JIASHU MIAO

✉ michaelmicmiao@aol.com 🌐 michaelmiaomiao.github.io/about/ 📞 +1-(310)-308-0058 in www.linkedin.com/in/michaelmiaomiao/ ⚙ michaelmiaomiao

Work Experience

Sony Pictures Inc., Sony Corporation
Los Angeles, CA
Data Analyst Intern, Application Development & Maintenance Jun. 2019 - Sep. 2019

- Deployed Jupyter workspace with **R and Python** environment, **SQL interface** and API on 10 **Linux** servers for **remote** access, big-data management & computing;
- Extended the service with **proxy, authenticator, Java Node.js** for **multi-user** access, collaboration and administration, **saving 30%+ time** than individual deployment;
- Conducted **anomaly detection** for **50k** Nielsen TV stream data in Sony's **Tableau warehouse** by extreme value analysis(Z-score) and **machine learning(ML)** models;
- Assisted Sony's data science initiatives with **SQL database manipulation**, **NLP** classification for TV genres and **time series** decomposition and **visualization** development;
- Created Python **automation solutions** to support business intelligence and presentations for platform deployment to **100+ audiences across teams.**

UCLA Health, Department of Medicine Statistics Core
Los Angeles, CA
Statistical & Analytical Research Assistant Sep. 2018 - Jun. 2019

- Assisted principal statistician with **database design, data management** and project development for 12k patients' information using **REDCap and SQL;**
- Engineered 12 **predictive models (**logistic, kNN, SVM, trees) applied in healthcare, which **optimized current algorithms** for hospital LACE index and readmission prediction;
- Implemented **model validations and evaluation** with multiple measurements (accuracy, AUC, ROC), and approved by senior leadership with a 15% increased performances.
- Conducted further analytical study and research and **initiated a paper thesis** in healthcare.

Forkaia
Irvine, CA
Data Science & Analytics Intern Mar. 2019 - Jun. 2019

- Assisted two data-science based projects to establish business plans, implement marketplace analysis for user information and **A/B testing for platform performance;**
- Conducted **data preprocessing** on 20+ raw datasets and addressed **data consistency** issue with error detection, **correction and imputation** in R;
- Collaborated on an **audio detection model** to identify and classify 400 audio instances in Spotify, and parsed solutions with ML, information retrieval, clustering algorithms.

USA-UES, Edtech Company
Los Angeles, CA
Automated Statistics & Software Engineer Intern Jul. 2018 - Dec. 2018

- Implemented **survey analysis** based on data collected from 200+ students, engineered 14 scaled factors to **automatically formulate customer reports,** saving 40% time;
- Optimized the education model, currently integrated into "USA-UES" iOS portal, by performing **feature selection**, **regularization,** and added 7 parameters for high-level automation, adjustment and testing;
- Designed and developed a front-end utility to achieve **integration of data storage and analysis solutions**, and added fuzzy matching, text entering & searching functionalities;
- Assisted the research team with improved **survey methodology implementation** and **statistical consulting** to develop software-based algorithms;
- Developed mechanisms for **data-fetching , web-automation and mapping** with R & Python, and managed on **Github** to facilitate 5 projects and assist 20 people's work.

Wilshire Financial Network
Beverly Hills, CA
Business Analyst Intern Mar. 2018 - Jun. 2018

- Conducted 5 cohort analysis of real estate bankruptcy data from 10+ California judge's calendars to **investigate potential customer patterns and several KPI's;**
- Optimized the customer database management and **assisted acquisition plans** with 15% increase in conversation rate.

Education

University of California, Los Angeles.
B.S in and Statistics & Applied Mathematics;
GPA 3.80/4.00; Expected 2020;
• RA at UCLA Halperin Big-data Genomics Lab.
• EY NextWave Data Science Global Competition 2019 — Country/Region's Finalist.
• Kaggle Competition: Predicting the Affordability of Houses in Ames — Overall Ranking: 2, Top 6%.

Technical Skills

•Python, R, SQL, C++, SAS, Matlab, Node.js (NPM), Bash, Git, HTML, RHEL& CentOS Linux

• Tableau, Plotly, PostgreSQL, pgAdmin4, JupyterHub & IPython Notebook, REDCap Database, MS Excel, Google Analytics, Citrix

• Linear Algebra, Mathematical Modeling, Data Mining and Modeling, Data Visualization, ML Algorithms, Database Management

Selected Projects

EY DATA CHALLENGE - LOCATION PREDICTION[R,PYTHON]
• Featured and tested 37 Classification models with ML algorithms: LDA, QDA, k-NN, SVM, Boosting, Forest, Ensemble, etc; evaluated and tuned for high performances.

DUE DILIGENCE CHECKING [PYTHON, R (NLP)]
• Developed an interactive BI mechanism using NLP R-packages and Python Selenium to fetch, analyze and process, mapping text data for background checking of new employees with extended automatic functionalities.

ZIPLINE UAV(UNMANNED AERIAL VEHICLE) FLIGHT DATA EXPLORATION [R, PLOTLY, TABLEAU]
• Combined time-series and geolocation study of 800+ datasets from 40 drones' flights over a month to conclude advice for business operation and identify engineering risks.

RUGBY 7S SPORTS ANALYTICS [R, TABLEAU]
• Studied Canadian Women's Rugby players' fatigue levels and performance by analyzing game data, athletes' surveys, and delivered executable advice to coaches.

DEMOGRAPHIC ANALYSIS WITH MAPS [SQL]
• Visualized analysis of demographic information for Seattle areas for three decades via Tableau interactive maps and plots to provide geolocation distribution, city plan insights, etc.

CLINICAL STATS CONSULTING [R, PYTHON, GPOWER]
• Designed experiments with GEE models and inference analysis to assist UCLA Stein Eye Institute to assess eye surgeries and treatments.