ANANYA TIWARI

+91-9305058289 | ananyatiwari1508@gmail.com | Linkedin | Github | Leetcode

Education

Bachelor of Technology in Computer Science and Engineering (Data Science) 2022 - 2026
Bennett University (Times Of India Group), Greater Noida CGPA: 9.0

Experience

Data Analyst | PolicyBazaar For Business Sept 2025 - Present

- **Owned** SalesOps and Marketing ad-hoc analytics, partnering with **5+ BU heads** on root-cause analysis and business decisions.
- Built and maintain **10+ KPI dashboards** tracking sales performance, pipeline conversion, and operational metrics.
- Built **RPAs and GenAI-assisted workflows** to eliminate recurring manual tasks.
- Developed **Apache Airflow DAGs** and converted manual **Python** scripts to automated reporting pipelines, reducing manual reporting effort by **50%.**
- Recognised with the **Star of the Quarter award.**

Data Science Intern | ZYOD June 2025 - Sept 2025

- Profiled and validated **20+ RDS** tables using **PostgreSQL** and **DBeaver** to assess data quality.
- Modeled **5+ business-critical** tables into **star schema** to support reporting; explored **S3-to-Snowflake** data loading.
- Authored **60+** raw and staging layer tables in **Snowflake**, enabling seamless **BI reporting**.
- Queried and manipulated **100K+** rows of structured data from relational databases to support **LLM** training.

Projects

SemanticSearch - AI Research Assistant

- Tech Stack: Python, Docker, ECR, FastAPI (CORS), Boto3, React, ECS Fargate, Lambda, Apache Cassandra (AWS Keyspaces), IAM, Amazon S3, Sentence Transformers (all-MiniLM-L6-v2), FAISS.
- Constructed a **scalable containerized service** with **React** frontend handling **1,000+** concurrent queries and achieving **92%** search accuracy.

ETL DataPipeline

- Tech Stack: Python libraries, Apache Airflow, AWS.
- Built a scalable **ETL** pipeline to process tweet datasets with over **100,000** tweet records, leveraging **AWS S3** and **EC2** for efficient storage and computation.
- Automated the **ETL** workflow with **Apache Airflow**, achieving a **98%** job success rate and reducing manual intervention by **50%**.

Avatar - Character Composer

- Tech Stack: Azure AI Custom Vision, OpenAI, Google Cloud Functions, API Gateway, Google Pub/Sub, Azure Cosmos DB, Azure Blob Storage, Azure Active Directory (AAD).
- Orchestrated AI character composer service, generating over **500** unique character images using Azure Custom Vision based on tagged data.
- Refined architecture to deliver the project within a **$28 budget on Azure services**, significantly reducing costs.

Awards & Achievements

Winner - Microsoft Azure CWB Hackathon'24 **June 2024**

- Solved PRODIGY9's problem statement by creating Project **Avatar** and enhanced workflows by **40%**.

Conference Paper Published - IEEE **Dec 2024**

- Published at ICCES 2024 (IEEE): Improving Deepfake Detection Using Deep Learning Techniques

Technical Skills

Languages: Python | SQL | HiveQL | MySQL | PostgreSQL | Athena | NoSQL
Libraries & Frameworks: NumPy | Pandas | Scikit-learn | Matplotlib | Seaborn
Tools & Platforms: Power BI | QuickSight | Excel | MS PowerPoint | AWS (S3 & EC2) | Hive | PySpark | Apache Airflow | Airbyte | Snowflake | Claude Code | Google Search Console | Meta Ads Manager
Core CS: Data Visualization | Data Querying | Data Manipulation | Data Wrangling | Data Modeling | Data Validation | Machine Learning | Regression | Classification | Clustering | Statistical Analysis
Soft Skills: Problem Solving | Analytical Thinking | Effective Communication | Adaptability | Team Player

Certifications

- IBM Machine Learning Professional.
- Google Data Analytics Professional.
- Mastering Big Data Analytics with PySpark.