

# Monish Gosar

monish.emailbox@gmail.com | +91 7045636928 | linkedin.com/in/monish-gosar | github.com/MonishGosar

## Education

NMIMS University - Mumbai, B.Tech in Data Science

Sept 2021 – May 2025

- **CGPA:** 3.46 / 4
- **Programming Languages:** Python, SQL.
- **Technical Tools:** Git, PowerBI, Tableau, AWS, Azure Cloud.
- **Frameworks:** Pandas, Numpy, Tensorflow, Keras, Scikit-Learn, OpenCV.
- **Databases:** MySQL, PostgreSQL, SQLite.

## Experience

Data Science Intern, *Indilabs.ai*

Nov 2024 – Ongoing

- Developed a **risk intelligence platform** that improved **collection efficiency** for major banks, engineered monitoring dashboards **reducing operational costs** and implemented **automated diagnostics** with behavioral scoring that increased **risk prediction accuracy by 14%**.
- Created **vintage analysis models** and **portfolio distribution** frameworks for enhanced recovery outcomes in high-risk segments, effectively managing a **\$1 billion lending portfolio** while providing comprehensive risk assessment capabilities.
- Built an **AI agent** for **debt collection analysis** using **Azure AI Studio**, **GPT-4o Mini**, custom prompt pipelines, and **Cosmos DB** to implement the three-component system (SQL, Analysis, and Visualization agents).
- Enhanced the system with **domain-specific knowledge** stored in Cosmos DB, creating a **two-level response framework** that improved handling of complex queries while maintaining clear boundaries for the AI's expertise.

Python Developer Intern, *RE Journal*

May 2024 – Aug 2024

- Developed **K-Means Clustering** model using Python/scikit-learn to analyze **1,000 user profiles**, achieving 85% clustering accuracy in identifying **3 distinct customer segments** based on **6 behavioral features**.
- Scraped **150,000+ real estate data points** from RERA using **Selenium** and **Beautiful Soup 4**; designed a scalable **SQL database** schema to efficiently manage the data across various cities.
- Built a **real estate analytics dashboard**, first in **Power BI** and then in **Streamlit**, to visualize data points, offering insights into **buyer behavior** and **real estate market trends**.

Data Science Intern, *Quantum Software*

May 2023 – July 2023


- Enhanced data efficiency using Python for **2G/4G KPIs analysis**; led model development and forecast improvement with **LSTM, ARIMA, SARIMA**.
- Utilized **AI4Bharat's Indic language** model for **sentiment analysis** of multilingual customer data. This solution achieved **99.3% accuracy**.

## Projects

Contractify: RAG-Powered QA Bot for Legal Contract Assistance. 

Oct 2024

- Built a **RAG-backed legal contract summarizer chatbot** using **FAISS** and **Nomic embeddings**, integrating **Llama3** and Google's Gemini Model for precise summaries.
- Developed a **LangChain-based Q&A** system with **custom prompt templates** and **Streamlit UI**, enabling interactive legal document querying.

Audio Classification - Industrial Steel. 

Sept 2024

- Implemented a **1D and 2D CNN** for steel pipe sound classification, achieving a **98% F1 score** using FFT-generated spectrograms to classify steel pipe sounds.
- Applied **data augmentation** (time shifting, pitch scaling) to improve model robustness.
- Engineered automated ML pipeline on **Azure** using **Data Factory** and **Blob Storage**, streamlining the training

workflow for audio classification models.

### Money Mule Detection.

Feb 2024

- Developed stacked ensemble ML model (**Random Forest, XGBoost, CatBoost**) for fraud detection on transaction dataset, focusing on handling class imbalance using SMOTE.
- Analyzed **100,000+** transactions, achieving **0.987 F1 Score**, **99% Accuracy**, and a **Brier score** of **0.0013**.
- Applied **SHAP** and **LIME** techniques to generate feature importance visualizations and explain individual model predictions.

### TECHNICAL & EXTRA CURRICULAR - ACHIEVEMENTS/CERTIFICATIONS/LEADERSHIP

---

- Won **2nd Place** in **LexisNexis Risk Solution Hackathon**.
- Ranked **125th** in **Amazon ML Challenge 2024**.
- **Finalist** at **Convolve Epoch 2 IDFC Hackathon**.
- **Sub-Head** of **R&D** at **Analytika, The Data Science Club**, leading a team of **5** members and organizing **3** technical events on **ML/AI fundamentals**.