

Mithilesh Chaurasiya

+91-9325732481

mithileshchaurasiya1627@gmail.com

My Portfolio

GitHub

LinkedIn

Professional Summary

Driven Machine Learning Engineer with 2.8 years of experience at Infosys, specializing in data engineering, ETL pipelines, and production ML systems. Hands-on experience building end-to-end Computer Vision and MLOps pipelines using Airflow, MLflow, and DVC. Passionate about designing scalable, reproducible, and real-world AI solutions.

Key Skills

- **Programming:** Python, Java MLflow, DVC, DagsHub, Docker, Minikube (Kubernetes), FastAPI, Streamlit, CI/CD (Jenkins, Azure DevOps)
- **Data Engineering:** Apache Airflow, Apache Spark, MongoDB, Snowflake, Hadoop
- **Data Analysis & Wrangling:** Pandas, NumPy, SQL Developer (Oracle)
- **Automation:** Selenium, UiPath
- **Machine Learning & Deep Learning:** PyTorch, Keras, Scikit-learn
- **Cloud:** Microsoft Azure
- **Version Control:** Git, GitHub
- **Deployment / MLOps:** Flask, HTML/CSS
- **Web Development:** Flask, HTML/CSS

Soft Skills

- **Problem-Solving:** Analytical mindset to optimize algorithms and data pipelines.
- **Collaboration & Adaptability:** Effective team player, quickly adapting to new technologies and collaborating across teams to integrate solutions in dynamic environments.
- **Client Interaction:** Experience gathering requirements and delivering data-driven solutions.

Education

Master of Computer Application (MCA) 2024 – Expected 2027
National Institute of Technology, Agartala

Bachelor of Science in Computer Science 2021
Thakur College of Science and Commerce, Mumbai
CGPA: 8.87/10

Higher Secondary Certificate (HSC) 2018
Utkarsha Vidyalaya & Junior College, Mumbai
Percentage: 64%

Secondary School Certificate (SSC)

2016

Jai Deep Vidya Mandir High School, Mumbai

Percentage: 75%

Professional Experience

Infosys Ltd.

June 2021 – January 2024

Senior System Associate (ETL Tester and Data Engineering)

- Supported production data pipelines powering ML and analytics systems, validating data integrity across Oracle, MongoDB, and Snowflake
- Built Python-based automation frameworks to validate data pipelines and APIs, reducing manual validation effort by 30
- Optimized end-to-end ETL pipelines for ML-ready datasets used in analytics and forecasting models.
- Validated APIs using Postman and supported model endpoint integration testing.
- Collaborated on model experiment tracking and dataset versioning workflows using MLflow and DVC to improve reproducibility of ML experiments.
- Participated in Agile/Scrum sprints as part of an 8-member engineering team; supported client interactions and feedback integration.

Selected Projects

Smart Traffic Management System (End-to-End CV + MLOps Project)

YOLO (Ultralytics), OpenCV, Airflow, MLflow, DVC, MongoDB, FastAPI, Streamlit

- Built a production-grade computer vision + MLOps system for automated traffic analytics.
- Implemented real-time vehicle detection and multi-object tracking for persistent ID assignment.
- Designed traffic intelligence modules for vehicle count, density, and flow estimation.
- Orchestrated complete ML lifecycle (data → training → inference) using Apache Airflow DAGs.
- Integrated DVC for large-scale dataset versioning (22GB IDD dataset) ensuring full reproducibility.
- Implemented MLflow-based experiment tracking with dataset hash → model version traceability.
- Added training deduplication logic to automatically skip duplicate experiments, reducing GPU cost.
- Built Streamlit dashboard for real-time traffic monitoring and analytics visualization.

- Designed modular, production-ready architecture following real-world ML system design principles.

Stock Market ETL Pipeline & Dashboard

Python, Airflow, MongoDB, Streamlit

- Built a complete ETL pipeline to fetch and clean stock market data, storing it in MongoDB.
- Automated daily workflows using Apache Airflow.
- Developed an interactive Streamlit dashboard showing stock price trends, SMA, and volume.
- Optimized MongoDB queries using indexing and caching for faster performance.

Real-Time Gender Detection

Python, TensorFlow, OpenCV, CNN

- Developed a CNN-based system to classify gender from facial images with over 90% accuracy.
- Implemented real-time face detection using OpenCV.
- Fine-tuned hyperparameters to optimize model performance.

MockAnalysis: Automated Mock Test Result Analysis System

Python, Selenium, Pandas, Matplotlib

- Automated scraping and analysis of mock test results.
- Visualized insights with Pandas and Matplotlib; reduced manual effort by 40%.
- Hosted the project on GitHub for portfolio demonstration.

BihariBlogger — Social Blogging Web App

Flask, Python, HTML/CSS, MongoDB, Bootstrap

- Developed a blogging platform with user authentication, CRUD operations, and responsive design.

Heart Disease Prediction ML Model

Python, Scikit-learn, Pandas, Matplotlib

- Built a machine learning model to predict heart disease using patient records.
- Performed feature engineering, classification, and performance evaluation.

Leadership & Extracurricular Activities

Machine Learning Team Lead — Developers and Coders Club (DCC)

National Institute of Technology, Agartala

- Leading the Machine Learning team responsible for mentoring junior students in ML, AI, and Data Science.
- Conducted sessions on Machine Learning fundamentals, Deep Learning, and Computer Vision.
- Guided students in building real-world ML projects and Kaggle-based learning.
- Coordinating with faculty mentors on research-oriented and project-based initiatives.
- Planning technical workshops, roadmap sessions, and ML awareness programs for NIT Agartala students.
- Volunteer, College Fest Committee — Assisted in planning and organizing technical events and cultural activities.

Certifications

- Accenture North America - Data Analytics and Visualization Job Simulation
- Infosys Certified Selenium with Python Automation Tester
- Infosys Certified Software Development Engineer in Test (SDET)
- Infosys Certified Digital Assurance Professional
- Infosys Certified Python Programmer
- Infosys Certified Data Management Associate
- Machine Learning Foundations — Great Learning
- Computer Vision Essentials — Great Learning
- Python for Machine Learning — Great Learning

Hobbies

- Reading
- Photography
- Traveling