



Jilludimudi Deepika Rani

jddeepikarani@gmail.com

[linkedin.com/in/deepika-ranij60bab0359](https://www.linkedin.com/in/deepika-ranij60bab0359)

International City, Dubai,

+971 545667963

Professional Summary

Big Data Engineer with 5 years of professional experience specializing in designing, implementing, and optimizing data solutions across large-scale distributed systems. Proficient in leveraging Big Data technologies including Hadoop, Apache Spark, Kafka, and NoSQL databases. Highly skilled in building efficient ETL pipelines, real-time data processing, and cloud-based data solutions. Recently relocated to the UAE and currently seeking opportunities to bring technical expertise and innovation to forward-thinking organizations.

Skills

- ❖ **Data - Eco System:** Hadoop, Sqoop, Hive, Apache Spark, Kafka, Hive
- ❖ **Cloud Skills:** AWS, Azure
- ❖ **Administration(Web and App) :** Websphere, WebLogic, HTTP, IIS, ApacheTomcat
- ❖ **Databases:** SQL, MySQL
- ❖ **Operating Systems:** Linux, Windows
- ❖ **Programming Languages:** Python, Scala, SQL

ACADEMIC-QUALIFICATIONS

- ❖ Bachelor of Computer Science and Engineering
Satya Institute of Technology and Management (GVP_S.I.T.A.M) 2016-2019
- ❖ Diploma in Metallurgical Engineering
Maharaja Ananda Gajapathi Raju(M.R.A.G.R) Government Polytechnic 2012-2015
- ❖ Secondary School Certificate
Vijnana Bharathi high school 2011-2012

Work Experience

Project Name: Nokia

Sep 2020 - Oct 2024

Project Role: Data Engineer

Responsibilities:

- ❖ Designed, developed, and maintained scalable and reliable Hadoop ecosystems to manage large datasets.
- ❖ Designed, developed, and maintained highly scalable and optimized Extract Transform Load(ETL) pipelines using PySpark within the Cloudera Data Platform (CDP).
- ❖ Led transformation and processing data efforts to clean, process, and convert huge datasets into usable formats for analytics purposes using PySpark.
- ❖ Maintained data quality and validation by implementing automated data quality checks and monitoring routines to verify accuracy and reliability of data pipelines.
- ❖ Automated data workflows through Apache Airflow for task scheduling and orchestration, ensuring smooth data pipeline operations and minimizing manual interventions.
- ❖ Worked on developing data models and SQL queries for reporting and analytics using Hive and Impala, directly in support of analytical needs across the business.
- ❖ Collaboration with cross-functional teams to define and deliver data solutions that support their business objectives.
- ❖ Experienced in using Sqoop to import and export data from and to cloud-based data storage services such as Amazon S3 or Microsoft Azure Blob storage.
- ❖ Knowledgeable about testing ETL processes in Big Data platforms, such as Hadoop or Spark
- ❖ Responsible for implementing Spark jobs for recommendation engines.
- ❖ Implemented PySpark jobs on AWS EMR to process structured and unstructured data from AWS S3.
- ❖ Used AWS Step Functions to sequence PySpark jobs running on AWS EMR and manage dependencies.
- ❖ Developed ETL pipelines using PySpark for the purpose of data transformation.
- ❖ Supported the data ingestion process and worked with Hadoop Distributed File System (HDFS) for efficient data storage.
- ❖ Conducted initial performance tuning of PySpark processes to enhance data pipeline efficiency.
- ❖ Developed and maintained scripting solutions for routine data processing tasks on a Linux environment.
- ❖ Collaborated with senior data engineers to implement data quality assurance best practices across pipelines.
- ❖ Designed and deployed highly scalable data architectures using AWS (S3, EC2, EMR) to provide cloud-based processing and storage solutions.
- ❖ Improved the performance of Apache Spark jobs by 40% to improve the speed of data processing.

Project Name:USAA

Sep 2019 – Aug 2020

Project Role: Middleware Administrator

Responsibilities:

- ❖ Handling Ticketing Operations on tools like Service now.
- ❖ Handling multiple applications like Websphere, Apache, Weblogic, Apache_Tomcat, Jboss, IIS, Iplanet
- ❖ Generating CSR
- ❖ Renewing the SSL Certificates
- ❖ Performing Patching
- ❖ Deployments like UAT, NAM, ASLM
- ❖ Handling STASK, CTASK, ITASK Incidents.
- ❖ Handling SSL Certificate Renewals.
- ❖ Handling File System troubleshooting (freeing space etc).
- ❖ Updating tickets with the regular intervals and Resolving/Completing them Within SLA.
- ❖ Handling Operations ofMiddleware(Start, Stop, Logs) during Server Operational Lifecycle.
- ❖ Experience in Deploying EmergencyReleases (Due to any build issue or emergency deployment).
- ❖ Experience to handle Weekly& Quarterlydowntime Activities like Application Maintenance.
- ❖ Experience in Troubleshooting deploy failure tickets in all environments.
- ❖ Troubleshooting and solving Urban Code Deploy (UCD) issues.
- ❖ Troubleshooting Production Issues and worked on resolving Application Server Hung Issues, High Memory and CPU Issues.
- ❖ Experience in using scheduling engine like control-m to monitor jobs and performing hold or release functionalities on various jobs.
- ❖ Good knowledge on tracking all the jobs information and job statuses such as success or failure.
- ❖ Able to modify JVM configurations, Port issues, Space issues.