ROHIT KUMAR DUBEY

(+44) 07393109551 | E-mail | LinkedIn | GitHub | Portfolio | Glasgow, United Kingdom

PROFESSIONAL SUMMARY

Recent Computing Science graduate with hands-on experience in **cloud platforms** (AWS, GCP, Azure), **data pipelines**, and **full stack development**. Skilled in **Python**, **SQL**, **Power BI**, and **Airflow**. Built scalable web applications using **Node.js** and **MongoDB**. Delivered robust **machine learning models** and contributed to **open-source** projects with **agile workflows**.

EDUCATION

Computing Science, M.Sc. | University of Glasgow **Sep. 2023 – Jun. 2025**

Coursework: Programming and Systems Development, Introduction to Data Science and Systems, Machine Learning and Artificial Intelligence for Data Scientists, Research and Professional Skills, Internet Technology, Human Centered Security, Deep Learning, Big Data: Systems, Programming and Management, Information Visualisation, Secured Software Engineering.

SKILLS

Programming & Tools: Python, SQL, Bash, Apache Airflow, Docker, Kubernetes, dbt, GitHub Actions, Git, Linux CLI, VS Code, Django, Terraform, Azure (AZ-900 certified), AWS (EC2, S3, Lambda, Glue), GCP.
Data Analysis & Visualization: Jupyter Notebook, Streamlit, Power BI, Matplotlib, Seaborn, Plotly, Pandas, NumPy.

EXPERIENCE

Software Developer Intern | StackRoute | Remote **Mar. 2022 – Aug. 2022**

- Completed an intensive **full-stack program** with hands-on training in **Java**, **SQL**, **Angular**, and **Google Cloud Platform (GCP)**. Built and deployed a web app as part of a **4-member team**.
- Delivered **end-to-end application** with **API integration**, **database design**, and **agile practices**, earning strong feedback for **technical proficiency** and **team collaboration**.

Full Stack Developer Apprentice | Arth - The School of Technologies | Remote **Sep. 2020 – Oct. 2021**

- Completed a part-time **Full Stack Apprenticeship**, building apps with **JavaScript**, **Node.js**, **MongoDB**, and deploying on **GCP**; trained in **AWS**, **Azure**, and cloud fundamentals.
- Contributed to **open-source projects** using **Git**, **GitHub**, and **agile workflows**; upskilled in **Bash**, **Ansible**, and **OpenShift** through expert-led sessions.

Cloud Intern | Sapio Analytics | Remote **May 2020 – Aug. 2020**

- Managed **IAM users and roles** on **AWS**, implementing **secure access policies** that **improved team efficiency and cloud security**.
- Provisioned and automated **EC2 instances** on Ubuntu 16.08, configuring MySQL, Nginx, and Gunicorn, leading to a **30% improvement in deployment efficiency**.

PROJECTS

Customer Churn Prediction | Python, Matplotlib, Seaborn, Scikit-learn **May 2025**

- Conducted thorough **Exploratory Data Analysis (EDA)** and applied **data preprocessing** techniques including **One-Hot Encoding** and missing value imputation to prepare telecom customer data for modeling.
- Developed and optimized **machine learning models** (Logistic Regression, Random Forest, and XGBoost), achieving **82% accuracy** with the tuned XGBoost model for predicting customer churn, enhancing **retention strategy insights**.

Fake News Detection: A Scalable Data Pipeline Approach | Python, Pandas, PostgreSQL, Airflow **May 2025**

- Built a robust ETL pipeline using **Python**, **Pandas**, and **SQLAlchemy** to ingest and transform over 10,000 news articles, loading them efficiently into a **PostgreSQL** database for downstream analytics.
- Automated dynamic table creation and timestamped batch insertion, enabling scalable integration with future **data pipelines**, dashboards, and machine learning systems.

Integration of MySQL and Wordpress through Terraform | Terraform, Amazon Web Services (AWS) **Nov. 2020**

- **Designed** and implemented secure **Virtual Private Clouds (VPCs)**, subnets, **NAT gateways**, and route tables on **AWS**, ensuring robust **database security** and reducing security risks by **25%**.
- Developed **Terraform** code to seamlessly integrate **MySQL** and **WordPress**, automating deployment and improving **deployment speed** by **40%** and **reliability** by **30%**.