**Weekly Report - 1**

**CSE 523 Machine Learning**

Group Name: - **Mavericks**

**Store Sales Forecasting**

| **Group Members** | **Roll No.** |
| --- | --- |
| Dhruv Prajapati | AU1940192 |
| Aaryan Mori | AU1940194 |
| Harshil Doshi | AU1940279 |
| Dhaval Chaudhary | AU1940180 |

1. **Task Performed in the first week:**

* Approaching a problem statement and understanding it in depth.
* Analyzing required dataset for building a model.
* Exploring various libraries which can be used.
* Analyzing the respective dataset and understanding the features of a model.

1. **Outcome of the task in the first week:**

We have analyzed the given data and the outcomes are listed below:

* There are **54 stores** and **33 product families** in the data. The time series starts from **2013-01-01** and finishes in **2017-08-31**. However, Kaggle provides us split data as train data and test data. The dates in the test data are for the **15 days** after the last date in the training data. Date range in the test data will be very important to us while we are defining a cross-validation strategy and creating new features.
* There are six types of data given:

1. Train
2. Test
3. Store
4. Transactions
5. Holidays and Events
6. Daily Oil Price
7. **Task to be performed in the next week:**

* Have a point to point understanding of data and catch some patterns or anomalies in the data.
* Importing the suitable libraries in the code.
* Start with Exploratory Data Analysis