Satyam Chandrakant Chatrola

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

EDUCATION

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

- Master of Science in Computer Science
- Coursework: **Machine Learning**, **Computer Vision, Deep Learning, Big Data**, Design and Analysis of Algorithms, **Applied ML in Finance**
- 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 at Rapidops **January 2022 – June 2023, Ahmedabad, India**

Face Recognition and Authentication System

- Spearheaded the development of an internal **face recognition and authentication system** employing a **10-megapixel** camera for **500+** employees, enhancing **productivity tracking** and **premises security**.
- Revamped face embedding generation with **Triplet loss** to generate distant embeddings based on **68 face landmarks** on live video stream**.**
- Optimized performance and embedding retrieval speed by **30%** with **vector databases** such as **Qdrant.**

AI-Powered Search and Recommendation System

- Developed AI-powered search & recommendations 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%**.
- Strengthened **recommendation system** with **market basket analysis** that boosted product interaction by **45%** and sales by **30%**.

Automatic Product Catalog and Store Conversion System

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

SKILLS

- Regression, Classification, **Gradient Boosted Trees**, Computer Vision, NLP, Text Processing, Word Embedding (Word2Vec, BERT).
- Neural Networks, Recommendation and Search Systems, Transformers, **LLMs (RAG, PEFT, LoRA)**, Prompt Engineering, Generative AI.
- **Python**, NumPy, Pandas, Polars, Matplotlib, Seaborn, Scikit-learn, **TensorFlow**, **PyTorch**, OpenCV, NLTK, Transformers, MLflow.
- **PySpark**, **Hadoop,** Tableau, SQL, **Statistical Modeling, A/B Testing, Hypothesis Testing,** FastAPI, Flask, REST APIs, MongoDB, PostgreSQL.
- **AWS** (**SageMaker**, Lambda, S3, AutoGluon, Autoscaling, IAM), Git, Docker, System Design, Apache Solr, **CI/CD (GitHub Actions)**, C, C++.

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

- Graduated from Udacity's **AWS Machine Learning Engineer** Nanodegree with top remarks.
- Certified for **Inferential Statistical Analysis with Python** by the **University of Michigan**.
- Successfully passed **University of Michigan** certified **Applied Machine Learning in Python**.

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%**.
- Accelerated model training by **50%** with **Parameter Efficient Fine Tuning (PEFT)** techniques like **quantization** and dynamic learning rate with cosine-annealing and warm-up while maintaining the same performance.

Migrating ETL Data Pipeline to Spark

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

Multiple Noise Source Identification in New York City

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

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

- Designed a 2 layered DCGAN with smooth convergence around a Saddle point, generating relevant images.

ACHIEVEMENTS

- Secured **1st 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 **1st 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**, facilitating **12** project contributions.