Kartikey Negi

📞 (+91) 7252911252 | ✉ kartikeynegi2002@gmail.com | in LinkedIn | ⚙ GitHub | 🐦 Twitter | ⊕ Portfolio

Experience

Yatra Online Ltd.   *Feb 2025 – Sep 2025*
Data Engineer/Analyst   *Gurugram*

Developed and maintained multiple data systems including hotel ranking, dynamic pricing, and customer analytics platforms. Designed ETL pipelines using Python, Airflow, and SQL for processing large-scale travel data. Collaborated with cross-functional teams to implement data-driven solutions improving business metrics. Implemented CI/CD and containerized services with Docker; instrumented monitoring and alerts for data jobs.

CRR Dashboard
(Python, Streamlit, Pandas, Requests, APIs, CSV)

Developed an ETL pipeline in Python to fetch and normalize hotel pricing data from multiple APIs (EYC, PWC, CIPLA) and benchmark it against competitor travel agencies. Built a secure Streamlit dashboard with LDAP-based login for role-based access and downloadable reports. Automated report generation and CSV exports; added caching to speed up repeated queries.

Hotel Ranking System
(Flask, PostgreSQL, Streamlit, PostGIS)

Architected hotel scoring system across multiple dimensions including location, pricing, and amenities. Integrated geospatial databases with PostgreSQL for precise proximity-based metrics. Built ETL pipeline for data transformation and Streamlit dashboard for analytics visualization. Optimized queries with spatial indexes and materialized views for faster map rendering.

Yatra Travel Chatbot
(FastAPI, Redis, Dialogflow, Streamlit)

Engineered intelligent chatbot handling complex flight and hotel queries using Dialogflow. Implemented backend webhook services with FastAPI and Redis for session storage. Developed Streamlit frontend and configured Dialogflow fulfillment mechanisms. Added rate limiting, logging, and retries to improve reliability of third-party calls.

YSF Email Automation System
(Python, Pandas, SMTP)

Automated partner-specific fare communication with custom filtering capabilities. Integrated multiple data sources with HTML email templates. Implemented scheduling through cron jobs with logging and error handling. Templated emails with Jinja2 and batched SMTP delivery to improve throughput.

Technical Skills

Languages: Python, SQL, Java, C
Frameworks & Libraries: Pandas, NumPy, Streamlit, FastAPI, Flask, Matplotlib, Seaborn, LangChain
Databases: MySQL, PostgreSQL, MongoDB, MS SQL, SQLite, Geospatial DB
Developer Tools: Git, GitHub, Docker, VS Code, Apache Airflow, Kafka, CronTab, Dialogflow, Apache Superset, Postman
Cloud: GCP, AWS (S3, EC2)
Concepts: Data Warehousing, Data Modeling, DBMS, Data Structures & Algorithms, OOP

Projects

TrustTrace Flow [GitHub]   *(FastAPI, DuckDB, ChromaDB, SentenceTransformers, Streamlit)*

Developed an ingestion and traceability pipeline for 25K+ patent records with SHA256 fingerprinting, dual storage in DuckDB, and semantic search using vector embeddings in ChromaDB. Exposed REST APIs with FastAPI and built a Streamlit dashboard for interactive search and API validation. Implemented data cleaning and preprocessing workflows with comprehensive error handling and logging.

BankPulse [GitHub]   *(AWS S3, Python, PostgreSQL, Pandas, Boto3)*

Designed a real-time banking transaction monitoring system simulating 10K+ transactions/sec. Built ingestion and detection mechanisms in Python, leveraged PostgreSQL for state management and deduplication, and implemented custom detection patterns (UPGRADE, CHILD, DEI). Batched results were written back to S3 for downstream analytics with comprehensive monitoring.

CSV EDA Agent [GitHub]   *(Streamlit, LangChain, Google Gemini, Pandas, Python)*

Developed an intelligent CSV data analysis tool using Streamlit and LangChain integration with Google Gemini 1.5 model. Built a user-friendly interface for uploading CSV files and performing natural language queries on datasets. Implemented pandas dataframe agent for intelligent data querying and visualization capabilities. Features include data preview, statistical analysis, and interactive Q&A interface for exploratory data analysis.

Education

Graphic Era Hill University   2025
Master of Computer Application (CGPA: 8.4/10)   *Bhimtal, India*

Graphic Era Hill University   2023
Bachelor of Computer Application (CGPA: 8.2/10)   *Bhimtal, India*