

Akash Sharma

AI Engineer

akashsharmaxxiv@gmail.com | linkedin.com/in/akash-sharma-01775b14a | github.com/CosmiX-6 | akashlabs.dev | Mumbai, India

PROFESSIONAL SUMMARY

AI Engineer with 4+ years of end-to-end ownership of production ML systems across forecasting, marketing science, propensity modeling, and data engineering infrastructure. Delivered 26+ production systems on an enterprise SaaS intelligence platform, from early research through operational scale. Combines rigorous statistical foundations with an engineering discipline focused on systems that are accurate, interpretable, and maintainable. Proven ability to translate ambiguous business problems into well-defined ML solutions and communicate results to non-technical stakeholders.

26+

PRODUCTION ML SYSTEMS

~52%

FORECAST ACCURACY GAIN

>99%

COMPUTE TIME REDUCTION

~80%

PIPELINE TIME REDUCTION

EXPERIENCE

AI Engineer

Dec 2024 - Present

Revsure AI

1.5+ yrs, Full Time

Direct employment at the Revenue Intelligence Platform. Continued ownership and advancement of the forecasting and marketing science stack with expanded engineering scope.

- Reduced booking forecast MAPE by ~52% via forecast-category feature engineering on the EOQ prediction model
- Engineered the full-stack Marketing Mix Modeling platform: BayesianRidge on STL residuals, Hill saturation curves (L-BFGS-B), adstock decay, response curve generation, and scenario planning for budget optimization
- Built dual-algorithm explainability: SHAP TreeExplainer for XGBoost paired with Ridge coefficient contribution, written to BigQuery audit tables
- Enhanced configurable multi-model ML framework to support 30+ runtime parameters and a full algorithm selection matrix
- Resolved production-critical bugs in the MMX platform: MinMax extrapolation, scenario cascade drift, and STL edge cases
- Authored customer-facing methodology documentation and model governance artifacts for enterprise stakeholders

Data Scientist

Apr 2022 - Dec 2024

ADA Asia

2 yrs 9 mos, Full Time

Developed ML systems for an enterprise SaaS intelligence platform, growing from individual model contributions to component-level ownership across forecasting, marketing science, and data engineering.

- Built the Revenue Forecasting Platform from a QTD heuristic through a production multi-layer XGBoost EOQ system with SHAP explainability and daily scoring across 4 quarter horizons
- Architected the Pipeline Projection Engine integrating 8+ ML model families into daily current and forward-quarter projections via a two-layer bottom-up and macro-scaling architecture
- Built Markov Chain multi-touch attribution enriched with firmographic and campaign metadata for daily touchpoint-level attribution across funnel stages
- Built account, lead, opportunity, and demand generation propensity models as a four-model scoring suite with daily production scoring and SHAP-based feature attribution
- Migrated revenue metrics pipeline from pandas to distributed PySpark, resolving critical implementation bugs and improving pipeline reliability at scale
- Reduced modeling pipeline runtime from 5-6 hours to 1 hour via joblib parallelization, processing 30+ customer accounts concurrently

SELECTED PROJECTS

Revenue Forecasting Platform

XGBoost, Ridge, STL, SHAP, SARIMAX, BigQuery, GCP Dataproc

Multi-layer EOQ system predicting pipeline and booking outcomes at daily frequency across 4 quarter horizons. Integrates time-decay adjustment, average-index fallback, and dual-algorithm SHAP explainability. Delivered ~52% reduction in booking model MAPE.

Marketing Mix Modeling Platform

BayesianRidge, Hill Curves, L-BFGS-B, STL, PCA, scikit-learn

Full-stack MMX platform with Hill saturation curve fitting (multi-start L-BFGS-B), adstock decay, STL residual decomposition, PCA-based channel attribution, isotonic response curve smoothing, and scenario planning engine for spend reallocation and budget optimization.

Pipeline Projection Engine

XGBoost, scikit-learn, PySpark, BigQuery, GCP Dataproc, Airflow

Central system integrating 8+ ML model families into daily CQ, NQ, and NQ+1 pipeline projections. Two-layer architecture: bottom-up propensity score aggregation combined with top-down macro forecast scaling for cross-layer reconciliation.

Multi-Touch Attribution System

Markov Chain, Python, pandas, scikit-learn, BigQuery

Probabilistic attribution system using Markov Chain removal-effect modeling — customer journeys represented as Markov states, transition matrix built from historical paths, enriched with firmographic and campaign metadata to explain which audience-campaign combinations drive conversion at each funnel stage.

TECHNICAL SKILLS

Languages	Python (expert), SQL (advanced), PySpark / Spark SQL (advanced)
Machine Learning	XGBoost, scikit-learn Pipelines, BayesianRidge, Ridge, LightGBM, CatBoost, Markov Chain
Statistical Modeling	STL Decomposition, SARIMAX, OLS Regression, Welch t-test, Cohen's d, Chi-squared, Difference-in-Differences, Isotonic Regression
Marketing Science	Marketing Mix Modeling, Multi-Touch Attribution, Incrementality Testing, Hill Saturation Curves, Adstock Decay, Scenario Planning
Data Engineering	BigQuery, PySpark, Spark SQL, Parquet, Apache Airflow, ClickHouse, Broadcast Joins, Window Functions
Cloud Infrastructure	Google Cloud Platform, Cloud Storage, Dataproc, Serverless Batch Jobs
Explainability	SHAP TreeExplainer, Ridge Coefficient Attribution, Feature Value Decomposition, Shapley Value JSON Packaging
Libraries and Tools	pandas, NumPy, statsmodels, SciPy, joblib, Splink, Click CLI, BERT Embeddings, Handlebars / Mustache

EDUCATION

Post Graduate Program in Data Analytics Imarticus Learning, Thane, India	Jan 2022
Bachelor of Information Technology T.Z.A.S.P. Pragati College, Dombivli, India	Nov 2020

CERTIFICATIONS

IBM Data Science Professional Certificate, IBM x Coursera (Oct 2021)	Machine Learning with Python, IBM x Coursera (Sep 2021)
Applied Data Science Capstone, IBM x Coursera (Oct 2021)	Data Analysis with Python, IBM x Coursera (Jun 2021)
Data Visualization with Python, IBM x Coursera (Sep 2021)	Databases and SQL for Data Science, IBM x Coursera (Apr 2021)