ADITI MISHRA

aditi.smile28@gmail.com | aditimishra28.github.io | github.com/aditimishra28 | linkedin.com/in/aditimishra-im/

SUMMARY

Results driven **Data Engineer** with experience at **IBM** and **ScaleUp Labs**, specializing in **AI agent** development, **LLM** fine-tuning (Hugging Face), and cloud-based data solutions. Strong background in **Python**, AWS/GCP/Azure, and technical leadership as former AI/ML Lead at **Google Developer Student Clubs**.

WORK EXPERIENCE

Internship, ScaleUp Labs   *May, 2025 - Present*
- Built and fine-tuned **AI agents** using Small Language Models with **Hugging Face's transformers**, focusing on real-world task automation and efficiency.
- Worked with pre-training data and **tokenization** to optimize model performance for startup-specific use cases.
- Leveraged cloud platforms (**AWS, Azure**) to deploy, test, and scale intelligent agents across AI-focused applications.

Data Engineer, IBM   *June 2022 - August 2024*
- Optimized **ETL** pipelines with **Hadoop**, **Spark**, and **SQL**, boosting data availability by 30% and cutting latency by 25%.
- Built real-time and batch data solutions on **AWS** (S3, EC2, RDS, Lambda), integrating diverse sources for better insights.
- Secure data exchange systems with **API** Gateway, **SFTP**, and Connect Direct, and optimized **SQL** queries to process billions of records with improved performance.

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

Programming & Data Engineering: Python, SQL, Docker, Kubernetes, Hadoop, Apache Spark, Git, Data Architecture, ETL Pipelines, Problem Solving
AI & Machine Learning: AI Agent Development, Small & Large Language Models (SLMs & LLMs), Model Pretraining & Fine-Tuning, Transformers, Hugging Face, TensorFlow, PyTorch, Scikit-learn, Boto3
Data Analysis & Visualization: Pandas, NumPy, SciPy, Seaborn, Matplotlib, Tableau, R, Excel, Predictive Analytics, Regression, Hypothesis Testing
Cloud & Databases: Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, BigQuery, MySQL, Oracle SQL

CERTIFICATIONS

- Cloud: 30 Days Google Cloud   *April 2022*
- Apache Spark: Databricks Certified Associate Developer Apache Spark3   *September 2022*
- Machine Learning: IBM Machine Learning with Python   *December 2022*
- AWS Certified Cloud Practitioner   *April 2025*

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.
- 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