VEMIREDDY SREENATH REDDY

Sr Big Data Developer

Passport No: N8548163

● **5.5+ years of** IT experience in **Big Data with Hadoop**, **Spark** and its components like Spark core, Spark SQL, Hive, HDFS and **Azure** (Data Factory, Blob Storage, SQL Database)**.**

● Excellent enthusiastic to work with new technologies, analytical skills, communication, interpersonal and strong ability to perform as part of a team and individual.

Work History

2020-07 - Current

Consultant

Ernst & Young, Chennai, Tamil Nādu, India

- Worked with multiple clients for data migration and reporting projects by using Cloud and On-Prem environments.

2019-05 - 2020-07

Software Engineer

Experis IT Pvt Ltd, Chennai, Tamil Nādu, India

- As per the client requirements, I have designed and developed the frameworks for generate the analytical reports for daily, weekly, monthly, yearly.

2017-03 - 2019-01

Engineer

RMSI Pvt Ltd, Hyderabad, India

- *Worked as a Data Engineer and collaborated with various teams & managed to understand the requirements and gave my best work.*

2016-06 - 2017-03

Trainee Engineer

TriGeo Technologies Pvt Ltd, Hyderabad, India

- *Stated my carrier as Trainee Engineer with Python and MySQL Technologies.*

Education

2011-06 - 2015-05

Bachelor of Technologies: Mechanical Engineering

Jawaharlal Nehru Technological University - Hyderabad

Contact

Address
Hyderabad, Telangana, India, 500072

Phone
+91 9989316535

E-mail
vemisreenath@gmail.com

LinkedIn
www.linkedin.com/in/vemireddy-sreenath-reddy-09bb9

Skills

Hadoop, Spark

Python, Scala, Core Java, Hive, Impala, SQL, Shell, Pig, Sqoop

AZURE (Data Factory, Blob Storage, SQL Database)

GitHub

PyCharm, IntelliJ, Eclipse, Jupyter Notebook

Projects

Project Title: CRM (Customer Relationship Management) Data Migration   **2022-01 - Current**

Description: Migrating the data from multiple sources to Azure cloud environment and generating the reports on SQL Database.

Client: ICICI

Technologies: Azure Data Factory, Blob Storage, SQL Database, Oracle, Key-Vault and Vertica.

Responsibilities:

- Created the ADF Pipeline to manage multiple sources and destinations.
- Migrating the data from Vertica On-premises and other environments to Cloud.
- Handling the team.

Project Title: H2H (Hana to Hadoop)   **2020-07 – 2021-12**

Description: H2H is a migration project from SAP-HANA to Hadoop environment. Need to migrate all possible hana reports into Hadoop with better approaches and as client's requests, hand overed the reports through mails, watch (UI web), different file types and hive tables.

Client: PayPal

Technologies: PySpark, HDFS, Hive, Python, GIT, Shell and Custom PayPal Frameworks.

Responsibilities:

- Prepared the business requirement, Design docs and Application docs.
- Converted the SAP-HANA reports into Hadoop system design and analysis.
- Involved in the Data loading and pre-processing the data using hive and spark.
- Contributed the data move to the mongo and watch reporting UI.
- Developed most of the reports using spark for better performance and approaches.

Project Title: **IAP (Integrated Analytics Platform)**   **2019-05 - 2020-07**

Description: Integrated Analytics Platform (IAP) is an analytical project to understand the patient's thoughts in the organization. As an initial task, we have taken the feeds from different applications of data from source systems to understand the data and ingested the same into the Hadoop environment. This project holds the 360-degree view of the patient activities and we are processing the events of patients into hive tables and analyzing the same.

Client: AbbVie

Technologies: Apache Spark, Hadoop, Hive, Scala, Impala, Shell, Autosys

Roles: Big Data Engineer

Responsibilities:

- Involved in Workflow and pipeline design.
- Ingestion of all source application data into IAP applications.
- Managed the different kinds of data as a data engineer.
- Involved in Spark application development.
- Written the hive and impala queries for transforming and analyzing the data.
- Developed Shell Scripts for handling spark jobs.
- Scheduled the applications using Autosys, so created JIL files for that.
- Managed all kind of report requirements using PySpark.