MOULIK SHAH

+1 646 209 1462 | mps10088@nyu.edu | Portfolio | moulik-shah | MoulikShah

Education

New York University **September 2024 – May 2026**
Masters in Computer Engineering **GPA: 4.0/4.0** *New York, NY*

- Coursework: Machine Learning, Natural Language Processing, Computer System Architecture

University of Mumbai **August 2019 – June 2023**
Bachelor of Technology in Electronics and Telecommunication Engineering **GPA: 8/10** *Mumbai, India*

- Coursework: Data Structures and Algorithms, Object Oriented Programming, Python, Mathematics, Database Management System, Computer Networks, Neural Networks, Big Data Analytics, Artificial Intelligence, Image Processing

Experience

MIKO **July 2023 – August 2024**
Data Scientist *Mumbai, India*

- Engineered a **recommendation engine** for skill-based games using **reinforcement learning** with **PyTorch**, employing **unsupervised clustering** techniques to segment users by engagement behavior. Conducted **A/B testing** to validate model efficacy, achieving a **25%** increase in user retention and reinforcing community safety through behavior-based recommendations.
- Led research and development of **multilingual NLP pipelines**, integrating **embedding models** and **Named Entity Recognition (NER)** for enhanced speech recognition. Refined pipeline through iterative experimentation with various embeddings, achieving a **30%** accuracy boost and supporting inclusive, language-diverse interactions.
- Enhanced **Open-Domain Question Answering system** with **Dense Passage Retrieval (DPR)** and **Retrieval-Augmented Generation (RAG)** models, achieving a **15%** reduction in unanswered queries through iterative model configuration and validation, supporting real-time NLP applications in **multilingual environments**.
- Integrated **LLMs** and **advanced embedding techniques** into chatbot functions, applying risk control and **anomaly detection** methods to analyze user behavior and uphold platform integrity by flagging suspicious interactions.
- Built a high-performing **song search engine** and **personalized recommendation system** using the iHeart Music dataset, using **feature detection** to analyze song attributes and elevate user experience in web-based applications, supporting platform safety by prioritizing user-friendly interactions.

Sykes and Ray Equities **October 2022 – December 2022**
Quantitative Research Intern *Mumbai, India*

- Collaborated on developing and tested trading models using **Monte Carlo simulations** and **Alpha Hedging**, conducting experimental **backtesting** and leveraging **graph theory** to analyze patterns—achieving a **12%** boost in portfolio performance.
- Automated portfolio strategies like Straddle and Iron Condor with **Pandas** and analyzed NSE datasets through **SQL** and **APIs**, achieving **85%** model accuracy through advanced **feature engineering** and **risk assessment strategies**.

Bipolar Factory **January 2022 – March 2022**
Machine Learning Intern *Mumbai, India*

- Architected end-to-end **deep learning** projects with **CNNs** in **TensorFlow**, optimizing model accuracy through **iterative tuning** and deploying on **AWS** for real-time, scalable applications.
- Developed **YOLOv4** and **SSD object detection models (89% accuracy)**, deploying via **Flask APIs** and validating for real-world performance, collaborating on **frontend** integration with **React.js** and **Node.js**.

Myraa Technologies **January 2022 – February 2022**
Deep Learning Engineer Intern *Mumbai, India*

- Fine-tuned **Random Forest** and **CNN** models for HR analytics, resulting in a **92%** increase in decision-making efficiency through enhanced data-driven insights.
- Implemented a **Binary Classifier** using **TensorFlow** to detect disaster tweets with **91%** accuracy, integrating **React.js** to create an intuitive UI for data visualization.

Ares Data Private Limited **December 2021 – January 2022**
Software Engineering Intern *Mumbai, India*

- Designed a **scalable microservices architecture** with **Spring Boot and Node.js**, boosting data retrieval speed by **25%** via **Redis** caching and load balancing. Integrated **MongoDB** to enhance backend performance in real-time, data-intensive applications.
- Implemented an automated testing suite with **JUnit** and **Mocha**, increasing code coverage by **30%** and optimizing the QA. Set up **CI/CD pipelines** in **Jenkins**, accelerating deployments and ensuring consistent quality.

DJS Phoenix **March 2020 – November 2021**
Head of Coding Department *Mumbai, India*

- Created an autonomous drone system with the help of **Python** and a **OpenCV** for my team's Immediate Aid Assistance Drone (IAAD) which helped in healthcare services and became the **finalist in the Indian Institute of Project Technology Competition**.
- Managed a team of 10 developers in designing and implementing an **autonomous object tracking system**, overseeing coding tasks and ensuring project milestones were met efficiently.

AIESEC **February 2020 – January 2021**
Corporate Relationship Strategist *Mumbai, India*

- Managed **B2B lead generation**, **sales**, and **client relations** for AIESEC's Incoming Global Talent program, collaborating with **MNCs** and **SMEs** to drive **successful internship placements**.
- Established **customer-centric standards** and efficient **communication channels** to **enhance client experience**, improve **team productivity**, and ensure alignment with **stakeholder needs** for seamless project execution.

Projects

Cross-Domain Adaptation through Soft Prompt Tuning in Low-Parameter Language Models **November 2024**

Python, Pandas, LoRA, LLMs, Linux, GitHub

- Optimized cross-domain adaptation for low-parameter language models using soft prompt tuning (e.g. prefix tuning) on **T5-large**, achieving over **90%** accuracy in sentiment analysis across multiple domains with less than **1%** of model parameters trained.
- Improved few-shot learning in resource-constrained models through techniques like **LoRA** and **soft prompting**, enhancing sentiment analysis accuracy while reducing the need for extensive model fine-tuning.

LLM-Powered Financial News Sentiment Trader **January 2024**

Python, Pandas, TensorFlow, BERT, UNIX

- Created LLMSentiments Trader with **DistilBERT** and **FinBERT** for real-time **NER** and sentiment analysis on financial news, supporting investment decisions for hedge funds. Integrated Point72's **CSP** library, improving stream processing efficiency.
- Implemented a sentiment-based trading strategy that yielded a **0.21%** return with a sharpe ratio of **0.84**, outperforming traditional methods by analyzing a **20GB dataset** of global news to track sentiment-driven stock movements.

Smart Interactive Marketing **February 2023**

Python, TensorFlow, CNN, React.js, Node.js

- Built a **CNN** model for fashion product classification using the DeepFashion Dataset **(800k** images), achieving **88%** accuracy and integrating it into a **React.js** application for real-time recommendations, which boosted product engagement by **30%**.
- Created an **Object Detection System** leveraging **ResNet50** and **VGG-16**, achieving **62%** accuracy and enhancing product recommendations by **20%** through effective image analysis and feature extraction.

Smart Mart **October 2022**

Python, TensorFlow, YOLO, React.js, Flask

- Developed a Smart Shopping Basket using **YOLOv3** for accurate product identification **81%** accuracy, connecting it with an e-commerce web platform via **React.js** and **Flask** to streamline user interaction.
- Deployed a real-time **recommendation system** that clustered user behavior and preferences, enhancing shopping experiences by **30%** through personalized product suggestions tailored to user segments.

Speech Emotions Recognition System **February 2022**

Python, TensorFlow, HuggingFace, Flask, Gradio

- Engineered a **CNN**-based speech-emotion recognition system with **82%** accuracy using **TensorFlow**, **MFCC features**, and **Wav2Vec2** from **HuggingFace**, classifying emotions like happiness, sadness, anger, and fear from speech.
- Implemented a real-time emotion classification via **Gradio** app and **Flask API**, increasing user engagement by **18%**, leveraging advanced **signal processing** techniques like MFCCs and Fourier Transform.

Extra-Curricular Activities

DJS Beats **Jan 2020 – Aug 2020**

Marketing Associate *Mumbai, India*

- Initiated contact with different companies to seek **sponsorship** for the events organised by our Student Forum.
- Coordinated **data tracking**, events, interviews, and managed **finances**, including budget monitoring.

Technical Skills

Languages: C, C++, C#, Java, Go, Python, R, SQL, HTML, CSS, JavaScript, TypeScript, Scala, Matlab, Julia, Kotlin, Swift
Frameworks: TensorFlow,PyTorch,Keras,OpenCV,YOLO(v3/v4),Django, Matplotlib,NumPy,Pandas,Flask,Roboflow,Gradio,Scikit-learn, React.js, Node.js, AngularJS,Three.js,Next.js,Tailwind CSS,Spring Boot,LangChain,Flutter,Bootstrap,Hadoop,BERT,RNN,CNN,Redux
DevOps & Cloud:Docker,Jenkins,Spark,AWS,Jira,Postman,GCP,Hive,NVIDIA CUDA,Snowfalke,Azure,Kubernetes,Airflow
Databases: MySQL, PostgreSQL, SQLite, Firebase, MongoDB, Oracle
Developer Tools: Linux, Lucid, Figma, HuggingFace, Tableau, Power BI, Microsoft Excel

Achievements

- Finalist at Indian Institute of Project Technology
- The Math Company Triathon Competition Winner
- Participated in State Level Competition DJ Spark
- Participated in Rubix 22- 43 Hour Online Hackathon which was organised by computer society of india student chapter of Thadomal Shahani Engineering College, Mumbai.
- Represented Maharashtra at National Level Throwball Competition