UTM UNIVERSITI TEKNOLOGI MALAYSIA

SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCST1043
Sem:…1…. Session:……2024/2025…

SECTION A: Project Information.

Program Name: **Masters of Science (Data Science)**

Subject Name: **Project 1 (MCSD 6215)**

Student Name: Liao Xiuya

Metric Number: MCS241025

Student Email & Phone: liaoxiuya@graduate.utm.my

Project Title: Predictions With Global Electric Vehicle Market

Supervisor 1:

Supervisor 2 / Industry Advisor(if any):

SECTION B: Project Proposal

Introduction:

Since 2010, driven by technological progress, government policy support and environmental awareness, the electric vehicle industry has developed rapidly and gradually occupied the market of the automotive industry. At present, the electric vehicle market is mainly popular with pure electric vehicles (BEV), plug-in hybrid electric vehicles (PHEV), fuel cell vehicles (FCEV) and hybrid electric vehicles (HEV). This study focuses on this trend and collects a comprehensive data set on global electric vehicle sales on authoritative platforms such as kaggle and the IEA Global EV Data Explorer. Focus on data from well-known countries with high electric vehicle penetration and market influence, including China, the United States, Germany and the United Kingdom.

These data cover historical sales data, charging infrastructure and other related indicators from 2010 to 2023. This study aims to use various data analysis tools, such as predictive models, machine learning, data analysis, etc., to analyze historical data and make a rough analysis and prediction of the future development trend of the electric vehicle industry.

Problem Background:

The global automotive market's shift toward electric vehicles (EVs) is driven by environmental concerns, technological advances, supportive government policies, and changes in consumer preferences. With the rapid growth of EVs since 2010, and the increasing variety of EV energy sources, which energy source will dominate the future EV market? This question is of particular concern to automotive companies, especially when it comes to their future development plans, and is also of concern to many consumers.

This study uses data science tools such as predictive modeling, machine learning, and time series analysis to analyze historical EV sales data from 2010 to 2020 and forecast future trends to 2030. The data for this study comes from Kaggle and IEA Global EV Data Explorer, focusing on major countries such as the United States, China, Germany, and Norway. This study provides stakeholders with valuable insights to respond to the changing EV market.

Problem Statement:

Here are some ways to formulate the problem of this project:

1. How to analyze historical sales data of various types of electric vehicles to identify the main trends and patterns that have emerged since 2010?
2. How to predict future trends: Accurate predictions of future electric vehicle development trends?

3. How to deliver strategic insights like advise automakers, policymakers, and investors to support informed decision-making and strategic planning?

Aim of the Project:

The goal of this study is to use the data sets of Kaggle and IEA Global EV Data Explorer to focus on major countries with high electric vehicle penetration and market influence in the world, so as to fully understand the dynamics of the electric vehicle market and effectively analyze and predict the future development of the electric vehicle industry.

Objectives of the Project:

1. To Analyze historical trends: Examine the sales data of various types of electric vehicles in major countries around the world from 2010 to 2020 to identify and analyze the main trend characteristics.
2. To Forecasting future sales: Leverage machine learning and time series analysis to develop forecasting models to forecast EV sales trends through 2030.
3. To Evaluate influencing factors: Evaluate the impact of technological advances, government policies, market dynamics, and consumer preferences on all types of electric vehicles.
4. To Deliver strategic insights: Advise automakers, policymakers, and investors to support informed decision-making and strategic planning.

Scopes of the Project:

1. Historical sales data: Analyze global EV sales from 2010 to 2020, with a focus on key country markets.
2. Technological Advancements: Assess the impact of energy options and vehicle design improvements.
3. Government policy: Evaluate the impact of policies, incentives, and regulations.
4. Market dynamics: Understand consumer preferences and economic factors.
5. Predictive modeling: Develop models to forecast EV sales trends through 2030.

Expected Contribution of the Project:

1. EV Market Trend Analysis: Provides a detailed analysis of global EV sales trends from 2010 to 2020, highlighting the trends driving changes in key markets such as the US, China, Germany, and Norway.
2. Future Sales Forecast: Develop forecasting models to predict EV trends in 2030, providing valuable forecasts.
3. Impact Analysis: Evaluate the impact of technological advancements, government policies, etc. on the EV market, and provide a comprehensive understanding of the various vehicle development trends in the EV market.
4. Strategic Advice: Provide valuable insights and recommendations to automakers, policymakers, and investors to support informed decision-making and strategic planning for the future electric market.

Project Requirements:

Software: R, Rstudio, Google Collab and Python

Hardware: ROG Strix G16 G614JVR

Technology/Technique/Methodology/Algorithm: Machine learning and time series analysis

Type of Project (Focusing on Data Science):

[X] Data Preparation and Modeling

[X] Data Analysis and Visualization

[ ] Business Intelligence and Analytics

[X] Machine Learning and Prediction

[X] Data Science Application in Business Domain

Status of Project:

[X] New

[ ] Continued

If continued, what is the previous title?

SECTION C: Declaration

I declare that this project is proposed by:

[X] Myself

[ ] Supervisor/Industry Advisor ( )

Student Name: LIAO XIUYA

LIAO XIUYA

Signature

16 November 2024

Date

SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

I/We agree to become the supervisor(s) for this student under aforesaid proposed title.

Name of Supervisor 1:

Signature **Date**

Name of Supervisor 2 (if any):

Signature **Date**

SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

Result:

[ ] FULL APPROVAL

[ ] CONDITIONAL APPROVAL (Minor)

[ ] CONDITIONAL APPROVAL (Major)*

[ ] FAIL*

***** Student has to submit new proposal form considering the evaluators' comments.

Comments:

Name of Evaluator 1:

Signature **Date**

Name of Evaluator 2:

Signature **Date**

CHAPTER 1

INTRODUCTION

1.1 Overview

In recent years, driven by technology and policy, the electric vehicle (EV) market has developed rapidly, and the world's major automakers, including Tesla, BYD and Nissan, have continued to innovate and create many high-tech, high-performance, and high-quality electric vehicle models, which are deeply loved by market consumers.

The widespread popularity of electric vehicles not only reduces greenhouse gas emissions and helps protect the environment, "and also stimulates the growth of the new energy industry and profoundly changes the way people live their lives. Therefore, numerous countries and regions have implemented various policies to promote the advancement of electric vehicles, such as car purchase subsidies, tax exemptions as well as the development of charging infrastructure, etc., so as to accelerate the swift advancement of electric vehicles industry technology, mainly reflected in: battery technology, motor and control system, automatic driving technology and other technological progress, so that the battery performance, system safety, driving experience of electric vehicles have been significantly improved, and a variety of models with different energy technologies have been born, such as: Plug-in hybrid electric vehicles (PHEVs), pure electric vehicles (BEV), fuel cell vehicles (FCEVs), and hybrid electric vehicles (HEVs).

In the future, with the further development of technology, the cost of manufacturing electric vehicles will be reduced, and high-tech technologies such as intelligent and autonomous driving technology will be more deeply integrated into various types of electric vehicles, and the driving experience of consumers will continue to be optimized and improved, so as to gain more consumers' love, and the market share of electric vehicles will be further expanded.

1.2 Problem Background

In recent years, the electric vehicle (EV) market has grown rapidly due to the continuous advancement and maturity of electric vehicle (EV) technology, and the rapid development of this market is mainly due to several factors: governments have implemented measures to foster the growth of electric vehicles, such as providing purchase incentives, tax exemptions, and financial support for the development of charging infrastructure; Continuous innovation in battery technology, motor and control systems, and autonomous driving technology has significantly improved the range and economy of electric vehicles. Increasing consumer concern for environmental protection and energy conservation is driving the demand for electric vehicles; The reduction in production and maintenance costs has made electric vehicles more affordable.

Today, the major models in the global electric vehicle market include: battery electric vehicles (BEVs), which rely on battery energy storage to achieve zero pollution emissions; Plug-in hybrid electric vehicles (PHEVs): integrate electric propulsion with internal combustion engines to achieve extended range and greater flexibility in fuel choices; Fuel cell electric vehicles (FCEVs): hydrogen fuel cells generate electricity, suitable for long-distance transportation, and the emissions are only water vapor; Hybrid electric vehicles (HEVs): Charging through energy recovery, with high fuel economy, each model has a unique development history and characteristics, and has played an important role in reducing pollution, improving energy efficiency, and promoting the development of the new energy industry.

Despite the promisingness of the EV market, there are still some challenges, such as battery recycling, the deployment and adoption of charging infrastructure, and how to reduce production and operating costs. It is anticipated that, through the combined efforts of governments, enterprises, and scientific research institutions, along with continuous technological advancements and cost reductions, the electric vehicle market will experience sustained rapid growth, and the market share of electric vehicles will keep increasing. Based on the past development trend of electric vehicles, this research project will take into account various factors such as policy, science and technology, and public opinion, predict the development trend of the dominant models in the electric vehicle market in the future, provide reference for the industry and policy makers, and contribute to the realization of global sustainable development goals.

1.3 Problem Statement

By analyzing the relevant historical data of the electric vehicle market sales since 2010, it is possible to analyze the trends and characteristics of the electric vehicle market, so as to predict the future development trend of electric vehicles. Based on this historical data, it can provide a strategic reference for automakers, policymakers, and investors.

1.4 Research Questions

The study focuses on the following research questions:

1. What is the trend of the development of the electric vehicle industry since 2010?

2. What are the trends and characteristics of the development of models in the historical development process?

3. Based on these trends, historical data, and national policies, what predictions can be made about the future?

1.5 Research Aim

The main objective of the research project is to comprehensively understand the historical development trend of the electric vehicle (EV) market by analyzing historical data from Kaggle and IEA Global EV Data Explorer, with a special focus on countries with high EV penetration rate and market influence, and to make predictions about the future development trend of various types of electric vehicles based on the changes in market dynamics, policy popularization, and iterative updates of science and technology in recent years, so as to provide reference for the industry and policymakers. Contribute to the achievement of the global Sustainable Development Goals.

1.6 Research Objectives

Objectives of this study:

a) Analyze historical trends: Analyze historical data on the electric vehicle market from 2010 to 2023 to derive historical market trends.
b) Forecast future sales: Use advanced data science techniques such as predictive modeling and machine learning to predict the development trends of major electric vehicle models by 2030.
c) Evaluate influencing factors: Evaluate the impact of factors such as technological advancement, government policies, market dynamics, and consumer preferences on the development trends of various types of vehicles.

1.7 Research Scope (Current Work)

The scope and limitations of this study are as follows:

a) This study will be limited to analyzing historical electric vehicle market data from 2010 to 2023, and using data science techniques such as forecasting models and machine learning to predict the development trends of various types of electric vehicles from 2024 to 2030.
b) This study will evaluate the impact of factors such as technological progress, government policies, market dynamics, and consumer preferences on the adoption of electric vehicles. Methods such as regression analysis and factor analysis are used to identify key drivers and barriers that promote or hinder the widespread adoption of electric vehicles.
c) This study will use limited analysis to refer to multiple influencing factors (such as technological progress, government policies, market dynamics, and consumer preferences) to maximize the accuracy of electric vehicle sales forecasts.

d) The dataset for this study will use data provided by authoritative databases such as Kaggle and IEA Global EV Data Explorer to ensure the reliability and comprehensiveness of the data.

1.8 Expected Research Contribution

a) Theoretical Contribution:

In-depth analysis of electric vehicle market trends, especially for major countries with high penetration rates. This study further enriches the theoretical research on electric vehicle market trends by analyzing historical sales data, evaluating influencing factors, and predicting future sales.

b) Methodological Contribution:

Use a variety of advanced data science analysis and prediction techniques (such as predictive modeling and machine learning) to improve the accuracy of electric vehicle sales forecasts as much as possible. Provide feasible methods for optimizing analysis and optimization of the electric vehicle market.

c) Practical Contribution:

Provide scientific basis for the industry and policymakers to formulate policies and measures to support the development of the electric vehicle market, help formulate strategic plans and optimize resource allocation. By analyzing various influencing factors, it helps to promote the sustainable development of electric vehicles.

1.9 Thesis Organization

The main sections of the thesis are structured as follows:

Chapter 2: Literature Review

Extensively review the existing literature related to electric vehicle market development, technological progress, government policies, and consumer behavior. Analyze the research background, explore the existing research gaps, and deepen the exploration of the research topics.

Chapter 3: Research Methods

Describe the methodology used in this study, including data collection, trend analysis, predictive modeling, and influencing factor assessment. Detailed information on data sources, analysis tools, and technical methods.

Chapter 4: Data Analysis and Results

Present the results of an analysis of EV sales data to identify key trends and patterns. Presents forecasts of future sales trends through predictive modeling and machine learning.

Chapter 5: Conclusions and Prospects

Summarize the main findings and contributions of the study, discuss the limitations of the study, and propose future research directions. Looking ahead to the development trends and potential impact of the electric vehicle market.

Chapter 2

Literature view

2.1 Introduction

In recent years, with the rapid growth of technology and the strong policy support and promotion by various countries, the electric vehicle (EV) market has grown rapidly. Leading automakers such as Tesla, BYD, and Nissan have created a series of high performance, high tech, long range models that are favored by consumers. In the fierce competition among major automakers, modern electric vehicle technology has developed rapidly, producing a series of mature and innovative high tech products, mainly in the areas of electric vehicle technology, battery technology, motor and control systems, and autonomous driving technology. With the development of these technologies, more and more EV models are emerging. The current market mainly includes “Battery Electric Vehicles (BEV)”, “Plugin Hybrid Electric Vehicles (PHEV)”, “Fuel Cell Electric Vehicles (FCEV)”, and “Hybrid Electric Vehicles (HEV)”. The emergence of an increasing number of models provides consumers with more affordable and diverse choices. In the future, with ongoing technological advancements and the support of global policies. the EV market is bound to continue its steady growth. This study aims to analyze and review historical data on the EV market to predict the future development of various EV models.

2.2 Historical literature on the electric vehicle market

The next discussion focuses on recent literature on growth of the electric vehicle (EV) market. The latest research suggests that a few primary models type will dominate future market trends, and will be determined by technological advances, alterations in environmental policy, and consumer preferences. These developments involve vehicle models that have much bigger batteries, more powerful smart technologies, and newer electric variants. Besides, yes, one would be accelerating from sustainable transportation solutions and infrastructure for charging that includes electric vehicles.

2.2.1 Type of Electric Vehicles

Today there are various types of electric vehicles available on the market, each with its own specifications and benefits. Here's a (relatively) simple overview of these primary types and their principles:

a) Battery Electric Vehicles (BEV)

Advantages:

i. Zero Emissions: Without tailpipe emissions, air quality could improve and pollution could decrease.
ii. Higher Efficiency: Electric motors have higher efficiency than their internal combustion counterpart.
iii. Low Maintenance: They have fewer mechanical parts than gasoline and hence easier and less to maintain.
iv. Just Awesome Ride: Silent & smooth drive.

Disadvantages:

stations, making long trips inconvenient.

Environmental Analysis:

Battery electric vehicles (BEVs) generate minimal environmental pollution, diminishing reliance on fossil fuels and cutting down greenhouse gas emissions. However, the production and disposal of batteries still require attention to environmental impacts.

b) Plugin Hybrid Electric Vehicles (PHEV)

Advantages:

i. Flexibility: Combines electric and internal combustion engine modes, highly adaptable.
ii. Reduced Emissions: Can run entirely on electric power for short trips, reducing tailpipe emissions.
iii. Long Range: The internal combustion engine provides additional range, avoiding battery depletion.

Disadvantages:

i. Complex Design: Requires two sets of power systems, making maintenance more challenging.
ii. High Cost: Generally more expensive than traditional internal combustion vehicles and BEVs.
iii. Increased Weight: Having both an internal combustion engine and a battery increases the vehicle's overall weight.

Environmental Analysis:

Effectively reduces tailpipe emissions on short trips, but relies on the internal combustion engine for long journeys. It is an important step towards zero emission vehicles.

c) Hybrid Electric Vehicles (HEV)

Advantages:

i. Fuel Efficiency and Reduced Emissions: Combines electric motors with internal combustion engines, reducing fuel consumption.
ii. Long Range: Internal combustion engine provides additional driving distance.
iii. Ready to Use: Does not require external charging, suitable for users without charging facilities.

Disadvantages:

i. High Cost: More expensive than traditional internal combustion vehicles.
ii. Complex Design: Requires two sets of power systems.
iii. Battery Life: Battery performance may degrade over time.

Environmental Analysis: HEVs perform well in reducing fuel consumption and emissions but still rely on fossil fuels, placing their environmental impact between traditional vehicles and BEVs.

d) Fuel Cell Electric Vehicles (FCEV)

Advantages:

i. Zero Emissions: Only emits water vapor, environmentally friendly.
ii. Long Range: Typically offers a range comparable to internal combustion vehicles.
iii. Quick Refueling: Hydrogen refueling time is short, similar to refueling with gasoline.

Disadvantages:

i. Limited Infrastructure: Insufficient hydrogen stations.
ii. High Cost: Hydrogen fuel cells and storage systems are expensive.
iii. Safety: Hydrogen storage and transportation require high safety standards.

Environmental Analysis: FCEVs do not produce pollutants during operation, only emitting water vapor. However, hydrogen production requires significant energy, especially when sourced from fossil fuels, which produces carbon dioxide. Developing renewable hydrogen production methods is crucial to realizing their environmental potential.

2.3 Important Factors Affecting the Development

In recent years, numerous studies have delved into the development trends of the electric vehicle (EV) market and the factors influencing these trends. Most of these studies focus on identifying the primary factors that shape the development trajectories of various types of EVs. However, there remains a gap in comprehensive research on the overall development trend of the EV market. This study aims to fill that gap by analyzing and synthesizing existing research on EV market trends. It will consider a wide range of factors affecting the development of different EV types and provide a detailed analysis and forecast of the overall trends in the current EV market.

Research on the market development trend forecast of various types of electric vehicles in recent years is as follows:

Authors	Related Paper	Summary	Market Analysis Predictions
Hertzke et al.	Dynamics in the global electric-vehicle market	Explores China's dominance in EV production due to strong policies and advanced manufacturing.	The global market is highly competitive, with China maintaining its dominant position due to scale and policy support.
Muratori et al.	The rise of electric vehicles: 2020 status and future expectations	Provides an overview of BEV growth as the primary zero-emission vehicle and global electrification trends.	BEVs will see steady market growth globally, driven by environmental policies and consumer preferences.
Usman et al.	Recent trends and future prospects in electric vehicle technologies	Reviews the role of HEVs as transitional technologies and PHEVs as flexible urban solutions.	HEVs will maintain steady demand, while PHEVs appeal to urban markets seeking flexible, efficient options.
Joao P. Trovao	Electric Vehicle Efficient Power and Propulsion Systems	It highlights the shift from traditional internal combustion engines to more efficient and eco-friendly propulsion systems like hybrid and all-electric vehicles.	Electrification will grow by 15% by 2030.
A.Olabi et al.	Battery Electric Vehicles: Progress, Power Electronic Converters, Strength (S), Weakness (W), Opportunity (O), and Threats (T)	The document explores the automotive industry's shift from internal combustion engines to electric vehicles (EVs).	Anticipated future trends in the electric vehicle (EV) market will likely prioritize “battery electric vehicles (BEVs)” and “plug-in hybrid electric vehicles (PHEVs)”.
Abdullah Dik	Electric Vehicles: V2G for Rapid, Safe,	This article explores the role of electric vehicles (EVs) in	Future trends in electric vehicles (EV) will likely center around “battery

	and Green EV Penetration	reducing carbon emissions and aiding the integration of renewable energy sources into the power grid.	electric vehicles (BEVs)" and "plug-in hybrid electric vehicles (PHEVs)".
Xiangyang Li	Industrial ripples: Automotive electrification sends through carbon emissions	The research paper discusses the importance of vehicle electrification in addressing climate challenges related to carbon emissions from the transport sector	The study predicts that electric vehicles will dominate the automotive market in the future, with hybrid vehicles acting as an intermediary step.
Conway et al.	A review of current and future powertrain technologies and trends	Examines the rise of FCEVs and powertrain efficiency improvements for long-distance applications.	FCEVs have significant potential in long-distance transport, particularly in markets like Japan.

Table 2.1 Future Trends and Market Analysis of Electrical Vehicles

The Table 2.1 above mentioned historical studies have roughly analyzed and predicted the market trends of some types of electric vehicles (EVs) in some countries and regions, which are specifically reflected in:

Regarding the global electric vehicle market: In the studies such as "Dynamics in the global electric-vehicle market" and "Electric Vehicle Efficient Power and Propulsion Systems", the authors mainly explored the global electrification trend and took China, which occupies a leading position in the electric vehicle market, as the research object, and concluded that the global market is rapidly transitioning to electric travel. Battery electric vehicles (BEVs) are becoming the flagship of zero-emission efforts. Their steady growth is driven by environmental regulations and increasing consumer demand (Muratori et al.), and with strong policy incentives and advanced manufacturing capabilities, their growth will be strongest in regions with sound infrastructure and government incentives, such as Europe, China and the United States (Muratori et al.), among which China leads in production, sales and technology, making electric vehicles more affordable (Hertzke et al.), which is a typical representative of the rapid transition of global electrification trends to electric travel in the future.

Battery and charging innovation: Lithium-ion batteries and solid-state batteries have always been a major issue hindering the development of electric vehicles. Today’s innovations in battery and charging technologies have increased battery energy density, reduced manufacturing costs, and significantly reduced charging time (A. Olabi et al.). At the same time, emerging wireless charging and vehicle-to-grid (V2G) integration are reshaping electric vehicle infrastructure, making it more practical and attractive (Hemavathi et al.).

Efficient powertrain: Advances in powertrains are revolutionizing the automotive sector, with many electric drive engines outperforming traditional internal combustion engines, and fuel cell electric vehicles (FCEVs) becoming a viable option for long-distance transportation (Conway et al.).

Sustainability and policy impact: Governments around the world are adopting stricter carbon emission regulations, which will directly promote the growth of the electric vehicle market, and hybrid electric vehicles (HEVs) serve as a transition technology in regions where electric vehicle adoption is slower (Xiangyang Li).

2.4 Classification

Classification is a data mining (machine learning) technique used to predict group membership for data instances. There are several classification techniques that can be used for classification purpose. (Aized and Arshad, 2017).

Machine Learning (ML) is a comprehensive interdisciplinary domain that draws from computer science, statistics, cognitive science, engineering, optimization theory, and various other fields of mathematics and science [1]. Among its numerous applications, data mining stands out as particularly significant [2]. ML techniques are broadly categorized into supervised and unsupervised learning.

In unsupervised machine learning, models make inferences from datasets that consist of input data without labeled responses [3]. Essentially, in unsupervised learning, there is no predefined output for the model to predict.

Supervised machine learning aims to uncover the relationship between input features (independent variables) and a target variable (dependent variable) [4]. These techniques are further divided into two main categories: classification and regression. In regression, the target variable is continuous, while in classification, the target variable takes on discrete class labels [5].

Classification is a machine learning approach used in data mining to predict the group membership of data instances [6]. Despite the variety of available machine learning techniques, classification remains the most widely used [7]. It is particularly popular in planning for future trends and knowledge discovery.

Research in machine learning and data mining has extensively studied classification problems [8]. Figure 1 illustrates a general model for supervised learning through classification techniques.

Despite its popularity, classification faces challenges such as handling missing data. Missing values in datasets can create problems during both training and classification phases. Reasons for missing data include non-entry due to misunderstandings, data deemed irrelevant at the time of entry, removal of data due to inconsistencies with other recorded data, and equipment malfunction [9].

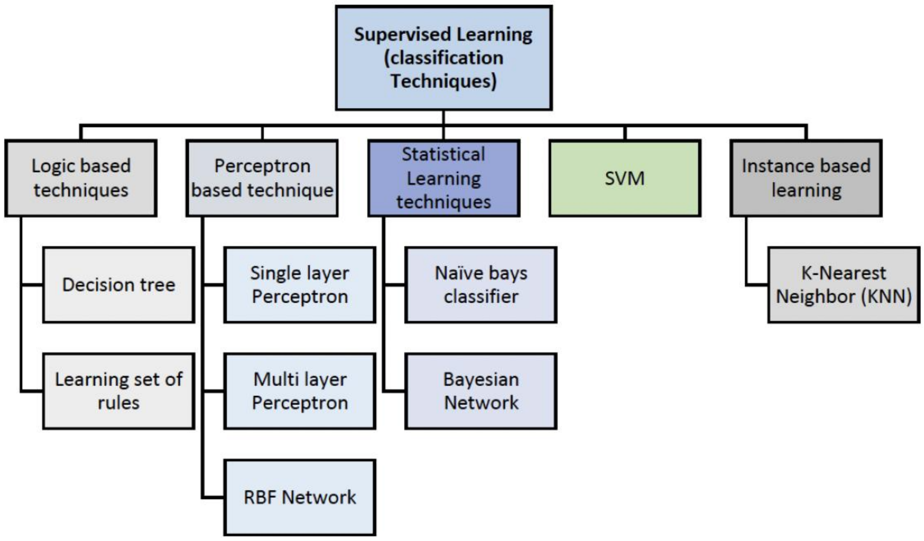

Figure 2.1 Supervised learning classification technique(Shoofi and Awan, 2017)

2.5 Market Trend Analysis and Forecast

Below is market trend analysis and forecasts for various electric vehicle types:

a. Battery Electric Vehicles (BEV)

Key factors affecting growth is Improvements in battery technology, like solid-state batteries, favorable government policies for the electric vehicle sector, and the development of charging infrastructure, including fast and wireless charging, are key factors propelling the electric vehicle market forward. Forecasts for the BEV market in past historical studies are BEVs will dominate the global electric vehicle market. Growth will be strongest in regions with well-established infrastructure and government incentives, such as Europe, China, and the United States (Muratori et al.).

b. Plug-in Hybrid Electric Vehicles (PHEV)

Key factors affecting growth are Urban consumers' demand for electric vehicle range and fuel support, gradual transition from internal combustion engines to fully electric solutions. Market forecast isPHEVs will maintain steady growth, especially in urban markets and regions where charging infrastructure is still developing, and their appeal lies in balancing efficiency and practicality (Usman et al.).

c. **Hybrid Electric Vehicles (HEV)**

Key factors affecting growth are Transition technology for regions not fully ready for electric vehicles, continued use in areas with low access to charging infrastructure. Market forecast is HEVs will gradually decline in favor of BEVs and PHEVs. However, their cost advantages make them remain relevant among cost-sensitive consumers (Xiangyang Li).

d. **Fuel Cell Electric Vehicles (FCEV)**

Key factors affecting growth are Use in long-distance and heavy-duty transportation, government investment in hydrogen infrastructure. Market Forecast: is FCEVs will find a place in specific markets, such as logistics and long-distance transportation, with growth potential in regions that are actively investing in hydrogen technology (Conway et al.).

e. **Charging Infrastructure and V2G Integration**

Key Factors Influencing Growth are Infrastructure development of fast charging networks and standardized protocols, integration with renewable energy systems for sustainability. Market Forecast is Growth in charging infrastructure will directly affect the speed of EV adoption, and investments in V2G technology will accelerate the integration of EVs with renewable energy grids (Hemavathi et al.).

Each type of electric vehicle will have a place in market development, especially BEVs are expected to lead the zero-emission revolution, while PHEVs will fill the gap during the transition period and FCEVs will be instrumental in meeting the specialized needs of long-distance transportation. The construction of battery technology and charging infrastructure will ensure sustainable growth for all electric vehicle types.

2.5.1 Market trend analysis and forecasts for various electric vehicle types:

a. Battery Electric Vehicles (BEV)

Key factors affecting growth:

Improvements in battery technology, like solid-state batteries, favorable government policies for the electric vehicle sector, and the development of charging infrastructure, including fast and wireless charging, are key factors propelling the electric vehicle market forward.

Forecasts for the BEV market in past historical studies:

BEVs will dominate the global electric vehicle market. Growth will be strongest in regions with well-established infrastructure and government incentives, such as Europe, China, and the United States (Muratori et al.).

b. Plug-in Hybrid Electric Vehicles (PHEV)

Key factors affecting growth:

Urban consumers' demand for electric vehicle range and fuel support, gradual transition from internal combustion engines to fully electric solutions.

Market forecast:

PHEVs will maintain steady growth, especially in urban markets and regions where charging infrastructure is still developing, and their appeal lies in balancing efficiency and practicality (Usman et al.).

c. Hybrid Electric Vehicles (HEV)

Key factors affecting growth:

Transition technology for regions not fully ready for electric vehicles, continued use in areas with low access to charging infrastructure.

Market forecast:

HEVs will gradually decline in favor of BEVs and PHEVs. However, their cost advantages make them remain relevant among cost-sensitive consumers (Xiangyang Li).

d. Fuel Cell Electric Vehicles (FCEV)

Key factors affecting growth:

Use in long-distance and heavy-duty transportation, government investment in hydrogen infrastructure.

Market Forecast:

FCEVs will find a place in specific markets, such as logistics and long-distance transportation, with growth potential in regions that are actively investing in hydrogen technology (Conway et al.).

e. Charging Infrastructure and V2G Integration

Key Factors Influencing Growth:

Infrastructure development of fast charging networks and standardized protocols, integration with renewable energy systems for sustainability.

Market Forecast:

Growth in charging infrastructure will directly affect the speed of EV adoption, and investments in V2G technology will accelerate the integration of EVs with renewable energy grids (Hemavathi et al.).

Conclusion:

Each type of electric vehicle will have a place in market development, especially BEVs are expected to lead the zero-emission revolution, while PHEVs will fill the gap during the transition period and FCEVs will be instrumental in meeting the specialized needs of long-distance transportation. The construction of battery technology and charging infrastructure will ensure sustainable growth for all electric vehicle types.

2.5.2 summary of factors affecting the future development trend of electric vehicles (EVs)

Here is a comprehensive overview of the factors influencing the future development trends of electric vehicles (EVs), derived from various research studies. including government policies, technological advancements, infrastructure development, environmental considerations, and economic factors.

Technology and infrastructure development

a. Battery technology and charging infrastructure:

i. Solid-state batteries: As the electric vehicle industry progresses, solid-state batteries are anticipated to become increasingly prevalent due to their superior energy density, expedited charging capabilities, and enhanced safety features compared to traditional lithium-ion batteries. These advancements are expected to significantly extend the driving range of electric vehicles and potentially optimize manufacturing costs. (Source: "Challenges, Solutions and Future Trends in Electric Vehicle Technology: A Review").

ii. Lithium-sulfur batteries: Studies on lithium-sulfur batteries indicate that these batteries offer superior energy storage capacity and may surpass current lithium-ion technology, thereby further enhancing the range and efficiency of electric vehicles, which may adversely affect the future development trends of electric vehicles. Quite an impact (source: "Efficient Power and Propulsion System for Electric Vehicles").

iii. Battery Management System (BMS): BMS innovation will further optimize battery life, performance and safety, and improve overall battery performance by more effectively monitoring and managing battery usage (Source: "Electric Vehicles: Battery Management Systems, Charging Stations, Comprehensive review of traction motors").

iv. Fast charging stations: Expanding fast charging networks is essential for shortening charging times and improving the convenience of long-distance travel. Advancements in ultra-fast charging technology are anticipated to further shorten the charging time of electric vehicles in the future. (Source: "Global Electric Vehicle Market Dynamics").

v. Wireless charging: Wireless charging solutions are currently under development and, if successfully developed, will eliminate the need for physical cables and connectors, thereby providing a more convenient and efficient charging option for electric vehicles. This technology can be integrated into parking lots and public areas to support mobile charging (Source: “Special Issue on Advanced Charging Technology for Next Generation Electric Vehicles”).

b. Vehicle-to-network (V2G) technology:

The purpose of V2G integration technology is to enable electric vehicles to supply power to the grid, enhance grid stability, and sell excess power to the grid through electric vehicles during peak power demand periods, allowing electric vehicle owners to monetize battery storage (Source: "Electric Vehicles: V2G enables fast, safe and green electric vehicle penetration”).

c. Smart grid:
   i. Energy management: Smart grids will be instrumental in managing the flow of energy between electric vehicles and the grid. Optimizing the use of renewable energy and ensuring efficient power distribution are top priorities for smart grids (Source: "Electric Vehicles: Technology" , integration, adoption and optimization).
   ii. Renewable energy integration: Combining electric vehicles with renewable energy sources like solar and wind will lead to a reduction in overall carbon emissions and help achieve a low-carbon economy (Source: "Forecasting Consumption and Emissions of Passenger Cars and Light Trucks to 2050").

Market dynamics and policy support

a. Government incentives and policy support:
   i. Government subsidies and tax incentives play a critical role in encouraging the adoption of electric vehicles. To reduce the purchase and operating costs of electric vehicles, governments worldwide are implementing various incentives and policies. These include car purchase subsidies, tax exemptions, and financial support for charging infrastructure development. Improve its market attractiveness (Source: "Global Electric Vehicle Market Dynamics").
   ii. Tighter vehicle emissions regulations and a ban on internal combustion engine vehicles are also speeding up the shift to electric vehicles. For instance, the European Union aims to completely prohibit the discontinuation of new internal combustion engine vehicle sales by 2035. (Source: "Assessment of the Impact of Electric Vehicle Mobility on Power Generation").

b. Economic factors:

i. Although the initial outlay for purchasing an electric vehicle is higher, their reduced operational and reduced maintenance expenses contribute to a lower overall Total Cost of Ownership (TCO) over their lifecycle. In the long run, electric vehicles prove to be more economical than internal combustion engine vehicles. costs (source: "Finding the most suitable vehicle type by integrating economic and environmental aspects through the AHP").
ii. Investment in charging infrastructure, battery production equipment and the integration of renewable energy sources are key to supporting the growing electric vehicle market (Source: “Overview of the future of electric and hydrogen cars and trucks”).

environmental and social impacts

a. Reduce carbon emissions:
   i. The popularization of electric vehicles will significantly reduce greenhouse gas emissions and help achieve the goal of protecting the global climate. For example, the actual benefits of China's popularization of electric vehicles are expected to reduce 54 billion tons of carbon dioxide equivalent by 2060 (Source: "The Feasibility of Carbon Neutrality in the Global and Chinese Transportation Sectors by 2060").
   ii. Research emphasizes considering the entire life cycle carbon emissions of electric vehicles, including manufacturing, use and disposal, to fully understand their impact on the environment (Source: "Assessing Europe's electric vehicle transition: Emissions during the manufacturing and use of electric vehicles").

b. Social and economic impact:
   i. The growing popularity of electric vehicles is anticipated to generate additional employment opportunities in fields like EV manufacturing, battery production, and charging infrastructure development. (Source: "Development of New Energy Vehicles under China's Carbon Peak and Carbon Neutral Strategy").
   ii. Increasing public awareness and acceptance of electric vehicles is critical to their widespread adoption. Educational campaigns and incentive programs can help address consumer concerns and expand the market for electric vehicles (Source: “Towards an Energy Future of Universal Electric Vehicles: Barriers and Opportunities”).

Conclusion:

The outlook for electric vehicles is determined by technological progress, policy support, market demand and environmental issues. The continuous progress of battery technology and charging infrastructure will play a key role in promoting the popularization and sustainable development of electric vehicles. Under the influence of various factors, it is expected that in the future, BEV will lead the zero-emission revolution, while PHEV and FCEV will provide transitional and professional solutions respectively. Continuous advancements in these fields will guarantee the sustainable growth and evolution of the electric vehicle market.

2.6 Methodologies for Ontology Development

A variety of methods have been explored in the literature. Through the analysis and comparison of various research methods, for this study: predicting the trend of the electric vehicle market, time series models such as ARIMA and SARIMA will be used to predict the electric vehicle market.

2.6.1 Sales Forecasting Techniques

ARIMA Model

SARIMA (Seasonal ARIMA) is an extension of ARIMA that accounts for seasonal differences and autoregressive/moving average components. This makes it particularly efficient in capturing seasonal swings in EV sales, such as increased demand. Purchases made during fiscal year-end reductions or new model releases. SARIMA accurately estimates monthly or quarterly sales. Validation with historical information ensures accuracy in detecting supply chain limitations and unanticipated demand surges.

Benefits in terms of the EV market:

By applying ARIMA:

a. Historical sales data are decomposed into trend, seasonal, and residual components.
b. The forecast accuracy can be verified using historical sales records.
c. The model is dynamically adjusted to respond to external influences such as changes in market conditions or policy changes.

ARIMA predicts future sales by modeling the relationship between past observations in a time series. It is denoted as ARIMA(p, d, q), where:

- p: Number of autoregressive terms.
- d: Number of non-seasonal differences.

q: Number of moving average terms.

The general equation of ARIMA is:

$$y_t = c + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots \emptyset_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

Where:

- y_t: Predicted value at time ttt.
- $\emptyset$: Coefficients of the autoregressive terms.

- θ_t: Coefficients of the moving average terms.
- ϵ_t: Error term at time ttt. : Error term at time

Steps in ARIMA Application:

a. Stationarity Check: The series is tested for stationarity using methods like the Augmented Dickey-Fuller (ADF) test. If non-stationary, differencing is applied.
b. Model Fitting: Based on parameter tuning (using criteria like AIC/BIC), the model is fitted to the data.
c. Forecasting: Predictions are generated for the future time points.

SARIMA Model

SARIMA: Seasonal ARIMA, besides everything that ARIMA does, includes seasonal differences and seasonal autoregressive/moving average components. This makes it really effective in capturing seasonal fluctuations of EV sales, such as increased purchases during end-of-fiscal-year discounts or new model launches. It also can be used to predict monthly or quarterly sales, thus providing effective forecasts. Validation with historical datasets ensures accuracy for anomalies such as supply chain constraints or unexpected demand surges.

Benefits in the context of the EV market: These models are particularly beneficial to stakeholders, including manufacturers and policymakers, to:

a. Forecast production needs.
b. Identify peak demand periods.
c. Design incentives for low demand seasons.
d. Best-selling vehicle prediction

To identify which types of EVs are poised to lead the market, harnessing machine learning-based classification models is essential. These models classify vehicle models based on their specifications and consumer preferences.

a. Modeling approach:
    i. Input data: Vehicle specifications (e.g., range, battery capacity, cost) and consumer preferences (e.g., affordability, charging speed, environmental impact).

ii. Model development: Using techniques such as logistic regression, random forest, or neural networks, the likelihood of a vehicle becoming a bestseller can be classified.
iii. Validation: Use historical data on sales performance and consumer feedback to improve model accuracy.

b. Applications:
   i. For example, a random forest model can analyze the weight of factors such as price vs. range in consumer purchase decisions.
   ii. Sensitivity analysis can reveal how specification changes (such as battery improvement.
c. Advantages in the context of the EV market: These models enable manufacturers to:
   i. Align product development with consumer priorities.
   ii. Identify market segments with unmet needs.
   iii. Optimize marketing strategies based on predicted high-demand features.

SARIMA introduces seasonal components to ARIMA. The model is denoted as SARIMA(p, d, q)(P, D, Q, s), where:

- P,D,Q: Seasonal autoregressive, differencing, and moving average terms, respectively.
- s: Seasonal period (e.g., 12 for monthly data).

The equation extends ARIMA with seasonal terms:

$$y_t = c + \sum_{i=1}^{p} \emptyset_i y_{t-1} + \sum_{i=1}^{q} \theta_i \epsilon_{t-1} + \sum_{j=1}^{p} \omega_j y_{t-js} + \sum_{j=1}^{p} \Phi_j \epsilon_{t-js} + \epsilon_t$$

Where:

- ω_j, Φ_j: Seasonal autoregressive and moving average coefficients.
- s: Seasonal period.

Seasonality Capture: SARIMA is particularly adept at capturing regular demand cycles, like annual or quarterly trends, making it an excellent choice for modeling EV sales that are influenced by policy changes or economic fluctuations.

2.6.2 Best-Selling Vehicle Type Prediction

Machine learning-based classification algorithms are critical for predicting which EV kinds will dominate the market. These models categorize vehicles based on characteristics and consumer preferences.

a. Modeling Approach:
   i. Input Data: Vehicle specifications (e.g., range, battery size, cost) and consumer preferences (e.g., affordability, charging speed, environmental impact).
   ii. Model Development: Techniques such as logistic regression, random forests, or neural networks can classify the likelihood of a vehicle being a best-seller.
   iii. Validation: Historical data on sales performance and consumer feedback is used to refine model accuracy.
b. Application:
   i. A random forest model helps determine which characteristics, such as price and range, have a greater impact on consumer purchasing decisions.
   ii. Sensitivity analysis can show how changes in specifications, such as battery upgrades, may impact consumer choices.

Supervised machine learning models use methods like logistic regression, random forests, and support vector machines (SVMs) to classify vehicles. The focus is on the mathematical underpinnings.

Logistic Regression

Logistic regression models the probability of a vehicle being a best-seller (P(y=1) based on features like price, range, and battery size:

$$P(y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}}$$

Where:

- ββ_0: Intercept.
- β_1, β_2, ⋯, β_n :Coefficients for features x_1, $x_2, \dots, x_n$.

- P(y=1): Probability of the vehicle being a best-seller.

The model estimates coefficients to maximize the likelihood function, fitting the data.

Random Forest

Random forests use an ensemble of decision trees to classify data. For a feature set X, each tree predicts a class label $\hat{y}$. The final prediction is based on majority voting:

$$\hat{y} = Mode(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m)$$

Where:

- $\hat{y}_i$: Prediction from the iii-th tree.
- m: Total number of trees in the forest.

Advantages for EV Prediction:

- Handles non-linear relationships in data.
- Avoids overfitting by aggregating results.

Support Vector Machines (SVM)

SVM finds a hyperplane that separates classes (e.g., best-seller vs. non-best-seller). For input XXX, the decision boundary is:

$$\boldsymbol{f(X) = \omega^T X + b}$$

Where:

- ω: Weight vector defining the hyperplane orientation.
- b: Bias term.

The goal is to maximize the margin MMM between classes:

$$\boldsymbol{M = \frac{2}{\parallel \omega \parallel}}$$

Kernel Trick: SVM applies kernels like radial basis functions (RBF) to handle non-linear separability.

Conclusion

By integrating these mathematical models, the paper can robustly forecast EV sales trends and predict market dynamics for best-selling vehicles. The ARIMA/SARIMA models provide reliable sales trend forecasts, while classification techniques like logistic regression, random forests, and SVM enable accurate predictions of consumer preferences for vehicle types. These methods collectively strengthen the analysis, aligning with the goals of the study "Electric Vehicle Market Forecast."

2.6.3 Software and Libraries

Python, PowerBI, Scikit-learn, Statsmodels, and Matplotlib/Seaborn for visualization.

2.7 Research Gap

Authors	Publication Date	Result	Research Gap
Sanjib Kumar Shil	10 Nov 2024	Accurate predictions of EV adoption trends in the USA.	Limited to the USA; does not address global or regional variations in EV adoption trends.
Shumo Cui	20 Oct 2024	EVs expected to constitute over 25% of global car stock by 2035.	Lacks specificity about regional market disparities and the factors influencing them.
Irvylle Raimunda Mourão Cavalcante	19 Sep 2024	Dataset covering scenarios up to 2050.	Absence of real-time adaptability for emerging trends and unforeseen disruptions in scenarios.
Fei Teng	31 Oct 2024	Enhanced accuracy in EV sales forecasting.	Lack of detailed data sources or validation methodologies for forecasting accuracy.
Ru Qi Yu	13 Nov 2024	Robust framework for forecasting EV sales in China.	Focused on China only; no comparative analysis with other key EV markets globally.
Ning Mao	14 Oct 2024	Impact analysis of EV penetration on ownership and energy demand.	Limited integration of social or behavioral factors impacting energy demand and vehicle ownership.
Yiying Liu	12 Sep 2024	Predicted development trends in the new energy EV industry.	Insufficient examination of the impact of evolving policy and regulatory frameworks.

Table 2.2 Summary of the specified research papers

The above studies have conducted in-depth discussions on the trend forecast, market analysis and related influencing factors of the electric vehicle (EV) market development, but there are still some gaps that need further research. These gaps are mainly concentrated in the following aspects:

a. Insufficient cross-regional comparative analysis

Although some studies focus on the world (such as Shumo Cui's study), there is a lack of in-depth analysis of market differences and their driving factors in different regions. For example, Ru Qi Yu's study focuses on China and does not compare data with other major markets (such as Europe and North America), and the research scope is not comprehensive enough.

b. Dynamic scenario modeling and real-time adaptability

Irvylle Raimunda Mourão Cavalcante provides trend forecasts covering 2050, but lacks dynamic adaptability in the face of rapidly changing policies, technologies and social trends. In addition, this static analysis cannot reflect the rise of emerging markets or the impact of unexpected events in a timely manner.

c. The interactive impact of social behavior and policy

Ning Mao and Yiying Liu discussed energy demand and industry development trends respectively, but the synergy analysis of social behavior factors (such as consumer preferences and acceptance) and policy/regulatory changes is insufficient. In particular, the impact of the evolution of energy policies and car purchase subsidies on consumer behavior has not been fully considered.

d. Forecasting methods and data verification

Fei Teng's research has improved the accuracy of forecasts, but the verification methods and data source transparency of the forecasting model are still insufficient. This deficiency may limit the generalizability of the model in other regions or situations.

e. Potential impact of future technological development

Current research rarely involves the potential impact of breakthroughs in electric vehicle technology (such as battery technology and new energy drive) on market trends and user acceptance.

Based on the above analysis, the following research directions need to be further explored to fill the key gaps in the current literature:

First, build a model that can comprehensively compare the differences between different regional markets (such as Asia, Europe, and North America), analyze the main factors that cause these differences, and conduct a comprehensive comparative analysis of global and regional markets. For the existing forecasting model, verify its applicability in different regional markets, and improve the transparency of its data sources and methods to achieve cross-regional verification and expansion of the forecasting model.

Secondly, develop an adaptive model that combines real-time data and scenario prediction, which can quickly respond to policy adjustments, new technological breakthroughs, and sudden market changes, achieve dynamic prediction and scenario adaptation, explore how policies (such as subsidies and tax incentives) and social behavior changes (such as environmental awareness and consumption habits) jointly affect the long-term development trend of the electric vehicle market, and build a synergistic model between social behavior and policy. Study how technological advances such as emerging battery technology and intelligent connected vehicles affect policy making and the future pattern of the electric vehicle market, and analyze the long-term impact of technological progress and policy interaction.

These research directions provide more comprehensive theoretical support and practical guidance for the global development and policy making in the field of electric vehicles.

2.8 Conclusion

Reasonableness of ARIMA/SARIMA and Classification Models:

1. **Suitability of Time Series Analysis (ARIMA/SARIMA):**

- Characteristics of EV Data: EV adoption trends and sales data often exhibit seasonality, trends, and irregular fluctuations, making time series models such as ARIMA (Autoregressive Integrated Moving Average) and SARIMA (Seasonal ARIMA) suitable tools for analysis.
- Advantages of ARIMA/SARIMA: These models effectively handle time dependence and seasonal patterns, providing a robust framework for short- to medium-term forecasting, especially in markets with mature patterns.
- Integration of policies and disruptions: By introducing exogenous variables (e.g., economic indicators, policy changes), ARIMA/SARIMA models can incorporate external shocks and address gaps in static, long-term scenario analysis.

2. **Role of Classification Models:**

- Market Segmentation: By dividing the EV market into segments according to sociodemographic, behavioral, or regional characteristics, classification techniques (such as logistic regression, random forests, or support vector machines) can uncover variability that conventional time series models are unable to capture.
- Policy and Consumer Interaction investigation: These models supplement the numerical projections of the ARIMA/SARIMA models by enabling investigation of the effects of various policy tools (such as taxes and subsidies) and consumer preferences on market adoption.

- Scalability and Adaptability: Classification models are well-suited for analyzing new trends in the quickly shifting EV industry because they can adjust to changing datasets.

3. Combination of ARIMA/SARIMA with Classification Models:

- Hybrid Approach: The integration of time series models into a classification framework improves the forecasts by jointly considering the historical data trend, consumer behavior, and policy dynamics.
- Improvement in Accuracy with Interpretability: Hybrid models leverage the strengths of both to provide accurate, interpretable, and adaptable forecasts that address the complexity issues noted in the literature.

This therefore will enable a comprehensively founded approach toward improving the research gaps identified earlier through using ARIMA/ SARIMA for time series analysis, classification models on behavioral and policy-driven segments.

Reference

1. Hertzke et al. (2024). Dynamics in the global electric-vehicle market.
2. Muratori et al. (2024). The rise of electric vehicles: 2020 status and future expectations.
3. Usman et al. (2024). Recent trends and future prospects in electric vehicle technologies.
4. Joao P. Trovao (2024). Electric Vehicle Efficient Power and Propulsion Systems.
5. A. Olabi et al. (2024). Battery Electric Vehicles: Progress, Power Electronic Converters, Strength (S), Weakness (W), Opportunity (O), and Threats (T).
6. Abdullah Dik (2024). Electric Vehicles: V2G for Rapid, Safe, and Green EV Penetration.
7. Xiangyang Li (2024). Industrial ripples: Automotive electrification sends through carbon emissions.
8. Conway et al. (2024). A review of current and future powertrain technologies and trends.
9. Sanjib Kumar Shil (10 Nov 2024). Environmental data, economic indicators, and policy-related data. Accurate predictions of EV adoption trends in the USA.
10. Shumo Cui (20 Oct 2024). Data from the International Energy Agency (IEA). EVs expected to constitute over 25% of global car stock by 2035.
11. Irvylle Raimunda Mourão Cavalcante (19 Sep 2024). Historical data on electric road mobility. Dataset covering scenarios up to 2050.
12. Fei Teng (31 Oct 2024). Enhanced accuracy in EV sales forecasting.
13. Ru Qi Yu (13 Nov 2024). Sales data for pure electric vehicles in China. Robust framework for forecasting EV sales in China.

14. Ning Mao (14 Oct 2024). Motor vehicle ownership and energy demand data. Impact analysis of EV penetration on ownership and energy demand.
15. Yiying Liu (12 Sep 2024). Market dynamics and technological advancements data. Predicted development trends in the new energy EV industry.Milad Haghani, Frances Sprei, Khashayar Kazemzadeh, Zahra Shahhoseini, Jamshid Aghaei (01 Oct 2023). Trends in electric vehicles research.
16. Irfan Ullah et al. (28 Feb 2023). Employing Bibliometric Analysis to Identify the Current State of the Art and Future Prospects of Electric Vehicles.
17. Ramy Kotb et al. (09 Feb 2023). Power Electronics Converters for Electric Vehicle Auxiliaries: State of the Art and Future Trends.
18. A. Olabi et al. (01 Sep 2022). Battery Electric Vehicles: Progress, Power Electronic Converters, Strength (S), Weakness (W), Opportunity (O), and Threats (T).
19. Yuefeng Liu et al. (01 Aug 2022). A review of lithium-ion battery state of charge estimation based on deep learning: Directions for improvement and future trends.
20. Joao P. Trovao et al. (24 May 2022). Electric Vehicle Efficient Power and Propulsion Systems.
21. Rui Martim Salgado et al. (26 May 2021). The Latest Trends in Electric Vehicles Batteries.
22. Amela Ajanovic et al. (18 Feb 2021). On the Historical Development and Future Prospects of Various Types of Electric Mobility.
23. S. Hemavathi et al. (05 Dec 2023). Present and Future Generation of Secondary Batteries: A Review.
24. Luis Nando Ochoa et al. (01 Nov 2023). New demand, new challenges, and new opportunities: Making EVs and the grid work together [Guest Editorial].
25. Monika Stoma et al. (23 Oct 2023). Future Challenges of the Electric Vehicle Market Perceived by Individual Drivers from Eastern Poland.
26. Lázaro V. Cremades (28 Jun 2023). Evaluation of the Power Generation Impact for the Mobility of Battery Electric Vehicles.
27. Wenyue W. Zhang et al. (11 May 2023). The alternative path for fossil oil: Electric vehicles or hydrogen fuel cell vehicles?
28. Aqib Zahoor et al. (01 May 2023). The carbon neutrality feasibility of worldwide and in China's transportation sector by E-car and renewable energy sources before 2060.
29. Manfred Dollinger et al. (21 Apr 2023). Physics-Based Prediction for the Consumption and Emissions of Passenger Vehicles and Light Trucks up to 2050.
30. Aiman Albatayneh et al. (03 Apr 2023). Future of Electric and Hydrogen Cars and Trucks: An Overview.

31. Cristian Domarchi et al. (02 Apr 2023). Electric vehicle forecasts: a review of models and methods including diffusion and substitution effects.
32. Steffen Link et al. (14 May 2024). Rapidly declining costs of truck batteries and fuel cells enable large-scale road freight electrification.
33. Bin Shui et al. (01 May 2024). Light-duty passenger vehicle electrification in China from 2021 to 2050 and associated greenhouse gas emissions: A dynamic fleet perspective.
34. Xiangyang Li et al. (01 Apr 2024). Industrial ripples: Automotive electrification sends through carbon emissions.
35. Andrea Di Martino et al. (01 Mar 2024). Increasing transportation sustainability from scratch: Small city context.
36. Haoyu Wang et al. (01 Feb 2024). Guest Editorial: Special Issue on Advanced Charging Technologies for Next-Generation Electric Vehicles.
37. Ugur Feslı et al. (01 Jan 2024). Electric Vehicles: A Comprehensive Review of Technologies, Integration, Adoption, and Optimization.
38. Ibham Veza et al. (01 Nov 2023). Electric vehicle (EV) and driving towards sustainability: Comparison between EV, HEV, PHEV, and ICE vehicles to achieve net zero emissions by 2050 from EV.
39. Seunghyun Cheon et al. (01 Nov 2023). Finding the most suitable vehicle type for projected years using analytic hierarchy process integrated with economic and environmental aspects.
40. Mohammad Mohammadi et al. (03 Sep 2023). Towards an Energy Future with Ubiquitous Electric Vehicles: Barriers and Opportunities.
41. G. C. Horvath et al. (17 Jul 2023). A Comprehensive Review of the Distinctive Tendencies of the Diffusion of E-Mobility in Central Europe.
42. Xia Li et al. (08 May 2023). Development of New-Energy Vehicles under the Carbon Peaking and Carbon Neutrality Strategy in China.
43. Wenbin Zhou et al. (01 Mar 2023). Cost, range anxiety and future electricity supply: A review of how today's technology trends may influence the future uptake of BEVs.
44. Chen Tang et al. (27 Dec 2022). Assessing the European Electric-Mobility Transition: Emissions from Electric Vehicle Manufacturing and Use in Relation to the EU Greenhouse Gas Emission Targets.
45. Lázaro V. Cremades et al. (02 Dec 2022). Analysis of the Future of Mobility: The Battery Electric Vehicle Seems Just a Transitory Alternative.
46. Martina Haase et al. (24 Oct 2022). Multi-criteria decision analysis for prospective sustainability assessment of alternative technologies and fuels for individual motorized transport.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This paper examines and reviews historical data from various types of electric vehicle markets, considering such major driving factors for the EV market: technology advancement, policy innovation, and consumer preference. This paper will apply ARIMA/ SARIMA to conduct appropriate forecasting of consumer preference for various types of vehicles with techniques like logistic regression, Random Forests, and SVM. These will provide useful insight into the trends of future development of various types of electric vehicles.

3.2 Research Framework

This research framework includes the following steps:

1. Problem Definition and Literature Review
2. Data Collection: Retrieve data from Kaggle using specific keywords.
3. Data Pre-processing: Cleaning and preparing data for further analysis.
4. Model Building: Building and training model.
5. Model Evaluation: Evaluation model performance using historical trends.
6. Trend Analysing: Exploring the trend of total annual sales of electric vehicles.

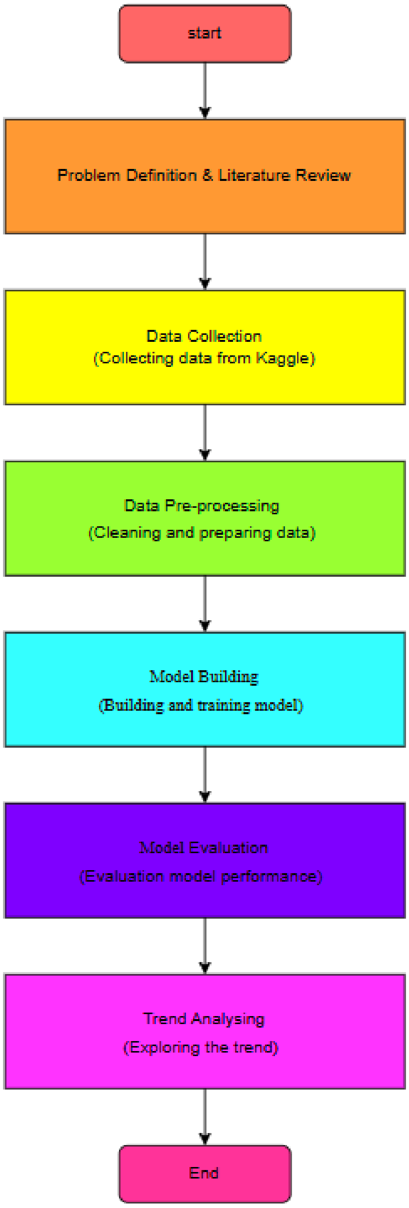

Figure 3.1 General Project Flow

3.3 Problem Formulation

This study will analyze the historical development trends of various electric vehicle (EV) markets and consider the main factors that affect the development trends, so as to make accurate predictions on the future development trends of the market. In order to ensure the accuracy of the prediction of the future development trends of the market, this study needs to solve the following two main problems:

a. Detailed analysis of the historical development trends of various electric vehicle (EV) markets, and then derive the general development trends;

b. Substitute historical data to verify the accuracy of the prediction results, and further revise and improve the prediction results by considering the influence of factors such as technology and policies.

3.4 Data Collection

The data required for this study comes from Kaggle, and through analysis and comparison, effective data cleaning is performed to obtain valid data.

The keywords used to capture the data are mainly:

a. “Development trend of various types of electric vehicles”

b. “Historical sales data of various types of electric vehicles in the world”

c. “Historical sales data of the global automobile market”

The captured data mainly includes the following ranges:

a. The time range covers the historical sales data of the electric vehicle market from 2010 to 2024

b. Detailed information includes detailed historical statistical data of various types of electric vehicles, including car brands, car types, and locations;

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	VIN (1-10)	County	City	State	Postal Cod	Model Yea	Make	Model	Electric Vel	Clean Alter	Electric Ra	Base MSRF	Legislative	DOL Vehicle ID	Vehicle Loc	Electric Uti	2020 Census Tract	
2	5YJYGDEE1	King	Seattle	WA	98122	2020	TESLA	MODEL Y	Battery Ele	Clean Alter	291	0	37	125701579	POINT (-1	CITY OF SE	5.3E+10	
3	7SAYGDEE	Snohomish	Bothell	WA	98021	2023	TESLA	MODEL Y	Battery Ele	Eligibility u	0	0	1	244285107	POINT (-1	PUGET SOI	5.31E+10	
4	5YJSA1E4X	King	Seattle	WA	98109	2019	TESLA	MODEL S	Battery Ele	Clean Alter	270	0	36	156773144	POINT (-1	CITY OF SE	5.3E+10	
5	5YJSA1E27	King	Issaquah	WA	98027	2016	TESLA	MODEL S	Battery Ele	Clean Alter	210	0	5	165103011	POINT (-1	PUGET SOI	5.3E+10	
6	5YJYGDEE5	Kitsap	Suquamish	WA	98392	2021	TESLA	MODEL Y	Battery Ele	Eligibility u	0	0	23	205138552	POINT (-1	PUGET SOI	5.3E+10	
7	3FA6P0SU	Thurston	Yelm	WA	98597	2017	FORD	FUSION	Plug-in Hy	Not eligibl	21	0	2	122057736	POINT (-1	PUGET SOI	5.31E+10	
8	1N4AZ0CP	Yakima	Yakima	WA	98903	2013	NISSAN	LEAF	Battery Ele	Clean Alter	75	0	14	150126840	POINT (-1	PACIFICOR	5.31E+10	
9	KNAGV4LD	Snohomish	Bothell	WA	98012	2018	KIA	OPTIMA	Plug-in Hy	Not eligibl	29	0	1	290605598	POINT (-1	PUGET SOI	5.31E+10	
10	1N4AZ0CP	Kitsap	Port Orcha	WA	98366	2015	NISSAN	LEAF	Battery Ele	Clean Alter	84	0	26	137322111	POINT (-1	PUGET SOI	5.3E+10	
11	5UXTA6C0	King	Auburn	WA	98001	2022	BMW	X5	Plug-in Hy	Clean Alter	30	0	47	240226332	POINT (-1	PUGET SOI	5.3E+10	
12	5YJYGDEEX	King	Seattle	WA	98144	2020	TESLA	MODEL Y	Battery Ele	Clean Alter	291	0	37	113323024	POINT (-1	CITY OF SE	5.3E+10	
13	WBY8P8C5	Kitsap	Bainbridge	WA	98110	2019	BMW	I3	Plug-in Hy	Clean Alter	126	0	23	228846642	POINT (-1	PUGET SOI	5.3E+10	
14	1G1FZ6S0	Yakima	Yakima	WA	98908	2021	CHEVROLE	BOLT EV	Battery Ele	Eligibility u	0	0	14	156686106	POINT (-1	PACIFICOR	5.31E+10	
15	WA1E2AF	Snohomish	Lynnwood	WA	98036	2021	AUDI	Q5 E	Plug-in Hy	Not eligibl	18	0	1	168371122	POINT (-1	PUGET SOI	5.31E+10	
16	1N4AZ0CP	King	Seattle	WA	98119	2015	NISSAN	LEAF	Battery Ele	Clean Alter	84	0	36	126304132	POINT (-1	CITY OF SE	5.3E+10	
17	1N4AZ0CP	King	Seattle	WA	98107	2013	NISSAN	LEAF	Battery Ele	Clean Alter	75	0	43	100938848	POINT (-1	CITY OF SE	5.3E+10	
18	1N4AZ0CP	Snohomish	Lynnwood	WA	98087	2013	NISSAN	LEAF	Battery Ele	Clean Alter	75	0	21	139800496	POINT (-1	PUGET SOI	5.31E+10	
19	1N4BZ0CP	Snohomish	Bothell	WA	98021	2017	NISSAN	LEAF	Battery Ele	Clean Alter	107	0	1	348979466	POINT (-1	PUGET SOI	5.31E+10	

Figure 3.2 Initial Dataset

	region	category	parameter	mode	powertrain	year	unit	value
2	Australia	Historical	EV sales sh	Cars	EV	2011	percent	0.0065
3	Australia	Historical	EV stock sh	Cars	EV	2011	percent	0.00039
4	Australia	Historical	EV sales	Cars	BEV	2011	Vehicles	49
5	Australia	Historical	EV stock	Cars	BEV	2011	Vehicles	49
6	Australia	Historical	EV stock	Cars	BEV	2012	Vehicles	220
7	Australia	Historical	EV sales	Cars	BEV	2012	Vehicles	170
8	Australia	Historical	EV stock sh	Cars	EV	2012	percent	0.0024
9	Australia	Historical	EV sales sh	Cars	EV	2012	percent	0.03
10	Australia	Historical	EV stock	Cars	PHEV	2012	Vehicles	80
11	Australia	Historical	EV sales	Cars	PHEV	2012	Vehicles	80
12	Australia	Historical	EV sales	Cars	PHEV	2013	Vehicles	100
13	Australia	Historical	EV stock	Cars	PHEV	2013	Vehicles	180
14	Australia	Historical	EV sales sh	Cars	EV	2013	percent	0.034
15	Australia	Historical	EV stock sh	Cars	EV	2013	percent	0.0046
16	Australia	Historical	EV sales	Cars	BEV	2013	Vehicles	190
17	Australia	Historical	EV stock	Cars	BEV	2013	Vehicles	410
18	Australia	Historical	EV stock	Cars	BEV	2014	Vehicles	780
19	Australia	Historical	EV sales	Cars	BEV	2014	Vehicles	370
20	Australia	Historical	EV stock sh	Cars	EV	2014	percent	0.014
21	Australia	Historical	EV sales sh	Cars	EV	2014	percent	0.16
22	Australia	Historical	EV stock	Cars	PHEV	2014	Vehicles	1100
23	Australia	Historical	EV sales	Cars	PHEV	2014	Vehicles	950
24	Australia	Historical	EV sales	Cars	PHEV	2015	Vehicles	1000
25	Australia	Historical	EV stock	Cars	PHEV	2015	Vehicles	2100
26	Australia	Historical	EV sales sh	Cars	EV	2015	percent	0.2

Figure 3.3 Initial Dataset

```
In [3]: # Load the dataset

from google.colab import drive
drive.mount('/content/drive')
file_path = '/content/drive/MyDrive/Colab Notebooks/IEA Global EV Data 2024.csv'
df = pd.read_csv(file_path)

# Display the first few rows
print(df.head())
```

```
In [ ]: df.head()
```

Figure 3.4 Loading and Showing the Initial Database

As shown in Figure 3.2, the original data contains more than 100,000 rows of data, including various electric vehicle sales data from 2010 to 2023 in various countries. This study will compare and analyze multi-source data to ensure the authenticity of the data.

3.4.1 Data Pre-Processing and Initial Analysis

Before data cleaning, it is necessary to complete the preliminary analysis and processing of the original data.

The data preprocessing stage includes:

a. 1. Visualize the original data, analyze the overall data, and understand the data characteristics.

b. 2. Data conversion: convert the data types of different sources after merging into the same type of data for subsequent processing.

CHECKING FOR DATA TYPES

```
df.dtypes
```

	0
TransactionID	object
Date	object
Time	object
CustomerID	int64
FirstName	object
Lastname	object
Shipping_State	object
ItemID	object
Description	object
Retail_Price	float64
Loyalty_Discount	float64
Discount	float64
Discounted_Price	float64

dtype: object

CONVERT DATA TYPES FOR 'TransactionID' TO STRING OBJECT

```
df.TransactionID = df.TransactionID.astype(str)
df.ItemID = df.ItemID.astype(str)
df.CustomerID = df.CustomerID.astype(str)
```

Figure 3.5 Analyzing and convert Database

3.4.2 Data Cleaning

This operation realizes data cleaning, removes information that does not meet the research scope and erroneous information, extracts the data required for the research, and finally exports it into new usable data.

INSERT SPLITTED 'Timestamp' COLUMNS INTO df, RENAMING AS 'Date', 'Time' AT LOCATION [1] AND [2] RESPECTIVELY

```
[ ] df.insert(loc = 1,column = 'Date',value = dfsplit[0])
    df.insert(loc = 2,column = 'Time',value = dfsplit[1])
    df
```

显示隐藏的输出项

DROP 'Timestamp' COLUMN FROM DF

```
df = df.drop(['Timestamp'], axis=1)
df
```

显示隐藏的输出项

```
df.rename(columns = {'Name1':'FirstName'}, inplace = True)
df.rename(columns = {'Surname':'Lastname'}, inplace = True)
df
```

显示隐藏的输出项

CREATE NEW COLUMN 'Discount' AND 'Discounted_Price' BY CALCULATING 'Retail_Price' AND 'Loyalty Discount'

```
[ ] df['Discount'] = df['Retail_Price']*df['Loyalty_Discount']
```

```
[ ] df['Discounted_Price'] = df['Retail_Price']-df['Discount']
    df
```

Figure 3.6 Extracting valid information

```
null_counts=df.isnull().sum()
null_counts
```

	0
TransactionID	0
Date	0
Time	0
CustomerID	0
FirstName	0
Lastname	0
Shipping_State	0
ItemID	0
Description	0
Retail_Price	0
Loyalty_Discount	0
Discount	0
Discounted_Price	0

dtype: int64

Figure 3.7 Checking missing values

```
df.describe()
```

	TransactionID	CustomerID	ItemID	Retail_Price	Loyalty_Discount	Discount	Discounted_Price
count	3455.000000	3.455000e+03	3.455000e+03	3455.000000	3455.000000	3455.000000	3455.000000
mean	111528.000000	1.797979e+08	5.276712e+09	58.526237	0.050457	2.968812	55.557323
std	997.516917	9.563412e+07	2.600486e+09	34.464217	0.032215	2.833184	32.728502
min	109801.000000	1.000000e+08	1.039855e+09	5.600000	0.000000	0.000000	5.040000
25%	110664.500000	1.000003e+08	2.963301e+09	31.800000	0.020000	0.777550	29.575000
50%	111528.000000	1.000009e+08	5.145202e+09	51.660000	0.050000	2.079000	49.810000
75%	112391.500000	2.000009e+08	7.645689e+09	79.800000	0.080000	4.374000	76.920000
max	113255.000000	4.000009e+08	9.916068e+09	159.800000	0.100000	14.382000	159.800000

Figure 3.8 Describing data frame

EXPORT DATAFRAME AS .XLSX AND .CSV

```
# Export transformed table to .csv
df.to_csv(f"Transaction.csv",index=False)
```

```
# Export transformed table to .xlsx
df.to_excel(f"OLTP.xlsx",index=False)
```

Figure 3.9 Export the New Dataset

3.5 Data Modeling

The processed data is further analyzed and processed, mainly including visual display and trend prediction.

```
[1] import pandas as pd
    from sklearn.model_selection import train_test_split
    from sklearn.preprocessing import StandardScaler
```

Generate | 10 random numbers using numpy | Close

```
# Load the dataset

from google.colab import drive
drive.mount('/content/drive')
file_path = '/content/drive/My Drive/content/dataset_ev/IEA_Global_EV_Data_2024_new.csv'
df = pd.read_csv(file_path)

# Display the first few rows
print(df.head())
```

```
Mounted at /content/drive
    region    category       parameter  mode powertrain  year      unit  \
0  Austria  Historical        EV stock  Cars        BEV  2010  Vehicles
1  Austria  Historical  EV stock share  Cars         EV  2010   percent
2  Belgium  Historical        EV stock Buses        BEV  2010  Vehicles
3  Belgium  Historical        EV sales  Vans        BEV  2010  Vehicles
4  Belgium  Historical        EV stock  Vans        BEV  2010  Vehicles
```

Figure 3.10 Loading Dataset

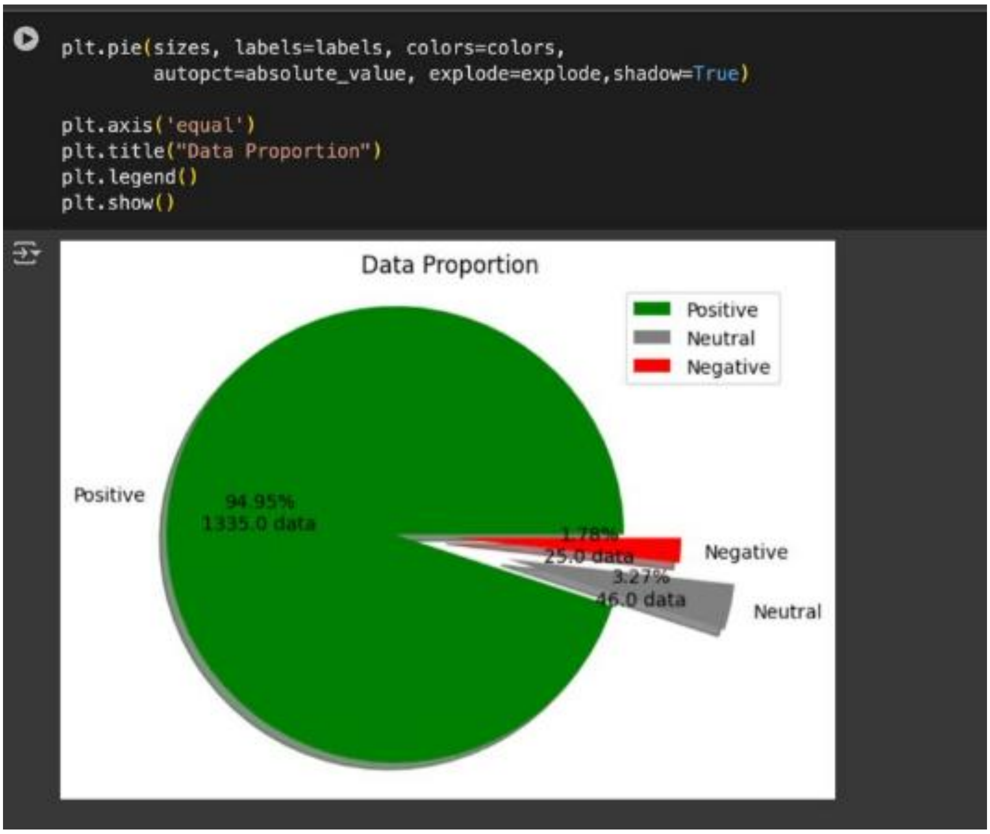

Figure 3.11 Data visualization

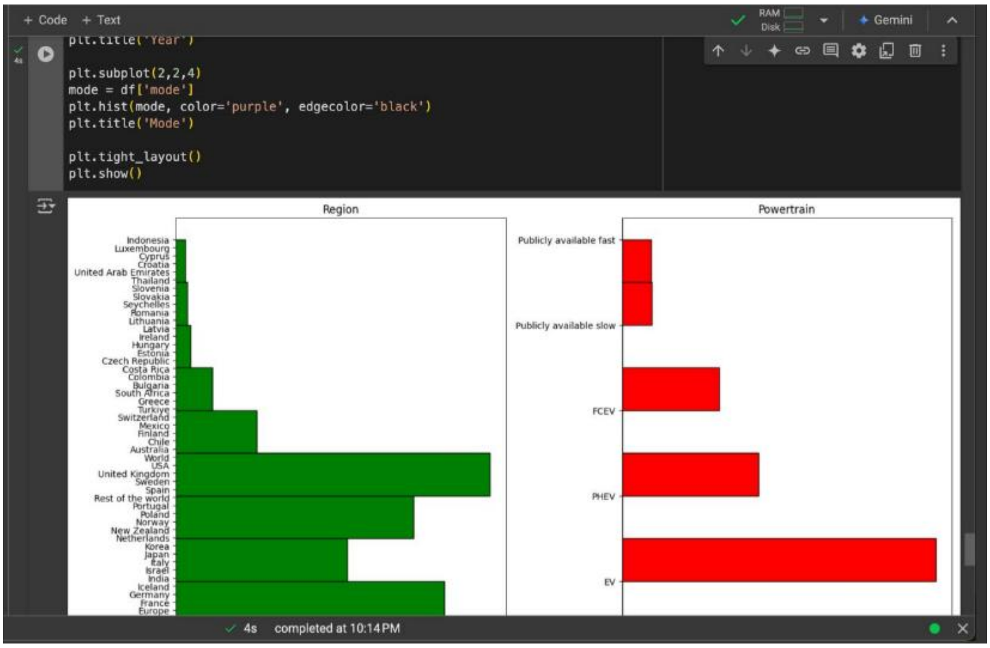

Figure 3.12 Data visualization

3.6 Summary

This chapter details the research process from data collection to classification model evaluation. This process demonstrates the implementation of specific research methods such as data collection and analysis, result modeling and visualization.

CHAPTER 4

INTRODUCTION

4.1 Overview

This section discusses the analysis and prediction of the market trends for various types of electric vehicles. This chapter will start with the identification of datasets, analyze datasets, establish historical trend models, and use machine learning techniques to build the results of the model. The mechanical learning techniques used include Long Short-Term Memory (LSTM), Deep Learning Models, and Hybrid Models. The results of the predictions based on these machine learning techniques on the trends in the EV market are presented in the following sections.

4.2 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is a method for exploring and understanding data that can help us spot data characteristics, trends, and outliers. When predicting the development trend of the electric vehicle market, the detailed analysis and processing of the data using the EDA method is helpful to make a more accurate prediction of the development trend of the electric vehicle market, and the EDA can be carried out through the following steps:

```
df
```

	region	category	parameter	mode	powertrain	year	unit	value	percentage
0	Austria	Historical	EV stock	Cars	BEV	2010	Vehicles	350	35000,00%
1	Austria	Historical	EV stock share	Cars	EV	2010	percent	789.999.961.853	78999996185300,00%
2	Belgium	Historical	EV stock	Buses	BEV	2010	Vehicles	3	300,00%
3	Belgium	Historical	EV sales	Vans	BEV	2010	Vehicles	7	700,00%
4	Belgium	Historical	EV stock	Vans	BEV	2010	Vehicles	62	6200,00%
...	...	...	...	...	...	...	...	...	...
12649	World	Projection-STEPS	EV sales share	Cars	EV	2035	percent	55	5500,00%
12650	World	Projection-STEPS	EV stock share	Cars	EV	2035	percent	31	3100,00%
12651	World	Projection-APS	EV charging points	EV	Publicly available fast	2035	charging points	9400000	940000000,00%
12652	World	Projection-APS	EV charging points	EV	Publicly available slow	2035	charging points	15000000	1500000000,00%
12653	World	Projection-STEPS	EV stock share	Trucks	EV	2035	percent	9	900,00%

12654 rows × 9 columns

Figure 4.1 Dataset

As shown in the figure above, the source dataset used in this study includes a total of 12,654 rows and 9 columns.

```
# One-hot encode categorical variables
X = pd.get_dummies(df.drop(columns=['value']), columns=['region', 'category', 'parameter', 'mode', 'powertrain', 'unit'])

# Separate the target variable
y = df['value']

print(X.head())  # Check the one-hot encoded features
```

```
   year    percentage  region_Australia  region_Austria  region_Belgium  \
0  2010  3.500000e+06             False            True           False
1  2010  7.900000e+15             False            True           False
2  2010  3.000000e+04             False           False            True
3  2010  7.000000e+04             False           False            True
4  2010  6.200000e+05             False           False            True

   region_Brazil  region_Bulgaria  region_Canada  region_Chile  region_China  \
0          False            False          False         False         False
1          False            False          False         False         False
2          False            False          False         False         False
3          False            False          False         False         False
4          False            False          False         False         False

   ...  powertrain_FCEV  powertrain_PHEV  powertrain_Publicly available fast  \
0  ...            False            False                               False
1  ...            False            False                               False
2  ...            False            False                               False
3  ...            False            False                               False
4  ...            False            False                               False

   powertrain_Publicly available slow  unit_GWh  unit_Milion barrels per day  \
0                               False     False                        False
1                               False     False                        False
2                               False     False                        False
3                               False     False                        False
4                               False     False                        False

   unit_Oil displacement, million lge  unit_Vehicles  unit_charging points  \
0                               False           True                 False
1                               False          False                 False
2                               False           True                 False
3                               False           True                 False
4                               False           True                 False
```

Figure 4.2 Encode Categorical Variables

Encode Categorical Variables are coded for subsequent data analysis cleanup.

```
import numpy as np

# Check for NaN values
print("NaN values in X_train:", np.isnan(X_train).sum())
print("NaN values in X_test:", np.isnan(X_test).sum())

# Check for infinite values
print("Infinite values in X_train:", np.isinf(X_train).sum())
print("Infinite values in X_test:", np.isinf(X_test).sum())
```

```
NaN values in X_train: year                        0
region_Australia                 0
region_Austria                   0
region_Belgium                   0
region_Brazil                    0
                                ..
unit_Milion barrels per day      0
unit_Oil displacement, million lge   0
unit_Vehicles                    0
unit_charging points             0
unit_percent                     0
Length: 83, dtype: int64
NaN values in X_test: year                         0
region_Australia                 0
region_Austria                   0
region_Belgium                   0
region_Brazil                    0
                                ..
unit_Milion barrels per day      0
unit_Oil displacement, million lge   0
unit_Vehicles                    0
unit_charging points             0
unit_percent                     0
Length: 83, dtype: int64
Infinite values in X_train: year                        0
region_Australia                 0
region_Austria                   0
region_Belgium                   0
region_Brazil                    0
                                ..
unit_Milion barrels per day      0
unit_Oil displacement, million lge   0
unit_Vehicles                    0
```

Figure 4.3 Check for Missing or Infinite Values

Ensure that there are no missing (NaN) or infinite values in your training and test datasets before training the model.

```
[23] from sklearn.model_selection import train_test_split

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 4.4 Split the Dataset

```
[25] from sklearn.preprocessing import StandardScaler

     # Standardize the features
     scaler = StandardScaler()
     X_train = scaler.fit_transform(X_train)
     X_test = scaler.transform(X_test)
```

Figure 4.5 Standardize the Features

```
[27] # Reshape data to fit the 1D CNN input
     X_train = X_train.reshape(X_train.shape[0], X_train.shape[1], 1)
     X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)

     df
```

	region	category	parameter	mode	powertrain	year	unit	value
0	Australia	Historical	EV stock share	Cars	EV	2011	percent	3.900000e-04
1	Australia	Historical	EV sales share	Cars	EV	2011	percent	6.500000e-03
2	Australia	Historical	EV sales	Cars	BEV	2011	Vehicles	4.900000e+01
3	Australia	Historical	EV stock	Cars	BEV	2011	Vehicles	4.900000e+01
4	Australia	Historical	EV stock	Cars	BEV	2012	Vehicles	2.200000e+02
...	...	...	...	...	...	...	...	...
12649	World	Projection-STEPS	EV sales share	Cars	EV	2035	percent	5.500000e+01
12650	World	Projection-STEPS	EV stock share	Cars	EV	2035	percent	3.100000e+01
12651	World	Projection-APS	EV charging points	EV	Publicly available fast	2035	charging points	9.400000e+06
12652	World	Projection-APS	EV charging points	EV	Publicly available slow	2035	charging points	1.500000e+07
12653	World	Projection-STEPS	EV stock share	Trucks	EV	2035	percent	9.000000e+00

12654 rows × 8 columns

Figure 4.6 Reshape Data for 1D CNN

Figure 4.4, Figure 4.5, and Figure 4.5 show the steps including Split the dataset into training and testing sets and standardize the features to ensure that they are on a similar scale. Then, reshape the data to be compatible with a 1D CNN.

```python
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv1D, Dense, Flatten, Dropout

# Build the model
model = Sequential()

# Add a 1D convolutional layer
model.add(Conv1D(filters=64, kernel_size=2, activation='relu', input_shape=(X_train.shape[1], 1)))

# Add a Flatten layer
model.add(Flatten())

# Add a Dense layer
model.add(Dense(128, activation='relu'))

# Add a Dropout layer for regularization
model.add(Dropout(0.2))

# Add the output layer
model.add(Dense(1, activation='linear'))

# Compile the model
model.compile(optimizer='adam', loss='mean_squared_error')

# Display the model's architecture
model.summary()
```

```
/usr/local/lib/python3.10/dist-packages/keras/src/layers/convolutional/base_conv.py:107: UserWarning: Do not pass an
  super().__init__(activity_regularizer=activity_regularizer, **kwargs)
Model: "sequential_2"
```

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 82, 64)	192
flatten (Flatten)	(None, 5248)	0
dense (Dense)	(None, 128)	671,872
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 1)	129

```
Total params: 672,193 (2.56 MB)
Trainable params: 672,193 (2.56 MB)
Non-trainable params: 0 (0.00 B)
```

Figure 4.7 Build the 1D CNN Model

This step builds a simple 1D CNN model for regression.

```
[ ] # Train the model
    history = model.fit(X_train, y_train, epochs=50, validation_data=(X_test, y_test), batch_size=32)

317/317 ──────────────────── 9s 12ms/step - loss: 66445782286336.0000 - val_loss: 8370079242649
Epoch 23/50
317/317 ──────────────────── 7s 19ms/step - loss: 55054304280576.0000 - val_loss: 8366858017177
Epoch 24/50
317/317 ──────────────────── 4s 13ms/step - loss: 27219581730816.0000 - val_loss: 8363263498649
Epoch 25/50
317/317 ──────────────────── 4s 12ms/step - loss: 80761742950400.0000 - val_loss: 8358953431859
Epoch 26/50
317/317 ──────────────────── 5s 16ms/step - loss: 34965416837120.0000 - val_loss: 8354764161024
Epoch 27/50
317/317 ──────────────────── 6s 19ms/step - loss: 45324504662016.0000 - val_loss: 8350851714252
Epoch 28/50
317/317 ──────────────────── 9s 27ms/step - loss: 37333755756544.0000 - val_loss: 8346719485952
Epoch 29/50
317/317 ──────────────────── 7s 15ms/step - loss: 55015263698944.0000 - val_loss: 8342290300928
Epoch 30/50
317/317 ──────────────────── 4s 12ms/step - loss: 59982254964736.0000 - val_loss: 8337708443238
Epoch 31/50
317/317 ──────────────────── 4s 14ms/step - loss: 25464619925504.0000 - val_loss: 8332803624140
Epoch 32/50
317/317 ──────────────────── 6s 17ms/step - loss: 47399196164096.0000 - val_loss: 8327621980979
Epoch 33/50
317/317 ──────────────────── 4s 12ms/step - loss: 27119182675968.0000 - val_loss: 8322425238323
Epoch 34/50
317/317 ──────────────────── 6s 14ms/step - loss: 55144095940608.0000 - val_loss: 8317090083635
Epoch 35/50
317/317 ──────────────────── 6s 17ms/step - loss: 28061691019264.0000 - val_loss: 8311794355404
Epoch 36/50
317/317 ──────────────────── 4s 12ms/step - loss: 41846394847232.0000 - val_loss: 8306336727040
Epoch 37/50
317/317 ──────────────────── 4s 12ms/step - loss: 20842754867200.0000 - val_loss: 8300655122841
Epoch 38/50
317/317 ──────────────────── 6s 20ms/step - loss: 22580278853632.0000 - val_loss: 8295339261952
Epoch 39/50
317/317 ──────────────────── 4s 13ms/step - loss: 30645900279808.0000 - val_loss: 8289512534835
Epoch 40/50
317/317 ──────────────────── 4s 13ms/step - loss: 27313483808768.0000 - val_loss: 8283814992281
Epoch 41/50
317/317 ──────────────────── 7s 19ms/step - loss: 18212183867392.0000 - val_loss: 8277769322496
Epoch 42/50
317/317 ──────────────────── 4s 12ms/step - loss: 85511347634176.0000 - val_loss: 8272130500198
```

Figure 4.8 Train the Model

Train the model using the training data.

```
# Evaluate the model
test_loss = model.evaluate(X_test, y_test)
print(f'Test Loss: {test_loss}')

80/80 ──────────────────── 1s 8ms/step - loss: nan
Test Loss: nan
```

Figure 4.9 Evaluate the Model

```
[ ] import matplotlib.pyplot as plt

    # Plot the training and validation loss
    plt.plot(history.history['loss'], label='Training Loss')
    plt.plot(history.history['val_loss'], label='Validation Loss')
    plt.title('Training and Validation Loss')
    plt.xlabel('Epochs')
    plt.ylabel('Loss')
    plt.legend()
    plt.show()
```

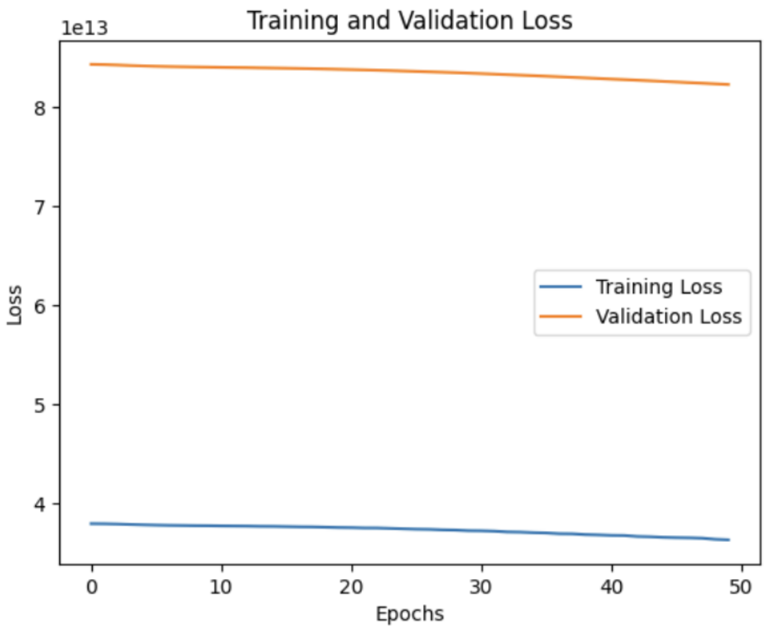

Figure 4.10 Visualize the Training Process

Visualize the training and validation loss to understand how well the model is training.

4.3 Explore the trend of total annual sales of electric vehicles

```
import pandas as pd
import matplotlib.pyplot as plt

a, b, c, d = plt.cm.Reds, plt.cm.Greens, plt.cm.Blues, plt.cm.Purples
mode = ['Buses', 'Cars', 'Trucks', 'Vans']
group_mode = data_sale.groupby(['mode'])['value'].sum()

# Convert 'value' column to numeric before grouping
data_sale['value'] = pd.to_numeric(data_sale['value'], errors='coerce')
group_mode = data_sale.groupby(['mode'])['value'].sum()

# Convert values to numeric, handling potential errors
mode_num = []
for i in mode:
    try:
        mode_num.append(int(group_mode[i]))
    except ValueError:
        print(f"Warning: Could not convert value for mode '{i}' to int. Using 0 instead.")
        mode_num.append(0)  # Or handle the error differently

plt.figure(figsize=(6, 6), dpi=300)
explode = [0.02, 0.02, 0.02, 0.02]
pie1, _, _ = plt.pie(mode_num, labels=mode, autopct='%1.2f%%', pctdistance=1.1,
                     labeldistance=0.8, radius=1.2, colors=[a(0.6), b(0.6), c(0.6), d(0.6)]
                     explode=explode, textprops={'fontsize': 10})
plt.setp(pie1, width=0.5, edgecolor='k')
plt.title('Proportion of various mode vehicles sales', fontsize=20)
plt.show()
```

Figure 4.11 Analyse the proportional composition of sales of vehicles

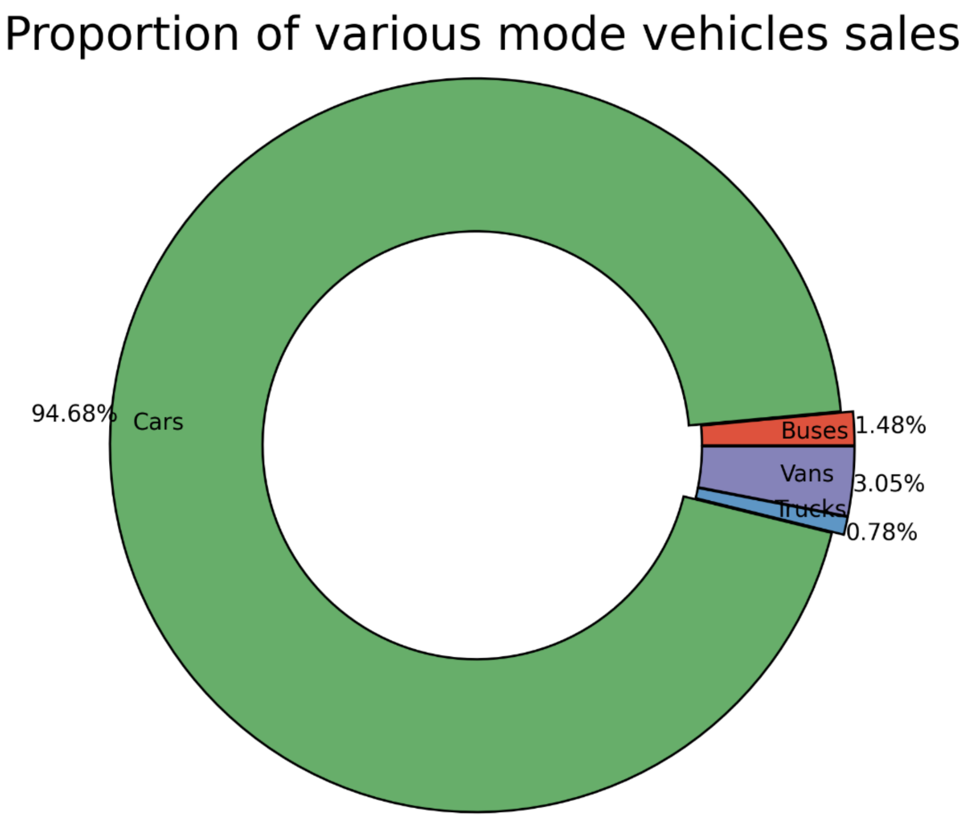

Figure 4.12 Proportion of various mode vehicles sales

As can be seen from the above chart, among the electric vehicles sold, cars account for the overwhelming majority, followed by minivans, buses, and trucks. Next, we will analyze the proportion of vehicle sales by different powertrains.

```
group_powertrain=data_sale.groupby('powertrain')['value'].sum()
plt.figure(figsize=(6,6),dpi=300)
explode=[0.02,0.02,0.02,]
pie,_,_=plt.pie(group_powertrain,labels=group_powertrain.index,autopct='%1.2f%%',pctdistance=1.1,labeldistance=0.8,ra
                                   explode=explode)
plt.setp(pie,width=0.5,edgecolor='k')
plt.title('Proportion  of  various  powertrain  vehicles  sales',fontsize=20)
plt.show()
```

Figure 4.13 Analyze the proportional composition of sales of vehicles

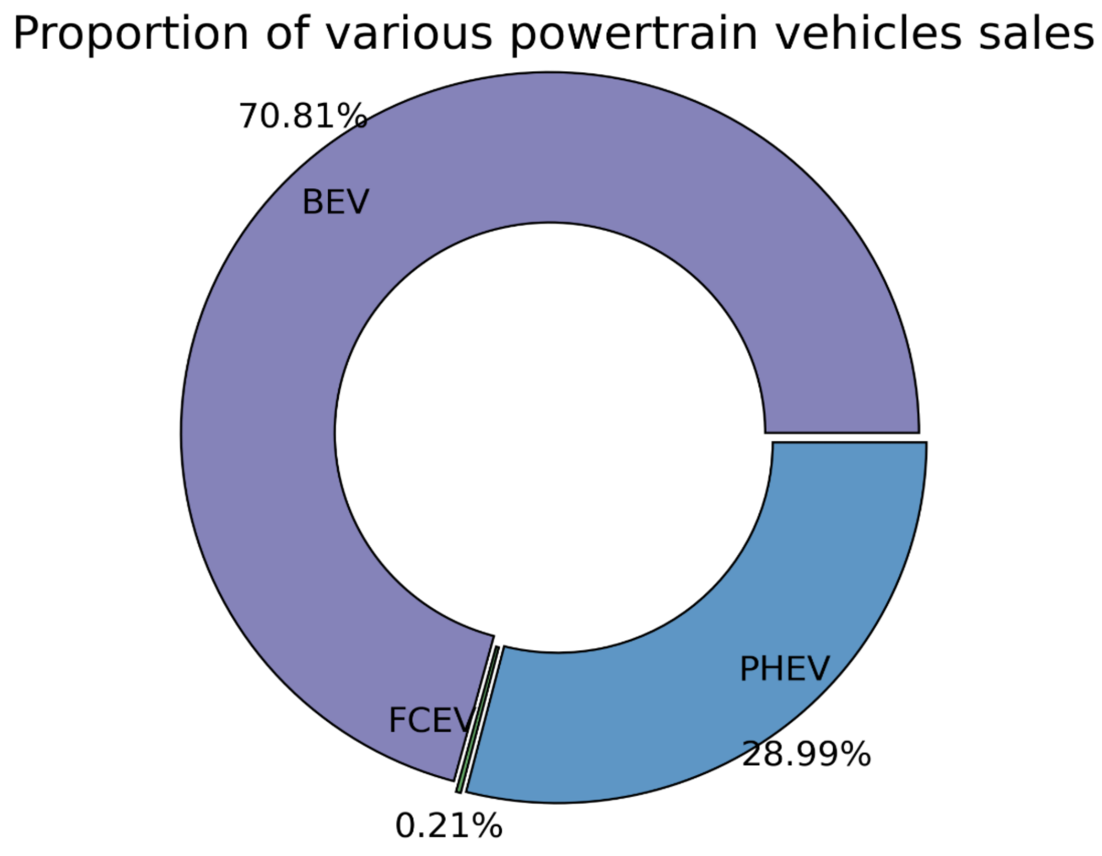

Figure 4.14 Proportion of various powertrain vehicles sales

As shown in the above chart, the sales proportion of Battery Electric Vehicles (BEV) accounts for more than one-third, Plug-in Hybrid Electric Vehicles (PHEV) are slightly lower than one-third, and Fuel Cell Electric Vehicles (FCEV) are negligible. Next, we will analyse the annual sales and proportion of Battery Electric Vehicles (BEV) and Plug-in Hybrid Electric Vehicles (PHEV).

```python
group_year_power=data_sale.groupby(['year','powertrain'])['value'].sum()
group_year_power=group_year_power.unstack(level=-1)
group_year_power['sum']=group_year_power.sum(axis=1)
xaxis=group_year_power.index
bev=group_year_power['BEV'].astype(int)
phev=-group_year_power['PHEV'].astype(int)
plt.figure(figsize=(9,6),dpi=300,constrained_layout=True)
plt.bar(xaxis,bev,color='yellowgreen',label='Battery Electric Vehicle')
plt.bar(xaxis,phev,color='palevioletred',label='Plug-in hybrid electric vehicle')
plt.legend(loc='best')
plt.yticks([])
for i ,j in zip(xaxis,bev):
        plt.annotate(j,(i,j),fontsize=10,ha='center')
for i ,j in zip(xaxis,phev):
        plt.annotate(-j,(i,j-500000),fontsize=10,ha='center')
plt.xticks(xaxis)
plt.xlabel('Year',fontsize=20)
plt.ylabel('Sales',fontsize=20)
plt.title('Annual various powertrain vehicles sales',fontsize=20)
plt.show()
```

Figure 4.15 Analyze the trends of electric vehicles.

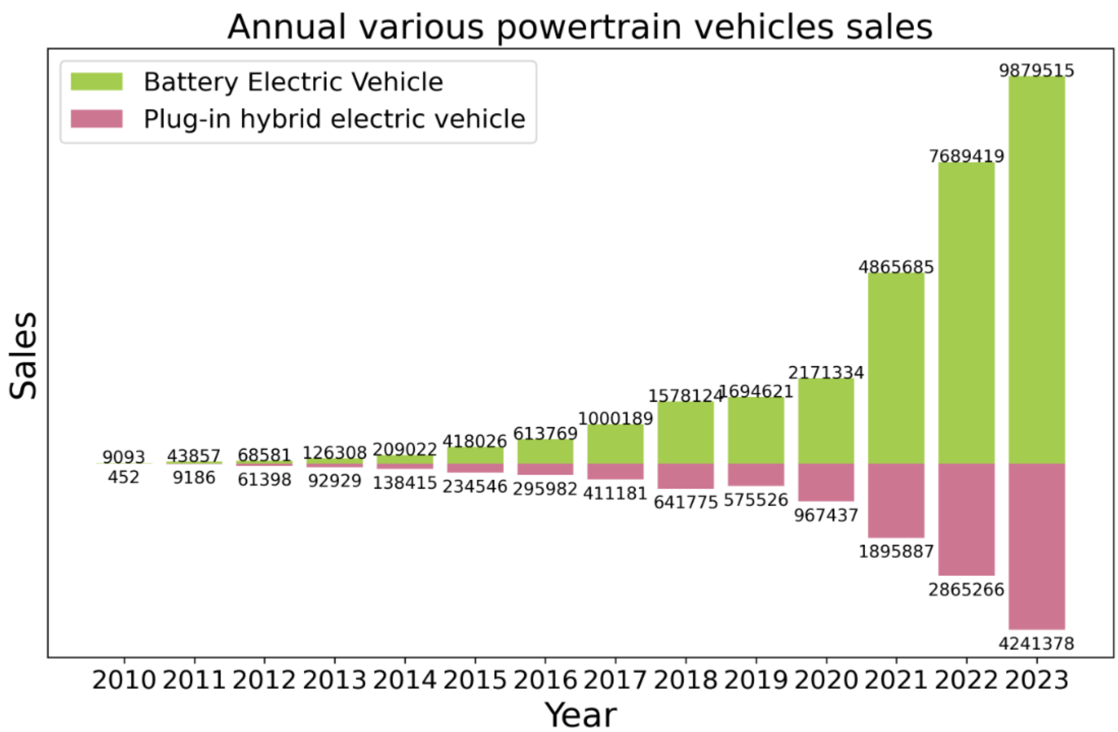

Figure 4.16 Chart of Annual various powertrain vehicles sales

This graph analyses the annual sales of various powertrain vehicles between 2010 and 2023, especially for battery electric vehicles and plug-in hybrid electric vehicles, during which sales of battery electric vehicles increased significantly, especially after 2018, when the growth rate accelerated. In 2023, sales of battery electric vehicles reached 9879515 units. In contrast, sales of plug-in hybrid electric vehicles are also increasing, but at a relatively slower pace, with 4241378 units sold in 2023.

```
fig,ax=plt.subplots(figsize=(9,6),dpi=300,constrained_layout=True)
ax.set_xlim(0,100)
bev_ratio=group_year_power['BEV']/group_year_power['sum']*100
fcev_ratio=group_year_power['FCEV']/group_year_power['sum']*100
phev_ratio=group_year_power['PHEV']/group_year_power['sum']*100
ax.barh(xaxis,bev_ratio,color='yellowgreen',label='BEV')
ax.barh(xaxis,fcev_ratio,left=bev_ratio,color='cornflowerblue',label='FCEV')
ax.barh(xaxis,phev_ratio,left=bev_ratio+fcev_ratio,color='palevioletred',label='PHEV')
for  i,j  in  zip(bev_ratio,xaxis):
        ax.annotate(f'{i:.2f}%',(i/2,j),va='center')
for  i,j  in  zip(phev_ratio,xaxis):
        ax.annotate(f'{i:.2f}%',(100-i+i/2,j),va='center')
ax.set_yticks(xaxis)
ax.set_ylim(2009,2025)
ax.set_xlabel('Sales  percent')
ax.set_ylabel('Year')
ax.set_title('Annual  various  powertrain  vehicles  sales  proportion',fontsize=20)
ax.legend(loc='upper  left',ncols=3)
plt.show()
```

Figure 4.17 Analyze Annual various powertrain vehicles sales proportion

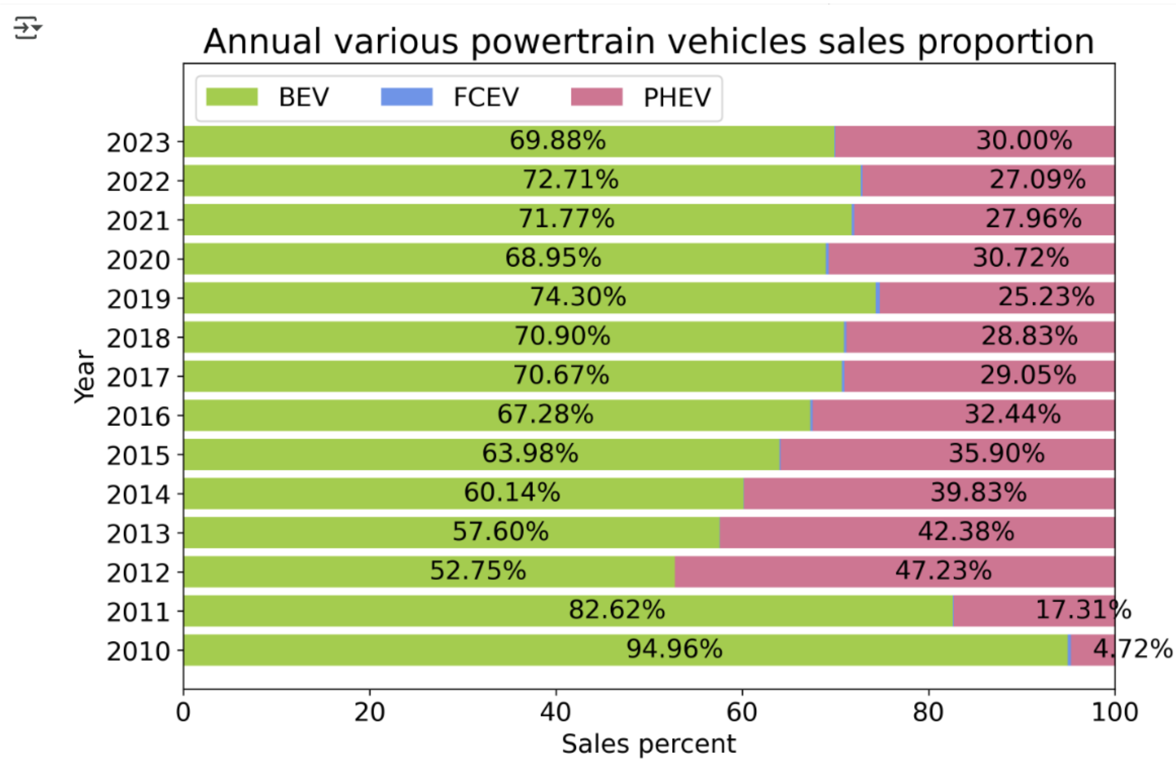

Figure 4.18 Chart about the proportional composition of sales of vehicles

This is the histogram of "Annual Percentage of Vehicle Sales by Powertrain", reflecting the sale percentages of three kinds of vehicles, namely BEVs, FCEVs, and PHEVs, from the period ranging between 2010 to 2023. Throughout these years from 2010 to 2023, BEVs were always at the top in the market. Though their market shares decreased, they have remained at a high level. The gradual growth in the market share of PHEVs demonstrates the interest and acceptance of hybrid technology in the market. FCEVs have not yet been able to capture a significant share of the market, reflecting challenges in their market acceptance.

```
[ ]  group_region=data_sale.groupby('region')['value'].sum().astype(int).reset_index()
     group_region.sort_values(by='value',ascending=False,inplace=True)
     group_region.reset_index(drop=True,inplace=True)
     new_row=pd.DataFrame([{'region':'rest of world','value':group_region['value'][10:].sum()}])
     group_region=group_region.drop(group_region[9:].index)
     group_region=pd.concat([group_region,new_row],ignore_index=True)
     print(group_region)
     fig,ax=plt.subplots(figsize=(6,6),dpi=300)
     pie,_,_=ax.pie(group_region['value'],labels=group_region['region'],autopct='%1.2f%%',pctdistance=0.8,radius=1,textpr
                      labeldistance=1.05,colors=plt.cm.Spectral(np.linspace(0,1,len(group_region['region']))))
     plt.setp(pie,width=0.5,edgecolor='k')
     ax.set_title('Proportion of total sales of electric vehicles',fontsize=20)
     plt.show()
```

Figure 4.19 Analyze the country with the highest annual sales of electric vehicles

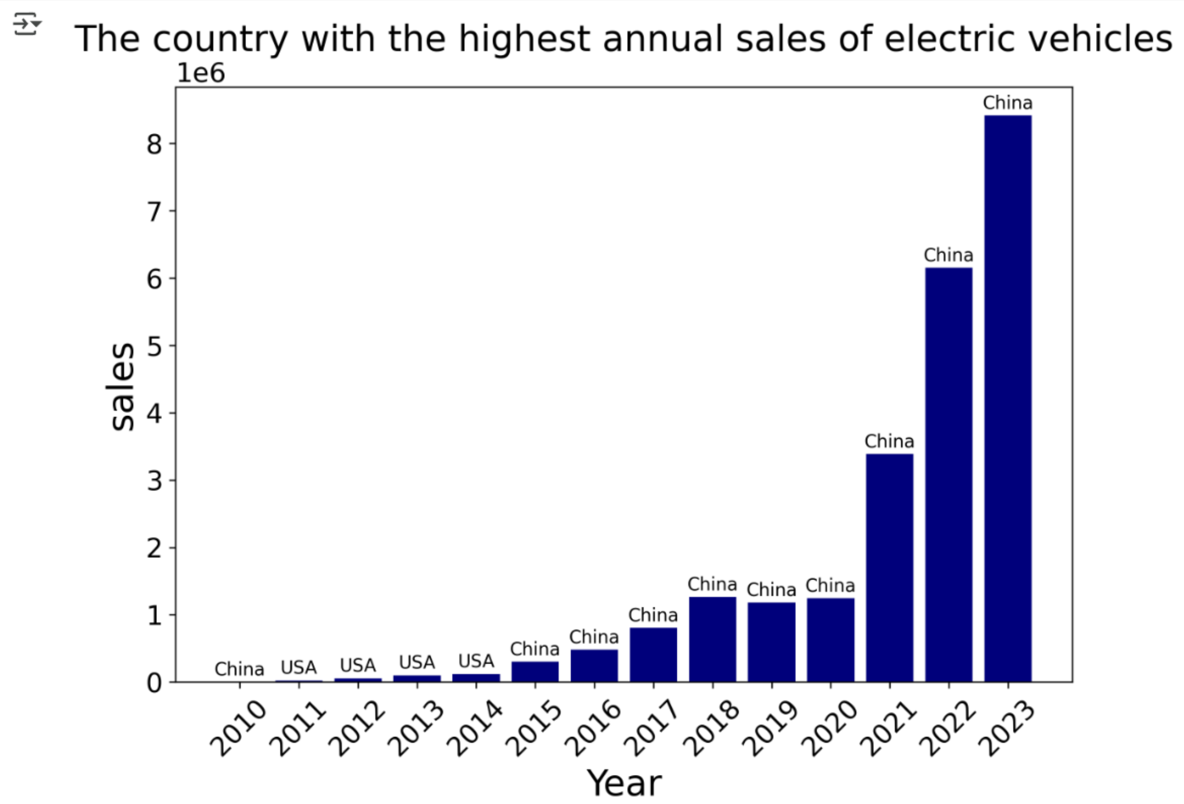

Figure 4.20 Chart about the trends of electric vehicles.

As can be seen from the above chart, except for the years 2012-2014, when the United States had the highest sales of electric vehicles, China has been the leading country in electric vehicle sales for the rest of the years. The reasons for this can be attributed to several factors:

1) China's vast population,

2) government policy support

3) the development of China's industrial technology.

```
[ ] group_region=data_sale.groupby('region')['value'].sum().astype(int).reset_index()
    group_region.sort_values(by='value',ascending=False,inplace=True)
    group_region.reset_index(drop=True,inplace=True)
    new_row=pd.DataFrame([{'region':'rest of world','value':group_region['value'][10:].sum()}])
    group_region=group_region.drop(group_region[9:].index)
    group_region=pd.concat([group_region,new_row],ignore_index=True)
    print(group_region)
    fig,ax=plt.subplots(figsize=(6,6),dpi=300)
    pie,_,_=ax.pie(group_region['value'],labels=group_region['region'],autopct='%1.2f%%',pctdistance=0.8,radius=1,textpr
                   labeldistance=1.05,colors=plt.cm.Spectral(np.linspace(0,1,len(group_region['region']))))
    plt.setp(pie,width=0.5,edgecolor='k')
    ax.set_title('Proportion of total sales of electric vehicles',fontsize=20)
    plt.show()
```

```
              region     value
0              China  23358308
1                USA   4770925
2            Germany   3012826
3             France   1666650
4     United Kingdom   1659853
5             Norway    879813
6        Netherlands    798918
7             Sweden    706166
8              Korea    676352
9      rest of world   4730788
```

Figure 4.21 Analyze the proportion of total sales of electric vehicles

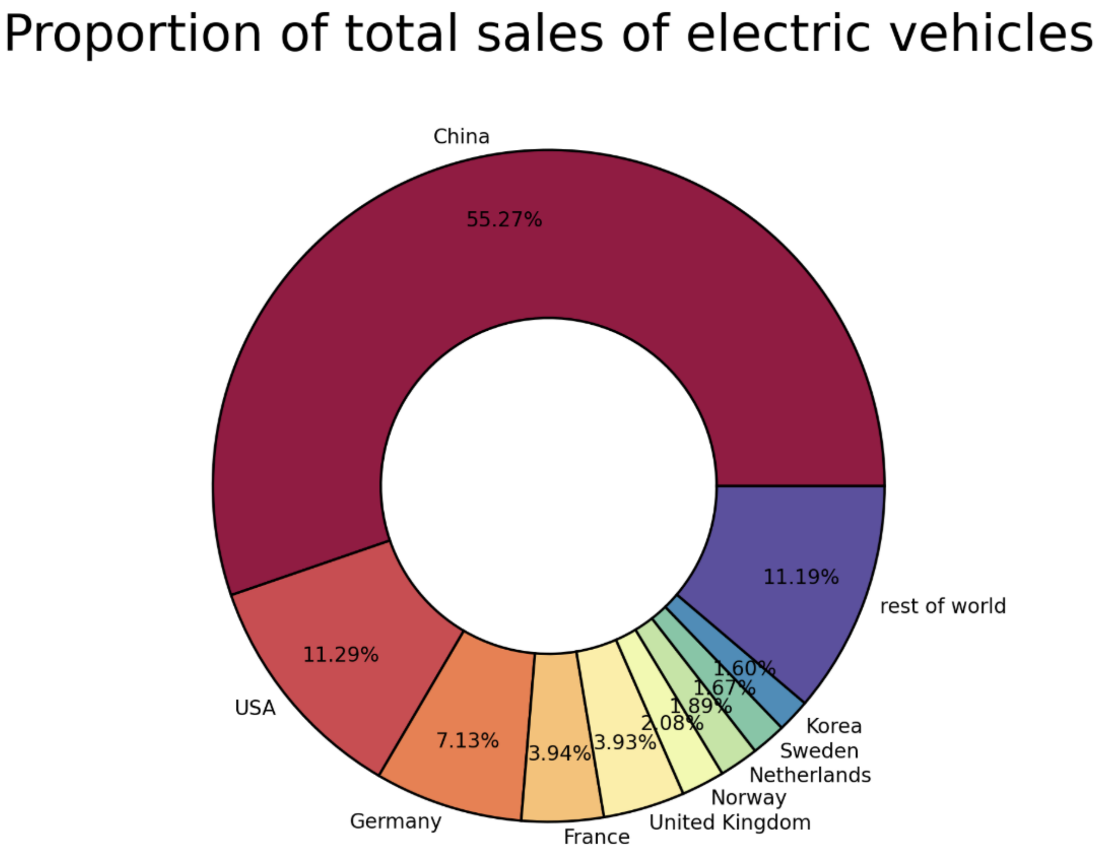

Figure 4.22 Chart about the proportion of total sales of electric vehicles

Among the total cumulative sales of electric vehicles over the years, China accounts for 55.27% of the market share, followed by the United States and Germany.

CHAPTER 5

INTRODUCTION

5.1 Summary

This paper presents the trend of the world EV market by applying machine learning techniques in combination with ARIMA/ SARIMA models. The study deals with annual and proportional sales of the various types of EVs and therefore conveys dynamics with regard to the adoption and market segmentation of EVs. Data from 2010 to 2023 will be used for projections on future trends with respect to powertrain technologies, regional market leaders, and categories of vehicles.

Key Findings and Analyses

1. Trend of Total Annual Sales of Electric Vehicles (EVs)

There has been a tremendous improvement in the yearly sales of electric vehicles within 20102023. Improved technological development of the batteries, supportive policy, and increasing demand spurred the sales, especially after 2018. BEVs lead in 2023, having the highest sale in that category at 9,879,515 units, with PHEVs at 4,241,378.

2. Proportional Composition of EV Sales by Vehicle Types

- Categories of Vehicles:

Cars continue to be the dominant EV market share, followed by minivans, buses, and trucks. It depicts that on the journey to electrification, passenger vehicles have greater consumer appeal over their commercial or utilitarian counterparts.

- Powertrain Technologies:

BEVs contribute to more than onethird of total EV sales, while PHEVs stand slightly behind. In contrast, FCEVs remain negligible due to their high cost, challenges in building infrastructure, and limited consumer acceptance.

3. Annual Trends of Various Powertrain Vehicles

- BEV sales have started showing exponential growth, especially post2018, riding on the back of rapid technological improvement, falling battery costs, and strong government incentives.

- While PHEV sales also increased, at a slower pace, they showed steady interest in the hybrid technology.

- FCEVs continue to struggle to secure a reasonable market share, which could be attributed to a number of barriers, such as high production cost and lack of adequate hydrogen refueling infrastructure.

4. Geographic Market Analysis

- Market Leaders:

Except for 20122014, when the United States was leading in EV sales, China has been dominating the global EV market.

- Reasons for China Leadership:

a) A large population base.

b) Strong government policies that include subsidies and mandates.

c) Advances in industrial and battery technologies.

- Market Share:

Cumulatively, China accounts for 55.27% of global EV sales, followed by the United States and Germany.

5. Proportional Sales Analysis

A histogram analysis is performed to show the percent composition of EV sales by powertrain technologies from 2010 to 2023. Although the latter has a constantly growing market share, BEVs are always on top, displaying their dominance. PHEVs have been growing gradually, showing the rise in consumer interest in hybrid solutions.

Methodology

- Machine Learning Models:

Classification and regression models analyzing the vehicle types and powertrain technologies predict market demand based on historical trends.

- ARIMA/SARIMA Models:

These models forecast the annual sales of EVs by detecting and utilizing seasonal and cyclic patterns to ensure precision in the predictions.

ARIMA: This represents the nonseasonal trend of annual sales.

SARIMA: This includes the seasonal variation and gives a closer look at the dynamics of the market.

Conclusion

The integration of machine learning and ARIMA/SARIMA models has shown the potential for effective forecasting of EV market trends. Some of the findings are:

1. Dominance of BEVs and steady growth of PHEVs.

2. China's leading role in global EV adoption, driven by favorable policies and technology advancements.

3. Further exploration of underperforming segments like FCEVs to address adoption barriers.

Conclusive insights from the study will aid manufacturers, policy changes, and other stakeholders in key decisions making positions to ensure further, sustainable growth for the EVs market.

5.2 Future Work

This study, in fact, provides a good platform for understanding the dynamics in the electric vehicle market; however, a few avenues can be explored further to add to these findings:

1. Incorporation of Emerging Data Sources

- Integrate real-time data from emerging sources like IoT-enabled vehicles and charging infrastructure networks, apart from consumer sentiment analysis, to fine-tune the forecasting models.

- Study the impact of imminent trends like autonomous EVs and V2G technologies on the dynamics of this market.

2. Exploration of FCEVs and Alternative Powertrains

- Research the reasons for the limited adoption of FCEVs, considering infrastructure challenges and cost aspects.

- Develop predictive models related to alternative powertrains such as hydrogen fuel cells and next-generation battery chemistries.

3. Regional Market Segmentation

- Extend the regional scope of analysis to cover developing markets, including Africa, South America, and Southeast Asia.

- Analyze the socioeconomic and policy-related factors influencing the adoption of EVs in those regions and outline market-specific strategies.

4. Advanced Machine Learning Techniques

- Use deep learning and ensemble methods to model non-linear relationships and interactions of variables in electric vehicle adoption.

- Combine with reinforcement learning to allow for optimized policy recommendations on incentives and subsidies.