Rounak Kumar

rounakgupta914@gmail.com | +91 7979886359
github.com/raw9k | linkedin.com/in/guptarounak | kaggle.com/raw9kk

Skills

Languages: Python, R, SQL, C++
Libraries & Frameworks: Scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, LightGBM, CatBoost, Pandas, NumPy, Flask, Streamlit, FastAPI (REST API development)
Tools: Docker, Git, GitHub Actions, MLflow, Power BI, Azure, AWS, DVC (versioning), Comet ML
Workflow Skills: Data Pipelines, ETL Workflows, Hyperparameter Tuning, Model Evaluation (Accuracy, F1, AUC), CI/CD Integration, Containerization, Monitoring & Logging

Education

Secondary Education | CBSE ***91.4% (2020)***
Higher Secondary Education | CBSE ***83% (2022)***

Birla Institute of Technology, Mesra Dec 2022 - Jun 2027
Integrated MSc. in Quantitative Economics and Data Science ***CGPA: 7.46/10***
Relevant Coursework: Probability & Statistics (I & II), Linear Statistical Models and Regression Analysis, Optimization Techniques, Stochastic Processes, Sampling Techniques and Design of Experiments, Statistical Machine Learning I.

Projects Experience

Hybrid Anime Recommendation **Link**

- Constructed a scalable anime recommendation system utilizing automated Azure Blob Storage ingestion and preprocessing pipelines to efficiently handle over 70 million user-anime ratings for robust and seamless large-scale data processing
- Developed an embedding-based hybrid recommender model as resulting in improved recommendation accuracy leveraging Keras neural network architecture with content-based and collaborative filtering, integrated with Comet ML for tracking.
- Deployed an end-to-end ML application as demonstrating systematized updates and continuous delivery via Azure Web App containers and GitHub Actions CI/CD pipelines, facilitating seamless model retraining and user-facing API availability.
- **Tech Stack:** Python, Keras, Comet ML, GitHub Actions, Docker, Azure Web App and Containers

Network Security System **Link**

- Constructed an end-to-end phishing detection pipeline as validated through validated schema checks and reproducible artifacts by implementing MongoDB ingestion, transformation, feature processing, model training, and a FastAPI /predict service.
- Instituted secure automation as demonstrated through on-push cloud deployments and signed image digests by containerizing the service, configuring GitHub Actions for Azure Container Registry, and injecting secrets via environment variables.
- Provisioned production hosting as validated via externally accessible API and consistent workflows by deploying to Azure Web App, adding a health endpoint, and implementing a responsive Jinja interface with cohesive theming.
- **Tech Stack:** Python, MongoDB Atlas, MLflow, GitHub Actions, Docker, Azure Web App and Containers

Student Performance Prediction Model **Link**

- Accomplished accurate math score prediction as reflected by R^2=0.87 by engineering a full-stack ML pipeline with robust validation, feature engineering, and cross-validation, ensuring reproducibility and delivering high predictive accuracy.
- Achieved 40% reduction in manual pre-processing evidenced through workflow metrics by automating ingestion, transformation, and model training using Python, Scikit-Learn, and CatBoost with reproducible pipelines and robust error handling.
- Delivered scalable public model hosting as measured by uptime and reach by deploying with Docker, GitHub Actions, and Azure Web App, enabling continuous delivery, mechanized builds, and secure containerized environments.
- **Tech Stack:** Python, Scikit-Learn, CatBoost, Pandas, Docker, GitHub Actions, Azure Containers and Web App

Leadership & Extracurriculars

Joint President, Society for Data Science – BIT Mesra April 2025 – Present

- Spearheaded the university's Data Science Society, curating and leading hands-on workshops, hackathons, and expert-led sessions on Generative AI, Machine Learning, and Data Science to foster a culture of innovation and technical excellence
- Mentored students in data-driven projects, guiding them through practical implementations and real-world applications, and actively fostered collaborations with industry experts to enhance learning, exposure, and professional networking.

Certificates

- **Career Essentials in Data Analysis** - Microsoft & LinkedIn
- **Machine Learning with Python** - freeCodeCamp
- **Learning Microsoft Power BI** - Infosys Springboard