Chandra Sekhar Chennupati

✉ cs.chennupati@gmail.com ☎ +91 9949385751 ⌖ Hyderabad, India 🔗 linkedin.com/in/chandra-sekhar-chennupati

SUMMARY

I am an experienced **Data Engineer** professional with around **4 years** in the field skilled in **Python**, Google Cloud Platform(**GCP**), **Pyspark**, **SQL** and **Airflow**. I specialize in building and optimizing data pipelines. I constantly commit to continuous learning and leveraging new technologies to enhance data solutions.

SKILLS

Programming Languages
Python, SQL, Shell Scripting

Cloud Services
Google cloud platform (GCP) - BigQuery, Google Cloud Storage(GCS), Dataproc, Cloud Composer, Cloud SQL, Cloud Functions, Vertex AI, Cloud Shell, Pub/Sub
Amazon Web Services (AWS) - S3, Lambda
Microsoft Azure - Azure Data Factory, Azure Databricks, Azure Synapse Analytics

Orchestration Tools
Cloud Composer, Control-M

Frameworks
Apache Spark (Pyspark), Apache Airflow

ETL Tools
IBM InfoSphere DataStage

DBMS
Teradata, postgres, oracle

PROFESSIONAL EXPERIENCE

Data Engineer   07/2021 – present | Hyderabad, India
Tata Consultancy Services

- Build ETL/ELT/EL scalable batch data pipelines to move data from on-premise to Data Warehouse (**BigQuery**) and vice-versa using **Python**, **Pyspark**, **SparkSQL** and **Apache Airflow**
- Iteratively design and develop a modular, reusable PySpark and Airflow based **framework** to create and orchestrate distributed data processing pipelines
- Fine-tuned and migrated oracle **stored procedures** and **packages** to Bigquery
- Optimize and performance tune big data pipelines to handle TBs of data using **Apache Spark** and **SQL**
- Designed and optimized **data models** to support analytical and transactional use cases, ensuring high performance and scalability.
- Worked extensively with **Amazon Web Services (AWS)**, utilizing **EMR, Kinesis, RDS, S3, and SQS** to build robust and scalable data solutions.
- Debug erroneous data pipelines including undocumented legacy pipelines and troubleshoot complex data issues and provide solutions
- Create **Bash** or **Python** scripts to **automate** routine tasks

PERSONAL PROJECTS

Smart AI Doc Miner
Generative AI, Flask, Python, Google Cloud Platform, pgvector, Google Gemini 1.5
- Implemented **RAG (Retrieval-Augmented Generation)** methodology to enable efficient document retrieval and analysis.
- Integrated **Google Gemini 1.5 model** for state-of-the-art natural language processing and understanding.
- Utilized **pgvector** for vector storage and similarity search, enhancing data retrieval capabilities.
- **Deployment**: Hosted the application on **Google Cloud Run**, ensuring scalability and seamless deployment.

EDUCATION

V R SIDDHARTHA ENGINEERING COLLEGE   2017 – 2021 | Vijayawada, India
Bachelor of Technology in Electronics and Communication Engineering
CGPA: 8.61

AWARDS

Employee of the Month x 3
- On three different occasions, I have been awarded Employee of the Month for my individual contributions

Project Performance Improvement Award x2
- As part of a 5 member team, helped the customer save more than $150k
- As part of a 2 member team, worked to meet Government's Data Regulatory policies within a strict deadline of 4 months

CERTIFICATES

- Google Cloud Certified Professional Data Engineer 🔗
- Google Cloud Certified Associate Data Practitioner 🔗
- Google Cloud GenAI L400 Certified 🔗

INTERESTS

Cricket | Volleyball | Badminton | Reading | Photography