Rounak Thakur

Mobile: +44 77681 44342 | Email: rounak26thakur@gmail.com | Location: London, United Kingdom

Summary

Results driven data professional pursuing an MSc in Big Data Science at Queen Mary University of London. Possess a robust foundation in data engineering, analytics, and deep learning. Experienced in building scalable data pipelines, developing enterprise data models, and deploying advanced machine learning models. Proficient in frameworks like PyTorch and TensorFlow, with hands on experience in transformer architectures (e.g., BERT, GPT) via Hugging Face for NLP tasks. Passionate about solving real-world challenges in AI, I thrive in collaborative environments and am committed to delivering innovative, data-driven solutions that drive measurable impact.

Professional Experience

Associate Developer 2 - Data Engineering
Tyson Foods IT Solutions
April 2022 – August 2024

- Designed and optimized enterprise data models, integrating multiple data sources to enhance system architecture and ensure seamless data access across the organization.

- Developed Power BI dashboards for financial and operational data, enabling real-time monitoring and strategic decision-making.

- Automated data pipelines using GCP BigQuery, improving data processing efficiency by 15%.

- Built datasets such as Production Cost, Accounts Payable, and Capacity Utilization, contributing to $25M in cost savings over time.

- Implemented data quality checks and governance frameworks, enhancing reporting accuracy and compliance.

AI Developer Intern
NxtGen Data Infinite Center
October 2021 – January 2022

- Developed a license plate recognition system using OCR, OpenCV, and DeepStream, enhancing vehicle identification accuracy.

- Built a real-time attendance tracking solution using live camera feed and machine learning-based people counting.

- Explored fundamental machine learning concepts, applying SciPy, NumPy, Pandas, and Matplotlib for data analysis.

Education

MSc Big Data Science – Queen Mary University of London, United Kingdom | 2024–2025

B.E Electronics & Communications – MVJ College of Engineering, Bangalore, India | 2017–2021

Projects

NLP Vector Space Semantics
GitHub: NLP-Vector-Space-Semantics

- Developed a character recognition model using Hugging Face Transformers to classify dialogue from the EastEnders TV script.

- Preprocessed text, created document embeddings, and evaluated performance using cosine similarity.

- Achieved high F1-score by optimizing tokenization and vector space separation.

CIFAR-10 Deep Learning Model
GitHub: CifarDeepLearning

- Implemented a convolutional neural network (CNN) to classify images from the CIFAR-10 dataset.

- Utilized TensorFlow and Keras for model development and training.

- Achieved significant accuracy improvements through data augmentation and hyperparameter tuning.

MLAudio Project
GitHub: MLAudio

- Explored machine learning techniques for audio data analysis.

- Applied feature extraction methods and classification algorithms to audio datasets.

- Demonstrated the potential of ML models in audio recognition tasks.

Skills

Programming & Tools:
Python, SQL, Power BI, Git, Google Cloud Platform (GCP), BigQuery, OpenCV

Machine Learning & AI:
Scikit-learn, Pandas, NumPy, Matplotlib, TensorFlow, PyTorch, DeepStream, OCR

NLP & Deep Learning:
Hugging Face Transformers, Tokenization, Text Classification, Transfer Learning, Model Fine-Tuning

Data Engineering & Visualization:
Data Modeling, ETL Pipelines, GCP Data Services, Power BI Dashboarding, Data Governance

Achievements

• Tyson Foods IT India Chess Tournament Winner

• Corporate Football Tournament Winner