Foundations of Data Science

Project Ideas

Prashaanth Jagannathan

July 1st,2017

Instructor: Julian Jenkins III

1. Project: **Comparing US based airlines performance using the tweets**
   1. **Project Problem:** Comparing the engagement level, problem-solving abilities of US based Airlines and sentiment analysis from the data retrieved from Twitter.
   2. **Dataset:** Data Extraction from Twitter using R packages or using the airline sentiment dataset from Kaggle <https://www.kaggle.com/crowdflower/twitter-airline-sentiment>
   3. **Solution:** Comparison of the data enables to understand the areas which airlines has to improve, implement the best practices across various attributes and provide better understanding of the customer
2. Project: **Analyze the amazon product review or ratings to predict which product will be purchased more in the future**
   1. **Project Problem:** Identifying the most popular product and predict their future purchases based on the review or rating of product
   2. **Dataset:** Data available for individual product from <http://jmcauley.ucsd.edu/data/amazon/> or Mp3 data from http://times.cs.uiuc.edu/~wang296/Data/
   3. **Solution:** Predicting the future sales of product based on review will enable amazon to optimize the product inventory and promotions
3. Project: **Project the sales for each department in each of the 45 Walmart stores.**
   1. **Project Problem:** Predicting the sales for each department in each store with historical sales data from 45 Walmart stores located in different regions
   2. **Dataset:** Data for predicting sales for each department is available at <https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data>
   3. **Solution:** Predicting the future sales for each department in each of the stores enables Walmart to manage the product better, optimize various marketing campaign and pricing of the product.