VENKATA VARAPRASAD

AZURE DATA ENGINEER

8142272545
varaprasadvenkata47@gmail.com

PROFILE

- Over all 4.7 Years of IT experience on data driven application design and development with 3.5 years of relevant hands-on- experience in various Azure data services.
- Proficient in Azure technologies such as Azure Data Factory (ADF), Azure Data Bricks (ADB),Azure Synapse Analytics, Azure Active Directory, Azure Storage, Azure data Lake Services (ADLS), Azure key vault, Azure SQL DB, Azure HDInsight.
- Having good hands-on experience on Azure Dev Ops (ADO) services like Repos, Boards, Build Pipelines (CI/CD), Ansible (yml scripting) for resource orchestration and code deployment. Hands on
- Experience developing data engineering frameworks and notebooks using Azure data bricks using Spark SQL, Scala, pyspark Experience in Apache Hadoop frame works such as Hdfs, Map Reduce, Hive etc.
- Experience in Microsoft Azure Cloud with Data Factory, Link Services, HDI Cluster, Data Lake Gen2and Data Bricks.
- Good knowledge on Azure synapse Analytics. Proficient in Big data ingestion tools like Kafka, spark streaming and Sqoop for streaming and batch data ingestion.
- Worked with Big Data distributions like Hortonworks (Hortonworks 2.1) with Ambari.
- Hands-on experience in application development using Java, RDBMS, and Linux shell scripting.
- Hands-on experience working with IDE tools such as Eclipse, NetBeans and Maven.

PROFESSIONAL SKILL

Programming language:	Python, Scala
Operating System:	Windows, Linux
Scripting Language:	Shell Script
Hadoop Ecosystem:	HDFS, Sqoop, Hive, HBase
Distribution Package:	Cloudera 5.13
IDE :	PyCharm, Jupyter notebook
Cloud Technologies :	AWS, Azure
Relational Database:	Oracle, MySQL
Tracking Tool:	Jira
Frame Work:	Spark, Spark-SQL, Pyspark
Version Control Tool:	Git & GitHub
SDLC Methodology:	Agile

EDUCATION

BABA INSTITUTE OF TECHNOLOGY AND SCIENCES
Jawaharlal Nehru Technological University, Kakinada

WORK EXPERIENCE

Azure Data Engineer

EDVENTURE SOFTWARE PRIVATE LIMITED	**Dec 2017-july 2021**
ACCENTURE CTH (RLABS Enterprise Services)	**Aug 2021- Dec 2021**
TCS	**June 2022-March2023**

PROJECT#1

Client:	Sanofi
Role:	Azure Data Engineer
Environment:	Azure data factory, Azure data bricks, Azure SQL db, ADLS, Pyspark

Description:

Sanofi, an innovative global healthcare company. We chase the miracles of science to improve people's lives. We provide potentially life-changing treatment options and life-saving vaccine protection to millions of people globally, while putting sustainability and social responsibility at the Centre of our ambitions.

Roles & Responsibilities:

- Developed pipelines in azure data factory to fetch data from different sources and loaded it into Azure synapse analytics.
- Extensively involved in writing transforming data in PySpark and pushing into ADLS.
- Worked on Azure Devops to setup infrastructure and to build the and deploy application.
- Implemented Incremental load strategy for loading on daily basis.
- Handled the releases for every deployment and the unit/integration testing.
- Responsible for creating Linked services, datasets, pipelines in Azure data factory.
- Created stored procedures using TSQL. Responsible for creating copy activity, lookup activity, Metadata activity, Monitor the pipelines are report fix the issues.
- Creating data flows to transforming the data to Azure using Azure data factory.
- Developed complete framework of the project end to end and met customer's deadlines proactively.

Client:	AIA Corporation
Role:	Azure Data Engineer
Environment:	PySpark, Python, Hive, Databricks , ADF, Azure, SQL

Description:

AIA empowers promotional products distributors to be boldly independent without being alone. A partnership with AIA enables distributors to focus on their business and engage the market with the confidence of a well-funded operational engine behind them. AIA provides access to a cloud-based platform that helps distributors do more— like process orders and invoices—in less time, so they can get paid faster.

Roles & Responsibilities:

- Involved in ETL workflow using Azure AD, Data bricks with PySpark as per business requirements, which includes extraction of data from Relation Database and loading it to Azure SQL DB.
- Extensively involved in writing transforming data in PySpark and pushing into ADLS Gen2.
- Storing the resulted data into Azure SQL DB for Power BI and Sport fire Consumption.
- Experienced in data migration from On-Prem to Azure Cloud with Databricks using Spark api.
- Attend daily scrum calls and update the ADO user stories on daily basis.
- Monitor the pipeline jobs and report/fix them if any issues.