Gourav Katoch
Kangra, Himachal Pradesh, 176094

📞 +918219011 ✉ gouravkatoch016@gmail.com 🔗 www.linkedin.com/in/gourav-katoch ⭘ github.com/gouravkatoch

Projects

Real Time Data Pipeline| *Apache Kafka, Spark Streaming, AWS S3, Tableau* — Jan' 25
- Designed and implemented a real-time data pipeline capable of processing over 50,000 events per second using Apache Kafka and Spark Streaming.
- Integrated AWS S3 for scalable data storage and efficient retrieval.
- Monitored and visualized real-time performance metrics using Tableau dashboards.
- Optimized pipeline throughput and ensured fault-tolerant, distributed data handling.

Hadoop Based Log Analysis System| *Hadoop, MapReduce, Hive, Pig* — Nov' 24
- Built a distributed system for processing large-scale log data, cutting down analysis time by 75% compared to conventional approaches.
- Leveraged Hadoop MapReduce for parallel log processing across nodes.
- Utilized Hive for efficient querying and Pig scripts for streamlined data transformation.
- Enabled faster insights and scalable handling of unstructured log files.

Cloud Data Warehouse Migration| *Snowflake, IBM Watson, Flask, Airflow* — Mar' 24
- Transitioned a legacy SQL Server database to Snowflake, enhancing query speed by 60% and lowering operational costs.
- Employed dbt to manage data transformations.
- Set up Airflow for reliable task scheduling and pipeline orchestration.
- Modernized the data infrastructure to support scalable, cloud-native analytics.

Certificates

Business Intelligence — Feb' 25
IBM — https://courses.lpu.skillsnetwork.site/certificates/0df07d7078154b1585bc846192d37e6c

Big Data 101 — Feb' 25
IBM — https://courses.lpu.skillsnetwork.site/certificates/460d9268b94c45609abe1597257fcfb0

Introduction to MongoDB (For Students) — Jul' 24
MongoDB University — https://learn.mongodb.com/c/wf8Z5JC1TqWTuyDURB51mQ

Technical Skills

Languages C++, Python, JavaScript

Technologies/Frameworks: Power BI, Tableau, Hadoop, Snowflake, SQL,

Skills: Data Structures and Algorithms, Problem-Solving, Attention to Detail, Data Visualization, Interpersonal

Data Engineering and Analytics: ETL Processes, Data pipeline design, real time data streaming

Education

Lovely Professional University Punjab — 2022 – 2026
Computer Science and Engineering — CGPA: 8.09 — *Jalandhar, Punjab*

DHT Kisumu Senior Academy — 2020 – 2022
Pearson International A 'Level — Percentage: 63.05% — *Kisumu, Kenya*

DHT Kisumu Senior Academy — 2016 – 2020
Pearson International GCSE — Percentage: N/A — *Kisumu, Kenya*