Tianyi(Toby) Wei

● tianyi.wei@rhsmith.umd.edu ●202-808-1275 ● https://www.linkedin.com/in/tianyi-wei-rhsmith/

EDUCATION

University of Maryland, Robert H. Smith School of Business   **College Park, Maryland**
Master of Science in Business Analytics (**A STEM-Certified Program**), Current **GPA**:3.78/4.00
                                                                                          Expected December 2019
- **Courses:** Database Management Systems(SQL, Tableau); Data Processing and Analysis in Python(Python); Data Mining and Predictive Analytics(Machine Learning, R); Big Data and Artificial Intelligence(Deep Learning, Hadoop, Hive, Pig, Spark)
- **Honors:** Terrapin Scholarship

Shanghai University of Finance and Economics, School of Economics   **Shanghai, China**
Bachelor of Economics, Mathematical Economics   September 2014-July 2018
- **Honors:** Second-Class Scholarship of Outstanding Exchange Student

WORK EXPERIENCE

Wood Mackenzie   **Annapolis, Maryland**
Summer Research Intern(Data Science & Business Intelligence)   June 2019- August 2019
- Achieved automatically feed and mark weekly model results data into SQL Server database using Python to implement report automation.
- Designed a weekly auto-update Tableau story to track the coal market trends and published on Tableau Server for a team of 20 analysts internally use.
- Extracted user level data and customized dimensions and metrics data from Google Analytics API using Python.
- Conducted logistic regression based on extracted user level data using Python and identified key factors that influence a web visitor will return or not.

Orient Securities   **Shanghai, China**
Equity Research Intern, Strategy Group   April 2018-July 2018
- Built spreadsheet templates based on Bloomberg and Wind functions to track macroeconomic and capital market data, analyzed global market environment changes, and finished regular market reports.
- Assisted to finish research reports using different stock selection methods to choose potential stocks and industry.
- Searched different industry's information and presented research topics to senior financial analyst.
- Employed time series analysis method to forecast basic economy and finance index using R.

Haier   **Shanghai, China**
Data Scientist Intern, Global Internal Control & Audit   January 2018-April 2018
- Visualized competitive product analysis among seven different home appliances by developing interactive platform with more than 30,000 records via R shiny.
- Crawled 500 research reports from research institute website and generated high frequency words using Python.

Nielsen   **Shanghai, China**
Data Analyst Intern, Retail Plus MSP Department   July 2017-November 2017
- Compiled data from monthly questionnaires, provided monthly and quarterly analysis report for clients to better understand retail stores' performance and improve brand management.
- Designed reader-friendly questionnaires to mystery shoppers, increased 30% of recording efficiency and accuracy.
- Directed a team of 16 to arrange connection between mystery shoppers and clients to conduct mystery shopping.

PROJECT EXPERIENCE

Car Model Recognition   February 2019-May 2019
Course Project for Big data and Artificial Intelligence for Business (**CNN, keras, scikt-learn, numpy, pandas, scipy**)
- Read Stanford Car Dataset in the .mat format and process the data for each image gave according label in Python
- Processed and resized 8144 images to appropriate size using opencv2
- Trained a car recognition model using CNN on 6108 labeled car images based on VGG16 imagenet and reached a 30.0% testing accuracy compared to 0.6% baseline accuracy

Hospital Readmission Rate Prediction   February 2019-May 2019
Course Project for Data Mining and Predictive Analytics
- Cleaned 38,221 records of healthcare data with 26 variables: included using kNN imputation to fill missing values in patients demographic data, integrated levels in variables with multiple levels by different means
- Implemented machine learning algorithm included: Logistic Regression, Regression Tree, Random Forest, Bagging, Gradient Boosting and achieved 77.91% accuracy on testing data

TMDB Movie Box Office Analytics and Prediction   February 2019-May 2019
Course Project for Data Processing and Analysis in Python (**scikt-learn, numpy, pandas, re, matplotlib, seaborn**)
- Cleaned the TMDB dataset with 23 attributes and 3000 instances: included transforming JSON format fields, filling missing values and extracted a cast table to store massive cast and related movie information
- Provided descriptive analysis to explore movie revenue trends, most productive movie genres, actors, etc and generated WordCloud of most frequent words used in movie titles
- Performed feature engineering kept top 25 elements within each column to avoid massive dummy variables based on each element frequency and used Linear Regression to predict movie revenue with 9.5% error rate on testing data

Twitter Sentiment Analysis   January 2019-Febuary 2019
Course Project for Big data and Artificial Intelligence for Business (**keras**)
- Trained fully connect Neural Network based on 100,000 records of Twitter Sentiment Analysis dataset using Python and achieved 0.769 accuracy.

Providing Customized Solutions of housing properties to Prospective Students   October 2018-December 2018
Course Project for Database Management Systems (**SQL, Tableau**)
- Led a team of four and built a database management system using Microsoft SQL server to provide customized solutions of housing properties to prospective university students
- Leveraged Tableau Map as an interactive frontend user interface to show customized results and summary of the properties in a user-friendly way

Customer Classification Using K-means Algorithm   November 2017-December 2017
Course Project for Data Mining
- Analyzed 60,000 passenger records of data with 20 features of an airline company via Weka and identified three key customer segments using K-means algorithm to develop new marketing strategy.

TECHNICAL SKILLS

- Proficient in Python, R, SQL, Tableau, Hadoop, Hive, Impala, Pig, Spark