



Jakeer Syed

Data Engineer

Phone: +4915171964230

Address: 52a Nußzeil, Frankfurt Am Main, Frankfurt - 60433, Germany

Email: syedjakeer1992@gmail.com

Highly competent Software Engineer with 5 years of experience which includes 2 years as Data Engineer with extensive work in developing and deploying Big data pipelines using Cloud Computing Services. Understands the complex needs of processing big data and has experience in developing code and modules to address those needs.

CORE QUALIFICATIONS

- Extensive hands-on experience with Hadoop Eco System, ETL tools like Sqoop and Data processing tools like Spark.
- Extensive experience in developing Spark applications using Python programming language.
- Hands-on experience with Spark SQL, Spark Streaming and other messaging platforms like Kafka.
- Hands on experience in UNIX, SQL.
- Extensive knowledge and hands-on experience with Google Cloud Platform and its services like Data Flow, Big Query, Pub/Sub, Data Proc and Big Table.
- Good knowledge on AWS services like S3, RedShift, EC2, EMR
- Good knowledge on agile methodology of Software Development.
- Experience in building CI/CD pipelines using Bamboo and Jenkins.
- Experience in building data pipelines launching cloud services using configuration management tools like Terraform with infrastructure as code
- Extensive experience with ITIL concepts like Incident management, Change management and Problem management.
- An efficient interpersonal communicator with sound analytical problem solving skills and management capabilities.
- Strive to keep the slope of learning curve high and able to quickly adapt new environments and technologies.

SKILLS

Big data Ecosystem: Hadoop, SGOOP, Hive, Pig, Flume, Spark, Kafka.

Languages: SQL, HiveQL, PIG, Python, SparkSQL, Shell Scripting.

Databases: Oracle 10g & Oracle 11g, MySQL, NoSQL (HBASE).

Repositories: Bit Bucket, GitLab.

Tools: Eclipse, Cloudera Manager, IntelliJ, PyCharm, JIRA, Bamboo, Terraform, Jenkins.

Cloud Platform: AWS EC2, S3, Redshift, Google Cloud Compute Engine, Data Flow, Big Query, Pub/Sub, DataProc, Big Table, Data Studio, Data Prep.

File Formats: JSON, Parquet, CSV, Text, Log

EDUCATION

- Bachelor of Engineering, Electrical and Electronics Engineering (2010-2014) – Osmania University, Hyderabad, India with 9 CGPA.
- Intermediate (2008-2010) - MPC from Narayana Junior College, Ongole, Andhra Pradesh, India with 97.3%.
- SSC (2007-2008) – Montessori High School, Ongole, Andhra Pradesh, India with 90.3%.

CERTIFICATIONS AND ACHIEVEMENTS

- **Google certified Cloud Professional Data Engineer**
- **Scrum certified Professional Scrum Master (PSM I)**
- Certified in DevOps and Agile methodologies from HCL Technologies LTD.
- Awarded HCL spot award for Outstanding Performance in consecutive years.

EXPERIENCE

Devoteam GMBH (August 2019-Till date)

Senior Data Engineer

Project Details:

Client: Online retailer

Main motto of this project is to implement a data pipeline using terraform on google cloud platform which reads the real time data from the Pub/Sub topic and launches a data flow job that performs sentiment analysis, writes the output data to Google Big Query which is later visualized using Data Studio dashboards to analyze the performance of the product in the market and actions required to improve the sales of the product.

Key Accomplishments:

- Worked on automation of building a data pipeline using terraform and GCP.
- Hands-on experience with multiple Google Cloud Platform services like Cloud Data Flow, Pub/Sub to consume the streaming data and pre-process the data for sentiment analysis.
- Hands-on experience with Data Prep to perform Data cleansing before pre-processing. Experience with Apache Beam in Python Programming language to develop the data transformation recipe.
- Hands-on experience in performing sentiment analysis using Google Cloud Auto ML services like Natural Language API.
- Stored the analyzed results to Big Query for easy visualization using Google Data Studio.
- Performing SQL operations to visualize the results on the data studio.

Devoteam GMBH (November 2018-July 2019)

Senior Data Engineer

Project Details:

Client: Porsche AG

Modern vehicles have a large number of sensors that help to increase the quality and safety of the vehicle. Porsche has the Smart Mobility analytics platform which receives the data from the IoT sensors. The goal of this project is to build a data lake in lambda architecture such that it handles both real time as well as batch data received from the vehicles. Data collection must be monitored and data stored and transformed so that further analysis of the data or integration of processes is efficiently possible.

Key Accomplishments:

- Developed Spark Streaming applications that read the data from different Kafka topics and post the data to rest endpoints after performing transformations.
- Developed Embedded RESTful Server application which uses spark to transform and post the data to multiple rest endpoints
- Developed Spark Streaming template using default interfaces which reads the data from the Kafka topics and write the data to different sinks like HDFS and Impala.
- Worked on performance testing the spark streaming application with different kinds of load.
- Deployment of code on different environments using CI/CD Bamboo pipeline
- Develop Unit test cases, Swagger UI for REST endpoints.

HCL GMBH, Frankfurt (Germany) (January 2018 - September 2018)

Software Engineer

Project Details:

Role: Big Data Engineer

Client: Deutsche Bank AG Frankfurt

Db-data Analytics is a global analytics platform developed by Deutsche Bank, which offers the capacity for secure storage as well as secure import and export of large amounts of data between the processing systems and also to perform analytics. This platform helps the bank to analyze the data and to take respective actions toward the insights made for multiplying its business in the market.

Key accomplishments:

- Created Sqoop jobs to import the incremental data from traditional RDBMS used by other internal applications and load it into the HDFS and Hive tables to use it for further analysis.
- Worked on Flume to load the streaming log data directly into HDFS.
- Created Hive queries that helped market analysts spot emerging trends by comparing fresh data with historical metrics.
- Implemented query performance tuning techniques in HIVE
- Scheduled Hive, PIG and Sqoop Jobs using OOOIE workflow scheduler.

HCL Technologies LTD, Bangalore (India) (January 2015 - January 2018)

Software Engineer

Project Details:

Role: Service Operations Specialist (Level-3)

Client: Deutsche Bank AG Frankfurt

Subject Matter Expert for Deutsche Bank's platinum payment processing applications. Being the point of contact for Business Operations, involved in handling major Application Outages (P1, P2, P3) as well as coordination and interaction with relevant stakeholders for Production Issue resolution.

Key accomplishments:

- Root Cause Analysis of reported incidents from system logs and exceptions.
- Application Release and Change coordination with L3/Infra /DBA and testing teams.
- Implementation of fixes/patch for application as per the client's/system requirement
- Preparing SQL queries for Payment reconciliation reports for business teams to confirm on the business impact in case of any outages.
- Deploying hot fixes and patches in production environment.
- Testing for the code fixes/failed payments in test environment.
- Troubleshoot and fix the issues with the continuous pipelines.
- Actively Participating in Defect review meetings with Vendor, Testing and Business Operations team on behalf of Application Management.
- Configuration of infrastructure required for different environments using Terraform
- Creating and updating knowledge database (KEDB) articles regularly
- Providing documentation and technical presentations to management.

LANGUAGES

Languages	Read	Write	Speak
English	Yes	Yes	Yes
German	Pursuing A2 level training		