Curriculum Vitae

Nivetha Subramaniyan
Mobile: +91 6379653933
Email: nivethasubramaniyan2913@gmail.com

Profile Summary:

- 3 years of professional IT experience and currently working as Data Engineer
- Having 2 years of experience in **AWS Services** like S3, LAMBDA, SSM,SNS, IAM and Cloudwatch
- 2 years of experience of working with **Azure Services** – ADF, Databricks, Azure Synapse Analytics
- 3 years of experience in Data Ingestion using Python
- Having 3+ year of Experience in Data Transformation using **Pyspark**
- Working Knowledge on Hadoop components like Map Reduce, HDFS and Hive
- Having 3 year of Experience in **SQL** (Joins, Window functions like Lead, lag, Rank, Dense rank)
- 3 year of working Knowledge on GIT for migrating the codes to higher environments
- POCS on implementing Spark using **Python** and **Spark SQL** for faster testing and processing of data which is responsible to manage data from different sources (pycharm)
- Excellent end-user Interfacing, Analytical thinking ability and Communication Skills
- Very quick learner with commitment to the team and to the time-bound targets
- Smart worker and result oriented team player inclined towards achieving career goals and enterprising objectives

Technical skills:

Big Data Technologies	Apache Spark, Map Reduce
Languages	Python, SQL
Cloud Technologies	AWS, Ms Azure

Professional Experience:

Company 1 : SPI Global pvt ltd, Trichy (June 2020 to December 2021)
Role : Data Engineer

Company 2 : Tiger Analytics chennai (December 2021 to Present)
Role : Data Engineer

Project #1

Project name : Data Analytics on Sensor API

Role	:	Data Engineer
AWS Services	:	S3, SNS, SSM, IAM, Lambda
Programming Language	:	Python
Tools	:	Pycharm, Excel, Github

Project Description:
This project is meant for extracting the data from API (Client specific) transforming and loading thedata in the form of Parquet on client specific S3 Bucket.

Roles & Responsibilities:

- Good working experience on AWS distributions and efficiently used **Python** and **Pyspark** for Data ingestions and Data Transformations respectively.
- Responsible for extracting the data from API and saving it in S3 (parquet)
- Responsible for monitoring the jobs
- Ability to build deployment on AWS, build scripts (Boto3 & requests) and automatedsolutions using **Python.**
- Implemented AWS solutions using **S3 and AWS lambda**
- Good experience in writing Spark applications using Python.
- Used Spark and **Spark-SQL** to read the parquet data and save the tables on S3
- Implemented **Spark** using Python and utilizing **Data frames** and Spark SQL API for fasterprocessing of data.
- Used Pycharm to process data stored in S3 via BOTO3.
- Responsible for creation & run of ETL Lambda Jobs on AWS Console.

Project #2

Project name	:	POC on Spark data processing.
Role/Project	:	Python/Pyspark Developer
AWS Services	:	S3, SNS, SSM, IAM, Lambda
Technology	:	PySpark, Python
Tools	:	Pycharm, Github

Project Description:
Purpose of the project was to Reading and extracting the Json data from different folder in S3 and merging them in the form of dataframe, then saving the entire data in the form of Parquet file on our local machine

Roles & Responsibilities:

- Reading and extracting the Json data from different folder in S3 with the access key.
- Merging them in the form of dataframe.
- Using Windowing operations like lead, Lag, Rank, Dense Rank.

- Aggregate operations like sum, average, Count.

Project #3

Project name	:	Pepsico Project Simulation in **AZURE**
Azure Services	:	ADF, Databricks, ADLS, Azure Synapse Analytics, Logic app
Database	:	SQL Database
Technology	:	PySpark, Python

Project Description:

Extract data from a source, such as product's customer details. The product includes snacks, cold beverages, etc. It is ingested into the data lake and the data is processed using ADF, Databricks, and a SQL database. Transform the data to Azure Synapse Analytics, then generate the report using PowerBI.

Roles & Responsibilities:

- Created Pipelines in ADF using Linked Services/Datasets/Pipeline/to extract, Transform and load data from different sources like Azure SQL, Blob storage, Azure SQL Data warehouse
- Data Ingestion to one or more Azure Services (Azure Data Lake, Azure SQL, Azure Storage, Azure Data Warehouse)
- Developed spark applications using Pyspark and spark SQL for data extraction, transformation and aggregation from multiple file formats for analyzing & transforming the data
- Responsible for estimating the cluster size, Monitoring and troubleshooting of the spark databricks cluster
- Hands on experience on developing SQL scripts for automation purpose

Academic Summary

COURSE	UNIVERSITY	MARK	PASSOUT
B.E (CSE)	MIET Engineering College	7.8 CGPA	2020
HSC	Periyar Maniyammai Girl's Hr.Sec.School	71.7%	2016
SSLC	Parumala st.geregorious matric hr sec school	92.4 %	2014