

Akhilesh Santosh Jadhav

New York — akhilesh.j@nextitjobs.com — +1 (680)910-9825 — [Portfolio](#) — [LinkedIn](#) — [GitHub](#)

PROFESSIONAL SUMMARY

- Certified Data Engineer with 3+ years of expertise in designing, developing, and maintaining scalable data platforms using cloud services (AWS, Azure, GCP) and big data frameworks like Apache Spark, Kafka, and Hadoop.
- Skilled in architecting end-to-end data pipelines and real-time analytics systems, managing both SQL and NoSQL databases to convert raw data into actionable business insights.
- Proficient in modern DevOps practices with containerization (Docker, Kubernetes) and CI/CD pipelines, and in deploying machine learning models using Python and cloud-based tools to ensure robust, cost-effective data solutions.

WORK EXPERIENCE

Data Engineering Intern

Jun 2024 – Present

Syracuse University

Syracuse, NY

- Optimized a scikit-learn model for Spotify's track popularity—part of the Spotify Popularity Prediction initiative—achieving an R^2 score of 0.85 through careful feature selection and preprocessing.
- Implemented a linear regression solution on Amazon SageMaker, enabling over 1,000 real-time inferences with auto-scaling.
- Developed a Flask API on AWS Elastic Beanstalk for the Prediction API, reducing prediction response times by 30%.

Data Engineer

Jul 2023 – May 2024

PTC

Boston, MA

- Transformed AWS Glue and PySpark ETL processes for the Sales Data Integration project, cutting execution time by 30% and reducing AWS Glue costs by 15%.
- Initiated a real-time streaming solution with Amazon Kinesis and Lambda for the Real-Time Analytics for Sales project, reducing data latency from hours to seconds.
- Streamlined deployments using AWS CloudFormation with CI/CD pipelines for the Infrastructure Automation project, decreasing setup time from days to under 3 hours.
- Devised data migration techniques converting datasets to Parquet format for the Data Warehouse Optimization project, yielding a 25% reduction in storage costs and a 40% improvement in query performance.

Junior Data Engineer

Jun 2020 – Jul 2022

Sigma Infosolutions Ltd.

Bangalore, IN

- Monitored and refined ETL pipelines using Azure Data Factory and HDInsight, reducing downtime by 15%.
- Enhanced data integrity with Python/PySpark validation scripts for the Data Quality Enhancement project, decreasing errors by 20%.
- Broadened data access by launching two new ETL workflows with Kafka and MapReduce as part of the Expanding Data Ingestion initiative, boosting availability by 67%.
- Automated routine tasks with Python and Azure Automation for the Workflow Automation project, increasing team efficiency by 25%.

TECHNICAL SKILLS

Programming Languages: Python (Pandas, NumPy, Scikit-learn), SQL, Scala, Java, R, Bash, C/C++

Big Data & ETL: Apache Spark, Airflow, Kafka, Airflow, DBT, Snowpipe, Hadoop, Iceberg, Trino

Cloud & DevOps: AWS (Redshift, S3, EMR, Glue, Lambda, SageMaker, CloudFormation, CloudWatch, IAM), Azure Data Factory, Google Cloud Platform, Docker, Kubernetes, Jenkins, Git

Databases: Snowflake, MySQL, PostgreSQL, SQL Server, Oracle, MongoDB, NoSQL, Cassandra, Redis, Vector Databases

Analytics & ML: Tableau, Power BI, TensorFlow, Machine Learning, Business Intelligence, Data Analytics, Jupyter, LLMs, LLM-Driven Data Engineering

EDUCATION

Syracuse University

Aug 2022 – May 2024

Master of Science in Computer Science

Vishwakarma Institute of Technology

Jul 2018 – Jun 2022

Bachelor of Technology in Computer Engineering

CERTIFICATIONS

- AWS Certified Data Engineer [\[Credentials\]](#)
- Databricks Fundamentals [\[Credentials\]](#)
- Azure Data Engineering Associate [\[Credentials\]](#)