Yanyu Chen

yanyuc@andrew.cmu.edu | (781) 654-9686 | linkedin.com/in/yanyuc

Education:

Carnegie Mellon University   **Pittsburgh, PA**
M.S. in Data Analytics for Public Policy & Management   **May 2023**
Courses: Software Construction, Web Application Development, Data Structures, Database Management, Deep Learning
Peking University   **Beijing, China**
B.S. in Economics   **June 2021**

Skills:

Programming Languages: Python, Java, TypeScript/JavaScript, HTML&CSS, PL/SQL, SQL, R, Stata
Frameworks & Libraries: Django, React, AJAX, jQuery, PyTorch, PySpark, Scikit-learn, Plotly, CoreNLP, Pandas, SciPy
Tools & Techniques: Git, Maven, NPM, MySQL, Oracle DB, Linux, AWS, GCP, Apache, GitHub CI/CD, MS Office Suite

Professional Experience:

Athena Accessible Technology   **San Mateo, CA**
Machine Learning Engineer Intern   **June 2022 – August 2022**
- Designed and implemented data analytics pipelines to overview the time series eye gaze movement data efficiently
- Developed an internal library for the company to store functions used in daily productions with detailed documentations
- Improved the accuracy of the eye gaze reading classification model by 20% using newly selected features and models by feature engineering such as feature extraction, selection, and model experiments

CITIC Securities Company   **Beijing, China**
Research Intern   **May 2020 – March 2021**
- Independently finished 3 analysis reports; participated in research of petrochemical related data using valuation models
- Developed a Python web crawler to gather oil and chemicals prices from static websites; transformed the information into useful data for the team to conduct research on different chemical companies
- Refined and analyzed datasets for variable correlation with Stata to more accurately predict trends of long-term oil price

Selected Projects:

Online Restaurant Reservation and Order System   **[*Python, Django, JavaScript, React, AWS, MySQL, Git*]**
- Worked in a group of four to construct a full-stack restaurant reservation and order system based on Django; deployed the system on AWS EC2 and it is available for use by real business with options like adding dishes and staffs
- Designed Django models to store various information and relationships among entities such as dishes and reservations
- Utilized React.js to build a floating shopping cart that updates in real time when dishes are added to an order

WebApp Spatial-Temporal Extraction and Analytics Framework   **[*Java, TypeScript, React, Node, NLP, Git*]**
- Collaborated with two teammates to implement a web-based data analytics framework to extract and display time and space data from various sources including plain texts, kept its extensibility for easily adding new plugins
- Designed APIs for data and visualization plugins, sketched domain and object diagrams for program structure
- Built a dynamic webpage UI using React for selecting plugins and rendering data visualizations with Plotly.js

Product Review Keywords Extraction Based on Aspect Sentiment Analysis [*Python, PyTorch, NLP, Deep Learning, Git*]
- Constructed neural networks to extract keywords triplets from over thousands of different product reviews
- Reproduced the key idea from a research to form a natural language processing transformer with grid tagging scheme; conducted experiments with data augmentations and a new encoder to achieve an improvement of 6 on the F1-score

Worldwide Life Expectancy Analysis: Prediction, Classification and Clustering [*Python, Scikit-learn, Machine Learning*]
- Built supervised and unsupervised models such as Linear Regression, Random Forest, Logistic Regression, K-means Clustering to gain insights from a cross-sectional healthcare dataset with more than 20 features