YUNXIN HONG

+1-2065655339 | yunxinh@outlook.com | https://github.com/rattlesyylz | Seattle, WA
linkedin.com/in/yunxin-hong-907369247/

EDUCATION

University of Washington **Sept 2023 - Jun 2025 (Expected)**
Master's, Data Science GPA: 4

University of Washington **Sept 2020 – Jun 2023**
Bachelor's, Applied Mathematics GPA: 3.94

- Relevant Coursework: Database System, Machine Learning, Deep Learning, Algorithms, Calculus, Probability, and Statistics

SKILLS

- Programming Skills and Tools: Python, MySQL, R, Java, HTML, CSS, JavaScript, Tableau, Power BI, Azure, GCP, AWS and Git
- Methodologies: Statistical Analysis, Predictive Analysis, Data Preparation, Data Visualization, Probability, and Machine Learning

PROFESSIONAL EXPERIENCES

Data Scientist Intern **Shanghai, China**
NIO | Autonomous Driving Department Jul 2023 - Sept 2023

- Built a construction scenario detection model with **CNN algorithm**, **Logistic Regression**, and **Decision Tree** model in **Python**, enhancing the identification accuracy from 73% to 81% and reducing the average time spent in the process by 90%
- Mined ~7000 data with **MySQL** queries to assess traffic cone scenarios in autonomous driving, built **statistical models** for 3 algorithm, collaborated with Algorithm Team to design and implement the optimal solution
- Created an automated online database on **Azure** to calculated **statistical matrices** and replace manual quality inspection workflows, resulting in the elimination of redundant manual checks and a substantial increase in time efficiency by over 80%
- Developed Machine Learning Models (**XGBoost**) to categorize ~300 user feedback, designed and implemented interactive dashboard with **Tableau**, analyzed and solved problematic user cases, and raised construction identification accuracy from 80% to 88% by improving data wrangling process and optimizing in-car video recording time

Data Scientist Intern **New York, NY**
Trove Quantitative Technology | R&D Department Jul 2022 - Sept 2022

- Built machine learning models (**logistic regression**, **Ada Boost**, **random forest**, **model blending**) to predict success likelihood of funding application, improving baseline accuracy by 30% through leveraging hyperparameter tuning
- Established a database and constructed a model to analyze US stocks using several indicators following Ned Davis' method and the Fab Five Component algorithm in **Python**, leading to a 15% increase in revenue
- Crawled 20-year data for 5 stocks, built time-series model **(LSTM)** and **anomaly detection** model to predict future returns and risk, optimized the risk management strategies, and improved the robustness of investment models by 65%
- Managed a team of 20 research assistants in daily exploration of factors, in-depth analysis of research reports, and execution of code for convertible bond strategy program, and found 3 new factors in the convertible bond strategy, resulting a 30% profit increase

Data Analyst Intern **Beijing, China**
Sky Limit Entertainment | SoReal VR R&D Department Jul 2021 - Sept 2021

- Developed **data pipeline** through **Python** to conduct weekly analysis of 15 games' usage frequency to curate three-game combined ticket packages on app, increased average daily ticket sale revenue by 30%
- Applied the **K-means** algorithm to analyze user profiles from 4 VR amusement parks, and estimated user profiles for the upcoming Shanghai Disney VR Park to design business strategies suitable for the 30+ age group with children
- Executed an **A/B test** for instant member registration at Chengdu's VR amusement park, with one location offering a discount coupon, leading to a ~500 increase in average monthly visitors compared to the location without the coupon incentive
- Collaborated through cross-functional communication with software engineers to improve the app, resulting in acquiring 100+ new members within a week

PROJECT EXPERIENCE

Deep Learning Research Assistant **Seattle, WA**
University of Washington eScience Institute | Carbon Neural Network Model Jul 2022 - Jan 2024

- Developed a mobile application to accomplish image recognition by using **MobileNets** to detect objects and output CO_2 value by linking backend deep learning models, providing users with detailed data of CO_2 consumption
- Trained 10 deep learning models (**Seq2seq** & **autoencoder)** with tunning hyperparameter and combining models to calculate CO_2 value, analyzed models' performance with metrics, where the application will be used in NGO and companies