PAVAN GEDDAM

Visakhapatnam, AP 530013 • 913 392 2809 • pavan970535@gmail.com

Professional Summary

4+ Years of Experience in **Python Developer** and **Data Engineer** in Big Data technologies like Hadoop, Spark, Hive. Including the works on GCP cloud, AWS cloud and Image Processing. Interested to learn and find insights on data.

Skills

- Languages   : Python, Java.
- Databases   : MySQL, PostgreSQL.
- Cloud   : GCP, AWS.
- Big Data   : PySpark, Hive, Sqoop.
- Web Frameworks   : Django.
- Front-End   : HTML, CSS, Bootstrap, Javascript.
- Image Processing   : OpenCV, PIL.
- Others   : Git.

Work History

Developer, 07/2021 to Current
Accenture – Hyderabad, Telangana

OSA (Operational Sales Analytics)

Worked on the product zone layer which is responsible for delivering the transformed data to the client which is used for the analytics.

My Contribution:

- Written python scripts for converting the raw data into required format which helps ingestion team to move the data to Big Query**.**
- After having the raw and harmonized data, we will create final data by transforming harmonized and raw data. The final data used in the semantic layer.
- Written the SQL Queries to transform the data**.**
- Written Stored procedures, which will be triggered when there is new data.

Developer, 05/2020 to Current
Prompt Cloud Technologies– Bangalore, Karnataka

Data Analytics on Scraped data

Scraping the data from different websites cleaning the data and used for the analysis as per the client requirement.

My Contribution:

- Scraping the data from different websites and storing it in the files.
- written the **Spark** scripts (**Python**) for cleaning and transforming the data as per the client requirement.
- Storing and extracting the data from **AWS** S3.

Developer, 04/2018 to 03/2020
Hallmark Solutions – Visakhapatnam, Andhra Pradesh

Data Analytics on Insurance data

Developed an application for the insurance data that contains patient details and the billing and pharmacy data.

My Contribution:

- Loaded and extracted the data using **Sqoop** from relational database into **HDFS.**
- Created **Indexes, Views, Bucketed, External, Partition, Managed tables** based on requirements
- Improving the performance and optimization of the existing algorithms in hadoop using **Spark.**
- Created hive schemas using **performance techniques** like **Indexes, partitioning and bucketing.**

OCR (Optical Character Recognition)

Developed the application to classify the documents, recognize and extracting the data.

My Contribution:

- Enhancing the images using the **Image Processing.**
- Written the backend logic to extracting and maipulate data using **Python.**
- Classifying the documents based on its type using the **Machine Learning.**

Education

B.Tech: Computer Science And Programming, **Percentage - 72** ,05/2017

Vignan Institue Of Information Technology – Visakhapatnam
Intermediate: **Percentage - 84**, 05/2013
10th Grade: **Percentage - 86**, 05/2011

Academic Achievments

All India Rank of 2500 in **Gate 2018.**