Yutong Jiang

(734)510-1308 | katieji1998@gmail.com

EDUCATION

University of Michigan-Ann Arbor Aug 2020 - Dec 2022
Master of Data Science Ann Arbor, MI
Current GPA: 3.8/4.0
Master of Applied Statistics
Current GPA: 3.8/4.0
Key Courses: Bayesian Modeling and Computation; Data Science and Analytics using Python; Exploratory Data Analysis; Principles and Practices in Effective Statistical Consulting; Data Manipulation and Analysis; Multivariate regression analysis

University of Colorado Boulder Aug 2016 - Jun 2020
Bachelor of Applied Mathematics Boulder, CO
GPA: 3.7/4.0
Honor: Dean's Honor List
Key Courses: Statistical Learning; Markov chains, Queues, and Monte Carlo Simulation; Introduction to Data Science with Python; Machine Learning

PROJECT EXPERIENCE

Galaxy Detection Challenge Sep 2020 - Dec 2020
University of Michigan-Ann Arbor
- Implemented a neural network using Pytorch that can detect galaxies in simulated telescope images.
- Trained the neural network to obtain 89% accuracy on unseen test data.

Factors of Affecting Fatal Aviation Accident by Bayesian Analysis Jan 2021 - Apr 2021
University of Michigan-Ann Arbor
- Constructed a Bayesian logistic regression model and a hierarchical regression model to analyze factors that contributed to fatal aviation.
- Applied Markov Chain Monte Carlo to the sampling probability distribution for each estimate
- Simulated the distribution close to the real one.

Analysis on the Factors of Affecting Job Change Jan 2021 - Apr 2021
University of Michigan-Ann Arbor
- Compared random forest, Adaboost, and logistic regression models to discover the most important features.
- Used the model to predict whether a candidate was willing to change a job.
- As a result, logistic regression performs best with AUC(Area under the curve) 0.7812.

Faculty Questionnaire Jan 2019 - Apr 2021
University of Colorado Boulder
- Applying model selection and feature selection using R and ended up using Quadratic discriminative analysis to make predictions of whether an instructor is a good instructor based on Faculty Questionnaire data.
- The accuracy of testing is more than 90%.

Analysis of the Relationship of Household Income and Covid-19 Sep 2021 - Oct 2021
University of Michigan
- Analyzed the connections between the household income level and the Covid-19 diseases by county in the US using SparkSQL
- Filtered out the useful data and aggregated the data to get the tendency and relationship for families' income levels and the diseases.
- Visualized the relations by applying seaborn in Python.

RESEARCH EXPERIENCE

CRISPR accessible DNA bases in random genomes Jan 2019 - Dec 2019
Undergraduate Researcher - Professor Manuel Lladser
- Computed the probability distribution of the fraction of bases in a memoryless DNA sequence called PAM motif that has known length that can be edited by CRISPR-Cas9.
- Applied generating functions, and combinatorial methods to compute the DNA distribution.
- Succeeded in determining this distribution when the PAM motif is NGG (i.e. AGG, CGG, GGG, or TGGG).

SKILLS LIST

- **Technical:** R, Python, SQL, Spark, SparkSQL, Tableau, Hadoop, Latex, Excel
- **Languages:** Chinese, English
- **Other:** Teamwork, communication, critical thinking, problem-solving, strong work ethic