INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

Área Departamental de Engenharia de Electrónica e Telecomunicações e de Computadores

Electric Vehicle X Driving Range Prediction – EV X DRP

David Alexandre Sousa Gomes Albuquerque

Licenciado

Projecto Final para obtenção do Grau de Mestre
em Engenharia Informática e de Computadores

Orientadores : Doutor David Pereira Coutinho
Doutor Artur Jorge Ferreira

Júri:

Presidente: Doutor Nuno Datia

Vogais: Doutor David Pereira Coutinho
Doutor Artur Jorge Ferreira
Doutor Gonçalo Duarte

Janeiro, 2022

INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

Área Departamental de Engenharia de Electrónica e Telecomunicações e de Computadores

Electric Vehicle X Driving Range Prediction – EV X DRP

David Alexandre Sousa Gomes Albuquerque

Licenciado

Projecto Final para obtenção do Grau de Mestre
em Engenharia Informática e de Computadores

Orientadores : Doutor David Pereira Coutinho
Doutor Artur Jorge Ferreira

Júri:

Presidente: Doutor Nuno Datia

Vogais: Doutor David Pereira Coutinho
Doutor Artur Jorge Ferreira
Doutor Gonçalo Duarte

Janeiro, 2022

Abstract

The electric vehicle (EV) use as a reliable and eco-friendly means of transportation has increased rapidly over the past few years. When choosing an EV, the vehicle's driving range capacity is a decisive factor to be taken into account as it minimizes drive's anxiety while driving. An EV's driving range prediction depends on multiple factors that must be factored in to accurately infer results. As machine learning has become a widely used approach for highly complex problems such as the EV driving range prediction, machine learning based solutions already provide accurate enough solutions for the problem. The present project explores the implementation of a machine learning based model to learn and adapt from existing data, while aiming to provide more real and accurate results.

Keywords: electric vehicle; range prediction; energy consumption; machine learning

Contents

List of Figures **ix**

List of Tables **xi**

Acronyms **xiii**

Glossary **xv**

1 Introduction **1**

2 State of Art **5**

2.1 EV Range prediction 5

2.2 Machine Learning and EV Range prediction 6

3 Future Work & Planning **7**

3.1 Intro 7

3.2 Planning 8

References **9**

List of Figures

1.1 System overview - Dataset generation. 2
1.2 System overview - Learning phase. 3
1.3 System overview - Estimation phase. 3

3.1 Project planning. 8

List of Tables

Acronyms

DT	decision tree.
eRange	electric range.
EV	electric vehicle.
GBRT	gradient boosted regression trees.
KNN	k-nearest neighbor.
LightGBM	light gradient boosting regression tree.
MLR	multiple linear regression.
RF	random forest.
RT	regression tree.
SOC	state of charge.
SOM	self-organizing map.
XGBoost	extreme gradient boosting regression tree.

Glossary

Big data	Big data is a field that handles with large datasets that are too big and complex for traditional data processing .
Dataset	A structure containing data for a model.
Machine learning	A branch of AI focused on learning from data.
Neural networks	A collection of connected synapse nodes, simulating a biological brain .
Python	A high level programming language.
Time series	A series of data points indexed by time.

1

Introduction

On today's day and age, the global concern on climate change has been a major focus on recent international agreements, such as the Paris Agreement (*Paris Agreement* 2015), incentivizing many car manufacturers to introduce electric vehicles (EVs) as the eco-friendly solution for sustainable transport for the future.

EVs have been growing in popularity in recent years and as a result, car manufacturers have increased competitiveness on vehicle's performance (Figenbaum *et al.*, 2015), namely the driving range capacity is a decisive factor for consumers (Egbue and Long, 2012).

The EV's autonomy also known as electric range (eRange), allows consumers to know an estimate of the remaining driving distance for the existing EV battery power, easing driver's anxiety for the duration of a trip to a charging station (Smuts *et al.*, 2017; Song and Hu, 2021).

The eRange can be estimated through many driving data parameters, such as vehicle design, driver's behavior, whether, road inclination and state of charge (SOC) estimation and its accuracy allows consumers to rely on its vehicle for longer travel time and efficient charging plans however, eRange estimation is a complex problem with multiple influencing factors (Varga *et al.*, 2019), fueling previous studies in the past to provide a solution for this challenge.

The rise in popularity of machine learning (Amershi *et al.*, 2019) has demonstrated its effectiveness in the past with a variety of fields such as big data (Condie *et al.*, 2013; Zhou *et al.*, 2017), pattern recognition analysis and data mining (Bose and Mahapatra, 2001). This is due to its nature of learning from previous data to gradually achieve better results making it a widely recognized tool for complex problems (Mitchell, 2006). As a result, machine learning has been chosen in the past as one of the adopted solutions for the eRange estimation problem (will be shown in section 2.2) making it a more accurate solution.

This project addresses the eRange estimation problem through the use of a machine learning based model, being comprised by three distinct phases: the Dataset generation phase, the Learning phase and the Estimation phase.

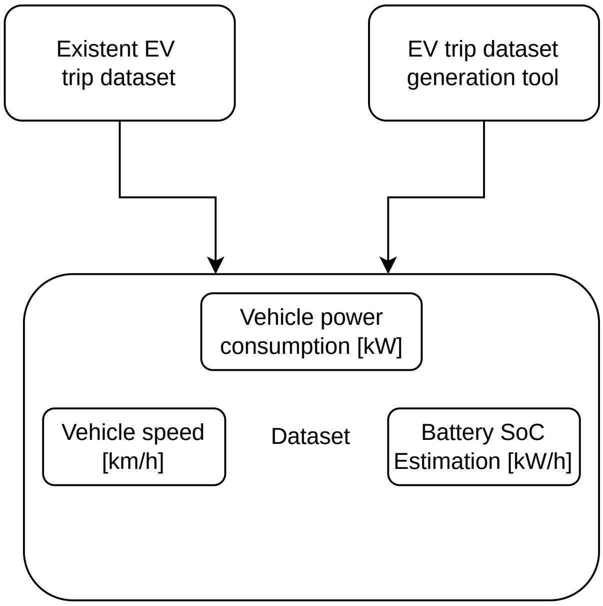


Figure 1.1: System overview - Dataset generation.

On the Dataset generation phase (figure 1.1) a dataset will be created from historical traffic data from personally recorded vehicle trips, as well as external existing datasets such as *VED* (Oh *et al.*, 2019) or *Emobpy*, an external library that provides the generation of EV trip and consumption time series (Gaete-Morales *et al.*, 2021). The resulting dataset contains multiple trips with their respective vehicle power consumption [kW] and the vehicle speed [km/h] in a time series format.

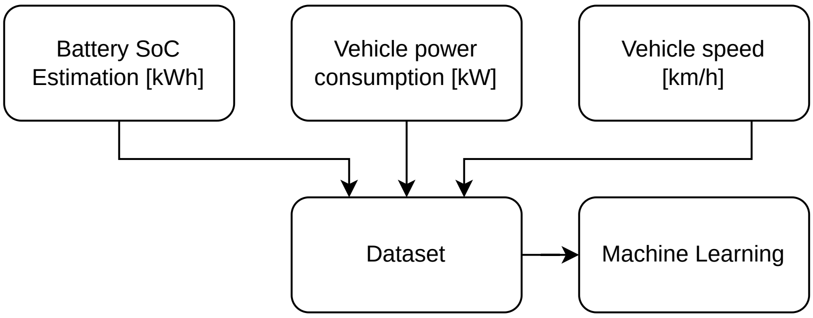


Figure 1.2: System overview - Learning phase.

The generated dataset will then be used to train the model on the Learning phase through machine learning, allowing it to learn the estimation of the eRange for each trip on the dataset as depicted in figure 1.2.

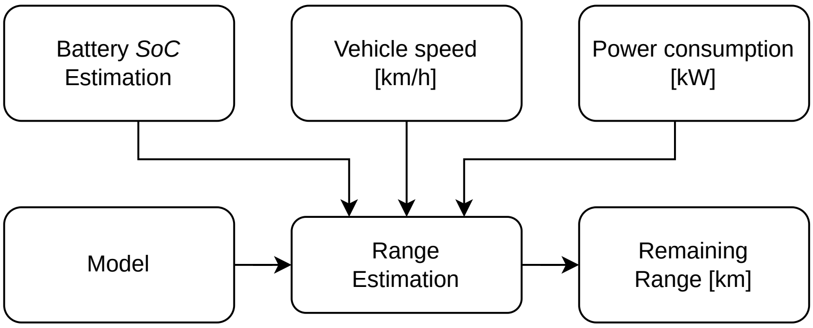


Figure 1.3: System overview - Estimation phase.

After training and learning the model with the dataset, the Estimation phase performs the eRange prediction on live SOC monitoring of a driving EV (figure 1.3).

The remainder of this document is structured as follows: Chapter 1 refers existing solutions of the eRange estimation problem and their reliance on available datasets (Section 2.1) while mentioning machine learning and its usability in existing eRange estimation solutions (Section 2.2). On Chapter 3 will be discussed the work state (Section 3.1) followed by future work and planning for the next stage in the development of this work (Section 3.2).

2

State of Art

2.1 EV Range prediction

Nowadays In today's day and age, EVs have spired multiple studies concerning related problems in this field, such as statistical measurement of charging (Brighente *et al.*, 2021), regenerative braking (Yoong *et al.*, 2010), charging topologies (Yilmaz and Krein, 2013) and eRange prediction (Varga *et al.*, 2019).

The eRange prediction is an important EV feature to present to consumers as it reduces driver's anxiety while driving. This has been previously studied before, prompting multiple ways to tackle the problem.

When proving a solution to the eRange prediction problem, valid EV driving data in the form of a dataset is required to learn and test the proposed model, and compare it to existing alternatives, making it indispensable in determining the effectiveness of the chosen solution.

To this end, existing datasets such as the *VED* dataset (Oh *et al.*, 2019) that although providing sufficient EV driving data for estimation, there are only three dinstinct EVs present on the dataset, all from the same model 2013 *Nissan leaf*.

Another dataset solution is the use of a Python tool *Emobpy* (Gaete-Morales *et al.*, 2021) that generates multiple trips for charging and driving studies on EVs, providing a dataset based on empirical mobility statistics and customizable assumptions.

Existing work has demonstrated the use of eRange estimation on EVs, showing the need different types of accuracy on eRange estimation depending on the SOC state, the proposed approach by Zhang *et al.*, 2012 minimizes the performance impact of minimum cost route searching from high accuracy eRange prediction.

Other studies have focused on delivering higher eRange estimation accuracy, making use of more complex models. The use of an adaptive history based model approach was proposed by Coutinho, 2021, which relies on past information about vehicle's instant consumption energy, to determine an adaptive average energy consumption.

2.2 Machine Learning and EV Range prediction

The use of machine learning for a multitude of cases (Amershi *et al.*, 2019) in fields such as big data (Condie *et al.*, 2013; Zhou *et al.*, 2017) and data mining (Bose and Mahapatra, 2001) has proven its robustness on solving complex problems.

As a result, some approaches for the eRange problem have already applied machine learning, most notably supervised learning to achieve the estimation. Decision trees (DTs), random forest (RF), and k-nearest neighbor (KNN), have been used in ensemble stacked generalization (ESG) approach (Ullah *et al.*, 2021) proving its effectiveness in yielding more acceptable values for proposed evaluation metrics. Recent models using gradient boosted regression trees (GBRT) have combined extreme gradient boosting regression tree (XGBoost) and light gradient boosting regression tree (LightGBM) to provide better predictive performance from these ensemble methods (Zhao *et al.*, 2020).

Approaches using unsupervised clustering of self-organizing maps (SOMs) for clustering big data into driving patterns (Lee and Wu, 2015). A hybridised version of SOM with regression trees (RTs) has used SOM's feature of its neurons storing related information being kept closely together to perform the RTs locally, avoiding bushy trees and improving upon previous solutions by keeping meaningful knowledge extraction on bushy trees (Zheng *et al.*, 2016).

Reinforcement learning in the form of neural networks has also been used for external energies disturbances on the speed profile of a driving profiles so that it could then be combined with multiple linear regression (MLR) for the estimation (De Cauwer *et al.*, 2017).

Although more complex than previous solutions, the use of machine learning for the eRange estimation problem had reduced the prediction error, and thus further justifying its usage in this project.

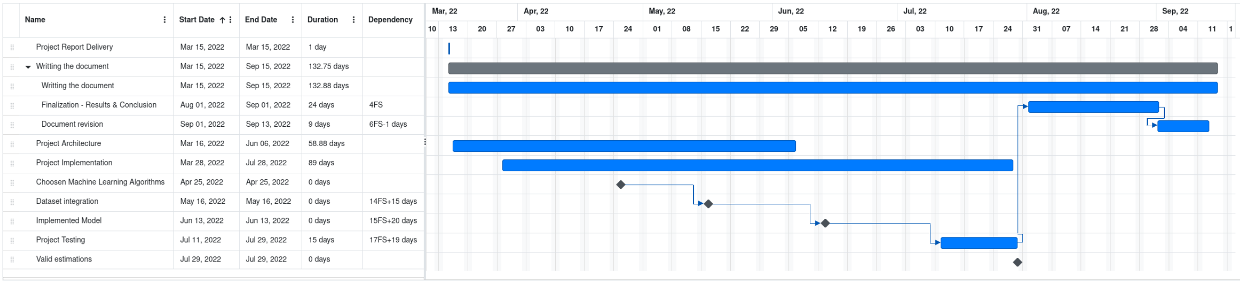
3

Future Work & Planning

3.1 Intro

The initial stage of this project focused on studying and understanding the problem, analyzing existing eRange solutions. From then, researching usable a dataset was prioritized as it would prove vital for testing the study's model. Settling in two equally valid datasets (Oh *et al.*, 2019 and Gaete-Morales *et al.*, 2021), the new focus for this study shifted on implementing an existing solution of the "basic", and adaptive history based model approaches (proposed by Coutinho, 2021) while also writing this report and thus, additional research on existing related works was conducted for better understanding of the problem regarding eRange estimation and its difficulties.

3.2 Planning



Name	Start Date	End Date	Duration	Dependency
Project Report Delivery	Mar 15, 2022	Mar 15, 2022	1 day	
Writting the document	Mar 15, 2022	Sep 15, 2022	132.75 days	
Writting the document	Mar 15, 2022	Sep 15, 2022	132.88 days	
Finalization - Results & Conclusion	Aug 01, 2022	Sep 01, 2022	24 days	4FS
Document revision	Sep 01, 2022	Sep 13, 2022	9 days	6FS-1 days
Project Architecture	Mar 16, 2022	Jun 06, 2022	58.88 days	
Project Implementation	Mar 28, 2022	Jul 28, 2022	89 days	
Choosen Machine Learning Algorithms	Apr 25, 2022	Apr 25, 2022	0 days	
Dataset integration	May 16, 2022	May 16, 2022	0 days	14FS+15 days
Implemented Model	Jun 13, 2022	Jun 13, 2022	0 days	15FS+20 days
Project Testing	Jul 11, 2022	Jul 29, 2022	15 days	17FS+19 days
Valid estimations	Jul 29, 2022	Jul 29, 2022	0 days	

Figure 3.1: Project planning.

The project development will be separated into three separate stages: Project Architecture, Project Implementation, and Project Testing. Project Architecture stage will plan the software architecture to develop, this stage will continue after Project Implementation has started, allowing future adjustments according to implementation demands. The required software to have a working model on this project will be developed during the Implementation stage, having the following milestones: Dataset integration, and Implemented Model. The Project Testing stage will ascertain the validity of the estimated results. During the previously mentioned project stages, the thesis will be continually worked until its completion, depending on Project Testing stage to reflect on the projects results and accomplishments.

References

[1] Murat Yilmaz and Philip T. Krein, "Review of battery charger topologies, charging power levels, and infrastructure for plug-in electric and hybrid vehicles", *IEEE Transactions on Power Electronics*, vol. 28, no. 5, pages 2151–2169, 2013. DOI: 10.1109/TPEL.2012.2212917.

[2] Indranil Bose and Radha K. Mahapatra, "Business data mining - a machine learning perspective", *Information & Management*, vol. 39, no. 3, pages 211–225, 2001, ISSN: 0378-7206. DOI: https://doi.org/10.1016/S0378-7206(01)00091-X. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037872060100091X.

[3] David Coutinho, "Classic ev x project driving range prediction", Draft version, Jul. 2021.

[4] Yang Song and Xianbiao Hu, "Learning electric vehicle driver range anxiety with an initial state of charge-oriented gradient boosting approach", *Journal of Intelligent Transportation Systems*, vol. 0, no. 0, pages 1–19, 2021. DOI: 10.1080/15472450.2021.2010053. eprint: https://doi.org/10.1080/15472450.2021.2010053. [Online]. Available: https://doi.org/10.1080/15472450.2021.2010053.

[5] Ona Egbue and Suzanna Long, "Barriers to widespread adoption of electric vehicles: An analysis of consumer attitudes and perceptions", *Energy Policy*, vol. 48, pages 717–729, 2012, Special Section: Frontiers of Sustainability, ISSN: 0301-4215. DOI: https://doi.org/10.1016/j.enpol.2012.06.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301421512005162.

[6] Carlos Gaete-Morales, Hendrik Kramer, Wolf-Peter Schill, and Alexander Zerrahn, "An open tool for creating battery-electric vehicle time series from empirical data, emobpy", *Scientific Data*, vol. 8, no. 1, page 152, 2021, ISSN: 2052-4463. DOI: 10.1038/s41597-021-00932-9. [Online]. Available: https://doi.org/10.1038/s41597-021-00932-9.

[7] Y. Zhang, W. Wang, Y. Kobayashi, and K. Shirai, "Remaining driving range estimation of electric vehicle", in *2012 IEEE International Electric Vehicle Conference*, ser. 2012 IEEE International Electric Vehicle Conference, 2012, pages 1–7. DOI: 10.1109/IEVC.2012.6183172. [Online]. Available: https://doi.org/10.1109/IEVC.2012.6183172.

[8] Martin Smuts, Brenda Scholtz, and Janet Wesson, "A critical review of factors influencing the remaining driving range of electric vehicles", in *2017 1st International Conference on Next Generation Computing Applications (NextComp)*, 2017, pages 196–201. DOI: 10.1109/NEXTCOMP.2017.8016198.

[9] Irfan Ullah, Kai Liu, Toshiyuki Yamamoto, Muhammad Zahid, and Arshad Jamal, "Electric vehicle energy consumption prediction using stacked generalization: An ensemble learning approach", *International Journal of Green Energy*, vol. 18, no. 9, pages 896–909, 2021. DOI: 10.1080/15435075.2021.1881902. eprint: https://doi.org/10.1080/15435075.2021.1881902. [Online]. Available: https://doi.org/10.1080/15435075.2021.1881902.

[10] Chung-Hong Lee and Chih-Hung Wu, "A novel big data modeling method for improving driving range estimation of evs", *IEEE Access*, vol. 3, pages 1980–1993, 2015. DOI: 10.1109/ACCESS.2015.2492923.

[11] Cedric De Cauwer, Wouter Verbeke, Thierry Coosemans, Saphir Faid, and Joeri Van Mierlo, "A data-driven method for energy consumption prediction and energy-efficient routing of electric vehicles in real-world conditions", *Energies*, vol. 10, no. 5, 2017, ISSN: 1996-1073. DOI: 10.3390/en10050608. [Online]. Available: https://www.mdpi.com/1996-1073/10/5/608.

[12] Erik Figenbaum, Nils Fearnley, Paul Pfaffenbichler, Randi Hjorthol, Marika Kolbenstvedt, Reinhard Jellinek, Bettina Emmerling, G. Maarten Bonnema, Farideh Ramjerdi, Liva Vågane, and Lykke Møller Iversen, "Increasing the competitiveness of e-vehicles in europe", *European Transport Research Review*, vol. 7, no. 3, page 28, 2015, ISSN: 1866-8887. DOI: 10.1007/s12544-015-0177-1. [Online]. Available: https://doi.org/10.1007/s12544-015-0177-1.

[13] Alessandro Brighente, Mauro Conti, Denis Donadel, and Federico Turrin, "Evscout2.0: Electric vehicle profiling through charging profile", *CoRR*, vol. abs/2106.16016, 2021. arXiv: 2106.16016. [Online]. Available: https://arxiv.org/abs/2106.16016.

[14] Tyson Condie, Paul Mineiro, Neoklis Polyzotis, and Markus Weimer, "Machine learning on big data", in *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, 2013, pages 1242–1244. DOI: 10.1109/ICDE.2013.6544913.

[15] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V. Vasilakos, "Machine learning on big data: Opportunities and challenges", *Neurocomputing*, vol. 237, pages 350–361, 2017, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2017.01.026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231217300577.

[16] Saleema Amershi, Andrew Begel, Christian Bird, Robert DeLine, Harald Gall, Ece Kamar, Nachiappan Nagappan, Besmira Nushi, and Thomas Zimmermann, "Software engineering for machine learning: A case study", in *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, 2019, pages 291–300. DOI: 10.1109/ICSE-SEIP.2019.00042.

[17] Liang Zhao, Wei Yao, Yu Wang, and Jie Hu, "Machine learning-based method for remaining range prediction of electric vehicles", *IEEE Access*, vol. 8, pages 212 423–212 441, 2020. DOI: 10.1109/ACCESS.2020.3039815.

[18] B. Zheng, P. He, L. Zhao, and H. Li, "A hybrid machine learning model for range estimation of electric vehicles", in *2016 IEEE Global Communications Conference (GLOBECOM)*, ser. 2016 IEEE Global Communications Conference (GLOBECOM), 2016, pages 1–6. DOI: 10.1109/GLOCOM.2016.7841506. [Online]. Available: https://doi.org/10.1109/GLOCOM.2016.7841506.

[19] Tom Mitchell, "The discipline of machine learning", Tech. Rep. CMU ML-06 108, 2006.

[20] *Paris agreement*, UN Treaty, United Nations, Dec. 2015. [Online]. Available: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-d&chapter=27&clang=_en.

[21] Bogdan Ovidiu Varga, Arsen Sagoian, and Florin Mariasiu, "Prediction of electric vehicle range: A comprehensive review of current issues and challenges", *Energies*, vol. 12, no. 5, 2019, ISSN: 1996-1073. DOI: 10.3390/en12050946. [Online]. Available: https://www.mdpi.com/1996-1073/12/5/946.

[22] M.K Yoong, Y.H Gan, G.D Gan, C.K Leong, Z.Y Phuan, B.K Cheah, and K.W Chew, "Studies of regenerative braking in electric vehicle", in *2010 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology*, 2010, pages 40–45. DOI: 10.1109/STUDENT.2010.5686984.

[23] G. S. Oh, David J. Leblanc, and Huei Peng, *Vehicle energy dataset (ved), a large-scale dataset for vehicle energy consumption research*, 2019. arXiv: 1905.02081 [physics.soc-ph].