

ADITI MISHRA

aditi.smile28@gmail.com | aditimishra28.github.io | github.com/aditimishra28 | 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. Research Assistant in AI-driven threat detection, enhancing **cybersecurity** with **AI**.

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%.
- Developed real-time and batch data solutions across **AWS (S3, EC2, RDS Aurora, Lambda, VPC)**, integrating structured and unstructured data.
- Optimized SQL query** performance, processing billions of records for faster insights.
- Built **APIs** with **API Gateway** and managed data transfers (**SFTP**, **Connect Direct**), ensuring secure and efficient data exchange.

Research Assistant, AI & Cybersecurity

March 2025 - Present

- Spearheading research on **AI-driven** threat detection and **anomaly detection** to enhance network security.
- Creating **AI models** for proactive defense and optimizing **cybersecurity** strategies.
- Collaborating with Professor to explore transformative **AI applications** in cybersecurity, contributing to groundbreaking solutions in **threat prevention**.

Lead of AI/ML/Data Science, Google

September 2021 - June 2022

- Led teams in building data pipelines and machine learning models using **Python**, **SQL**, and **Google Cloud Platform**.
- Developed efficient ETL workflows and utilized **Tableau** for impactful **data visualization**, enabling data-driven decisions.

TECHNICAL SKILLS

Data Engineering: Python, JavaScript, HTML, Docker, Kubernetes, Hadoop, Spark, Git, Data Architecture, Problem Solving
Data Analysis & Visualization: Matplotlib, Tableau, R, Excel, Predictive Analytics, Hypothesis Testing, Regression, Power BI
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