Pei Yingchi

Resume

Blk 149 Mei Ling Street, Singapore, 140149

☏ +65 86699823

✉ yingchi.pei@gmail.com

🖥 www.peiyingchi.com

Education

2018 – 2020	**Computer Science, Artificial Intelligence (Master)**, *NUS*. ○ CAP: 4.83/5.0
2013 – 2017	**Business Analytics (Bachelor)**, *National University of Singapore (NUS)*. ○ CAP: 4.91/5.0, Honours with Highest Distinction ○ Winner of Lee Kuan Yew Gold Medal (Top student of the batch) ○ Five times Dean's Lister

Work

2019 Apr- Present — **Data Scientist**, *Indeed.com Ltd.*, Singapore.

Work in salary estimation and career experience team.

- Build models (XGBoost) for salary estimation using structured and unstructured data like job description text.
- Build Python modules for text summarization and ranking to generate representative content items, using NLP techniques such as TextRank and Word2Vec.
- Prototype exploration and exploitation pipeline for dynamic ranking.

2018 Aug - 2019 Jan — **Data Scientist**, *Bitmain Technologies Ltd.*, Singapore.

Work in btc.com's data team.

- Provided data insights for cryptocurrency mining platforms and blockchain explorers using Airflow scheduled Spark jobs.
- Developed a transaction fee prediction engine using Neural Networks and Generalized Linear Models, building the end-to-end process from acquiring real-time data (Python parser with Redis and MySQL) to training and evaluating models.
- Generated internal data reports using Spark SQL, Hive and graph databases like neo4j.

2017 Jul - 2018 Jul — **Data Scientist**, *DataSpark, Singtel*, Singapore.

Work in the application team.

- Researched on footfall analytics with telco data using machine learning algorithms such as Naive Bayes, Logistic Regression, and Random Forests. Implemented and productionized models into our data analytics platform using Python. Submitted two research papers based on that with one published.
- Designed and develop the network planning application for telco operators to reduce upgrading cost while improving customer experience. The application was built with Scala and deployed in a big-data environment with Hadoop and Spark.

2016 May-Nov — **Data Scientist (Intern)**, *DataSpark, Singtel*, Singapore.

Part of the product team at DataSpark.

- Conducted geolocation data analysis projects to undercover new features and improve model accuracy by running Hadoop and Spark jobs; implemented reproducible code using R Markdown and Python for the projects.
- Built interactive data visualizations (Web apps) using JavaScript, Node.js and React for internal and external clients.

2015 May-Aug — **Market Research Analyst (Intern)**, *Millward Brown*, Shanghai, China.

Work on a long-term marketing research project for Budweiser.

- Prepared Budweiser's 2015 Q1 report which was well received by the client; discovered unusual patterns from data and initiated deep dive research to find explanations.
- Collected and complied the consumer survey data weekly using SPSS Survey Reporter.

Publications

2017 Ng, Y., **Pei, Y.**, Jin, Y. (2017).
"Footfall Count Estimation Techniques Using Mobile Data", 2017 IEEE 18th International Conference on Mobile Data Management (MDM), Daejeon, South Korea, 2017, pp. 307-314. doi: 10.1109/MDM.2017.49

Projects

2015 **Bankruptcy Prediction Using 10-K Reports**, *Bachelor Capstone Project*.
- Gathered data from 3 large financial databases; investigated different methods such as resampling and cost-sensitive learning to deal with the imbalanced dataset.
- Incorporated text-mining features such as fog readability and sentiment score to the traditional prediction model. The best model achieves testing AUC score 0.9.

2014 **Twitter Follower Analysis For "@TopShop"**, *Data Mining Project*.
- Collected data of 20,000 Topshop followers using R through the Twitter API.
- A classifier was built to classify tweets into categories. Compared various methods such as Logistic Regression, Random Forests, SVM, and K-Nearest Neighbours.

Activities

2014 – **Teaching Assistant**, *IS1112 E-Business Essentials, NUS*.
2015
- Crafted over 100 questions and provided assistance for the students throughout the course.

2014 **Program Organizer**, *Egyptian Society for Integrated Development*, Cairo, Egypt.
Summer
- Planned and conducted activities such as lectures and field tours for underprivileged children.
- Survived a bomb attack.

2013 – **Vice President**, *Chinese Society Chinese Orchestra, NUS*.
2014
- Organized an annual welcome event, which resulted in a membership increase of 45%.
- Expanded the influence of CSCO by organizing an annual music concert with more than 200 audiences. Initiated new events including music talks and cultural exchange programs.

Additional Information

Languages Fluent in English and Mandarin. Know basic German (A1 level)

Skills Statistics background with 12 statistics modules taken. Programming background with knowledge in Python, Scala, R, Spark, SQL and SAS

Interests Music: Play the piano, the cucurbit flute and the Erhu

Sports: Taekwondo Black Belt and Karate