

Mohit Appari

P: +1 8503001952 | mohitt.appari@gmail.com | www.linkedin.com/in/moh1tt | moh1tt.xyz | github.com/moh1tt

SUMMARY

Data Scientist researching systematic trading strategies, grounded in academic factor literature (Jegadeesh-Titman momentum, Fama-French/Carhart factors, Hamilton regime-switching models) and validated through rigorous backtesting. Built a Hidden Markov Model-based market regime detector that outperformed static strategies on a walk-forward out-of-sample test (Sharpe 0.769 vs. 0.641, drawdown -25.6% vs. -33.7%), and a full-stack quantitative research platform combining a multi-factor scoring model with a backtesting engine, regime filter, and live paper-trading execution. Background includes 3+ years of production ML/data engineering (40+ pipelines, 300K+ records, 83% latency reduction).

WORK EXPERIENCE

APD Florida

Tallahassee, FL

Data Scientist

Jun 2025 - Current

- DPerformed forecasting and budgeting analysis on ~\$200M in monthly financial transactions, producing charts, supporting data, and reports used for rate-change decisions and financial planning.
- Conducted research and trend analysis across a 60,000+ client base (state waiver program), tracking client movement between services, service utilization patterns, and data integrity issues, resolving discrepancies to maintain accurate reporting.
- Built and maintained 40+ automated data pipelines processing 300K+ records (Python, SQL Server, Azure Data Warehouse, Docker), ensuring reliability for downstream forecasting and financial analysis.
- Optimized SQL-based analytical workloads through indexing, partitioning, and caching, reducing query latency by 83% (60 min → under 10 min), enabling faster iteration on forecasting models.

Florida State University

Tallahassee, FL

Research Assistant

Jun 2024 - Sep 2024

- Developed and benchmarked deep generative modeling approaches (GANs, VAEs, RNNs) to synthesize large-scale molecular datasets, evaluating architecture tradeoffs for predictive performance on structured, high-dimensional data.
- Benchmarked and fine-tuned deep learning models (NequIP) via transfer learning across an 11,000+ material dataset, improving prediction consistency by 15% through iterative experimentation and hyperparameter optimization.
- Designed and executed controlled experiments across multiple model architectures using statistical hypothesis testing and A/B comparison methods, tracked via MLflow, applying rigorous model validation and out-of-sample evaluation practices directly analogous to backtesting methodology in quantitative research..

SP Global

Bangalore, IN

Software Developer

Jan 2022 - Jun 2023

- Built a production-grade data capture platform processing 10,000+ financial and business documents/day, using OCR and CNN-based classification to convert unstructured filings and scanned documents into structured data feeding S&P Global's downstream financial analytics and research products.
- Engineered cloud-based data pipelines (AWS S3, Athena, BigQuery) to ingest, transform, and serve large-scale datasets, enabling near real-time data availability for downstream financial data consumers.
- Designed a modular, reusable Python extraction framework with CDC-based updates and 80% test coverage, ensuring reliable, scalable data delivery across multiple business domains within a financial data organization.
- Developed REST API integrations with schema validation and CI/CD (GitHub Actions), improving data quality and deployment reliability for downstream analytics teams.

LG

Bangalore, IN

Software Developer Intern

May 2021 - Aug 2021

- Developed a Python-based computer vision and IoT prototype for logistics monitoring using YOLO, REST APIs, PostgreSQL, and Raspberry Pi, implementing real-time tracking and driver drowsiness detection achieving 90% accuracy; secured 3rd place in LG Campus Ideathon.

ITC

Bangalore, IN

Software Developer Intern

Jan 2021 - Mar 2021

- Developed a Master Data Management (MDM) application using Flask, MySQL, and MERN stack with REST APIs, role-based authentication, and automated workflows to improve data management efficiency.

PROJECTS

RegimeSense — Adaptive Market Regime Detection & Strategy Allocation | Python, hmmlearn, Alpaca API

- Built a Gaussian HMM-based regime detector and soft-allocation strategy pool, validated via walk-forward testing (Sharpe 0.769 vs. 0.641, drawdown -25.6% vs. -33.7% for pure momentum), deployed to live paper trading via Alpaca.

Seeking Alpha — Full-Stack Quantitative Research Platform | Python, FastAPI, Next.js, DuckDB, Alpaca API

- Built a factor-based equity scoring model (momentum, technical, fundamental, PEAD/event factors, cross-sectionally normalized) with a backtesting engine, regime filter, and live paper-trading execution across 500+ tickers.

EDUCATION

Florida State University, Masters in Data Science, Tallahassee, FL

Aug 2023 - May 2025

Coursework: Machine Learning, Deep Learning, Statistical Inference, Time Series Analysis, Big Data Analytics, Cloud Computing, Business & Finance

Christ University, Bachelors in Computer Science, Bangalore, IN

Aug 2018 - May 2022

SKILLS

Programming: Python, SQL, R, PyTorch, TensorFlow, scikit-learn, XGBoost, LightGBM, Pandas, NumPy, Statsmodels, MLflow, C++, Scala

Quant: yfinance, FRED (macro data), Finnhub, SEC EDGAR, Alpaca API, Options Fundamentals (Akuna Capital Options 101)

Databases & Cloud: FastAPI, Next.js, DuckDB, PostgreSQL, Snowflake, BigQuery, AWS, GCP, Azure, Docker, Kubernetes, yfinance, Alpaca API