**Team 5: U.S. Electronic Vehicle Sales and Technical Paper Citations Analysis**

UH SPE Machine Learning Bootcamp First Project: Non-Linear Modelling

Dung Nguyen, Xu Yang, Charles White, Jose Benavides

Abstract

Sales of electric, hybrid electric, and plugin hybrid vehicles have increased dramatically in the last decade. This is due to multiple advances in lithium ion battery technologies. These technological advances have their roots in technical papers that are written for multiple industry journals. These papers represent the origin of the technologies that will later appear in the vehicles that are being sold. By establishing a metric for understanding the impact that a technical paper has, a prediction can be made as to the number of electronic vehicles that will be sold in the coming years after the paper is released. Understanding future vehicle sales will allow vehicle manufacturers to know the number of vehicles that will need to be produced in a given year.

The dataset for our independent variable will be the time series of technical publications that are related to lithium ion batteries technologies. This will be done by scraping down to a local machine by using Publish or Perish 7 software. The dataset for our dependent variable will be the time series of include electric, hybrid electric, and plugin hybrid vehicles which will be gathered from AFDC.Gov. Once all data is gathered it will be evaluated and all unnecessary data columns will be dropped according to domain knowledge. The organized columns will contain keywords containing Booleans to ask as separate feature vectors.

Most likely, due to the nature of the assignment, the techniques to be used will be: a neural network and support vector machining if the data allows.