Capstone Project Ideas – Anushree Srinivas

Mentor: Srdjan Santic

The three ideas that I was looking at for the capstone projects are as follows:

1) **Walmart store sales forecasting:**

Objective and Project Overview:
The objective of this project is to predict the store sales and to also predict which departments are affected and the extent of the impact by using Walmart data set across 45 different regional Walmart Stores and different departments along with the holiday markdowns.

Historical sales data for 45 Walmart stores located in different regions are provided. Each store contains a number of departments, and you are tasked with predicting the department-wide sales for each store.

In addition, Walmart runs several promotional markdown events throughout the year. These markdowns precede prominent holidays, the four largest of which are the Super Bowl, Labor Day, Thanksgiving, and Christmas. The weeks including these holidays are weighted five times higher in the evaluation than non-holiday weeks. Part of the challenge presented of this project is modeling the effects of markdowns on these holiday weeks in the absence of complete/ideal historical data.

2) **Instacart Market Basket Analysis:**

Objective and Project Overview:
The main objective of the project is to predict which previously purchased products will be in a user's next order or their first order using Instacart's anonymized customer order data over time. The ability to identify which products the customers are likely to purchase again, and automatically adding those to cart through obtained predictions or provide a seamless interface for doing so will enhance their user experience. These predictions of the products the customers are likely to order can be send in personalized communications to customers reminding them to order again, by highlighting the predicted products in those communications.

The dataset for this project is a relational set of files describing customers' orders over time. The dataset is anonymized and contains a sample of over 3 million grocery orders from more than 200,000 Instacart users. For each user, information related to

between 4 and 100 of their orders, with the sequence of products purchased in each order are provided. The week and hour of day the order was placed, and a relative measure of time between orders is also available. The data for this project can be accessed here at https://www.kaggle.com/c/instacart-market-basket-analysis/data.

3) **TFI Revenue Prediction:**

Objective and Project Overview:

With over 1,200 quick service restaurants across the globe, TFI is the company behind some of the world's most well-known brands: Burger King, Sbarro, Popeyes, Usta Donerci, and Arby's. They employ over 20,000 people in Europe and Asia and make significant daily investments in developing new restaurant sites.

Right now, deciding when and where to open new restaurants is largely a subjective process based on the personal judgement and experience of development teams. This subjective data is difficult to accurately extrapolate across geographies and cultures.

The objective of this project is to predict the annual revenue sales of TFI's 100,000 regional locations Using TFI's demographic, real estate and commercial data to. Finding a mathematical model with these predictions will help TFI increase the effectiveness of investments in new restaurant sites.

TFI has provided a dataset with 137 restaurants in the training set, and a test set of 100000 restaurants. The data columns include the open date, location, city type, and three categories of obfuscated data: Demographic data, Real estate data, and Commercial data. The revenue column indicates a (transformed) revenue of the restaurant in a given year and is the target of predictive analysis.