

CASE STUDY

Leveraged Databricks to Develop a Comprehensive CLTV and Churn Analysis System



OVERVIEW

The client has a comprehensive multi-channel marketing and sales platform designed to elevate customer experiences and marketing efforts. The platform integrates and organizes extensive customer information, providing the necessary tools needed to enhance strategic decision-making and optimize marketing strategies.

CHALLENGES

Despite the platform's great capabilities, the client faced challenges in balancing their investments between acquiring new customers and retaining existing ones. Specifically, they struggled with predicting customer lifetime value (CLTV) and identifying customers at risk of churning. These difficulties led to inefficiencies in resource allocation and missed opportunities to enhance customer engagement. The client aimed to refine their use of the platform to better forecast customer retention and value, ultimately improving their overall marketing effectiveness and customer satisfaction.

INDUSTRY

- Retail
- Sales & Marketing

SERVICES

- CLTV Prediction
- Churn Analysis

TECHNOLOGY

- Azure Databricks Accelerator
- Azure Data Factory
- Delta Lake
- Python
- Blob Storage



SOLUTION

To build a churn prediction system, we used Databricks to integrate the client's dataset with their B2C data via pre-built connectors. This unified dataset, enriched with demographic and behavioral insights, was then prepared using Databricks' automated ETL pipelines. Data cleaning and standardization were handled with Databricks' data quality framework.

We engineered features relevant to churn prediction using Databricks' tools and applied automated feature selection to identify the most impactful variables. The Kaplan-Meier Fitter (KMF) model was trained in Databricks Runtime for Machine Learning, leveraging its scalable resources and optimized libraries.

Deployment was streamlined with Databricks' collaborative notebooks and deployment accelerators. Interactive dashboards, integrated with BI tools, provided customizable churn predictions. Model performance was evaluated using metrics such as the concordance index and calibration plots, with automated cross-validation ensuring robustness.

For LTV prediction, we integrated the client's dataset with their B2C data using Databricks' connectors, enhancing it with additional insights. Data preparation was managed with Databricks' automated ETL pipelines, and features like recency, frequency, monetary value, and customer lifetime duration were engineered.

The BetaGeo and GammaGamma models were trained in Databricks Runtime for Machine Learning, benefiting from its scalable compute resources. Deployment followed a similar process as for the churn model, using Databricks' collaborative notebooks and deployment accelerators.

Interactive dashboards allowed the client to filter and view LTV predictions, supporting targeted marketing and strategic decisions. Model evaluation included advanced metrics and automated cross-validation to ensure accurate forecasting. This integrated solution provided a comprehensive view of customer behavior and enhanced decision-making capabilities.



BENEFITS

- **Optimized Resource Allocation** - Balanced spending between customer acquisition and retention based on predictive insights.
- **Enhanced Customer Retention** - Identified at-risk customers to enable targeted retention efforts and improve retention rates.
- **Improved Long-Term Value Understanding** - Provided clear LTV forecasts to prioritize high-value customers and refine marketing strategies.
- **Efficient Resource Utilization** - Prevented wasted resources by focusing on high-risk customers and optimizing marketing spend.

About Contata

Contata Solutions is a trusted leader in technology and digital innovation. Through our work in data engineering, data analytics, machine learning, marketing automation and app development, we deliver solutions that address complex problems in ways that are simple, insightful and impactful.

Our promise and value proposition to our customers is simple: we leverage our deep technical expertise and global presence to bring software products and data-driven decision capabilities to life.

Founded in 2000, Contata is a privately-held company headquartered in Minneapolis that serves clients globally from offices in the United States and India.