Malhar Dhopate

dmalhar.24@gmail.com github.com/MDhopate (954)-536-4846
linkedin.com/in/mdhopate mdhopate.vercel.app

Summary

Data Scientist with 3+ years of experience building scalable **cloud-native solutions** in analytics, automation, and backend systems. Skilled in **Python**, **SQL**, **Snowflake**, **AWS**, and ML, with cross-functional impact across research, operations, and BI. Focused on delivering ethical, **data-driven solutions** to real-world problems.

Education

Indiana University Bloomington   Bloomington, IN
Master of Science (MS) in Data Science, **GPA: 3.90/4.00**

Lynn University   Boca Raton, FL
Bachelor of Science in Data Analytics, **GPA: 3.91/4.00**

Certifications:

- **AWS**: Certified Machine Learning Engineer - Associate
- **Stanford University**: Machine Learning Specialization

Technical Skills

Languages: **Python**, **SQL** (Postgres, MySQL, SQLite), **R**, Java

BI & Analytics: **Tableau**, **Power BI**, Matplotlib, Seaborn, Altair, Plotly

Cloud & DevOps: **AWS** (EC2, **SageMaker**, S3, API Gateway, **Snowflake**, Step Functions, Amplify), Docker, GCP

Machine Learning: **TensorFlow**, **PyTorch**, Scikit-learn, BERT, spaCy, Generative AI, Embeddings

Experience

Finance Intern - Process Automation | SAP, Celonis, Power Automate, Microsoft Suite   May 2025
Hoerbiger Service, Inc.   Deerfield Beach, FL

- **Automated** financial reports for the Controller, reducing preparation time by **85%** & improving forecast accuracy.
- **Built Advanced Macros** & Power Automate workflows to eliminate manual data consolidation, saving 30+ hrs weekly for the finance team.
- Leveraged process mining to analyze SAP workflows, identify inefficiencies, & recommend changes increasing reporting speed by **35%.**
- Partnered with finance leadership to enhance budget tracking, enabling faster & more informed strategic decisions.

Gen AI/NLP Research Assistant | LangChain, ChromaDB, SQLite, RAG   May 2024 – Apr. 2025
IU College of Arts & Sciences   Bloomington, IN

- Developed & deployed an end-to-end **NLP pipeline** using Pinecone, & VAD for real-time conversations.
- Set up a **Retrieval-Augmented Generation** (RAG) system with speech input, vector search & response generation.
- Designed **consent-aware** architecture using SQLite for ethical model training on community-driven narratives.
- Analyzed 150+ user interactions to optimize model prompts & **reduce** idle session **errors by 80%**.

Data Analyst | Process Automation, Tableau   Aug. 2021 – Apr. 2023
Nationwide Automotive Services   Boca Raton, FL

- Implemented an **ML pipeline in Python** to model customer behavior, improving client engagement & retention.
- Automated billing operations with VBA, **reducing billing time by 72%** & **improving** overall efficiency by **35%**.
- **Reduced** overdue accounts by **$900k** by data-driven segmentation strategies & customized payment plans.
- Led a team of 5 analysts to **integrate ML** & automation tools into business operations, improving reporting accuracy & team **productivity** by **52%**.

Academic & Technical Endeavors

Ensemble Machine Learning Framework for Comment Classification | Python, TensorFlow, Git

- Implemented an **ensemble ML framework** in TensorFlow and Python to classify comments using LSTM, SVM, Random Forest, Convolutional Neural Networks & Logistic Regression.
- Improved model F1-score by **18%** through advanced NLP preprocessing & **custom feature engineering**.

Cloud Computing Platform for Encrypted File Transfer | Python, Node.js, AWS - S3, EC2, Amplify

- Built a **scalable system** (S3 + EC2) for **encrypted** file transfer; incorporated anomaly detection for file access logs.
- Streamlined role-based **access control** & token validation for enhanced **data security** across AWS-hosted services.