

Software Engineer (Python) with 3+ years of experience in product development, expertise in building and consuming web services, (RESTful), Python, Pyspark, Pandas for data science. Having excellent analytical, debugging, and problem-solving skills. Able to meet deadlines in a fast-paced environment both as an individual contributor and as a team player.

## TECHNICAL SKILLS

**Python, Pandas, Pyspark, Sql**  
**Django-Rest, Flask-restful**

## AWS SERVICES

**Aws Glue, Aws Lambda, Aws Api Gateway, Aws S3, CodeCommit**

## TOOLS

**VS-code, Pycharm, GitHub, Git, Docker.**

## OPERATING SYSTEM

**Ubuntu, Windows**

## DATABASES

**Postgresql**  
**MySQL**

## EDUCATION

**KGR CET, Hyderabad, TS.**  
**B-TECH | 2014 - 2018**

## CERTIFICATIONS

**AWS Certified Developer - Associate**

## WORK EXPERIENCE

### Fission Computer Labs | Hyderabad, TS

May-2019 – Aug-2021 | Software Engineer | 2.3 years exp

### Accenture | Hyderabad, TS

Sep-2021 – Present | Software Engineer | Application Development Analyst

### MT - Research(serverless) | Python, Aws Api gateway, Aws lambda

MI - Research application helps to store the research papers in cloud s3 using user interface platform.

- Built Api gateways for upload, Get, Download functionalities
- Built Aws lambdas to process the files from one layer to another.

### CP - Data Migration(serverless ETL) | Python, Pyspark, AWS Crowler, Sql, AWS S3, AWS Glue Job, AWS Lambda

Data Migration ETL(Extract Transformation Load) serverless pipeline to process the data between Source system(S3) to Target system(Salesforce).

- Built AWS glue job in pyspark for Transformation Logic.
- Built AWS glue job in pyspark for Normalization Logic.
- Built AWS glue job in pyspark to process data(source system to Target logic).

### Commcare ETL Pipeline | Django, Postgresql, Pyspark, Apache Airflow, Kafka, HDFS

Replace the existing data processing pipeline with a performant pipeline. This involves ingesting a continuous stream of transactional/observational data from Kafka topics into HDFS for further processing. The processing of data is taken care of by Spark, Data is processed, intermediate results are stored in HDFS, and then data is pushed into the staging layer in Postgresql. The data in the staging layer is merged with the final reporting tables in Postgresql.

- API Creation on top of Data Lake and Data Warehouse and added the swagger doc.
- Process the data in HDFS using Pyspark Batch and Streaming scripts.
- Analysis of Input Data that has been produced into Kafka.

### Commcare Health Project | Django, Postgres, Javascript

US Health is a recent project in development by Commcarehq due to the Covid-19 situation, this project is deployed in various US states with features like contact tracing and symptoms checker.

- Developing specific features like geocoder implementation for the geological location of users and designing a periodic data collection interface.
- Developed a bulk case importer to onboard existing cases on to the system.
- Developed overwriting case list/details in bulk.