MIRAT SHAH

shah.mir@northeastern.edu | +1 (617) 888-6436 | linkedin.com/in/mirat-shah | github.com/miratshah27

EDUCATION

Northeastern University — **Boston, USA**
Master of Science in Computer Science – GPA: 4.0/4.0 — *Expected - May 2026*

- **Teaching Assistant:** Object Oriented Design (CS 3500)
- **Courses:** Natural Language Processing, Machine Learning, Database Systems, Algorithms, Programming Design Paradigm

Sardar Patel Institute of Technology — **Mumbai, India**
Bachelor of Technology in Information Technology – GPA: 3.6/4.0 (Awarded 80% Tuition Scholarship) — *Graduated - May 2024*

- **Courses:** Machine Learning, Probability and Statistics, Data Visualization, Data Science, Software Engineering, Data structure, Operating System, Database, Object Oriented Programming, Cloud Computing, Data Mining, Distributed Computing

SKILLS

Languages: Python, C++, C, MATLAB, Java, JavaScript/TypeScript, HTML, CSS
Database: SQL/MySQL, PostgreSQL, MongoDB, DynamoDB, Oracle, AWS RDBMS, Redis
Machine Learning and Deep Learning: PyTorch, Tensorflow, Caffe, NumPy, Pandas, Scikit-Learn, Keras, OpenCV, Large Language Models, Llama, BERT, NLTK, Hugging Face, JAX, Tableau, MLOps, SVM, CNN, RNN, LSTM, Transformers, LangChain, RAG, AB Testing
Tools: Linux/Unix, Git/Github/GitLab, Visual Studio, Jupyter, Microsoft Office, Google Apps
Others: AWS EC2, Sagemaker, S3, PySpark, Hadoop, Azure, Containerization, Docker, Kubernetes, Agile, CI/CD, REST API, Junit, Kafka

EXPERIENCE

Amazon Web Services — **East Palo Alto, USA**
Machine Learning SDE Intern (Incoming) — *May 2025 – Aug 2025*

Indian Institute of Technology (IIT) Bombay [Link] — **Mumbai, India**
Applied ML Research Assistant at SPANN Lab — *Jan 2023 – Jun 2024*

- **Led a team of 4** to develop lightweight Multiresolution and Deep Learning-based skull-stripping and brain tissue segmentation model using **Python and Tensorflow**
- Designed an architecture that reduced trainable parameters from **1.4 million to 22 thousand** and FLOPS from **192 million to 14 million**
- Achieved Dice scores of **0.96 and 0.97** on IBSR and NFBS datasets, ensuring **99.8% shift-invariance** via Adaptive Polyphase Sampling
- Pioneered the only library by developing a **Python wrapper for C++ codebase**, enabling seamless integration of Shearlet Transform with CNN, enhancing geometric feature extraction
- Conducted Explainable AI studies using a statistical modeling and model-agnostic occlusion-based Shapley Additive Explanations (SHAP) method, validating the model's decisions across diverse brain slices
- Developed a **scalable data pipeline** for model training, processing **10TB+ of medical imaging data** across multiple GPU clusters. The project resulted in a publication at IEEE International Conference on Acoustic Speech and Signal Processing (**ICASSP**)

Carnegie Mellon University [Link] — **Pittsburgh, PA, USA**
Applied ML Research Assistant at Xu Lab, Computational Biology Dept. — *Mar 2023 – Mar 2024*

- Enhanced Cryo-ET subtomogram classification by applying **Transfer Learning**, using 3D-CNN and Video Vision Transformer (**PyTorch**) pre-trained on large-scale video datasets (Kinetics 400)
- Achieved **highest 87.38% accuracy and 99.05 AUC** on simulated dataset and **99.57% accuracy and 99.99 AUC** on real Cryo-ET dataset
- Fine-tuned 3D-ResNet-34, resulting in a **19.19% increase** in accuracy compared to random initialization. **Leveraged Pandas** for data analysis, evaluating model performance (accuracy, AUC) and generating statistical summaries, confusion matrices, and ROC curves
- **Reduced training effort by 50%** using pre-trained weights from video datasets, enhancing model performance and efficiency
- Verified with Grad-CAM that video-initialized models more accurately capture subtomogram regions than randomly-initialized models

Cyphertree Technologies [Link] — **Pune, India**
Software Engineer Intern — *Apr 2022 – Jul 2022*

- Automated real-time updates for a data collection tool by integrating Google Sheets and Trello APIs, using **NodeJS and AWS Lambda**
- Implemented Slack API integration using webhooks and **AWS Lambda triggers** to facilitate instant notifications in a RESTful architecture, **achieved 100% automated workflow** using **ELK (ETL) stack**, and building a **CI/CD pipeline** for streamlined deployment
- **Reduced user response time by 3 seconds** with **NodeJS** scripts through dynamic webhook generation
- Modernized form handling with **ReactJS and Material UI**, building a modular frontend interface that rolled out in beta to **1000+ users**

PROJECTS

Llama2-QLoRA: Efficient Multilingual Sentiment Analysis *(Tech: PyTorch, LLM, Llama, Transformers, Hugging Face, Numpy)* [Link]

- **Fine-tuned Llama2-7B** using **QLoRA** for multilingual sentiment analysis, leveraging **PyTorch** and understanding of **Transformers** to achieve a 30% increase in Test AUC with only 2% of the dataset, optimizing performance through **Parameter Efficient Fine-Tuning** (PEFT) and **4-bit quantization**. Utilized **Numpy** for data preprocessing and model evaluation

Myers-Briggs BERT: Classify & TextGen *(Tech: PyTorch, BERT, Hugging Face, NLTK, spaCy, BeautifulSoup)* [Link]

- Finetuned a **BERT** model for MBTI personality type prediction with **0.48 exact match** and **0.86 for partial match** accuracy. Scraped a dataset of **68,733 posts**. Implemented a BERT-based model for generating personality-specific text**,** achieving a **loss of 0.02**