

Bharath Chandra Reddy B

☎ 909-376-6816 | ✉ bharathbcr@gmail.com | in [Linkedin](#) | 🐙 [Github](#)

EDUCATION

California State University, San Bernardino

Master of Science, Computer Science

San Bernardino, CA

Aug 2022 – May 2024

Jawaharlal Nehru Technological University, Hyderabad

Bachelor of Technology, Electronics and Communication Engineering

Hyderabad, India

Aug 2015 – Jul 2019

EXPERIENCE

Senior Software Engineer

Apr 2021 – Jun 2022

Discover Financial Services (Client)

Hyderabad, India

- Migrated Oracle workloads to **Azure Synapse Analytics** by leveraging **Azure Data Lake Storage** and Synapse Serverless SQL Pools, achieving **\$678,000** in annual cost savings and a **14% boost** in performance.
- Designed and implemented a real-time data pipeline to process semi-structured data, integrating **150 million** raw records from **30+ sources** using **Azure Event Hubs** and **Azure Databricks (PySpark)**.
- Conducted data analysis for credit risk and fraud detection using **Azure Databricks**, Power BI, Azure ML, and Python (Pandas, Scikit-learn), driving an **8%** fraud rate reduction.

Software Engineer

Mar 2020 – Mar 2021

Discover Financial Services (Client)

Hyderabad, India

- Maintained **99.8%** pipeline uptime while ingesting streaming and transactional data across 8 primary sources using **Azure Databricks (Spark)**, **Azure Synapse Analytics**, **Azure Data Lake Storage**, and Python.
- Automated ETL workflows across millions of rows in Synapse and Databricks, reducing manual workload by **29%**.
- Ingested data from disparate sources via **SQL, REST APIs** (including Google Analytics and Salesforce) using Python, creating unified data views for business intelligence tools like **Tableau**.

Data and Analytics Consulting Intern

Jul 2019 – Feb 2020

Pressmart Media Ltd

Hyderabad, India

- Designed and implemented a custom **ETL pipeline** to ingest transactional and event data from a web app with **2,000+** daily active users, eliminating external vendor dependency and saving **\$15K annually**.
- Collaborated directly with clients to gather requirements and deliver actionable **Tableau dashboards**, automating manual reporting processes and saving **17+** hours per week.
- Leveraged **PySpark** to parallelize data processing across large-scale streaming datasets, boosting ingestion performance by **67%**.

PROJECTS

Spotify Data Pipeline — Real-Time Analytics with Snowflake and AWS | *Python, Snowflake, AWS*

- Designed and deployed a scalable **ETL pipeline** to automate ingestion and analysis of Spotify's Top 100 Global trending songs, enabling **near real-time** reporting and cutting manual data retrieval to zero.
- Built **AWS Lambda functions** to extract, transform, and load song, album, and artist metadata from the **Spotify API** into **S3**, achieving sub-minute data availability and ensuring cost-effective serverless scaling.
- Optimized Snowflake data models and leveraged **Snowpipe** for continuous ingestion from **S3**, reducing query latency by **40%** and supporting advanced SQL analytics for trend forecasting.

TECHNICAL SKILLS

Programming Languages: Python (PySpark, Pandas, NumPy), SQL (T-SQL, PL/SQL), Java, C/C++

Cloud: Azure (Data Factory, Synapse, Blob Storage, Databricks), AWS (EC2, S3, EMR, Redshift, Glue)

Tools: Apache Airflow, Databricks, Hadoop, Hive, Snowflake, Azure Synapse Analytics

Databases & Storage: MS SQL Server, MySQL, Oracle, Snowflake, NoSQL Databases

Data Integration & ETL: Azure Data Factory, AWS Glue, PySpark, REST APIs

Data Visualization: Power BI, Tableau, MS Excel

DevOps & CI/CD: Docker, Jenkins, GitHub Actions, Git

CERTIFICATIONS

Azure Data Engineer Associate (DP-203)

Google Cloud Professional Data Engineer