Yifei Wu

Phone: +1(614) 9616980 | E-mail:yw924@georgetown.edu | Linkedin:https://www.linkedin.com/in/yifei-wu-20576b229/

EDUCATION

08/2022-05/2024	**Georgetown University** Current GPA:3.95/4.0	Washington DC,USA

- Major: Master of Science in Data Science&Analytics
- Selected Courses: Prob modeling/Stat Computing,Database and SQL,Data Science and Analytics,Advanced Data Visualization, Statistical Learning for ANLY, Time Series, Big Data and Cloud Computing

Scholarship Owner （1 time）

09/2019-05/2022	**The Ohio State University** Overall GPA: 3.58/4.00	Columbus, USA

- Major: Bachelor of Science in Mathematics (Financial Math Specialization)
- Dean List Owner（3 times）
- Selected Courses: Calculus III, Foundations of Higher Mathematics , Probability, Introduction to Math Statistics I , Introduction to Math Statistics II, Differential Equations and Their Applications, Introduction to Financial Math,Linear Algebra, Math Statistics, Scientific Computing, Partial Differential Equations for Science and Engineering, Financial Economics for Actuaries

SKILLS

Languages: Chinese (Native), English (Fluent)
Programming: Python(pandas,numpy,seaborn,scikit-learn,scipy),R,SQL,C++,HTML,MATLAB.
Tools: Excel,Access,Tableau
Soft Skills: Communication, Business acumen, Workplace Etiquette

INTERNSHIP EXPERIENCE

05/2021-08/2021	**Baidu, Inc.** *(Data Analysis Intern)*	*Shanghai, China*

- Crawled business information and data of the new energy vehicle enterprises and conducted a competitive analysis which brought a clear view for the R&D Department to better understanding their competitors
- Researched and analyzed the markets of e-commerce platforms, including Taobao and Pinduoduo, to write a marketing analysis report,saving the group time for more in-depth research.
- Cleanned and rearranged the data that we collected(Company information, Marketing Performance,Margin Value,Sales),combined 35features from 13 different tables, conduct feature engineering such as normalization.

05/2023-08/2023	**Topsperity Securities.** *(Analyst)*	*Shanghai,China*

- Mining private public placement product data in the database, organizing and cleaning the data, analyzing the product data, and finally ranking and screening the products on the line.
- Conducted trading tests on brokerage firms' back-office trading systems; wrote modifications to quantitative trading strategies using quantitative trading systems and optimized some strategies to meet requirements.
- Perform attribution analysis through machine learning, develop mining factors based on product performance, market data, etc., find effective factors and combine them, write quantitative strategies for backtesting, and reproduce some research reports.
- Prepare departmental weekly reports and research reports.

RESEARCH EXPERIENCE

04/2021-06/2021	**Data Science and Deep Learning in AI(Image Recognition)**	Columbus, USA

Undergraduate Researcher

- Trained a model with CV2 which canbe used for the parking lot to realize the recognition

of the vehicle licence plate number, and we can reach an accuracy around 78%.

- Processed image data, converting pixels into numeric tensors for subsequent analysis and modeling. Through CV2, pytesseract to train model which can access the text through the Image we supply, and detect the language of the text.Such model can achieve a accuracy around 64%.

01/2023-05/2023	**Us Bank Performance Analysis (Data Visualization)**	Washington, USA

Group project

- Using the data from FFIEC(Federal Financial Institutions Examination Council Cntral Data Repository) , grab all kinds of data about banks over the past six years, explained through each of the eight different visualizations, comparing four different questions
- Using Altair,plotly,matplotlib and other packages to create the visualization digram such as interactive plot,linked plot, time series plot,scatter plot,distribution,maps which are able to visualize and explain problems.
- A visual analysis of the operations of all banks in the United States over the past six years to determine the reasons for some of the bank failures.

08/2022-12/2022	**Championship Predictions**	Washington DC,USA

Personal project

- Using the Twitter api to obtain text and user data, train four classification models to predict the winner of the 2022 World Cup and to determine the impact of messages from different Twitter users on the final result. The confusion matrix was plotted using seaborn and scikit-plot to determine the best model. We used HTML/CSS and Quarto to create a web page to present the results of the project.
- Nltk was used to count the frequency of tweets and specific words written by users, apriori was used to find the relationship between words to find out the relationship between certain words such as "country" and "champion", and finally to draw a conclusion.