

# ADIN RAJA

## Data Scientist / Machine Learning Engineer

[Portfolio](#) | [GitHub](#) | [LinkedIn](#) | [Medium](#) | [Email](#) | [9653669698](#)

---

### TECHNICAL SKILLS:

- **Programming Languages** : Python, SQL, Java Script, Html, CSS
  - **Machine & Deep Learning Frameworks** : Scikit-learn, Pytorch, TensorFlow, XG-Boost, MLflow
  - **Data Analysis & Visualization** : Pandas, NumPy, Seaborn, Matplotlib, Power BI
  - **Web Scraping & API's** : Requests, Flask, FastApi, BeautifulSoup
  - **Tools** : Git & GitHub, Jupyter, Vscode, My SQL, Docker, Streamlit, AWS , Excel
- 

### EDUCATION:

**Bachelor of Engineering In Information Technology** | Nov 2022 – May 2025  
University of Mumbai

**Diploma in Information Technology** | Sep 2019 – June 2022

---

### WORK EXPERIENCE:

**Data Science and Machine Learning Intern** | [Certificate](#)

Infotact Soltuions

- Implemented async I/O in data scraping pipeline reducing data collection time from ~23 hrs → ~40 mins.
- Deployed [FlickerFlix](#) – an end to end movie recommendation engine built from scratch on freshly scraped 170,000+ movie metadata, applying lemmatization and Tf-Idf vectorization.

**Machine Learning Intern** | [Certificate](#)

Prodigy Infotech

- Built and deployed multiple ML models across regression and classification tasks, gaining hands-on experience in data pre-processing, model tuning, and performance evaluation.

**Data Science and Analytics Intern** | [Certificate](#)

Tamizhan Skills

- Conducted end-to-end data analysis — from preprocessing to visualization — using Python (Pandas, Seaborn , Matplotlib) and Power BI for data driven decision making.
- 

### PROJECTS:

**RiskRadar – Credit Risk Modelling** | [GitHub](#)

**Classification | Streamlit | Render**

- Designed a production ready CIBIL-like credit risk scoring system for financial institutions to classify loan applicants into different risk bands to support smarter lending decisions while minimizing default risk.
- Prevented data leakage and Engineered key features like loan to income ratio, delinquency ratio and average days past due.
- Handled class imbalance using SMOTE, fine tuned a logistic regression model using optuna, achieving 0.94 recall on the default class (1) and auc score of 0.983.

**Mumbai Property Price Estimator** | [GitHub](#)

**Data Collection (API) | Flask | Render**

- Built an end-to-end machine learning solution for real-time property price estimation in Mumbai by web scraping 50K+ live listings from the MagicBricks API.
- Generated impactful features like area per bedroom and location-based target encoding, reducing model error margin (> 10) from 8% → 6% and improving R<sup>2</sup> from 0.95 → 0.99.
- Fine-tuned an XGBoost regression model and deployed it via a flask api endpoint.

**DentAware – AI Powered Car Damage Detector** | [GitHub](#)

**CNN | ResNet50 | PyTorch**

- Deployed a deep learning (CNN) based system using a fine tuned ResNet50 model to detect and classify vehicle damage from images across 6 categories with ~80% validation accuracy.
  - Implemented data augmentation and transfer learning to handle limited labelled data (~2.3K images)
- 

### CERTIFICATIONS

- Microsoft Certified (PL-300 – Power BI Data Analyst Associate) , LeetCode SQL 50
- 

### PUBLICATIONS:

**FraudShield – Online Payment Fraud Detection System** | [GitHub](#)

**Docker**

- Published a Research paper in IJARST (International Journal of Advanced Research in Science & Technology)
- Developed a machine learning based fraud detection pipeline to help banks and payment platforms to detect fraudulent online transactions in real time.
- Trained a Random Forest Classifier achieving 99% recall and 94% precision on the fraudulent class.