Yeyang Zu

An innovative Data Analysis Engineer with researching experiences on empirical model construction, regression and classification prediction, and machine learning algorithms

(+86) 189-139-20241

zuyeyang@ucla.edu

yeyangzu-benson.info

Education

Bachelor in Environmental Science, Minors in **Statistics** and Civil Engineering
University of California, Los Angeles
2017-2021, GPA 3.89/4.00

Coursework on Statistics:
- Machine Learning
- Computation
- Experiment Design
- Optimization

Coursework on Math:
- Probability
- Statistical Algorithm
- Linear Algebra
- Polynomial Calculus

Honors:

2017-2021 Dean's List
2021 Student Researcher Spotlight

Technical Skills

Programming / Visualization:

Python • R language • ArcGIS

SQL • MATLAB • Git • EXCEL

Tableau • C++

Machine Learning:

• MLPClassifier • sklearn • GridSearchCV • Pandas • Numpy • Matlibplot • Seaborn

Research Experience

Mar 2019 – Present: **Independent Researcher Intern** — UCLA CEE Department with Prof. Sanjay Mohanty
Engineering Critical Zone Lab
- Applied **PCA** to test most significant main effects of Biochar properties in E. coli removal capacity, help client in Biochar Production Methods.
- Applied **PLS regression** to prove the hypothesis of minimal interaction effects and develop empirical models of E. coli's growth-death-index.

Sep 2019 – Jun 2020 : **Lead Data Analyst** — UCLA Institute of Environmental Science
Patagonia Reservation Team
- Extracted 70GB all cloud free light bands raster and DEM imagery data over Sentinel-2 Satellite by **Sen2Cor** platform
- Horned a supervised classification of willows by **random forest model in R script** with logic rules, improved kappa values by 300% to 0.84.

Data Analysis Projects

June 2021 **Supervised Penguins Classification by Python** — UCLA Math Department
- Applied **f_classif** and **mutual_info_classif** to select the 3 most correlated features out of 25 features of different penguins species
- Applied **GridSearchCV** in optimizing all **Decision Tree, Logistic Classification, and SVM,** resulting with best test accuracy of 0.97

.

May 2021 **Optimization of Nozzles Design by R** — UCLA Stats Department
- Built cost-effective forecasting using 8 parameters to Nozzle's ejection using only 2^{8-4} **factorial factional design** to save budget from full 2^8 runs.
- Optimized the ejection ranges while minimize the consumption usage.

April 2021 **Car Price Market Forecast by R** — Kaggle Competition
- Used **LeapForwardSelection** and **Linear Regression** to train top 5 of the 25 properties of cars to help the clients in car price prediction
- Winning **7th rank out of 100 competitors** with a robust R score of 0.96.

May 2020 **Supervised Wine Quality Testing by MATLAB** — UCLA CEE Department
- Used **Artificial Neural Network** and **Linear Regression** to predict a quantitively evaluations of Wine quality, avoid subjective human-testing.

Professional Development

March 2021-Present **Team Lead in Deep Tech Track** — Stanford CEO
Global Chinese Entrepreneurs Alliance
- Evaluating early-stage investors and provide **due diligence report** and select valuable BPs for collaborated Capital Funds.
- Organizing 6 hours weekly industry analysis and hosting weekly 2-hours presentation across topics including AI software and Deep Tech.

Publication

Valenca, R., Le, H., **Zu, Y**., Dittrich, T., Tsang, D., Dutta, R., Sarkar, D., Mohanty, S.K. "Nitrate Removal Uncertainty in Stormwater Control Measures: Is the Design or Climate a Culprit?" *Water Research, Pergamon,* 24 Dec. 2020

Valenca, R., Borthakur, A., **Zu, Y**., Matthiesen, E. A., Stenstrom, M., Mohanty, S. K. "Biochar Selection for Escherichia Coli Removal in Stormwater Biofilters." *Journal of Environmental Engineering, ASCE,* 24 Nov. 2020