Satyam Chandrakant Chatrola

+1 (973) 905-1864 sc10247@nyu.edu linkedin.com/in/satyamchatrola github.com/Nightshade14 Portfolio GCP Profile

EDUCATION

New York University — New York City, NY
Master of Science in Computer Science (CGPA: 3.61 out of 4.00) — *September 2023 – May 2025*

Gujarat Technological University — Gujarat, India
Bachelor of Engineering in Computer Engineering (CGPA: 3.79 out of 4.00) — *June 2018 – June 2022*

EXPERIENCE

Machine Learning Engineer / Data Scientist — January 2022 – June 2023
Rapidops — *Ahmedabad, India*

Face Authentication and Authorization System *(Python, PyTorch, YOLO, MTCNN, FaceNet, Triplet Loss, Qdrant, OpenCV)*

- Spearheaded a face authentication and authorization system with **4 environments and 2 access points** while **leading a team of 3**, from **data collection through on-premises deployment and monitoring**; enhanced productivity tracking and premises security.
- Architected a **scalable microservice** utilizing YOLO, MTCNN, fine-tuned FaceNet model, and Qdrant to achieve **low-latency** authentication and authorization on **live video streams**, delivering a **0.93 F1 score** without need to retrain for new individuals.

AI-Powered Search and Recommendation System *(Python, PyTorch, Hybrid Recommenders, Apache Solr, Docker, FastAPI)*

- Boosted **conversion rate** by **23%** and **click-through rate** by **42%** with **collaborative, content, and market-basket** recommenders.
- Engineered search & recommendations with **custom tagger**s & **LTR** techniques with **Apache Solr** to serve results in **25ms**.

Natural Language (English) to SQL query generation *(Python, TensorFlow, Keras, T5, BERT)*

- Researched and benchmarked **State of the Art (SOTA) AI models** generating SQL from Natural Language with **76% EMA**.
- Experimented with Transformers (**T5**, **BERT**) generating **73% Exact Match Accuracy (EMA)** with **36%** faster inference.

SKILLS

Languages and DBs: **Python, SQL**, C, C++, Java, PostgreSQL, MySQL, MongoDB, Elasticsearch, Apache Solr, Qdrant (vector database)

AI and ML: SVMs, Gradient Boosted Trees, Clustering, Computer Vision, NLP, ANN, CNN, RNN, **Recommendation and Search Systems, Transformers**, Prompt Engineering, Generative AI (GANs, Diffusion models), **LLMs (RAG, PEFT, LoRA)**.

Data Science: NumPy, Pandas, Polars, Matplotlib, Seaborn, SciPy, Scikit-learn, **PyTorch, TensorFlow, Transformers**, AutoGluon, Optuna, Tableau, Statistical Modeling, **Hypothesis Testing, A/B Testing**.

MLE and Data Engineering: **AWS SageMaker, ONNX, TensorRT**, MLflow, Evidently, Apache (Hadoop, Spark, Airflow, Kafka).

SWE and Others: **AWS (EC2, Lambda, Step Function, S3, Autoscaling, IAM)**, REST APIs, Git, **Docker, Kubernetes, CI/CD (GitHub Actions, CircleCI)**, System Design, HTML, CSS, JavaScript, Node.js

RESEARCH EXPERIENCE

Approaches to Type 2 Diabetes Mellitus Prediction with Machine Learning and Deep Learning

- Researched AI techniques for Type-2 Diabetes Mellitus classification with **95.8% precision, 95.8% recall**, and **99.4% specificity**.

CERTIFICATIONS

- Graduated from **Udacity's AWS Machine Learning Engineer Nanodegree** with top remarks.
- Earned **Inferential Statistical Analysis with Python** Certification from the **University of Michigan**.
- Attained official **University of Michigan** Certification in **Applied Machine Learning using Python**.

PROJECTS (with Embedded GitHub Repository URLs)

Microservices webapp: Essay Evaluator *(Python, PyTorch, ONNX, TensorRT, FastAPI, AWS, MLflow, Evidently, Frontend)*

- **Calibrated** Transformers (**BERT**) and LLMs (**GPT-2**) with **PEFT techniques** (quantization), cosine-annealed learning rate and warm-up to assess essays, attaining a **Kappa Score of 81.7%** and **surpassing the Benchmark** score by **5.7%**.
- Orchestrated **2 fault-tolerant microservices** & inferred with **ONNX models** and **TensorRT** via **asynchronous REST APIs**.

ETL Data Pipeline migration to Spark *(Python, PySpark, AWS, MinIO, Databricks.Koalas)*

- Migrated **ETL pipeline** to **Spark** for **NYU's Open-Source** wildlife trafficking prevention project with **160% speedup**.

AWS Dog Breed Classifier *(Python, PyTorch, AWS (SageMaker, S3, Lambda, EC2, CloudWatch, IAM), Code Profiling, Monitoring)*

- Developed a **133** dog breeds image classifier with **92%** accuracy on AWS with **concurrent** and **auto-scaled** inference endpoints.

ACHIEVEMENTS

- Clinched **1st place** in INDRA-9 (**renewable energy contest**) through **smart irrigation system** for sustainable agriculture.
- Secured **1st runner-up** in the Kaggle O Predictor challenge, leveraging cutting-edge **predictive modeling techniques**.
- Delivered **Git - GitHub** and **Machine Learning workshops** for **23 summer interns**, **propelling 12 project contributions**.