Hongpeng Jin

(469) 543-7960 hopenjin@gmail.com linkedin.com/in/hongpeng-jin

SUMMARY

THIS IS A DATA PERSON, AND HE DREAMS TO WORK AS AN ANALYST.

EDUCATION

The University of Texas at Dallas, School of Management Dallas, TX
M.S., ITM, Business Intelligence and Analytics track Aug. 2016 – May. 2018
Nanjing Normal University, School of Geography Sciences Nanjing, China
B.S., Tourism Management, E-Business for Tourism track Sept. 2012 - Jun. 2016

ANALYTICS OR TECHNICAL SKILLS

Statistical Models: Regression Analysis, Hypothesis Test, Correlation Analysis, Cross-tabulation Analysis, ANOVA, Logistic Regression, Count data models
Machine Learning: Decision Trees, Random Forest, Naive Bayes, Support Vector Machine, Clustering(K-mean/ Hierarchical), Principal Component Analysis, NLP(TF-IDF/Word2Vec)
Programming: **Python**(Pandas/Sklearn/Seaborn), R(ggplot2, ctree, lmtest), **SQL**, Java, HTML, CSS, JavaScript
Tools: **Tableau**, **Power BI**, **Excel(VLOOKUP/PivotTable/Macros/VBA)**, **MS SQL**, SPSS, ERP system(Epicor)

Certifications: Advanced Google Analytics, Google Analytics for Beginners, Master SQL for Data Science

BUSINESS EXPERIENCE

Samsung Electronics America Plano, TX
QA Engineer(MQL IMS team) Mar. 2019 – Present

- Tested Data throughput, Voice Quality Testing, comparing throughput performance, and identifying major call-related problems (call drop, interference issue, handover failure scenario, RF performance scenario etc.) and other technical issues.
- Expertise in verifying call connectivity and device performance under different voice and data network coverage like LTE, VoLTE, CDMA, UMTS & GSM
- Performed Different Sanity, Regression, Smoke, Integration, UI, Performance, Stress and UAT Test and used excel sheet and PLM as project management tool to track the case execution details and prioritize testing activities based on requirements
- Created effective and detail oriented bug reports for many different types of bugs during project for QA and developers to solve the issue.

ZTE USA Inc. Dallas, TX
QA Analyst / Software Development Engineer in Test Apr. 2018 – Mar. 2019

- Executed test cases, including black box testing and manual handset tests focusing on user requirements
- Adjusted test cases (like Built new cases or deleted outdated cases) in **test cases database using SQL**
- Wrote weekly report or project reports to summarize the situation of manual test result(such as, rate of pass, rate or fail, the number of the cases we tested and left, reasons of fail test cases, plans for this fail cases, etc.)
- **Analyzed data and logs** in fail test cases to find out the reason or the most likely cause, and then collaborated with R&D team to decide further test, verification and analyzation
- Wrote and maintained **python scripts** for Devices Automatic Testing (AIO project) including 43 scripts for Tacfone test cases and 71 scripts for Verzion using Python/Android Studio/GitLab/Jenkins/UiMap
- Built testing software tools to optimize previous test processes, such as, an integrated testing tool with GUI for reducing CMD command operation, a scraping tool for catch data from some website, etc. **using Python/ Selenium/Pyinstaller/Wxpython**

IT department of FACSS-UTD, UT Dallas Dallas, TX
Web Database Developer (Intern) Aug. 2017 – Dec. 2017

- Assisted to build conceptual model (ER diagram) for user management and implement database using **MySQL**
- **Visualized the data** in our server, such as, the trend of user registration, user online behavior characteristics, daily kernel active users, and then make dashboards and stories **using Tableau**
- Developed Web admin interface to help our activities specialist to **operate database online** (involving insert, delete, update or query user info) using PHP/MySQL/HTML/CSS
- Developed functions or web boards for our organization (like activities' or festivals'(e.g. new year) special demands) or for students(like the online market board buy or sell used stuff) using PHP/HTML/CSS

Nanjing Duohuo Network Technology Corporation Nanjing, China
Data analyst (Intern) Aug. 2013 – Mar. 2014

- Built and maintained **a customer dataset from various sources**, including each university's user data collected by our campus agents, the data on our social media pages, and the data in our app server**, using Excel(vlookup/Pivot Table)**
- **Filled the missing data**, set one or several data indexes to represent some complicated data. To get better

data quality, help our campus agents to get more useful and summarized information
- Analyzed the data to dig out useful insights, including **calculating their statistical metrics** (mean, median, variance, etc.), and then made visual weekly report or presentation (including **stacked bar, box plot**, etc.) to help our campus agents and managers know students better and compete with other companies effectively

UT DALLAS DATA ANALYTICS COURSE PROJECTS(SELECTED)

Sentiment Classification of Yelp Restaurant Reviews Feb 2018 – Apr. 2018

- Parsed almost **1 million** restaurant reviews across 10 states (AZ, OH, TX, etc.) which was filtered from Yelp **JSON** dataset using **Python Pandas**
- Pre-processed review texts (remove HTML tags, punctuation, lowercase) and then **tokenized** them (delete stop words and lemmatize within word tokens)
- Transformed **unstructured data** (review texts) into the dense vectors (60/100/120 dimensions) using **Word2Vec/Doc2Vec** techniques which are based on **Neural Network Models** and **Python**/**Gensim** library
- Built a couple of classifiers with **Python Scikit-learn** library, such as **Logistic Regression**, **Support Vector Machine**, and **Random Forests**, utilizing the vectors from **Word2Vec** as the predictors
- Optimized and compared the classifiers considering different **evaluation metrics**, like **accuracy**, **precision**, **recall**, **ROC curve**. Also, the hypermeters (vector dimensions, windows size, etc.) in Word2Vec were adjusted according to the performance of classifiers

Customer Booking Prediction from Expedia Clickstream Data Sep. 2017 – Nov. 2017

- Pre-processed 50 thousand customer demographic and web visiting data, including **features scaling** and imputation of missing values
- Reduced the 45 features to 6 to 10 factors using **PCA methods**
- Resample the training data to address the imbalanced dataset (13.4% booked)
- Built different models, **Logistic Regression**(**L1**, **L2**), **Support Vector Machines, Random Forest**, and **AdaBoost**, for the tasks of booking prediction.
- Optimized parameters of different classifiers, and selected the best classifiers using cross-validation methods

The Popularity Analysis of Online News at www.mashable.com Jan. 2017 – Mar. 2017

- **Preprocessed the dataset** including 39 thousand records and 61 attributes, such as, deleting useless attributes (like URL), **Scaling**, **One Hot Encoding**, and then converted this into a binary classification problem(large shares or small shares)
- Trained and evaluated **Logistic Regression**/ **SVM**/**Decision Tree**/**Neural Network Models** Using **Python Scikit-learn** library
- Improved the accuracy using **Parameter Tuning**, **Feature Selection**, **PCA**, **ICA**, etc.

COURSES AT UT DALLAS

Applied Machine Learning(**Python**/**R**), Big Data Analysis(Hadoop/Hive/Spark), Advanced Business Analytics with R, Data Management(T-SQL/SQL Server), Statistics and Data analysis(Excel), Applied Econometrics (R), Data Visualization(**Tableau**/**QlikView**), System Analysis and Project Management(MS Visio/Lucidchart), Financial Management, Business Economics, International Finance, Economics Principles, Accounting for Management(Excel)

ADDITIONAL INFORMATION

Eligibility: Work for any employer in the U.S. for up 36 months without sponsorship
Personal Website: https://hopenjin.github.io/myBetterResumePage