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 – Dec 2022*

Chu Hai College of Higher Education (CHCHE) Hong Kong, China

Bachelor of Business Administration, Finance *Sep 2012 – Dec 2016*

Project Experience

Amazon Review Classification *Spring 2022*

- Cleaned data by removing URLs, emojis, and 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**

University Information Retrieval Chatbot *Spring 2022*

- Scraped FAQs on the university websites with **BeautifulSoup**
- Cleaned and parsed conversations using **NLTK** and **spaCy**
- Extracted textual feature via **GloVe** word embedding and **Seq2Seq LSTM** using **TensorFlow**

Personal Loan Acceptance Classification *Fall 2021*

- Feature engineered on household income zipcode using **uszipcode**
- Eliminated one highly correlated feature and some outliers through a well-developed **EDA**
- Obtained 0.96 at macro F1-score by the designed random forest model

Housing Prices Competition *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 achieved 2.7% improvement in performance
- Ranked **Top 5%** in the leaderboard with the tuned gradient boosting model

Work 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

Skillset

Programming Languages:	**Python**, SQL, VBA, R, HTML5, JavaScript, AIML
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, BeautifulSoup, **TensorFlow**, Linux