Pankaj Chouhan

(850) 339-6651 | pankajchouhan93552@gmail.com | linkedin.com/in/pankaj | github.com/PC-FSU

Education

Ph.D in Computational Science | Florida State University — Aug. 2019 – June 2024
- **Specialization:** Machine learning applications for engineering.

Master's in Physics | Indian Institute of Technology, Ropar — Aug. 2016 – May 2018

Bachelor's in Physics | Delhi University — Aug. 2013 – May 2016

Technical Skills

Programming : Proficient in Python, C/C++, MySQL. Familiar with MATLAB and shell scripting.

Libraries : PyTorch, PyTorch-Lightning, TensorFlow, scikit-learn, Numpy, Pandas, Matplotlib, HuggingFace, pyTest, unittest.

Developer Tools : Docker, CI/CD (Github Actions), VIM, Linux, SLURM, Power BI, AWS(ec2, S3, Lambda)

Experience

Research Assistant — Aug 2019 – May 2024
Florida State University — *Tallahassee, FL*
- Developed a machine learning model that predicts the best chemical composition given desired polymer properties. Achieved a $5x$ speed-up in computation time over the existing physics-based model using a Gaussian process regression model.
- Achieved an 11x speed-up in training time compared to the previous model by utilizing reduced order models based on Gaussian processes and neural networks. Models inform sparse training data applications - 100 samples were used in training.
- Performed a comparative study between Neural Networks (vanilla MLP and multi-task MLP) and Gaussian Processes, concluding that Gaussian Processes performed better for limited data applications.

Data Analyst — Feb 2019 – Aug 2019
Alpha Six Capital — *Gurgaon, India*
- Increased the efficiency of in-house trading algorithms by optimizing code, resulting in a 35% speed improvement.
- Established a MySQL database and designed an ETL pipeline for storing end-of-day (EOD) data for downstream tasks.
- Created interactive Python and Power BI dashboards for EOD analysis, facilitating comparison of stock market indicators.

Teaching Assistant — Jan. 2021 – Jan. 2024
Florida State University — *Tallahassee, FL*
- Principal instructor for undergraduate courses "Introduction to Scientific Computing" and "Introduction to Data Science."
- Topics included Python, data visualization, exploratory data analysis, hypothesis testing, SQL, and machine learning.

Machine Learning Research Intern | Bayesian optimization — June 2021 - Aug 2021
Argonne National Lab — *Chicago, IL*
- Developed a ML routine that reduces the cost of running expensive-high-accuracy simulations by using information from cheap-low-accuracy simulations. The PyTorch-based routine uses Bayesian optimization and Gaussian process regression.

Deep Learning Research Intern | computer vision — June 2020 – Aug 2020
Argonne National Lab — *Chicago, IL*
- Developed a Python pipeline to de-noise and predict the progression of 3-D quantum Monte-Carlo simulation time series. Treated the simulation like a video and assumed that each frame represents a slice of the simulation. Used OpenCV for video de-noising and a CNN-LSTM neural network for forecasting the simulation's future behavior.

Publications

- Pankaj Chouhan and Sachin Shanbhag. Surrogate modeling with Gaussian processes for an inverse problem in polymer dynamics. Int. J. Comput. Methods, 20(8):2143003, 2023.
- Pankaj Chouhan and Sachin Shanbhag. Surrogate modeling of sparse functional data obtained from molecular models of polymer viscoelasticity. Submitted to the journal: **EAAI**.

Projects

Cancer Detection Engine | computer vision | *Python, PyTorch-Lightning, Pandas, MLFlow, SHAP, DVC, Docker, FastAPI, CI, PyTest.*
- Develop a customized VGG16-CNN architecture to detect tumors in CT scan images by leveraging transfer learning.
- Implemented data versioning (DVC), performed exploratory data analysis (Pandas), and experiment tracking (MLFlow).
- Ensures reliability via continuous integration (Github actions), functionality tests (PyTest), error analysis (SHAP), and a production-ready server (FastAPI).