

VIVEK KUMAR VERMA

Contact No.: (+91)7008954530

Email: kumarvermavivek870@gmail.com

BIG DATA ANALYST - DEVELOPER

Domain Expertise **AZURE, SPARK, SCALA, PYSPARK, PYTHON, ADF, ADB, AWS, SQL.**

EXPERIENCE SUMMARY:

- Having 2 years of experience in Big Data technology for Batch Jobs development and 5 months of experience in Data Science with overall 2.5 years of IT experience.
- Experienced in Azure, AWS, Python, Spark, Pyspark, Scala, AI/ML algorithms, Chatbot development
- Experience in design, development and implementation in Spark 2.x.
- Developed/Implemented data lake solution on Azure.
- Have hands on experience in data extraction (JSON responses) from various sources using Azure Databricks and implemented Data Lake in ADLS Gen1 and Data Lake in ADLS Gen2.
- Have extensively used Azure Data Factory to orchestrate Databricks jobs and other ADF activities.
- Insightful knowledge of developing and maintaining an application with major and minor enhancements with fixing the real time issues.
- Experience in understanding and analysing client needs/requirements.
- Developer in Spark Development team for Cleansing and Analysing the Huge data to create of Customer specific Report/Dashboard in Azure environment.
- Have hands on experience in creating machine learning model and deep learning model and Implementation of NLP as well.
- Have good Experience on converting the business requirements into technical solutions.

SKILLS:

Technology: SPARK

Primary Skill(s): Python, Spark, Pyspark, Scala, SQL, Azure Services – Data Lake, Data Bricks, Data Factory, Blob Storage, AWS – S3, AWS Glue, AWS Athena

Secondary Skill(s): Machine Learning, NLP, Analytics, Neural Networks

PROJECT EXPERIENCE

Project/Product 1:

NOV 2019 – July 2020 with MINDTREE LIMITED, Software Engineer

Project: Unilever Data Lake (UDL)

Client: UNILEVER [Global Marketing]

Involved in Data Pull and processing in Marketing Domain

Description:

UDL project aimed is to bring Unilever's data from different sources to one place (ADLS) and to analyse that data to get business insights out of it.

Azure platform is used in UDL. Spark is used to process and transform the data so that it can be used to get business insights.

Tools & Technologies: ADF, ADLS, BLOB, Azure Databricks, Spark, Pyspark, Scala, SQL, SSMS, VSTS, GIT, CICD, DevOps

Role Played: Software Engineer

Responsibilities / Contribution

- Design/Development of batch jobs to extract data to Azure Data Lake.
- Developed Spark- Scala Based Application using ADB and ingest data from various sources using ADF.
- Creating a scheduling mechanism to trigger the Azure pipeline to download the report to the ADLS environment.
- Worked in data extraction from various sources using Azure Data bricks and implemented Data Lake in ADLS Gen1 and ADLS Gen2.
- Experience in creating Azure Cluster managing the Nodes and Resource.

Project/Product 2

AUG 2020 - OCT. 2020 with MINDTREE LIMITED, Software Engineer

Project: BMW Onboarding

Client: BMW

Involved in Data Pull and processing in Marketing Domain

Description:

BMW project aimed is to bring BMW data from different sources to one place (S3) and to analyse that data to get business insights out of it.

AWS platform is used in BMW. Python is used to process and transform the data so that it can be used to get business insights.

Tools & Technologies: AWS – S3, AWS Glue, AWS Athena, crawler, Python, SQL.

Role Played: Software Engineer

Responsibilities / Contribution

- Developed data Pull and Processing (ETL) in BMW ONBOARDING for different datasets.
- Understood the functional domain of the application and worked on various tools, technology and application like Python, AWS tools (AWS – S3 Bucket, AWS Glue, AWS Athena, Crawler).
- Worked on Identifying, Understanding and Analysing technical problems upfront, given requirements / change requests / defects & design for different datasets and prepare the required documentation like Logical Data Model (LDM) and Physical Data Model (PDM).
- Zero defects in the code and documentation as participated in requirements discussion sessions and understanding the adequate requirements.

Project/Product 3

Nov 2020 till with MINDTREE LIMITED, Senior Engineer

Project: Unilever Data Lake (UDL)

Client: UNILEVER [CUSTOMER DEVELOPMENT]

Involved in Data Pull and processing in Customer Development

Description

UDL project aimed is to bring Unilever's data from different sources to one place (ADLS) and to analyse that data to get business insights out of it.

Azure platform is used in UDL. Spark is used to process and transform the data so that it can be used to get business insights.

Tools & Technologies: ADF, ADLS, BLOB, Azure Databricks, Power BI, Spark, Scala, Python, SQL, SSMS, VSTS, GIT, CICD, DevOps

Role Played: Senior Engineer

Responsibilities / Contribution

- Design/Development of batch jobs to extract data to Azure Data Lake.
- Developed Spark- Scala Based Application using ADB and ingest data from various sources using ADF.
- Creating a scheduling mechanism to trigger the Azure pipeline to download the report to the ADLS environment.
- Worked in data extraction from various sources using Azure Data bricks and implemented Data Lake in ADLS Gen1 and ADLS Gen2.
- Experience in creating Azure Cluster managing the Nodes and Resource.

Project/Product 4

AIMinds

Description

Train AI Model to identify the user query/search and provide the suggestion based on user Search.

Tools & Technologies: Python, Spark, AI/ML, ReactJS

Role Played: Software Engineer

Responsibilities / Contribution

- Created various AI/ML model using python.
- Created UI to interact with users easily using (React js, Bootstrap, Axios).
- Created API to interact with UI as a backend using (flask).
- Used axios to setup the connection between UI and API

Education:

Degree	Major Subject	Institution
B. Tech (81%)	CSE	CVRCE
12 th (80%)	Science	DAV Public School, CRRC, GAYA
10 th (85.5%)		AL- Momin International School

ACHIEVEMENTS

- Introduction to Data Science in Python - COURSERA Certified
- Machine Learning - COURSERA Certified
- Azure Data Engineer Associate – MICROSOFT Certified
- Outstanding Performer Award in MINDTREE
- Got two A-Team Award for displaying exemplary team spirit in BIG-DATA Project (Client UNILEVER)

PERSONAL DETAILS

- Strength: Quick learner, Self-motivated, Hardworking, Team Worker.