Carmen W. Matar

818-219-9949 | carmenwmatar@gmail.com | linkedin.com/in/cwmatar | github.com/C-Matar | Los Angeles, CA

EDUCATION

University of California, Berkeley
Master of Molecular Science and Software Engineering
August 2024 – May 2025
GPA: 3.9/4.0
Relevant Coursework: Machine Learning Algorithms, Applications of Parallel Computers, Programming Languages for Molecular Sciences: Python and C++, Numerical Algorithms applied to Computational Quantum Chemistry, Software Engineering for Scientific Computing, Software Engineering Fundamentals for Molecular Sciences, Principles and Techniques of Data Science, Ethics in Molecular Science and Software Engineering

University of California, Berkeley
B.A. Molecular and Cell Biology
August 2021 – May 2023
Relevant Biology Coursework: Molecular Therapeutics, Molecular Endocrinology, Physiology and Cell Biology Lecture & Lab, Genetics-Genomics-and-Cell-Biology, General Biology I/II
Relevant Chemistry Coursework: General Chemistry I/II, Organic Chemistry, Biochemistry
Relevant Mathematics Coursework: Single Variable Calculus, Multivariable Calculus, Statistics, Linear Algebra, Physics I/II

Los Angeles Pierce College
A.A. Science, Technology, Engineering and Mathematics | A.A. Social and Behavioral Sciences
August 2018 – June 2021

TOOLS

Languages: Python, SQL, R, C++
Development Tools & Platforms: Git, Bash, Github, Docker, Kubernetes, AWS, Jupyter, Linux, VSCode, Powershell
ML Frameworks: PyTorch, TensorFlow, Scikit-learn, Keras, Torchvision
Visualization Tools: Matplotlib, Scikit-image, OpenCV, Seaborn, TensorBoard, UMAP, Plotly
Scientific Computing Tools: ScanPy, Seurat, Pandas, NumPy, RDKit, CMAP, GSEApy, SciPy, Biopython
Parallel Computing and HPC Tools: MPI, OpenMP, CUDA

PUBLICATION

Acute Lymphoblastic Leukemia in Children and Adults: A Review of the Differences in Biology, Treatment, and Prognosis *April 2022*
https://escholarship.org/uc/item/0z62945j

EXPERIENCE

AI/ML Engineer Intern
Violet Therapeutics | Boston, MA (Remote) | *Python, R, PyTorch, Scikit-learn, Harmony, GSEApy, CMAP*
January 2025 – June 2025
- **Led the end-to-end development of an ML pipeline** for biomarker discovery and drug target prioritization in multiple sclerosis using scRNA-seq data.
- Identified DEGs across disease states; integrated **ScanPy** and **Seurat** for robust cross-platform analysis and cell-type annotation.
- Developed and trained **Ridge, Lasso, and ANN models**, achieving **95% accuracy** for chronic active and **97%** for chronic inactive MS classification.
- **Built** an **end-to-end framework** to prioritize druggable gene targets, incorporating **GSEApy** and **CMAP**-based drug repurposing.

Administrative Assistant
High-Tech Electric Co. | Los Angeles, CA
August 2018 – August 2021
- Managed intake, scheduling, and documentation, improving client satisfaction and streamlining operations.

PROJECTS

Machine Learned Conformer Energy Prediction via Approximating Pairwise Potential (C++, Python, HDF5, NumPy, Pandas, Matplotlib)
March 2025 – May 2025
- Led a team of two to build a **C++ neural network model** to predict molecular conformer energies using atomic pairwise potentials.
- Trained on ANI-1 dataset; Test Set **MAE of 6.67 Hartree greatly outperformed Lennard-Jones MAE of 2039 Hartree**.

Generative Radiological Image Captioning via Encoder-Decoder Architecture (PyTorch, Data Processing, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Natural Language Processing (NLP))
October 2024 – December 2025
- **Led a team of four** to develop a **CNN-RNN pipeline** for automatic captioning of radiological images.
- Used Inception-v3 for feature extraction and an LSTM decoder for generating medical captions.
- Trained on 80K-image ROCO dataset with extensive preprocessing, hyperparameter tuning, and loss tracking.