

Manikant Goutam

manikantgautam3@gmail.com • 8982243462 • Block-C, 4th Floor, 403, Geet SkyVille, Nabibagh, Bhopal, MP, India • [linkedin.com/in/Manikant Goutam](https://www.linkedin.com/in/Manikant Goutam)

SUMMARY

Highly motivated and results-oriented Senior Data Engineer with 5+ years of experience in designing, developing, and deploying scalable data pipelines for Big Data processing. Proven expertise in Apache Spark, Airflow, cloud platforms, and various databases. Passionate about leveraging data to drive business insights and optimize processes.

EXPERIENCE

Nielsen Media India Pvt. Ltd, Senior Software Engineer - Data

Jun 2023 - Present

Project: Audience Measurement and Impressions Analytics

Description:

- Maintained and used Scala and Spark for all projects where implemented scalable Data Pipelines with help of Airflow for 2 years here
 - Aggregating the daily campaign viewership into monthly, then weekly then annually.
 - Equipment Airflow, EMR, Spark, and prestoDB for pipeline management and workflows.
 - Was Managing Airflow DAGs for multiple pipelines.
 - Spark Workflow optimization.
 - Automation, Scripting, Product Solution and Architecture, Cloud, Data Processing, Big Data Analysis, Engineering, and Deployment.
 - Developed and maintained data documentation and metadata to ensure data quality and consistency.
 - Collaborated with data scientists and analysts to provide clean and organized data for modeling and analysis.
- Technologies: Python, Scala, Spark, EMR, Athena, S3, Postgres, Airflow, GitLab Pipelines, Jira, Rest APIs.

Autodesk India, Software Development Engineer - Data

Aug 2022 - Jun 2023

Project: Partner Incentives and Analytics

Description:

- Worked within a cross-functional intensive program team to support the final customer.
 - Gathered and analyzed financial data to support decision-making and process improvement.
 - Generated financial reports and presentations for internal and external stakeholders.
 - Collaborated with various teams to ensure the timely delivery of accurate financial data.
 - Assisted in identifying areas for process improvement and implementing changes to increase efficiency.
 - Utilized SQL to perform complex data manipulations, aggregations, and join operations.
 - Collaborated with data scientists and data analysts to provide clean and organized data for modeling and analysis.
 - Developed and maintained data documentation and metadata to ensure data quality and consistency.
- Technologies: Matillion ETL, Python, SQL, Data modeling, Data transformation, and integration, Data quality control, Collaboration, and teamwork.

IDFC First Bank, Data Engineer

Jan 2022 - Aug 2022

Project: Reporting System/Engine

Description: As part of this project, I have worked in a team and we have implemented PySpark Engine, where Spark was hosted in the K8s cluster, and in order to execute Spark jobs this engine fetches submit jobs to the cluster for dynamic jobs. This complete process helped to generate various reports which helped to run bank statements smoothly, still working on other projects.

Technologies: Spark, EMR, Glue, Athena, S3, Cassandra, Postgres, Airflow, Kinesis, ReactJS, NodeJS, Bitbucket Pipelines.

Zee Entertainment Enterprises Limited, Big Data Engineer

Feb 2020 - Jan 2022

Overview: Maintained and developed Scalable Data Pipelines, analytical and Audience Builder Engine system using Spark and Scala for 2 years.

Project: In-house Audience Builder for Zee Platform

Description: Part of in-house Audience Builder for AdTech, and CLM on top of Data Lake. Handled around 400 million active users. The system provides audience creation through segmentation and a rule-based engine (recency and frequency basis).

Fresh data availability to end systems (Ad server and CLM tool) hourly.

Developed Audience Builder Spark Pipeline based on frontend JSON rule and in Scala from scratch.

Technologies: Spark, EMR, Glue, Athena, S3, Cassandra, Postgres, Airflow, Kinesis, ReactJS, NodeJS, Bitbucket Pipelines.

Project: Hipi Analytical Pipeline

Description: AWS EMR provisioning for PySpark Pipeline, extracting user basis on Activity (view, skip, play, etc.) and generating final reports.

Technologies: Python, Spark, EMR, SQS, Airflow, Postgres, AWS Athena.

Project: Near real-time custom targeting implementation for Ad Server (GAM)

Description: Part of this project in which we prepared users' demographic, interest-based, location, and user's audience data availability to Advertisement API and DFP (Google Ad server). It also helps provide real-time pin code-based targeting.

Technologies: Elasticsearch, Spark, EMR, SQS, Airflow, DynamoDB, Postgres, Python, SNS, Redis.

Project: Developed user's content consumption distribution pipeline which I used to share watch consumption of users to their belonging operators like IDEA, VODAFONE, TATA SKY, and AIRTEL.

Technologies: Python, Spark, EMR, SQS, Airflow, Postgres, AWS Athena.

Project: Enrichment project

Description: Set up of systems to enrich user's data, by implementing gender prediction model, interest identification using advertisement logs, age group prediction based on watch history, age, and gender tagging by enabling third-party data and user acquisition data.

Other Tasks:

- Compiled, cleaned, and manipulated data for proper handling.
- Developed, implemented, supported, and maintained data analytic protocols, standards, and documentation.
- Analysed complex data and identified anomalies, trends, and risks to provide useful insights to improve internal controls.
- Designed and developed Data Pipelines and analytical data structures.

*Linuxworld Pvt. Ltd, **DevOps Trainee***

May 2019 - Jul 2019

Project: Automated Basic Configuration Tasks using Ansible by researching core Linux concepts. Integrated with OWN LINUX CLOUD architecture. Providing data nodes with flexible storage services.

Project: Automated Deployment of CLOUD COMPUTING VIRTUALIZATION with CONTAINERISED DOCKER Integration on Linux system using PYTHON.

Project: Developed Software Service, Storage Service, Platform Service, and Container Service automated and deployed by Docker and Ansible.

Other Tasks:

- Promoted a positive public image for the company and worked to control the dissemination of information.
- Maintained an eagerness to add to my knowledge and skills.
- Effectively participated in cross-team strategizing efforts to achieve optimal goals.
- Worked well independently and on a team to solve problems.

EDUCATION

National Institute of Technology (NIT), Calicut

M.C.A.

Jiwaji University

B.SC. (MATHS-CS)

LICENSES & CERTIFICATIONS

RHCSA (190-129-148)

Red-Hat •

RHCoE (Ansible Automation: 192-129-148)

Red-Hat •

SKILLS

Python • Scala • SQL • Shell Script • RedHat Linux • Ansible • Apache Spark • Kafka • Airflow • Pandas