

Akash

India | +91-7899678022 | akash.hiremath25@gmail.com | linkedin.com/in/akash-hiremath25/

SUMMARY

Results-driven **Data Analyst / Data Engineer** with hands-on experience in **Python, SQL, ETL pipeline development, AWS cloud fundamentals, and data automation**. Proven ability to design, optimize, and maintain **scalable data pipelines** that transform large, complex datasets into reliable, analytics-ready assets. Experienced in **data ingestion, cleansing, validation, and modelling**, supporting end-to-end analytics and business intelligence workflows. Demonstrated success in **automating data processes and optimizing SQL performance**, reducing manual processing efforts by up to **50%**. Adept at collaborating with cross-functional teams to deliver **data-driven solutions**, dashboards, and insights across operational and business domains.

WORK EXPERIENCE

Zoyinfinity Hospitality Pvt. Ltd. — Data Engineer Trainee

Jan 2025 – Jan 2026

- Designed and maintained **end-to-end data pipelines** using **Python and SQL** to process hospitality domain data, including bookings, customer transactions, and operational metrics.
- Built and optimized **ETL workflows** to ingest data from multiple sources, perform data cleaning, validation, and transformation, and load analytics-ready datasets into relational databases.
- Developed **high-performance SQL queries, views, and aggregations** to support business intelligence and reporting use cases, improving data accessibility and query efficiency.
- Assisted in deploying and managing **AWS cloud infrastructure** and implemented **automation/build scripts** for scheduled batch jobs, ensuring scalable and reliable data processing.
- Ensured **data quality and pipeline reliability** by implementing validation checks, monitoring jobs, documenting data workflows, and collaborating with product, operations, and analytics teams to deliver business-aligned data solutions.

Coders Boutique (Remote) — Data Engineer Intern

Oct 2024 – Dec 2024

- Engineered **WebScrapIn**, a Chrome extension that automated LinkedIn profile extraction, increasing data collection throughput by **40%**.
- Developed **batch processing pipelines** and **live activity logging** using **SQL**, reducing manual intervention by **30%**.
- Initiated development of **ApplyIn**, an AI-driven resume parser and job-application autofill engine leveraging **NLP**; projected to reduce application time by **50%**.
- Utilized **Python, BeautifulSoup, Selenium, SQL, and REST APIs** to streamline ETL workflows and enhance data reliability.

Analyttica Datalab, Bengaluru — Data Analyst Intern

Feb 2024 – Jun 2024

- Improved classification accuracy by **20%+** across multiple projects by designing, training, and fine-tuning ML models including **Logistic Regression, Decision Trees, and DNNs**.
- Built and deployed an **intelligent ticket categorization system**, reducing triage time by **40%** and boosting classification precision by **25%**.
- Performed **EDA** on high-volume datasets (100,000+ rows), uncovering insights that drove **10+ strategic business decisions**.

- Designed interactive **Tableau dashboards**, conducted cohort analysis, and enabled data-driven retention strategies, reducing decision-making cycles by **15%**.
 - Collaborated with cross-functional teams to deliver **data pipelines, feature engineering workflows, and model evaluation reports**.
-

PERSONAL PROJECTS

Scraping & Classification Web App — Streamlit, MySQL, Celery, NLTK

Feb 2024

- Built a full-stack system that automatically extracts, preprocesses, and classifies news content with **93% model accuracy** using **Multinomial Naive Bayes**.
- Reduced data processing time from **5 minutes to 6 seconds** by implementing asynchronous workflows with **Celery, Redis**, and optimized preprocessing pipelines.
- Integrated **modular ETL components**, improving scalability and adaptability for real-time classification tasks.

RAG Model for Multi-Document QA — Python, FAISS, Gemini API

December 2023

- Developed a **Retrieval-Augmented Generation (RAG)** pipeline enabling multi-document semantic search and contextual Q&A using FAISS vector indexing.
 - Enhanced retrieval speed by **30%**, reducing manual document lookup time by **25%**.
 - Implemented **embedding generation, chunking, ranking, and retrieval logic**, making knowledge search faster and more efficient.
-

SKILLS

- **Programming & Querying:** Python, SQL
 - **Data Engineering Concepts:** ETL / ELT Pipelines, Data Ingestion, Data Transformation, Data Modelling, Batch Processing
 - **Databases & Data Warehousing:** PostgreSQL, Snowflake
 - **Big Data & Distributed Processing:** Apache Spark (PySpark), Databricks
 - **Workflow Orchestration & Integration:** Apache Airflow, Azure Data Factory, Apache Kafka
 - **Cloud Platforms:** AWS, Microsoft Azure
 - **Data & Analytics Libraries:** Pandas, NumPy
 - **APIs & Data Integration:** REST API Integration, Data Automation
 - **DevOps & Engineering Tools:** Git, Docker, CI/CD
-

EDUCATION

- **University of Mysore**
Master of Computer Applications (MCA)
Feb 2022 – Sep 2023
 - **Sharnbasva University, Karnataka**
Bachelor of Computer Applications (BCA)
Aug 2018 – Nov 2021
-

CERTIFICATIONS

- **IBM Data Science Certificate**
- **LinkedIn–Microsoft Data Analytics Certificate**