gender equity: Insights from Santiago, Chile. In *Urban mobility and social equity in latin america: Evidence, concepts, methods* (p. 103-129). Emerald Publishing Limited. doi: https://doi.org/10.1108/S2044-994120200000012009

Sanchez, T. W., Shen, Q., & Peng, Z.-R. (2004). Transit mobility, jobs access and low-income labour participation in US metropolitan areas. *Urban Studies*, *41*(7), 1313-1331. doi: 10.1080/0042098042000214815

Saneinejad, S., & Roorda, M. (2009). Application of sequence alignment methods in clustering and analysis of routine weekly activity schedules. *Transportation Letters*, *1*(3), 197-211. doi: 10.3328/TL.2009.01.03.197-211

Sasirekha, K., & Baby, P. (2013). Agglomerative hierarchical clustering algorithm-a. *International Journal of Scientific and Research Publications*, *83*, 83.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., . . . Lin, C.-T. (2017). A review of clustering techniques and developments. *Neurocomputing*, *267*, 664-681. doi: https://doi.org/10.1016/j.neucom.2017.06.053

Scheiner, J., & Holz-Rau, C. (2012). Gendered travel mode choice: a focus on car deficient households. *Journal of Transport Geography*, *24*, 250-261. (Special Section on Theoretical Perspectives on Climate Change Mitigation in Transport) doi: https://doi.org/10.1016/j.jtrangeo.2012.02.011

Schneider, F., Ton, D., Zomer, L.-B., Daamen, W., Duives, D., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2021). Trip chain complexity: a comparison among latent classes of daily mobility patterns. *Transportation*, *48*(2), 953–975.

Scholten, C. L., & Joelsson, T. (2019). *Integrating gender into transport planning: From one to many tracks*. Springer.

Shamshiripour, A., Rahimi, E., Shabanpour, R., & Mohammadian, A. K. (2020). How is COVID-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. *Transportation Research Interdisciplinary Perspectives*, *7*, 100216. doi: https://doi.org/10.1016/j.trip.2020.100216

Shao, Q., Zhang, W., Cao, X., Yang, J., & Yin, J. (2020). Threshold and moderating effects of land use on metro ridership in shenzhen: Implications for tod planning. *Journal of Transport Geography*, *89*, 102878. doi: https://doi.org/10.1016/j.jtrangeo.2020.102878

Shiftan, Y. (1998). Practical approach to model trip chaining. *Transportation Research Record*, *1645*(1), 17-23. doi: 10.3141/1645-03

Sidorov, G., Gelbukh, A., Gómez-Adorno, H., & Pinto, D. (2014). Soft similarity and soft cosine measure: Similarity of features in vector space model. *Computación y Sistemas*, *18*(3), 491–504.

Simma, A., & Axhausen, K. W. (2001). Structures of commitment in mode use: a comparison of switzerland, germany and great britain. *Transport Policy*, *8*(4), 279–288.

Song, Y., Ren, S., Wolfson, J., Zhang, Y., Brown, R., & Fan, Y. (2021). Visualizing, clustering, and characterizing activity-trip sequences via weighted sequence alignment and functional data analysis. *Transportation Research Part C: Emerging Technologies*, *126*, 103007. doi: https://doi.org/10.1016/j.trc.2021.103007

Stanley, J., Hensher, D. A., Stanley, J., Currie, G., Greene, W. H., & Vella-Brodrick, D. (2011). Social exclusion and the value of mobility. *Journal of Transport Economics and Policy (JTEP)*, *45*(2), 197–222.

Statistics Canada. (2021a). *Annual Demographic Estimates: Subprovincial areas.* https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710014201. (accessed: 01.31.2021)

Statistics Canada. (2021b). *Low income cut-offs (LICOs) before and after tax by community size and family size, in current dollars.* https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1110024101. (accessed: 01.31.2021)

Staub, K. E., & Winkelmann, R. (2013). Consistent estimation of zero-inflated count models. *Health Economics*, *22*(6), 673-686.

Strathman, J. G., Dueker, K. J., & Davis, J. S. (1994). Effects of household structure and selected travel characteristics on trip chaining. *Transportation*, *21*(1), 23–45.

Sultana, Z., Mishra, S., Cherry, C. R., Golias, M. M., & Tabrizizadeh Jeffers, S. (2018). Modeling frequency of rural demand response transit trips. *Transportation Research Part A: Policy and Practice*, *118*, 494-505.

Suthaharan, S. (2016). Support vector machine. In *Machine learning models and algorithms for big data classification: Thinking with examples for effective learning* (pp. 207–235). Boston, MA: Springer US. doi: 10.1007/978-1-4899-7641-3_9

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2003, September). *Analyzing the determinants of transit ridership using a two-stage least squares regression on a national sample of urbanized areas* (University of California Transportation Center, Working Papers). Los Angeles: University of California Transportation Center.

Tiznado-Aitken, I., Lucas, K., Muñoz, J. C., & Hurtubia, R. (2020). Understanding accessibility through public transport users' experiences: A mixed methods approach. *Journal of Transport Geography*, *88*, 102857. doi: https://doi.org/10.1016/j.jtrangeo.2020.102857

Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press. doi: 10.1017/CBO9780511805271

Tran, M., Draeger, C., Wang, X., & Nikbakht, A. (2022). Monitoring the well-being of vulnerable transit riders using machine learning based sentiment analysis and social media: Lessons from COVID-19. *Environment and Planning B: Urban Analytics and City Science*, *0*(0), 23998083221104489. doi: 10.1177/23998083221104489

Turner, T., & Niemeier, D. (1997). Travel to work and household responsibility: New evidence. *Transportation*, *24*(4), 397–419.

Turrell, G., Haynes, M., Wilson, L.-A., & Giles-Corti, B. (2013). Can the built environment reduce health inequalities? a study of neighbourhood socioeconomic disadvantage and walking for transport. *Health & Place*, *19*, 89-98. doi: https://doi.org/10.1016/j.healthplace.2012.10.008

Tyndall, J. (2017). Waiting for the R train: Public transportation and employment. *Urban Studies*, *54*(2), 520-537. doi: 10.1177/0042098015594079

Vance, C., & Iovanna, R. (2007). Gender and the automobile: Analysis of nonwork service trips. *Transportation Research Record*, *2013*(1), 54-61.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer New York. Retrieved from `https://books.google.ca/books?id=EoDSBwAAQBAJ`

Walks, A. (2018). Driving the poor into debt? automobile loans, transport disadvantage, and automobile dependence. *Transport Policy*, *65*, 137-149. (Household transport costs, economic stress and energy vulnerability) doi: https://doi.org/10.1016/j.tranpol.2017.01.001

Wan, Q., Li, Z., Qi, Y., Yu, J., Pu, Z., Peng, G., & Liu, Q. (2019). Comparing uncertainties in travel mode choice decisions for various trip chains. *Advances in Mechanical Engineering*, *11*(4), 1687814019835102.

Wang, F., & Ross, C. L. (2018). Machine learning travel mode choices: Comparing the performance of an extreme gradient boosting model with a multinomial logit model. *Transportation Research Record*, *2672*(47), 35-45. doi: 10.1177/0361198118773556

Wang, Q. R., & Suen, C. Y. (1984, July). Analysis and design of a decision tree based on entropy reduction and its application to large character set recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-6*(4), 406-417. doi: 10.1109/TPAMI.1984.4767546

Wilson, W. C. (1998). Activity pattern analysis by means of Sequence-Alignment methods. *Environment and Planning A: Economy and Space*, *30*(6), 1017-1038. doi: 10.1068/a301017

Xie, C., Lu, J., & Parkany, E. (2003). Work travel mode choice modeling with data mining: Decision trees and neural networks. *Transportation Research Record*, *1854*(1), 50-61. doi: 10.3141/1854-06

Yan, X., Liu, X., & Zhao, X. (2020). Using machine learning for direct demand modeling of ridesourcing services in Chicago. *Journal of Transport Geography*, *83*, 102661. doi: https://doi.org/10.1016/j.jtrangeo.2020.102661

Yang, L., Shen, Q., & Li, Z. (2016). Comparing travel mode and trip chain choices between holidays and weekdays. *Transportation Research Part A: Policy and Practice*, *91*, 273-285. doi: https://doi.org/10.1016/j.tra.2016.07.001

Ye, X., Pendyala, R. M., & Gottardi, G. (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B: Methodological*, *41*(1), 96-113. doi: https://doi.org/10.1016/j.trb.2006.03.004

Yousefzadeh Barri, E., Farber, S., Kramer, A., Jahanshahi, H., Allen, J., & Beyazit, E. (2021). Can transit investments in low-income neighbourhoods increase transit use? exploring the nexus of income, car-ownership, and transit accessibility in Toronto. *Transportation Research Part D: Transport and Environment*, *95*, 102849. doi: https://doi.org/10.1016/j.trd.2021.102849

Zhang, N., Jia, W., Wang, P., Dung, C.-H., Zhao, P., Leung, K., . . . Li, Y. (2021). Changes in local travel behaviour before and during the COVID-19 pandemic in hong kong. *Cities*, *112*, 103139. doi: https://doi.org/10.1016/j.cities.2021.103139

Zhang, Q., Clifton, K. J., Moeckel, R., & Orrego-Oñate, J. (2019). Household trip generation and the built environment: Does more density mean more trips? *Transportation Research Record*, *2673*(5), 596-606. doi: 10.1177/0361198119841854

Zhang, Y., Brussel, M., Thomas, T., & van Maarseveen, M. (2018). Mining bike-sharing travel behavior data: An investigation into trip chains and transition activities. *Computers, Environment and Urban Systems*, *69*, 39-50.

Zhang, Y., & Xie, Y. (2008). Travel mode choice modeling with support vector machines. *Transportation Research Record*, *2076*(1), 141-150.

Zhao, J., Wang, J., & Deng, W. (2015). Exploring bikesharing travel time and trip chain by gender and day of the week. *Transportation Research Part C: Emerging Technologies*, *58*, 251-264. (Big Data in Transportation and Traffic Engineering)

Zhao, X., Yan, X., Yu, A., & Van Hentenryck, P. (2020). Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behaviour and Society*, *20*, 22 - 35. doi: https://doi.org/10.1016/j.tbs.2020.02.003

Zhou, X., Wang, M., & Li, D. (2019). Bike-sharing or taxi? modeling the choices of travel mode in Chicago using machine learning. *Journal of Transport Geography*, *79*, 102479. doi: https://doi.org/10.1016/j.jtrangeo.2019.102479

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). Zero-truncated and zero-inflated models for count data. In *Mixed effects models and extensions in ecology with R* (pp. 261–293). New York, NY: Springer New York. doi: 10.1007/978-0-387-87458-6_11

APPENDICES

APPENDIX A.1 : Feature importance algorithm
APPENDIX A.2 : Partial Dependence Plot (PDP) algorithm
APPENDIX A.3 : Individual Conditional Expectation (ICE) algorithm
APPENDIX A.4 : Local interpretable model-agnostic explanations (LIME) algorithm
APPENDIX B.1 : Detailed comparison of classification models' performances
APPENDIX B.2 : Detailed comparison of regression models' performances

APPENDIX A1: Feature importance algorithm

Algorithm A.1 describes a model-agnostic permutation-based feature importance technique introduced by Fisher, Rudin, and Dominici (2019). In this study, the permutation-based feature importance of RF algorithm is computed using `scikit-learn` package in `Python` platform.

Algorithm A.1: Feature importance algorithm

Data: Trained model $\hat{f}$, feature matrix X, outcome y, error measure $L(y, \hat{f})$.
Fit the model on a train data with real features and calculate the actual model performance (e.g. RMSE for a regression model);
for *Each feature* $j \in \{1, \ldots, p\}$ **do**
  Permutate values of feature j and generate a new feature matrix;
  Fit the model on the modified data and estimate the permutated model performance;
  Compute the difference between the actual model performance and the permutated model performance
end
Rank features according to the differences between their permutated model and the actual one

---

APPENDIX A2: Partial Dependence Plot (PDP) algorithm

The feature space x is divided into subgroups j and C. j includes the feature on which the partial dependence function $\hat{f}_j$ is applied, and C corresponds to the remaining attributes in the dataset. x_j and x_C define the values of features in j and C, respectively. The partial dependence function estimates the relationship between x_j and the targeted variable by keeping the feature values in subgroup C unchanged. Therefore, a function is generated depending only on feature j and the average effect of other features in C (Casalicchio et al., 2018). The partial dependence function $\hat{f}_j$ on x_j is

$$\hat{f}_j(x_j) = \mathbb{E}_{X_C}[\hat{f}(x_j, X_C)] \qquad x_s \in j, x_c \in C \tag{A.1}$$

$$\hat{f}_j(x_j) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}(x_j, x_C^{(i)}). \tag{A.2}$$

Accordingly, PDP can be constructed using Algorithm A.2 as follows.

Algorithm A.2: PDP algorithm

Data: Feature matrix X, Set of unique values $\{x_{j1}, \ldots, x_{jk}\}$ in feature j.
Select feature j on which the partial dependence function is applied (C is the set of the remaining features).
for *Each feature* $j \in \{1, \ldots, p\}$ **do**
  for *Each unique value* $x_{ji} \in \{x_{j1}, \ldots, x_{jk}\}$ **do**
    Replace all original values x_j of the selected feature j with the constant value x_{ji} (i is the number of observations);
    Keep the values x_C of complement features C unchanged;
    Fit the model and compute the predicted response $\hat{f}_j$ for the modified dataset;
    Average the predicted value to obtain $\bar{f}_j(x_{ji})$;
  end
  Plot the pairs $\{x_{ji}, \bar{f}_j(x_{ji})\}$.
end

APPENDIX A3: Individual Conditional Expectation (ICE) algorithm

Similar to the PDP algorithm, ICE iterates on the set of unique values for each feature; however, it plots the output per instance instead of averaging it for all observations. Accordingly, n estimated response curves each of which corresponds to the value of i-th observation $x_j^{(i)}$, the prediction of i-th observation $\hat{f}_j^{(i)}(x_j^{(i)})$ while the value of the i-th instance for the features in $x_C^{(i)}$ is unchanged, are plotted. Therefore, ICE plots include curves $\hat{f}_j^{(i)}$ for each observation in $\{(x_j^{(i)}, x_C^{(i)})\}_{i=1}^{n}$ (Molnar, 2020; Casalicchio et al., 2018). In other words, the ICE plot is a disaggregated form of PDP.

APPENDIX A4: Local interpretable model-agnostic explanations (LIME) algorithm

In order to demonstrate LIME interpretation, these steps are followed:

1. Selecting the person for whom an explanation for the black-box model (e.g., Neural Networks) is required.
2. Perturbing the dataset to get the predictions for the black-box model using these new points.
3. Weighting the new perturbed samples based on how close they are to the person.
4. Training a weighted, interpretable model (Linear Regression in our case) on the new dataset.
5. Explaining the prediction by analysing the local model.

Since the experiment is a regression task, the linear regression model is employed as the interpretable model in LIME.

APPENDIX B1: Detailed comparison of classification models' performances

Table B.1 presents a detailed statistical comparison of models' performances. The comparison includes the best performing model together with the models whose performances are not statically different than the best one. Also, the results show the logistic regression's average rank in terms of different metrics. According to all threshold-dependent and -independent metrics, ML algorithms, specifically RF, XGB, and NN, are frequently among the top models in predicting the probability of taking transit trips by low-income individuals.

Table B.1 : Comparing the performance of the best classifiers and the baseline classifier using the Friedman Aligned Ranks test and its *post hoc* analysis.

Metric	**Friedman Aligned Ranks test (p-value)**	**The best model and possible ties**	**Log Regression's mean rank**
Accuracy	6.659e-10*	b: RF (1.4) t: XGB (1.9) & NN (3)	6
Precision (C_0)	5.132e-09*	b: RF (1.9) t: XGB (2.5) & NN (2.5) & DT (3.5)	6.9
Recall (C_0)	2.046e-07*	b: RF (1.65) t: XGB (2.2) & SVM (3.3) & NN (3.9) & DT (4.65)	5.4
F1-Score (C_0)	9.449e-10*	b: RF (1.4) t: XGB (1.9) & NN (3.1)	6
Precision (C_1)	1.745e-08*	b: RF (1.5) t: XGB (2.1) & SVM (3.5) & NN (3.8)	5.8
Recall (C_1)	1.347e-08*	b: RF (1.9) t: XGB (2.55) & NN (2.7) & DT (3.5)	7
F1-Score (C_1)	5.771e-10*	b: RF (1.4) t: XGB (2) & NN (2.8)	6
AUC-ROC	3.458e-10*	b: RF (1.7) t: XGB (2.2) & NN (2.1) & SVM (4)	5

b: the best model; t: possible ties with insignificant difference;
*: statistically significant based on $\alpha = 0.05$

APPENDIX B2: Detailed comparison of regression models' performances

Table B.2 shows a detailed statistical comparison of regressors' performances. It lists the best-performing regressor and statistically tied models. It also The average ranks of the models according to each metric are written within parentheses. According to the Bergmann-Hommel *post hoc* test, ML models, including tree-based algorithms and NN, are the best models for estimating the number of transit trips for vulnerable groups. Moreover, traditional models, e.g., ZINB and Hurdle, have the lowest predictive power. Accordingly, utilizing ML algorithms to model either a classification or a regression travel-mode problem is recommended.

Table B.2 : Comparing the performance of the best regressors and the baseline regressors using the Friedman Aligned Ranks test and its *post hoc* analysis.

Metric	Friedman Aligned Ranks test (p-value)	The best model and possible ties	Traditional models' mean rank
R_Squared	1.18e-11*	b: RF (1) t: NN (2.2) & XGB (3.2)	LinR (8) & ZINB (6) & Hurdle (7)
RMSE	1.08e-01*	b: RF (1) t: NN (2.2) & XGB (3.2)	LinR (8) & ZINB (6) & Hurdle (7)
MDAE	8.76e-12*	b: DT (1) t: SVM (2.1) & NN (3.1)	LinR (8) & ZINB (4.1) & Hurdle (7)
RRSE	1.06e-11*	b: RF (1) t: NN (2.2) & XGB (3.2)	LinR (8) & ZINB (6) & Hurdle (7)

b: the best model; t: possible ties with insignificant difference;
*: statistically significant based on $\alpha = 0.05$

CURRICULUM VITAE

Name Surname : Elnaz YOUSEFZADEH BARRI

EDUCATION :

- **B.Sc.** : 2008, Islamic Azad University of Tabriz, Faculty of Art and Architecture, Department of Architecture
- **M.Sc.** : 2016, Islamic Azad University - Science and Research Branch Tehran, Faculty of Civil Engineering, Architecture and Art, Department of Urban Planning

PROFESSIONAL EXPERIENCE AND REWARDS:

- **2019 - 2021** - Visiting Researcher, Department of Geography & Planning, University of Toronto, CA.
- **2021 - 2022** - Course Lecturer, School of Urban & Regional Planning, Toronto Metropolitan University, CA.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Yousefzadeh Barri, E.**, Farber, S., Kramer, A., Jahanshahi, H., Allen, J., & E. Beyazit. (2021). Can transit investments in low-income neighbourhoods increase transit use? Exploring the nexus of income, car-ownership, and transit accessibility in Toronto, *Transportation Research Part D: Transport and Environment, 95*, 102849.

- **Yousefzadeh Barri, E.**, Farber, S., Kramer, A., Jahanshahi, H., Allen, J., & E. Beyazit (November 2020). *The Myth of the Captive Transit Rider: Do Transit Investments in Low-Income Neighbourhoods Matter?*. Paper presented at 60th ACSP Annual Conference.

- **Yousefzadeh Barri, E.**, Farber, S., Kramer, A., Jahanshahi, H., Allen, J., & E. Beyazit (January 2021). *The Myth of the Captive Transit Rider: Do Transit Investments in Low-Income Neighbourhoods Matter?*. Paper presented at 100th annual meeting of the Transportation Research Board.

- **Yousefzadeh Barri, E.**, Farber, S., Jahanshahi, H., & E. Beyazit (January 2022). *Understanding Transit Ridership of Low-Income Carless Individuals through a Comparison of Statistical and Machine Learning Algorithms*. Paper presented at 101th annual meeting of the Transportation Research Board.

- **Yousefzadeh Barri, E.**, Farber, S., Jahanshahi, H., Tiznado-Aitken, I., & E. Beyazit. How income and car ownership shape travel behaviour: Exploring daily activity patterns through clustering trip chain sequences, *(Under review)*.

- **Yousefzadeh Barri, E.**, Farber, S., Jahanshahi, H., & E. Beyazit. Understanding Transit Ridership of Low-Income Carless Individuals through a Comparison of Statistical and Machine Learning Algorithms, *(Under review)*.

OTHER PUBLICATIONS:

- **Yousefzadeh Barri, E.**, Z.S.S. Zarabadi (January 2016). *Evaluating the Factors in Locating Underground Spaces using ANP Model (Case Study: Tehran Azadi Square)*. Paper presented at International Conference in applied research on Civil Engineering, Architecture and Urban Planning.

- **Yousefzadeh Barri, E**, Z.S.S. Zarabadi (March 2016). *Conservation of Historic City Centers using Urban Underground Space Development*. Paper presented at International Conference of Architecture, Restoration, Urbanism, and Sustainable Environment.

E. YOUSEFZADEH BARRI

EVALUATING TRAVEL MODE DECISIONS AND TRANSPORT MODELS IN UNDERSTANDING TRANSIT EQUITY: THE CASE OF GREATER TORONTO AND HAMILTON AREA

2022