4-2024

Optimizing V2V Energy Exchange using AI

Marwa Alghawi
msa2268@rit.edu

A Thesis submitted In Partial Fulfillment Of the Requirements for the Degree of Master of Science in Electrical Engineering

Title: Optimizing V2V Energy Exchange using AI

Author: Marwa Alghawi

Date: April 2024

Supervised by:
Dr. Jinane Mounsef, Assistant Professor
Department of Electrical Engineering and Computing
Rochester Institute of Technology-Dubai Campus
United Arab Emirates

Marwa Alghawi

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Electrical Engineering
Department of Electrical Engineering and Computing Sciences

Approved By:

Dr. Jinane Mounsef — Date
Thesis Advisor/ Assistant Professor of Electrical Engineering
Department of Electrical Engineering and Computing Sciences

Dr. Abdulla Ismail — Date
Thesis Committee Member / Full Professor of Electrical Engineering
Department of Electrical Engineering and Computing Sciences

Dr. Ioannis Karamitsos — Date
Thesis Committee Member / Assistant Professor of Data Analytics
Department of Graduate Programs and Research

Acknowledgment

First and foremost, I would like to express my humble gratitude to Allah, the Almighty, for all of His blessings.

I would also like to thank my supervisor, Dr. Jinane Mounsef, for giving me her time, sharing with me her knowledge, and providing me with guidance, support, and encouragement along this path.

I am also grateful to the members of the examination committee, Dr. Ioannis Karamitsos, Dr. Abdulla Ismail, for their feedback, comments, and guidance that helped me to improve for thesis and my academic journey.

I owe a debt of gratefulness to my family for their love, encouragement, support, prayers, and patience throughout my academic journey.

Finally, I would like to express my appreciation to everyone who has supported me along the way, including my friends, colleagues, and fellow students. They absolutely made this journey delightful.

Abstract

A key factor in promoting sustainable economies, industries, and society is the use of electric vehicles (EVs). They are an important step toward a greener future since they reduce greenhouse gas emissions, improve air quality, and encourage energy independence. The demand for charging stations (CSs) will undoubtedly rise as the global EV market continues to grow over the next few decades, requiring significant investments from both public and private players.

By creating a cutting-edge AI model, this work tackles the crucial problems of EVs range prediction and charging optimization decisions. The contributions are summarized as the below:

1. Developed AI models to predict estimated range for EVs without explicitly stating the use of data.

2. Implemented a binary decision-making process to determine the necessity of EV charging, enhancing energy management strategies and mitigating range anxiety.

3. Proposed a comprehensive solution incorporating vehicle-to-vehicle (V2V) and grid-to-vehicle (G2V) energy sharing, selecting optimal charging stations (CS1, CS2) or vehicles (V2, V3) based on factors like waiting time, distance, and energy provision.

4. Curated a dataset highlighting essential data variables crucial for optimizing AI models for V2V energy sharing, facilitating the development of sustainable transportation solutions.

The technology, when integrated into the V2V framework, creates a strong foundation for self-governing energy and enables effective energy sharing between EVs. These developments have important ramifications for encouraging the use of EVs, improving customer satisfaction, and furthering sustainable transportation programs.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

To meet the Paris Agreement's goal of limiting global warming to 1.5°C, emissions must be reduced by 45% by 2030 and achieve net-zero emissions by 2050 [1]. Around 75% of all emissions from mobility worldwide are currently produced by cars, trucks, and other road vehicles [2]. As per data from The World Bank, the transportation industry accounts for approximately 64% of the world's oil consumption and contributes to 23% of energy-related CO_2 emissions [3].

In 2020, an L.C.A. study for the British Department of Transport found that EVs reduced pollutants by around 65% when compared to comparable internal combustion vehicles. By 2030, Battery Electric Vehicles (BEV) will reduce greenhouse gas emissions by 76 percent and a potential 81 percent reduction by 2050 based on anticipated improvements in battery production and continued decarbonization of the U.K. electric grid [4].

Currently, customers enjoy a more comprehensive array of options when considering the purchase of an EV. As gasoline prices surge and EV costs decline, the popularity of EVs is on the rise [5]. The biggest auto markets will transition to all-EVs by 2035, offering both a significant economic opportunity and a glimpse of a more environmentally friendly future [6].

Recent research and solutions focus on optimizing EVs energy resources. The Bidirectional Bipolar Junction Transistor, or rB-TRAN, is ideal for use in EVs, EV chargers, and other devices that help us get closer to a net-zero future, applicable where energy efficiency is essential. Two or more traditional power switches are replaceable with a single B-TRAN thanks to its innovative bidirectional architecture [7].

However, one limitation of EVs compared to internal combustion engine (ICE) vehicles is the accessibility of refueling options. While traditional fuel stations are widespread, CSs for EVs are not as prevalent, leading to range anxiety in users, particularly in urban areas [8, 9]. It is possible that the current electric infrastructures were not intended to handle this significant rise in power use. As the global EV market continues to expand over the coming decades, the demand for CSs will inevitably surge, necessitating substantial investments from both governments and private investors along with support of the public [10].

CS availability, especially in residential areas with multiple occupants, presents a significant challenge. Apartments often lack dedicated charging infrastructure, leading to range anxiety and parking concerns. Residents may face uncertainties about finding an available charging spot and the time required for a full charge. Moreover, the scheduling factor adds complexity, as residents may need to coordinate and compete for limited charging resources [11, 12].

In this context, V2V energy sharing emerges as a transformative solution that can offer EV owners flexibility and energy security beyond the reliance on conventional CSs [13]. V2V represents a form of crowdsourcing where those needing energy are the ones making requests, individuals providing energy are the workers, and the task at hand is the request for charging initiated by the energy requester [14]. The most critical stage in V2V energy sharing, where all EV user data analysis, processing, and EV matching, selection, scheduling, and routing take place. Within the framework for V2V energy sharing, modules and their roles are presented, displaying the importance of AI agents for energy usage monitoring, demand prediction, and smart adaption. Under the physical layer of a V2V energy-sharing framework, an AI agent can post a charging request even before the State of Charge (SoC) drops below the crucial limit, which can lessen the range anxiety within the EV users [13].

With the help of cutting-edge AI solutions like the one this study suggests, V2V energy sharing has the potential to completely transform how we view and use energy in the transportation industry. A 2016 poll provides important insights into the potential of this technology by illuminating the perceived benefits and drawbacks of V2V applications. The study's users recognized the many advantages that V2V technologies offer, such as improved cost-effectiveness, efficiency, comfort, and safety. V2V systems facilitate communication between cars, which can optimize traffic flow, minimize congestion, and ultimately improve road safety. Furthermore, by lowering carbon emissions and our reliance on fossil fuels, the smooth sharing of energy amongst vehicles can make a substantial contribution to the sustainability of our transportation networks [15]. As the study made clear, it is imperative to address worries about security, control, and privacy. By putting strong security measures in place and protecting user privacy, we can allay these worries and realize the full potential of V2V energy sharing. V2V energy sharing has the potential to transform our transportation networks and open the door to a more efficient and sustainable future with careful planning, teamwork, and the adoption of cutting-edge AI technologies. In a separate survey conducted in 2014 across the US, UK, and Australia, the advantages of V2V applications were identified as fewer crashes, reduced crash severity, improved emergency response, less traffic congestion, lower vehicle emissions, and lower insurance rates. On the flip side, respondents expressed concerns and disadvantages related to safety implications in the event of equipment or system failures, legal liabilities for drivers or owners, system security vulnerabilities (from hackers), interactions with pedestrians and bicyclists, and the potential for drivers to overly rely on the technology [16]. A study was conducted to develop a matching algorithm for Spatio-Temporal Non-Intrusive Direct V2V Charge Sharing Coordination. This research included a survey of 153 EV owners in the US, revealing that 40% of participants are interested in both receiving and supplying energy in the V2V charge exchange network. Additionally, 72% prefer to go to the supplier's location for V2V activities, while 46% are willing to pay 5$ for supplier EV services. Moreover, 40% would prefer to have 10-30 EVs participating in V2V communities in their respective areas [17]. In terms of cost considerations for EV users, a simulation study demonstrated that integrating both V2G and V2V systems outperforms scenarios involving only V2G and independent energy charging. This integration led to a cost reduction of 9.17% for residential car parking space and 12.58% for shopping center car parking space [18]. It is worth noting that the number of V2V charging modes depends on the participating EVs and the type of charging option available [19].

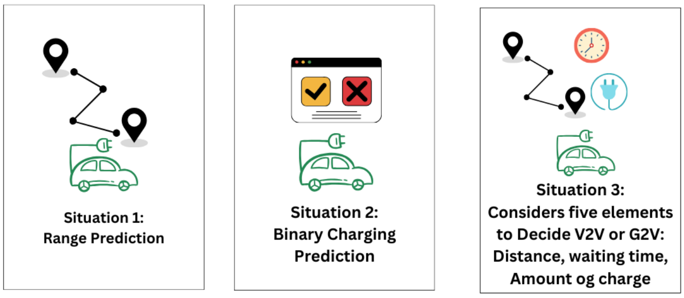


Figure 1.1: Proposed EV model prediction.

The Aim of this thesis is to develop AI models that predict the estimated range for electric vehicles (EVs) and optimize charging decisions to enhance energy management and reduce range anxiety. First, it implements a binary decision-making process to determine the necessity of EV charging, enhancing energy management strategies, and mitigating range anxiety. It then proposes a comprehensive solution incorporating V2V and G2V energy sharing, selecting optimal charging stations (CS1, CS2) or vehicles (V2, V3) based on factors like waiting time, distance, and energy provision. The study curated a dataset highlighting essential data variables crucial for optimizing AI models for V2V energy sharing, facilitating the development of sustainable transportation solutions. These elements are crucial in determining the most effective approach for managing energy distribution and ensuring the efficient operation of EVs. Figure 1.1 represents the proposed solution. These predictions are made internally by the AI agent within the EV, without requiring data from other vehicles or external sources. These decisions serve as the foundation for subsequent actions in a V2V energy sharing framework. It serves the physical layer, where an AI agent can anticipate charging needs based on learned behaviour or post charging requests preemptively to alleviate range anxiety. The related research questions that is driving this study is to explore the key findings, arguments, research methods used, and limitations. Here are the questions you have to look in it:

- Which data variables are essential for optimizing AI models in V2V energy sharing?
- How can the curated dataset facilitate the development of sustainable transportation solutions?
- What are the most critical factors to consider for V2V energy Sharing?
- What methodologies can be employed to enhance the predictive accuracy of AI models for EV range estimation without relying on explicit external data sources?

Through innovations in vehicle-to-vehicle (V2V) and grid-to-vehicle (G2V) energy sharing, the study seeks to contribute to the creation of a more efficient, reliable, and sustainable EV ecosystem.The underlying objectives are to split into:

- Foundation:

1. To create AI models that estimate the range of EVs without explicit data from external sources.
2. To implement a binary decision-making process to determine the necessity of EV charging.

- Broad Impact:

1. To propose a comprehensive solution for V2V (vehicle-to-vehicle) and G2V (grid-to-vehicle) energy sharing, optimizing the selection of charging stations and vehicles based on waiting time, distance, and energy provision.
2. To curate a dataset with essential variables for optimizing AI models in V2V energy sharing, supporting sustainable transportation solutions.

- Implicit Nature:

1. Enhancing the reliability and user confidence in electric vehicles to promote their adoption and support environmental sustainability.

This thesis follows a comprehensive structure to address the research goals and findings in a systematic manner. Beginning with an Introduction, it provides background context, problem statements, research objectives, and an overview of the thesis structure. Subsequently, the Literature Review delves into relevant literature and theoretical frameworks, identifying gaps in current research and advancing the understanding of the subject matter. The Methodology section outlines how the data is processed, transmitted, and analyzed within the machine learning algorithm, including model architectures tailored for different scenarios, data collection methods, and limitations. Following this, the Data section elaborates on the data collection process, driving patterns, parameters included, and any constraints encountered. Model Implementation and Evaluation are then discussed, covering aspects such as dataset splitting, model training, evaluation metrics, and the presentation of results, along with a critical assessment of the model's limitations. Proceeding, navigating Implementation Challenges addressed considering V2V Energy Sharing Framework and AI implementation in V2V energy sharing. The thesis also explores avenues for Future Work and Publications, identifying potential research directions and showcasing resultant publications. Finally, the Conclusion encapsulates key findings.

Chapter 2

Literature Review

An overview of significant studies in the fields of forecasting the range of EVs based on various input features to determine the likelihood of a car requiring charging, thereby aiding in efficient charging EV management and finally, a localized final decision introduces a decision-making component where the model must choose the optimal charging station for the EV based on factors like proximity, waiting time, and available energy. This chapter explores the key findings, arguments, research methods, and limitations presented in the current research and market status for the same concerning topic of "EV distance prediction", "V2V and EV Matching".

Efficient energy utilization among EVs is required for better battery management. Paper [20] presents Peer-to-Peer Car Charging (P2C2), which enables EVs to share charge while in motion through cloud-based coordination. This solution targets individual EV-focused predictive analysis and optimization to alleviate range anxiety and optimize charging behavior by introducing Mobile Charging Stations (MoCS), which are high-battery-capacity vehicles used to replenish the overall charge in a vehicle network using a simulator based, developed by using a simulation framework and performed a set of quantitative analysis. SUMO (Simulation of Urban Mobility)12 is a traffic simulator to support peer-to-peer BEV charging on-the-go, MoCS, and MoCS hubs. It also manages the charging behavior of BEVs based on their battery levels, mobility, and interactions with charging providers. [21] introduces distributed heuristic algorithms for V2V charge sharing, aiming to optimize the matching between energy providers and requesters while minimizing overhead and waiting time. The focus is on enhancing the efficiency of energy exchange networks among EVs. The paper [22] proposes Multi-Agent Reinforcement Learning (MARL) approaches for EV charging coordination with V2V energy exchange, primarily focusing on the effective management of EV charging stations from a macroscopic perspective, considering the overall energy balance and optimization of charging processes across multiple EVs within the charging station. It does not directly address individual EV needs, such as estimating distance or deciding whether charging is required. Instead, it operates at a higher level, optimizing the charging and discharging processes to minimize energy costs and ensure efficient charging station operation. [23] presents a data-driven matching protocol for managing V2V energy exchange to address environmental concerns and alleviate charging overload issues in power systems. The protocol utilizes a deep reinforcement learning (DRL) approach to learn the long-term rewards of matching actions based on a formulated Markov decision process (MDP) offline. Additionally, to protect the privacy of EV owners, a federated learning framework is proposed, enabling collaboration between EV aggregators without sharing sensitive information. An optimization model is established and converted into a bipartite graph problem at the online matching stage for efficient computation. The work in [24, 25] focus on managing energy between EVs using V2V energy sharing, optimizing system-wide energy allocation and matching. In [26], the authors utilize machine learning techniques to predict energy demand for EV networks for G2V systems by proposing an Energy Demand Learning (EDL) algorithm, where a Charging Station Provider (CSP) gathers information from all CSs to predict energy demand for the area. Further, Federated Learning (FL) can be utilized to provide security and privacy insurance. In the meantime, [27] emphasizes comprehending how energy is distributed and consumed in EVs in a smart grid setting. By maximizing energy transfer to charging stations based on consumption patterns, it seeks to provide insights for planning EV CSs and sustainable energy policy usage.

The work in [28] proposes new machine learning techniques for predicting energy demand in EV networks. They introduce Energy Demand Learning (EDL) and Federated Energy Demand Learning (FEDL) approaches to enhance prediction accuracy while addressing privacy concerns and reducing communication overhead. The proposed approaches rely on data collected from multiple CVs in the EV network to improve energy demand prediction accuracy. Both references [26, 28] involve machine learning techniques for predicting energy demand in EV networks. They also explore the application of federated learning to leverage distributed data sources while preserving privacy. However, while the work in [26] focuses on predicting energy demand for EV networks using federated energy demand learning, the work in [28] applies federated learning to probabilistic prediction models for EV driving range. Both approaches aim to improve prediction accuracy and efficiency by leveraging data from distributed sources within the EV network.

Predicting a BEV's remaining range and the minimal charge needed to finish a trip safely is the primary emphasis of [29]. It uses recurrent neural networks (RNNs) to estimate, and even with daily route variations, it exhibits good accuracy. On the other hand, the work in [30] investigates the application of machine learning-based regression models for EV range prediction, namely ensemble stacking generalization and linear regression. It draws attention to the difficulties in locating publically accessible datasets and demonstrates how appropriate linear regression is for precise range estimation.

The work in [31] and [34] concentrate on using deep reinforcement learning (DRL) algorithms for energy management system (EMS) optimization for effective energy distribution among groups of prosumers participating in V2V energy trading. These studies address price mechanisms and privacy preservation issues with unique approaches, such as federated learning frameworks and Markov Decision Processes (MDP). The potential of V2G and V2V technologies to improve grid flexibility and lessen environmental effects [32] and [33]. In order to enable resource utilization and job offloading over vehicle networks, they develop data-driven matching protocols and collaborative fog computing systems, which eventually improve system efficiency and service throughput. Moreover, the work in [35] emphasizes the use of EVs to create greener transportation systems and suggests methods to maximize energy trading volume and societal welfare by using double-sided auctions and fog-based architectures to coordinate V2V traffic. In order to encourage the widespread adoption of EVs, these publications highlight the significance of match-maximization tactics and effective resource allocation systems. Finally, the work in [36] and [37] use DRL-based optimization techniques and data-driven matching algorithms to tackle the problems associated with V2V energy management. In order to improve V2V energy trading and raise local energy self-sufficiency, they emphasize how important it is to take long-term benefits, privacy protection, and computing efficiency into account while designing matching algorithms. By raising the bar for energy management and optimization in the context of plug-in vehicles EVs and P2P trading systems, these articles collectively advance the state-of-the-art and open the door to future developments in more efficient and sustainable energy solutions.

Table 2.2 discusses the below points:

- System Data: A broader information set characterizes the environment in which the proposed adaptive task offloading framework operates. Including world vehicular environment data and information about the physical environment where vehicles operate, including road conditions, traffic patterns, and infrastructure layout. Vehicular communication network data: Data on the communication capabilities and behaviors of vehicles, such as transmission range, signal strengths, and connectivity patterns. Task generation data: Information about generating computational tasks within the vehicular environment, including task sizes, frequencies, and types. Vehicle characteristics and states: Data about individual vehicles participating in the network, such as their positions, speeds, energy levels, computational capacities, and current task loads. Environmental factors: Additional data related to external factors may influence task offloading decisions, such as weather conditions, time of day, and geographic location.

- ML Utilization or Other Algorithms: Indicates whether the paper incorporates machine learning, artificial intelligence, or other algorithms to develop its solution or relies solely on traditional algorithms or simulations.

- Privacy Concerns: Reflects whether the paper addresses privacy concerns associated with EV users' data and interactions within the proposed framework, including implementing measures to protect users' privacy.

- V2V Framework: Specifies the involvement of the proposed framework within the V2V communication layer, distinguishing whether it operates primarily at the physical layer for local data exchange or involves higher-level communication network protocols for matchmaking and coordination.

Table 2.1: Studies on EV Range Prediction.

Aspect	**Reference [29]**	**Reference [30]**
Prediction of EV Range	Proposes a novel approach using recurrent neural networks (RNNs) to predict remaining range and minimum charge required to complete a trip safely.	Explores different machine learning algorithms to improve existing methods for electric range (eRange) estimation.
Use of Local Data	Utilizes historical drive cycle data collected from the vehicle to train predictive models.	Utilizes publicly available datasets to train predictive models.

Table 2.2: Studies on V2V Energy Sharing.

Paper	System Data	Using ML/Algorithms	Privacy Preservation	V2V Layer
[31]	Local data, where historical drive cycle data collected from the vehicles within the cluster to train predictive models.	Multi Double Deep Q-Network (DDQN) agents based DRL algorithm is implemented	N/A	Physical Layer
[32]	Characteristics and dynamics of the vehicular network, communication infrastructure, and task generation mechanisms.	Adaptive Learning-based Task Offloading (ALTO) algorithm, which is an extended version of the multi-armed bandit algorithm	N/A	Application Layer
[33]	Primarily obtained locally, further supported by communication and coordination among vehicles facilitated by the simulation environment to enable the V2V energy exchange process	double-sided auction mechanism for matching EVs	N/A	Network and communication layer
[34]	The paper utilizes local data from EVs to understand their behaviors, preferences, and charging patterns.	The paper utilizes Deep Reinforcement Learning (DRL) methods to evaluate the long-term reward of V2V matching actions based on the Markov decision process (MDP).	Federated learning framework is employed to protect the privacy of EVs, where different aggregators collaborate without exchanging private datasets.	Network and communication layers
[35]	includes EVs, meeting points (M Ps), and a central control management platform.	game theory, specifically the Gale-Shapley matching game, to match consumers and providers	N/A	Physical and Communication Layer
[36]	The system model comprises EVs, smart houses, charging stations, power/communication infrastructures, and a data control center.		max-weight V2V matching algorithm	N/A, Communication Layer
[37]	The proposed V2V charge sharing system involves a centralized server and mobile apps used by EV drivers to communicate with the server via cellular infrastructure.	algorithms for matching demander EVs to suppliers for charging, considering various factors.	N/A	Physical Layer

Table 2.3: Model Data utilization in Different Papers

Paper	Model Data	Model Parameters
[31]	online-available datasets	required energy loads, solar energy generation, EV status signals, and energy consumption during driving cycles
[32]	historical data, simulations, or experiments designed to replicate real-world scenarios.	energy demand, charging rates, battery capacities, energy trading strategies, and V2V communication protocols
[33]	combination of real data, assumptions, simulation environment	number of available EVs, unit of charge, EVs' speed, available power for supply energy consumption, and top-up power needed.
[34]	Simulation data is used to validate and test the proposed V2V matching protocol	departure time, arrival time, daily mileage, and charging requirements
[35]	Real-life dataset based on commercially available EVs	car model, battery capacity, energy consumption rate, and charging time.
[36]	real-life dataset based on commercially available EVs	Car model, battery capacity, energy consumption rate, and charging time
[37]	trip-based probabilistic EV charging behavior model and a survey for incentives and interest	energy consumption, grid load, and supplier EV battery cycling

1 Advanced State of Art

Uncertainty is defined as the inability to fully understand all options or the ramifications of each, which complicates scenario interpretation and decision-making. AI and other intelligent technologies can help human decision makers with predictive analytics in two ways: (1) by identifying relationships among numerous factors, they can help human decision makers gather and act upon new sets of information more effectively; and (2) by generating new ideas through probability and data-driven statistical inference approaches. Generating fresh data and forecasts about consumers, assets, and operations is one of predictive analytics' main purposes. AI technologies and humans can work together to handle various decision-making tasks. AI is probably in a good position to handle difficult problems by applying analytical methods [38]. This brings us to this solution and how this solution addresses two key challenges in the realm of EVs: mitigating Range Anxiety among users and enhancing the flexibility of charging infrastructure. It consists of three core elements: improving human comfort by alleviating range anxiety, implementing V2V energy sharing to optimize EV user experience and reduce strain on charging stations, and promoting the transition from ICE vehicles to EVs to encourage environmental sustainability. Overall, it offers a comprehensive approach to improving the EV user experience and contributing to environmental conservation. In this section, we compare our proposed Situations (Situation 1, Situation 2, and Situation 3) with existing research, particularly papers [29, 30] and [34]. We highlight the key differences and advancements provided by our solutions.

1.1 Situations 1 and 2

Our Situation 1 and Situation 2 leverage multiple machine learning algorithms to predict the range of EVs using historical data obtained from EV users' daily commutes. This approach ensures the privacy of users by not explicitly accessing their individual data, similar to the approach in [29].

In contrast, the work in [30] utilizes publicly available datasets to train predictive models, which may raise concerns regarding data privacy. Additionally, both papers [29, 30] focus solely on predicting range without considering the need for charging or not during a trip.

1.2 Situation 3

Our Solution 3 expands beyond range prediction to incorporate charging decisions, offering a comprehensive approach to energy management for EVs. We compare this solution with a relevant paper that shares similar concepts but does not fully address charging decision-making.

For example, the work in [34] encompasses various aspects of EV behavior, preferences, and charging patterns, utilizing DRL methods to evaluate V2V matching actions based on the MDP. While the work in [34] addresses some aspects of charging decision-making, it may lack the depth and comprehensiveness of Solution 3, which integrates both range prediction and charging decisions seamlessly into the decision-making process, that works as a feedback to the AI, enhancing the decision making.

Chapter 3

Methodology

This chapter describes how the data is processed, transmitted, and analyzed within the machine learning algorithm, influencing its efficiency and effectiveness in solving problems [39]. In simple terms, the choice of a machine learning algorithm and the basic structure or design of the machine learning model together define what is known as the model architecture [40]. Figure 3.1 illustrates how each methodology in the specified Situation uses data inputs to develop the model architecture.

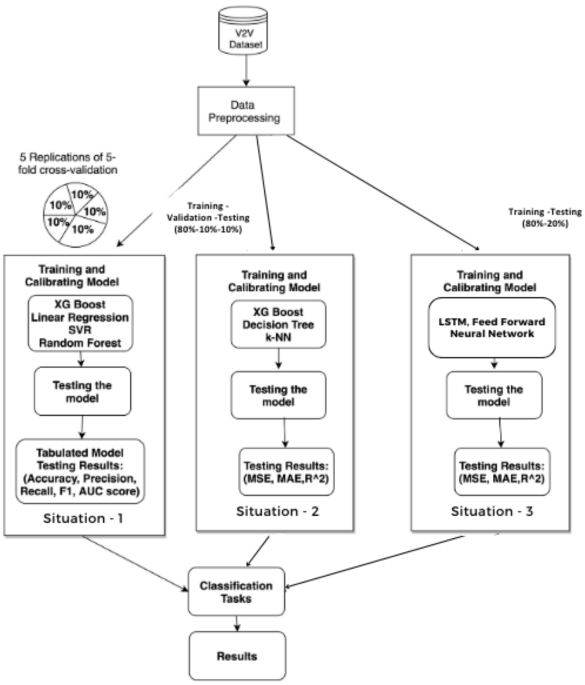


Figure 3.1: Methodology for Situations 1, 2 and 3.

Figure 3.2 shows that the models perform three main functions, the first of which is the Estimated Range task. This task uses AI models to predict the driving range in kilometers for EVs, helping drivers plan their trips more efficiently. Second, Decision 1 provides a binary output (1 or 0) to indicate whether a car needs charging within the day, which helps in energy management and alleviates range anxiety for EV owners. Third, the Comprehensive Solution task incorporates V2V and G2V energy sharing strategies. Considering factors like waiting time, distance, and energy provision, the model recommends the best charging approach, choosing between charging stations (CS1, CS2) or vehicles (V2, V3), to support sustainable transportation.

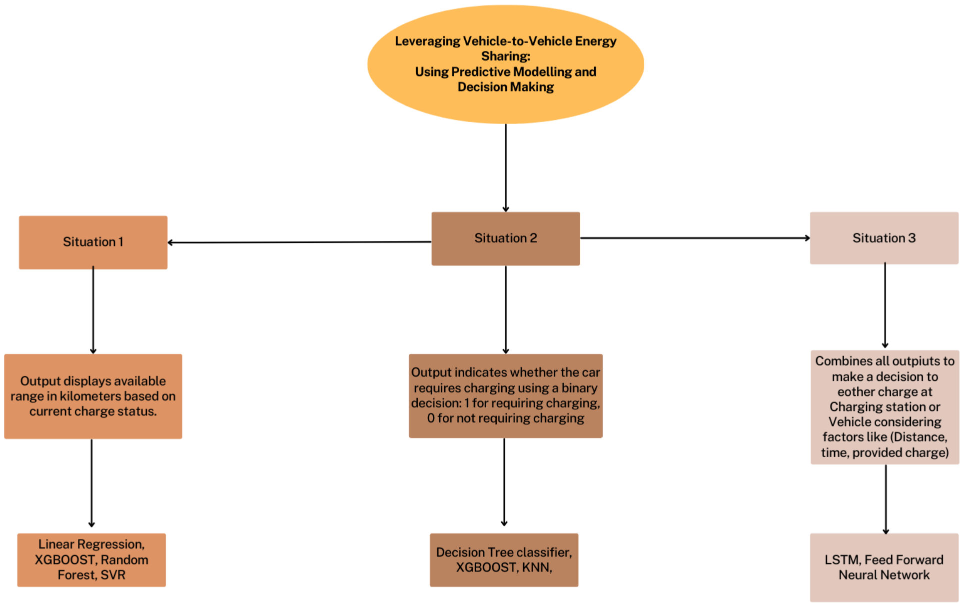


Figure 3.2: Model Situation 1,2 and 3 Architecture.

Additionally, the rationale behind the chosen architecture and its alignment with the specific requirements of V2V optimization tasks is discussed, shedding light on how the model addresses the complexities inherent in the problem domain.

1 Model Architecture - Situation 1

Situation 1 is considered a regression problem type. The following machine learning models are considered for this problem:

- Linear regression is one of the most straightforward and widely employed machine learning algorithms. Essentially, this approach models the connections between dependent and independent variables, progressing from analysis and learning to current training outcomes [51].
- XGBoost regressor and gradient-boosting decision trees are considered and used by data scientists and researchers to optimize their machine learning models.
- Random Forest captures non-linear relationships and interactions between features, which may be beneficial if the relationship between available energy and distance is not strictly linear.

- Support Vector Regression (SVR) is also implemented to capture non-linear patterns in the data, and the kernel trick allows it to model complex relationships. SVR is adequate in high-dimensional spaces and is particularly useful when dealing with non-linear relationships. It uses a kernel trick to transform the input features into a higher-dimensional space.

2 Model Architecture - Situation 2

In Situation 2, the expected outcome is binary: 1, which indicates a need to charge, or 0, which indicates no need to charge. The chosen models are the Decision Tree classifier, XGBOOST, and KNN. Decision trees are intuitive and can capture non-linear relationships. They are well-suited for binary classification tasks and can handle interactions between features. Decision trees are also interpretable, making it easier to understand the decision-making process. KNN is a simple and effective algorithm for classification. It classifies instances based on the majority class among their k-nearest neighbors. KNN is non-parametric and can adapt to different data distributions. It might be effective when the decision boundary is non-linear.

3 Model Architecture - Situation 3

In Situation 3, the methodology involves addressing multiple decisions handled by a separate neural network model. Situation 3 is inherently more complex comparing to Situation 1 and 2, involving more intricate patterns and temporal dependencies in the data. To accurately capture and process this complexity, advanced models are necessary. Models are selected based on below:

- LSTM Model: The Long Short-Term Memory (LSTM) model is designed to handle sequential data, making it suitable for tasks involving time series or sequence prediction. The architecture comprises several layers: the input layer accepts the input data, formatted as sequences with a single feature. This input is then fed into the LSTM layer, which consists of 50 LSTM units. These LSTM units can retain information over long sequences, making them practical for capturing temporal dependencies in the data. The LSTM layer processes the sequential input data and extracts relevant features. Following the LSTM layer, a dense layer with 32 neurons and a ReLU activation function is added to introduce non-linearity and extract features from the LSTM output. Finally, the output layer consists of 5 neurons, corresponding to the five output variables of the model. This layer produces the model's predictions for these variables.

- FFNN: Unlike the LSTM model, the Forward Neural Network (FFNN) operates on tabular data without considering any sequential information. The architecture of the FFNN is as follows: the input layer accepts the tabular input data directly without considering sequence or time-related information. This input is then passed through two dense layers. The first dense layer consists of 64 neurons with a ReLU activation function. This layer is responsible for capturing complex patterns and relationships present in the input data. The second dense layer follows, comprising 32 neurons with ReLU activation function. This layer further refines the features extracted by the previous layer. Like the LSTM model, the output layer of the

FFNN consists of 5 neurons, corresponding to the five output variables of the model. This layer produces the model's predictions for these variables.

The Scikit-learn (Sklearn) library is employed, recognized as one of the most essential and robust libraries for machine learning in Python. It offers a wide range of effective tools for machine learning and statistical modeling, such as classification, regression, clustering, and dimensionality reduction, all accessible through a uniform interface in Python [52].

Chapter 4

Data

The dataset marks the initial phase of constructing a machine learning model. It constitutes an M×N matrix, where M denotes the columns (features), and N represents the rows (samples). Excel spreadsheet program is utilized to generate and organize the data [41].

A bespoke dataset was meticulously constructed for this study in response to the absence of a standardized V2V data frame containing detailed specifications and travel-related information. This section outlines the methodology and parameters involved in creating the custom dataset, emphasizing the inclusion of V2V charge transfer scenarios and leveraging legitimate car specifications from authoritative sources. Table 2.3 compares further the utilized data in different papers discussing the same concerns.

1 Data Collection

Comprehensive research on various car models sourced from reputable automotive websites, ensuring the legitimacy of technical specifications such as make, model, range, and kWh, are considered from [42]. Additionally, V2V charge transfer scenarios were incorporated based on established standards and protocols from recognized industry sources.

2 Driving patterns

The worldwide average daily driving distance for passenger cars, as the study indicates, ranges from 25 to 50 kilometers. However, it is crucial to recognize the substantial geographic variations in driving patterns. For instance, European countries may exhibit lower average distances [43, 44] compared to regions like the United States [45] or the Middle East, where commuting can reach up to 139km [47]. These differences include personal commuting habits, urban planning, and cultural factors. In Dubai, where commuting distances are notably higher, personal driving habits such as the observed kilometers in daily commuting can exceed the global average. The validation and reliability of the dataset were ensured through rigorous checks and referencing real-world experiences. The dataset was meticulously crafted based on thorough research and real-world references to closely replicate real-life EV situations, ensuring its accuracy and relevance.

3 Parameters Included

Input Parameters

The self-created V2V data frame comprises the following key parameters using the program Excel:

- Car Model: It is extracted from the database with a total number of 360 car models [42].

- Range: It is based on the car model specifications in km [42].

- Electric Energy: It is based on the car model specifications in kWh [42].

- Daily commute (km): It is selected for each car model randomly between 6-139 km. It reflects a diverse range of driving distances from Worldwide Daily Driving Distance, by insights from personal experiences and known commuting patterns [43, 45].

- Available Charge: It is determined as 20% of the total electric energy capacity of the car model, assuming that vehicles are within the range anxiety zone. Charging is optimal when the battery level is between 20% and 30%, as it helps avoid deep discharge damage in lithium-ion batteries and ensures efficient recharging [46]. Scenarios where cars often fall into this charge percentage include car owners without home charging facilities, long highway journeys, and unexpected daily commutes.

- Estimated Distance or Output: It is the amount of range according to the available charge of each car model. It is calculated based on a simple mathematical Formula of ratio and proportions, where:

$$a : b :: c : d \tag{4.1}$$

$$range : kWh :: x : kWh(available) \tag{4.2}$$

- SC-1, SC-2 km: They represent the allocation of the distance to each car's first CS option. They vary significantly, ranging from a minimum of 5.6 km to a maximum of 275 km. The minimum distance is based on data indicating that, on an average day, whether working or non-working, individuals in Greece typically travel only 5.6 km. Conversely, the maximum distance considered is 275 km, which reflects the spacing of CSs in remote areas of Canada, accommodating the range capacities of currently available EVs [42][47].

- V2, V3 km: It is the distance traveled to reach or find an EV in km. The range is between 0.4 to 10 km to travel to another vehicle for charge exchange.

- SC1 [Level 2 - km/30 min]: The CS power capacity is level 2, utilized in most standard CSs. The time considered for charging is 30 minutes, providing 20.12 capacity [48].

- SC2 [Level 3 - km/30 min]: The CS power capacity is level 3, utilized also in most standard CSs as fast charger. The time considered for charging is 30 minutes, providing 72.42 capacity [48].

- V2: It represents the power capacity that can be shared from Vehicle 2, considering V2 as the GMC 2022 Hummer model [42].

- V3: It represents the power capacity that can be shared from Vehicle 3, considering V3 as the LUCID DREAM Air performac 19 [42].

- W.T for SC 1,SC 2: It is important to note that these values may vary depending on the geographical location, population, and charging infrastructure. The worst-case scenario is waiting for a CS in the queue, which could last for months. The considered scenario is big cities with more CS infrastructure available. London has an average of 2 hours and 37 minutes for each car while the minimum waiting time was considered at the Outer Hebrides in Scotland due to the less amount of EVs and more available CSs, with a typical waiting time of around 1 hour and 40 minutes [47].

- W.T V2, V3: It is the time considered in minutes until V1 finds V2 or V3. It is calculated based on a simple physics formula:

$$Speed = Distance/Time \tag{4.3}$$

$$Time = Distance/Speed \tag{4.4}$$

  If we have distance and speed, we can figure out the time. Constant value of speed is considered 60km/hour, assuming the car is driving at this speed limit.

- SC1 +: It is the amount of km distance added if CS1 is selected.

- SC2 +: It is the amount of km distance added if CS2 is selected.

- V2+: It is the amount of km distance added if V2 is selected.

- V3+: It is the amount of km distance added if V3 is selected.

Output Parameters

- Solution 1: Estimated Distance or Output km is the amount of range according to the available charge of each car model.

- Solution 2: Decision 1 is a fundamental binary decision determining if a car requires charging. The binary output is either 1, indicating the need for charging, or 0, signifying no charging is necessary. It also identifies urgent use cases, prioritizing situations where the remaining range is below 10 km, in addition to cases where it is below zero.

- Solution 3 : It includes several indices that guide decision-making:
  - INDEX 1: Prioritizes the charging station that requires the shortest driving distance.
  - INDEX 2: Selects the charging option that offers the maximum energy within a 30-minute period.
  - INDEX 3: Chooses the charging source with the shortest waiting time in the queue.

– INDEX 4: Considers the previous outputs (SC1+, SC2+, V2+, V3+) and selects the option that provides the maximum additional charge to the vehicle.

– INDEX 5: Represents the final decision, choosing among the four charging options based on the highest frequency of previous selections (CS1, CS2, V2, V3).

4 Limitations

The intent is to construct a comprehensive EV for range prediction and a charging decision dataset, considering crucial parameters such as car specifications, driving patterns, and environmental influences. Notably, this dataset comprises simulated data rather than real-life observations, serving as a valuable resource for academic research and model development. With a size of 7 parameters for each of the 360 car models, this dataset will provide a robust foundation for analyzing key determinants of EV range and refining predictive models. Additionally, it will serve as a benchmark for evaluating prediction techniques and guiding policy decisions to advance sustainable transportation initiatives. This dataset has been publicly available on the Github platform to facilitate its use and inspire further research endeavors. Researchers and practitioners are encouraged to access and utilize this dataset for various applications, leveraging its insights and potential to drive innovation in the field of EVs and related domains [49].

Chapter 5

Model Implementation and Evaluation

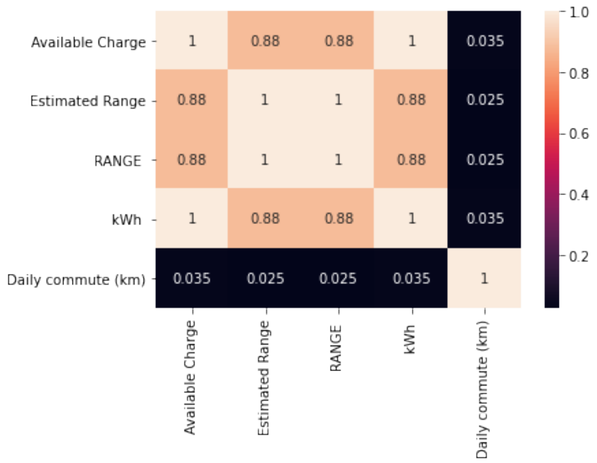
The model implementation and evaluation are integral components of the model execution and assessment phase. The methodology employed is illustrated in Figure 3.1. Initially, before designing the model architecture, Exploratory Data Analysis (EDA) is conducted to derive insights from the data beyond formal modeling or hypothesis testing. This process helps in understanding the dataset variables and their interactions more thoroughly [50].

Beginning with Situation 1, the task involves predicting the remaining range in kilometers for the car after a daily commute. This provides the driver with a primary estimate of the remaining distance they can travel based on their daily commuting patterns.

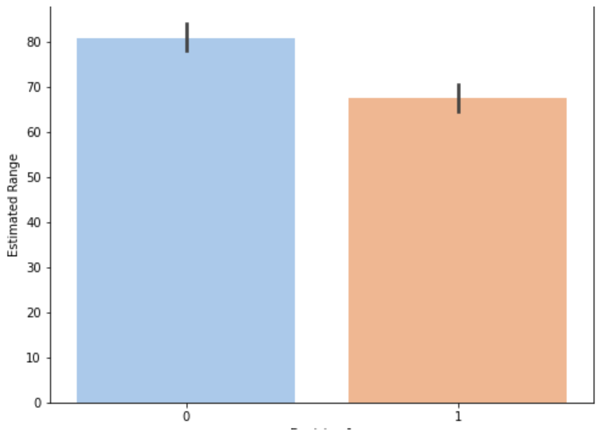
As shown in Figure 5.1a, the attention was directed towards scrutinising the correlation between all inputs and the output for Situation 1. A heat map analysis of the dataset for Situation 1 reveals significant correlations between various input parameters and the estimated range output. Strong positive correlations are observed between estimated range and both the kWh parameter and available charge, highlighting the influence of electric energy levels on range estimation. Conversely, the daily commute parameter shows a weaker correlation, suggesting a lesser impact on the estimated range compared to energy-related factors. Additionally, the perfect correlation between the range parameter and the estimated range indicates a precise alignment between specified and calculated range values.

The bar graph in Figure 5.1b shows the distribution of cars based on their charging needs, categorized as either requiring charging (labeled as 1) or not requiring charging (labeled as 0) as part of Situation 2. This visualization provides insights into the frequency of charging needs within the dataset. Given that Situation 2 is a binary categorical variable, visualizations like scatter plots or heatmaps would not effectively display its relationship with the Estimated Range. Consequently, a bar graph is used for its clarity and effectiveness in illustrating the prevalence of charging requirements among the cars.

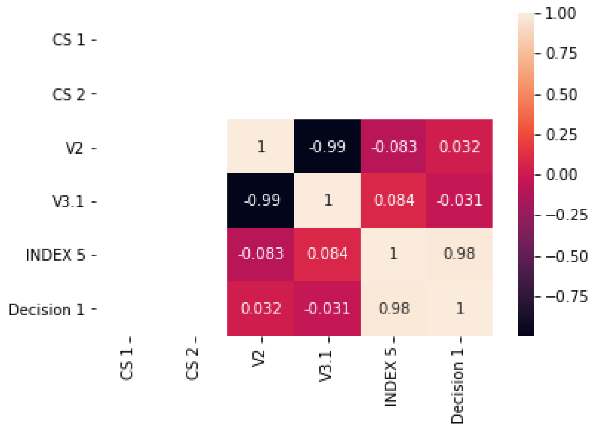
For Situation 3, Figure 5.1c presents a correlation heatmap that illustrates the correlation coefficients among various features. Notably, Decision 1 demonstrates a strong positive correlation with other features, evidenced by a coefficient of 0.98. In contrast, V2 indicates a weak negative correlation (-0.083), and V3.1 shows a slight positive correlation (0.084). Both CS1 and CS2 are constant features and therefore do not exhibit any correlation with other features.



(a) Heat Map - Situation 1



(b) Bar chart - Situation 2



(c) Heat Map - Situation 3

Figure 5.1: Exploratory Data Analysis.

1 Dataset Splitting

1.1 Train-Test Split

In machine learning, data splitting is a widely employed technique that involves dividing a dataset into distinct subsets, typically for training, testing, and sometimes validation purposes. This method is instrumental in identifying optimal model hyper parameters and estimating the model's ability to generalize to new, unseen data [56, 55].

In machine learning model development, ensuring that a trained model generalizes to new, unseen data is crucial. To facilitate this, the dataset is typically divided into two subsets through a process known as data splitting, resulting in a train-test split. Generally, the larger subset, about 80% of the data, is used for training, while the smaller subset serves as the testing/validation set.

For Situations 1 and 2, the approach involves an 80-10-10 split, where 80% of the data is used for training, 10% for validation, and the remaining 10% for testing. In contrast, for Situation 3, the data is split differently, with 80% allocated for training and 20% for testing.

1.2 Model Training

When training a machine learning model, the first step is to select the best hyperparameters for the algorithm to learn the optimal parameters. These optimal parameters are used to correctly map the input features (independent variables) to the labels or targets (dependent variable) in order to achieve a certain level of intelligence [57]. Hyperparameters are significant for machine learning algorithms since they directly control the behaviors of training algorithms and have a significant effect on the performance of machine learning models. [58] For both Situations 1 and 2, the critical parameter model is highlighted. Aside from that, these critical parameters are based on hyperparameter tuning, which is a crucial step in the process of training machine learning models. The tuning technique used for Situations 1 and 2 is Grid Search, a systematic approach that evaluates the model's performance across a predefined grid of hyperparameter values.
In Situation 3, hyperparameter tuning is conducted using a Keras Tuner. This library optimizes hyperparameters for TensorFlow neural networks by identifying the most effective settings. The model utilizing the Keras Tuner undergoes adjustments to its hyperparameters, making it a hyper-tuned model [59].

For situation 1, an N-fold cross-validation technique is commonly employed to optimize the utilization of available data. This method involves partitioning the dataset into N folds, typically using 5-fold. The dataset is set up for 5-fold cross-validation, with each fold using 10% of the data for validation and the remaining data for training. This setup allows for thorough model evaluation and training. For Situation 2, where the task involves binary classification, using ROC AUC for model evaluation instead of k-fold cross-validation stems from several considerations. The receiver operator characteristic (ROC) curve is often used in binary outcome analysis to illustrate how well a model or algorithm performs [13]. Unlike k-fold cross-validation, which may not be directly applicable to binary classification tasks, ROC AUC is specifically designed to assess the discriminatory power of binary classifiers.

1.3 Model Evaluation Metrics

The evaluation of machine learning models involved several key metrics. Due to the nature of Situation 1 and Situation 3 outputs, Absolute Error (MAE) and Mean Squared Error (MSE) were used to quantify the accuracy of predictions, measuring the average difference and squared difference between predicted and actual values, respectively. Additionally, the R-squared score was employed to assess the overall goodness of fit of the models, indicating the proportion of the variance in the dependent variable explained by the independent variables. Furthermore, training and testing scores were considered to evaluate the model's performance on seen and unseen data, providing insights into overfitting or underfitting issues. These metrics provided a comprehensive understanding of the model's predictive capabilities and generalization performance.

In Situation 2, the models are evaluated against binary output, which necessitates the adoption of specific performance metrics tailored to such scenarios. Unlike Situation 1, where metrics like MAE, MSE, and R-squared are commonly employed for regression tasks, Situation 2 involves classification tasks with binary outcomes. In binary classification, predicting the correct class label is pivotal, and metrics such as accuracy, precision, recall, and F1 score gain prominence.

1.4 Validation and Testing: Feature Selection

Feature selection is an essential task in machine learning-related problems because irrelevant features, used as part of the training procedure of different prediction systems, can increase the cost and running time of the system and make its generalization performance much poorer [54].

The techniques used for feature selection are Univariate selection using the chi-squared (chi^2) statistical test and a correlation heat map to illustrate the relationships between features when interdependencies exist among their values. After performing feature selection for Situations 1,2 and 3, the selected features are selected based on the Chi Score for each feature, as shown in Table 5.1.

Table 5.1: Summary of Feature Selection Techniques and Selected Features for Different Models in Various Situations.

Model	**Feature Selection Technique**	**Selected Features**
Situation 1	Chi-Squared, Correlation Heatmap	SN ,RANGE ,kWh ,Available Charge
Situation 2	Chi-Squared, Correlation Heatmap	RANGE, kWh, Daily commute (km), Available Charge, Estimated range
Situation 3	Chi-Squared, Correlation Heatmap	Decision 1, W.T for SC 2, SC-2 km, V2 km, V3 km, SC1 [Level 2 - km/30 min], SC2 [Level 3 - km/30 min], V2, V3, W.T for SC 1

In Situation 1, as demonstrated in 5.2, the 'RANGE', 'Available Charge', and 'kWh' features consistently have the highest importance scores for all models (Linear Regression, XGBoost, Random Forest, and SVR), indicating their significance in predicting the target variable. The 'SN' feature also shows relatively high importance across all models, albeit slightly lower than the aforementioned features. The 'Decision 1' feature has moderate importance in all models except for Random Forest, where it has a relatively higher importance score. The 'Daily commute (km)' feature generally has lower importance scores compared to other features across all models.

In Situation 2, as demonstrated in 5.3, the 'Daily commute (km)' feature has the highest importance score across all models, indicating its significance in predicting the target variable. For XGBoost and Decision Tree models, 'RANGE' and 'Estimated Range' also show notable importance. 'kWh' and 'Available Charge' exhibit moderate importance in most models, while 'Output: km' has relatively lower importance in KNN.

Finally for Situation 3, the feature selection score results are shown in 5.4, where:

- Index 1:

  The features SC1 [Level 2 - km/30 min], SC2 [Level 3 - km/30 min], V2, V3, CS 1, and CS 2 are not included in the feature selection because they do not exist as part of the inputs. The features with NaN chi-squared scores likely indicate that they are constant and do not show significant variability, making it hard for the model to find correlations with the target variable.

- Index 2:

  Only SC1 [Level 2 - km/30 min], SC2 [Level 3 - km/30 min], V2, and V3 are considered inputs, while the rest of the features are not included. Similar to Index 1, NaN values for some features suggest they might be constant and lack variability.

Table 5.2: Feature Selection - Situation 1.

Model	**Feature Score**
Linear Regression	
RANGE	4.040613
Available Charge	1.360677
kWh	1.353712
SN	0.451674
Decision 1	0.051224
Daily commute (km)	0.025044
XGBoost	
RANGE	4.042569
kWh	1.356742
Available Charge	1.337432
SN	0.445424
Decision 1	0.038128
Daily commute (km)	0.027359
Random Forest	
RANGE	4.034815
Available Charge	1.358558
kWh	1.355766
SN	0.443531
Decision 1	0.070825
Daily commute (km)	0.010862
SVR	
RANGE	4.028888
Available Charge	1.364626
kWh	1.359623
SN	0.447285
Decision 1	0.051869
Daily commute (km)	0.017922

- Index 3:

  Features SC1 [Level 2 - km/30 min], SC2 [Level 3 - km/30 min], V2, V3, CS 1, and CS 2 are not included in the feature selection, likely because they are not part of the inputs. The chi-squared scores for W.T V2 and W.T v3 are relatively high, indicating they are important for predicting Index 3.

- Index 4:

  Features SC1 +, SC2 +, V2+, and V3+ are the only considered parameters. Similar to previous indices, NaN values for some features suggest they might be constant and lack variability.

- Index 5:

  Features CS 1 and CS 2 are not considered in the feature selection due to the previous selection/decisions made on the primary "Indices". V2 and V3 play significant roles in predicting Index 5, as indicated by their high chi-squared scores. The final decision, 'Decision 1', is highly correlated with V2 and V3, suggesting that these charging providers are crucial factors in determining the outcome.

Table 5.3: Feature Selection - Situation 2.

Model	**Feature Score**
XGBoost	
Daily commute (km)	0.424092
Estimated Range	0.071514
RANGE	0.048868
Available Charge	0.039488
kWh	0.021221
KNN	
Daily commute (km)	0.419184
Output: km	0.045381
RANGE	0.039185
kWh	0.023200
Available Charge	0.003726
Decision Tree	
Daily commute (km)	0.417476
RANGE	0.084425
Estimated Range	0.060110
kWh	0.051876
Available Charge	0.034930

Overall, the feature selection process identifies relevant features for each target variable, with some features being constant or not included due to their absence in the input data. The selected features reflect the key factors influencing each index, providing insights into their predictive relationships.

As shown in Table 5.5, it is evident that feature selection helps regularization to reduce overfitting or underfitting, which substantially impacts model performance. In Situation 1, before feature selection, models like XGBoost and SVR frequently show good training scores but suffer from overfitting, as seen by substantial differences in training and test scores. Even though scores may drop after feature selection, models such as Random Forest and Linear Regression show better generalization and less tendency toward overfitting, as seen by more minor score differences.

Meanwhile, in Table 5.6, Situation 2 models like KNN and Decision Tree exhibit great accuracy before feature selection but show a trade-off after feature selection. KNN and Decision Tree demonstrate better generalization even with lower scores, as seen by more minor differences in training and test results. This highlights how crucial feature selection is in maintaining robustness against overfitting or underfitting by striking a balance between model complexity and performance for improved regularization.

In Situation 3, Table 5.7 illustrates the performance of two cases, before and after feature selection. Initially, both the LSTM and FFNN models showed comparable performance in terms of MSE and R-squared values, although the FFNN model recorded a slightly higher MAE than the LSTM model. Post feature selection, the LSTM model's performance enhanced notably—demonstrating a significant reduction in MSE and MAE, and an improvement in R-squared value. Conversely, the performance of the FFNN model deteriorated significantly after feature selection, with marked increases in MSE and MAE and a reduction in R-squared value. Therefore, while feature selection markedly benefited the LSTM model, it adversely affected the FFNN model.

Table 5.4: Feature Selection - Situation 3 (LSTM and FFNN).

Target Variable	Feature	Chi-squared score
Index 1	SC-1 km	1.1349
	SC-2 km	0.9050
	V2 km	18.6311
	V3 km	20.1375
Index 2	SC1 [Level 2 - km/30 min]	NaN
	SC2 [Level 3 - km/30 min]	NaN
	V2	NaN
	V3	NaN
Index 3	SC-1 km	0.0159
	SC-2 km	0.2272
	V2 km	0.0211
	V3 km	0.1461
Index 4	SC1 +	NaN
	SC2 +	NaN
	V2+	NaN
	V3+	NaN
Index 5	CS 1	NaN
	CS 2	NaN
	V2	26.1785
	V3.1	31.1816
	Decision 1	126.0

2 Results

In the assessment of models for Situation 1, SVR proved to be the top performer, excelling in MAE, MSE, and R-squared metrics. In Situation 2, XGBoost emerged as the best model, showing vital accuracy and predictive strength. These results highlight the effectiveness of SVR and XGBoost in their respective predictive tasks, indicating their potential for practical application in enhancing V2V energy sharing in the physical layer. For situation 3, the LSTM model exhibited significant performance improvement.

The study addresses the pressing need for efficient energy management in EVs to mitigate range anxiety and promote sustainable transportation practices. Three distinct scenarios were considered, each requiring tailored AI models to address specific challenges.

- Situation 1: Estimated Range Prediction

  For predicting the estimated range of EVs, feature selection played a crucial role in enhancing model performance. After employing Chi-Squared and correlation heatmap techniques, the selected features (RANGE, kWh, Daily commute, Available Charge) provided a refined input set for AI models. Notably, models such as Random Forest and Linear Regression demonstrated improved generalization and reduced overfitting post-feature selection, highlighting the importance of regularization techniques. It is important to note that both linear regression and SVR, scored 1 in R-square. When there is no discrepancy between the anticipated and actual values, the model fits the data exactly, and the R-square value is equal to 1. After employing feature selection, a validation set, k-fold cross-validation the score is still consistent. This can be explained with, in linear regression, If the

Table 5.5: Comparative analysis of model performance in Situation 1 - Before and after feature selection.

Before Feature Selection			
Method	**MSE Test**	**MAE**	**R-squared (R^2)**
XGBOOST	0.04347	0.99992	0.10759
Linear Regression	0.00000	0.00000	1.00000
SVR	0.00209	0.03619	1.00000
Random Forest	0.53401	0.24986	0.99902
After Feature Selection			
XGBOOST	1.02887	0.22977	0.99810
Linear Regression	0.00000	0.00000	1.00000
SVR	0.00200	0.03521	1.00000
Random Forest	0.34531	0.18667	0.99936

Table 5.6: Comparative analysis of model performance in Situation 2 - Before and after feature selection.

Before Feature Selection					
Method	**Accuracy**	**Precision**	**Recall**	**F1**	**AUC Score**
XGBOOST	0.98148	0.95833	1.0	0.97872	0.98597
KNN	0.87037	0.83333	0.86957	0.85106	0.87027
Decision Tree	0.96296	0.92	1.0	0.95833	0.96774
After Feature Selection					
XGBOOST	0.98148	0.95833	1.0	0.97872	0.99158
KNN	0.77778	0.68966	0.86957	0.76923	0.78962
Decision Tree	0.81481	0.70968	0.95652	0.81481	0.83310

input variables have a high correlation, the linear regression will overfit the results. Due to the fact that the data is curated based on mathematical formula, it is foreseen that the input variables have a high correlation [60]. In SVR, Finding the optimal fit that properly predicts the target variable while lowering complexity to prevent overfitting is the aim of support vector machines (SVR). SVR naturally solves multicollinearity by employing the Support Vector Machine (SVM) method to ignore duplicated variables and concentrate on the most essential ones [61]. In summary, multicollinearity and the innate mathematical relationships in the curated data cause linear regression to overfit, while SVR can manage these problems well. Despite SVR's resilience, overfitting can still occur because to the comparatively small size of the dataset, as demonstrated by the flawless R-Squared values in both models.

- Situation 2: Charging Necessity Decision

  Various machine learning algorithms were evaluated to determine whether an EV requires charging within the day. Despite the initial high accuracy observed with models like KNN and Decision Tree, a trade-off was evident after feature selection. However, models showcased improved generalization post-feature selection, suggesting a balanced approach between model complexity and performance for effective regularization.

- Situation 3: Comprehensive Solution with V2V and G2V Energy Sharing

  Integrating V2V and G2V energy-sharing strategies aimed to optimize charging decisions based on waiting time, distance, and energy provision. The feature selection highlighted critical variables (Decision 1, Waiting Time for Charging Sta-

Table 5.7: Comparative analysis of model performance in Situation 3 - Before and after feature selection.

Before Feature Selection			
Method	**MSE Test Data**	**MAE**	**R Squared**
LSTM	0.01236	0.01930	0.55030
FFNN	0.01041	0.03983	0.59283
After Feature Selection			
LSTM	0.00558	0.00908	0.57755
FFNN	0.12636	0.17564	0.19746

tions, Vehicle Distances) essential for model performance. Post feature selection, the LSTM model exhibited significant performance improvement, while the FFNN model showed degradation, emphasizing the effectiveness of feature selection in enhancing model efficacy.

This study underscores the importance of tailored AI models and feature selection techniques in addressing diverse challenges in EV energy management. By leveraging machine learning algorithms and optimizing feature sets, the proposed solutions demonstrate promising results in predicting range, determining charging necessity and optimizing energy-sharing strategies. These findings pave the way for more sustainable and efficient transportation practices, contributing to the transition towards a greener future.

2.1 Limitations of the Model

While predictive models are valuable tools for making informed decisions and predictions based on data, it is essential to recognize that they simplify complex systems and may only partially capture some relevant factors or nuances [62].

Understanding the limitations of a model is essential for interpreting its results correctly and avoiding overreliance on its predictions. By acknowledging and addressing these limitations, researchers and practitioners can make more informed decisions about when and how to use the model effectively.

- Data Source Limitations: The data used to build the model relies on information from reputable automotive websites and established standards and protocols. However, the accuracy and completeness of these sources may vary, which could introduce biases or inaccuracies into the model.

- Geographic Variations: Driving patterns and commuting distances vary significantly across regions and countries. While efforts were made to account for these differences, geographic variations could still impact the model's generalizability, especially in regions with unique commuting habits, such as Dubai.

- Assumptions and Interpretations: The model construction involves making assumptions and interpretations based on available data and research. These assumptions may only partially capture some of the complexities and nuances of real-world driving scenarios, leading to potential oversimplifications or inaccuracies in the model predictions.

- Limited Scope: The model may have a limited scope regarding the factors considered and the scenarios modeled. For example, it may not account for certain variables or external factors influencing electric vehicle performance or charging requirements, such as weather conditions or infrastructure limitations.

- Uncertainty and Future Changes: Like any predictive model, the predictions are uncertain, especially when extrapolating future trends or scenarios. Additionally, the automotive industry is rapidly evolving, with advancements in technology and changes in consumer behavior impacting the relevance and accuracy of the model over time.

Chapter 6

Navigating Implementation Challenges

Within the dynamic landscape of transportation systems, the concept of Vehicle-to-Vehicle (V2V) energy sharing emerges as a transformative force, transcending traditional boundaries to redefine the way vehicles interact and collaborate on the road. Beyond the realms of mere communication, V2V energy sharing operates at the intersection of various layers, encompassing physical, communication, and network domains. This multi-faceted approach brings forth a myriad of challenges and opportunities, as vehicles evolve into interconnected nodes in a vibrant energy ecosystem. In this chapter, first a comprehensive exploration of V2V energy sharing, traversing through the intricacies that span across different layers of architectures is explained followed by the implementation issues and mechanisms with an emphasis on the local environment utilising AI for V2V Energy Sharing.

1 V2V Energy Sharing Framework

Following the published paper that provides an overview of the V2V energy framework [64], the V2V energy sharing framework can be conceptualized as a multi-layered system, with each layer addressing key aspects essential for its successful implementation. The challenges within the V2V energy sharing framework can be summarized as follows:

- Physical layer:
  - Numerous power conversion stages leading to power loss and reduced efficiency.
  - Increase cost associated with external DC-DC converters compared to on-board charging solutions.
  - Durability, power, and safety of batteries used in EVs.
- Communication Layer:
  - Systems like Bluetooth and ZigBee, which have short transmission ranges and poor data rates, are not suitable for dynamic charging.
  - Characteristics required for a real-time-in motion communication system.
  - Creating a successful information and communication system for EV dynamic charging.

- Allowing for efficient V2V connection with little latency and excellent dependability.

- Application and Data Layer:
  - When EV participants are connected, accurate verification of transferred energy and possible security breaches are made.
  - Dependence on reliable and secure communication is essential for dynamic pricing and congestion management in the V2V charge allocation protocol for VANETs.
  - Insufficient security measures in V2V message dissemination systems (digital platforms).
  - Persuading EV vendors to participate in V2V charging.
- Interoperability and Standards Layer:
  - Charging Connectors and Protocols.
  - Payment Systems.
  - Grid Integration.
  - CS Network Accessibility.
- Network Layer:
  - Ensuring data security in the transfer of data energy between participants.
  - Limited network bandwidth and high content demand in V2V communication.
  - Enabling cooperative NOMA in V2V networks.
  - Ensuring network security and privacy in V2V communication.
  - Deployment of 6G Infrastructure.

2 Local Environment Considerations for AI Implementation in V2V energy sharing

Implementing AI-driven V2V energy sharing systems requires careful adaptation to the specific characteristics and constraints of the local environment.

- Data Collection and Management:

  Data is crucial for AI. Assessing local data availability(e.g., vehicle battery levels, driving patterns, traffic conditions, weather data) which can be establishing partnerships for data sharing. Along this, a robust data privacy and management practices shall be in line with local regulations.

- ML Algorithms and Models:

  Customize AI models to reflect local driving patterns and conditions (Machine learning models for range estimation, Optimization algorithms for charging decisions, Importance of real-time processing and decision-making). This can be accomplished in collaboration with local AI researchers and institutions.

- Communication Infrastructure:

  Map out existing communication infrastructure and identify gaps. This is evaluated by the reliability of local communication networks and plans for potential upgrades. Should be in line with with local cybersecurity standards.

- Interoperability and Standards:

  Adoption of international standards while considering local adaptations. Working with local standards organizations to adopt relevant protocols for privacy and security. Not to forget the involvement between local/international vehicle manufacturers (Ensuring compatibility with different vehicle types) and local service providers (Ensuring compatibility with different charging stations).

- Integration with Physical and Network Infrastructure:

  Conduct a thorough assessment of the local charging infrastructure and grid capabilities (Real-time monitoring and control systems). Develop strategies to integrate AI systems. by ensuring AI systems can interact seamlessly with local charging stations and grid infrastructure. Assessing and addressing any specific local challenges such as power outages or grid instability.

- Financial and Incentive Mechanisms:

  Researching existing financial programs and propose new initiatives to support AI deployment in V2V energy exchange. Considering new incentives tailored to local economic conditions and policy targets.

Chapter 7

Future Work and Publications

1 Future Work

The plan involves exploring further optimization techniques for AI models to predict EV range more accurately, which may include considering additional input variables or alternative modeling approaches. Additionally, there's a focus on investigating advanced machine learning algorithms or hybrid models to improve the accuracy and efficiency of determining EV charging necessity.

Furthermore, the aim is to extend the comprehensive solution to fully incorporate G2V energy sharing, tackling challenges such as infrastructure integration and scalability. As part of these future plans, collaboration with EV manufacturers and other stakeholders in the car industry, as well as local authorities with data on EV users and charging station details, is crucial. This collaboration aims to obtain real-world data that can enhance the determination of V2V and G2V charge exchange status.

2 Publications

This work has resulted into two scholarly publications. The first is a journal review paper that provides an overview of V2V energy sharing, presenting the conceptual framework, key findings, and potential applications:

Marwa Alghawi*, Jinane Mounsef, "Overview of Vehicle-to-Vehicle Energy Sharing Infrastructure", IEEE Access, April 2024.

The second publication is an accepted paper at the Artificial Intelligence Applications conference. It focuses on the results of solutions 1 and 2 (EV range prediction and charging necessity determination), showcasing the efficacy of the developed AI models. The intent is also to publish the obtained results of solution 3, comprehensive V2V and G2V energy sharing separately, emphasizing its significance in advancing sustainable transportation solutions:

Marwa Alghawi*, Jinane Mounsef, Ioannis Karamitsos, "Optimizing Vehicle-to-Vehicle Energy Sharing with Predictive Modeling", Artificial Intelligence Applications and Innovations (AIAI), Corfu, June 27-30, 2024.

Additionally, part of the future plan involves conducting a survey among EV users to gather data on their willingness to participate in V2V energy sharing and incorporate

AI into their EV decision-making process. This survey serves as a platform for collecting valuable insights and data from EV owners, which can inform the development and implementation of V2V energy sharing initiatives.

Chapter 8

Conclusion

In this study, we conducted an in-depth analysis of EV range prediction and charging optimization. Our investigation involved reviewing significant literature in the field and proposing a novel AI model to address the challenges associated with EV range estimation and charging decisions. Through our analysis, we identified key methodologies and approaches employed in existing research, including predictive modeling, machine learning algorithms, and coordination strategies for energy exchange among EVs.

We found that SVR demonstrated the best performance in Situation 1, with the lowest MSE and highest R-squared value among the tested models. In Situation 2, XGBoost emerged as the optimal model, exhibiting the highest accuracy and precision compared to other algorithms. For Situation 3, the LSTM model, after feature selection, showed the lowest MSE and highest R-squared value, indicating its effectiveness in predicting EV range and optimizing charging decisions.

The implications of our findings extend to various real-world applications in the automotive industry and sustainable transportation sector. By accurately predicting EV range and optimizing charging decisions, our AI model can significantly enhance the user experience and promote widespread adoption of EVs. Moreover, our model's integration within the V2V and G2V framework establishes a critical infrastructure for autonomous energy management and facilitates efficient energy exchange among EVs. These advancements have the potential to accelerate the transition towards sustainable mobility and reduce greenhouse gas emissions. While our study has made significant contributions to the field of EV range prediction and charging optimization, several avenues for future research and improvements remain unexplored. One potential direction for future work involves refining the accuracy and robustness of our AI model through additional data collection and validation processes. Additionally, exploring advanced machine learning techniques and reinforcement learning algorithms could enhance the adaptability and efficiency of our model in dynamic environments. Overall, ongoing research efforts in this area are crucial for addressing emerging challenges in EV adoption and advancing sustainable transportation initiatives.

Bibliography

[1] ”Net zero coalition.” United Nations. Accessed: Jan. 15, 2023. [Online]. Available: https://www.un.org/en/climatechange/net-zero-coalition

[2] 3 ways of ensuring a sustainable electric vehicle transition. World Economic Forum. (n.d.). https://www.weforum.org/agenda/2023/06/3-things-policy-makers-need-to-collaborate-on-for-sustainable-electric-vehicle-transition-amnc23/

[3] Krishna, G. (2021). Understanding and identifying barriers to electric vehicle adoption through thematic analysis. Transportation Research Interdisciplinary Perspectives, 10, 100364. https://doi.org/10.1016/j.trip.2021.100364

[4] Taub, E. A. (2022, October 19). E.V.s start with a bigger carbon footprint. but that doesn't last. The New York Times. https://www.nytimes.com/2022/10/19/business/electric-vehicles-carbon-footprint-batteries.html

[5] Matulka, R. (n.d.). The history of the Electric Car. Energy.gov. https://www.energy.gov/articles/history-electric-car

[6] How electric vehicles will shape the future. McKinsey Company. (n.d.). https://www.mckinsey.com/featured-insights/themes/how-electric-vehicles-will-shape-the-future

[7] How these emerging tech names are shaping the future of the EV market. Nasdaq. https://www.nasdaq.com/articles/how-these-emerging-tech-names-are-shaping-the-future-of-the-ev-market

[8] P. Mahure, R. K. Keshri, R. Abhyankar, and G. Buja, ”Bidirectional Conductive Charging of Electric Vehicles for V2V Energy Exchange,” in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, pp.2011-2016, 2020, doi: 10.1109/IECON43393.2020.9255386.

[9] E. Bulut and M. C. Kisacikoglu, ”Mitigating Range Anxiety via Vehicle-to-Vehicle Social Charging System,” in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, Sydney, NSW, Australia, pp.1-5, 2017, doi: 10.1109/VTC-Spring.2017.8108288.

[10] Veneri, O., Ferraro, L., Capasso, C., and Iannuzzi, D. (2012b). Charging infrastructures for EV: Overview of technologies and issues. 2012 Electrical Systems for Aircraft, Railway and Ship Propulsion. https://doi.org/10.1109/esars.2012.6387434

[11] Access to EV charging stations in Europe: How do countries compare?. euronews. (n.d.). https://www.euronews.com/next/2023/09/18/access-to-ev-charging-stations-in-europe-is-a-significant-concern-how-do-countries-compare

[12] Sircar, N. (2024, January 29). “range anxiety”, rushing to malls to charge cars: EV owners in UAE list challenges. Khaleej Times. https://www.khaleejtimes.com/uae/range-anxiety-rushing-to-malls-to-charge-cars-ev-owners-in-uae-list-challenges

[13] Shurrab, M., Singh, S., Otrok, H., Mizouni, R., Khadkikar, V., Zeineldin, H. (2022). An efficient vehicle-to-vehicle (V2V) energy sharing framework. IEEE Internet of Things Journal, 9(7), 5315–5328. https://doi.org/10.1109/jiot.2021.3109010

[14] Ibrahim, Y. S. (2022). Crowdsourcing Based Framework for Vehicle-to-Vehicle Power Transfer (thesis). Khalifa University of Science and Technology , Abu Dhabi.

[15] T. Schmidt, R. Philipsen, P. Themann, and M. Ziefle, “Public perception of V2X-technology - evaluation of General Advantages, disadvantages and reasons for data sharing with connected vehicles,” *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, doi:10.1109/ivs.2016.7535565.

[16] B. Schoettle and M. Sivak, “A survey of public opinion about connected vehicles in the U.S., the U.K., and Australia,” *2014 International Conference on Connected Vehicles and Expo (ICCVE)*, 2014, doi:10.1109/iccve.2014.7297637.

[17] Bulut, M. C. Kisacikoglu, and K. Akkaya, “Spatio-temporal non-intrusive direct V2V charge sharing coordination,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9385–9398, 2019, doi:10.1109/tvt.2019.2931954.

[18] A. Khele, C. Jiang, and H. Wang, “Fairness-aware optimization of vehicle-to-vehicle interaction for Smart EV charging coordination,” *2023 IEEE/IAS 59th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, 2023, doi:10.1109/icps57144.2023.10142094.

[19] H. A. Gabbar, ”V2V Charging,” in *Fast charging and resilient transportation infrastructures in Smart Cities*, Cham, Switzerland: Springer, 2022, pp. 256-260.

[20] Chakraborty, P., Parker, R., Hoque, T., Cruz, J., Du, L., Wang, S., Bhunia, S. (2022). Addressing the range anxiety of battery electric vehicles with charging en route. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-08942-2

[21] Chatterjee, P., Majumder, P., Debnath, A., Das, S. K. (2022). Distributed decision making for V2V charge sharing in Intelligent Transportation Systems. 2022 19th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). https://doi.org/10.1109/secon55815.2022.9918576

[22] Fan, J., Wang, H., Liebman, A. (2023). Marl for decentralized electric vehicle charging coordination with V2V Energy Exchange. IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society. https://doi.org/10.1109/iecon51785.2023.10312315

[23] Tao, Y., Qiu, J., Lai, S., Sun, X., Wang, Y., Zhao, J. (2023). Data-driven matching protocol for vehicle-to-vehicle energy management considering privacy preservation. IEEE Transactions on Transportation Electrification, 9(1), 968–980. https://doi.org/10.1109/tte.2022.3188766

[24] Shurrab, M., Singh, S., Otrok, H., Mizouni, R., Khadkikar, V., Zeineldin, H. (2022a). A stable matching game for V2V energy sharing–A user satisfaction framework. IEEE Transactions on Intelligent Transportation Systems, 23(7), 7601–7613. https://doi.org/10.1109/tits.2021.3071449

[25] Kim, O. T., Tran, N. H., Nguyen, V., Kang, S. M., Hong, C. S. (2018). Cooperative between V2C and V2V charging: Less range anxiety and more charged evs. 2018 International Conference on Information Networking (ICOIN). https://doi.org/10.1109/icoin.2018.8343205

[26] Saputra, Y. M., Hoang, D. T., Nguyen, D. N., Dutkiewicz, E., Mueck, M. D., Srikanteswara, S. (2019). Energy demand prediction with Federated Learning for Electric Vehicle Networks. 2019 IEEE Global Communications Conference (GLOBECOM). https://doi.org/10.1109/globecom38437.2019.9013587

[27] Ramoliya, F., Trivedi, C., Darji, K., Kakkar, R., Gupta, R., Tanwar, S., Polkowski, Z., Alqahtani, F., Tolba, A. (2024). ML-based energy consumption and distribution framework analysis for evs and charging stations in smart grid environment. IEEE Access, 12, 23319–23337. https://doi.org/10.1109/access.2024.3365080

[28] Thorgeirsson, A. T., Scheubner, S., Funfgeld, S., Gauterin, F. (2021). Probabilistic prediction of energy demand and driving range for electric vehicles with Federated Learning. IEEE Open Journal of Vehicular Technology, 2, 151–161. https://doi.org/10.1109/ojvt.2021.3065529

[29] Eagon, M. J., Kindem, D. K., Panneer Selvam, H., Northrop, W. F. (2022). Neural network-based electric vehicle range prediction for smart charging optimization. Journal of Dynamic Systems, Measurement, and Control, 144(1). https://doi.org/10.1115/1.4053306

[30] Albuquerque, D. (2022a). Electric Vehicle X Driving Range Prediction – EV X DRP (thesis).

[31] Yavuz, M., Kivanç, Ö. C. (2024). Optimization of a cluster-based energy management system using deep reinforcement learning without affecting Prosumer Comfort: V2X technologies and peer-to-peer energy trading. IEEE Access, 12, 31551–31575. https://doi.org/10.1109/access.2024.3370922

[32] Chopra, A., Rahman, A. U., Malik, A. W., Ravana, S. D. (2022). Adaptive-learning-based vehicle-to-vehicle opportunistic resource-sharing framework. IEEE Internet of Things Journal, 9(14), 12497–12504. https://doi.org/10.1109/jiot.2021.3137264

[33] Yassine, A., Hossain, M. S. (2023). Match maximization of vehicle-to-vehicle energy charging with double-sided auction. IEEE Transactions on Intelligent Transportation Systems, 24(11), 13250–13259. https://doi.org/10.1109/tits.2023.3265870

[34] Tao, Y., Qiu, J., Lai, S., Sun, X., Wang, Y., Zhao, J. (2023a). Data-driven matching protocol for vehicle-to-vehicle energy management considering privacy preservation. IEEE Transactions on Transportation Electrification, 9(1), 968–980. https://doi.org/10.1109/tte.2022.3188766

[35] Shurrab, M., Singh, S., Otrok, H., Mizouni, R., Khadkikar, V., Zeineldin, H. (2022a). A stable matching game for V2V energy sharing–A user satisfaction framework. IEEE Transactions on Intelligent Transportation Systems, 23(7), 7601–7613. https://doi.org/10.1109/tits.2021.3071449

[36] Zhang, R., Cheng, X., Yang, L. (2019). Flexible energy management protocol for cooperative EV-to-EV charging. IEEE Transactions on Intelligent Transportation Systems, 20(1), 172–184. https://doi.org/10.1109/tits.2018.2807184

[37] Bulut, E., Kisacikoglu, M. C., Akkaya, K. (2019). Spatio-temporal non-intrusive direct V2V charge sharing coordination. IEEE Transactions on Vehicular Technology, 68(10), 9385–9398. https://doi.org/10.1109/tvt.2019.2931954

[38] Jarrahi, M. H. (2018). Artificial Intelligence and the future of work: Human-AI symbiosis in organizational decision making. Business Horizons, 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

[39] Kufel, J., Bargieł-Łaczek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., Gruszczyńska, K. (2023). What is machine learning, artificial neural networks and deep learning?—examples of practical applications in medicine. Diagnostics, 13(15), 2582. https://doi.org/10.3390/diagnostics13152582

[40] What is a model architecture?. Hopsworks. (n.d.). https://www.hopsworks.ai/dictionary/model-architecture: :text=What

[41] Kupek, T. (2023, January 4). Getting your data in shape for machine learning. Stack Overflow. https://stackoverflow.blog/2023/01/04/getting-your-data-in-shape-for-machine-learning/

[42] Ev database. EV Database. (n.d.-a). https://ev-database.org/

[43] Passenger mobility statistics. Eurostat Statistics Explained. (n.d.) `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Passenger_mobility_statistics&stable=0#Distance_covered`

[44] Iea. (n.d.). Trends in charging infrastructure – Global EV Outlook 2023 – analysis. IEA. https://www.iea.org/reports/global-ev-outlook-2023/trends-in-charging-infrastructure

[45] Solar On EV. (2021, November 16). Worldwide daily driving distance is 25-50km? what about Au, US, UK, EU, and... https://www.solaronev.com/post/average-daily-driving-distance-for-passenger-vehicles

[46] (2024, February 23). Is it better to charge your electric car when the SOC is below 20? https://wattsaving.com/blogs/knowledge-base/charge-your-car-at-soc-20

[47] EV queue: Leasing Options. EV Queue — Leasing Options. (n.d.). https://www.leasingoptions.co.uk/news/blog/ev-queue-the-longest-wait-times-in-the-uk/8895

[48] Electric vehicle charging. Mission Electric. (n.d.). https://www.missionelectric.org/electric-vehicle-charging.html

[49] Alghawi, M. (2024). V2V Data frame - Situation 1, 2 [Data set]. https://github.com/MarwaAlghawi/V2V-data-frame/tree/main

[50] What is exploratory data analysis?. IBM. (n.d.). https://www.ibm.com/cloud/learn/exploratory-data-analysis

[51] Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. Journal of Applied Science and Technology Trends, 1(4), 140–147. https://doi.org/10.38094/jastt1457

[52] Scikit learn tutorial. Tutorialspoint. (n.d.). `https://www.tutorialspoint.com/scikit_learn/index.htm`

[53] Nantasenamat, C. (2021, June 4). How to build a machine learning model. Medium. https://towardsdatascience.com/how-to-build-a-machine-learning-model-439ab8fb3fb1

[54] Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., & García-Herrera, R. (2018). Feature selection in machine learning prediction systems for Renewable Energy Applications. Renewable and Sustainable Energy Reviews, 90, 728–741. https://doi.org/10.1016/j.rser.2018.04.008

[55] Birba, D. (2020). A comparative study of data splitting algorithms for ... http://kth.diva-portal.org/smash/get/diva2:1506870/FULLTEXT01.pdf

[56] Gillis, A. S. (2022, April 15). What is data splitting and why is it important?. Enterprise AI. https://www.techtarget.com/searchenterpriseai/definition/data-splitting

[57] Nyuytiymbiy, K. (2022, March 28). Parameters and hyperparameters in machine learning and Deep Learning. Medium. https://towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac

[58] Jia Wu, Xiu-Yun Chen, Hao Zhang, Li-Dong Xiong, Hang Lei, Si-Hao Deng, Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb, Journal of Electronic Science and Technology, Volume 17, Issue 1, 2019, Pages 26-40, ISSN 1674-862X, https://doi.org/10.11989/JEST.1674-862X.80904120.

[59] Shanmugavadivel, K., Sathishkumar, V. E., Kumar, M. S., Maheshwari, V., Prabhu, J., Allayear, S. M. (2022). Investigation of applying machine learning and hyperparameter tuned deep learning approaches for arrhythmia detection in ECG images. Computational and Mathematical Methods in Medicine, 2022, 1–12. https://doi.org/10.1155/2022/8571970

[60] Ansari, Y. (2023, June 14). Linear Regression (an overview). Medium. https://medium.com/@novus$_afk/linear-regression-an-overview-13d37a6bc4dd$

[61] Gore, K. (2018). GENERAL LINEAR REGRESSION. IJRAR , 06(01), 44–46. https://doi.org/E-ISSN 2348-1269

[62] Danishuddin, Kumar, V., Faheem, M., Woo Lee, K. (2022). A decade of machine learning-based predictive models for human pharmacokinetics: Advances and challenges. Drug Discovery Today, 27(2), 529–537. https://doi.org/10.1016/j.drudis.2021.09.013

[63] Muschelli, J. (2020, October). Roc and AUC with a binary predictor: A potentially misleading metric. Journal of classification. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7695228/

[64] Alghawi, M., Mounsef, J. (2024). Overview of vehicle-to-vehicle energy sharing infrastructure. IEEE Access, 12, 54567–54589. https://doi.org/10.1109/access.2024.3388088