Preet Patel

13 Vauxhall Pl, Chapel Hill, North Carolina patelpb96@gmail.com 6307156148 https://www.linkedin.com/in/patelpb96/ My Website

SKILLS

Certifications: Stanford Online & Deeplearning.ai – Supervised Machine Learning
Hard Skills: Git version control; high performance computing, Python and data analysis libraries (Numpy, Scipy, Matplotlib, Pandas/Polars); big data analytics, visualization, and machine learning (ML) analysis (sklearn, TensorFlow); time-series analysis (LSTM, ARIMA), post-graduate mathematical skills; LaTeX; Linux/Unix; SQL; OpenMP/MPI
Soft Skills: English communication skills, strong presentation skills, motivated and independent self-learner, persuasive writing, critical thinking, curiosity, teamwork, adaptability
Other: English, Gujarati, working proficiency in Spanish, Graphic Design (Photoshop, Cinema 4D)

PROJECTS

Python Module Development and Implementation (Element Tracers) - Github
• Implemented new models and made them accessible to entire collaboration (200+ people), enabling new and ongoing projects.
• **Optimized existing code** for element tracer processing from 2 classes into 1, **with a measured speedup of over 50%** when compared to previous routines. Prepared package for use with **HPC**s via shell-scripting on **Linux**-based systems.

Machine Learning: Maximum Likelihood Estimation and MCMC - Github
• Wrote data analysis pipeline to analyze raw observational data with simple means and variance with fits to a gaussian profile; a maximum likelihood estimation after constructing a log-likelihood function; an MCMC ML-algorithm to convergence.
• Successfully determined the wavelength of light at which the emission occurs, identifying the source for calibration.

Machine Learning: Metropolis-Hastings Algorithm - Github
• Manually implemented a Metropolis-step algorithm to determine the best fit model to galaxy-catalog data, in M-L space.
• Successfully converged on the galaxy mass-luminosity relation, confirming our understanding of the universe.

Predictive Analytics
• **Project 1:** Utilized social listening data from a collection of perfume reviews to drive business decisions for investments. Reduced the time that goes into deciding whether to invest in a product by over 50%.
• **Project 2:** Organized and analyzed over 4 years of service level data to predict service level in upcoming quarter. **Result:** developed an ARIMA class model which predicts service level for 90 days with low RMSE.

EDUCATION

M.S. Physics	**University of California, Davis**	**2020 - 2023**
B.S. Physics	**University of Michigan, Ann Arbor**	**2015 - 2019**
B.S. Astrophysics	**University of Michigan, Ann Arbor**	**2015 - 2019**
Minor in Statistics		

EXPERIENCE

Graduate Researcher (Astrophysics) — March 2021 - September 2023
University of California - Davis
• Utilized **Python**, parallel processing, supercomputers, Linux systems, advanced mathematics**,** ML methods (supervised learning), scaling analysis, and hydrodynamic simulations to complete multiple projects with the FIRE collaboration.
• Parsed through several **petabytes** of simulation data stored as HDF5s across national supercomputers. Additionally **optimized runtime by 50%**. Result: **1st author publication**, with additional authorlist publications in prep. A subset of this data publicly available at https://fire.northwestern.edu/.

Teaching Assistant (TA) — October 2020 - April 2023
University of California - Davis
• Used **data visualization** and diverse communication skills teach both technical and non-technical students about complex physical phenomena across various subfields of physics.
• (example: quantum mechanics for non-STEM majors, with detailed lectures, office hours, grading, and homework assistance). Class sizes: 30-250 students, for 1 to 3 hours per session.

Bluewaters Student Intern — May 2018 - May 2019
University of California - Davis
• Created my own computing cluster using laptops, and optimized programs on **Linux** HPC systems with CUDA, OpenMP/MPI.
• Explored parallelization based on job type and architecture (**GPU vs CPU**) to create n-body (10^5) galaxy simulations.