

# Rohanraje Bhosale

+1 (508)-509-5706

therohanrajbhosale@gmail.com

www.rohanraje.com

in LinkedIn

*Machine Learning Engineer with hands-on experience designing, building, and deploying production-grade ML systems that process large-scale sequence and language data. Proficient in Python, SQL, PyTorch, and Spark, with a strong foundation in building scalable data pipelines and leveraging foundation models (LLMs) for complex NLP tasks. Eager to apply deep analytical skills and experience with end-to-end model development to solve challenging problems in speech recognition and delight millions of users.*

## Work Experience

### Sahyadri Polytechnic Sawarde

June 2022 – January 2024

#### Assistant Professor (ML & Data Systems Focus)

- Developed and scaled automated systems using Python (Pandas, NumPy) and SQL for grading and plagiarism detection across a 150+ student cohort, improving feedback efficiency by 60%.
- Led 12+ student teams in building and shipping ML-powered projects for hackathons and research expos, focusing on practical application and model performance evaluation.
- Engineered a centralized data tracking system to monitor and analyze student engagement, improving resource allocation for technical workshops by 40%.

### Trismus Healthcare Technologies

January 2022 – June 2022

#### Machine Learning Engineer Intern

- Engineered scalable preprocessing pipelines for large-scale, high-resolution image datasets, reducing training time by 35% and ensuring data compliance for production workflows.
- Designed and evaluated CNN architectures (PyTorch) for classification, translating model performance metrics into actionable business insights for domain experts.
- Built robust quality assurance and data audit tools to maintain dataset integrity and model reproducibility, a critical component for shipping reliable ML products.

## Projects

### Interactive Resume with Conversational AI Assistant

www.rohanraje.com

- Engineered a voice-enabled conversational AI assistant for my personal portfolio using the Web Speech API for speech-to-text and text-to-speech functionalities.
- Implemented a custom rule-based NLU engine in JavaScript to parse user queries and map them to a knowledge base designed to answer professional inquiries about my skills and projects.
- Designed the system to provide an engaging, hands-on demonstration of my skills in NLP, conversational flow, and building user-centric AI applications.

### Agentic AI for Customer Support (Natural Language Understanding System)

- Built a production-ready conversational AI system using Retrieval-Augmented Generation (RAG) and foundation models (GPT-4) to automate 85% of Tier-1 customer queries.
- Engineered custom routing logic and fine-tuned prompts to ground model outputs, cutting resolution time and reducing human escalation rates by 40%.
- Implemented a continuous learning data pipeline to capture user feedback, enabling automated retraining and evaluation of LLM agents to improve contextual understanding over time.

### Agentic AI-Driven Market Intelligence (Large-Scale Data Processing)

- Designed a scalable agent framework (LangGraph, LangChain) leveraging **Spark** and cloud services to process and analyze 10K+ documents daily from 200+ real-time data sources.
- Employed RAG pipelines to ground LLM (GPT-4, Claude 3) outputs with real-time external data, achieving 92% accuracy in forecasting market shifts 3–5 weeks in advance.
- Delivered insights through an interactive dashboard connected to a 15TB+ Google BigQuery warehouse, demonstrating experience in handling large-scale data for analytical systems.

### ML Framework for Skin Disease Detection (End-to-End Model Development)

- Developed a complete end-to-end R-based ML pipeline for early skin cancer detection, demonstrating proficiency across the entire model lifecycle from preprocessing to performance evaluation.
- Modularized the framework using **caret** and **mlr** to enable plug-and-play integration of new classifiers, improving diagnostic precision by 22% and showcasing skills in building extensible tools.

## Education

### University of Massachusetts Dartmouth

Dartmouth, MA

Master's in Data Science (Expected)

January 2026

### Symbiosis International University

MH, India

B.Tech in Computer Science and Engineering

June 2022

## Technical Skills

**Languages:** Python, SQL, R, MATLAB, JavaScript  
**ML Frameworks:** PyTorch, TensorFlow, Scikit-learn, Hugging Face, LangChain, LangGraph  
**Large-Scale Data:** Apache Spark, AWS (S3, Glue, Lambda, Athena, Redshift), Azure, GCP  
**Foundation Models:** OpenAI API (GPT-4), Claude, RAG, FAISS, Pinecone, Prompt Engineering  
**Databases & Querying:** PostgreSQL, MySQL, MS SQL Server, MongoDB, Snowflake, Redshift

## Publications

Bibliometric Analysis of Machine Learning and Text Mining Algorithms for Diagnosis of Leukemia. (Available on LinkedIn)