Data Science Models

09/03/2025 1:11 pm EDT

Bluecore's data science models use AI and machine learning models to optimize your campaigns. The models help target specific customers, recommend the best products to them, and learn to do both at the right time.

Data models like Predictive Audiences group your customers to optimize their messages. Product Recommendation Models use catalog data to match and recommend the best products to customers. Email Send Time Optimization adjusts each campaign's send time to send the message to every individual at their best time.

Training and Setup

All Bluecore intelligence models and properties require historical behavioral and engagement data for training. The required backfill data is collected as part of the onboarding process, separate data does not need to be provided for training.

For more, see setting up data ingestion.

Frequency

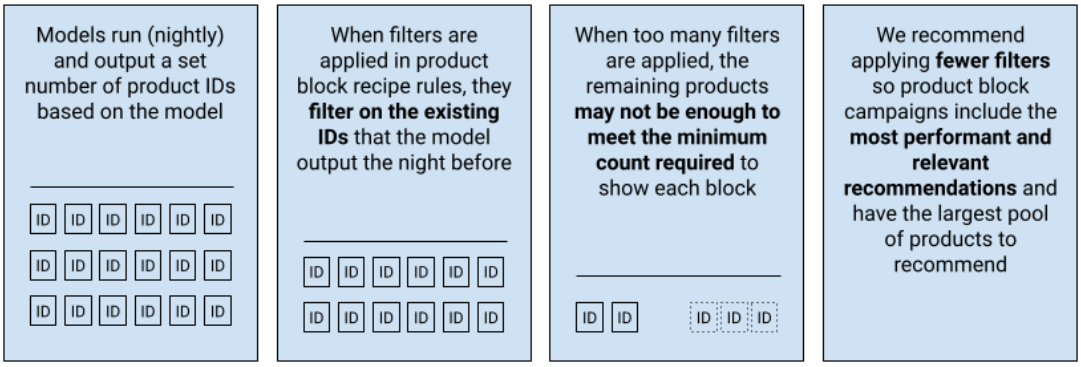
Most Bluecore product and customer models are calculated nightly; models are never updated more than once a day, regardless of usage.

For a subset of Bluecore models, the platform automatically stops calculating scores for that specific model if it hasn't been used within its activity window (outlined for the relevant models below). In these cases, the models will be updated once the data is used again or when the data is considered stale (whichever occurs first).

Filtering applied to model outputs

Once calculated, the recommendations are made available for use in Dynamic Product Blocks for personalized messaging and in Audience Builder for segmentation.

Anytime filters are applied to the models, the models do not produce another set of recommendations; the filters are applied on top of the existing most recent recommendations. This means that the more filters are applied to recommendations, the smaller the pool of products available to include and the less accurate the recommendations may be.



Lifecycle Stages model

The Lifecycle Stages model groups customers based on which of four buying cycle stages they fall into:

1. Non-Buyers

2. Active Buyers

3. At-Risk Buyers

4. Lost Buyers

Three of the predictive stages are for people who purchased something in the past - at risk, lost, and active). Non-buyers is a rule-based stage which corresponds to people who have never made a purchase.

If a Customer Profile is an email address and no history of purchases for that email address, then it is considered as a non-buyer.

For buyers, the model learns the portfolio of buying cadences for a given client by fitting probability distributions on the history of purchases across all customers. Once the model is built, when looking at an individual user, it estimates where they fall in that portfolio and given their own history of purchases, it computes a probability that the user will make a new purchase in the future.

Active Buyers have the highest probability, followed by At-Risk Buyers, then Lost Buyers. Each customer falls into exactly one bucket.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	2 years minimum, leverages up to 4 years

Predicted Customer Lifetime Value Model

The Predicted Customer Lifetime Value (PCLV) is an estimation of how much a customer will spend in the future.

Like the lifecycle stages computations, it follows a similar learning pattern which looks at the different spectrum of customers, computing a portfolio of buying patterns, and comparing each individual customer to this portfolio

to learn where they fall.

The PCLV breaks down into two values - the predicted number of future orders and the predicted average value of orders. The PCLV is the product of these two values.

The predicted number of orders is computed using past behavior. It looks at the buying pattern, and for customers who don't have a buying pattern (Non-Buyers), it looks at the browsing behavior (when the customer was online for the last time, when they joined for the first time, how many products they browsed, etc).

The predicted value of orders is based on the historical values of orders as well as the browsing history. For example, if a customer browses expensive products, the value of orders would increase as well.

The PCLV is an expected value. If someone has a 10 percent chance of making one purchase in the future, and is likely to spend $10, their PCLV would be equal to $1.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	2 years minimum, leverages up to 4 years

Likelihood to Convert

Bluecore's Likelihood to Convert model captures this short-term nuance, allowing marketers to design campaigns that nudge customers along in their journey toward a purchase.

The model uses numerical features that capture each user's interaction with your site. Specifically, it takes into account statistics on: on-site browsing and cart behavior, search history, and purchase history.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year
Product Search events	1 year

Likelihood to open/click on email

Predicts how likely a Customer Profile is to open or click an email.

Bluecore samples data from customers who are subscribed or unsubscribed. The models observe various signals: behavior on the website, purchases, and email interaction. The model tells between click/no-click and open/no-open based on the same signals.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Delivered emails	1 year
Opened emails	1 year

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Clicked email	1 year
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year

Likelihood to unsubscribe on email

This model predicts how likely a Customer Profile is to unsubscribe from emails.

It samples data from customers who are subscribed or unsubscribed, looking at various signals: behavior on the website, purchases, and email interaction. It uses those as signals to predict how likely a Customer Profile is to unsubscribe.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year
Email Delivered events	1 year
Email Opened events	1 year
Email Clicked events	1 year
Email Unsubscribe events	1 year

Channel Preference

Channel preference uses a predictive, learning-based model to filter audiences and target communication based on the implicitly preferred channel of your customers, either mobile or desktop.

The feature identifies the avenue that customers are most likely to be engaged in based on interactions on the channel, such as clicks.

⚠ This model is only available if Bluecore is your platform for both email and mobile sends.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year
Email Delivered events	1 year
Email Opened events	1 year
Email Clicked events	1 year
SMS Delivered events	1 year
SMS Clicked events	1 year

Discount Affinity

This model examines to what extent a markdown price motivates customers to purchase.

The markdown affinity score looks at the history of purchases and the history of views of a customer. For each of the products viewed or purchased, the model looks at the history of prices of that product, and it assumes that the full price refers to the historical maximum price of this product.

The model can compute a ratio between the price where the product was viewed or purchased and the full price. For example, for a user who bought a product for $90 with a maximum price historically known to be $100, this ratio would be equal to 90 percent.

The score for a customer is an average of all these scores across all the products ever viewed/purchased. Since the products purchased are more significant than the products viewed, the average weighs more towards the purchase scores.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year

Product Affinity

Although this is sometimes referred to as Category Affinity or Brand Affinity, it can handle many types of affinities, and could be set-up to calculate affinities for any type of product attribute.

Category Affinity looks at a year's worth of viewed product, add to carts, purchases and through collaborative filtering using matrix factorization, learns an affinity space where both products and customers live. Bluecore gives higher weights to purchases than carts or views.

Once the model is trained, it:

1. Defines a category affinity filter using product fields and their value
2. Finds all products that match those values
3. Takes the average of all these products in the affinity space resulting in a single point in the product/customer affinity space
4. Computes each customer's similarity score to this average product point, which measures the customer's affinity to this category, higher score meaning more affinity.

All customers that have some preference are then ordered with a percent-rank from 0-100 representing their ordering, with those with higher percent-rank having higher affinity.

- "Some Preference" → percent-rank >=0
- "Medium Preference" → percent-rank >= 20
- "Medium-High Preference" → >= 40
- "High Preference" → percent-rank >= 60
- "Very High Preference" → percent-rank >= 80

Setup

For all types of product affinities, the client needs to provide direction for which categories or values should be included when determining affinities. This can be set up in the Bluecore UI or a list can be provided.

- If the product affinity is set up as a category affinity, the results are added or updated on Customer Profile records each time the model runs, but the same is not true for other types of affinities added. Any additional affinities added are only accessible through Audience Builder, and not via standard exports.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year

Email Send Time Optimization (STO)

Determines the optimal times during a specified time period to send a campaign. Over time, Bluecore's intelligence optimizes to each individual based on previous click engagement times to individually customize the delivery times of future campaigns. Each send is fully personalized on a per individual basis to reach their personal highest engagement window within a 24-hour period.

The model analyzes online activity such as clicks and conversions to determine the best time to reach an individual. Up to three campaigns in a 24-hour day can be sent using STO, while Bluecore decides what times of day and order to be sent to optimize for clicks over the day.

As the STO model optimizes the timing of personalized campaigns, it's important to note that the scheduling is based on the timezone set for the namespace. This means if a campaign with STO is scheduled to send on Sept. 1, it could actually send on Sept. 2 for some recipients, depending on their timezone.

If an STO-enabled campaign is not sent in the past 15 days, the platform automatically stops running these calculations until the data is used again or the data is considered stale (whichever occurs first).

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	100 days
Purchase events	100 days
Email Delivered events	100 days
Email Open events	100 days
Email Click events	100 days

Email Auto Prioritizer (Smart Automatic Frequency Capping)

Auto Prioritizer uses AI and reinforcement learning to determine the right frequency and priority of emails to each Customer Profile, potentially ignoring usual frequency capping.

As it learns, the Bluecore Auto-Prioritizer continuously analyzes results from these decisions and optimizes its exploration to the areas where the most engagement has occurred until ultimately finding the best choice.

After the learning phase, Bluecore Auto-Prioritizer does not stop learning but instead will continuously explore away from the best choices to either re-affirm the current decision or to find a better option if that choice has changed.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Purchase events	60 days
Email Delivered events	60 days
Email Open events	60 days
Email Click events	60 days

Product Recommendation Models

Next Best Purchase

The Next Best Purchase model predicts a customer's next purchase. Next Best Purchase provides marketers with a dynamic and continuously learning recommendation strategy that can be used across their entire customer list.

This model leverages Bluecore's unique data assets, specifically its full historical and real-time understanding of how customers interact with the website, for example what products they view, cart or purchase as well as how they've engaged with product recommendations in past campaign sends. This unique feedback loop is combined with Bluecore's deep and real-time understanding of the retail catalog to continuously explore and optimize the products and categories that should be recommended to each customer.

If this model isn't used in a campaign in over seven days, the platform automatically stops running these calculations until the recommendation data is used again or the data is considered stale (whichever occurs

first).

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	1 year
Purchase events	1 year
Add-to-cart events/Viewed cart events	1 year
Email Click events	1 year
Valid Product Catalog data	1 year

Product Groups

*Requires “Next Best Purchase” model

This model is a derivation of the next best purchase model. Next Best Purchase can recommend products that come from different categories, or “product groups” since it is not designed to output a “coherent” set of recommendations. Product groups solve for this by only recommending products from the most likely product group the customer is going to purchase from next.

If this model isn’t used in a campaign in over seven days, the platform automatically stops running these calculations until the recommendation data is used again or the data is considered stale (whichever occurs first).

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Purchase events	30 days

Cross-Sells

*Requires “Next Best Purchase” model

This model is a derivation of the product groups model. It is aimed at driving cross purchase (purchase in a different category)

How it works:

This model outputs a coherent set of recommendations, from a product group that is new to the customer.

If this model isn’t used in a campaign in over 7 days, the platform automatically stops running these calculations until the recommendation data is used again or the data is considered stale (whichever occurs first).

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Purchase events	30 days

Co-Recommendations

The Co-Recommendations model suggests products using Bluecore’s recommendation engine, which applies collaborative filtering to make recommendations based on similar customers’ actions.

By default, Co-Recommendations can be based on what similar customers viewed, purchased, added to cart, or searched; with technical services, these can also be based on custom events.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Viewed Product events	30 days
Purchase events	30 days
Add-to-cart events/Viewed cart events	30 days

Replenishable Products & Replenish Buyers

There are two main steps for identifying Replenishable Products and Categories. First, customers who have purchased a specific product (or category) in a pattern that implies replenishment (at least 50 people have purchased from a grouping at least three times) are identified. Then, the personal replenishment cadence for these products is determined. The exact details of the pattern may be customized, but generally, the model looks for a minimum number of repeat purchases that are spaced out over time.

Once these customers (Replenishers) and the products they replenish are identified, the next step is to determine which of these products have a minimum number of Replenishers. These products are now identified as Replenishable Products.

Once the Replenishable Products and their Replenishers are identified, the Replenishers are targeted at their dynamically calculated personal replenishment cadence each time a new purchase is detected. In addition, all new buyers of Replenishable Products are targeted at an optimized cadence based on the average of all the personal replenishment cadences of the Replenishers.

This way, the highly valued Replenishers are messaged at their own personal cadence while motivating new buyers to become replenishment buyers at a timing that is automatically customized for each product.

Type of Data	Amount of Historical Data Required for Maximum Effectiveness
Purchase events	Minimum 30 days, will use all purchase history available