IDENTIFYING THE TRAVEL PATTERNS OF ON-DEMAND TAXI TRIPS THROUGH INFERRED TRIP PURPOSES

Durage Dineth Dhananjaya

218014X

Degree of Master of Science

Department of Transport and Logistics Management

University of Moratuwa

Sri Lanka

SEPTEMBER 2022

IDENTIFYING THE TRAVEL PATTERNS OF ON-DEMAND TAXI TRIPS THROUGH INFERRED TRIP PURPOSES

Durage Dineth Dhananjaya

218014X

Degree of Master of Science

Department of Transport and Logistics Management

University of Moratuwa

Sri Lanka

SEPTEMBER 2022

DECLARATION OF ORIGINALITY

I declare that this is my own work, and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 23/09/2022

UOM Verified Signature

STATEMENT OF THE SUPERVISOR

The above candidate has carried out research for the Degree of Master of Science under my supervision.

Name of the supervisor: Dr. T. Sivakumar

Signature of the Supervisor: Date: 23/09/2022

UOM Verified Signature

ABSTRACT

Activity-based modeling has become the backbone behind transportation planning, and trip purposes that can be inferred from large GPS datasets of different travel means are paving the path to augmenting its accuracy. In this context, the trip purpose inference problem has emerged since the GPS is unable to capture the trip purposes explicitly. This problem has not been thoroughly addressed in developing countries despite the fact that the applicability of a trip purpose inference model extremely depends on the land use context. Hence, this study attempted to ameliorate an accurate model (base model) for a chosen study area in Colombo District, Sri Lanka using the on-demand taxi trips data.

Point of Interest (POI) data is often accompanied by the purpose inference models as it provides a complete insight into the land use around origin and destinations. In the pursuit of the main objective of the study, machine-learning-based text classification was tested to improve the number of informative POIs and its outcome indicated that the Support Vector Machine (SVM) classifier can be utilized effectively and efficiently. The designed trip purpose inference model referring to a base model is proposed as a three-layer trip purpose inference framework in which a method to impute purpose based on trip regularities and the method to identify residential trips was included as two layers before using the base model. The validity of the proposed model was evaluated with the assistance of household travel survey data and with respect to the purpose proportions and division level R-square. Furthermore, travel patterns of the on-demand taxi trips were assessed in terms of temporal regularity, trip lengths, and spatial dynamics. It is recommended to conduct further studies to assess the applicability of other parameters such as trip origin context and trip time with the assistance of unsupervised learning methods.

Keywords:

Travel behavior, trip purpose inference, human mobility, taxi trips, GPS data, POI data, semantic enrichment, developing countries.

ACKNOWLEDGEMENTS

I would like to convey my heartfelt gratitude to Dr. T. Sivakumar for providing me an opportunity to conduct this study under his supervision and for assisting me to drive this towards success from the very beginning. The guidance he has given me assisted to improve the quality of this work gradually and ultimately add a valuable context of knowledge for the community regarding the study area. I would like to extend my appreciation to Senior Professor Amal S. Kumarage for further helping me to identify the sections that required improvements in making this work more worthwhile for both academics and industry.

My warm thankfulness goes to Dr. Amal S. Perera for helping me in data collection and also providing me the additional knowledge from a computer science background which was highly useful during the data processing task. I would also like to thank all the members who worked on the research project “Data-Driven Transport Analytics” and the member of the “DataSearch” research center for the assistance and comments on the work. Furthermore, I would like to convey my sincere gratitude to Accelerating Higher Education Expansion and Development (AHEAD) Operation for financially assisting this study through the above-mentioned project.

Most importantly, none of this would be possible without the love and care of my family. I am grateful to my parents and my brother for strengthening me throughout this study period and helping me to start again with stamina during the critical periods which motivated me to complete this study at level best.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

AFC - Automatic Fare Collection (AFC)

API - Application Programming Interfaces (API)

AGPS - Assisted GPS (AGPS)

AIM - Activity Inference Model (AIM),

CDR - Caller Detail Records (CDR)

DBSCAN - Density-Based Spatial Clustering of Applications with Noise (DBSCAN)

DOP - Drop off Point (DOP)

DMR - Dirichlet-Multinomial Regression (DMR)

DS - Division Secretary (DS)

DT – Decision Tree

KDE - Kernel Density Estimation (KDE)

KNN – K Nearest Neighbors

GPS - Geographical Processing System (GPS)

HG - Human Geography (HG).

HDOP - Horizontal Dilution of Precision (HDOP)

LBS - Location Based Services (LBS)

LR – Logistics Regression

MNB – Multinominal Naïve Bayes

ML - Machine Learning (ML)

NLP - Natural Language Processing (NLP)

NSAT - Number of sync SATellites (NSAT)

OSM - Open Street Map OSM

OSRM - Open Street Routing Machine (OSRM)

POI - Point of Interest (POI)

PUP - Pick Up Point (PUP)

R - Walking Radius (R)

SML - Supervised Machine Learning (SML)

SVM – Support Vector Machine

TRID - Transport Research International Documentation (TRID).

TS - Transportation Science (TS)

UML - Unsupervised Machine Learning (UML)