**Mallaiah B**

**Email: mallaiah.dataengineer@gmail.com**

**Mobile: +91-7842570604**

|  |
| --- |
| **CAREER OBJECTIVE** |

As a Data Engineer, experienced in Big Data ecosystems, Cloudera and MS Azure Cloud Services having a strong background in a Big Data arena. A collaborative engineering professional with substantial experience designing and executing solutions for complex business problems involving large scale reporting solutions and Predictive Analytics.

|  |
| --- |
| **PROFILE SUMMARY** |

* 5.4 yearsof IT experience in a variety of industries, this includes hands on experience in Big Data Analytics and Development.
* 3+ years of experience in MS Azure Cloud Services like Azure Databricks, Azure Data Factory , Azure SQL DB, Azure Data Lake Storage and Azure Blob Storage .
* Experience in creating data driven workflows for data movement and transformation as part of ETL using Azure Data Factory.
* Implemented business logic as per the requirement using Databricks with PySpark and Spark .
* Experience in manipulating/analyzing large datasets and finding patterns and insights within structured and unstructured data.
* Experienced in handling different file formats like CSV, JSON, Parquet and Avro.
* Experienced in using agile approaches, including Extreme Programming, Test-Driven Development and Agile Scrum.
* Knowledge of Source Control Management System using GIT.
* Good team player with excellent communication skills with strong attitude towards learning new technologies.

|  |
| --- |
| **TECHNICAL EXPERTISE** |

|  |  |
| --- | --- |
| Big Data Ecosystems | Apache Spark. |
| Microsoft Azure Cloud Services | Azure Databricks, Azure Data Factory, Azure SQL DB, ADLS Gen2 and BLOB Storage. |
| Programming Languages | Pyspak, SQL and Python |
| Data Bases | SQL DB, DB2/400 |
| Development Methodologies | Agile/Scrum. |

|  |
| --- |
| **PROFESSIONAL EXPERIENCE** |

* Software Engineer in Genesis Infocom Pvt Ltd, Bangalore from Apr 2018 to till Date.

|  |
| --- |
| **EDUCATIONAL CERTIFICATION** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Examination** | **Institute / Board / University** | **Year of passing** | **Percentage** |
| B. Tech  (ECE) | NBKR Inistitute of Science and Technology, Vidyanagar ,SPSR Nellore(AP) | 2009 | 75.3 |
| Intermediate | SBN Junior College, Board of Intermediate Education ,Venkatagiri, Tirupati(AP) | 2005 | 92.0 |
| SSC | APRS,Venkatagiri, Tirupati (AP) | 2002 | 85.3 |

|  |
| --- |
| **PROJECTS SUMMARY** |

**PROJECT#2**

Project : Cloud Sphere

Client : TotalEnergies SE

Duration : Apr-2020 to till date

Domain : Mining

Role : Data Engineer

Organization : Genesis Infocom Pvt Ltd

Team Size : 8

Methodology : Agile

Environment : Apache Spark, Python, Azure Data Factory, Azure Databricks,

Azure SQL DB, Azure Data Lake Storage Gen2, Blob Storage.

**Description:**

TotalEnergies SE is a French multinational integrated energy and Petroleum Company founded in 1924 and is one of the seven super major oil companies. Its businesses cover the entire oil and gas chain, from crude oil and natural gas exploration and production to power generation, transportation, refining, petroleum product marketing, and international crude oil and product trading. TotalEnergies is also a large-scale chemicals manufacturer.

**Roles & Responsibilities:**

* Communication with Business Owner & Onsite Team in daily stand up and scrum call.
* Extract data from SAP systems using appropriate extraction methods or tools, ensuring compatibility with Azure Delta Lake.
* Design and develop data transformation processes to map and align the extracted SAP data with the target schema and data model in Delta Lake.
* Moved all Meta data files generated from various source systems to ADLS for further processing.
* Imported data from different sources like ADLSs and BLOB for computation using Spark.
* Implemented Spark using Python in Databricks utilizing Data Frames and Spark-SQL API for faster Processing of data.
* Used Avro, Json and Parquet data formats to store data into ADLS.
* Created data driven work flows for data movement and transformation using Data Factory.
* Extracted huge files by using Azure Storage Explorer from the Data Lake.
* Involved in all the Acquisition, Transformation and Model phases of the project.
* Monitoring the Spark Jobs in the cluster environment debugging for the failed jobs.
* Interacted closely with business users, providing end to end support.

**Project#1**

Project Name : Orbit

Client : Premier Farnell

Duration : Apr 2018 – Mar 2020

Domain : Retail

Role : Software developer

Organization : Genesis Infocom Pvt Ltd

Environment : AS/400, RPG/400, RPGLE, QUERY/400, CLLE,

CL/400, DB2/400, OS/400, SQL/400, SQLRPGLE

AS/400 Tools : PDM, SEU, SDA, RLU, FTP

Change Management Tools : Aldon-CMS

**Description:**

Premier Farnell plc is a leading multi-channel, high service distributor supporting millions of engineers and purchasing professionals globally. They market and distribute a comprehensive range of products and services and solutions throughout Europe, North America and Asia Pacific. Orbit is an in-house ERP product of Premier Farnell, is used for business processing which includes different functionalities like order processing, invoicing, Warehouse management, Sales statistics, etc.

**Roles & Responsibilities:**

* Understanding business requirements
* Participate in Design Meetings for design issues and providing solutions.
* Preparing the Test Cases and Involved in unit testing.
* Interacting with clients in weekly and monthly calls
* Doing code review and Peer reviews.
* Issue analysis for Production support tickets.
* Perform unit/module testing of the software developed to find errors and confirm.

|  |
| --- |
|  |