UDIT GUPTA

Data Engineer

@ udgupta155@gmail.com   📞 +91 9981335244 | +91 6260419774   📍 Indore, Madhya Pradesh

in https://www.linkedin.com/in/guptaudit15/   📅 15th May 1994

PySpark Developer and Data engineer with 4.3 years of comprehensive experience in areas like data analytics, data warehousing, designing, implementing, unit testing, data management, analyzing requirements, data operations like Live data streaming and ETL (Extract, Transform, Load). Working Experience includes PySpark, BitBucket, Google Cloud Platform, Azure DevOps, Python and Unix/Shell Scripting.

EXPERIENCE

PySpark Developer & Data Engineer
Cognizant Technology Services

📅 Jan 2022 – Present   📍 Pune, Maharashtra

- Creating Data Pipeline using PySpark and Hadoop.
- Created various automation in the existing projects with the help of the latest technologies.
- Optimized existing process load timing by 70-80 percent. Worked on large-scale data set having 20 billion records.
- Work within the product development team, to an agile methodology.
- Provide technical leadership to coworkers, testers, and other teams for coding/project/technical work.

ETL & Kafka Developer
Tata Consultancy Services

📅 Oct 2018 – Jan 2022   📍 Indore, Madhya Pradesh

- Creating Kafka Components including Producers, Consumers, Topics, and Data Streaming.
- Creating Data Pipeline using Source Connector and Sink Connectors.
- Azure DevOps CI/CD Pipeline creation for Kafka.
- Instance Management and Admin for Kafka.
- Developing End-to-End ETL Pipelines using Informatica and Snaplogic.

EDUCATION

Master of Computer Application
UIT RGPV, Bhopal, MP

📅 2016 – 2018

Bachelor of Science in Computer Science
KCPS College, Ashta (BU Bhopal)

📅 2012 – 2015

XII (Senior Secondary- Science)
Saraswati Vidya Mandir, Ashta (MP Board)

📅 2011 – 2012

X (Secondary)
Saraswati Vidya Mandir, Ashta (MP Board)

📅 2009 – 2010

TECHNICAL SKILLS

PySpark   Python   Pandas   SQL/PLSQL   Confluent Kafka   Azure DevOps   HDFS   Unix

PROJECTS

Domain: BFSI
Role: PySpark Developer

📅 Jan 2022- Present

- Data Pipeline Creation using Spark (PySpark), HDFS, Testing, and Deploy.

Domain: Health Care
Role: Kafka and ETL Developer

📅 May 2019- Jan 2022

- Data Streaming using Confluent Kafka. Create Data Pipelines using Connectors in Kafka. Manage and Monitor GCP Instances.

Domain: EAS
Role: ETL Developer

📅 Oct 2018- Apr 2019

- Data Pipeline Creation using Informatica. Create Stored Procedures using PLSQL.

Academic Projects

- Major Project "Baskets Lounge" (Android Project)
- Implementation of TIC-TAC-TOE game (Python)
- Data Mining Image Classification.

CERTIFICATIONS

🏆 AZ-900 Azure Fundamentals Certification

✳️ Python 101 for Data Science from IBM

✳️ Confluent Fundamentals Accreditation