ZHELU（Jerry）MAI

Denton, Texas | 940-595-9706 | ZHELUMAI@my.unt.edu | Personal Website | GitHub

Education

University of North Texas	Denton, TX
Master of Science, Business Analytics	*Jan 2021 – Nov 2022*
Chu Hai College of Higher Education (CHCHE)	Hong Kong, China
Bachelor of Business Administration, Finance	*Sep 2012 – Dec 2016*

Projects

Amazon Review NLP Project   *Spring 2022*

- Cleaned data by removing URLs, emojis, non-ascii characters using **regular expression**
- Extracted linguistic features based on tokenization, POS tagging, and parsing sentence via **spaCy**
- Tuned parameters via 10-fold cross-validation. SMOTE was used to do **oversampling**

Personal Loan Project   *Fall 2021*

- Abstracted new key features like household incomes from digging zip code using **uszipcode**
- Eliminated one highly correlated feature and some outliers through a well-developed **EDA**
- Obtained 0.96 at Marco f1-score by the designed random forest model

Housing Prices Competition in Kaggle   *Summer 2021*

- Gained 5% improvement in accuracy by removing the outliers based on the boxplot via **Matplotlib**
- Merged similar features for adjusting the skewed distribution that achieve 2.7% improvement in performance
- Ranked **Top 5%** in the leaderboard with the tuned gradient boosting model

Working Experience

Operation Officer   *Nov 2019 – Dec 2020*
Bank of China (Hong Kong)   Hong Kong

- Built automatic programs via **VBA** to deal with 10 daily reports and parts of weekly reports, saving 19% working time, therefore, I was assigned to help the payroll team in managing payroll issuance every day

Operation Analyst   *Nov 2017 – Jul 2019*
Dollar General Global Sourcing Limited   Hong Kong

- Improved productivity by 25% through developing **VBA** automatic programs connecting to the **Access**, generated 5 daily reports and 7 weekly reports. By doing so I had more time to address ad-hoc projects, etc.

Skillset

Programming Languages:	**Python**, SQL, VBA, R, HTML5, JavaScript
Data Visualization:	**Tableau**, Seaborn, Matplotlib, RStudio, QlikView
Database:	**MySQL**, SQL Server, MongoDB
Big Data:	Hadoop, Spark, Hive, Pig
Tools:	scikit-learn, Pandas, **spaCy,** GitHub, PyCharm