

Lead Engineer

EXECUTIVE SUMMARY

A Versatile Data Engineer with the ability to adapt to any situation and proven potential to grow self and others, with knowledge and experience in Data agnostic pipelines, analytics, business insight and in developing cloud-native applications based on serverless architecture, AWS services, humongous contributor for achieving customer and team goals. Passionate towards grooming new resources.

AREAS OF EXPERTISE

- AWS Lambda (Serverless)
- Java
- Python
- Serverless Framework
- Spring Boot
- Glue ETL
- AWS EMR
- AWS SNS
- Eventbridge
- Pyspark
- AWS API Gateway
- AWS SES
- AWS Chime
- AWS Step Functions
- AWS Pinpoint
- AWS Comprehend
- AWS Personalize
- Cognito
- Angular
- React
- Ionic
- Flutter
- TypeScript
- AWS Quicksight
- Looker Dashboard
- Firebase
- Athena
- AWS Kinesis Data Analytics
- AWS Kinesis Firehose
- AWS Kinesis Analytics
- AWS Kinesis Data Streams
- AWS Rekognition
- AWS Transcribe
- AWS SageMaker
- AWS RDS
- AWS DynamoDB
- AWS Neptune
- Azure ML Studio
- AWS S3
- AWS AppSync
- AWS CloudFront
- AWS CloudWatch
- Node.js
- Azure Applied AI Services
- AWS Secrets Manager
- AWS CloudFormation

SKILLS OVERVIEW

- Experienced in employing NLP techniques for tasks such as chatbot development, language translation, and content generation.
- Experienced in Event Data Modelling and Data Pipeline tools.
- Experienced in working with Graph Databases and Relational and Non Relational Databases
- Experienced on Data Visualization tools such as Quicksight, lookr
- Strong advocate for implementing cloud services for business problems.
- Experienced in working with SCM/Build Tools.
- Worked in multiple streaming tools for real time data analysis
- Exposure to Data Analytics, Big Data & ML Services

in AWS

- Skilled in AWS services for managing serverless resources
- Experienced in implementing Agile methodology for all projects
- Expertise in working with Serverless architecture on AWS

RELEVANT EXPERIENCE

Apr 2023 Onwards
Lead Engineer
Research and Development

Exploration on Latest AI Technologies

- Established a robust knowledge base using OpenAI, enhancing the chatbot's ability to provide accurate and contextually relevant information to users.
- Integrated AWS Bedrock services to build a resilient infrastructure for chatbots, ensuring scalability and high availability to meet varying user demand
- Implemented Azure Open AI to enhance the intelligence of chatbots, enabling them to understand user queries and deliver precise responses through natural language processing
- Explored Google Vertex AI on the Chatbot implementations
- Conducted knowledge-sharing sessions and training programs to upskill team members.
- Explored and implemented Large Language Models (LLM) such as OpenAI's GPT-3, fine-tuning their capabilities and code synthesis

Mar 2023 to Apr 2023
Senior Data Engineer
Global hospitality and entertainment company

Develop monitoring metrics for AWS services

- Successfully set up PagerDuty for alert notifications, enabling efficient incident management.
- Configured escalation policies and on-call schedules to ensure proper coverage and response.
- Integrated monitoring systems with PagerDuty to trigger alerts for critical events and facilitate timely incident response.
- Proficient in processing SQS queues to monitor and identify any stagnant messages.
- Implemented automated mechanisms to trigger PagerDuty alerts for critical log events.
- Developed automated processes to detect and alert PagerDuty if expected files were missing.

Oct 2021 to Feb 2023
Senior Data Engineer
Design and make technology

Develop ETL pipeline and Personalization

- The user interactive data are obtained from collector and processed for good and bad events.
- The collected events were enriched in enricher ECS services.
- Implemented custom enrichment to add additional values to the existing events
- Created external tables based on the data requirements in hive
- Used Apache Airflow to orchestrate the EMR jobs to load and transform the data in and from external tables.

- Created cloudwatch alarms to monitor the efficiency of events flow
 - Visualised the data based on the requirements in looker dashboard.
 - Use of Feature Groups based on the domain structuring.
 - Used Sagemaker Processing jobs to load the processed data to feature groups
 - Use of lambdas and Appsync to orchestrate the personalization API calls.
 - Use of Neptune to map the logged in user and non logged in user.
-

Jul 2021 to Oct 2021

Data Engineer

Outsourced Customer Support

Develop ETL pipeline and predict reason(attrition)

- Use of Amazon Data Migration Services (DMS) the data are pulled from the client's data warehouse and stored in raw S3
 - Use of Amazon Glue jobs, the csv files are converted to parquet and stored in refined S3
 - Use of SageMaker notebooks to explain the R compatibility in notebooks and performed operations over the refined dataset using R
 - Use of Amazon Quick Sight Dashboard to visualize the data from the refined bucket
 - Created advanced calculations, custom expressions, and data aggregations in QuickSight to derive meaningful metrics for many purposes.
 - Utilized the SPI dataset extensively to extract, transform, and load data.
-

Jun 2021 to Jul 2021

Data Engineer

Transportation Services

Develop ETA solution for Supply chains

- The GPS locations obtained from blackberry sensors were pulled using Amazon ECS.
 - The latitude and longitude were pushed into RDS using Amazon Kinesis Firehose
 - Based on the latitude and longitude of the order the SageMaker multi-model predicts the ETA of that order according to its route.
 - Use of AWS AppSync and Pinpoint to make real-time subscriptions on creation or updating of ETA in RDS.
 - Use of AWS AppSync and Amazon Location Services to show the exact location of the order.
 - Authorization and Authentication is done with the help of AWS Amplify (Cognito)
-

Feb 2021 to May 2021

Application Developer

Media Broadcasting Provider

Develop social marketing solution

- Use of AWS Personalize to find the trending list of videos from given data set and also to find the activity factor of video to calculate the credits
 - Data population into RDS from various external services
 - Use of AWS AppSync to make real-time subscriptions on creation or updating of record in a DynamoDB table.
 - Authorization and Authentication is done with the help of a third-party identity provider
 - Setup CloudWatch events with Lambda and trigger SES notification for sending payment invoice and also to store the invoice pdf in S3 buckets.
 - Purchasing of credits was done through the third party payment gateway.
-

Oct 2020 to Jan 2021

Application Developer

Educational Institution

Develop high-performant video conferencing app

- Integration of Zwibbler whiteboard with AWS Chime.
 - Created an online library as a custom plugin for Moodle LMS, included calendar patches to schedule the recurring meeting.
 - Performed setup of Kinesis Firehose to convert raw data to parquet files and store them in AWS S3
 - Setup CloudWatch events with Lambda and trigger SES notification for sending weekly attendance report.
-

Apr 2020 to Sep 2020

Data Engineer

Media Broadcasting Provider

Develop real-time video analysis model leveraging

- Implemented a microservice architecture using Lambda and API gateway.
 - Orchestrated a series of workflows involving Lambdas using Step Functions
 - Worked on managing user pools with AWS Cognito for human review workflows in A2I and user authentication. Worked with DynamoDB on handling time se
 - Created glue tables and glue crawlers to crawl data source in s3 and update glue table automatically.
 - Connected the glue data catalog to amazon Athena to query the logs using SQL
 - Created QuickSight analysis to visualize chime events in beautiful and informative charts and graphs to identify issues and errors.
 - Implemented drill down functionality in all graph to provide detailed information to the user
 - Implemented checking the quality and correctness of the HLS stream from some of the AWS services below,
 - Rekognition Custom Labels API – To predict the logo for a particular TV channel
 - Comprehend – To predict the TV show from the subtitle text. SageMaker – To create a decision model from audio, video and textual data.
 - A2I – To develop a web driven UI for training ML models. Transcribe – To extract text from HLS audio stream.
 - Developed and optimized SQL Athena queries to perform complex data transformations, aggregations, and joins on large-scale datasets.
 - Developed and maintained complex QuickSight dashboards to track key performance indicators (KPIs) and provide real-time business insights.
-