

Akhilesh Santosh Jadhav

New York | akhilesh.j@nextitjobs.com | +1 (315)760-4039 | [Portfolio](#) | [LinkedIn](#) | [GitHub](#)

PROFESSIONAL SUMMARY

- Seasoned Data Engineer with over three years of experience, excelling in AWS, Azure, and Big Data technologies like Hadoop and Apache, focused on enhancing data warehouse capabilities and optimizing distributed systems for insightful business intelligence.
- Proficient in designing and implementing robust ETL/ELT pipelines, integrating complex data systems with Amazon Redshift, S3, and Azure Data Factory to improve data accessibility and processing efficiency across projects.
- Skilled in deploying and managing data visualization and DevOps tools including Power BI, Tableau, Docker, Kubernetes, and Jenkins, leading to streamlined operations and agile deployment capabilities that support large-scale data environments.

WORK EXPERIENCE

Syracuse University

Data Engineering Intern

Syracuse, NY

Jun 2024 – Present

- Designed predictive model using scikit-learn, achieving an R^2 score of 0.85 by rigorously selecting features and preprocessing data, demonstrating high accuracy for model.
- Spearheaded deployment of the machine learning model on Amazon SageMaker, configuring auto-scaling and optimized instance types to manage over 1,000 real-time inference requests with low latency.
- Integrated a Flask API with SageMaker endpoints on AWS Elastic Beanstalk, enhancing prediction response times by 30% through efficient API management and load balancing, ensuring scalability and robust user interactions.
- Implemented comprehensive security measures and monitoring protocols using AWS IAM and CloudWatch, ensuring adherence to data integrity and protection standards.

PTC

Data Engineer

Boston, MA

Jul 2023 – Jun 2024

- Streamlined AWS Glue and PySpark ETL processes, reducing execution time by 30% and cutting AWS Glue costs by 15% through code refactoring, efficient data transformations, and resource tuning.
- Established a real-time analytics solution using Amazon Kinesis and AWS Lambda, decreasing data availability latency from hours to seconds and enabling immediate insights.
- Introduced Infrastructure as Code with AWS CloudFormation and CI/CD pipelines, reducing environment deployment time from days to under 3 hours and ensuring consistent setups.
- Migrated data to Parquet format and optimized partitioning, achieving a 25% reduction in data storage costs and a 40% improvement in query execution times in Amazon S3 and Redshift by leveraging Redshift Spectrum.
- Directed data security initiatives by managing AWS KMS encryption keys, implementing least privilege IAM policies, and utilizing AWS CloudTrail and GuardDuty for monitoring, ensuring GDPR compliance.
- Mentored team of 4 members and conducted workshops on AWS services and coding best practices, increasing team productivity and enhanced code quality through regular peer reviews.

Sigma Infosolutions Ltd.

Junior Data Engineer

Bangalore, IN

Jun 2020 – Jul 2022

- Executed ETL processes using Azure Data Factory and Hadoop HDInsight, implementing Azure Monitor for real-time performance tracking and reducing pipeline downtime by 15% through issue detection and resolution.
- Devised Python scripts integrated with Apache Hive and Cassandra for automated data validation and cleaning, decreasing data errors by 20% and maintaining data integrity for downstream analytics.
- Collaborated in configuring ETL workflows using Azure Data Factory, Apache Kafka, and Hadoop MapReduce, constructed two new data pipelines and increased data availability by 67%.
- Crafted automation scripts with Python and Azure Automation for file transfers and daily report generation, increasing team efficiency by 25% and minimizing manual workload through error-free automation.
- Architected SQL queries and built dashboards with Apache Hive and Power BI, reducing report generation time by 75%.
- Pioneered data storage strategies with Azure Blob Storage, Apache Hadoop, and Apache Cassandra, improving query times by 30%.

CERTIFICATIONS

- Academy Accreditation – **Databricks Fundamentals** [\[Credentials\]](#)
- Azure Data Engineer Associate (DP-203)**: Design and Implement Data Storage [\[Credentials\]](#)
- Microsoft Technology Associate: Python** [\[Credentials\]](#)

EDUCATION

Syracuse University

Master of Science in Computer Science

Syracuse, NY

Aug 2022 – May 2024

Vishwakarma Institute of Technology

Bachelor of Technology in Computer Engineering

Pune, IN

Jul 2018 – Jun 2022

TECHNICAL SKILLS

Languages:	Python, Scala, SQL, R, Unix Shell Scripting, Bash
Big Data:	Hadoop (HDFS, YARN, MapReduce, Hive, Sqoop, Oozie, HBase), Apache Spark, Kafka, NiFi, Flink, Cassandra, Airflow
Cloud Platforms:	AWS (S3, EC2, EMR, Redshift, Lambda, DynamoDB), Azure (Data Factory, Blob Storage, HDInsight, Databricks), GCP
Databases:	MySQL, PostgreSQL, SQL Server, Oracle, Snowflake, MongoDB, NoSQL, Cassandra, Redis
ETL/ELT Tools:	Informatica PowerCenter, Talend, SSIS, SAS Data Management, Matillion, Azure Data Factory, dbt
Data Visualization:	Power BI, Tableau, Qlik, Looker, SSRS
DevOps and CI/CD:	Docker, Kubernetes, Jenkins, Git
Compliance:	GDPR, CCPA, Data Encryption, Data Masking
ML Tools & dev:	TensorFlow, PyTorch, Scikit-Learn, Keras, MLflow, Flask, Django