Satyam Chandrakant Chatrola

Machine Learning Engineer

+1 (973)-905-1864 sc10247@nyu.edu https://www.linkedin.com/in/satyamchatrola/ https://nightshade14.github.io/satyamchatrola/
https://www.kaggle.com/satyamchatrola https://www.cloudskillsboost.google/public_profiles/8f8d70ce-809c-45e2-b6ba-022e3bc4cf9f

EDUCATION

New York University **September 2023 – May 2025, New York, USA**

- Master of Science in Computer Science
- Coursework: **Machine Learning**, **Computer Vision, Deep Learning, Big Data**, Design and Analysis of Algorithms, Computer Networking
- Cumulative GPA: 3.611 / 4.00

Gujarat Technological University **June 2018 – June 2022, Gujarat, India**

- Bachelor of Engineering in Computer Engineering
- Coursework: Data Structures and Algorithms, DBMS, Object-Oriented Programming, Artificial Intelligence, Data Visualization
- Cumulative GPA: 8.54 / 10.00 (WES course-by-course evaluation CGPA: 3.79 / 4.00)

WORK EXPERIENCE

***Data Scientist* @ Rapidops** **January 2022 – June 2023, Ahmedabad, India**

Smart Face Recognition-based Attendance System

- Spearheaded the development of a **smart face recognition-based attendance system** using a **10-megapixel camera** for **500+ employees**, enhancing productivity tracking and authentication processes.
- Leveraged **OpenCV, dlib** and **PyTorch** to train model with **Triplet loss** to generate distant embeddings based on **68 face landmarks.**
- Matched **face vector embeddings** based on similarity among database of faces and experimented with **vector databases** like **qdrant**.

AI-Powered Search and Recommendation System

- Developed AI-powered search & recommendation with **custom taggers** & **LTR** techniques with Apache Solr to serve results in **10ms**.
- Designed advanced data and **machine learning pipelines** with PySpark and reduced model training time by **60%**.
- Engineered a **recommendation system** with **market basket analysis** that boosted product interaction by **45%** and sales by **30%**.

Automating product catalogs and store for BigCommerce and Shopify

- Engineered **automation scripts** for BigCommerce and Shopify **store conversion** and **reduced manual labor** by **95%**.

SKILLS

- ***AI and Big Data:*** Python, TensorFlow, PyTorch, **Generative AI**, **Recommendation and Search Systems**, Diffusion Models, MLflow, **Transformers**, **LLMs with RAG, Prompt Engineering**, Computer Vision, Natural Language Processing, **PySpark**, **Hadoop**, OpenCV, dlib.
- ***Others:*** SQL, Git, Docker, **AWS [SageMaker, Lambda, S3, AutoGluon, Autoscaling, IAM]**, System Design, REST APIs, MongoDB, PostgreSQL, Apache Solr.

RESEARCH EXPERIENCE

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

- Authored a research paper on Machine Learning and Deep Learning techniques for predicting Type-2 Diabetes Mellitus, achieving a classification accuracy with **95.8% precision** and recall, and **99.4% specificity** using BRFSS data.

CERTIFICATIONS

- Udacity **AWS Machine Learning Engineer** Nanodegree.
- **Inferential Statistical Analysis** with Python offered by the **University of Michigan**.
- **Applied Machine Learning** in Python offered by the **University of Michigan**.

PROJECTS AND OPEN-SOURCE CONTRIBUTION

Automated AI-based essay evaluation with Transformers and fine-tuned Large Language Models (LLMs)

- Leveraged **Transformers like BERT** and **fine-tuned LLMs** like **GPT-2** to evaluate essays with a Kappa Score of **81.7%**.
- Utilized **Parameter Efficient Fine Tuning (PEFT)** techniques like **quantization** and dynamic learning rate with annealing and warm-up to **reduce training time by 50%** while maintaining the same performance.

Multiple Noise Source Identification in New York City

- Developed an AI model with ensemble of Machine Learning techniques to identify **10 noise sources** of NYC with **86%** accuracy.

Migrating ETL Data Pipeline to Spark

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

Training Deep Convolutional Generative Adversarial Network (DCGAN) to generate images of clothes

- Designed a DCGAN with 2 convolutional layers to **converge** smoothly around a **Saddle point**; generating relevant images.

ACHIEVEMENTS

- Secured **first place** in the prestigious **India's Next Development Renewable Energy & Astronomy (INDRA-9)** competition, presenting an innovative smart irrigation system in sustainable agricultural technology, outperforming over **100** competing teams.
- Achieved **first runner-up** in the **Kaggle O Predictor**, **a data science competition**, showcasing analytics and predictive modeling skills.
- Delivered **workshops** on Git/GitHub and Machine Learning, to **23 summer interns**, which facilitated their project contributions.