UTS UNIVERSITY OF TECHNOLOGY SYDNEY

A DATA-DRIVEN DECISION SUPPORT SYSTEM FOR MOBILE TELEMATICS

by Mohammad SiamiNamini

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

under the supervision of Dist. Prof. Jie Lu, and Dr. Mohsen Naderpour

University of Technology Sydney
Faculty of Faculty of Engineering and Information Technology

October 2020

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Mohammad SiamiNamini declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Computer Science, Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 4/11/20

DEDICATION

To my darling wife for her passion and patience and to my beloved parents for their encouragement which enabled my dreams to come true.

And

To my lovely little son Ariyan, who made me stronger, happier and more fulfilled than I could have ever imagined. I love you to the moon and back.

ACKNOWLEDGEMENT

I would like to express my appreciation to Distinguished Professor Jie Lu, who supervised this Ph.D. research, for her knowledgeable comments, valuable support and guidance during hard times, and constructive suggestions along the way. I want to thank Jie for her wonderful scientific and spiritual support during this journey and her willingness to allow my research to follow my interests. I am also grateful to Dr Mohsen Naderpour, my co-supervisor, for his knowledgeable suggestions and valuable technical advice throughout my PhD study.

Looking back over my PhD journey, I see my wife, Anahita, truly standing shoulder to shoulder with me. Thank you very much for being so supportive in all circumstances throughout the four years of this Ph.D., for understanding the stress I was subjected to, for sacrificing your time for me while you were also completing a Ph.D. and for caring for our little son, Ariyan, to give me the freedom to follow my scientific interests. I could not have accomplished this without your constant love and support.

I would like to thank my parents, the first ones who taught me, for their encouragement and support despite the geographical distance. Pursuing a PhD would not have been possible without their love and assistance. I express my gratitude to all my friends and colleagues in the Decision Systems & e-Service Intelligence (DeSI) Laboratory for their help and valuable comments during my study.

I would also like to express my appreciation to the Faculty of Engineering and Information Technology and the Centre for Artificial Intelligence (CAI) at the University of Technology Sydney (UTS) for their provision of conference registration and travel funds during this research. This research was also supported by

an International Research Scholarship (IRS) and the UTS President's Scholarship (UTSP) funded by UTS.

Last, but not least, a special thank you goes to Ms Sue Felix, Ms Jemima Moore, and Ms Michele Mooney for helping me to identify and correct grammar and syntax problems in my publications.

ABSTRACT

Mobile telematics is an emerging technology that collects data on human behaviour using smartphones. All smartphones have internal sensors with the capability to record and transmit data to an external server. This emerging technology is easy to use, the initial cost is very low, and generates a massive amount of data which are noisy, complex, and uncertain. This opens many opportunities for data-driven decision making such as driving behaviour risk analysis, usage-based insurance, remote sensing, and fleet management. Traditional decision-making techniques are not able to work with this type of unstructured and complex data. Thus, new techniques are needed based on advanced analytics to analyze mobile telematics streams.

This research develops a big data-driven decision support system (DSS) for mobile telematics. The research relies on the capabilities of advanced analytics techniques, machine learning, and fuzzy logic. The research presents an innovative analytical system for mobile telematics which consists of four major components: 1) a data preparation component that prepares a trajectory dataset to a new and ready-for-analysis format; 2) a driving style pattern recognition that extracts hidden human patterns in mobile telematics using unsupervised learning and unlabelled data; 3) a fuzzy risk assessment is proposed to assess risk of drivers by fuzzy logic using extracted patterns by unsupervised learning; and 4) a missing data imputation component which is a novel Choquet Fuzzy Integral Vertical Bagging (CFIVB) algorithm to classify large labelled mobile telematics stream datasets.

The proposed models were evaluated on two real-world mobile telematics datasets, namely an unlabelled dataset collected by a usage-based insurance company

containing 500,000 journeys of 2500 drivers, and an anonymized driving behaviour dataset consisting of streaming data of 408 trips of 310 unique drivers. Various validation measures were used to evaluate the performance of the proposed models. The area under a curve (AUC) and accuracy are used to evaluate the classification algorithms and the Davis-Boulding index, the Calinski-Harabasz index, execution time, and mean square error are utilized to evaluate clustering algorithms and find the optimal number of clusters. The sensitivity analysis results show the proposed model is consistent across different variations of the model.

The proposed DSS can be applied on all stream data risk assessments. Moreover, 29 unique driving styles were extracted from mobile telematics data and these patterns can be applied as labels for supervised learning modelling. In addition, performance measures depict the CFIVB algorithm performs well in this domain, and it can be applied for similar problems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS