

Ananya Mehra

@linkedin | @github | mehrananya1620@gmail.com | 6397604517

Data Engineer

PROFESSIONAL SUMMARY

Experienced data engineer proficient in Python, Shell scripting, and Django REST framework. Skilled in Agile methodologies and Git, with a focus on optimizing systems and delivering scalable solutions, transitioning to software engineering roles.

EDUCATION

DIT University	Bachelor of Technology in Computer Science CGPA : 7.92	August 2018-May 2022
----------------	---	----------------------

EXPERIENCE

- Data Engineer** | Optum July 2022 –Present
- **Automated Data Comparison and Validation Pipeline:** Engineered a robust automation solution to eradicate manual data comparison tasks, reducing manual verification time by 85 percent. I developed Python and Shell scripts to meticulously verify file-to-database integrity and ensure precise volume accuracy. This project significantly enhanced the precision and efficiency of the validation process within 6 months. *Technologies used: Python, Shell Scripting, Apache Airflow, Teradata, Git*
 - **Designed and Developed End-to-End Data Integration Pipeline:** Architected a comprehensive process to integrate fresh, user-invisible data using DataStage for ETL logic, enabling access to new data for over 500 users. I developed Python and Shell scripts with customizable config files to meet user needs and created a custom DAG in Apache Airflow for structured task orchestration. This project applied software engineering principles to create scalable and maintainable solutions, completed within 4 months. *Technologies used: Python, Shell Scripting, Apache Airflow, Teradata, DataStage, Git, Jenkins*
 - **Enhanced Existing Data Integration Functionality:** Improved ETL logic, DAG scheduling, and Shell scripts to incorporate long-invisible data, increasing data access coverage by 30 percent. I used Teradata for testing and validation, demonstrating my ability to enhance and maintain complex software systems. These enhancements were implemented over a 3-month period. *Technologies used : Shell Scripting, Apache Airflow, Teradata, DataStage, Git, Jenkins*
 - **Led System Enhancement Project:** Integrated two new fields into a critical team system, refining raw data using DataStage, Python, and Shell scripting, and processed over 63 million records across seven files. I managed complex system testing and leveraged Teradata for robust database validation, showcasing leadership in software development and system enhancement. This project was successfully completed within 5 months. *Technologies used : Datastage, Shell Scripting, Python, Git, Teradata, Jenkins*
- Backend Developer Intern** | Moxtain Jan 2022- May 2022
- Designed and implemented RESTful APIs for an eCommerce website, enhancing application functionality and user experience. Designed RESTful APIs for an eCommerce website using Spring Boot, Git, and PostgreSQL.
 - Utilized tech stack including Spring Boot, Git, GitHub, Postman, PostgreSQL, and AWS.

PROJECTS

- Blog API** | Python, Django Rest Framework, Token Authentication, Pagination, Filtering, Browsable API 2024
- * Developed resilient API endpoints for managing blog posts using Python and Django Rest Framework.
 - * Implemented token-based authentication to secure API endpoints and restrict certain actions to unauthenticated users.
 - * Incorporated pagination to efficiently handle large datasets and enhance performance and integrated filtering options, allowing users to narrow down search results by category, date, or user.
 - * Enabled DRF's browsable API for easy testing and exploration of endpoints, facilitating developer interaction and debugging.
- Ecommerce Website** | Python, Django Rest Framework, Pytest, SwaggerUI 2023
- * Implemented resilient API endpoints using Python and Django Rest Framework.
 - * Conducted Pytest-based testing to ensure optimal performance and reliability.
 - * Integrated SwaggerUI for streamlined API documentation.

SKILLS

Data Engineering : Python, ETL, Shell Scripting, Apache Airflow
Software Development : Java, Python, Spring Boot , Django, Django Rest Framework
Database Technologies :PostgreSQL, MySQL, Snowflake, Teradata
Programming Languages :Java, Python
Version Control & DevOps : Git, GitHub, Jenkins, CI/CD, Docker, Kubernetes
Other :AWS, Kafka, Spring Boot, Restful APIs, Spark, Statistical Analysis, Machine Learning, Model Deployment

AWARDS AND CERTIFICATIONS

Diamond Award Optum Uhg	August 2023
Data Science Coursera	2020
Machine Learning Coursera	2020
Deep Learning Coursera	2020