SANDEEP MULLANGI

LEAD DATA ENGINEER

📍 NALLAGANDLA, HYDERABAD

📞 9177216380

✉ Sandeep.mullangi2@gmail.com

PROFILE

A data fanatic enabling data and product speak the same language. I enjoy building scalable applications to support data processing, implementing machine learning models and product analytics.

Over the years I've gained experience working in early to mid-range startup environments, established companies and professional services organization.

EDUCATION

BACHELORS - (2007-2011)
UNIVERSITY NAME – Vignan University, Hyderabad

MASTER - (2014-2015)
UNIVERSITY NAME – California State University, Fullerton

EXPERIENCE

Cognitive Scale, Hyderabad, Telangana   **Sep2018 – present**

Senior Data Engineer.

- Developed Mutual Fund Data pipeline parsing source xml files and storing information in MongoDB using python.
- Used docker and Flask to build microservices which were then deployed to Kubernetes.
- Used Python to do all scripting, data wrangling and data analysis using popular libraries like Pandas, NumPy.
- Developed complex SQL queries to transform the data based on requirements and also after discussions with data scientist team.
- Planned all data activities and associated releases and database designs.
- Worked closely with customer success and support teams to solve various data related issues once the customers have been transitioned from onboarding stage.

Berkshire Hathaway Energy, Las Vegas, Nevada   **Nov 2016 –Aug2018**

Senior Data Engineer.

- Provided various promotions to customers of NVEnergy after calculating eligibility criteria of each customer using Hive, Hbase, R, java. This enabled a more engaging and relevant online experience for the customer.
- Developed application that will be able to handle high volume lookups using columnar databases like HBase using java.
- Developed various customer level KPIs as requested by data scientist team to predict call volume to customer service center thereby provided recommendations to reduce call volume by 40%.
- Maintain the Data lake in Hadoop by building data pipeline using Oracle, SAP-HANA, Sqoop, Hive and PySpark.
- Wrote complex SQL queries in oracle and SAP-HANA and then ingested the data to Big Data Platform for analysis.
- Implemented and automated Various Data Science Models (Logistic Regression, ARIMA, and Weighted Moving Average) after the inputs from Data Scientist team.
- Used python to automate few tasks like pulling data from weather website to big data platform on daily basis, enriching call reason based on text data entered by CSR in the table.
- Developed few Dashboards in Tableau for visualization as per specification.

T-Mobile, Seattle, Washington   **Sep 2015 – Oct 2016**

Data Engineer.

- Built variety of data pipelines like streaming and batch to Data Lake, supporting advanced features like schema evolution, streaming using Hadoop technologies like Hive for performing various Transformations, ActiveMQ as Messaging System and Storm as Processing Engine.
- Written Storm Topology to accept events from ActiveMQ producer and emit to either Hbase/HDFS bolts based on the requirement.
- Created various shell scripts to automate tasks like compression of Hbase tables, Converting source Oracle DDLs to Target Hive DDLs, creation of Hive tables in the Hadoop cluster.
- Created several Pig UDF's like generating Hbase row key based on hashing algorithm, Date Conversion, MD5 Checksum using java.

California State University, Fullerton (June 2014 – August 2015)

Technical Student Assistant.

- Find the inconsistencies in the Donor entity database by writing SQL queries and then cleaning the database by loading the batch file to the Advance tool.
- Create reports on request from various departments by writing SQL queries.
- Microsoft Excel (VLOOKUP, PIVOT TABLE, FORMULAS, MACROS, ETC) was used to analyze the data.
- Load Alumni, employment, address change details to the database received from Admissions and records department

Infosys Technologies Ltd (Oct 2011- Dec 2013)   **Client: CareFirst Insurance**

Systems Engineer.

- Skilled in Analysis, Design and Developing test cases for SQL objects verification and testing.
- Using Application Life Cycle Management software (QC) to manage test cases, defects, QTP scripts for regression tests.
- Creating and updating tables, views, working with triggers, stored procedures.
- Wrote complex SQL queries to test the data.
- Responsible to develop the custom tools as per the client needs.

C E R T I F I C A T I O N S

1. **HDPCD Certified Hadoop Developer**

2. **Microsoft HTML5 Certified**