JUN ZE HE

(626)-592-2842 | junzehe977@gmail.com | https://www.linkedin.com/in/junze-he-9146a3267/ | https://github.com/JunJul

EDUCATION

University of California, Los Angeles (UCLA), Los Angeles, CA — Expected June 2025
Bachelor of Science, Data Theory

- GPA: 3.324
- Relevant Coursework: Python, R programming, Data Analysis, Computational Statistics, Mathematics for Data Theory, Statistical Optimization in Machine Learning, and Data Mining

Break Through Tech at Cornell University, New York, NY — May 2024 – Aug 2024
Certificate in Machine Learning Foundations (eCornell)

- Relevant coursework: Machine Learning, Computer Vision, Natural Language Processing, Language Models

CAMPUS INVOLVEMENT / EXTRACURRICULAR ACTIVITIES

National Student Data Corps @ UCLA, Los Angeles, CA — Jan 2024 – June 2024
Project Lead

- **Led** 5 teammates to complete the credit score and the heart disease classification project by Python
- **Managed** data understanding, feature engineering, modeling, and model evaluation in both projects by Trello
- **Accomplished 90%** true positive rate by XGB in the Credit Score classification project and 95% true positive rate by Random Forest in the heart disease classification project

DataRes @UCLA, Los Angeles, CA — Jan 2024 – June 2024
Data Journalist

- **Proposed** research questions on LACrime and California K-12 disciplinary dataset from Kaggle in a team of 5
- **Extracted** insights using Python (Seaborn, Matplotlib, Pandas), **presented** my findings to other club members, and **published** 2 journal medium articles to explain my findings

PROJECTS

Honkai Star Rail's Player Reviews Analysis, Los Angeles, CA — Aug 2024 - Sep 2024

- **Spearheaded** end-to-end NLP analysis of 19,000+ player reviews for *Honkai Star Rail* to identify actionable insights for game improvement
- **Collected** data using Python and SerpAPI, streamlining review aggregation from the Google Play Store
- **Utilzied** LatentDirichletAllocation, Top2Vec, and BERTopic to cluster unstructured text and **uncovered** game difficulty spikes (35%), login instability (28%), and controller support requests (20%).

WORK EXPERIENCE

Accenture, Los Angeles, CA — Aug 2024 – Dec 2024
AI Studio Program Intern, Machine Learning Engineer Intern

- **Developed** a Technology News Insight Engine to uncover industry trends from various technology articles, supporting business development for tech startup clients
- **Automated** text mining and article categorization pipelines using NLP techniques, Groq API, and BERTopic for efficient data processing
- **Designed** an interactive recommendation system leveraging Retrieval-Augmented Generation (RAG) to answer client queries, reducing client query resolution time by 15%

Mt San Antonio College, Walnut, CA — June 2022 - Aug 2022
Research Assistant

- **Assisted** a computer science lab in University of California, Irvine in **integrating** a health watch with their Institute of Future Health platform
- **Analyzed** the watch's data collection accuracy through data visualization and statistical measurements in Python
- **Demonstrated** minimal variance between watch data and medical devices, with a 2.96 mean absolute error (MAE), 5.34 mean squared error (MSE), and statistically significant p-values (≤ 0.05)

SKILLS

- Programming: Python, R, SQL, HTML, CSS, JavaScript
- ML/Could: TensorFlow, PyTorch, AWS, Azure
- Tools: Git, GitHub, Google Suite, Microsoft Office, Trello
- Languages: Mandarin (Fluent), Cantonese (Fluent)