Rohith Chippa

Tempe, AZ | chipparohith12@gmail.com | +1 314 398 5654 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Data Engineer/Machine Learning | Cloud-Native ETL Pipelines | AWS, Spark, SQL

PROFESSIONAL SUMMARY

Data Engineer with **1.5+ years of experience** in designing and implementing **scalable data pipelines**. Proficient in **cloud-based ETL processes, machine learning integration.** Successfully transformed raw data into actionable insights, driving informed decision-making for business and engineering teams. Passionate about utilizing cutting-edge technologies and fostering collaboration to deliver impactful data solutions.

EDUCATION

GRAND CANYON UNIVERSITY – M.S. in Data Science | GPA: 3.89/4.00   **APRIL 2025**
Courses: Data Mining, Big Data Architecture, Machine Learning, Data Visualization

TECHNICAL SKILLS

Languages & Tools: Python, SQL, JavaScript (ReactJs, NodeJs, ReactThreeFiber), Shell/Bash, OOPs, DSA
ETL & DataEngineering: Airflow, AWS Glue, Kinesis, Kafka, Terraform, Databricks.
Big Data & Databases: Apache Spark, PySpark, Flink, DynamoDB, AWS (Redshift, RDS, EMR, Lambda, S3), PostgreSQL, MySQL, Oracle, SQL Server, dbt.

Visualization & ML/AI: Power BI, QuickSight, Docker, Selenium, Pandas, NumPy, PyTorch, Keras, Transformers, RAG.

RELEVANT EXPERIENCE

GRAND CANYON UNIVERSITY   **Phoenix, Arizona**
STUDENT INSTRUCTIONAL ASSISTANT   Oct 2024 – Mar 2025

- Collaborated with subject matter experts to design course curriculum, assignments, and supplementary learning materials for graduate-level courses: Applied Data Visualization (DSC-527) and Practical Applications of Time Series Forecasting (DSC-577).
- Provided academic support to students by clarifying course concepts, curating resources, and facilitating practical application of tools such as Power BI, Python (Matplotlib, Seaborn), and time series libraries (ARIMA, Prophet).

TATA CONSULTANCY SERVICES   **Hyderabad, India**
ETL DEVELOPER   Aug 2021 – May 2023

- Designed and implemented scalable ETL pipelines using **AWS Glue** and **Apache Iceberg on Amazon S3** to ingest CRM data (Siebel) into a centralized Customer 360 data model, enabling ACID-compliant data lakes and improving data availability by 30%.
- Built and optimized dimensional and normalized data models in **Amazon Redshift** using **dbt**, supporting 30+ curated tables with clearly defined staging, intermediate, and mart layers, resulting in 25% faster dashboard load times.
- Collaborated with cross-functional teams to integrate real-time data analytics, enhancing decision-making capabilities and driving a 20% increase in operational efficiency.

PROJECTS

DeFtunes Data Pipeline And Analytics ([Github](#))

- Engineered and deployed a comprehensive data pipeline on AWS using a medallion architecture with S3 data lake and Redshift data warehouse.
- Implemented robust ETL processes via AWS Glue for data extraction (RDS PostgreSQL, REST API) and transformation to Apache Iceberg format.
- Developed a star schema data model and analytical views in Redshift using dbt, orchestrating daily runs with Apache Airflow and AWS Glue Data Quality checks for data integrity.

AI Wardrobe – Fashion Item Attribute Prediction ([Github](#))

- Developed a data-driven wardrobe management system with a **FastAPI** backend, utilizing a **ResNet50-based deep learning model** for clothing item classification from images.
- Engineered data pipelines for image preprocessing and multi-label classification, achieving ~95% training accuracy.
- Built interactive dashboards with **React** to visualize wardrobe data and enabled CSV/PDF report generation via APIs for insights like seasonal trends.

HIRE HUB([Github](#))

- Developed a full-stack web platform using React.js for the frontend and Node.js with Express.js for the backend, facilitating job discovery and application for job seekers and job posting/application review for recruiters.
- Implemented a MySQL database to manage user profiles, job listings, and applications, securing user authentication and resource access with JSON Web Tokens (JWT).

CERTIFICATIONS

- Data Engineering [Professional Certification](#) by DeepLearning.AI & Amazon Web Services.
- NPTEL Online [Certification](#) Programming, Data Structures and Algorithms using python.
- Responsive Web Design [Certification](#) from [freeCodeCamp.org](#).