Saumit Paul

+1 (984) 335-8991| sspaul3@ncsu.edu | LinkedIn | GitHub | Raleigh, NC

Applied ML Engineer with 3+ years of experience in building FDA-approved computer vision models for medical imaging.

EDUCATION

North Carolina State University 08/2025 - 05/2027
Masters of Science, Electrical Engineering GPA : 3.88
Related Coursework: Generative AI For Systems, Data Science For Signal Processing, Neural Networks

Manipal Institute of Technology 07/2018 - 08/2022
Bachelor of Technology, Electrical & Electronics Engineering GPA: 8.15

PROFESSIONAL EXPERIENCE

Data Scientist, DeepTek.ai 06/2022 - 07/2025

- Developed chest X-ray computer vision models (UNet, XCeption, FPN, DenseNet) using TensorFlow/Keras, achieving over 90% AUROC for pleural effusion detection.
- Led regulatory and clinical AI validation for US FDA, Thai FDA, HSA, and CE, including an MRMC study with 24 radiologists, demonstrating statistically significant improvement in classification and localization (wAFROC).
- Built and directed ML experiment evaluation platform (Fermi), integrating inference, metrics, and MLflow-based experiment tracking, reducing evaluation time from hours to under 5 minutes.
- Maintained MongoDB database of over 1.8M chest X-ray scans, regularly importing new annotations, enforcing schema consistency, and ensuring annotation quality.

Automation Intern, Anheuser-Busch InBev 01/2022 - 05/2022

- Designed Power BI Dashboard on employee work patterns and built SQL Database.
- Performed data wrangling on Task Mining API data with over 1M records weekly in Azure Data Pipeline for analysis.

Data Science Intern, iQGateway 07/2021 - 08/2021

- Developed model diagnostic visualization for AutoML Pipelines, enhancing interpretability.

Machine Learning Intern, Prime Focus Technologies 08/2020 - 10/2020

- Wrote Python scripts for automated audio processing and feature extraction.
- Evaluated deep learning models for audio source separation to guide production model selection.

PROJECT EXPERIENCE

FunSearch - Cache Replacement (ECE 592 GenAI For Systems)

- Built LLM-driven evolutionary search (OpenAI API, Ollama) to generate and iteratively refine C++ cache replacement policies, achieving 0.414 average IPC against a subset of SPEC CPU 2006 benchmark workloads.
- Utilized RAG over SQLite with graph-structured JSON memory (policy nodes + IPC-delta edges) to guide branching exploration and designed surrogate heuristic to prune weak candidates before full evaluation reducing evaluation time per rejected policy from ~45 to ~30 minutes.

ARoS Lab Volunteering (NCSU)

- Developed transformer-based methods (wav2vec2, WavLM, DistilHuBERT) for dementia and speech impairment detection from audio, achieving 85% F1 score with <5% standard deviation on DementiaBank Pitt corpus.
- Trained classifiers including SVM, random forest, and neural networks on transformer embeddings, optimizing for edge deployment without ASR transcripts.

PUBLICATIONS

- A Comprehensive Evaluation of DeepTek CXR Analyzer in Detecting and Localising Suspicious Findings in Chest X-rays (https://dx.doi.org/10.26044/ecr2024/C-23322)
- Artificial Intelligence as a Proficient Tool in Detecting Pulmonary Tuberculosis in Massive Population Screening Programmes: A Case Study in Chennai, India (https://doi.org/10.2185/jrm.2024-015)

TECHNICAL SKILLS

Programming Languages: Python, R
ML & AI: TensorFlow, Keras, PyTorch, Transformers, Ollama, Groq, LangChain, Scikit-Learn, Machine Learning, Deep Learning, Computer Vision, NLP, Generative AI, RAG
Data & Pipelines: Pandas, NumPy, SQL, MongoDB, Azure Data Factory, Matplotlib, Seaborn, OpenCV, MLFlow
Version Control & MLOps: Git, DVC