WANG RUIJUE

Mobile: 85871467 Email: wangruijue527@outlook.com

EDUCATION

National University of Singapore (NUS) **Aug 2020 – May 2024**
- Bachelor of Science (Data Science and Analytics), Honours
- Second Major in Economics, Minor in Computer Science

SM2 Scholarship **Aug 2019 – May 2024**

INTERNSHIP

Beavo Investment, Strategy Research Intern **May 2022 – Aug 2022**
- Converted, cleaned and merged raw data from zip files to standard pickle files
- Modified the previous linear model for predicting the stock price in the next 1 minute via feature selections, time period selections and regularizations. Initiated a new model (combination of logistic model and linear model) for prediction.
- Improved the efficiency of model code and strategy code via parallel design with the help of other computer engineers, and ran codes in windows and Linux environment

PROJECT

Consulting project of Feasibility Analysis and Financial Forecast of Blind Box **Dec 2020 – Jan 2021**
- Did research on industry, market and Part A company, interviewed some consumers, and used SMOT model to analyze the present opportunity and challenges
- Based on economic calculation about the cost and income, and compared with current business model, provided suggestions on improving business model and strategy plans for next 3 years.

Machine Learning project of comparing model performance **Feb – Nov 2022**
- Image data with only ML methods: Compared the performance of Logistic, Ridge, and Lasso model on predicting the gender of given image. Tried different preprocessing techniques \such as varying resolution, channel dimension, image color, denoising and contrast enhancement, partial face. Based on accuracy and AUC curve to analyze the effectiveness of different methods.
- Tabular data with ML and DL methods: Compared the performance of baseline machine learning (Random Forest, XGBoost, LightGBM) with deep learning (Multi-layer Perception, radial basic function network, attention-based model) on tabular data. Analyzed and concluded that ML (0.74) may perform better than DL (0.51) on small and simple tabular datasets.
- Text data with ML and DL methods: Compared the performance of ML (logistics model, SVM) and DL (CNN, LSTM) on sarcasm detection. Tried different NLP preprocessing methods, such as expanding contractions, removing stopwords, tokenization, TF-IDF and undersample. Did EDA on Subjectivity and Polarity Analysis, Profanity relationships. Tested the hypothesis by using other features like the length, number of unique words. Gave limitations and improvements of the methods.

Front-end project of AI assistant app **Oct 2022 – Nov 2022**
- Designed a chatbot that allowed medical sales representatives to find brief answers to their questions, check the confidence of the answers provided, and explore other possible answers
- Provided "uploading" tab to update our dataset in real time based on the file uploaded by users and "stats" tab to show success rate of users' interested device so that they can make comparison easily
- Docked app to connect UI interface with the backend datasets and model, used AWS to test application.

ADDITIONAL INFORMATION

- **Technical Skills:** Python, R, SQL, Microsoft office, Java (Basic), MATLAB (Basic), Hadoop (Basic), Spark (Basic), BDA Certificate (Junior)
- **Languages:** Fluent in Mandarin and English (Both spoken and written)
- **Co-Curricular Activities:** NUS Chinese Drama Club (General Secretary), NUS Statistics and Data Science Society (Publicity), NUS Chinese Students' Committee (Events), NUS Viva Latinus
- **Interests:** Latin & Ballroom Chinese Opera & Drama, Broadcasting & Hosting and outdoor activities