

Amogh P

+91 7019627172 | amogh.pwork@gmail.com

linkedin.com/in/amogh-prakash-93b585273 | amoghpportfolio.netlify.app | github.com/Amoghprakash2k4

SUMMARY

Software Engineer with strong foundations in C++, Python, SQL, and distributed systems. Hands-on experience building, benchmarking, and optimizing real-time data pipelines built with Kafka, Spark, and scalable data stores. Skilled in building automation workflows, validating system behavior under load, analyzing failures, and modeling production-like workloads — aligned with core engineering roles in high-performance, distributed database environments.

TECHNICAL SKILLS

Languages: Python, C++, C, SQL.

Distributed Systems & Data: Apache Spark, Kafka, Airflow, Hadoop, AWS.

Databases: MySQL, PostgreSQL, SQLite, MongoDB.

Concepts: Distributed Systems, Stream Processing, Data Modeling, ML Pipelines, ETL, System Design, DSA.

Dev & Collaboration: GitHub, VS Code, Jupyter, Google Colab, Notion, Linux.

SELECTED ENGINEERING PROJECTS

UrbanMind — Distributed Event Pipeline Validation | *Kafka, PySpark, SQL*

- Built a low-latency **Kafka-PySpark** ingestion pipeline, improving real-time civic event availability by **30%**.
- Designed automated checks to validate ETL correctness, reducing data transformation issues by **25%**.
- Analyzed system behavior under burst-heavy workloads and optimized reliability bottlenecks in streaming paths.

Distributed Realtime Event Streaming Engine | *Kafka, PySpark, JavaScript*

- Engineered and load-tested a scalable sentiment-processing pipeline handling high-volume event streams with stable throughput.
- Benchmarked **batch vs. streaming** execution paths, reducing end-to-end latency by **42%**.
- Instrumented dashboards to track system lag, throughput, and model outputs for performance debugging and tuning.

MCP-TerminalAI: Automation Agent | *Python, LangChain, Playwright*

- Developed a CLI automation agent with **testable workflow orchestration** using natural-language commands.
- Integrated APIs (DuckDuckGo, Airbnb) and built automated validation for response correctness and error recovery. Improved execution reliability using targeted prompt-engineering and failure-pattern analysis.

Multi-Pollutant Analytics — Data Integrity & System Validation | *ML, GIS*

- Performed **data integrity testing** across multi-source pollutant datasets (correlations, imputation checks, temporal consistency).
- Built GIS visualizations to validate spatial distribution correctness and hotspot detection across regions.
- Developed a stacked model (MLP + CatBoost + FT-Transformer) reaching **76% accuracy**; validated robustness using cross-validation and error analysis.

STRENGTHS & ACHIEVEMENTS

- Built reliable, high-performance Kafka-Spark-SQL pipelines for real-time data processing.
- Strong in DSA, OOP, and System Design for developing and scaling distributed systems.
- Participated in a national-level hackathon developing an AI storytelling system with real-time content generation.
- Solved 290+ LeetCode problems; earned **HackerRank Problem Solving (Intermediate)** certificate.
- **Captain of district and zonal-level cricket team**, demonstrating leadership, collaboration, and ownership.

EDUCATION

PES University

Nov 2022 – May 2026

B.Tech in Computer Science and Engineering

GPA: 8.02

Siddaganga PU College

Aug 2020 – May 2022

Higher Secondary Education

Percentage: 94%