Customer retention forecasting model

Nutritional supplements provider

Partnered with client to build a multi-variate timeseries forecasting model to enhance the accuracy of retention rates forecasts

Nutritional supplements provider needs customer retention forecasting model

Picture this…

You're looking to analyze current and historical data to forecast retention rate using advanced analytics model and enhance the accuracy of the forecasts. You were using 8-week rolling averages method to forecast customer retention rates which resulted in big variances inaccuracies in the forecast

You turn to Accordion.

We partner with your team to build a multi-variate timeseries forecasting model to enhance the accuracy of retention rates forecasts, including:

1) Evaluating the impact of multiple attributes such as marketing spend, discount etc., controlled by the company and external factors sch as covid cases, federal interest rates etc.
2) Comparing customer retention rate forecasts from multiple models such as FB-Prophet, Neural-Prophet, 8-week rolling average etc. to identify the most accurate model for optimizing retention forecast
3) Customizing the model to improve forecasting agility by providing flexibility to update input parameters for each of the shortlisted variables.
4) Deploying the model and scheduling the code to run on a weekly cadence to extract data and store the results in the data warehouse

Your value is enhanced.

- Improved the forecast accuracy (MAPE) by 10 percentage points in comparison to the current 8-week rolling average method during the testing period
- Integrated the model output to existing reports and reduce the efforts to create financial projections, business health and marketing reports

CUSTOMER RETENTION FORECASTING MODEL

KEY RESULT

- Improved the forecast accuracy by 10 percentage points
- Current 8-week rolling average method during the testing period

VALUE LEVERS PULLED

- Anaconda (Python)
- Azure
- Databricks
- Notebook
- Azure Data Factory
- SSMS

Customer retention forecasting

CUSTOMER RETENTION FORECASTING MODEL

Situation

- Client was using 8-week rolling averages method to forecast customer retention rates which resulted in big variances inaccuracies in the forecast
- Partnered with the client to analyze current and historical data to forecast retention rate using advanced analytics model and enhance the accuracy of the forecasts

Accordion Value Add

- Evaluated the impact of multiple attributes such as marketing spend, discount etc., controlled by the company and external factors sch as covid cases, federal interest rates etc.
- Compared customer retention rate forecasts from multiple models such as FB-Prophet, Neural-Prophet, 8-week rolling average etc. to identify the most accurate model for optimizing retention forecast
- Customized the model to improve forecasting agility by providing flexibility to update input parameters for each of the shortlisted variables.
- Deployed the model and scheduled the code to run on a weekly cadence to extract data and store the results in the data warehouse

Impact

- Improved the forecast accuracy (MAPE) by 10 percentage points in comparison to the current 8-week rolling average method during the testing period
- Integrated the model output to existing reports and reduce the efforts to create financial projections, business health and marketing reports

Methodology / Approach

01

Data Acquisition – Transformed the transactional level dataset to weekly acquisition cohort level with period wise retention rates. Utilized additional factors in the analysis that could potentially impact the retention rates like marketing spend, discount etc., controlled by the company and external factors like covid cases, federal interest rates etc.

02

Feature engineering (Shortlisting of appropriate variables) – Correlate the above variables with customer churn to ensure appropriate variables are selected and create combined variables to ensure appropriate features for modelling. dataset was cleaned from extreme outliers occurring due to promotional campaigns that could not be quantified

03

Model selection – Tested various statistical and ML based forecasting techniques such as Weighted Moving Average, ARIMA, SARIMA, XGBoost, Fb-Prophet and Neural-Prophet. Multiple iteration of these models were run with various combination of input variables and selected Fb-Prophet and Neural-Prophet methods as the final models had the lowest MAPE (mean absolute percentage error)

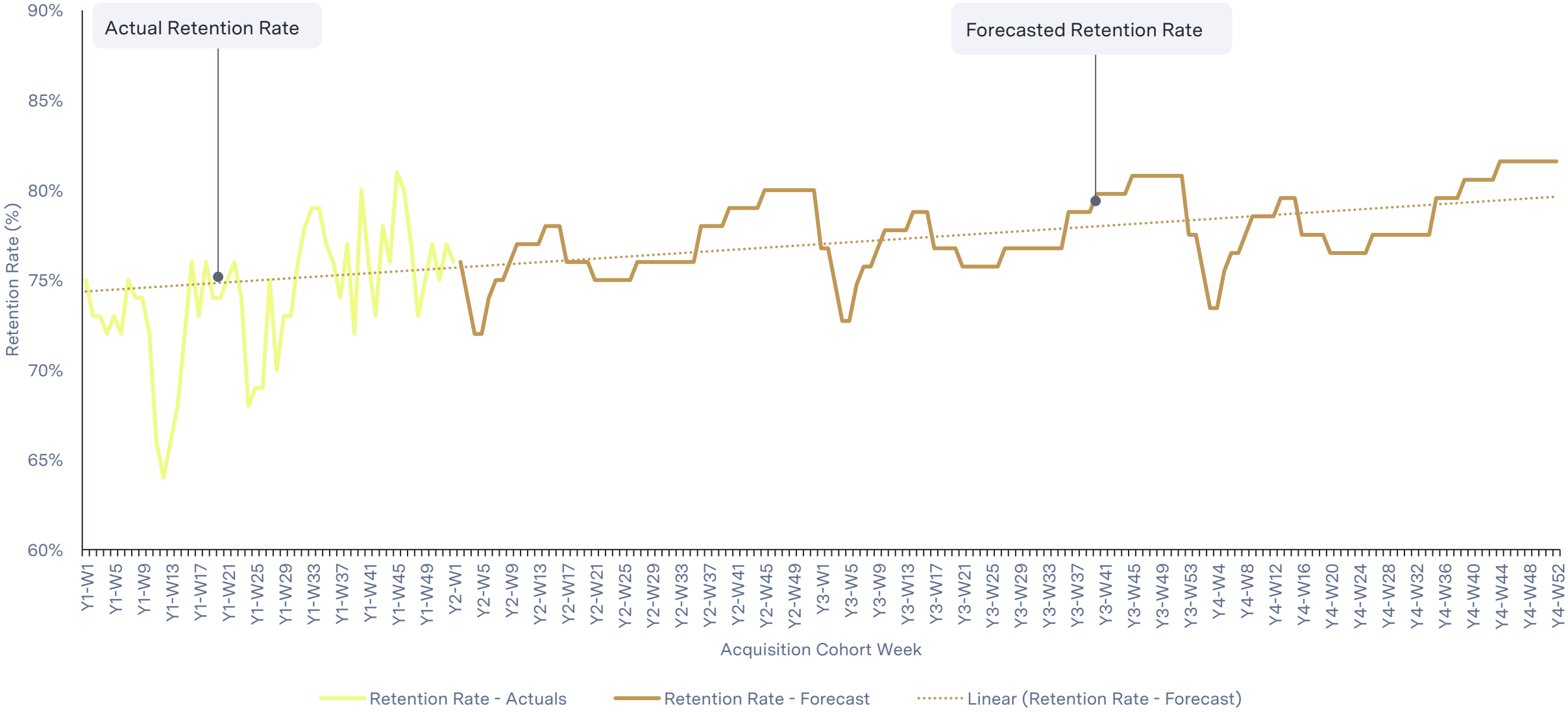
04

Model iteration – Most recent data is withheld from training and used to evaluate models, and their parameters based on the forecast accuracy. MAPE (mean absolute percentage error) is used for evaluation the model with the lowest MAPE is selected.

05

Model deployment – Deployed all the required bash code in Azure Databricks notebook and scheduled the code to run on a weekly cadence using Azure Data Factory. The Databricks notebook interacts with the database to extract data and store the results in a DWH table.

Model forecast



Deployment architecture in Azure ecosystem