

Pranesh KVA

Erode, Tamil Nadu, 638002

+91 9384386991 ✉ praneshkva@gmail.com [in linkedin.com/in/praneshkva](https://www.linkedin.com/in/praneshkva) github.com/Pranesh2006

Internships

AI Data Analyst Intern | Pandas, Numpy, SQL, Dashboard, Power BI, AWS Athena, Influx DB

Apr' 2026 - present

Rochester Institute of Technology (RIT)

- Interning at **RIT**, spearheading the **integration of AI-driven analytics** into **RIT's research** workflows via Python to automate **complex data transformations** and reduce manual processing time by **32%**.
- Developing and deploying **predictive statistical models** to analyze large-scale academic datasets, successfully identifying **hidden feature correlations** to improve forecasting accuracy by **18%**.
- Engineering **automated reporting pipelines** and **dynamic dashboards using SQL** and **advanced visualization** techniques, accelerating **stakeholder access** to critical, AI-generated insights by **40%**.
- Optimizing data ingestion and **cleaning protocols** specifically for AI training sets, **systematically eliminating** over **10,000** anomalous records to ensure **high-fidelity data feeds** for downstream applications.

Data Science Intern | NumPy, MLflow, ETL Pipelines, Parallel Processing, Data Partitioning

Jun' 2025 – Aug' 2025

LaunchED Global

- Interned at **LaunchED Global** and engineered **high-fidelity data quality infrastructure** by deploying **automated validation** checks and **monitoring alerts** for **ETL** pipelines, increasing **dataset reliability** by **25%** for downstream model ingestion.
- Optimized **scalable data processing workflows** using **parallel processing** and **strategic partitioning**, achieving a **45% boost** in **throughput** while reducing **cloud compute costs** by **20%**.
- Integrated **automated data monitoring mechanisms** for **early detection** of **pipeline anomalies**, ensuring the delivery of stable and **reliable datasets** for engineering teams.

Projects

Real-time AWS Data Lakehouse | Apache Kafka, Apache Spark, AWS S3, Docker, GitHub Workspaces, Medallion Architecture

Mar' 2026

- Architected a **real-time data ingestion pipeline** using Apache Kafka, processing **1.2M+** daily events with sub-second latency.
- Engineered a **Medallion architecture data lakehouse** (Bronze, Silver, Gold layers) on (Amazon-Web-Services) **AWS S3**, systematically refining **data quality and cutting downstream query** execution time by **30%**.
- Developed automated **Apache Spark ETL workflows**, executed and tested within **GitHub Workspaces**, to seamlessly transform unstructured data and reduce **processing overhead** by **45%**.
- Containerized pipeline components using **Docker** to implement **robust fault-tolerance** and **automated recovery protocols**, achieving **99.9%** availability with zero data loss.
- Repository: [Github](#)

Dynamic Parallel – Parallel Compute Engine | C++, Qt6, Windows API, CMake, std::atomic, multi-threading

Feb' 2026

- Architected a C++ **master-worker engine** orchestrating **12** processes, using **lock-free std::atomic** work-stealing to maintain **100% CPU utilization** without mutex bottlenecks.
- Implemented **hardware-aware core pinning via Windows API** (SetProcessAffinityMask) for AMD Ryzen, eliminating **thread migration** and drastically reducing **L1/L2 cache misses**.
- Integrated a **zero-latency IPC layer** via Windows File Mapping for shared memory, enforcing strict **RAII** compliance for **kernel handles** to guarantee **0** memory leaks.
- Deployed a **real-time Qt6/QtCharts** telemetry dashboard to dynamically visualize cluster throughput velocity and hardware-level performance.
- Repository: [Github](#)

- Engineered a **Retrieval-Augmented Generation** (RAG)-powered sentiment system using API inference to model audience perception, resolving complex textual ambiguity with **92%** precision.
- Adapted **hybrid Natural Language Processing** (NLP) pipelines (VADER, POS tagging) to **handle volatile** Instagram content, dynamically interpreting evolving platform slang, emojis, and sarcasm.
- Architected a **scalable Python backend** integrating **Qdrant vector database** for **high-speed semantic retrieval**, processing **10,000+ real-time interactions** with sub-**200ms** latency.
- Operationalized production-grade brand monitoring workflows, outperforming baseline polarity models to boost granular emotion detection accuracy by **15%**.
- Repository:** [Github](#)

Certifications

Oracle Cloud Infrastructure 2025 Certified Foundation Associate Oracle University - Certificate	Feb'2026
Microsoft Power Bi Proficiency 365DataScience - Certificate	Nov' 2025
Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate Oracle University - Certificate	Oct' 2025
Generative AI: Business Transformation and Career Growth Coursera - Certificate	Sep' 2025
Google Data Analytics Specialization Coursera - Certificate	Sep' 2025
MCP: Build Rich-Context AI Apps with Anthropic DeepLearning.AI - Certificate	Sep' 2025
Career Essentials in Data Analysis Microsoft & LinkedIn - Certificate	Sep' 2025
SQL (Advanced) verified proficiency HackerRank - Certificate	Jul' 2025
Machine Learning Operations for Generative AI Google Cloud - Certificate	Jun' 2025

Skills

Languages: Python, SQL, C/C++, Java
Technologies / Frameworks: Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Regex, NLTK, Qdrant, FastAPI, LangChain, Power BI, MySQL, Apache Spark, Apache Kafka, Docker, Kubernetes, PostgreSQL, AWS S3, AWS EC2
Domain Skills: Data Structures & Algorithms, Distributed Systems, Real-Time Data Processing, Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), System Architecture (MVC), Statistical & Risk Modelling, A/B Testing, Model Context Protocol (MCP), ETL Pipeline Design

Education

Lovely Professional University <i>Bachelor of Technology - Computer Science and Engineering — CGPA: 8.56</i>	Aug' 2023 – present <i>Phagwara, Punjab</i>
Jaycees Matriculation Higher Secondary School, Erode <i>12th — Percentage: 96.7</i>	Jun' 2022 – Mar' 2023 <i>Erode, Tamil Nadu</i>
Jaycees Matriculation Higher Secondary School, Erode <i>10th — Percentage: 100</i>	Jun' 2020 – Mar' 2021 <i>Erode, Tamil Nadu</i>