Yu-Ting Shen

405.200.2633
YuTing.Shen-1@ou.edu
linkedin.com/in/yu-ting-shen-6b730b160
github.com/ytatus94

SKILLS

Programming Language: C/C++, Python (pandas/numpy/scipy/scikit-learn/matplotlib/seaborn), R, BASH shell script
Machine Learning: data scraping, feature engineering, classification, linear/logistic regression, decision tree, clustering, NLP
Statistics: Regression (LSE/MLE), Confidence intervals, Monte Carlo method, Hypothesis test, A/B test, bootstrapping, Bayesian
Version Control: Git • Database: SQL

EXPERIENCE

The European Organization for Nuclear Research (CERN) Research Assistant — Geneva, Swiss | 2015 - 2018

- Analyzed 30M+ events to improve the true positive rate of isolated electrons from 0.85 to 0.93 and the calibrated electron isolation efficiency became the standard version of the collaboration.
- Implemented the decision tree based methods to maximize the significance and optimized the data selection processes.
- Built data cleaning and processing pipeline to analyze the data and performed an unbinned maximum likelihood fit to extract signal within 95% confidence interval.

Academia Sinica Research Assistant — Taipei, Taiwan | 2009 - 2011

- Discovered serious bugs and redesigned Monte Carlo simulation programs using C++ to improve the accuracy.
- Used decision tree based methods to classify background sources and reduced the major background events by 50%.

Taiwan Semiconductor Manufacturing Company R&D engineer — Hsinchu, Taiwan | 2006 - 2009

- Completed an urgent project 2 months ahead of the normal schedule and saved the contract.
- Programmed measurement system to automate data collection processes and reduced 30% of manual operations time.
- Led and supported 7 projects to build device models using regression, visualized and documented the results for customers.

PROJECTS

Real Lepton Efficiency

- Classified particles using decision tree based methods to improve the real lepton identification efficiency from 80% to 98%.
- Designed data filtration algorithms in C++ reducing the data size from ∼400 TB to less than 200 GB.
- Developed a new framework to reduce the run time from one week manual operations to 5 hours automation.

Capital Bikeshare Rental Demand

- Performed feature selection and engineering to remove 3 standard deviations outliers and created new datetime features.
- Visualized the weather information and imputed missing data using the random forest model based estimation.
- Used random forest model to predict the total count of bikes rented during each hour of the last 10 days in each month.

Upstart 3-year Terms Loan Charged-off Rate Prediction

- Analyzed 50,000 loans with different repayment status within two years to extract the monthly default rate.
- Predicted the charged-off rate by the time all 3-year terms finished using linear regression.

INSPIRE HEP PostDoc Position in North America

- Scraped website using Beautiful Soup and converted results into SQL database using sqlalchemy.
- Leveraged natural language processing (NLP) and machine learning to predict institution participants in experiments.

EDUCATION

University of Oklahoma Ph.D. in Physics | Cum. GPA: 3.47 / 4.00 — Norman, OK | 2011 - 2018

- Published 4 top journal papers and 5 conference notes, and presented in 2 conferences.
- 6 semesters teaching assistant for college and graduate level courses.

National Taiwan University M.S. in Physics | Cum. GPA: 3.86 / 4.00 — Taipei, Taiwan | 2003 - 2006

- Won 3^{rd} place graduate student thesis awards, The Physical Society of Taiwan, 2006.