Jianyang Zhou

831-295-3884 | jz432@duke.edu | Durham, NC

EDUCATION

Duke University

- M.Eng in Computer Science (Data Analytics & ML Concentration) **08/2023 - 05/2025**
- GPA: 4.00/4.00

Cornell University

- B.S in Operations Research & Information Engineering **09/2020 - 12/2022**
- GPA: 3.53/4.00
- Coursework: Data Structure and Algorithms, Optimization, Machine Learning, Statistical Data Mining, Simulation Modeling and Analysis, Probability and Stochastic Process, Supply Chain Analysis

WORKING EXPERIENCE

Research Assistant | Duke University | Durham, NC **08/2023 - Present**

- Using reinforcement learning to find the optimal solution of the treatment schedule for HIV patients
- Improved the Q-learning algorithm for the model, and increased cure rate by 20% after simulation

Operations Research Intern | Norfolk Southern | Atlanta, GA **01/2023 - 04/2023**

- Extracted and cleaned 1,000,000 data of cargo network in IBM DB2, developed automated report scripts of cargo volume's trend for 50 terminal yards using Python, SQL, and Spotfire, supporting developers in improving prediction model
- Analyzed the cargo arrival time from customers on an hourly basis, considered multiple factors' impact, including origin, destination, shipment type, etc., using the random forest to classify the time pattern
- Implemented usability test report for yard scheduling system

Data Scientist Intern | Kingsoft Cloud | Beijing, China **05/2021 - 08/2021**

- Evaluated SaaS platform user portrait by K-means clustering model and PCA to classify 500k users into 4 tiers with different objectives and backgrounds, visualized the data by Tableau in 20 different views
- Implemented ETL for 10000 business competitor's bid-winning projects, analyzed data in multiple dimensions by k-means, SVM, and random forest in Python with Pandas, Sklearn, Matplotlib packages

PROJECTS

Scheduling Patrol Officers in Tompkins County **08/2022 - 12/2022**

- Researched algorithms and evaluation models to enforce traffic laws on the roads in Tompkins County
- Used the K-means clustering algorithm to group these accidents overtime and determined two-hour time intervals where accidents have the highest probability of happening and made a schedule for patrol officers
- Developed an evaluation model to test the algorithm considering the priority of different types of incidents and how likely these incidents would be reduced due to the presence of an officer

Domestic Flight Delays and Cancellations Analysis **01/2022 - 05/2022**

- Built models in Python to investigate a dataset of 5,000,000 flights delays and cancellations in the U.S
- Predicted future airplane delays with A/B testing, exponential smoothing, seasonality, and other machine learning techniques, successfully predicted the delay time in next year with 90% accuracy in 3 minutes
- Published advice for passengers regarding choosing flight through intuitive visualization by SQL and Tableau, and demonstrated the correlation between airplane delays and factors such as weather, time, etc

NY State Helicopter Emergency Transportation System **10/2021 - 12/2021**

- Led a five-student team to model and implement a New York State helicopter rescue system
- Analyzed historical data, and performed thinning method simulation to generate random data
- Developed an algorithm using a heuristic method to compute the optimal location arrangement for 5 to 12 helicopters. Optimized the average response time to 0.42 hours, and the available rate at 79.5%

SKILLS

Programming: Python, Java, C, C++, R, HTML, CSS, JavaScript, SQL
Tools: Unix, Git, Eclipse, XCode, VS Code, LaTeX, DB2, Teradata, Gurobi, AMPL, Tableau, Spotfire