Saiprakash Nalubolu

snalubolu@binghamton.edu | (425) 364-9845 | linkedin.com/in/saiprakashnalubolu | github.com/prakashnalubolu | Portfolio

Summary

An MS CS(AI/ML) grad with 3+ years of industry experience, delivered agentic AI, RAG applications, built reliable MLOps/CI/CD pipelines, conducted research in histopathology imaging and distributed GPU training/inference.

Education

Binghamton University, State University of New York   *Jan 2024 - Dec 2025*
M.S. in Computer Science(AI track)   CGPA: 3.84/4.00
Relevant Coursework: Machine Learning, Deep Learning, Artificial Intelligence, Human Computer interaction, Design patterns.
Vellore Institute of Technology (VIT), Vellore   *Jul 2016 - Jun 2020*
B.Tech in Computer Science and Engineering
Relevant Coursework: Data Structures & Algorithms, Operating Systems, Database Management services, Software Engineering.

Experience

Software Development Engineer in Test (Intern), MiHIN (Remote, Michigan)   Sep 2025 – Present
- Building Selenium/Behave and API test suites for MiHIN's Health information Exchange, validating interoperability on AWS.

Research Assistant, Binghamton University   Sep 2024 – May 2025
- Saved **33%** of processing time by automating histopathology image segmentation pipeline(contrast enhancement, H&E deconvolution, watershed segmentation) for high resolution WSI images.
- Developed a deep learning survival model for lung cancer by fusing TCGA clinical data with histopathology features; applied SHAP to explain drivers(TNM stage, ECOG, NLR/PLR, morphology).
- Leveraged HPC GPU cluster for distributed training and inference of large scale histopathology models, optimizing runtime with PyTorch(FSDP) and CUDA on a cluster containing NVIDIA A100, H100, T4 GPUs.

Software Engineer in Test Automation, CITI bank (Through TCS)   Mar 2021 – Dec 2023
- Boosted automation test coverage from 66% to 87% by engineering Selenium Java frameworks with Cucumber(BDD) and embedding them into Jenkins/Maven CI/CD pipelines.
- Prevented 400+ production defects by developing 1000+ automated tests for Citi's microservices based servicing platform.
- Led a team of 5 to ensure accurate communication to 30M+ users by automating customer servicing message workflows.

Projects

Kitchbot: Agentic Meal Planner Github Link   Jun 2025 - Aug 2025
- Built a ReAct agent with LangChain + Gemini (LLM) for reliable tool routing and faster responses.
- Enforced typed tool contracts with Pydantic to ensure structured I/O, prevent invalid calls, and simplify debugging.
- Kept conversations fast and consistent by summarizing older history and storing shared agent state, cutting token use and latency.

Samethalu.com: RAG search for Telugu Proverbs website   Mar 2025 - May 2025
- Live RAG service handling 100-250 daily queries; maps free text situations to proverbs via a low latency API + web UI.
- Built embeddings + vector search with metadata filters(translation, meaning, usage) for precise, explainable top k matches.
- Docker containerized API + web UI with usage analytics and multilingual search for relevance tuning.

Social Media Toxicity & Engagement Analysis (Reddit + 4chan) Github Link   Aug 2024 - Dec 2024
- Analyzed 9M+ Reddit/4chan posts and comments with Transformers(BERT/RoBERTa) + ModerateHateSpeech API.
- Found 2.3× higher toxicity in top engagement posts and visualized trends via Flask/Plotly dashboards.

Handwritten Digit & Text Recognition (Android, PyTorch Mobile) Github Link   Aug 2024 - Sep 2024
- Designed PyTorch models(custom CNN vs VGG16 fine tuned), selecting the best accuracy-latency profile for mobile OCR.
- Prototyped on device OCR for Android using PyTorch Mobile + CameraX; achieved 0.9 FPS (1.1 s/frame)
- Trained DL models in Google Colab on MJ Synth(300k); compared GRU, LSTM variants. Tuned optimizers/lr to reduce overfitting.

Technical skills

Languages: Python, Java, SQL, Git, Linux, Gherkin
AI/ML & LLM: PyTorch, TensorFlow/Keras, scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Transformers(BERT, RoBERTa), LangChain, RAG, embeddings/vector search, CUDA/cuDNN, OpenCV, SHAP, CNN/CRNN, PyTorch Mobile
Platforms & Delivery: Docker, REST APIs, Flask, Streamlit, PostgreSQL, Faktory; CI/CD (Jenkins, GitHub Actions, Maven), AWS; CameraX; Selenium WebDriver,Cucumber/Behave(BDD)

Achievements

- Finalist, XFoundry Horizons challenge(NASA NEXPLORE 2040): Defined a problem statement that has huge impact on Earth and Mars/Moon. Provided reasonable solutions to NASA leaders, Industry experts and Inventors.
- Shipped a TSA Checkpoint Advisor web app in 7 hours; instant search for items allowed in luggage for all sorts of travel.