CPTS475/575 Data Science

Project Proposal

Harsha Maddipati - 11633702 | Nikhil Chakravarthy **-** 11655295

Proposal 1: Project Idea 7.3: Store Item Demand Forecasting Challenge

We are given 5 years of store-item sales data. Our goal is to predict 3 months of sales for 50 different items at 10 different stores.

Our Approach:

The data set consists of fields “date”, “store”, “Item” and “sales”, We plan on obtaining time series plots of the data to analyze sales of each item at a particular date.

We are going to be pooling the stores together to conduct our analysis and also, model the stores separately to see if we can find any interesting patterns.

Proposal 2: House Prices: Advanced Regression Techniques

We are given 79 features that influence residential home prices in the city of Ames, Iowa. Our goal is to predict these prices by feature engineering.

Our Approach:

To find out which features to use, we plan on gaining some domain knowledge about the city Ames. We are also going to avoid features that have a lot of missing values. Plot SalePrice against features to observe the trends. Figure out the number of outliers in ‘SalePrice’ and check for ways to deal with it. Finally, we hope to make a sound prediction of ‘SalePrice’