

# Rohit Sonawane

[linkedin.com/in/rohitsonawane10](https://www.linkedin.com/in/rohitsonawane10) | (213) 913-7400 | [sonawane.rohit251@gmail.com](mailto:sonawane.rohit251@gmail.com) | Los Angeles, California

## PROFESSIONAL SUMMARY

Master's student with **hands-on experience as an AI/ML engineer**, specializing in Natural Language Processing (NLP), Deep Learning, and multi-modal Artificial Intelligence, developed scalable solutions for finance, healthcare and legal domains. Delivered impactful projects, including an NLP pipeline reaching **91% accuracy**, an unsupervised model detecting **23% erroneous entries** and a multi-modal diagnostic system with **92% accuracy**, published researcher in applied Artificial Intelligence

## SKILLS

**Technical Skills:** Numpy, Pandas, Tensorflow, HuggingFace, Keras, FastAI, Scikit-learn, PyTorch, LangChain

**Job Specific Skills:** LLMs, RAG, Deep Learning, Generative AI, Information Retrieval, Communication

## WORK EXPERIENCE

### CRIF Highmark (part of Orient Technologies)

Pune, India

Machine Learning Intern (Intern Consultant)

June 2024 – December 2024

- Developed and built a refined **unsupervised clustering model** to group customer addresses, allowing better data quality and customer management, and allowing banks to identify **~23% incorrect/duplicate entries**.
- Modelled and deployed an **NLP pipeline for transaction classification** by analyzing transaction descriptions, resulting in **91% categorization accuracy** and reducing manual efforts completely.

### Translab Technologies

Mumbai, India

Data Engineering and Analytics Intern

December 2022 – May 2023

- **Formulated a data pipeline/framework** allowing **100,000+ records** of raw data from various databases to be consolidated into a centralized system for effective processing and analysis.
- **Analyzed and predicted growth trends** with the Analytics team, steering client teams to enhance decision-making, ultimately benefiting **1 million+** general consumers.

## PROJECTS

### DermaVLM: Multi-modal Skin Disease Diagnosis – as a student in Mumbai

January 2025 – May 2025

- Prepared a **custom database for 20 unique dermatological conditions** by combining dermatology datasets (SKINCAP and SCIN) to achieve a **92% classification accuracy** for diverse skin conditions.
- **Optimized model efficiency and scalability (ResNet101 CNN and Llama 3.2 11B VLM)**, lowering inference time by **48%**, improving semantic understanding with medical images by **44%**.

### NyaayAI: RAG-powered Chatbot for Legal Assistance – as a student in Mumbai

January 2024 – May 2024

- Collaborated with 2 other peers to **design and deploy a RAG-based chatbot**, combining **nomic embeddings, ChromaDB** and a **fine-tuned LLM** to interpret Indian Law documents with domain-specific accuracy.
- **Improved retrieval accuracy and response quality** by optimization, attaining a **training loss of 0.0762 with AdamW** while ensuring computational efficiency.

## EDUCATION

### University of Southern California

May 2027

Master's in Computer Science (Artificial Intelligence)

### SVKM'S NMIMS Mukesh Patel of Technology Management and Engineering, Mumbai

May 2025

Bachelor of Technology (with Diploma) in Computer Engineering with Distinction (**GPA: 3.86 / 4.0**)

## PUBLICATIONS

Saket Sultania, Rohit Sonawane, Prashasti Kanikar, "**Machine Learning based Wildfire Area Estimation Leveraging Weather Forecast Data**," International Journal of Information Technology and Computer Science(IJTCS), Vol.17, No.1, pp.1-15, 2025. DOI:10.5815/ijitcs.2025.01.01

- Designed and implemented an **advanced wildfire area estimation tool**, utilizing AutoGluon to train machine learning models on weather data, reaching a **final RMSE of 1.84 km<sup>2</sup>**.