

Pauras Ajay Raut

Galway Ireland | 0894145552 | rautpauras000@gmail.com | [GitHub](#) | [Portfolio](#)

Summary: -

Machine Learning Engineer with 1+ years of experience in AI and Data Science across energy, healthcare, and finance, skilled in predictive modeling, workflow automation, and large-scale data insights. Pursuing a Master's in Genomic Data Science at NUI Galway with a focus on ML, visualization, and genomics. Hands-on feedback systems, customer purchase analysis, medical image classification, and LLM-based chatbots, with strong foundations in programming, preprocessing, visualization, and deep learning.

EDUCATION: -

Masters in Genomic Data Science from National University of Galway, Ireland (2.1) Sept 2024 – Sept 2025

- **Key modules:** - Machine Learning (ML), Introduction to Molecular and Cellular Biology, Data visualization.
- **Project: -**
Chatbot using LLM:-
 - Collaborating on cBioPortal, an open-source platform with over 5M+ cancer genomic profiles used by global research institutions.
 - Built an LLM-based chatbot (LangChain + OpenAI) to streamline data queries and improve accessibility, reducing manual search time by 60% and enhancing data interpretation for researchers and clinicians.

Bachelor of Engineering in Computer Science from ST John College of Engineering (1.1) Aug 2020 – May 2024

- **Key modules:** - Artificial Intelligence (AI), Data Structures, Social Media Analytics, Cybersecurity and Laws, Data Warehousing and Mining, Data Science.
- **Project: -**
Mental Health Analysis:-
 - Developed a mental health analysis system using OpenAI API + SVM, achieving ~92% accuracy in classifying depression severity from 200+ user responses.
 - Integrated standardized Mental Health America questionnaires and semantic diary analysis (1K+ entries) to detect emotional trends, improving early-intervention recommendations by 25%.

Work Experience: -

Aug 2022 – Jan 2023

ADANI POWER GROUP LIMITED, ML Engineer

- **Feedback Systems:**
 - Created and optimized feedback systems using automated dashboards to track performance trends with 99% reliability.
 - Streamlined audit workflows, automating 10+ manual tasks to improve efficiency by 35% and reduce reporting errors by 25%.
 - Conducted knowledge transfer sessions for 15+ team members, ensuring seamless integration of new releases across departments.
- **Customer and Sells:**
 - Conducted a market-basket analysis project under Adani Power's internship program to explore product placement optimization for local retail shops.
 - Processed 7,500+ transaction records and applied the Eclat algorithm (using the *apyori* library) to discover high-support product pairings.
 - Built an association-rule-mining model in Python with Pandas, NumPy, and Matplotlib for data preprocessing and visualization.
 - Generated insights that informed inventory grouping and promotional strategies, demonstrating how analytics could improve sales efficiency and customer satisfaction.

Jan 2022 – Jul 2022

Tata Steel, Data Engineer Intern

- Developed a predictive maintenance model using Python, SQL, and Pandas, analyzing 10K+ industrial sensor records to forecast equipment failures with ~85% accuracy.
- Applied time-series forecasting and anomaly detection to monitor machine health, helping reduce potential downtime by ~20%.
- Built interactive Power BI dashboards that enable data-driven maintenance planning and improved decision-making speed by ~30%.

Skills Profile: -

TECHNICAL SKILLS: -

- Programming Languages: SQL, Python, HTML, JavaScript.
- Libraries and Frameworks: Pandas, Matplotlib, Sci-kit Learn, Numpy, Tensorflow, LangChain, Hugging Face Transformers.
- IDE and Tools: PyCharm, Visual Studio Code, NetBeans, Google Colab, Jupyter Notebook, MS Word, MS PowerPoint, Power BI.
- Database Management System: MySQL, ChromaDB.

SOFT SKILLS: -

- Problem-Solving: Strong analytical thinking to develop effective and innovative solutions.
- Adaptability: Quick learner with the ability to thrive in dynamic and fast-paced environments.
- Collaboration: Proven ability to work effectively in team-oriented settings and cross-functional project.

CERTIFICATIONS: -

- Foundations: Data, Data, Everywhere (Coursera).
- Machine Learning A-Z: AI, Python (Udemy).
- Python for Data Science (Udemy).
- Data Science expert track with Python and Machine Learning (Interface).

Personal Projects: -

- Breast Cancer Detection: -
 - Implemented a logistic regression model to classify breast cancer cases, achieving an accuracy of ~90% through cross-validation for reliable predictions.
 - Pre-processed and analyzed a dataset of 500+ instances by handling missing values, scaling features, and splitting data into training and testing sets for robust evaluation. Achieved approximately 90% accuracy through cross-validation, ensuring reliable results.
 - Utilized Python libraries such as scikit-learn, pandas, and NumPy for data analysis, preprocessing, and model development.
- Bank Stock analysis: -
 - Extracted and analyzed stock market data for Bank of America using Python libraries such as yfinance, processing over 5 years of historical data to identify trends.
 - Cleaned and pre-processed financial datasets, ensuring 100% accuracy for technical analysis and data visualizations. Performed technical analysis to evaluate stock trends, market behavior, and key indicators, using Python-based tools for actionable insights.
 - Leveraged Python libraries like pandas and matplotlib to create comprehensive visualizations and insights, supporting data-driven decision-making.
- Travelpia: -
 - Designed and implemented an AI-powered travel assistant using Retrieval-Augmented Generation (RAG) to deliver personalized itineraries across 32 Irish counties, improving search accuracy by ~85% through contextual retrieval.
 - Scraped and processed 50+ Wikivoyage travel pages, generating 10,000+ text embeddings stored in a FAISS vector database for efficient semantic similarity queries.
 - Integrated LangChain and OpenAI API for natural conversation flow, itinerary generation, and location-specific recommendations with sub-second response time.
 - Deployed an interactive Streamlit dashboard on Hugging Face Spaces, enabling users to query and explore travel destinations dynamically.