Mukund Yadav

+1 (540) 558 3922 | mukundy8@gmail.com | Personal Website | GitHub | LinkedIn | VA, USA (open to relocation)

Education

Virginia Tech — Blacksburg, VA
Master of Science in Computer Engineering — *Jan 2023 – Dec 2024*
- MS Thesis: Machine Learning, Image Optimization (Computer Vision, NLP, LLMs).
- TA: ECE5505 VLSI/Digital Testing, CS2506 Computer Architecture, ECE2564 Embedded Systems.
- Courses: Deep Learning, ML on Graphs, Optimization for Machine Learning, Stochastic Analysis, Computer Architecture.

Purdue University — West Lafayette, IN
Bachelor of Science in Electrical and Electronics Engineering — *Aug 2018 – May 2022*
- Capstone: Industrial Internet of Things Sensors & Devices (C++).
- TA/Tutor: Multivariate Calculus, Linear Algebra, Multivariate Statistics, Differential Equations.

Experience

Software Engineer-Machine Learning — Oct 2025 – Present
Microsoft — *Blacksburg, VA*
- Implement cloud based solutions with **C++**, **Python** and **PyTorch** for machine learning models, optimizing latency and memory usage.
- Optimized the performance of **LLMs** and **Foundational Models**, reducing latency in production.

Machine Learning Researcher — Jan 2024 – Dec 2024
Synergy Lab, VT, Publication Link — *Blacksburg, VA*
- Developed and deployed **Deep Learning** (DenseNet and Convolutional) models and **end-to-end machine learning** pipelines using **Python-PyTorch** for large scale image enhancement, leading to a journal publication in IEEE.
- Enhanced model performance by optimizing multi-level VGG loss functions, improving SSIM and MS-SSIM from 0.707 to 0.9135; further applied **transfer learning** to reach 0.9914 (40.2% improvement).
- Integrated PyTorch's **DDP** library for **distributed parallel processing** and training on multiple GPU node clusters to reduce training time from 42 hours to 23 hours (45.2% reduction).
- Publication: M . Yadav, W. Feng, G. Cao, "Impact of Loss Function on COVID-19 CT Image Denoising," IEEE Transactions on Radiation and Plasma Medical Sciences, 2024.

Machine Learning Engineer — Jan 2023 – Dec 2023
Wireless Dept, VT — *Blacksburg, VA*
- Created a **full-stack machine learning** solution via **Reinforcement Learning** with **Python** and **C++** to optimize **time series signal** detection, increasing SNR from 0.19 to 0.34 (79% increase) over classical methods.
- Engineered a stochastic algorithm leveraging the GLRT, successfully reducing false alarm probability by 60%.
- Derived the formula for the threshold of false alarm probability and applied **Proximal Policy Optimization**, increasing agent's total rewards by 25% and improved detection-to-false alarm ratio by 30%.
- Revamped the PPO training pipeline with **C++ multithreading-mutex/Eigen**, cutting training time by 45%.

Machine Learning Engineering Intern- NLP — Aug 2022 – Dec 2022
IUPUI — *Indianapolis, IN*
- Spearheaded the development of a voice assistant using **Python** and **NLP**, enhancing **Automatic Speech Recognition**.
- Devised a data-driven and full stack ML solution strategy by increasing training data size utilizing **pandas** and **numpy** by 30% resulting in overfitting reduction by narrowing accuracy gap by 10%.
- Constructed and trained an **LSTM Deep Neural Network** in PyTorch and **NLTK**, raising accuracy by 22%.
- Pushed the model on AWS **(EC2, S3, Lambda)**, providing accurate and timely medical information to 100+ senior citizens.

Projects

C++ Low Latency Trading System | C/C++ | *GitHub Link* — May 2024 – Aug 2024
- Designed a low latency trading system in C++, boost.asio, processed market data, and applied strategies like moving average crossover, risk management and order execution modules to monitor trading positions with trading latency of 1 millisecond.

Contradictory, My Dear Watson | NLP/Sentiment Analysis/Python | *Kaggle Link* — Jan 2023 – May 2023
- Developed an NLP-based XLM transformer model for a Kaggle competition using RoBERTa for sentiment analysis.
- Attained the model's accuracy of 93%, ROC-AUC score of 0.95 and finished in the 99th percentile of Kaggle leaderboards.

Technical Skills

ML: PyTorch, TensorFlow, End-to-end ML, Recommendation System, Transformers, Computer Vision, NLP, On Device ML
Cloud & DevOps: AWS (EC2, S3, RDS, EKS), Docker, Kubernetes, Distributed Systems / Large-scale ML Systems, CI/CD
Systems & Tools: Linux, Git, multithreading, Eigen, REST APIs,Data Structures and Algorithms
Mathematics: Linear Algebra, Probability & Statistics, Stochastic Calculis, Time Series Analysis
Languages: Python, C, C++, Bash, SQL