

KOMPALA LOKESH

+91 83282 60402 | kompalalokesh123@gmail.com | [linkedin.com/in/kompalalokesh](https://www.linkedin.com/in/kompalalokesh) | github.com/KOMPALALOKESH

EXPERIENCE

Data Analytics Intern

Jan 2024 – Apr 2024

Mentored Minds

- Developed and optimized a scalable Twitter data pipeline using Apache Airflow for automated ETL processes, ensuring data accuracy and reliability.
- Implemented real-time data ingestion from Twitter API and managed workflows on AWS EC2, leveraging Amazon S3 for efficient data storage.
- Utilized Docker for containerized deployment, transforming raw data with Pandas and enhancing overall ETL performance and efficiency.

PROJECTS

End-to-End E-commerce ETL Pipeline | Python, Spark, Kafka, Streamlit

- Developed an e-commerce data pipeline using FastAPI for ingestion, Kafka for streaming, and Apache Spark for processing.
- Engineered optimized MongoDB storage for efficient key-value indexing and seamless data retrieval.
- Designed interactive Streamlit dashboards with PyMongo integration, implementing scalable architecture for high performance and reliability in real-time data visualization and insights.

Weather API - Airflow Orchestration | Python, Airflow, PostgreSQL

- Designed and optimized an Airflow DAG for weather data ETL, enhancing pipeline efficiency and scalability.
- Developed Python scripts for API data extraction, transformation, and PostgreSQL loading, with robust logging for progress tracking and error identification.
- Utilized XCom for efficient task communication and leveraged Docker for consistent deployment environments.

Salesdb data modelling | Python, SQL, Pandas

- Created a comprehensive EER diagram to represent the sales data model entity relationships and attributes.
- Utilized Python, SQLAlchemy, and SQL connectors to develop an efficient Extract, Transform, Load (ETL) process.
- Employed pandas and other data processing libraries to perform extensive data cleaning and schema validation, ensuring data integrity and consistency throughout the pipeline.

TECHNICAL SKILLS

Languages: Python, SQL

Technologies: GitHub, Databricks, AWS, Docker, Git, Linux

Databases: MySQL, PostgreSQL, MongoDB

Data Science: Numpy, Pandas, Matplotlib, Scikit-learn, TensorFlow, Keras, Deep Learning, NLP, Gen AI

Data & Analytics: Spark, Kafka, Excel, Tableau, Airflow, Pipeline Design and Security, Data stores, Lambda/Kappa architecture

EDUCATION

Bachelor of Technology in Computer Science

Sri Venkateswara University College of Engineering, Tirupati

Aug. 2020 – Apr. 2024

CGPA: 8.36

Intermediate - M.P.C

Narayana Junior College, Nellore

Jun. 2018 – Apr. 2020

CGPA: 9.69

Tenth Class

Narayana High School, Venkatagiri

Apr. 2018

CGPA: 10.0

RELEVANT COURSEWORK

- Data Structures
- Data Science
- Machine Learning
- Big Data & Analytics

ACCOMPLISHMENTS

- Ranked Top 30% worldwide in LinkedIn Python skill test.
- Ranked Top 30% worldwide in LinkedIn Machine Learning skill test.
- Secured Runner up position in Ignite coding club Data Science contest conducted in college.