SAARANSH PANDEY

Philadelphia, PA | saaransh@seas.upenn.edu | (215) 433-4966 | https://www.linkedin.com/in/saaransh-pandey

EDUCATION

University of Pennsylvania, School of Engineering and Applied Science, United States **May 2025**
Master of Science in Electrical and Systems Engineering, Major in Data Science GPA: 3.96/4.0
Coursework: Applied Machine Learning, Big Data Analytics, Statistics for Data Science, Deep Generative Models

Delhi Technological University, India **June 2020**
Bachelor of Technology in Engineering Physics, Major in Electronics GPA: 8.36/10
Coursework: Machine Learning Foundations, Pattern Recognition, Computational Methods

WORK EXPERIENCE

Aeronautical Systems Incorporated Virginia, United States
Data Scientist *(Jun 2024 – Present)*

- Led a team of interns to develop a Manufacturer/Distributor Predictor using Python and GPT-4 that automated manual analysis, reduced processing time from hours to seconds, and achieved over 90% accuracy.
- Developed the Competition Score Predictor using Azure AutoML to assist customers in making strategic bids, automating a previously manual process through deployment in Azure ML Studio.
- Built the Similar NSNs Identification Tool to find product alternatives or justify pricing, leveraging Pinecone for efficient data retrieval and deployed through Azure Endpoint to streamline operations.

UnitedHealth Group – Optum Gurugram, India
Machine Learning Engineer *(Jun 2020 – Jul 2023)*

- Led a data science initiative to detect early-stage opioid addiction by developing and deploying a machine learning classification model that achieved a recall of 0.63 and specificity of 0.75, thereby improving patient care.
- Developed a Python automation script using Levenshtein-based fuzzy matching to validate and match facility data, reducing process time from 14 days to under 10 minutes - a 99% reduction in processing time – while enhancing database accuracy.
- Spearheaded the Member Roster Management Project as Technical Lead, developing a Power BI dashboard for enhanced data visualization and oversight, coupled with automating SAS module files to streamline roster management processes.

RESEARCH EXPERIENCE

University of Pennsylvania | *Research Assistant – Dr. Masao Sako* *(Nov 2024 – Present)*

- Analyzing sky coverage efficiency for NASA's Roman Space Telescope by modeling overlaps in observed regions using Python with AstroPy and Healpy for pixel mapping and rotations, ensuring optimized coverage for supernova detection.

Indraprastha Institute of Information Technology Delhi (IIIT-Delhi) | *Research Assistant – Dr. Vivek Bohara* *(May 2019 – Jul 2019)*

- Developed a data-driven toolkit for simulating Visible Light Communication systems using Python and GNU Radio, providing an interactive platform for analyzing and comparing different modulation techniques.
- Published this work in the 2019 IEEE International Conference on Advanced Networks and Telecommunications Systems.

PROJECTS

Advancing Text-to-Image Generative Model Capabilities | *Python, PyTorch, GPU/TPU* *(Oct 2024 – Dec 2024)*

- Enhanced text-to-image generative models by developing an edit-friendly P2P method combining the Stable Diffusion 2-1-base model with a custom LLaMA 3b v2 encoder, achieving a SSIM of 0.73 and LPIPS of 0.06 for localized editing.
- Integrated consistency model to boost global editing performance, reducing runtime from 30 minutes to 15 seconds while reaching a peak SSIM of 0.82 and LPIPS of 0.12, balancing editing quality with computational efficiency.

Neural Network Model Compression and Optimization | *Python, PyTorch, GPU/TPU* *(Aug 2024 – Sep 2024)*

- Developed and evaluated model compression techniques, including global, layer-wise, and channel pruning on the VGG16 model, with global pruning achieving up to 45% sparsity while maintaining over 70% top-1 accuracy on the ImageNet validation dataset.
- Implemented iterative pruning and extended to transfer learning, achieving optimized performance on the melanoma dataset by retraining pruned models in resource-constrained environments.

Credit Card Fraud Detection | *Python, TensorFlow, PyTorch* *(Mar 2024 – May 2024)*

- Addressed a significant class imbalance in a dataset of 24 million rows, with only 0.1% fraudulent transactions, by applying various methods including undersampling and SMOTE, finding undersampling to be the most effective.
- Achieved optimal performance with the XGBoost model on the undersampled dataset, attaining a test recall score of 0.86 and an ROC AUC score of 0.96, thereby significantly improving the detection of fraudulent transactions.

TECHNICAL SKILLS

Programming Languages: Python, SQL, Java, C, C++

Other Tools: Microsoft Azure, AWS, Google Cloud Platform, Hugging Face, Cloud Dataflow, TensorFlow, PyTorch, Pinecone, LangChain, Redis, PySpark, MS SQL, DBVisualizer, Pandas, JMP, Microsoft Power BI, OpenCV, Kaggle, Heroku, Postman, Jupyter Notebook, MATLAB

Skills: Model Deployment, MLOps, Feature Engineering, Data Structures and Algorithms, Object-Oriented Programming, Statistical Analysis, Data Analysis, Data Visualization, Machine Learning algorithms, Computer Vision, Natural Language Processing