NIKHIL DUBEY

Aspiring Data Engineer & Data Scientist | SQL | Python | Cloud Computing | Passionate About Big Data, AI & Analytics

Email - nikhil.dubey@avantika.edu.in • https: //linkedin.com/in/nikhildubey1 • Avantika Univeristy India

Summary

Programming Languages: Python, Java, C++, SQL, Scala
Data Engineering: ETL Pipelines, Data Warehousing, Apache Spark, Hadoop, Airflow
Data Science & Analytics: Machine Learning, Deep Learning, Statistics, Data Visualization (Power BI, Tableau, Matplotlib, Seaborn)
Cloud Computing:

- **AWS:** S3, EC2, Lambda, RDS
- **GCP:** BigQuery, Dataflow
- **Azure:** Synapse, Data Factory

Big Data Technologies: Hadoop, Spark, Kafka, Snowflake
Software Development: Object-Oriented Programming (OOP), System Design, REST APIs, Agile Methodologies
DevOps & Tools: Docker, Kubernetes, Git, Jenkins, CI/CD Pipelines

Key Achievements

Optimized Etl Pipeline
Implemented ETL pipeline improving data processing by 30% using Apache Spark.

Automated Data Reporting
Automated data visualization workflow speeding reporting by 40% with Python.

Enhanced Ml Model
Developed machine learning model increasing accuracy by 20% in predictions.

Cost-effective Cloud Solution
Built cloud solution reducing costs by 15% with AWS and GCP integration.

Education

Avantika University India	
Bachelor of Technology - BTech, Computer Science	08/2022 - 08/2026
Central Board of Secondary Education	
12, PCM	07/2021 - 07/2022
Central Board of Secondary Education	
10	07/2019 - 07/2020

Skills

Software Development • Data Science • Analytical Skills • Problem Solving • English

Projects

Movie Revenue Prediction using Machine Learning
Avantika University 2025

Developed a predictive model using regression techniques to estimate box office revenue based on movie metadata.
- Used features like budget, genre, cast, and marketing trends.

AI-Based Course Recommendation System
Avantika University 2025

Designed a recommendation engine for Udemy courses based on user preferences and browsing behavior.
- Implemented collaborative filtering and content-based filtering techniques.