Michael Baluja

Machine Learning and Data Science

✉ michael@michaelbaluja.com

📞 760-330-6385

📍 San Diego, CA

🌐 michaelbaluja.com

in linkedin.com/in/michaelbaluja

🐦 twitter.com/mlbaluja

github.com/michaelbaluja

EDUCATION

M.S. Electrical and Computer Engineering
University of California, San Diego

08/2021 - 06/2023   *La Jolla, CA*

- Focus: Machine Learning and Data Science

B.S. Cognitive Science
University of California, San Diego

07/2019 - 06/2021   *La Jolla, CA*

- Focus: Machine Learning and Neural Computation
- Major GPA: 3.912

WORK EXPERIENCE

Data Assistant
UC San Diego - Library Data Curation Team

06/2021 - Present   *La Jolla, CA*

- Analyze eight public research repository APIs, integrating into single search tool.
- Create and publish adaptive open source user interface for greater scientific community to mass collect research data.
- Formalize methodologies for journal submission and data curation conference presentations.
- Collaborate with team of three data curators to understand best practices for reusable research.

Research Assistant
UC San Diego - CMRR STAR Group

01/2020 - Present   *La Jolla, CA*

- Improve deep reinforcement learning framework to design error protection weight masks for text classification networks trained on two different classification tasks for multiple data sets.
- Increase network model performance by up to 300% in the presence of simulated weight errors.
- Communicate experimental results in conference papers, posters and presentations for four undergraduate conferences.

Instructional Assistant
UC San Diego

10/2020 - 12/2020   *La Jolla, CA*

- Familiarized upward of 100 students with data science topics including python, git, statistics, and machine learning.
- Co-instructed 18 discussion sections covering course material for groups of up to 10 students.
- Assessed individual course final projects across more than a dozen data science questions.

SKILLS

Python MATLAB SQL git NumPy
Pandas Scikit-learn PyTorch Unix
selenium tkinter Deep Learning
Reinforcement Learning

PERSONAL PROJECTS

Deep Reinforcement Learning for 2048 (03/2021 - 06/2021)

- Created python-based player environment of popular game 2048 using NumPy.
- Engineered custom three layer neural network to learn value of in-game moves based on game states and actions.
- Implemented state-of-the-art learning algorithms such as Monte Carlo and Temporal Difference techniques to simulate game play and learning process.

Empirical Classifier Comparison for NLP (10/2020 - 12/2020)

- Tested four machine learning classification techniques on three different metrics for performance ranking.
- Analyzed and explored three data sets for usability to ensure proper functionality of trained models.

Python Lyrical Analysis (05/2020 - 09/2020)

- Integrated python libraries to aggregate lyric information, extract necessary features, and present user-chosen sentiment analysis for tens of thousands of artists.
- Designed pipeline for data collection, intake, cleaning, analysis, and graphical presentation of results.

Job Automation Analysis (04/2020 - 06/2020)

- Led team of four through data analysis to analyze three potential correlates for job automation in the United States for over 700 occupations across 25 year period.
- Utilized statistical techniques such as geospatial analysis, hypothesis testing, and multiple linear regression.

ORGANIZATIONS

California Louis Stokes Alliance for Minority Participation in Science, Engineering and Mathematics (06/2020 - Present)

INTERESTS

Human-centered data science

Computer Vision

Natural Language Processing