Pakhi Tyagi

London, UK | Email: pakhi1411j@gmail.com | Phone: +44 7352 847130 | LinkedIn:
https://www.linkedin.com/in/pakhi-tyagi-0ba27b230/ | GitHub: https://github.com/pakhi-1411-2002

SUMMARY

Results-driven Data Scientist with 10 months of hands-on experience delivering production-grade machine learning models, full-stack data products, and analytics dashboards that drove 20–80% efficiency gains for enterprises and government clients. Proven expertise in computer vision, NLP, explainable AI (SHAP/LIME), A/B testing, and scalable Python systems.

WORK EXPERIENCE

Mphasis, India
Software Engineering Trainee- Data Analytics and HR Technology, Dec 2024 – May 2025
- Designed and delivered a production Django + SQL HR analytics dashboard that automated 80% of manual time-tracking, employee-release workflows and analyse employee productivity trends
- Built and integrated XGBoost attrition-risk and attendance-anomaly models (92% precision), powering real-time alerts that cut unexpected turnover in teams
- Developed automated data pipelines with feature engineering that reduced weekly HR reporting time from 5 hours to 1.5 hours (70% faster) and real-time alerts and reminders
- Tech: Python, Django, SQL, XGBoost, Scikit-Learn, SHAP, Pandas, Celery, Docker, Git

IG Drones, India
Python Development Intern- Data Science Team, Jul 2024– Oct 2024
- Built and deployed a real-time helmet-detection computer vision model using drone footage for Odisha Traffic Police (TensorFlow + OpenCV + CNNs), containerized the full training & inference workflow with Docker
- Engineered an automated annotation pipeline that accelerated data labelling by 33% and doubled model training cycles per sprint
- Reduced system inefficiencies by 38% through structured optimisation and documented operational improvements.
- Tech: Python, TensorFlow, OpenCV, Scikit-Learn, Pandas, NumPy, Docker, Git

PROJECTS

Credit-Risk Explainer Dashboard
- Built a credit coaching engine that sorts over 1 million daily transactions into needs and wants, helping give simple credit tips and raising user engagement by 35%.
- Created a synthetic dataset of 500,000 anonymized transactions, clearly labeled by category (groceries, shopping, EMI payments), improving model training accuracy by 20%.
- Implemented a rule-based + ML classifier with 92% precision to spot essential vs. non-essential spending and detect risky habits like high credit use or missed payments.

Public Art Style Intelligence
- Built an end-to-end system that performs style classification, CLIP embeddings, <50 ms vector search, clustering, and Grad-CAM explanations.
- Trained an Azure classifier on 80k+ art images and generated 512-dim CLIP fingerprints for similarity & micro-genre discovery.
- Deployed a Streamlit app returning style, embeddings, similar artworks, cluster ID, and explanation heatmap in ~1.5 s.
- Tech: Azure Custom Vision, PyTorch, CLIP, FAISS/Qdrant, Streamlit, scikit-learn, Grad-CAM

Automated Data Analysis & ML Pipeline
- Built a reusable pipeline that ingests data, cleans, performs EDA, trains models, generates automated reports.
- Integrated clustering, anomaly detection, regression, and automated experiment comparison with MLflow tracking.
- Deployed live inference via FastAPI and automated scheduled runs using Airflow/Prefect with CI/CD & drift monitoring.
- Tech: Python, MLflow, FastAPI, Airflow/Prefect, scikit-learn, Pandas, Docker, GitHub Actions

EDUCATION

University of Southampton	**Southampton, UK**
Master of Science, Data Science	*Graduation Date: Sep 2026*

St. Joseph's University   **Bangalore, India**
Bachelor of Science, Computer Science & Mathematics   *Graduation Date: Jun 2024*
- GPA: 3.8/4.0

TECHNOLOGY

- Languages: Python (expert), SQL, R, C++, C
- ML/DL: Scikit-Learn, TensorFlow, XGBoost, SHAP, OpenCV, CNN
- Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Tableau
- Data: Pandas, NumPy, Matplotlib, Seaborn, Power BI
- Tools: Git, Docker, Streamlit, Jupyter, Django, Azure (AZ-900)

Certifications

- Microsoft Certified: Azure Fundamentals (AZ-900) – 2025
- PCAP Certified Associate Python Programmer- 2025