Praveen Kalyan Puppala

Data Engineer

+91 7036694336 | praveenpuppala99p@gmail.com | linkedin.com/in/praveen-kalyan | Hyderabad

Summary

With over **3+ Years of professional experience in software development**, experienced Data Engineer adept at transforming complex data into meaningful data and turning into valuable insights. Skilled and knowledge in Azure cloud service such as **Databricks, Azure cosmos DB and SQL** to collect, process, and analyze data. Skilled in **PySpark, SQL, Apache Spark, and Statistical data analysis**. Seeking new challenges in a dynamic, growth-oriented environment.

Skills

Programming Languages / Scripting: Python, UNIX commands

Framework: PySpark, Apache Kafka, Spark

ETL Tool: Apache NiFi

Relational Databases: SQL, Oracle 11G & 12C database, PostgreSQL

No SQL Databases: Apache HBase

Cloud: Azure, AWS, Databricks

Tools: JIRA, Confluence, GIT, Putty, WinSCP, BMC Remedy ITSM

Education

Bachelor of Technology in Information Technology, CGPA **7.75/10**   **Jul,2017-Jul,2021**

College: SRKR Engineering College, Bhimavaram, Andhra Pradesh, India.

Professional Experience

Senior Systems Engineer, Infosys Limited, **Hyderabad**, Telangana, India. **Dec,2022 - Present**

Project - Data Engineer | IASO project | March,2024– Present

Domain - Telecom

- Developed and implemented a comprehensive data cleansing framework in **Python** and SQL, resulting in 60% reduction in manual effort required for cleaning data.
- Developed and implemented ETL pipeline in **PySpark** transforming **CSV data** as a source and loading the meaningful data into **PostgreSQL** database in **Azure cosmos DB**.

- Developed and implemented ETL pipeline in PySpark transforming **JSON payloads** coming as a source and transforming the data into tables using data frames and storing the meaningful data in **Apache HBase**.

Project – Data Migration Support Engineer | Compass project | Sept,2023 – Feb,2024

- Developed an **Apache NiFi** data pipeline to migrate billions of records from an **Oracle Database** of a legacy application, serving as the source system, to **PostgreSQL** and **Apache HBase databases** as new destinations for cloud-based application.
- Using Apache NiFi, I drive daily data migrations from source to destination systems as part of our routine activities.
- We use **Cloudera Manager** to manage all Hadoop technologies. In which I have knowledge in **YARN** as a resource management layer that manages spark jobs.

Project – L2 Data support Engineer | C-Cubed project | Jan,2023 – August,2023

- Expertise in managing **Apache Spark** execution life cycles, changes in configuration files via **Putty** and including restarting, killing running jobs and stopping Spark jobs and schedulers.as per business demands.
- Fixing data issues in production.
- Expertise in writing complex SQL queries to fix data issues and analytical purposes.
- Resolving Incidents within SLA's.
- Expertise in **Change** management, **Problem** management and **Incident** management using **BMC remedy ITSM tool**.
- Experience in **Change advisory board (CAB)** meetings.
- I have knowledge in developing **pyspark** Spark jobs using **Python**.

Assistant System Engineer, TCS, Hyderabad, Telangana, India.   **Aug,2021-Dec,2022**

Project – Oracle SQL Developer | AP Industries | Aug,2021-Dec,2022

- I have experience working as an SQL developer with **Oracle 11g** database.
- Expertise in writing complex SQL queries to meet business requirements.
- Developed database schema and tables for a newer application based on business demand.
- Conducted data analysis to prevent data issues and ensure alignment between upstream and downstream processes, resolving any data issues impacting day – to – day business operations.
- Expertise in Joins, Nested Queries, CRUD operations and Views.

SIGNATURE

Place: Hyderabad   **Praveen Kalyan Puppala**