Singodia A

Senior BigData Engineer, 8+ years experience
Maharashtra Mumbai, India
https://augmntx.com/profile/senior-bigdata-engineer-singodia-a/axo22370

+91 9820045154 | hr@augmntx.com

8+ years Experienced , Motivated, self-taught, and teamwork-oriented software engineer with a can-do attitude. Has significant experience in developing high- volume applications.

INDUSTRIES: Enterprise Software

TECHNICAL SKILLS

• Python	8 years
• ETL	8 years
• AWS S3	8 years
• SQS	8 years
• SNS	8 years
• Pyspark	8 years
• AWS Lambda	8 years
• SQL	8 years
• AWS Glue	7 years
• Hadoop	5 years
• AWS RedShif	5 years
• AWS Step Function	5 years
• AWS Sagemaker	5 years
• SnowFlake	3 years
• Cassandra	3 years
• DynamoDB	3 years
• Kafka	3 years
• Kinesis	1 years

PROJECT DETAILS

• Takana Ikkan ETL December-2020 - September-2021

Project Description: Building and scaling reliable systems, ensuring data is managed properly
Responsibilities: Experience to performance tune and optimize highly scalable distributed applications and enable their vertical/ horizontal scaling. Proven experience in delivering solutions that cover the full lifecycle of application in a qualify and sustainable way, including resiliency, latency, DR and failover. Familiarity with big data technologies such as Apache Hadoop, Apache Spark, and Apache Kafka for processing and analyzing large volumes of data.
Technologies: Python, AWS Glue, AWS Lambda, AWS RDS, AWS GatewayApi,Pyspark,AWS DMS
Industry: Enterprise Software

• Patrika Report Dashboard February-2019 - November-2020

Project Description: Building dashboard to provide key insights, provide reliable & actionable insights
Responsibilities: Perform the software code creation, code compiling and testing, version control and release management activities. Manage production support teams to make sure service levels are maintained and any interruption is resolved in a timely fashion. Ensuring high quality though out the development process good grasp of technology and tools used for development. Developed and maintained ETL pipelines using PySpark and Apache Spark to process large volumes of data, ensuring data quality and accuracy. Implemented real-time data processing solutions using Spark Streaming and Kafka, enabling timely decision-making for business operations. Designed and optimized data models in Cassandra for effcient storage and retrieval of large datasets, ensuring high availability and fault tolerance. Collaborated with cross-functional teams to understand business requirements and translate them into technical specifications for data solutions. Utilized AWS Glue and AWS Lambda to automate data ingestion, transformation, and loading processes, reducing manual efforts and improving e ciency. Conducted performance tuning and optimization of Spark jobs, improving query performance and reducing processing time. Monitored data pipelines and performed troubleshooting to identify and resolve issues, ensuring continuous data flow and availability. Documented technical specifications, data architecture, and ETL processes for reference and knowledge sharing within the team.
Technologies: Python, Pyspark, Spark Streaming, AWS glue, AWS Lambda, Cassandra, Kafka
Industry: Enterprise Software

EMPLOYMENT HISTORY

• Senior BigData Engineer July-2023 - September-2024

Description: ● Conducted quantitative analysis and developed custom reports for clients in the healthcare industry. ● Collaborated with cross-functional terms to develop and deliver client presentations. ● Developed and executed sql queries to extract and manipulate healthcare data from various sources. ● Designed and implemented statistical models to identify patient risk factors and predict advance events. ● Familiarity with big data technologies such as Apache Hadoop, Apache Spark, and Apache Kafka for processing and analyzing large volumes of data. ● Tech used - Python, AWS Glue, AWS Lambda, AWS RDS, AWS GatewayApi,Pyspark,AWS DMS,AWS Sagemaker

• Senior Bigdata Engineer September-2021 - July-2023

Capgemini
Description: ● Data ingestion layer: A connection with data sources such as Actual Digital Sales, NDDC Forecasted Demand, Promo information, Product information, ELC Inventory and other sources. ● Data (pre-processing) layer: Responsible for processing the ingested data and makes sure the structure and quality is usable for model input and business review. ● Model development/Training layer: This layer is connected to the data layer and provides an independent place to train and improve models. It also includes a parameter tuning function to optimize the model’s hyper parameters. Once a model is trained it should be passed on to the back testing layer for validation. ● Tuning/back-testing layer: This layer is responsible for back testing a model and comparing it to another model over time. Once the model is back-tested and the results are as desired, the newly created model can be pushed into (pre)-production. ● Automated ML pipeline/ Inference layer: This layer is used to make the predictions. This layer will use a created model and use production data to make predictions. ● Designed and developed ETL pipelines using AWS Glue for ingesting, cleaning, and transforming data from various sources into Redshift and Snowflake data warehouses. ● Conducted performance tuning and optimization of SQL queries in Redshift and ● Snowflake to enhance data processing e ciency. ● Utilized Python and R for data analysis and statistical modeling, enabling data-driven decision-making for business stakeholders. ● Python, PySpark, AWS Glue, AWS Sagemaker, AWS Lambda, Snowflake, AWS RDS,Spark ML.

EDUCATION

• Bachelors degree in B.Tech(Computer Science) from JNIT, Jaipur 2012 - 2016

LANGUAGES

• English - advanced

END OF DOCUMENT
This PDF is automatically generated by AugmntX on 21 Nov 2026
Profile URL:https://augmntx.com/profile/senior-bigdata-engineer-singodia-a/axo22370