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 focus on ML, visualization, and genomics. Hands-on in feedback systems, customer purchase analysis, medical image classification, and LLM-based chatbots, with strong foundations in programming, preprocessing, visualization, and deep learning.

EDUCATION: -

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

- **Key modules: -** Machine Learning (ML), Introduction to Molecular and Cellular Biology, Data visualization,
- **Project: -**
  Chatbot using LLM: -
  - Currently contributing to cBioPortal, an open-source platform that provides visualization and analysis of cancer genomics data.
  - Working on developing a chatbot using Large Language Models (LLMs) to enhance user interaction, streamline data queries, and improve accessibility to genomic insights. This project aims to simplify complex data interpretation for researchers and clinicians.

Aug 2020 – May 2024 **Bachelor of Engineering in Computer Science from ST John College of Engineering**

- **Key Modules: - Artificial** Intelligence (AI), Data Structures, Social Media Analytics, Cybersecurity and Laws, Data Warehousing and Mining, Data Science
- **Project: -**
  Mental Health Analysis: -
  - The project aimed to raise awareness about mental health and promote proactive, self-driven management of emotional and psychological health.
  - Developed a mental health analysis system using OpenAI API and Support Vector Machine (SVM) to classify user responses into different levels of depression severity.
  - Integrated questionnaires from Mental Health America to assess users' mental health status and provide tailored solutions for mild to moderate cases.
  - Implemented semantic analysis on personal diary entries to monitor emotional trends and recommend professional consultation for severe depression cases.

Work Experience: -

Aug 2022 – Jan 2023 **ADANI POWER GROUP LIMITED,** ML Engineer

- **Feedback Systems:**
  Developed and optimized feedback systems by creating dashboards and reports to track trends, identify discrepancies, and ensure system reliability through automation. Streamlined audit workflows and automated key tasks, improving efficiency, reducing errors, and enhancing intra-departmental communication.
  Implemented software solutions aligned with architecture requirements, and conducted knowledge transfer sessions for new releases, ensuring smooth adoption and integration

- **Customer and Sells:**
  Internship at Adani Power focused on optimizing product placement for local society shops using the Eclat algorithm. Processed 7,501 transactional data entries to identify frequently purchased product combinations. Developed an association rule-mining model with the Eclat algorithm using the Apoyri library. Discovered top product pairings with high support values for better inventory and marketing strategies. Utilized Python libraries such as Pandas, NumPy, and Matplotlib for data preprocessing, analysis, and visualization. Provided actionable insights that contributed to increased sales and customer satisfaction for local society shops.

Jan 2022 – Jul 2022 **Data Science Intern – Tata Steel**

- Developed a predictive maintenance model to forecast equipment failures, reducing downtime and improving operational efficiency.
- Processed and analyzed large-scale industrial datasets using Python, SQL, and Pandas.
- Applied time-series forecasting and anomaly detection techniques to monitor equipment health.
- Built interactive dashboards in Power BI to visualize real-time performance metrics for decision-makers.
- Collaborated with cross-functional teams to ensure deployment feasibility and scalability of ML solutions.

Skills Profile: -

TECHNICAL SKILLS:

- Programming Languages: SQL, Python.
- Libraries and Framework: Pandas, Matplotlib, Sci-kit Learn, Numpy, Tensorflow, Django.
- IDE and Tools: PyCharm, Visual Studio Code, NetBeans, Google Colab, Jupyter Notebook, MS Word, MS PowerPoint, Power BI.
- Database Management System (DBMS): MySQL.

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: -

- Brain Tumour Detection using Python and scikit-learn: -
  - Developed a machine learning model for brain tumor detection using Support Vector Machine (SVM) and Logistic Regression, achieving an accuracy of 95% with the SVM model on a benchmark medical image dataset.
  - Pre-processed and scaled a dataset of 5,000+ medical images, conducted exploratory data analysis (EDA), and created visualizations to identify key features and patterns.
  - Optimized model performance through hyperparameter tuning and cross-validation, ensuring high precision and reliability for tumor classification in medical images.
- Breast Cancer Detection: -
  - Developed 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