

MRUDULA DESHMUKH

Contact: +1 (716) 390-8431 | mrudulad25@gmail.com | [LinkedIn](#) | [GitHub](#) | Buffalo, NY

WORK EXPERIENCE

- Endeavor Health Services | Data Analyst | Buffalo, NY** Sept 2025 – Dec 2025
- Developed and operationalized deep learning models (TensorFlow, Keras) on 500K+ healthcare records to forecast clinical outcomes with 92% accuracy, enabling real-time predictions for 50+ clinicians
 - Engineered automated ETL pipelines using Python, SQL, and Airflow to process complex healthcare datasets with 99.9% data quality, reducing processing time by 70%
 - Built production ML inference pipelines with Flask REST API and Docker hosted on GCP infrastructure, reducing inference time to under 100ms and increasing predictive accuracy by 15% through hyperparameter tuning
 - Created Power BI dashboards integrating EMR data to track patient outcomes and clinical KPIs across 5+ departments, enabling \$2M+ in identified cost savings
- University at Buffalo | ML Research Assistant | Buffalo, NY** Apr 2025 – Aug 2025
- Designed and optimized deep learning models in TensorFlow/Keras for behavioral pattern detection, reducing manual neuroscience analysis time by 70%
 - Constructed real-time ML inference pipelines with Python signal processing for closed-loop optogenetics experiments maintaining real-time processing (<100ms) while improving detection performance by 15%
 - Engineered scalable analytical pipelines for time-series data processing using Scikit-learn with cross-validation frameworks for reproducible modeling
- AI Adventures | Data Analytics | Pune, MH** May 2022 – Jun 2023
- Trained and evaluated classification models (Random Forest, XGBoost) in Scikit-learn, achieving 92% accuracy through k-fold cross-validation and feature engineering for churn prediction
 - Implemented end-to-end analytics pipelines with Python, SQL, and Apache Airflow, reducing report generation time by 60% and improving data reliability
 - Analyzed 100K+ customer records using SQL and Python to identify behavioral drivers, improving downstream model performance by 18%

PROJECTS

- Foodborne Illness Forecasting** Jan 2024- Apr 2025
- Developed time-series forecasting pipelines using ARIMA, Prophet, and ensemble methods, reducing prediction error by 21 points (RMSE: 281.2 → 260.2, MAE: 237.0 → 203.2)
 - Trained Random Forest classification models with geotemporal feature engineering, achieving 83.3% accuracy and 100% sensitivity for outbreak detection
 - Applied statistical techniques, including the PELT algorithm for structural break detection and segmented ARIMA modeling to support evidence-based public health decision-making
- Synapse Street - AI Multi-Agent Stock Analysis** Nov 2025
- Architected a multi-agent AI system using LangGraph processing 3,000+ entities with Qdrant vector search and embeddings, delivering a production-ready system in 24 hours
 - Implemented distributed data pipelines using Spark and PySpark for large-scale dataset processing, enabling production-grade experimentation and model evaluation
 - Optimized ML models through hyperparameter tuning and cross-validation, achieving 18% improvement in back-tested ROI
- Alzheimer's Disease Detection using Deep Learning** Oct 2023- Apr 2024
- Constructed CNN-based diagnostic models using TensorFlow/Keras on medical imaging datasets, achieving 92% classification accuracy with 14% improvement in diagnostic recall
 - Launched production ML solution on GCP using Docker and Flask REST API with model serving endpoints and monitoring for clinical evaluation workflows
 - Applied model explainability techniques (SHAP, LIME) for CNN decision interpretation, enabling clinician trust and adoption in regulated healthcare environments

EDUCATION

- University at Buffalo – State University of New York** Buffalo, NY
Master of Science in Data Science Aug 2024 – Dec 2025
- Savitribai Phule Pune University, AISSMS Institute of Information Technology** Pune, MH
Bachelor of Engineering in Electronics & Telecommunication Aug 2020 – May 2024

SKILLS

Programming: Python, SQL, R
Data Science & ML: Scikit-learn, TensorFlow, Keras, PyTorch, Classical ML (Random Forest, XGBoost), Time Series Forecasting (ARIMA, Prophet), Feature Engineering, Model Evaluation (Cross-Validation, RMSE, MAE)
Data Engineering & Pipelines: Apache Airflow, ETL/ELT Pipelines, Spark, PySpark, dbt
Databases & Warehousing: PostgreSQL, Snowflake, BigQuery
APIs & Deployment: Flask, FastAPI, REST APIs, Docker, GCP
Analytics & Visualization: Power BI, Tableau
Vector Search (Project): Qdrant, Embeddings, LangGraph
Healthcare Data (Domain): EMR/EHR Analytics, Epic, Cerner, HIPAA Compliance, Clinical KPIs, HEDIS/Stars Metrics