Simardeep Singh Mehta

+1-934-222-8000 | mehtasimardeepsingh@gmail.com | linkedin.com/in/simarnyu | github.com/simarmehta | New York City

Education

New York University — Brooklyn, NY
M.S. Computer Science; GPA: 3.94/4.0 — *Sept. 2023 – May 2025*

- Selected Coursework: Big Data, Cloud Computing, Machine Learning, Design and Analysis of Algorithms
- NYU Academic Achievement Award – Top 5% in graduate program — Merit-Based Scholarship Recipient

Amity University, Uttar Pradesh — Noida, India
Bachelors in Computer Science; GPA: 3.77/4.0 — *July 2019 – June 2023*

Experience

Machine Learning (MLOps) Intern — New York, USA
American Institute for Research — *May 2024 – Aug. 2024*

- Architected and executed a Gen AI pipeline using Flask and ChromaDB, integrating with Open source Llama model and Azure OpenAI for an ensemble Retrieval Augmented Generation (RAG) system, resulting in a 25% improvement in document retrieval efficiency.
- Orchestrated end-to-end CI/CD workflows with automated testing and Kubernetes deployments, reducing deployment times by 20%.
- Designed Helm charts for Kubernetes, facilitating version-controlled, templated deployments of the microservices architecture optimizing environments for reliability and performance.

Data Science Intern — Gurugram, India
Ernst & Young (EY) — *June 2022 – Aug. 2022*

- Built an ETL pipeline using AWS Glue and Lambda to process 3TB+ of financial data daily with 99.9% accuracy.
- Implemented a distributed data processing system using PySpark on EMR, reducing data processing time by 40%.
- Developed a pipeline using Python and Pandas to analyze SAP transactional data, implementing rule-based and statistical methods to detect anomalies.

Software Engineering Intern — Bengaluru, India
OnePercent Software LLC — *April 2022 – June 2022*

- Collaborated with Data Scientists and Machine Learning team to implement and scale a reinforcement learning model using TensorFlow to dynamically adjust game difficulty in a mobile game, leading to a 25% increase in user retention.
- Engineered and deployed a microservices architecture for backend using Docker and Kubernetes, improving scalability and reducing server costs by 20%.
- Enhanced efficiency and user satisfaction by 15% through bug fixes, unit testing, feedback mechanisms, and developing a real-time analytics pipeline using the ELK (Elasticsearch, Logstash, Kibana) stack.

Technical Skills

Programming Languages: Python, C++, SQL (PostgreSQL, MySQL), JavaScript, Dart, HTML/CSS
Frameworks: PyTorch, TensorFlow, Scikit-learn, Flask, Django, FastAPI, OpenCV, Keras, ResNet, VGG16
Developer Tools: AWS, Docker, ChromaDB, MongoDB, Kubernetes, Helm, Terraform, Git, Elasticsearch, Logstash, Kibana
Libraries: Pandas, NumPy, Matplotlib, Seaborn, BeautifulSoup, NLTK, HuggingFace, SpaCy, FastAI

Projects

EaTexas - Food Delivery App | *AWS Step Functions, Lex, SageMaker*

- Implemented a multi-modal ML pipeline using AWS Step Functions, integrating Lex with BERT for text analysis while leveraging AWS Rekognition alongside ResNet for advanced image processing of food items.
- Designed a real-time monitoring system using AWS CloudWatch and Grafana, setting up alerts and dashboards for the admin panel.
- Engineered a scalable, secure, and serverless architecture using AWS Lambda, API Gateway, and SQS, efficiently handling 1000+ concurrent users.

FashionGPT - Product Recommendation System | *FastAI, ResNet50, VGG16, CNN*

- Engineered an AI-driven fashion recommendation system combining ResNet50, VGG16, collaborative filtering, and content-based methods, improving product match accuracy by 30%.
- Implemented A/B testing framework using Python and Bayesian methods to optimize recommendation algorithms, resulting in a 15% increase in click-through rates.
- Automated data collection for 100,000+ items using BeautifulSoup, cutting manual entry by 70%.