**Electric Motor Speed Prediction**

**Section 1 (About the Data Set):**  The Dataset is available on Kaggle and is meant to predict the speed of the electric vehicle given several components like current parameters, voltage parameters etc.

It comprises of several sensor related data. **All recordings are sampled at a frequency of 2Hz**. The dataset’s importance is to claim and show how well the electric motors are performing with such less effort. In a world full of growing technology, electric motors are a must now.

**Features:**

**Ambient** - Ambient temperature as measured by a thermal sensor located closely to the stator.

**Coolant** - Coolant temperature. The motor is water cooled. Measurement is taken at outflow.

**u\_d** - Voltage d-component

**u\_q** - Voltage q-component

**Motor\_speed** - Motor speed

**Torque** - Torque induced by current.

**i\_d** - Current d-component

**i\_q** - Current q-component

**pm** - Permanent Magnet surface temperature representing the rotor temperature. This was measured with an infrared thermography unit.

**Stator\_yoke** - Stator yoke temperature measured with a thermal sensor.

**Stator\_tooth** - Stator tooth temperature measured with a thermal sensor.

**Stator\_winding** - Stator winding temperature measured with a thermal sensor.

**Profile\_id** - Each measurement session has a unique ID. Make sure not to try to estimate from one session onto the other as they are strongly independent.

**Section 2 (Exploratory Data Analysis):** To better understand the data, we perform EDA. For my dataset, i have done the following:

1. **Taking a look our data is oriented:**

A screenshot of a cell phone

Description automatically generatedChecking the Data Types and no of records

**** Shape of the dataset

1. **Basic Descriptive Statistics of the dataset:**

A screenshot of a cell phone

Description automatically generated

1. **Null Value Checker:**

A screenshot of a cell phone

Description automatically generatedMy dataset has no null values

1. **Correlation Check using Heatmap:A screenshot of a video game

   Description automatically generated**

We can hardly find any high positive correlation so we will use all the features instead of selected features for prediction of motor speed.

**Section 3 (Data Pre-processing):**

Our data requires no scaling since there is no large difference between values, but it surely requires outlier removal due to so many outliers.

Using IQR outlier removal method we achieve the new shape of (814996,12).

No of outliers: 183,074

Now that our data is not skewed and is normalized, we can move forward to machine learning.

**Section 4 (Machine Learning):**

A screenshot of a cell phone

Description automatically generated

The following data frame shows us how our different regression models performed on test set. The data wasn’t overfitting and was normalized giving us excellent results.

A screenshot of a cell phone

Description automatically generated

Polynomial regression after being plotted seemed to give us accurate values which means that the residual error was very less. The features we chose were of second degree.

A close up of a map

Description automatically generated

**Arnav Oberoi, 18CSU032**