EXPERIMENTAL STUDY ON DRIVING SCENARIOS AND DRIVER BEHAVIOURS IN MALAYSIA BY USING MACHINE LEARNING TECHNIQUES

SALEM ABDULLAH SALEM GARFAN

UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021

EXPERIMENTAL STUDY ON DRIVING SCENARIOS AND DRIVER BEHAVIOURS IN MALAYSIA BY USING MACHINE LEARNING TECHNIQUES

SALEM ABDULLAH SALEM GARFAN

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

THESIS PRESENTED TO QUALIFY FOR A DOCTOR OF PHILOSOPHY

FACULTY OF ART, COMPUTING AND CREATIVE INDUSTRY
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

UNIVERSITI PENDIDIKAN SULTAN IDRIS
SULTAN IDRIS EDUCATION UNIVERSITY

Please tick (√)
Project Paper
Masters by Research
Master by Mixed Mode
PhD

INSTITUTE OF GRADUATE STUDIES

DECLARATION OF ORIGINAL WORK

This declaration is made on the ....10th....day of....June....20..21..

i. Student's Declaration:

I, Salem Abdullah Salem Garfan, ID:P20171001174, FSKIK (PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled Experimental Study on Driving Scenarios and Driver Behaviours in Malaysia by Using Machine Learning Techniques is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

Signature of the student

ii. Supervisor's Declaration:

I Dr.Bilal Bahaa Zaidan (SUPERVISOR'S NAME) hereby certifies that the work entitled Experimental Study on Driving Scenarios and Driver Behaviours in Malaysia by Using Machine Learning Techniques (TITLE) was prepared by the above named student, and was submitted to the Institute of Graduate Studies as a * partial/full fulfillment for the conferment of Doctorate of Philosophy in Artificial Intelligence (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.

DR. BILAL BAHAA ZAIDAN
Faculty of Art Computing and Creative Industry, UPSI
hp: 010 420 4377

13/6/2021
Date

Signature of the Supervisor

INSTITUT PENGAJIAN SISWAZAH /
INSTITUTE OF GRADUATE STUDIES

BORANG PENGESAHAN PENYERAHAN TESIS/DISERTASI/LAPORAN KERTAS PROJEK
DECLARATION OF THESIS/DISSERTATION/PROJECT PAPER FORM

Tajuk / *Title*: EXPERIMENTAL STUDY ON DRIVING SCENARIOS AND DRIVER BEHAVIOURS IN MALAYSIA BY USING MACHINE LEARNING TECHNIQUES

No. Matrik /*Matric's No.*: P20171001174

Saya / *I* : Garfan Salem Abdullah Salem

(Nama pelajar / *Student's Name*)

mengaku membenarkan Tesis/Disertasi/Laporan Kertas Projek (Kedoktoran/Sarjana)* ini disimpan di Universiti Pendidikan Sultan Idris (Perpustakaan Tuanku Bainun) dengan syarat-syarat kegunaan seperti berikut:-
acknowledged that Universiti Pendidikan Sultan Idris (Tuanku Bainun Library) reserves the right as follows:-

1. Tesis/Disertasi/Laporan Kertas Projek ini adalah hak milik UPSI.
   The thesis is the property of Universiti Pendidikan Sultan Idris
2. Perpustakaan Tuanku Bainun dibenarkan membuat salinan untuk tujuan rujukan dan penyelidikan.
   Tuanku Bainun Library has the right to make copies for the purpose of reference and research.
3. Perpustakaan dibenarkan membuat salinan Tesis/Disertasi ini sebagai bahan pertukaran antara Institusi Pengajian Tinggi.
   The Library has the right to make copies of the thesis for academic exchange.
4. Sila tandakan (√) bagi pilihan kategori di bawah */ Please tick (√) for category below:-*

☐	**SULIT/*CONFIDENTIAL***	Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub dalam Akta Rahsia Rasmi 1972. / *Contains confidential information under the Official Secret Act 1972*
☐	**TERHAD/*RESTRICTED***	Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan ini dijalankan. / *Contains restircted information as specified by the organization where research was done.*
☑	**TIDAK TERHAD / *OPEN ACCESS***	

(Tandatangan Pelajar/ Signature)

DR. BILAL BAHAA ZAIDAN
Faculty of Art Computing and Creative Industry, UPSI
hp: 010 420 4377
Date: / /20

(Tandatangan Penyelia / *Signature of Supervisor*)
& (Nama & Cop Rasmi / *Name & Official Stamp*)

Tarikh: 13th June 2021

Catatan: Jika Tesis/Disertasi ini **SULIT** @ **TERHAD**, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan ini perlu dikelaskan sebagai **SULIT** dan **TERHAD**.

Notes: If the thesis is CONFIDENTAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction.

ACKNOWLEDGEMENT

First and Foremost all praise and glory to ALLAH, the Almighty, the greatest of all, on whom ultimately we depend for sustenance and guidance. I would like to thank Almighty Allah for giving me the opportunitie, determination and to do my research. May his continuous grace and mercy be with me throughout my life and give me all power, wisdom, knowledge and right guidance to spread the right knowledge endlessly.

I would like to express my special appreciation to each person has contributed in any way to the work was done in this thesis. First of all, I would like to express all my thanks and gratitude to my supervisor Dr.Bilal Bahaa Zaidan for accepting me into his group since our first meeting. I thank him for believing in me, regardless of his doubt at the begining due to my education background ☺, and I hope to be one of his students whom makes him proud of us. Since we've met, you became a friend and a big brother for me, not just a mentor. To be honest, I could not have imagined having a better supervisor for my study. May ALLAH bless you, your family and all your beloved ones always.

I would like also to thank my co-supervisor Dr.Suliana Binti Sulaiman for her kindness and support from the beginning of my study. I would like also to express my warm gratitude and appreciation to Dr.Aws Alaa Zaidaan and Dr.Osamah Shihab Albahri for their limitless guidance, mentoring and support. I would like also to thank all participants for giving me their valuable time and be part of this thesis.

I would like also to thank all my friends whom I've gained during my study. We've been as family and shared together our happiness and sadness. We've stayed together all days and nights studying and motivating each other to accomplish our tasts. May our friendship last forever and may ALLAH bless you always.

I cannot thank enough my parents and siblings for believing in me and their endless support. Your were my biggest supporters and the main reason to be as I am now. I hope I've made you proud. May ALLAH bless you and wish you endless good health and happiness.

Last but not least, I would like to express my warm thanks to my dearest friend and brother Abdullah Hussein Alamoodi for making my study life easier. We've begun our study journey together from the bachelor degree to the master degree and the Ph.D. I've been admiring your continued hard working and motivation. May ALLAH bless you and I wish you continued success in your life.

DEDICATION

I dedicate this work to my loving family for their endless sacrifies, constant love and support. This thesis is dedicated to my father, the man whom raised me and taught me how to be the man I am today; to my mother for her non-stop prayers, support, and love throught my life; and to both my sisters and brother for their endless support and encouragements.

To me dear friends, whom I was blessed by knowing each one of you, I dedicate this thesis to you. Your support was a major role to finish this work; whether by direct assistance or by just standing by me always. Some of you have played a major part of this work and I couldn't be able to accomplish this work as it is now without your help and assistance.

To the love of my life, my wife, whom I haven't known you yet, this thesis is dedicated to you. You should know that this was not an easy work to do, and I was getting motivated whenever I was thinking how proud will you be after I finish this work.

To my beloved kids, whom I haven't met your mom yet, I dedicate this work to you. I've been working hard to achieve this level due to the importance of education and for better living life ahead and to make you proud of your father.

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

ABSTRACT

The increasing number in annual road fatalities has caused a major challenge in many countries. Minimising fatalities and improving safety are the top priorities of different countries. This study aimed to analyse driver behaviours in Malaysia and the impacts of practising eco-driving to improve safety, reduce fuel consumption and green gas emission by using smartphone sensors and OBD2 (ELM327) adapter based on event thresholds and machine learning algorithms. In the experimental study, 30 drivers had participated, which were 17 novice drivers (7 males and 10 females) and 13 experienced drivers (8 males and 5 females). A Honda Civic 2019 car was used in the experiment. A specific route was selected for all drivers, which consisted of two types of road (highway and urban), with a total distance of 20.6 km. The analysis of driving behaviour was based on threshold events and machine learning algorithms. This was to classify the different driving scenarios. In the driver's profiling, driving behaviour was categorised into three driving behaviours, such as safe, normal, and aggressive driving. Random Forest model was selected for the classification after being compared to other different machine learning algorithms (Decision Tree, Support Vector Machine, K-Nearest Neighbour, and Naïve Bayes models). The results of this experiment showed that a remarkable reduction in terms of fuel consumption and CO_2 emission of up to 30% less was achieved when participants followed the eco driving techniques. Moreover, aggressive events were notably reduced in eco driving as compared to normal driving. Furthermore, the selected machine learning model was able to differentiate and classify different driving scenarios with high classification accuracy of up to 100 %, such as identifying male and female drivers, novice and experienced drivers, and driving in the highway or city.

Keywords: driver behavior, machine learning, smartphone, OBD2, classification

KAJIAN EKSPERIMEN PENGAMBILAN SCENARIOS DAN PERINGKAT PEMANDU DI MALAYSIA MENGGUNAKAN TEKNIK PEMBELAJARAN MESIN

ABSTRAK

Peningkatan dalam jumlah tahunan kematian di jalan raya memberikan cabaran besar di kebanyakan negara. Salah satu keutamaan negara - negara ini adalah untuk mengurangkan kadar kematian dan meningkatkan keselamatan. Kajian eksperimen ini bertujuan untuk menganalisa tingkah laku pemandu di Malaysia dan kesan terhadap pemanduan secara berhemah dari segi peningkatan keselamatan, pengurangan penggunaan bahan bakar, dan pelepasan gas rumah hijau dengan menggunakan sensor telefon pintar dan penyesuaian OBD2 (ELM327) berdasarkan peristiwa yang ditetapkan dan algoritma pembelajaran mesin. Seramai 30 pemandu terlibat dalam kajian eksperimen, iaitu 17 pemandu pelatih (7 lelaki dan 10 perempuan) dan 13 pemandu berpengalaman (8 lelaki dan 5 wanita). Kereta Honda Civic 2019 digunakan dalam percubaan. Laluan tertentu yang dipilih terdiri daripada dua jenis jalan raya, iaitu lebuh raya dan jalan bandar dan merangkumi jarak 20.6 km. Analisa terhadap tingkah laku pemandu diklasifikasikan mengikut senario berbeza dan berdasarkan peristiwa yang ditetapkan dan algoritma pembelajaran mesin. Tingkah laku pemandu dikategorikan kepada tiga tingkah laku, iaitu pemanduan yang selamat, normal dan agresif. Model Random Forest dipilih untuk mengklasifikasikan setelah dibandingkan dengan algoritma pembelajaran mesin lain (*Decision Tree, Support Vector Machine, K-Nearest Neighbour, and Naïve Bayes models*). Hasil eksperimen ini menunjukkan penurunan kadar luar biasa dengan pengurangan sebanyak 30% dari segi penggunaan bahan bakar dan pelepasan CO_2 ketika peserta mengikuti teknik memandu secara berhemah. Sementelahan itu, peristiwa agresif dapat dikurangkan dengan pemanduan berhemah berbanding dengan pemanduan biasa. Selain itu, model pembelajaran mesin yang dipilih dapat membezakan dan mengklasifikasikan senario pemanduan berbeza dengan mencapai ketepatan hingga 100%, seperti mengenal pasti jika pemandu ialah lelaki atau wanita, pemandu pelatih atau berpengalaman, dan memandu di lebuh raya atau di bandar.

TABLE OF CONTENTS

CHAPTER 2 LITERATURE REVIEW

TABLE OF FIGURES

TABLE OF TABLES

TABLE OF EQUATIONS

LIST OF ABBREVIATIONS

AHP	Analytic Hierarchy Process
ANN	Artificial Neural Networks
BN	Bayesian Network
CO	Carbon Monoxide
CO_2	Carbon Dioxide
CO_2e	Carbon Dioxide Equivalent
CDS	Context-Based Driver Score
CNN	Convolution Neural Network
CSV	Comma-separated Values
DNN	Deep Neural Network
DT	Decision Tree
DTW	Dynamic Time Warping
Eco-driving	Economic Driving
ECU	Electronic Control Unit
FN	False Negative
FT	False Positive
g/km	Grams Per Kilometer
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HDV	Heavy Duty Vehicles
HMM	Hidden Markov Model
Hz	Hertz

IMS	Intelligent Mechatronic Systems Inc.
IoT	Internet of Things
IoV	Internet of Vehicles
K-NN	K-Nearest Neighbors
Km	Kilometer
KT	Metric Kiloton
MIROS	Malaysian Institute of Road Safety Research
ML	Maximum Likelihood
MEMS	Micro-Electromechanical Systems
M/S^2	Meter per second squared
MT	Million Tonnes
MVN	Multivariate Normal Distribution
NLPCA	Non-Linear Principal Components Analysis
NO_2	Nitrogen Oxides
OBD	On Board Diagnostics
PCA	Principal Component Analysis
PM	Particulate Matter
RBF	Radial Base Function
RF	Random Forest
RNN	Recurrent Neural Networks
RPM	Rotation Per Minute
S	Seconds
SSAE	Stacked Sparse Auto-Encoder

SVM	Support Vector Machine
SAX	Symbolic Aggregate Approximation
SO_2	Sulphur Dioxide
TN	True Negative
TP	True Positive
UW	University of Waterloo
UBI	Usage-Based Motor Insurance
UI	User Interface
UPSI	Univirsiti Pendidikan Sultan Idris
V2V	Vehicle-to-vehicle
VDOT	Virginia Department of Transportation
WoS	Web of Science

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The contents of this chapter explain different aspects of this thesis. The aspects covered in this chapters are Research Background, Problem Statement, Research Objective, Research Question, Research Scope, and Thesis Outline. In Research Background, the reader will be having a bright insight about the origin of the topic. In Problem Statement, the problem of the research will be intensively explained. The objectives of this research to be done will be stated in Research Objectives. In Research Questions, the main questions that need to be answered in this research are provided. Research Scope provides details on the areas and factors that this research will include and exclude. In thesis Layout, the structures of the thesis are provided.

1.2 Research Background

Driving is one of the necessary tasks in most of daily peoples' life, and at the same time it is considered as association with quality of life in different cases (Vasconcelos et al., 2017). The number of vehicles usage on road is increasing due to the increasing in population. According to the report of Malaysian Institute of Road Safety Research (MIROS), the population number and registered vehicles number are increasing year by year ("Road Facts,"). Hence, the accidents number is increasing. The following Table 1:1 is extracted from MIROS report which is showing the increasing number of accidents in Malaysia.

Table 1.1

Accidents Number in Malaysia

Year	Registered Vehicles	Population	Road Crashes	Road Deaths
2013	23,819,256	29,947,600	477,204	4,915
2014	25,101,192	30,300,000	476,196	6,674
2015	26,301,952	31,190,000	489,606	6,706
2016	27,613,120	31,660,000	521,466	7,152

The increasing number of car accidents is a major challenging in many countries. The consequences of accidents can harm individuals and governments in term of social and economic loss which impact the society. Traffic road accidents are causing more than a million deaths per year over the world, and it is predicted to be the fifth leading cause of death by 2030 (Jarret Engelbrecht, Marthinus Johannes Booysen, Gert-Jan van Rooyen, & Frederick Johannes Bruwer, 2015). According to the

Department of Statistics of Malaysia, transports accidents recorded the fourth highest cause of death in 2016 ("Statistics on Causes of Death, Malaysia," 2017). According to MIROS report conducted on Malaysian motorcyclists, the age group between 16 – 25 years old is involving in more than 40% of the total fatalities (Azzuhana Roslan, 2017). Improving traffic safety today is one of the top priorities list of different countries around the world (Saiprasert, Thajchayapong, Pholprasit, & Tanprasert, 2014). Understanding the main causes of accidents will assist finding solutions to prevent or reduce traffic accidents. According to (Camlica, Hilal, & Kulić, 2016), human errors, road conditions, and vehicle failure are the main contributing factors to road traffic accidents. Human error (e.g. lack of attentions, incorrect control, aggressive driving) is the most influential factor in contributing to traffic accidents. As stated in (Allamehzadeh & Olaverri-Monreal, 2016), human error is causing around 90 percent of traffic accidents. Thus, analyzing and understanding driver behavior is the key factor to prevent or reduce traffic accidents. The term of driver behavior can be defined as different concepts related to driving mannerisms and driving actions of a driver which introduce everlasting variables (Meiring & Myburgh, 2015). Studying driving behavior not just assist in road safety improvement, it also assist in safe environment by finding solutions to reduce fuel consumptions and gas emissions.

According to the department of environment in Malaysia, the increasing number of air pollution due to the increasing trend of industrial sources and numbers of vehicles. Vehicles' emission is the major source of the air pollution especially in the urban areas (Environment, 2019). Based on the environment's department of Malaysia, the overall air pollutant's accumulation emission load in 2018 was 2,210,634 metric tons of carbon monoxide (CO), 95.6% of the emission source was from motor vehicles; 889,890 metric

tons of nitrogen oxides (NO_2), 25% of the emission source was motor vehicles; 257,457 metric tons of Sulphur dioxide (SO_2), 6% of it from motor vehicles; and 26,789 metric tons of particulate matter (PM), 15% of the emission load was from motor vehicles (Environment, 2019). Based on Global GHG and CO2 Emissions report, carbon dioxide (CO_2) emissions in 2018 reached up to 257.84 metric kiloton (kt) in Malaysia, 24.69% of the emission load was from transport sector (Crippa et al., 2019). The following table 1-2 illustrates the CO2 emissions in Malaysia from 2014 to 2018 in kt.

Table 1.2

CO2 Emissions in Malaysia from 2014 to 2018

Year	Total CO2 emissions (kt)	CO2 emissions in transport sector (kt)	CO2 emissions per capita (mt)
2018	257.84	63.66	8.05
2017	246.77	62.06	7.80
2016	246.36	62.80	7.90
2015	249.58	61.41	8.12

Reducing fuel consumption and gas emissions are primary goals in different countries (Günther, Rauh, & Krems, 2017). The increasing number of vehicles is one of the main causes of environmental pollutions due to the gases emissions from the vehicle engine while burning the fuel such as carbon dioxide (CO2) and acid rain, as well as these greenhouse effects increase the possibility to be infected with some lung diseases such as lung cancer (Magaña & Muñoz-Organero, 2016). Hence, it is necessary to minimize the usage of fuel in order to reduce the impact of vehicles on the environment. According to a study (Zaid, Myeda, Mahyuddin, & Sulaiman, 2015),

greenhouse gas emissions is rapidly growing in Malaysia due to the growth in industrial and transportation sectors, with a total of 194 million tons in 2011 to predicting to reach over 285 million tons in 2020. Practicing economic driving (eco-driving) strategies is one of the best solutions in minimizing fuel consumption, greenhouse gas emissions, and improving traffic safety (Husnjak, Forenbacher, & Bucak, 2015). The term of eco-driving can be defined as applying set of rules during driving aimed to minimize fuel consumption and improve safety (Rionda et al., 2014). Based on this definition, it can be found that modifying driving style could improve the efficiency of driving which leads to improve safety and reduce energy consumption. The following table 1-3 illustrates the potential impact of practicing eco-driving on different areas of activities. Encouraging drivers to adopt eco-driving strategies can be through different ways such as prior education, post-drive statistics, or real-time in-vehicle feedback. Among these ways, it has been found that real-time in-vehicle feedback is working well (Thill & Riveiro, 2015).

Table 1.3

The Potential Impact of Eco-Driving (Husnjak, Forenbacher, Et Al., 2015)

Activity area	**Potential Impact**
Safety	Road safety improvement Driving skills improvement
Environment	Greenhouse gas emission reduction Reduction of noise
Economic	Fuel consumption reduction Reduction of maintenance costs Reduction of traffic accidents costs
Social	Driving awareness improvement Reduction of stress during driving Increasing ride comfort

Measuring and understanding behavior of driver while driving can be achieved through deploying in-vehicle sensors. These sensors aimed to measure different aspects of daily activities in vehicle operation (Wallace et al., 2015). In the last decades, the technology has been rapidly advanced especially in the areas of smartphone and wearable devices. Smartphone and wearable devices are low cost, widely available, and easily accessible (Saiprasert, Pholprasit, & Thajchayapong, 2017). The improvement of the capabilities of smartphone devices by the variety of embedded powerful sensors included such as global positioning system (GPS) and accelerometer opened the opportunity to measure driving behavior using these devices instead of the high cost installation of external hardware in the vehicles (Saiprasert et al., 2014). Additionally, On Board Diagnostics (OBD) adapters provide an access to read the status of different in-vehicle sensors such as fuel rate and oxygen, and at the same time OBD adapters enable a transparent connectivity between the smartphone device and electronic control unit (ECU) of the vehicle (Meseguer, Toh, Calafate, Cano, & Manzoni, 2017).

It has been mentioned in a study that eco-driving is a very cost-effective strategy aimed to reduce fuel consumptions up to 7%, reduce NOx emissions up to 5%, reduce PM2.5 emissions up to 7%, and reduced CO2e around 700 metric tons per year. In a transit service bus, the amount of fuel savings around 208,200 Liters of diesel per year (Y. Xu, Li, Liu, Rodgers, & Guensler, 2017).It has been stated in another study collected different studies' results that eco-driving contributed to less accidents up to 35% during the test drives(Valdemars, Atstaja, & Vasiljeva, 2015).

1.3 Problem Statement

On the basis of the above, Malaysia, as many countries, is facing major challenges related to road safety and a concern is arising towards increasing the rate of accidents and fatalities year after year. Furthermore, the effects of the increasing rate of CO_2 emission on the environment is concerning. Vehicles' emission is one of the major sources of air pollution. Hence, finding practical solutions to reduce air pollutions and improve safety is important.

Understanding drivers' behavior is essential to develop drivers modeling (Hansen, Busso, Zheng, & Sathyanarayana, 2017), (Sysoev, Kos, Guna, & Pogačnik, 2017) and to improve drivers' driving efficiency (Tselentis, Yannis, & Vlahogianni, 2017), (Oren Musicant & Lotan, 2016), (Jo, Kim, Park, & Yoon, 2015). However, achieving effective ways to maximum fuel efficiency without obstructing the internal vehicle's structure and providing at the same time accidents' prevention or reduction system is challenging (Khanapuri, Shastri, D'souza, & D'souza, 2015). According to (Vaezipour, Rakotonirainy, Haworth, & Delhomme, 2017), the lack of practicing eco-driving by drivers is decreasing their driving safety and increasing their risk of occurring accidents. Complete solutions related to eco driving using smartphone are lacking (Husnjak, Forenbacher, et al., 2015). Utilizing the new technologies such as artificial intelligence, machine learning, and deep learning is important to find innovative ways to overcome the current challenges. However, with the limited availability of data, it is challenging to develop effective solutions (Simmons, Caird, & Steel, 2017), (Huang et al., 2016), (Aichinger, Nitsche, Stütz, & Harnisch, 2016), (Brezger & Albers, 2013).

On the basis of the systematic review was running on the driver behavior domain with the scope of smartphone, massive challenges were extracted from the related articles. However, the critical analysis was concentrating on the challenges and limitations within the field of eco-driving due to its impacts on improving road safety and reducing CO_2 emission. Hence, several gaps and issues were identified related to eco-driving. The following figure 1-1 illustrates the major challenges obsereved from the literature review.

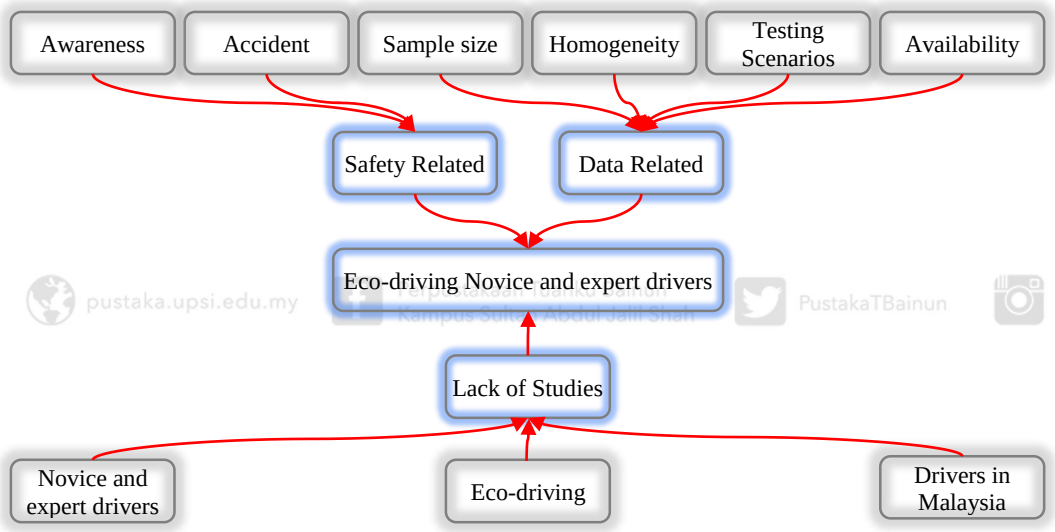

Figure 1.1. Challenges Observed from the Literature

As seen in the above figure 1-1, the challenges observed in the literature were classified into three major groups, safety related, data related and lack of studies. The first challenges group is *Safety related challenges*. Lack of safety was one of major issues in the literature. As found in the literature, driving aggressively is a common phenomenon that needed to be considered in order to improve safety as pointed out in (Dobbins & Fairclough, 2017), (Pratama, Ardiyanto, & Adji, 2017). Moreover, drivers' errors is highly involved in causing traffic accidents as mentioned in (Alaybeyoglu &

Senel, 2017), (Osafune et al., 2017). Moreover, lack of awareness was a major challenge in the literature such as low level awareness of drivers about traffic (Alaybeyoglu & Senel, 2017), awareness about potential results of risk driving (Feraud, Lara, & Naranjo, 2017), maintaining self-awareness of drivers (Bi et al., 2017), and health condition awareness of drivers (Natpratan & Cooharojananone, 2015). The second group of challenges is *Data related challenges*. Different issues were found on the studies' sample. Many studies in the literature stated that the small sample size and population is an issue as mentioned in (Simmons et al., 2017), (Günther et al., 2017). Other studies pointed out different issues related the sample as the sample selection (Vaezipour et al., 2017), (Ryder, Gahr, Egolf, Dahlinger, & Wortmann, 2017), homogeneity of the sample (Ryder et al., 2017), limited tested scenarios (Eboli, Mazzulla, & Pungillo, 2016), (Orfila, Saint Pierre, & Messias, 2015), and datasets availability (Alam, Hariz, Hosseinioun, Saini, & El Saddik, 2016), (Romera, Bergasa, & Arroyo, 2016). The third group of challenges is *Lack of studies.* The studies on the effectiveness and efficiency of eco-driving are still lacking (Günther et al., 2017), (Y. Xu et al., 2017). There is a lack of studies related to driving behaviors in Malaysia, Novice driver behavior and eco-driving behavior with respect to Malaysian drivers and novice Malaysian drivers. Apart from that, the exploration of machine learning usage with different scenarios related to novice and expert Malaysian driver behavior.

Based on the critical analysis of the systematic review in section 2.5, different real-time smartphone systems were introduced for eco-driving evaluation based on fuel consumptions and drivers' classification. However, none of the proposed systems for drivers' classification have classified aggressive drivers based on the whole trip, the proposed systems were just calculating the defined aggressive events. Moreover, most

of the proposed systems have not mentioned the aggressive thresholds clearly of the defined events, except for (Magaña & Muñoz-Organero, 2016). Further, different factors were not considered while evaluating the proposed systems in term of fuel consumptions such as the fuel consumption is differ from vehicle to another.

Based on the systematic review conducted in the domain of driver behavior using smartphone, few articles only have been done in the area of eco-driving; most of these articles have been conducted in Spain, and none of these articles were in Malaysia or have studied the Malaysian drivers. Moreover, there is no homogeneity in their sample of the study which means there is no specific age range or driving experience was concentrated in the study. Furthermore, most of the scenarios in the experiments were limited and some studies did not mentioned the scenarios while collected the data such as the weather and geographic location. Further, the sample size of most of the studies was very small. Moreover, no study was claiming the aggressive driver based on the whole trip, the previous studies were only showing the aggressive events. Hence, machine learning needs to be utilized and modulated to discover if it is possible to classify aggressive driving based on the whole trip. These issues are claimed mainly from 2.5. Therefore, this research aims to explore the impacts of eco-driving on improving driving efficiency and the possibilities to identify normal and eco driving based on the whole trip. The following figure 1-2 illustrates a summary of the different issues and concerns with the respect to eco-driving behavior.

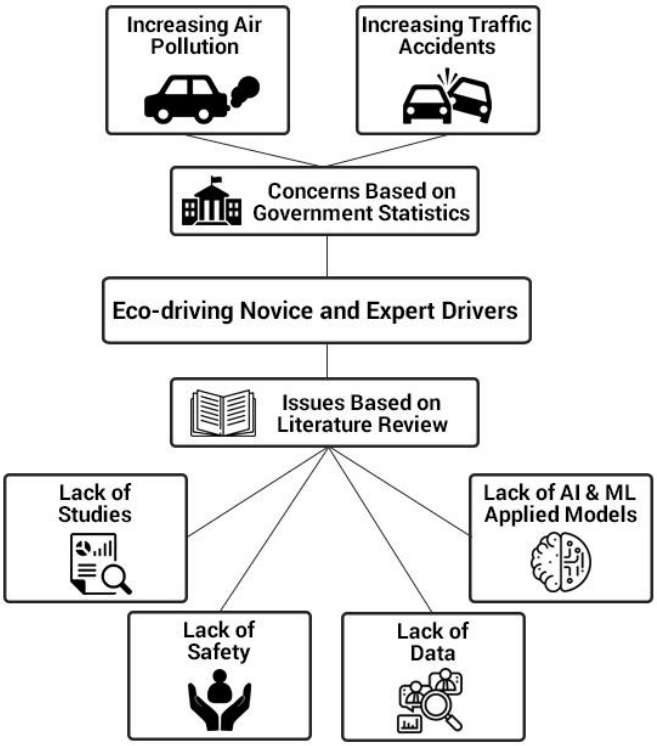

Figure 1.2. A Summary of the Current Issues and Concerns

This results a lack of information with respect to novice and experienced driving behaviors with different driving style (normal and eco-driving). In addition to the investigation of the capacity of machine learning to handle this classification problem, this research is investigating the mentioned issues and attempt to develop a machine learning based driver behavior analysis.

1.4 Research Objectives

The aim of this research is to explore novice and expert drivers' behaviors in Malaysia based on the collected data and study the impact of eco-driving on improving road safety, fuel consumption, and CO2 emission. In addition to that, to identify the possible

ways to encourage drivers practicing eco-driving techniques in order to improve their driving efficiency. The main objectives of this research are, as the following:

1) To investigate the current literature of driver behavior using smartphone systematically.
2) To explore experimentally the driving behavior of novice and expert drivers and the impacts of eco-driving in Malaysia using smartphone.
3) To analyse drivers behavior using drivers' profile via threshold events and the impacts of eco-driving on safety, fuel consumption, and CO_2 emission.
4) To modulate and optimize a machine-learning-based classifier to differentiate between normal and eco driving behavior.
5) To evaluate the different driving scenarios based on machine learning (e.g. male vs female drivers, normal vs eco driving).

1.5 Research Questions

The main questions of this research, are as following:

1) What are the current state of art in regards with the studies of driver behavior using smartphone in the academic literature?
2) What is the driving behavior of novice and expert drivers in Malaysia, and to what extent eco-driving could improve their driving behavior?
3) To what extend eco-driving could impact drivers' safety, environment, and economic?

4) Is it possible to distinguish between normal and eco driving within the study case based on the given data?
5) Is there any significant differences among the different driving scenarios (i.e. highway vs city)?

1.6 Research Scope

Few aspects need to be taken into account regarding the scope of this research as the following:

- This study is concentrating on private light car drivers only; all other types of vehicles such as heavy vehicles, motorbikes, and bikes are out of our scope.
- This study is focusing on both novice and expert drivers to be able to compare between them in term of fuel saving, CO2 emission, and aggressive driving.
- Weather and driving time factors (day or night) are included in our scope of study.
- The experiment of this study is taking a place in Malaysia.
- This study is measuring the CO2 emission only. The other greenhouse gas emissions such as CO, NOx, and PM are excluded due to the need to install extra sensors on the vehicles and there is no benchmark on the required sensors to measure them.

1.7 Research Significance

The results of this research will contribute toward solving issues in different areas. First of all, this research should contribute for better understanding and provided a vision of novice and expert drivers' behaviors in Malaysia, which will assist the government to apply or provide suitable solutions in order to reduce or limit traffic accidents and at the same time this would open a market opportunity for any other party to apply or provide their solutions in different areas such as to reduce traffic accidents or reduce fuel consumption. As mentioned in the above Table 1-2, practicing eco-driving would contribute in the areas of safety, environment, economic, and social. From safety perspective, practicing eco-driving by high segment of the society would improve driving efficiency and improve road safety. Moreover, eco-driving from environment point of view would contribute in the reduction of noise and greenhouse gas emission. Economically, practicing eco-driving will contribute in reducing fuel consumption, cost of maintenance, and cost of traffic accidents. Eco-driving would contribute to the social by increasing the awareness of drivers while driving, ride comfort, and stress reduction while driving.

1.8 Thesis Layout

This research is consisting of seven chapters and they are as follow:

Chapter One: a slight background description of the domain of driver behavior with smartphone and eco-driving are provided. Additionally, the gap of this research is stated

followed with research objectives and research questions. Then, the scope of this research is defined. After that the significance of the research is provided.

Chapter Two: a comprehensive investigation is carried out on the domain of driver behavior with smartphone. In this investigation, the terms of the search on the current literature are defined. The taxonomy is drawn up based on the results of the systematic literature review protocol that adapted in this research to review and analyze the literature. Moreover, different important elements are extracted from the related articles and explained such as challenges, motivations, and recommendations for the domain of driver behavior with smartphone. Moreover, different methodological aspects are provided followed with a critical analysis on the current literature of the domain of driver behavior with smartphone.

Chapter Three: the research methodology and the flow of the research are designed and explained in this chapter.

Chapter Four: an overview of the dataset's attributes is provided. Followed with comprehensive details about the experiment phase and the process of it. Additionally, features extraction rules for drivers' profile are provided. Moreover, machine learning features are presented and explained.

Chapter Five: The results of the collected data are provided in this chapter. It starts from eco driving learning process. And then fuel consumption, CO2 emission, and aggressive driving results are provided for each driver in normal and eco driving mode.

Chapter Six: This chapter presents machine learning results. Machine learning process is highlighted. Then different machine learning models are addressed and compared to select the highest accuracy. After that, different scenarios are presented and compared based on the selected machine learning model.

Chapter Seven: A discussion of the research findings are provided. Then, research achievements and research limitations are provided. At the end of the chapter, the conclusion and future works are presented.

1.9 Chapter Summary

In this chapter, a concern about increasing road accidents and effecting the environment by increasing the CO2 emission caused by vehicles were provided. Then, the problem statement of the research was discussed and provided. After that, the research aims and research objectives were presented. Next, the research questions and research scope were given. Then, research significances were highlighted. At the end, thesis layout was explained. The following figure 1-2 highlights the research flow chart.

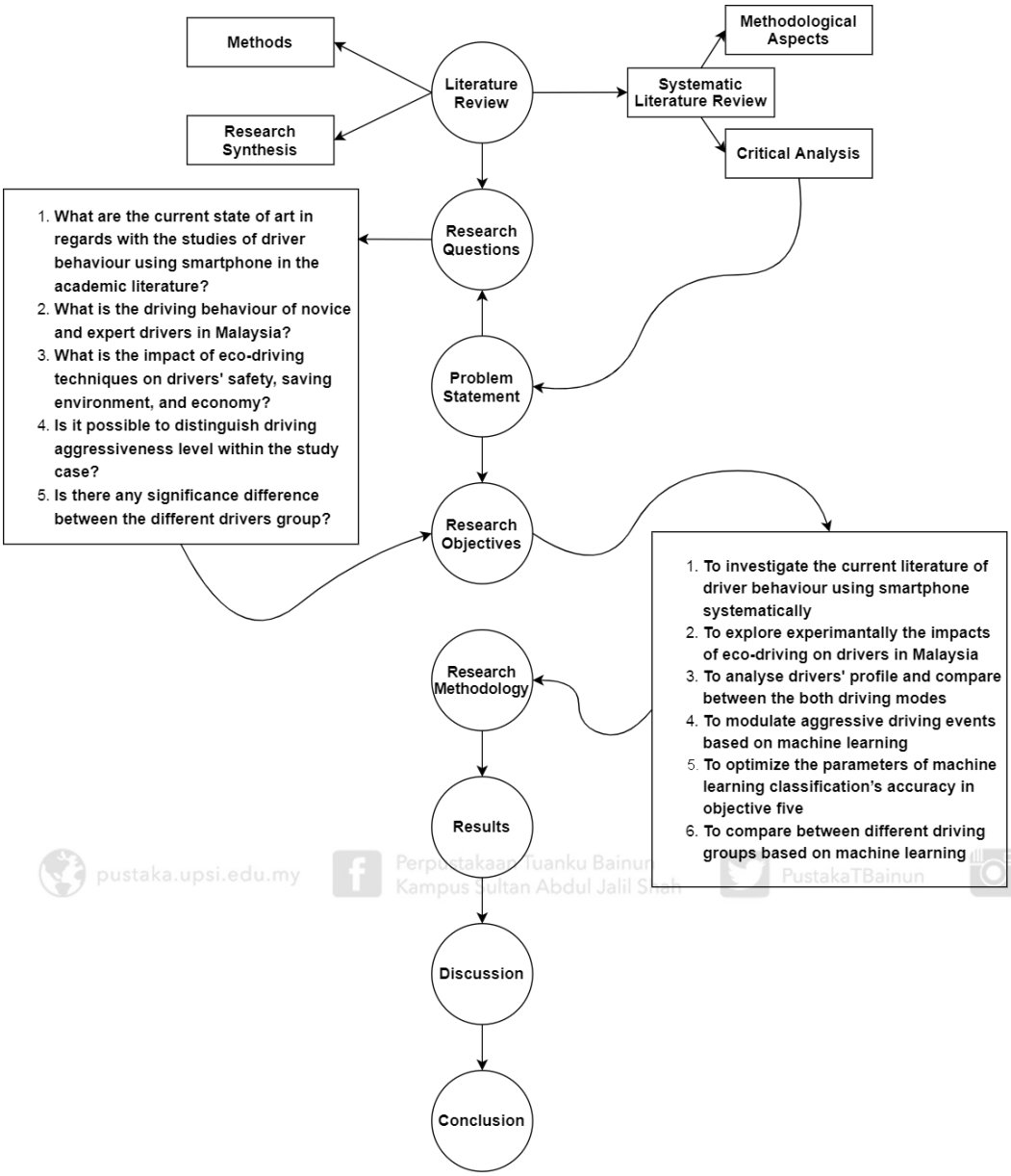

Figure 1.3. Research Flow Chart