KEERTHI S │ AWS BIG DATA ENGINEER

Chennai, India

Mobile: (+91) 904 779 0431

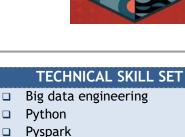**E-Mail**: keerthik889@gmail.com

Skype: Keerthi Savadamuthu

LinkedIn: www.linkedin.com/in/keerthi-savadamuthu-data-engineer

SUMMARY	TECHNICAL SKILL SET
❑ Motivated and highly productive IT professional with 5+ years of experience as a backend developer ❑ Expertise in Big data engineering, Python, Pyspark, Apache Spark, AWS S3, SQL, AWS Redshift, AWS EMR, AWS EC2, OOPS, Hive, Hadoop etc ❑ Proven background in delivering sensible business technology solutions on time while working as a team member ❑ Databricks certified Data Engineer with in depth knowledge of big data frameworks and AWS Cloud ❑ Deft in building data pipeline and ETL batch processing ❑ Adept at building and maintaining effective working relationships with co-workers and clients through outstanding interpersonal skills ❑ Passionate and hardworking with penchant for developing customized interfaces that factor in unique demands for accessibility, researchablity and security	❑ Big data engineering ❑ Python ❑ Pyspark ❑ Apache Spark ❑ AWS S3 ❑ SQL ❑ AWS Redshift ❑ AWS EMR ❑ AWS EC2 ❑ OOPS ❑ Database Management

EXPERIENCE

Organization	Designation	Duration
Altimetrik, India	Engineer-Data Engineer	May 2022 - Ongoing
Infosys Limited, India	Technology Analyst	May 2018 - May 2022

@ Altimetrik, India

Role	Engineer –Senior Data Engineer
Key Deliverables	❑ Compiled and analyzed data, processes, and codes to troubleshoot problems and identify areas for improvement ❑ Collaborated with the front-end developers and other team members to establish objectives and design more functional, cohesive codes to enhance the user experience Developing ideas for new programs, products, or features by monitoring industry developments and trends ❑ Was responsible for writing clean codes to develop the data pipeline ❑ Collaborating with onsite team in gathering business requirements and data modelling

@ Infosys Limited, India

Role	Technology Analyst
Key Deliverables	❑ Built reusable codes and libraries for future use ❑ Liaison with developers, designers and system administrators to identify new features ❑ Recorded data and reported it to appropriate parties, such as clients or leadership ❑ Participated in continuing education and training to remain current on best practices, learning new programming languages, and better assist other team members ❑ Troubleshoot and debug applications ❑ Provided training and support to internal teams

PROJECTS

Verizon - Logging and Metrics

- Performing Automation to fetch the files from S3 logs and split based on config inputs. The config inputs are further copied to the redshift table using psycopg2
- Performing Automation to split the S3 log files to chunk files and store it in Redshift table
- Creation of Audit tables for monitoring and execution of jobs in redshift
- GitHub is used for version control and collaboration
- Actively collaborated in scrum, sprint planning, demos to customers and documentation and reporting
- Devised solutions to improve the job efficiency and the execution time of jobs

Verizon - Optix Billing

- Implemented new approach in display of bills and adjustments that need to be provided
- Logged bugs identified and validated bug fixes
- Assisted in Pattern Analysis for reducing the Tickets count in Machine Learning
- Devising new instances and spinned instances for AMI rehydration in AWS
- Handling of installation and maintenance of kits in DevOps boxes
- Proposed Change Requests and conducted health checks during deployment activity
- Closely coordinated in scrum, sprint planning, demos to customers and documentation and reporting

Novartis – CESAAR

- Building Data Pipelines to ingest data from multiple data sources to oracle database.
- Developing Databricks Notebooks in Pyspark and spark SQL to transform the ingested Data and create generic notebooks for SCD1, SCD2 types for different batch requirements Proposed Change Requests and conducted health checks during deployment activity
- Closely coordinated in scrum, sprint planning, demos to customers and documentation and reporting
- Collaborating with onsite team in gathering business requirements and data modelling
- Working experience in CI CD pipeline to build and deploy the code to test and Production Environment.
- Collaborating with team to create branching strategies in Devops for development and release to higher environments

Ancestry – International Analytics

- Building API classes for the exposure of performance media data programmatically.
- Google SA360 API is built to generate an ads report using doubleclick scope.
- Taboola API is built to call the Taboola Top Campaigns API endpoint.
- Facebook API is built to call the facebook report insights API using accounted endpoint.
- Geneanet sales data is migrated from MySQL to redshift.
- This project was handled individually and data are moved to redshift table for business use.
- Working experience in airflow pipeline and CI CD pipeline.

EDUCATION

- Bachelor of Engineering in Electronics and Communication Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi, India, 2018

TRAININGS/CERTFICATIONS

- Databricks Certified Data Engineer Associate
- Infosys Global Agile Developer Certification
- Infosys certified Python Programmer