# Leander Antony A

Chennai, India • +91 8610317213 • antony.leander@gmail.com
GitHub: https://github.com/LEANDERANTONY

## Professional Summary

AI/ML practitioner with a strong foundation in engineering and a transition into advanced machine learning, generative AI, and neural networks. Experienced in solving real-world problems using end-to-end ML pipelines, vector databases, and multimodal AI systems. Adept in deploying models, handling imbalanced datasets, and building robust GenAI systems.

## Technical Skills

Languages & Tools: Python, SQL, C++, R (basics), Git
Libraries/Frameworks: Scikit-learn, XGBoost, TensorFlow, PyTorch, LangChain, HuggingFace Transformers
Data: Pandas, NumPy, Matplotlib
AI/ML: Classification, Clustering, SMOTE, Hyperparameter Tuning, Cross-Validation
GenAI/LLMs: RAG pipelines, ChromaDB, Cross Encoders, Prompt Engineering
Deployment: Flask, Streamlit, GitHub Pages

## Projects

***Ongoing***

- **Multi-Modal Deep Learning for Early Pancreatic Cancer Detection**
    - Designed multi-modal AI pipeline integrating CECT imaging, urinary biomarkers, and serum miRNAs.
    - Preprocessed 1,000+ CT scans (TCIA) and 500+ biomarker samples (Kaggle).
    - Developed dual-path architecture using ResNet50 for imaging and MLP for structured data.
    - Currently implementing feature and decision level fusion strategies with stratified 5-fold CV.

***Completed***

- **Credit Card Fraud Detection**

Built a fraud classification system on a highly imbalanced dataset (only ~500 fraud cases in 282,000+ entries). Used SMOTE/ADASYN to handle class imbalance. Tuned XGBoost with cross-validation and achieved ROC-AUC ≈ 0.97. Custom threshold optimization improved TPR significantly with minimal FPR.

- **Generative AI QA RAG System for Insurance Policy documents**

Designed a retrieval-augmented generation system using LangChain and ChromaDB. Engineered chunking and cross-encoder reranking for top-k document retrieval. Used OpenAI LLMs with few-shot prompts to extract insights from Insurance policy documents.

- **Gesture Recognition Using Deep Learning**

Developed a model to recognize hand gestures from video sequences. Employed CNNs and video generators to handle temporal image data efficiently. Achieved >85% accuracy.

- **Automatic Ticket Classifier (NLP)**

Applied NMF on 5000+ unlabeled complaints; final classifier achieved >80% test accuracy. Mapped unstructured support tickets into product-based clusters and trained classifiers for prediction.

- **Telecom Churn Prediction**

Preprocessed customer-level telecom data with 70,000+ rows and 172 features. Applied feature engineering, imputation strategies, and classification modeling. Model achieved 97.9% sensitivity and 96.3% precision on test set.

- **Bike Demand Prediction**

Performed EDA and feature analysis to model bike sharing demand using linear regression. Dataset spanned 700+ days; achieved $R^2$ ≈ 0.80 on test set. Visualized weather/seasonality impact and tuned predictors for optimal RMSE.

- **Lending Club Loan Default Risk Analysis**

Built a predictive model to assess loan default risk using financial and demographic features. Modeled using 10,000+ loan records; achieved logistic regression ROC-AUC ≈ 0.89. Cleaned and imputed real-world lending data.

## Education

Master of Science in AI/ML from Liverpool John Moores University • Oct 2024 – Oct 2025

Executive PG Program in Machine Learning & Artificial Intelligence with specialization in Generative AI IIIT Bangalore • Oct 2023 – Oct 2024

B.Tech in Mechanical Engineering
Manipal Institute of Technology • Aug 2015 – May 2019

## Publications

First-author publication in Solar Energy (Elsevier) "Influence of Stepped Cylindrical Turbulence Generators on the Thermal Enhancement Factor of a Flat Plate Solar Air Heater"

## Certifications

- Data Science: R Basics — HarvardX (edX), Dec 2019
- Programming for Everybody (Python) — University of Michigan (Coursera), May 2018
- Solar Energy: Photovoltaic Conversion — Delft University of Technology (edX), May 2020