**Weekly Report - 5**

**CSE 523 Machine Learning**

Group Name: **Mavericks**

**Store Sales – Time series Forecasting**

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

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

* Analysing deeply the relationship of the oil prices with the sales.
* Examining the Sales for the different stores and finding the correlation among them.
* Observing the total sales of the stores.

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

* Analysing deeply the relationship of the oil prices with the sales:

1. We divided the data into two clusters. Such that over 70 and under 70. Which seems pretty understandable.
2. From this we found that all of the plots almost contain same pattern.
3. When daily oil price is under about 70, there are more sales in the data. This made us to extract new pattern for daily oil price.

* Examination over the sales for the different stores:

1. When we examined them with correlation matrix, we found that most of the stores were similar to each other. Some stores, such as 20, 21, 22 and 52 may be a little different.

* Daily total sales of the stores:

1. While observing, we realized some unnecessary rows in the data while we were looking at the time series of the stores one by one.
2. Also, if we take the stores from above, some of them have no sales at the beginning of 2013. And this can be seen in the stores 20, 21, 22, 29, 36, 42, 52 and 53.
3. So, we decided to remove those rows before the stores opened.
4. **Task to be performed in the next week:**

* Finding the particular products that aren’t sold in the particular stores using zero forecasting and removing them from the data.
* Creating the new data frame for the product families which never got sold.
* Exploring the best fit algorithm for the model.