Smart Energy In The Information Age Project Proposal

Electric Car Range Prediction and feature analysis using Machine Learning

Harshith Reddy Kallu	Viraj Kamat	Rohit Rawat
112674629	112818603	112963417

Abstract

Given a dataset of electric cars containing information about their battery capacity, electric range, weather, and other exogenous data regarding the car we wish to use Machine Algorithms along with deep learning to predict electric car range and how to better use these predictions in improving electric car travel range. We also wish to provide a novel approach on top of existing machine learning approaches to provide a more accurate model for our study.

Intuition & Motivation of our Project

Considering current global pollution and scarcity of nonrenewable fuels, we are moving towards renewable fuels which are environment-friendly options. In the vehicle industry, one of the good options is Electric Vehicles. There are many companies that are working in this direction and making good progress but one challenging factor which is important in this scenario is the range for batteries of Electric vehicles. So, we are planning to use different Machine Learning models to predict battery life for Electric Vehicles. Additionally, we will also explore the interesting dependent insights based on those predictions using data on the charging stations, etc.

Dataset Information

We have gathered a few datasets:

1. SpritMonitor Crawler:https://github.com/armiro/crawlers/tree/master/SpritMonitor-Crawler This crawler crawls spritmonitor.de in the vehicle fueling records; extracting useful data (e.g. trip_distance, driving_style, etc.) and appends a record to vehicle_name.csv.
2. Weather dataset: We are going to incorporate the weather factors.
3. Plug-In EVerywhere Charging Station Network: This dataset contains records for different locations of the charging stations.

Challenges

We face challenges due to vehicle design, driver characteristics, and the external environment when it comes to range prediction. Some of the parameters are constant such as mass, weight, transmission type, etc and we also have variable parameters such as traffic flow, battery, and external weather. We need to consider factors in vehicle design and how a vehicle in motion is under the influence of different forces. The prediction of the range depends on the battery state of charge - which is influenced by the hardware sensors and software calculation of the state of charge which depends on temperature and other factors. Another challenge is that we have to somehow parameterize driver behavior, for example, how aggressive he is will affect the battery discharge and in turn, influence the range of the EV. The field is still developing so we face the challenge of data scarcity.

Approach

1. We are planning to do web crawling to gather the data. Then, Data cleaning and preprocessing will be done before implementing models.
2. Using Principal Component Analysis(PCA) or Multidimensional Scaling(MDS), we will try to come up with the most significant features that affect battery life.
3. In our objective of predicting the range of an electric car with machine learning we will be implementing the following classes of machine learning algorithms; namely :
   A. Regression
   B. Long Short Term Memory Neural Model (L.S.T.M)
   C. Ensembling multiple models (XGBoost & L.S.T.M)
4. While the above list is not exhaustive, we will be using other machine learning algorithms and also exploring novel approaches that could help improve the accuracy of our predictions.
5. We are still working on building a novel solution while also attempting to incorporate as much exogenous data as possible that would help with our prediction accuracy.
6. In the end, we will be implementing optimization algorithms to optimize the models for better performance like Online gradient descent, Online balanced descent, etc.

Next Steps & Road-Map

1. Clean the dataset, filter outliers along with imputing missing values (Data preprocessing)
2. Apply various machine learning models to perform our analysis.
3. Incorporate exogenous data to improve our prediction accuracy.
4. Research into novel/combined approaches to address the prediction problem in a new way.
5. Testing & bug-fixing our models.