

# ADITI MISHRA

[aditi.smile28@gmail.com](mailto:aditi.smile28@gmail.com) | [aditimishra28.github.io](https://aditimishra28.github.io) | [github.com/aditimishra28](https://github.com/aditimishra28) | [linkedin.com/in/aditimishra-im/](https://linkedin.com/in/aditimishra-im/)

## SUMMARY

Experienced Data Engineer with **2+** years of expertise in **ETL** development, data warehousing, and Business Intelligence. Skilled in **AWS** services and **Big Data** technologies to build and optimize scalable data pipelines. Strong in **SQL**, **Tableau**, and **cloud**-based data solutions, delivering high-performance analytics and business insights.

## WORK EXPERIENCE

### Data Engineer, IBM

June 2022 - August 2024

- Designed and optimized large-scale **ETL pipelines** using **Hadoop**, **Spark**, and **SQL**, improving **data availability** by 30% and reducing **latency** by 25% for advanced analytics.
- Developed and maintained scalable **data processing solutions** integrating structured and unstructured data for **real-time** and **batch workloads** across cloud platforms like AWS (**S3**, **EC2**, **RDS Aurora**, **Lambda**, **VPC**).
- Optimized SQL queries** and **data models**, enhancing query performance and processing billions of records for faster business insights.
- Built **APIs** using **API Gateway** and developed **Managed File Transfers** (SFTP, Connect Direct) to streamline data exchange and processing. Implemented **secrets management** using AWS **Secrets Manager**, securing sensitive data for authentication and API access.

### Lead of AI/ML/Data Science, Google

September 2021 - June 2022

- Led students in building **data pipelines** and **machine learning models** for real-world applications using **Python**, **SQL**, and **Google Cloud**.
- Guided teams in designing and implementing **ETL workflows**, ensuring **efficient data processing** and **integration** for AI-driven projects.
- Data visualization** using **Tableau**, enabling teams to present **insights effectively** and drive data-driven decision-making.
- Collaborated on **end-to-end data projects**, applying **big data technologies** and best practices to enhance **model performance** and **scalability**.

## TECHNICAL SKILLS

**Data Engineering:** Python, JavaScript, HTML, Docker, Kubernetes, Hadoop, Spark, Git

**Data Analysis & Visualization:** Matplotlib, Tableau, R, Excel, Predictive Analytics, Hypothesis Testing, Regression

**Database & Cloud:** MySQL, PostgreSQL, Oracle SQL, BigQuery, Google Cloud, Microsoft Azure, Amazon Web Services(AWS)

**Frameworks:** TensorFlow, Scikit-learn, Pandas, NumPy, Scipy, Matplotlib, Seaborn, Boto3

## CERTIFICATIONS

- Apache Spark: Databricks Certified Associate Developer Apache Spark3 September 2022
- Cloud: 30 Days Google Cloud April 2022
- Machine Learning: IBM Machine Learning with Python December 2022

## PROJECTS

### Exploring Fuel Efficiency (MPG) | Feature Engineering, Regression Analysis

September 2024 - October 2024

- Conducted feature engineering and regression analysis, achieving 90% accuracy and identifying key factors.
- Optimized model via preprocessing, outlier detection, and feature selection for better interpretability.

### KMeans Visualization | Clustering, Python, Matplotlib

July 2024 - August 2024

- Implemented the KMeans clustering algorithm from scratch, partitioning randomly generated data points into specified clusters.
- Created an animated visualization to demonstrate centroid convergence and data separation in clustering.

### Password Manager | JavaScript, Python

May 2024 - June 2024

- Designed a secure, three-tier password management system integrating a Flask-based backend, Chrome extension, and Expo mobile app.
- Implemented encrypted credential storage and seamless authentication, enhancing security and accessibility for users.

## EDUCATION

Clark University, Worcester, United States of America

August 2024 - December 2025

**Masters of Science in Data Analytics**

**CGPA: 3.7/4**

Relevant Courses: Machine Learning, Deep Learning, Natural Language Processing, Big Data Technologies, Cloud computing

Dr. Abdul Kalam Technical University, Uttar Pradesh, India

August 2018 - June 2022

**Bachelor of Technology in Electronics Engineering**

**CGPA: 3.4/4**

Relevant Courses: Digital Signal Processing, Communication Systems, Data Structures & Algorithms, Probability & Statistics