



Venigalla Sai Kumar

Big Data Engineer

Hadoop & database Developer with 3 years of experience in developing with data models , data mapping, configuring and leveraging the Hadoop ecosystem and databases to glean meaningful insights from structured, semi-structured and unstructured data by following Agile Scrum methodology.



venigallasaikumar@gmail.com



+91-8801447788



Visakhapatnam, India



venigalla-saikumar.web.app/



linkedin.com/in/venigallasaikumar

SKILLS

Hadoop

Hbase

Hive

Sqoop

SQL

SQL DBA

Docker

PostgreSQL

MySQL

MongoDB

D3.js

Pyspark

Python

HTML

PL/SQL

Pig

AWS(RDS,S3,ECS,EC2)

Azure Services

LANGUAGES

English

Full Professional Proficiency

Telugu

Native or Bilingual Proficiency

Hindi

Professional Working Proficiency

INTERESTS

Data Analysis

Exploring new places and interested to meet new people.

WORK EXPERIENCE

Data Engineer

EGS Infotech Pvt Ltd

09/2019 - Present

Vizag, Andhra Pradesh, India

Achievements/Tasks

- Worked for a Health Care project ('EGS Health Care Product' – Provides the detailed information & calculations related to MIPS (Merit-based Incentive Payment System) score. Developed dashboards with interactive data visualization charts using D3.js and PowerBi.
- Created Hadoop instance along with Hbase (NoSQL) and migrated data from RDBMS to Hadoop using pyspark and sqoop. Developed multiple Data Models and written PL/SQL programs according to the client requirements and also for product development.
- Developed an AI machine learning behavior tool using Microsoft azure cognitive services and python. Created Docker images of python Flask API project and pushed into azure docker instance for making it serverless API.
- Built ETL pipelines using Azure Data Factory to extract data from the blob storage and transform load data into RDBMS by keeping triggers at scheduled events.
- Extensive experience in giving POC Demos to multiple Clients.

Big Data Developer

Tata Consultancy Services

06/2018 - 09/2019

Chennai, Tamil Nadu, India

Achievements/Tasks

- Worked in a Digital Merchandising project ('OPTUMERA'- Harnessing the power of Big Data Analytics, and prescriptive algorithms to respond to evolving consumer and market trends - <https://www.tcs.com/optumera-digital-merchandising-suite>) took part in Assortment Optimization and Cluster Analytics modules with hybrid architecture consisting both RDBMS (PostgreSQL) and NoSQL (Hbase).
- Database disaster recovery system and Version control using Liquibase / Flyway and integrating them with Jenkins. Replication and creation of a stable environment in the cloud (GCP and Azure). Performed large scale data loading and processing using Spark. Automated and scheduled recursive jobs via python and shell-script.
- Better querying ability owing to work nature in PostgreSQL, Phoenix and Hive. Skilled in writing PL-SQL, HQL, and Shell-scripts. Worked for retail clients like Office Depot and Petco.
- Created PL/SQL, HP/SQL scripts and tuned query performance. Apart from development designed ETL tasks using Talend and Hive respectively. Installed and maintained the Hadoop ecosystem manually and also via Ambari.

CERTIFICATES

SQL for data science by University of California, Davis (05/2020 - Present)

Microsoft Certified: Azure Fundamentals (04/2021 - Present)

EDUCATION

Bachelor of Technology - Mechanical

GITAM Deemed to be University, Rushikonda, Visakhapatnam, India.

06/2014 - 04/2018

CGPA - 7.58/10

ACHIEVEMENTS

- Built reusables in form of automation and contributed to project.
- An Active GUSAC (GITAM University Science Activity Center) Volunteer & Organizer.