Saaransh Pandey

Austin, TX
LinkedIn

Email: saaransh@seas.upenn.edu
Mobile: +1 (215) 433-4966
Portfolio

Education

University of Pennsylvania, School of Engineering and Applied Science — Philadelphia, PA
- *Master of Science in Electrical and Systems Engineering, Major in Data Science* — *May 2025*
  Courses: *Applied Machine Learning, Big Data Analytics, Statistics for Data Science, Deep Generative Models* — GPA: 3.97

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

Work Experience

- **Aeronautical Systems Incorporated** — Virginia, United States
  Data Scientist — *Jun 2024 - Present*
  - Led a team to develop a Manufacturer/Distributor Predictor using Python and GPT-4, automating manual analysis, reducing processing time from hours to seconds, and achieving 90% accuracy.
  - Built a Competition Score Predictor with Azure AutoML, deployed via Azure ML Studio, automating strategic bid processes for customers.
  - Created a Similar NSNs Identification Tool using Pinecone for efficient data retrieval, deployed via Azure Endpoint, enhancing pricing justification and product alternative identification.
- **UnitedHealth Group – Optum** — Haryana, India
  Machine Learning Engineer — *Jul 2020 - Jul 2023*
  - Developed a classification model to detect early-stage opioid addiction, achieving a recall of 0.63 and specificity of 0.75, improving patient outcomes.
  - Built a Python automation script with Levenshtein-based fuzzy matching, reducing facility data validation time by 99% (14 days to 10 minutes) and enhancing database accuracy.
  - Led the Member Roster Management Project, developing a Power BI dashboard and automating SAS modules for streamlined roster oversight.
- **EshopBox Pvt Ltd** — Haryana, India
  Software Developer Co-Op — *Dec 2017 - Dec 2018*
  - Engineered a data pipeline with Apache Beam on Cloud Dataflow and a TensorFlow linear regression model for shipment/inventory predictive analysis in BigQuery
  - Developed a warehouse management app using NativeScript-Angular, launched on Google Play Store, reducing shipment processing time by 90% via barcode scanning.

Research Experience

- **University of Pennsylvania** — Pennsylvania, United States
  Research Assistant – Dr. Masao Sako | *Github* — *Nov 2024 - May 2025*
  - Built custom Simlib exposure libraries to enable accurate simulation of Roman supernova detection yield
  - Analyzed sky coverage efficiency for NASA's Roman Space Telescope using Python, AstroPy, and Healpy for pixel mapping, optimizing supernova detection.
- **Indraprastha Institute of Information Technology Delhi (IIIT-Delhi)** — Delhi, India
  Research Assistant – Dr. Vivek Bohara | *Github* | *Paper* — *May 2019 - Jul 2019*
  - Developed a Python and GNU Radio toolkit for simulating Visible Light Communication systems, published in the 2019 IEEE ANTS Conference.

Projects

- **WUDAC Spring 2025 Datathon (Won 2nd Place)** *Github*: Built an end-to-end analytics pipeline on 16 million+ sessions and transactions to segment users into four personas with K-means clustering, then trained an XGBoost model that delivered a 28% lift in conversions and 2.6× ROI for the top decile cohort (ROC–AUC 0.87). Tech: Python, Pandas, scikit-learn, XGBoost (Mar '25 – Apr '25)
- **Neural Network Model Compression (Model Pruning, Transfer Learning)** *Github* : Developed and evaluated global, layer-wise, and channel pruning on the VGG16 model, with global pruning achieving 45% sparsity while maintaining over 70% top-1 accuracy on the ImageNet validation dataset. Implemented iterative pruning and extended to transfer learning, retraining pruned models for optimized performance on the melanoma dataset in resource-constrained environments. Tech: Python, PyTorch, GPU/TPU (Aug '24 – Sep '24)
- **Credit Card Fraud Detection (Imbalanced Classification)** *Github*: Addressed a dataset of 24 million transactions with only 0.1% fraudulent entries by applying undersampling and SMOTE, finding undersampling most effective. Achieved a test recall of 0.86 and an ROC AUC of 0.96 using the XGBoost model, significantly improving fraud detection rates. Tech: Python, TensorFlow, PyTorch (Mar '24 – May '24)

Technical Skills

- **Languages**: Python, SQL, Bash, JAVA, C, C++
- **Tools & Frameworks**: Azure, AWS, Google Cloud Platform, TensorFlow, PyTorch, Pinecone, LangChain, PySpark, Power BI, OpenCV, Apache Beam, Cloud Dataflow, Pandas, Jupyter Notebook, MATLAB, Scikit, NLTK, Django, Flask, NodeJS
- **Skills**: Model Deployment, MLOps, Feature Engineering, Statistical Analysis, Data Visualization, Machine Learning, NLP, Computer Vision