AASTHA SINGH

aasthasingh2329@gmail.com
+91 82182 74722
linkedin.com/in/aastha-singh05 | github.com/Aastha0523

Profile

Data Scientist with 3 years of experience building forecasting and ML solutions end-to-end. I'm curious by nature, committed to continuous learning, and enjoy turning messy data into clear, decision-ready insights. Recently, I have been deepening hands-on work in GenAI (RAG and agentic systems) to bring reliable AI into enterprise workflows.

Professional Experience

Data Scientist — **Accenture** 09/2024 – Present
Bengaluru

Built an **enterprise-grade forecasting platform** for a **leading European industrial technology conglomerate** to predict core business KPIs such as **Orders, Revenues, and EBITA** across **Businesses and management hierarchies**. The system enabled rolling planning on a **€84.1B annual order intake** providing **confidence bands** and **probability-of-hitting-target** insights.

- Built a **Data Preparation & Outlier Treatment module** handling business sign anomalies, ML outlier detection (Isolation Forest/LOF/OCSVM), and imputation.
- Implemented automated **EDA framework** (univariate + multivariate) to analyze seasonality, stationarity, lag effects, and driver relationships; published Streamlit dashboards for self-serve diagnostics.
- Designed **Feature Selection + Drift Monitoring** pipeline combining Pearson/Spearman, Mutual Information, and PPS with optional drift checks.
- Developed **Univariate Forecasting framework** (Holt-Winters, SARIMA, Prophet, Orbit DLT/LGT) to forecast external drivers (GDP/PMI/interest rates) with prediction intervals.
- Architected **Multivariate Modelling module on AzureML** with walk-forward CV, Optuna tuning, and unified model factory (LightGBM/XGBoost/RF/SARIMAX/SVR/Linear/Prophet); improved median MAPE by **18–25%** vs baselines.
- Built a **Post-Modelling validation layer** including residual diagnostics, variance & level sanity checks, STL/Prophet decomposition review, and SHAP explainability.
- Delivered a **Distribution & Probability module** using residual bootstrapping + KDE to generate sigma-bounds and **probability-of-hitting-target** metrics.

Reporting & Data Analyst — **HCL** 02/2023 – 04/2024
Noida

- Developed ARIMA/ETS/Prophet forecasting and Isolation-Forest anomaly detection, cutting inventory variance by **15%**.
- Built SQL ETL from ServiceNow and Snowflake; achieved **3x** runtime improvement via indexing, partitioning, and optimized joins.
- Automated Python dashboards (Matplotlib/Plotly) with MAPE/RMSE/sMAPE validation; reduced reporting time by **60%**.
- Co-designed risk monitoring combining predictive models and rule engines; expanded issue-detection coverage by **25%**.

Recent Learning & Hands-on Projects

- Built a **Finance RAG assistant** using **SharePoint as the knowledge base** (policies, FP&A decks, KPI definitions, FAQ). Implemented chunking, embeddings, and vector search for grounded answers with traceable citations.
- Orchestrated **tool-using agent flows** on top of RAG for multi-step tasks (document retrieval, summary, Q&A, and action routing), including memory and guardrails for safe enterprise usage.

Skills

Machine Learning: Regression & Classification, Time-Series Forecasting (ARIMA/ETS/Prophet/SARIMAX), Tree Models (RF/GBDT/XGBoost/LightGBM), Model Selection & CV, Feature Engineering, Explainability (SHAP)
GenAI: RAG pipelines, embeddings & vector search, agentic workflows, evaluation for groundedness/factuality
Programming: Python , SQL
Data Engineering: SQL optimization (indexing, partitioning, query tuning), ETL/ELT, data quality checks
Cloud & MLOps: AzureML, **GCP** (BigQuery, Cloud Storage, Vertex AI basics), MLflow, Docker, CI/CD

Visualization: Matplotlib, Plotly, Streamlit
Databases & Tools: Snowflake, MySQL **Version Control:** Git, GitHub

Education

B.Tech, Computer Science Engineering, College of Engineering Roorkee 2017 – 2021

Awards

Winner — EY HackPions 4.0 (2022); Zonal Winner — Smart India Hackathon (2020)