

# CARLOS PAREJA

(562)551-1298

cpareja3025@gmail.com

[LinkedIn](#), [GitHub](#)

## EDUCATION

---

**University of California, San Diego**

*(expected) December 2023*

B.S. in Mathematics-Computer Science with a Specialization in Machine Learning

**Pasadena City College**

*2018-2021*

Associates in Physics, Mathematics, Natural Sciences

## EXPERIENCE

---

**Machine Learning Research Intern (UC San Diego)**

San Diego, CA

*Departments of Computer Science and Physics (Duarte Lab)*

*January 2022 - Present*

- I'm currently developing the JetNet Library for Machine Learning (ML) and High Energy Physics (HEP). JetNet is a library for developing and reproducing jet-based ML projects.
- My role is to expand the library by adding new particle physics datasets to the library and reformatting these datasets into PyTorch tensors and NumPy arrays. The package allows the user to specify their desired train/test splits, jet features, particle features, etc.
- JetNet is currently averaging 2000 downloads per month and my goal is to continue developing package in order to facilitate with research in the research cross disciplines of ML and HEP.

**Software Engineer Research Intern**

Pasadena, CA

*NASA's Jet Propulsion Laboratory (JPL)*

*April 2022 - September 2022*

- I implemented and tested Markov Chain Monte Carlo (MCMC) algorithms to facilitate in the research and analysis of data from upcoming galaxy redshift survey telescopes such as the Nancy Grace Roman Space Telescope and SPHEREx. The MCMC algorithms that I implemented from scratch and utilized were able to make a tight constraint on Cosmological parameters in a highly dimensional parameter space.
- Overall, writing code for two large scale NASA missions allowed me to assist the research division of JPL and being to achieve constraints of cosmological parameters will give us further insight to the physics that were play at the beginning of the Universe.
- I wrote my code in the Julia and Python languages and worked with GitHub to collaborate with my colleagues and tested my MCMC algorithms on the TACC Supercomputer.

**Information Technology Services (ITS) Kualu Build Support Technician**

San Diego, CA

*UC San Diego ITS*

*March 2022- June 2022*

- I supported the implementation of the Kualu Build platform by training and guiding new users on how to operate the platform.
- I captured clients current processes and requirements in order to translate them into a more effective and efficient business process utilizing the Kualu Build platform. I also recommended process improvement ideas that helped the business process run smoothly in the new environment

**Lamat Computational Astrophysics Fellowship (NSF REU)**

Santa Cruz, CA

*Departments of Applied Mathematics and Astrophysics*

*April 2021-January 2022*

- I worked on high-order Gaussian Process methods for Computational Fluid Dynamics (CFD) simulations with the GP research group at UCSC.
- My research consisted of comparing and implementing three traditional finite differencing methods against the new GP modeling method. Our current work consists of expanding to include more numerical methods and I am implementing these new methods as well.
- I conducted my data analysis in Python, calculated the error for each method and conducted a grid convergence study where I analyzed how error decreased when increasing the grid resolution of each numerical approximation. I demonstrated that with GP modeling we can achieve a highly accurate numerical approximation for CFD simulations.
- My research has have wide-ranging applications, most notably in improving numerical accuracy in larger scale CFD simulations in astrophysics and other fields of science and engineering.

## EXTRA PROJECTS

---

### Implementing a CNN to an LSTM model for Image Captioning of the COCO Dataset and ResNet50 to LSTM model

November 2022

*University of California, San Diego*

- I implemented a Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) network in order to generate captions for images of the COCO dataset. I also performed transfer learning by replacing my CNN with a ResNet50. The best model was with ResNet50 and I received a BLEU1 score of 54.5.

### Implementing a Multi-Layer Neural Network from scratch for Multi-Class Classification of the CIFAR-10 Dataset

October 2022

*University of California, San Diego*

- I implemented a Multi-Layer Neural Network from scratch, without the use of any deep learning frameworks. I only used the Python NumPy library to construct the layers, implement back-propagation, and several optimizing algorithms from scratch such as Stochastic Gradient Descent (SGD) with Momentum and the Adam optimizer. I performed Multi-class classification on the CIFAR-10 Dataset with a Softmax output layer and our best model achieved an accuracy of 61%.

## RELEVANT COURSES

---

**1. Computer Science / Data Science:** Data Structures and Algorithms (Java) , Mathematical Software, Data Science in Practice, Deep Learning, Systems Programming, Supervised Machine Learning Algorithms, Machine Learning in Physics, Intro to Computer Science I (C++), Fundamentals of Computer Science (Python)

**2. Mathematics:**

Calculus I, Calculus II, Honors Multi-variable Calculus, Linear Algebra, Honors Differential Equations, Mathematical Reasoning, Statistical Methods, Discrete Math and Graph Theory

## SKILLS

---

**Programming Languages:**

Python, Java, Julia, R, C++, C, L<sup>A</sup>T<sub>E</sub>X, Matlab, SQL

**Tools/Frameworks:**

Git, Linux/UNIX Command Line, Jupyter, VsCode, Eclipse, R Studio, Slack, NumPy, Pandas, Matplotlib, Seaborn, Tensorflow, PyTorch, JUnit, Zeus, Scikit-Learn, AWS

**Languages:**

English, Spanish

**Soft Skills:**

Customer Service, Teamwork, Leadership, Time Management, Communication, Problem Solving, Accountability