

# Michael Baluja

Machine Learning and Data Science



[michael@michaelbaluja.com](mailto:michael@michaelbaluja.com)



[linkedin.com/in/michaelbaluja](https://linkedin.com/in/michaelbaluja)



[michaelbaluja.com](https://michaelbaluja.com)



[github.com/michaelbaluja](https://github.com/michaelbaluja)

## Education

### M.S. in Electrical Engineering

University of California San Diego (September 2021 – March 2023)

**Focus: Machine Learning and Data Science**

### B.S. in Cognitive Science

University of California San Diego (June 2019 – June 2021)

**Specialization: Machine Learning and Neural Computation**

## Experience

### Data Scientist Intern

Experian (June 2022 – September 2022)

- Analyze and refactor existing development code to deliver up to 92.6% speed increase.
- Assemble Python package for statistical analysis across six metrics for measuring disparate impact.
- Integrate explainable machine learning algorithms to ensure fairness across model predictions.

### Data Assistant

University of California San Diego (June 2021 – June 2022)

- Create ETL pipeline to manage 125k research object records across eight research APIs.
- Maintain and debug open-source user interface for big data collection via supported repositories.
- Collaborate within data curation team to integrate user feedback into application development.

### Research Assistant

University of California San Diego (January 2020 – June 2021)

- 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 model performance by up to 300% in the presence of simulated weight errors.
- Communicate experimental results in papers, posters, and presentations for Great Minds in STEM and LSAMP undergraduate conferences.

## Technical Skills

- Programming: Python, C++, SQL, MongoDB
- Machine Learning: scikit-learn, PyTorch
- Data Analysis: Pandas, Selenium, Numpy, Scipy
- Computing: Linux/Bash, Docker, git

## Projects

### Autonomous Vehicle (September 2022 – December 2022)

- Program NVIDIA Jetson Nano for controlling vehicle via neural network with real-time camera input.
- Integrate popular computer vision frameworks to pilot autonomous vehicle around track obstacles.
- Utilize OpenCV for real-time traffic detection and speed limit/stop sign recognition.

### PyCurator (June 2021 – Present)

- Develop data pipeline and API queries via Object-Oriented multiple inheritance approach.
- Design and publish user interface to allow streamlined access to eight online research repositories.
- Implement best practices to ensure computational simplicity, reducing runtime by up to 72%.