Preet Patel

Chapel Hill, North Carolina | patelpb96@gmail.com | Phone No. by Request | https://www.linkedin.com/in/patelpb96/ | My Website

Technologist with multidisciplinary experience using big data to tackle difficult & undefined problems. I will use my background in statistics & programming, alongside my experience with using ML/AI to create solutions in any role.

PROJECTS

Python Module Development and Implementation (Element Tracers) - GitHub

• Streamlined **Python** code to merge **data pipelines** into one class, resulting in a **50%+ speed increase** over previous methods; packaged the solution for compatibility with High-Performance Computing clusters operating on Linux shell.
• Made these pipelines available for use to 200+ collaborators.

Machine Learning: Maximum Likelihood Estimation and MCMC - Github[1] , Github[2]

• Wrote **data analysis** pipeline to analyze raw observational datasets and find emission wavelengths by fitting Gaussian profiles by minimizing a log-likelihood function and running **ML** algo (MCMC) to convergence. Queried observational data with **SQL**.
• Successfully determined the wavelength of light at which the emission occurs, identifying the source for calibration.

Kaggle Competitions - Profile

• Abalone Dataset: predict the number of rings in the abalone shell based on other physical characteristics.
• Result: ensemble model (LGBM, Catboost, and XGBoost) with lower RMSE than **82% of competitors** (483/2606).

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:** produced an ARIMA class model which predicts service level for 90 days with low RMSE.

EXPERIENCE

Computational Astrophysicist (Grad Researcher) — 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 in MNRAS**, detailing it at multiple conferences. Additional author-list publications in prep. A subset of this data publicly available at https://fire.northwestern.edu/.

Teaching Assistant (TA) – Physics/Quantitative Courses — 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.

High-Performance Computing Intern — May 2018 - May 2019

Bluewaters Student Intern, 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.

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		

SKILLS

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