

Chetan Kumar Tyagi

Delhi, India • +91 9250920926 • iamcktyagi@gmail.com • linkedin.com/in/iamcktyagi • github.com/iamcktyagi

PROFESSIONAL SUMMARY

Data Engineer & Backend Engineer with 5+ years building production-grade **ETL/ELT pipelines**, distributed data systems, and high-reliability backend infrastructure. Specialises in Python, **PySpark (Core / SQL / Streaming)**, **Airflow-orchestrated pipelines**, AWS-deployed systems, and real-time data ingestion at scale. Experienced in performance tuning data pipelines for speed and correctness — including a 65%+ reduction in options data ingestion time through parallelisation and profiling-driven refactoring. Mentors junior engineers on code quality, testing, and pipeline design. Additionally builds LLM-integrated tooling and multi-agent systems, bringing AI engineering capability to data and backend roles.

TECHNICAL SKILLS

Languages: Python (expert), Rust (performance modules), SQL, Bash

Data Engineering: ETL/ELT pipelines, Airflow, Pandas, Polars, NumPy, Parquet, tick/OHLC data, PySpark (Core, SQL, Streaming), Hive, HDFS, Hadoop ecosystem, Snowflake

AI & LLM: Multi-agent orchestration, LLM harnesses, Vertex AI (Gemini), Claude API, OpenAI API, Ollama, MCP (server+client), subagent pipelines

Backend: FastAPI, Django, WebSockets, Socket.IO, REST APIs, AMQP, MQTT (RabbitMQ, Mosquitto)

Databases: MySQL, PostgreSQL, MongoDB, Redis, SQLite, DuckDB

Infrastructure: AWS (EC2, EKS, S3, Lambda, RDS, Athena, Glue, SQS), Docker, Git, CI/CD, Linux

Trading / Domain: Financial data pipelines, broker API integration, options pricing (SABR, Black-Scholes), event-driven systems

PROFESSIONAL EXPERIENCE

Software Development Engineer | [QuantInsti Quantitative Learning Pvt Ltd](#) | Mumbai, Maharashtra *Jul 2024 – Present*

- Built and maintained production-grade data ingestion pipelines for OHLC, tick, and fundamental data across equities, futures, options, and indices — spanning 1-minute to quarterly frequencies on the Blueshift platform.
- Designed ETL/ELT systems parsing data from APIs and flat files, applying transformation, validation, and deduplication — improving dataset reliability and reducing manual correction cycles.
- Implemented sanity-check modules ensuring data integrity across live trading datasets; caught upstream API inconsistencies before they propagated to strategy execution.
- Performance-tuned the options data ingestion pipeline end-to-end — profiled bottlenecks, parallelised I/O-bound stages, and refactored core library internals — reducing ingestion time from ~90 min to 25–30 min (65%+ improvement).
- Orchestrated data workflows via Airflow, cronjobs, and internal schedulers ensuring fault-tolerant, timely updates across environments.
- Built and maintained broker integration libraries standardising APIs between brokers and Blueshift, bridging execution and research layers.
- Implemented SABR and Black-Scholes options pricing models, extending the platform's quantitative analytics capability for derivatives data.
- Developed and executed Spark SQL jobs for large-scale dataset transformation and validation across historical options and equity data.
- Mentored and code-reviewed junior engineers on pipeline design, data-quality practices, and Python best practices; established team-wide testing and debugging protocols.
- Deployed and monitored production systems on AWS (EC2, EKS, S3, CloudWatch) under Ubuntu — maintained uptime and reduced manual intervention.

Senior Software Engineer | [Neam Caps Pvt Ltd \(sharpely / finpeg\)](#) | Thane, Maharashtra *Oct 2023 – Jun 2024*

- Systematically profiled and performance-tuned database queries and legacy systems — applied index optimisation, query restructuring, and caching — achieving 80% faster query results and 50% reduction in resource usage.
- Conducted large-scale backtesting for equity trading strategies using QuantConnect Lean (Python) on local infrastructure; collaborated with market researchers to surface actionable signals.
- Extended live trading engine to support full execution beyond paper trading, including custom margin models for capital efficiency.
- Built data scraping, cleaning, and ingestion pipelines integrated into databases for seamless strategy access.
- Automated ETL processes using Python and shell scripts with cronjob scheduling on EC2.

Software Engineer | [WealthWisers Technologies Pvt Ltd](#) | Noida, Uttar Pradesh *Apr 2021 – Oct 2023*

- Engineered TradeTez — a multi-broker, multi-account trading backend supporting NSE and NYSE markets across brokers (IIFL, SMC, Zerodha, 5Paisa).
- Built end-to-end trading bots covering backtesting, optimisation, and live execution using broker APIs; system handled concurrent multi-account order routing reliably.
- Designed real-time data pipelines and REST APIs with Django/Flask for market data ingestion and routing.
- Developed a publish-subscribe system using WebSockets and Socket.IO for live trade distribution across clients.
- Led ETL development for clients' contract note ingestion using AWS Lambda, RDS, EC2, and Flask APIs.

HIGHLIGHTED PROJECTS

Multi-Agent Trading Research System (Personal) *2024 – Present*

- Built a multi-agent orchestration system using Google Vertex AI (Gemini) and Claude as intelligent orchestrator, with DAG-based task routing and fallback model chains — applied to trading strategy research and signal analysis.
- Designed CLI harnesses supporting Claude, Gemini, OpenAI, and local models via Ollama; built cost-aware model routing with subagent pipelines for parallel workloads.

Automated Scalper (Personal) *2022 – 2024*

- Fully automated trading engine with three modules — Data Adapter, Position Manager, Executioner — using Kotak Neo APIs and AMQP-based messaging for distributed real-time execution.
- Implemented logic-based portfolio manager handling individual and paired trade management across positions.

TradeTez Backend (WealthWisers) *2021 – 2023*

- Multi-broker unified execution backend across NSE/NYSE; asynchronous APIs with high-speed data routing and webhook-triggered order execution backed by Redis state management.

EDUCATION & CERTIFICATIONS

Bachelor of Computer Applications — CCSU, Meerut, UP (2021)

- Deep Learning Specialisation — deeplearning.ai / Coursera
- SQL for Data Science — University of California, Davis / Coursera
- HackerRank 5-Star ratings — Python & SQL
- Cloud Computing (Azure) & Big Data Hadoop with Python — CETPA Infotech