Yagnaraman Madhavan

Post Graduate Program in Big Data and Machine Learning

2017-2018



Key Skills and Tools

Spark

Python

R

Machine Learning

Regression Analysis

Classification

Recommender Systems

Deep Learning - Tensorflow

Data Visualization - Tableau

COURSE PROJECTS

Jun'18

**Dimension ​Reduction​ ​through​ ​Principal​ ​Component​ ​Analysis​ ​on​ ​the​ ​Wine​ ​dataset**

**Course**Unsupervised Learning

This project involved step by step approach to implement Principal Component Analysis(PCA) using Python. A wine dataset with 13 attributes was used to achieve a new feature-set with lower number of dimensions that could be used for further analysis.

**Skills and Tools**

Principal Component Analysis, Unsupervised Learning, Python

May'18

**Exploratory Analysis On Vintage Art Data Using Tableau**

**Course**Visualization of Big Data

This Project involved application of Tableau in visualization of the Vintage Art data to understand the features of paintings and trends of painting purchase patterns across USA. Hive Tables of data was created and then connected to Tableau using ODBC driver. Finally, interactive Tableau dashboards were created and used for storytelling in the final report.

**Skills and Tools**

TABLEAU, Data Visualization, dashboards

May'18

**Identification of potential customer churn of a mobile service provider**

**Course**Unsupervised Learning

Applied K-Means clustering to identify different segments of customers for a telecom company. Based on analysis of the customer clusters, attributes that drive churn were identified which enabled the company to improve their services in a strategic manner.

**Skills and Tools**

K Means Clustering, Cluster Analysis, Consumer Behavior

Apr'18

**Product Recommendation Systems**

**Course**Recommendation Systems

This project involved building recommendation systems for Amazon products. A popularity-based model and a collaborative Filtering model were used and evaluated to recommend top-10 products for a user.

**Skills and Tools**

Collaborative Filtering, Recommender Systems, Python

Apr'18

**Movie Recommendation Systems**

**Course**Recommendation Systems

This Project involved building recommendation systems using user and item-based collaborative filtering methods on a movie rating dataset. Top k recommendations were made after evaluation of different models.

**Skills and Tools**

Recommender Systems, Python, Collaborative Filtering

Apr'18

**Analyzing customer churn problem in the Telecom industry**

**Course**Big Data on Spark

The purpose of this project was to identify the key reasons that drive customer churn for a telecom company. Spark (PySpark API) was used to analyze the relationship between customer churn & multiple independent variables affecting churn. Recommendations for improvements in service were suggested based on the results of the analysis.

**Skills and Tools**

Spark, Consumer Behavior, Python

Mar'18

**Case Study - Campaign for selling personal loans**

**Course**Supervised Learning and Ensembles

The project involved identification of potential customers for the personal loan product of a bank. Ensemble modeling techniques were leveraged to predict the propensity of a prospect to purchase the loan. This enabled the bank to devise targeted marketing campaigns to increase the conversion rate with a minimal budget.

**Skills and Tools**

Machine Learning, Ensembles Learning, Supervised\_Learning, Python

Mar'18

**Sentiment Analysis on Twitter in Real Time**

**Course**Real-time data Analytics

Applied sentiment analysis to understand user opinions & identify trends on different topics in using real-time streaming data from Twitter. Spark Streaming was used to analyze influential people or POI (person of Interest), trending topics and geo-based trends.

**Skills and Tools**

Spark, Real-time Data, Spark-Streaming

Jan'18

**Salary and Job Analysis using Python**

**Course**Python for Data Science

This project involved analysis of salaries dataset using python. Exploratory data analysis was done to find out different patterns of salaries and job scenarios in the industry and to find the correlation between different variables.

**Skills and Tools**

Python, EDA, Data Visualization

Oct'17

**Onion Sale Analysis, Visualization and Prediction**

**Course**Structuring & Visualizing Analytics problems

This project involved predicting the daily price of onions for the next 30 days for across states in India. Onion Price and Quantity analysis were done on State and Day Level using the scrapped data on Market Arrivals and Prices from The National Horticultural Research and Development Foundation (NHRDF) website (nhrdf.org). Used R to visualize the data in different forms to get insights.

**Skills and Tools**

Data Visualization, Predictive Modeling, R