P i n a k i   P a n i

Cherrywood, Dublin, Ireland | +353 899652806 | mkumarpani@gmail.com | LinkedIn | GitHub | Portfolio

PROFILE SUMMARY

- Data Analyst with 3.5 years of professional experience and a Master's degree in Data Analytics from NCI, Dublin. Skilled in data wrangling, visualization, ETL automation, and statistical/machine learning modeling. Experienced in turning complex datasets into actionable insights to support decision-making in healthcare and transport analytics.
- Recently built and deployed a "Dublin Transit Stress Analysis" application — a real-time data analytics dashboard integrating GTFS-RT feeds, weather APIs, and vehicle tracking to measure transit stress across Dublin. Strong technical foundation in Python, SQL, Power BI, and end-to-end project delivery.

CORE SKILLS

- **Data Analytics:** Data Cleaning, Predictive Analytics, Anomaly Detection, Time Series, Forecasting, A/B Testing
- **Visualization & BI:** Power BI, Streamlit, Matplotlib, Seaborn, Plotly
- **Programming:** Python, SQL, Shell Scripting
- **ETL & Databases:** Data pipelines, Relational Databases
- **Version Control & Tools:** Git, Linux, Windows

PROJECTS

➢ **Dublin Transit Stress Analysis (**Dashboard App**, 2025)**
- Designed and deployed a real-time transit stress dashboard for Dublin integrating GTFS-RT, weather (OpenWeatherMap), and traffic conditions.
- Engineered nightly ETL pipelines to collect, clean, and store millions of rows of transit and weather data.
- Developed metrics for vanished vehicles, delays, and stress intensity across routes.
- Built an interactive dashboard in Streamlit + Deck.gl with animations, filters, and route aggregation.

Tech Stack - **Python, Pandas, Pydeck, Streamlit, APIs, Data Engineering, Geospatial Analysis**

➢ **A Systematic Evaluation of Vision Transformers for Galaxy Classification** (Master's Thesis, **2024**)
- Conducted a comparative analysis of **Vision Transformer** models **(ViT Base, Swin Transformer,** and **DeiT Transformer)** for classifying galaxy morphologies using the **Galaxy10 DECals** dataset.
- Implemented **deep learning pipelines** for data preprocessing, augmentation, model training, and evaluation.
- Optimized model performance using transfer learning (pretrained on ImageNet) and hyperparameter tuning.
- Achieved **86.7% classification accuracy** with ViT Base model, outperforming ResNet50 while analyzing computational trade-offs.

Tech Stack - **Python, PyTorch, TensorFlow, Scikit-Learn, OpenCV**

➢ **Covid-19 Insights and Analytics (**Covid 19 Project**, 2023)**
- Developed a comprehensive **data analysis pipeline** to explore the global impact of Covid-19 using real-world datasets.
- Conducted **time-series analysis** on infection rates, recovery trends, and mortality rates across multiple regions.
- Built **interactive visualizations** to track pandemic progression using **Python.** Performed **correlation analysis** between Covid-19 cases and socio-economic factors to identify patterns.

- Applied **statistical modeling and forecasting techniques (ARIMA, Exponential Smoothing)** to predict future case trends.
Tech Stack: Python, Pandas, Matplotlib, Plotly, Jupyter Notebook, SQL, Machine Learning (ARIMA). **GitHub**

WORK EXPERIENCE

01 / 2022 – 06 / 2023 | Infosys Limited   **Technology Analyst**
- Designed interactive Power BI dashboards for Adverse Events Monitoring and others, enhancing decision-making for 8+ on premise servers by presenting Key-performance indicators and key trends.
- Led data quality analysis, identifying majority data issues proactively, enabling faster resolution and improving data integrity for CDARS (Clinical Data Analysis and Reporting System)
- Analyzed server log dataset(Server uptime, downtime, latency etc.) and incident report datasets to perform root cause analysis for several issues
- Performed multiple Extraction, Load and Transformation of data from multiple sources using Python.
- Automated 25% of routine data processing tasks using Python, cutting manual workload by 20+ hours per month and increasing efficiency
- Involved in development and deployment of ML models to predict ETL job delays and detect anomalies in processing patterns, reducing operational downtime and improving pipeline stability by 35%
- Documented workflows, data pipelines, and automation processes for knowledge sharing and business continuity

12 / 2019 – 12 / 2021 | Infosys Limited   **Systems Engineer**
- Responded to data pipeline and server alerts based on error logs and job monitoring, maintaining uptime and meeting internal SLAs
- Participated in triage calls and worked cross-functionally to resolve P1 incidents, ensuring minimum disruption to clinical systems
- Conducted data anomaly detection, improving data validation processes and increasing data accuracy for key business reports utilizing advanced Excel and Python scripting
- Supported multiple applications with data-fetch or retrieval issues, feature enhancement and bug fix issues along with user permission related tasks
- Monitored and maintained the health of upstream and downstream applications to ensure end-to-end system stability and uninterrupted data flow across the CDARS ecosystem
- Collaborated with cross-functional teams (business analysts, data engineers, and compliance teams) to streamline reporting processes
- Created and maintained technical documentation, including design specifications, user manuals, and deployment guides

---

EDUCATION

MSc in Data Analytics | National College of Ireland (Sep 2023 – Nov 2024) | **Grade: 2:1**
Key Courses: Data Cleaning, Data Visualization, Predictive Analytics, Machine Learning, Deep Learning, Business Intelligence, Statistics, Data Mining, Research Project

B.E. in Information Science | Acharya Institute of Technology (2015 – 2019) | **GPA: 7.0/10**