**STAT 542 Project 2 Report**

Ankush Agrawal (NetID : ankusha2)

Following is a step by step guide to my project:

1. I developed two initial models: the naïve and a seasonal naïve. I figured seasonal naïve is giving me a good error rate. Both the models were used for each store and department combination. Missing values for the weekly sales were imputed with zeros, then the last value of the weekly sales during the train period was used to predict each new week in the testing period.
2. For the third model, I first went through the resources on Piazza and also a blog post (<https://itnext.io/understanding-the-forecasting-algorithm-stlf-model-29d74b3a0336> ) which brought me up to speed with STLF concepts. I figured I needed to decide the hyperparameters and for this I decided to leverage ETS method, with information criteria as aic. This gave me an average error rate of 1683 on 10 folds. I did change the parameters a bit but since every run was taking 1294 seconds I decided to go with my best model.

Also, since SLTF requires minimum of 24 months of data (2years worth), I decided to run seasonal naïve model for the first 6folds (=12mths) and then SLTF model for the remaining folds. This ensured I had at least 24months of training data.

1. I denoised the data by running SVD on the training data. I am taking the first 12 components which will explain the maximum variance in my training data for every department.
2. The overall run time of the evaluation script 1294.16 seconds.

Here are the details of my computer: R version 3.5.1 (2018-07-02)

Platform: x86\_64-w64-mingw32/x64 (64-bit)

Running under: Windows >= 8 x64 (build 9200)

|  |  |  |  |
| --- | --- | --- | --- |
| Fold | Naïve | Seasonal Naïve | SLTF with method=ETS and ic = 'AIC' |
| 1 | 2043.411627 | 2214.901312 | 2214.901312 |
| 2 | 2551.221926 | 1742.84044 | 1742.84044 |
| 3 | 2223.819097 | 1740.698244 | 1740.698244 |
| 4 | 2772.356836 | 1662.676699 | 1662.676699 |
| 5 | 5147.754553 | 2383.332851 | 2383.332851 |
| 6 | 4190.999148 | 1626.172716 | 1626.172716 |
| 7 | 2225.885634 | 2019.327139 | 1597.863606 |
| 8 | 2103.462799 | 1673.862006 | 1344.068181 |
| 9 | 2194.095896 | 1649.59516 | 1275.366176 |
| 10 | 2320.814505 | 1620.930529 | 1249.0123 |
| Overall Average | 2777.382202 | 1833.43371 | 1683.693253 |