Ayan Ruzdan

📞 +91-78391-32077 ✉️ ayanruzdan@gmail.com 💼 linkedin.com/in/ayan-ruzdan ⭕ github.com/AyanRuzdan

EDUCATION

Lovely Professional University November 2021 – July 2025
Bachelor of Technology, CSE - 8.37 CGPA *Phagwara, Punjab*

Sunbeam School Lahartara March 2018 – March 2020
Senior Secondary, PCM - 90.4% *Varanasi, Uttar Pradesh*

EXPERIENCE

DataGrokr Analytics June 2024 – December 2024
Data Engineering Intern *Bengaluru, KA*
- Developed CI/CD pipelines using GitHub Actions and Jenkins to automatically run unit tests on Node.js and Python projects with Mocha and Pytest, respectively, whenever new changes were introduced.
- Developed ETL pipelines using Python, SQL, Pandas, and PySpark to extract, transform, and load data efficiently, enabling seamless data processing and analysis. Utilized Power BI to create interactive dashboards and visualizations for actionable insights.
- Designed and implemented cost-efficient, fault-tolerant ETL pipelines on Google Cloud Platform (GCP) to ingest data from APIs, Cloud SQL, Pub/Sub, and Cloud Storage into BigQuery. Leveraged Cloud Functions, Datastream, and Dataflow for seamless data transformation and ingestion while ensuring continuous data flow with Pub/Sub streaming and batch processing.

PROJECTS

Deep Learning for News Clustering and Retrieval | *Python, NLP, Streamlit, SentenceTransformers, Gemini* **July 2023**
- Built a deep learning pipeline to cluster and retrieve 1,440 science news articles using SentenceTransformer embeddings, K-means, and LDA.
- Optimized K-means clustering (k=5) using the elbow method, achieving 0.04 silhouette score and 0.51 cosine similarity; enhanced topic coherence with LDA (Cv=0.44).
- Visualized temporal keyword trends and topic clusters using PCA and Streamlit, improving interpretability of topic distributions.
- Deployed a Retrieval-Augmented Generation (RAG) chatbot with FAISS and Gemini Pro for semantic, real-time question-answering on news archives.

LeetCode Stats Tracker | *Python, AWS Lambda, S3, EventBridge, GraphQL* **October 2024**
- Developed a serverless pipeline using AWS Lambda to fetch LeetCode user statistics via a GraphQL API and store them in an S3 bucket.
- Configured AWS EventBridge to trigger the Lambda function periodically, ensuring automated data collection.
- Implemented data transformation using Python to extract rankings and problem-solving statistics, storing only the most recent 1,500 entries.
- Integrated GitHub Pages to render the JSON file from S3, enabling a public-facing dashboard for tracking coding progress.

CERTIFICATES AND CERTIFICATIONS

GCP Associate Cloud Engineer Google — Dec' 24
React Basics Meta — March' 24
MERN Stack CipherSchools — July' 23

TECHNICAL SKILLS

Languages: C, C++, Python, SQL, Java, JavaScript
Developer Tools: Pandas, PowerBI, PySpark, Jenkins, Docker, Git, GitHub, GCS, Pub/Sub
Technologies/Frameworks: Node.js, React, Express, HTML, CSS, GCP, AWS, Kubernetes
Databases: MySQL, MongoDB, SQLite, BigQuery

ACHIEVEMENTS

Leetcode: Solved 800+ questions on topics like DSA, SQL, and Pandas
HackerRank: 4-star rating in C++, 5-star rating in Python and SQL
Codechef: Highest rating of 1600 (3-star) with a total of 350+ problems solved
Quill Club Writers: Co-authored **Here Comes the Sun** in a collaborative book-writing workshop.