AKSHARA KADAM
PYTHON DEVELOPER/DATA ENGINEER
Plano, TX • aks.kdm129@gmail.com • +1 551 273 9478 • LinkedIn

SUMMARY

- A motivated Engineer with a **master's degree** and **6 years** of relevant experience with emphasis in web application, AWS, and ETL.
- Proficient in leveraging AWS cloud services to optimize infrastructure, enhance performance, and drive cost efficiency.
- Expertise in **Python programming, JSON, HTML, Flask framework, PostgreSQL, HTTP, RESTful API services, AWS Lambda, S3, CloudWatch, EC2, and RDS**.
- Proficiency extends to ETL processes and data pipelining, where I leverage advanced technologies like **AWS Glue**, **AWS EMR**, and **Apache Spark** to architect and optimize data workflows. Experience in Data processing using **ETL** tools like **Databricks**, and data visualization using **Power BI**.

TECHNICAL SKILLS

- **Language/Database:** Python, SQL, MySQL
- **Methodologies:** SDLC, Agile/Scrum
- **Frameworks:** Flask, Spark
- **AWS Services:** AWS Lambda, S3, EC2, Step Functions, SQS, SNS, RDS, Glue, EMR
- **Management Tools:** Git
- **CI/CD tool:** Jenkins
- **Data Lakehouse Platform:** Databricks
- **Data Warehousing and BI tools:** Snowflake and Power BI

PROFESSIONAL EXPERIENCE

Metier Enterprise LLC, Plano, TX ***August 2023 – Present***
Software Developer

- Develop the backend to work on a service provider application that handles the user requests for the service, and maintains the request cycle of the service. Use Python modules on Flask Framework by implementing the MVC architecture. Fix bugs, provide production support, and enhance application features like automated notifications for the web application.
- Deploy and manage ETL jobs on AWS Glue to enhance data transformation workflows and increase processing performance. Import diverse datasets in formats like CSV, JSON, and Parquet using Pyspark. Execute data transformations, converting the raw user data into desired consumable data by the business intelligence team for advanced analytics and reporting.
- Create data pipelines using Apache Spark on EMR for streaming data ingested from Kafka, to analyze and transform large amounts of data, leading to a 30% increase in processing efficiency.
- Develop and maintain data processing pipelines using Apache Airflow, automating ETL workflows to process large scale datasets with efficiency and reliability.
- Integrate the application with AWS services using Boto3, including S3, Lambda, RDS, enabling automated cloud resource management and data handling.

Global Infocom Networks LLP, Mumbai ***October 2022 – June 2023***
Python Developer

- Worked on a retail company web application to collect the user click data and perform transformations to validate, catalog, and enrich individual datasets to save to AWS S3.
- Data ingestion: User clicks in the front-end application trigger data to be sent to Kinesis via a Lambda function exposed publicly. The dataset was as large as 20M records therefore made use of AWS EMR cluster to continuously read data from the Kinesis stream using Spark Streaming. Processed the data and then wrote to S3 in Parquet format for efficient storage.
- Data processing: Scheduled EMR clusters to run daily via cron expressions, retrieving data from S3 buckets and processing it using Spark jobs before storing it back in designated S3 folders. A Glue Catalog was created on top of the processed data folders in S3 for easy data discovery and querying. Athena was used to query the processed data stored in S3.
- Job scheduling: Job scheduling for EMR clusters was achieved by using AWS EventBridge to define a schedule. When the event triggered, it invoked a Lambda function which then used the AWS CLI to launch an EMR cluster and submit the job for execution.
- Led the development of Delta Live Table pipelines and orchestrated complex ETL workflows in Databricks, enhancing streaming data processing and reliability for business-critical operations. Leveraged Power BI as a data visualization tool to create interactive and insightful dashboards, integrating seamlessly with existing data processing and ETL pipelines to enhance data-driven decision-making.

Capital One, Plano TX ***May 2021 – September 2022***
Sr. Python Developer

- Worked on a Data reporting and analyzing application that enabled users to extract raw data from Snowflake, combine it with new editable columns, and display a user-friendly table in the form of a report. Designed and developed RETSTful API and services using the 'Requests' library to interact within the microservices and with the front end.
- Used Python scripting and worked on standard Python packages like NumPy, Pandas, SQLAlchemy to support the application. Created PyUnit test cases for unit testing and developed a test framework using Python. Managed code with unit tests and GitHub.
- Used AWS Lambda to automate the reports to query data from Snowflake, S3 to store the lambda zip files, and CloudWatch to log the invocations. Used RDS to store the source and editable data and effectively made use of Cron jobs to refresh the report to ingest updated data.
- Worked on two major feature enhancements like saving custom views of the report and automated notification system which served the users for better analytical purposes.
- Utilized Splunk to monitor application logs, supporting the application by identifying and troubleshooting errors in the backend of the web application.

T. Rowe Price, Baltimore MD ***February 2020 – April 2021***
Python AWS Developer

- Worked on product-to-product data migration for an internal reporting application used for data analysis and financial report purposes.
- Built Lambdas to transform the data and AWS Step functions for workflow orchestration of the Lambdas so that when there was data in SQS it triggered the Step function and carried out the data processing. Made use of AWS S3 for processed data storage and AWS DynamoDB for metadata storage.
- Deployed the AWS services using Unity Deploy platform by creating different strategies and nodes. Configured the environmental variables to run the Lambda with versioning.
- Performed unit testing using PyUnit for the lambda functions to cover positive and negative test case scenarios. Made use of AWS CLI to interact with AWS services in the development environment using the commands in the command-line shell.
- Developed comprehensive Python scripts using Boto3 to automate and manage AWS resources. This includes managing data insertion and querying in DynamoDB.

Xtreme Intelligence, Inc, MD ***October 2019 – January 2020***
Python Developer

- Worked in an Agile environment to develop an e-commerce website hosted on AWS using Python scripting, RESTful services, and store objects in S3 buckets utilizing AWS lambda.
- Used Step functions for workflow orchestration by developing and maintaining state machines using JSON-based ASL enabling clear and structured workflow definitions.
- Responsibilities included but not limited to main Jira, creating new tickets gathering required information from business, and forwarding them in technical teams to the development team. Coordinated with a team of 6 developers offshore and gave them necessary roles and planned sprints till the launch. Participated in code review and buddy testing before moving the code to QA for testing. Ensured all the technical issues were ticketed, responding to Tier 1 tickets and escalating issues at Tier 2.
- Architected and deployed scalable, resilient cloud infrastructure using Terraform, integrating with CI/CD pipelines for automated testing and deployment of Python-based microservices; optimized resource utilization and implemented robust monitoring with custom Python scripts, resulting in a 20% reduction in operational costs and enhanced system reliability.

Global Infocom Networks LLP, Mumbai ***August 2015 – August 2017***
Software Engineer

- Developed and designed Python based API (RESTfulWeb Service) to interact with the web application. Successfully implemented Flask framework to design server applications. Wrote Python code and actively participated in the procedure to automate processes.
- Build and test functionality within a production pipeline. Involved in building database models, APIs, and views utilizing Python to build an interactive web-based solution.
- Designed and managed API system deployment using fast http server and Amazon AWS architecture. Gathered requirements and translated the Business details into technical design.
- Worked with Amazon AWS Cloud Services, (EC2, S3, EBS, ELB, Cloud Watch, Elastic IP, RDS, SNS, SQS, Glacier, IAM, VPC, Cloud Formation, Route53) and managing security groups.
- Implemented Amazon Cloud Services to automate data processing workflows, resulting in a 50% reduction in manual effort and improved data accuracy.

EDUCATION

Master of Science: Telecommunications Systems and Networks | Pace University, New York (2019)

Bachelor of Engineering: Electronics and Telecommunication Engineering | Mumbai University, India (2015)

CERTIFICATIONS

- **AWS Certified Solutions Architect – Associate SAA-CO3 (**T6TGSDX2FNB4Q5GD**)**
- **Databricks Certified Data Engineer – Associate (100383677)**