

# Rohan Patnaik

Product Data Scientist



+91-8283832094



rohanpatnaik1997@gmail.com



Github

## SUMMARY

- Data scientist skilled in machine learning, predictive modeling, and software engineering,
- Proficient in Python, SQL, and statistics. Specialized in financial data classification and fraud prediction.
- Expert in communicating technical concepts to non-experts with commendable teamwork and leadership abilities.

## TECHNICAL SKILLS

**Data Science and Analytics:** Pytorch, Keras, Scikit-Learn, Machine learning (ML), TensorFlow, Deep learning, Neural networks, A/B testing, Excel, Clustering, Tableau, NLP, Computer vision, Time series analysis, GenAI

**Database related:** MySQL, PostgreSQL, Trino, Redis, MinIO

**Cloud Platforms (IaaS, PaaS, SaaS):** AWS (S3, EC2, RDS, Lambda), Google Cloud, Heroku

**Software Development:** Python, NumPy, Pandas, Flask, Jenkins, NLTK, Jupyter, TypeScript, Docker, Kubernetes, Git, PySpark, GitHub Actions, Trello, JIRA, Kubeflow

## WORK EXPERIENCE

### National Payment Corporation of India (NPCI) - Euclid Innovations

Current position

#### Senior Data Scientist

- Led the development of fraud detection for the **AePS** (Aadhar Enabled Payment System) by engineering raw data features to create predictive models to reduce fraud incidents by **1.2million** working on more than **4TB data**.
- Utilized python to implement supervised machine learning techniques for time series anomaly detection as well.
- Contributed to a pioneering project on **UPI**, developing a federated model for **financial fraud detection**, showcasing expertise in handling complex, real-world financial data for new onboarding Indian banks..

### SKIPCHECKR AI

April 2022 – Jan 2024

#### Data Scientist

- Spearheaded the creation and development of an innovative **self-checkout system** powered by machine learning, overseeing the project from its initial concept to a **minimum viable product** (MVP).
- Showed strong presence in leadership and **project management**, driving sub projects to completion.
- Applied data science using **computer vision** in retail, bettering efficiency and convenience by targeted market research and **A/B testing**.

### SIEMENS

July 2019 – April 2022

#### Data Scientist/ Software Engineer

- Developed and analyzed sound and numeric data for the client assets using computer vision; Outcome: useful insights using RUL classification, with an **87% accuracy** for fault detection in the assets.
- Designed **dashboard** views with API calls to display asset health of 97 assets; performed ML analysis.
- Led team communications to address key security issues generated from tokenization for the user app.

## PROJECTS

### RAG WITH LLAMA2

- Customizing and fine-tuning LLM using **Llama2** as base model to generate articles for multiple categories of people with varied length of outputs involving langchain leveraging as well. Deployed the app on streamlit for better reach and efficiency.

### ZERO SHOT IMAGE CLASSIFICATION

- Developed an ML application using **EfficientB0** to predict food presence in images, achieving a 98% accuracy; data sources: Tiny Imagenet and Food101 dataset. Deployed the app on GCloud for storage efficiency.

### INSURANCE COST PREDICTION

- Leveraged a neural network-based machine learning model to accurately predict insurance costs from medical history data, achieving a remarkable mean absolute error reduction to 3169, enhancing cost quantification precision.

## EDUCATION

SRM UNIVERSITY

Bachelor of Technology in Computer Science and Engineering

2015 – 2019