SHIYU LIU

shiyuliu001@gmail.com • (949)294-0851 • linkedin.com/in/shiyu-liu1

EDUCATION

M.S. in Data Science, Brown University   To be conferred May 2021
• GPA: 4.0 out of 4.0
• Coursework: Machine Learning, Data Engineering, Statistical Learning, Deep Learning, Computer Vision.
B.S. in Cognitive Sciences; Statistics Minor, University of California, Irvine   March 2019
• GPA: 3.85 out of 4.0; Award: Magna cum Laude, Dean's Honor List

WORKING EXPERIENCES

Machine Learning Engineer Intern   June 2019-August 2019
Empyrean Software, San Jose, CA
• Carried out the development, testing, parameter tuning and optimization of a machine learning product aiming to efficiently predict process corners values in semiconductor wafer based on PVT and corner feature values.
• Utilized neural networks and different boosting algorithms to predict the corner values with an average accuracy over 0.97; Improved algorithms and data structures to increase the program running speed 30 times as the original testing speed.
Course Assistant   September 2018-January 2019
UCI Paul Merage School of Business, Irvine, CA
• Answered over 300 students' questions in class, during office hours, on website and during exams; Graded students' weekly assignments, quizzes, discussions and examinations; Held office hours.
Financial Assurance Intern   June 2018-August 2018
PricewaterhouseCoopers, Beijing, China
• Participated in the interim review of a public Real Estate Group with business coverage over leasing, emporium, hotel operations and multiform investments; Obtained, inspected and analyzed over 200 financial reports, financial agreements, commercial contracts and credentials.
• Conducted the confirmation procedure for most categories on financial statement; Collected and verified client's statistics through obtaining third-party documents, recalculation and necessary statistical methods.
• Conducted preliminary auditing tasks for the year-end auditing and the internal control procedures.

PROJECTS AND RESEARCH EXPERIENCES

Brown Datathon 2020, *First Place Winner (Dassault Systèmes)*   February 2020
• Used regression methods, Deep Neural Networks and Boosting algorithms to predict Additive Manufacturing(AM) temperatures and melt dimensions given lase speed, power, direction and edge effects; Achieved accuracy score over 0.999 for multiple tasks.
• Implemented feature engineering methods based on thermodynamics and PDE, bootstrapping and customized parameter tuning strategies to avoid data leakage problem and improve algorithm stability.
Sports Data Visualization and Virtual Match Simulation   December 2019
• Obtained NBA match statistics from websites automatically and periodically with web crawler; Visualized, compared and ranked players' performances and capabilities based on users' queries.
• Built data engineering pipeline with Python and MongoDB; Developed web user interface using *Dash by Plotly*; Stored and retrieved data in MongoDB before applying algorithms for virtual match results score calculation.
Student Researcher   August 2018-July 2019
Department of Statistics, University of California, Irvine
• Clustered over 100 rat neurons based on neural firing activities observed at over 700 time points using k-means clustering, hierarchical clustering and Mixture of Gaussians; Visualized the clustering results using heat maps.
• Classified wine quality**,** Z-Alizadeh Sani and breast cancer cell data using seven machine learning algorithms with hyperparameter tuning process and cross validations; Achieved accuracy score over 0.98 for each dataset.

SKILLS

• **Programming Language:** Proficient in Python, R, Julia, SQL and SPSS; Basic knowledge of MATLAB, C and Linux.
• **Statistical Analyses and Machine Learning:** GLM, PCA, Factor Analysis, Regression, MANOVA, Decision Tree (Boosting) Algorithms, KNN, LDA, SVM, EM, KNN, Causal Inference, RNN, CNN, LSTM, VGG.
• **Language:** Fluent in English and Chinese Mandarin; Conversational in French.