

Hitesh Kaushik

Software Engineer 2

Address Pune, Maharastra 411013

Phone 851-166-0226

E-mail hiteshkhk0105@gmail.com

Accomplished SDE2 troubleshooting and resolving issues by debugging programs and managing data product releases every quarter . Increase proficiencies through scalable technologies Automating data product quarterly releases using AWS cloud architecture. Offering expertise in Data pipeline creations, AWS migration, Querying trends, Using Python and Java programming coupled with Spark and AWS to create scalable and proficient frameworks.

Skills

Python programming	<div><div></div></div> Advanced
AWS cloud	<div><div></div></div> Upper intermediate
Spark (Pyspark and EMR architecture)	<div><div></div></div> Upper intermediate
SQL integration	<div><div></div></div> Upper intermediate
Java	<div><div></div></div> Intermediate
GIT	<div><div></div></div> Upper intermediate
Data extraction and synchronization	<div><div></div></div> Upper intermediate
Data Analysis	<div><div></div></div> Upper intermediate
Agile Software Development	<div><div></div></div> Upper intermediate
Cloud-based software development	<div><div></div></div> Upper intermediate

Work History

2023-06 - Current

Software Development Engineer 2

Precisely Data, Pune

- Currently working on creating a Central data repository, that is a generalized framework to move around 40 diverse data products into **Snowflake** and **Databricks**. Devising **AWS** step functions to download, convert to relevant formats and upload to warehouses. Exploring **Python** multithreading and to enhance the data loading cadence and minimizing the data pipeline turnaround time.
- Worked on an NLP model to find out duplicates in addresses. Used **REACT** hooks, **Python Flask API** and **Mongo db** to create a framework and UI that takes in account distance between two addresses and uses a decision tree

2021-05 - 2023-05 based model to find if two address are duplicated

Software Engineer 1

Precisely, Pune

- Developed an algorithm to interpolate addresses, scaled it using **Python Dask** and automated end to end interpolation with a two pointer algorithm, developed from scratch, that is generalized to find values in any alphanumeric range. Used **AWS Redshift**. Improved total Addresses count after interpolation by **40%**.
- Extended automation framework to **22 data products**, all with different set of schemas and requirements and **GB's of data** . Automation Framework is now used in data pipelines of product release lifecycle. Improved efficiency by **40%** with database independence, and this framework is used to test any data product.
- Worked on implementing and creating data product (Flood data) which involved Geospatial operations and technologies like (**Spark ,java , GIS, AWS** data pipelines, EMR clusters) over different datasets. Optimized **AWS EMR** cluster improving **CPU utilization by 30%** for each node in each cluster
- Framework coded in **Python, numpy** and **pandas** using **AWS (EC2 and AWS RDS)** resources with **PostgreSQL** as database to run continuous queries over 19M records and store trends into Tableau . Improving efficiency by 40% .Point in polygon (PIP) application using **Spark** and **AWS** architecture.
- Prepared and submitted reports on **Tabeleu** and other documentations to assist development team members, with product releases

2019-11 - 2021-05

Software Development Intern

Precisley, Pune

- Devised generalized unique scalable **Python** framework that is implemented with **SparkSQL** with **pyspark** as programming language to test and analyze huge data in any data product . Results in **tableau** displaying different trends to customers by Sales team
- Collaborated effectively with members of software development team and personnel in other departments.
- Developed processes to automate tasks, resulting in **40%** increase in efficiency. and Turn around time for product QA

Education

2015-05 - 2019-04

Bachelor of Technology (B.Tech): Information and Communication Technology

Dhirubhai Ambani Institute of Information And Communication Technology(DAIICT) - Gandhinagar Gujarat

2013-05 - 2015-05

Secondary (XII): Science

MP & EV School (C)BSE - Visakhapatnam, AP

2003-05 - 2013-04

Primary School (10th): ICSE

DE PAUL SCHOOL(ICSE) - Visakhapatnam