Aditya Thakur

Roll No.:22104
B.Tech - Computer Science Engineering
Indian Institute Of Information Technology, Una

+91-9805594941
adityathakurzm@gmail.com
github.com/adityathakur96
linkedin.com/in/adityathakur9617

EDUCATION

Degree/Certificate	Institute	CGPA/Percentage	Year
B.Tech	Indian Institute of Information Technology, Una	7.44 (Current)	2022-Present
Senior Secondary	Holy Heart Public School, Nahan	88.6%	2021
Secondary	D.A.V Public School, Nahan	86.7%	2019

PROJECTS

- **Olympic Data Analytics**   *November 2024 – December 2024*
  Apache Spark, Python, Microsoft Azure, MySQL   GitHub
  - Developed an end-to-end data engineering solution on Microsoft Azure for analyzing 2021 Olympics data.
  - Analyzed Olympics data using a CSV file containing nation-wise information on athletes, their respective fields, and medal statistics for comprehensive insights.
  - Utilized Azure Data Factory for data ingestion, storing data in Data Lake Gen2 for efficient processing and access.
  - Employed Apache Spark and Python in Azure Databricks for data processing and transformation.
  - Stored the final processed data in Synapse Analytics and MySQL for querying and reporting.
  - Created interactive dashboards and reports to analyze athlete performance, medal trends, and nations achievements.
  - Developed skills of data analysis ,visualization and able to develop the ETL pipelines for the given datasets for better results.

- **Apple Sales Data Analysis**   *December 2024 – January 2025*
  Apache Spark, PySpark, Databricks, DBFS, Delta Lake   GitHub
  - Developed an end-to-end ETL pipeline for analyzing Apple sales data using Databricks File System (DBFS).
  - Processed data stored in CSV, Parquet, and Delta formats, leveraging PySpark for efficient transformations.
  - Implemented transformations using Spark SQL to extract sales insights, such as identifying customers who purchased AirPods after buying an iPhone and those who bought only iPhone and AirPods.
  - Loaded transformed data into Delta Lake and Data Lake for scalable and reliable storage, ensuring efficient access for analytics.
  - Utilized Apache Spark to perform complex data transformations and prepare data for analytics and reporting.
  - Future work includes integrating advanced analytics and predictive modeling to forecast sales trends.

TECHNICAL SKILLS

- **Languages:** Scala, Python, C++,
- **Database:** MySQL, PostgreSQL, Firebase, MongoDB
- **Data Engineering Tools:** Apache Spark, Databricks
- **Developer Tools:** VS Code, IntelliJ IDEA, Git, GitHub, Google Colab
- **Operating Systems:** Windows

KEY COURSES TAKEN

- **Mathematics:** Linear Algebra, Basic Calculus, Discrete Maths, Probability & Random Processes
- **Software Development:** Data Structures and Algorithms, Object-Oriented Programming (C++),Software Engineering, Relational Database Management System (MySQL)
- **Systems and Infrastructure:** Operating Systems, Computer Networks, Computer Organization

POSITIONS OF RESPONSIBILITY

- **Design Head** of EUNOIA (Cultural Society), IIIT Una   *June-2024 – Present*
- **Executive Member** of Eunoia, IIIT Una   *Sep-2023 – Present*