eBook:

How Machine Learning Can Deliver Personalization for Video Workflows

Drive customer engagement with personalized recommendations

Contents

Chapter 1

Introduction to personalization and collaborative filtering

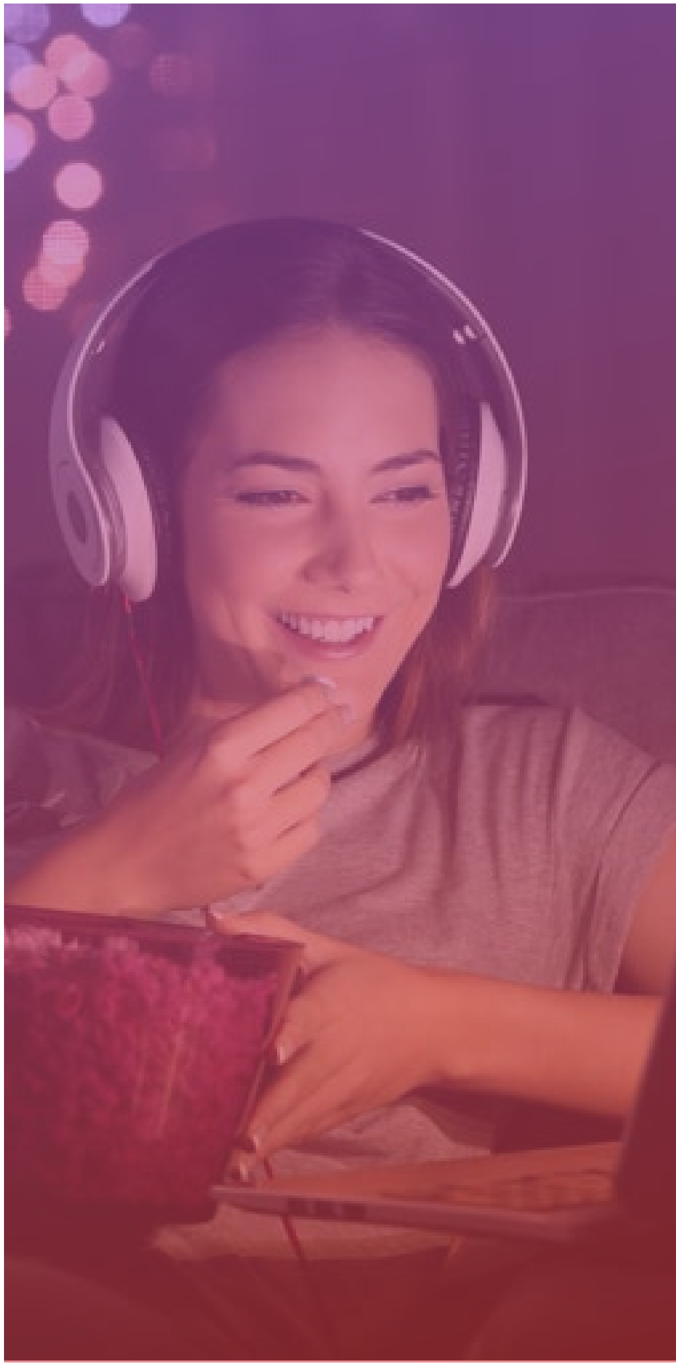
“Customers who watched this item also watched …”

The maturing of the online video market with over-the-top (OTT) and direct-to-consumer (DTC) services leads to a seemingly endless amount of content to be consumed. All that video on all those channels creates “surf fatigue” for viewers. It’s overwhelming and the numbers bear it out: The average U.S. adult spends over 7 minutes deciding what to watch.[1]

As a way to quickly surface video that might be of interest to its customers, streaming services serve up content recommendations based on machine learning (ML) algorithms utilizing user data. Such dialed-in recommendations around, for example, watch history, give the viewer the sense that the content has been curated especially for them.

To maximize views on your video streaming service, the quicker you can get visitors watching videos that they may enjoy, the more likely they are to stay engaged and continue viewing. In this eBook, discover what’s behind delivering tailored recommendations for videos using artificial intelligence (AI) and ML and how it can be used to drive better customer experiences for your viewers.

1. Neilsen, The Nielsen Total Audience Report, Q1 2019



The history of personalization: 2003

In 2003, three Amazon employees wrote a paper entitled "Amazon.com Recommendations: Item-to-Item Collaborative Filtering" that lays out the current state of personalization, as well as its future: "A good recommendation algorithm is scalable, over very large customer bases and product catalogs, requires only sub-second processing time to generate online recommendations, is able to react immediately to changes in a user's data, and makes compelling recommendations for all users regardless of purchases and ratings."

The piece was so influential that in 2017 the journal IEEE Internet Computing named it the single paper that had withstood the "test of time" over the course of the publication's twenty-year history.[2] Amazon machine learning and applied scientists continue to make contributions in the field, with the recent paper "Temporal-Contextual Recommendation in Real-Time"[3] winning the best-paper award in 2020 at the Association for Computing Machinery's (virtual) Knowledge Discovery and Data Mining conference (KDD).[4]

2. Amazon, The history of Amazon's recommendation algorithm, November 22, 2019
3. Amazon, Amazon researchers win best-paper award at KDD, August 26, 2020
4. KDD Accepted Papers, Temporal-Contextual Recommendation in Real-Time, August 2020

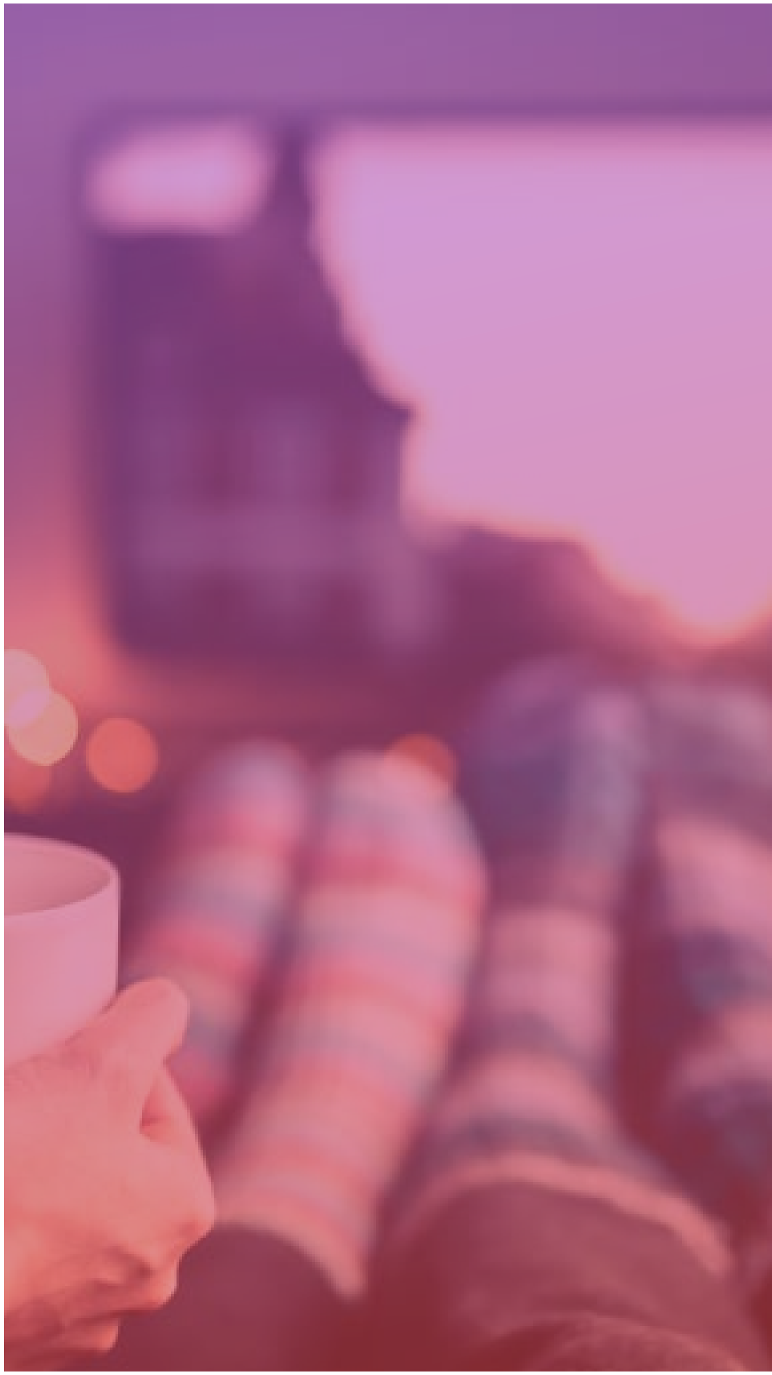
Collaborative filtering versus content filtering

The concept of collaborative filtering has persisted as a means to leverage user experiences to inform personalization. The idea behind collaborative filtering is that users with similar tastes are more likely to have similar interactions with something they've never seen.

The other major approach to recommendation engines is content filtering. Compared to strictly content-based filtering that relies on user preferences provided explicitly or implicitly, collaborative filtering takes into account other users' experiences.

The collaborative filtering approach has an advantage over content-based filtering. It produces better results around diversity (how dissimilar recommended items are); serendipity (a measure of how surprising the successful or relevant recommendations are); and novelty (how unknown recommended items are to a user).[5]

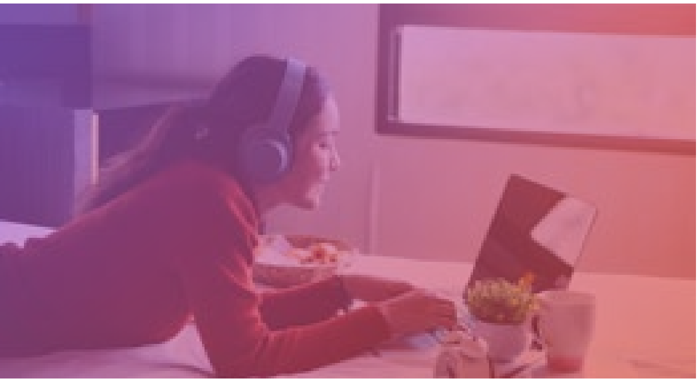
5. Brent Rabowsky and Liam Morrison, What's new in recommender systems, Amazon Web Services, November 17, 2020



Deep learning with Amazon Personalize

Amazon Personalize benefits

- Deliver high quality recommendations, in real-time
- Easily implement personalized recommendations in days, not months
- Personalize every touchpoint along the customer journey
- Data privacy and security



Deep learning moves beyond ML, attempting to simulate the way the brain learns and processes information, producing models with more accurate insights and predictions.

Amazon Personalize is a fully managed, deep-learning system for personalized recommendations. It incorporates approaches from recent research while lifting the burden from data science and developer teams of managing a recommender system at scale.[6] Through deep learning, Amazon Personalize can analyze nonlinear, co-interactions of large sets of data and produce recommendations that normal statistical algorithms cannot.

Amazon Personalize is an application programming interface (API)/AI service that allows you to host ML models for recommendations and personalization. It utilizes interactions from your website—user events such as clicking, watching, rating, and liking—and stores them.

Using Amazon Personalize, you can import information about your video titles as well add metadata about the titles and users. This metadata is important for training your model in order to aid the training algorithm to find insight into similarities in preferences and behavior of users.

6. Brent Rabowsky and Liam Morrison, What's new in recommender systems, Amazon Web Services, November 17, 2020

Chapter 2

How Amazon Personalize works

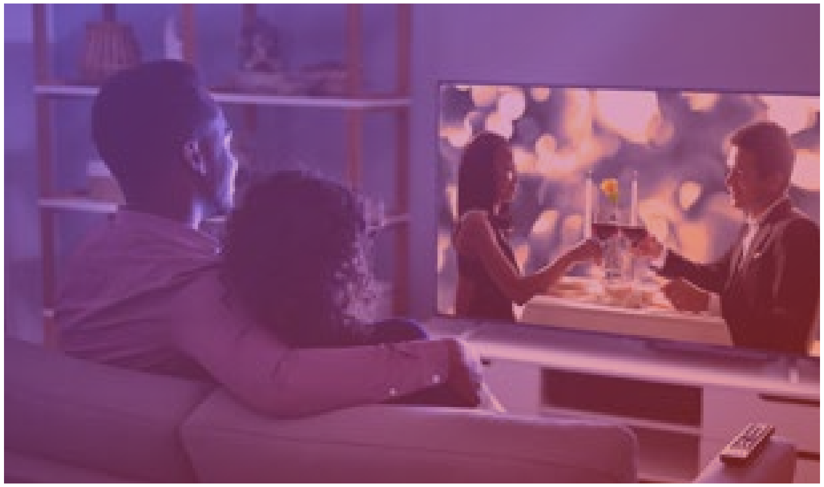
How Amazon Personalize works

Amazon Personalize makes recommendations based on an ML model which is trained using your data. Over time, the original dataset can grow to include new metadata as well as the consumption of real-time user event data.

Of course, to train a model, you need data, which is stored in Amazon Personalize datasets. The data can be provided by importing historical data from an Amazon Simple Storage Service (Amazon S3) bucket, or event data can be recorded as it is created.

Amazon Personalize has an Amazon Web Services (AWS) Management Console that can be used to create, manage, and deploy solution versions. Or you can use the AWS Command Line Interface (AWS CLI) or one of the Amazon Personalize SDKs.

The steps to setting up Amazon Personalize



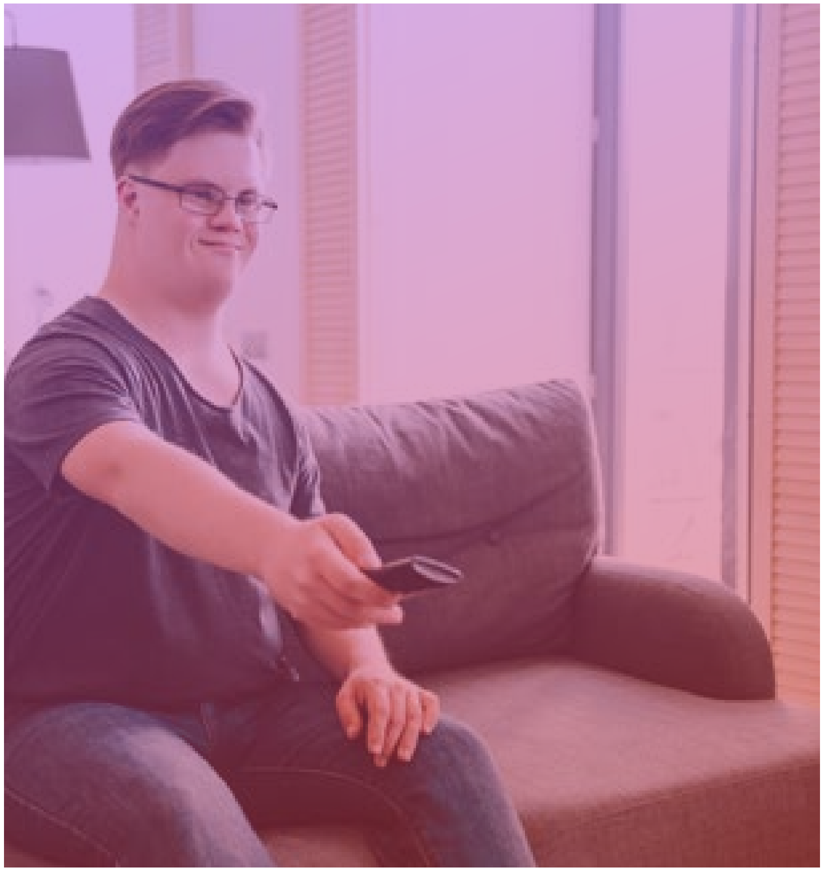
On a platform for watching streaming content, a campaign can display recommendations based on prior viewing habits that were part of the dataset. Beyond that, there is modeling of the historic data.

If someone has repeatedly searched for kids' educational shows in the morning but sit-coms later in the evening, the time of day will be taken into account for which recommendation to serve. Likewise, event filters will automatically move shows from the top recommendations in the list the instant they are watched.

How does Amazon Personalize get to the point of delivering such curated, real-time recommendations? Here's the workflow for training, deploying, and getting recommendations from a campaign:

- Create related datasets and a dataset group
- Get training data
- Import historical data to the dataset group
- Record user events to the dataset group
- Develop a solution version (trained model) using a recipe
- Evaluate the solution version using metrics
- Launch a campaign (deploy the solution version)
- Provide recommendations for users

How datasets and dataset groups function



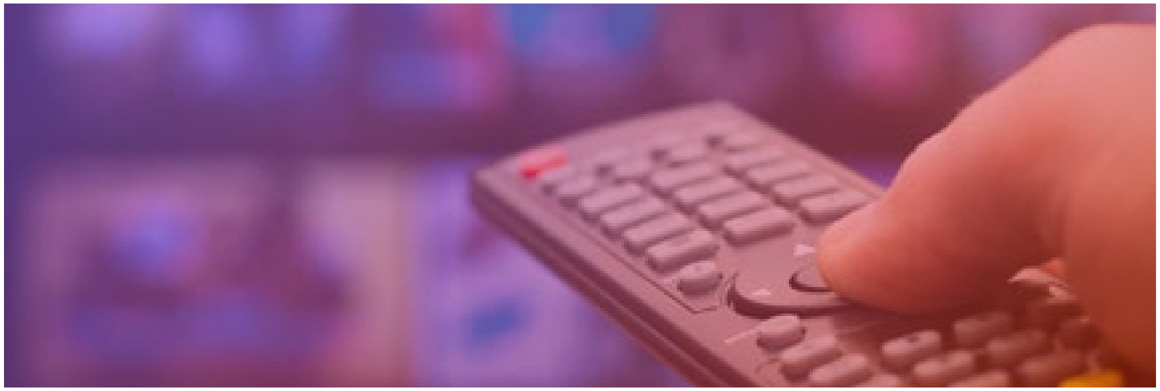
Dataset groups contain related datasets. Three types of historic datasets can be created: users, items, and interactions.

These dataset groups can do different things. For example, if you have an application that provides recommendations for movies and another for television programming, in Amazon Personalize, each application will have its own dataset group.

The minimum data requirements to train a model are:

- 1,000 records of combined interaction data (after filtering by eventType and eventValueThreshold, if provided)
- 25 unique users with at least 2 interactions each

A recipe that cooks up personalization



When there is enough data in the interactions datasets (historical and live events), it can be used to train a model (a solution version).

That model is trained using a recipe, which is available in Amazon Personalize. The recipes are based on an algorithm and data processing steps that optimize a solution for a certain type of recommendation based on input data.

Amazon Personalize supports a number of predefined recipes—it can automatically choose the most appropriate one based on its analysis of the training data. Or you can choose which recipe to use to train the model.

There is a different use case for each recipe and it's just a matter of deciding which one is best suited to your needs.

Amazon Personalize supports a number of predefined recipes … or you can choose which recipe to use to train the model.

Using metrics to dial in personalization

With a solution version based on the chosen recipe, it is ready for a deployment as a campaign.

After the solution version has been created, it's time to evaluate the metrics from the training.

These metrics give an indication of the solution version's performance. They are displayed in the console and you have the ability to create a new solution version, as needed. The most likely scenario is training your model with multiple recipes and using the one that results in the metrics with the best performance.

With a solution version based on the chosen recipe, it is ready for a deployment as a campaign, which allows Amazon Personalize to make recommendations for your users.

Get recommendations as users browse videos

Amazon Personalize campaigns can produce recommendations in real time or, using purely historical data, as part of a batch workflow. Real-time recommendations update as users use your applications.

On a site hosting exercise videos, for example, someone might be browsing cardio workouts. Amazon Personalize will serve up similar titles—while eliminating yoga videos—as the user chooses different videos.

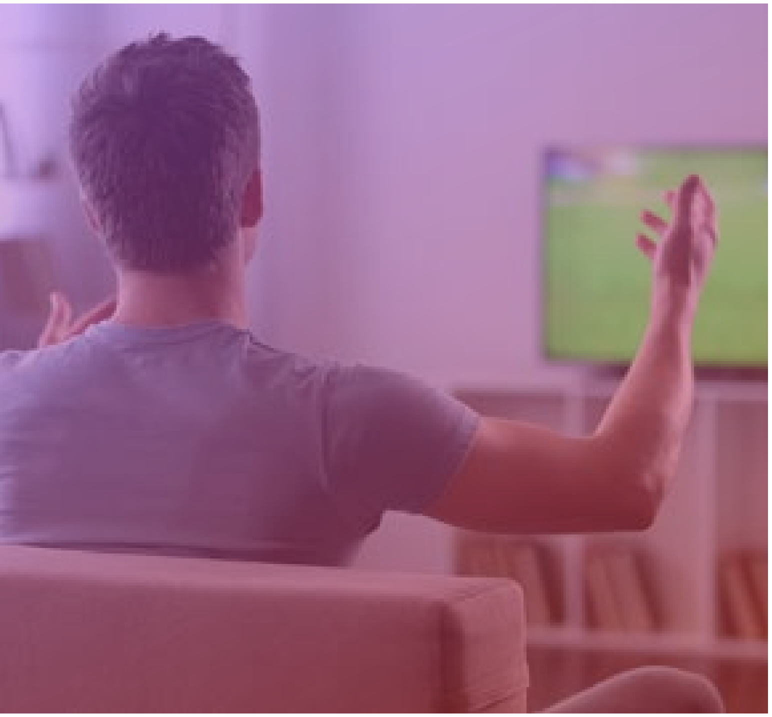
Batch recommendations are used when there are large datasets that don't require real-time updates, like sending a recommendation on your most popular videos to all the users on an email list.

Amazon Personalize will serve up similar titles … as the user chooses different videos.

Chapter 3

How lack of real-time personalization can frustrate viewers

How lack of real-time personalization can frustrate viewers



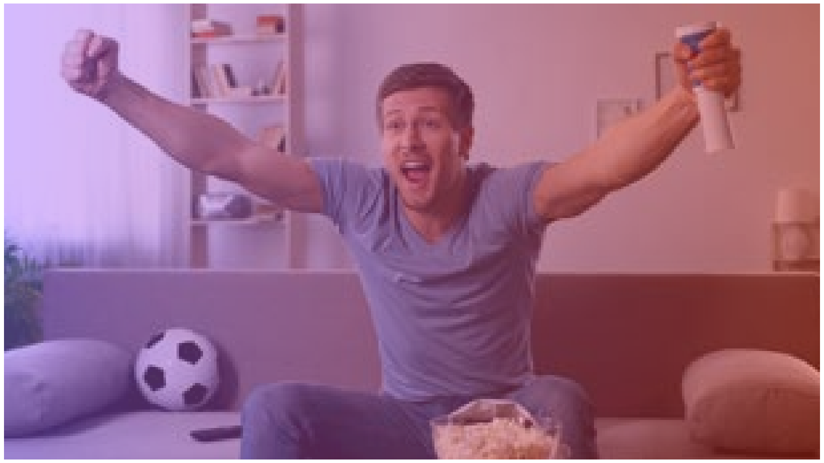
Oliver is a 20-year-old Brit and one of the millions of fans of a premier European football club. His favorite player—let's call him Ronen—has been loaned from Oliver's club to its archrival.

But Oliver is loyal to his hero and continues to root for Ronen and follow his exploits with the hated opposition—although he doesn't dare tell his mates. Problem is, when he visits the rival's website to watch a Ronen highlight, it's immediately followed by videos of goal after goal by his teammates. *Arrrgh!*

As a superfan, Oliver wants a Ronen highlight reel, not just one video, which is all he gets of Ronen. That's because the videos Oliver sees after Ronen's epic goal are the ones published most recently and are not personalized for him in any way.

This creates a very unsatisfying user experience for Oliver, rubbing it in that his beloved striker now plays in the bad guys' stadium.

Deliver on-target video recommendations



The reason Oliver has such a poor experience is because collaborative filtering systems have a cold-start user problem: they simply don't know what to do with someone they know very little about. He gets served what's recent, not what's relevant.

Amazon Personalize solves the cold-start problem of not being able to serve up relevant recommendations to new users with real-time event ingestion and altering recommendations based on those events. It does this by recommending new products to users who, by clicking, have positively engaged with similar items in the past.

Also, Amazon Personalize recommends new items with exploration based on metadata information. These recommendations update in real time, so a new user like Oliver begins to see personalized Ronen content after only a few clicks. *Gooooal!*

There's another way for the club to zero in on Oliver and his Ronen videos. Dynamic filters in Amazon Personalize allow you to apply business rules to your recommendations on the fly, making sure recommendations that are generated include a favorite player based on each individual user's preferences.

Amazon Personalize solves the cold-start problem of not being able to serve up relevant recommendations to new users.

Chapter 4

Case study:

Pulselive and real-time personalization goals

Case study

Pulselive teams up with Amazon Personalize

Wanting to cater to the specific interests of fans like Oliver, the football club tapped its digital partner, Pulselive, to increase fan engagement with its videos.

Pulselive faced two major challenges: The company didn't employ any data scientists and was looking for a solution its engineers with minimal ML experience could work with, yet still produce measurable results.

Given these considerations, Pulselive selected Amazon Personalize as a recommendation engine for the team for these reasons:

- Low barrier to entry, technically and financially
- Ability to quickly A/B test to demonstrate the value of Amazon Personalize
- Simple proof of concept (POC) caused minimal disruption to the existing site
- Clear focus on impact and improving results rather than needing to know the inner workings of the recommendation engine[7]

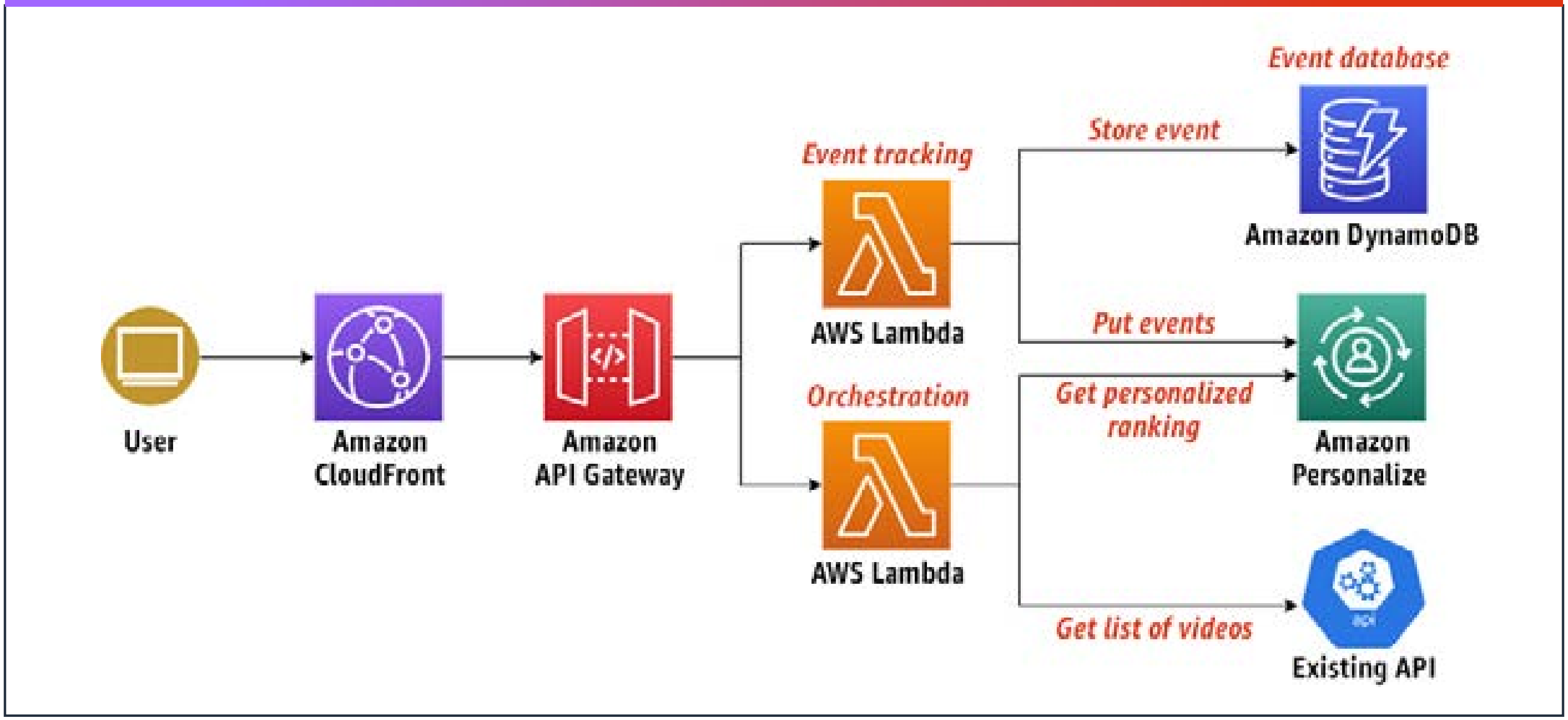
7. Mark Wood, Increasing engagement with personalized online sports content, Amazon Web Services, July 22, 2020

Getting a recommendation solution up and running—fast

Pulselive started out with A/B testing in a POC that was spun up in just a few days.

The company worked with the Amazon Prototyping team to determine a range of options for the first integration that required minimal changes to the current website and provided easy A/B testing.

After examining all the points where a user is shown a list of videos, it was clear that re-ranking the list of videos to watch next would be the quickest way to bring in personalized content.



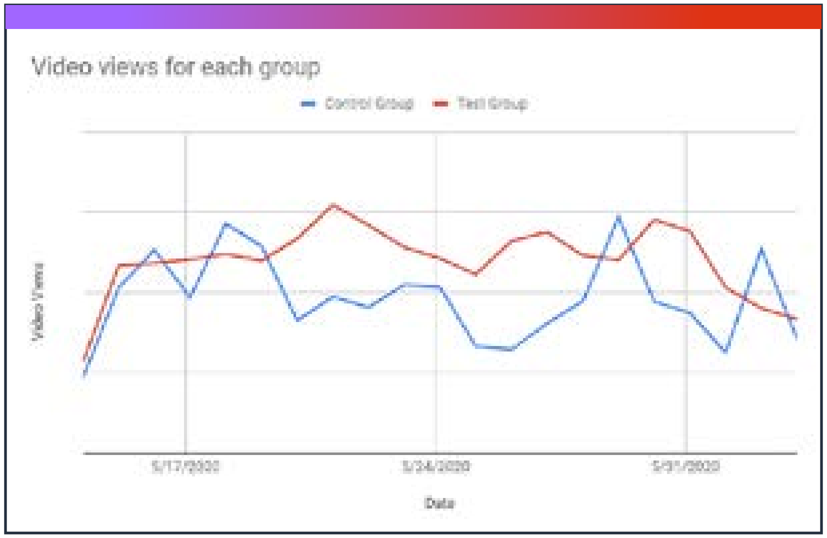
Video views up 20 percent through personalization

The video re-ranking experiment was tested over a two-week period, ensuring there were days with multiple matches, as well as those with none. Users were randomly assigned to the control group or the test group. Video views and what prompted them were tracked for everyone.

Initially, Pulselive didn't think re-ranking the video list would make much of a difference but were pleasantly surprised to see an increase in views per user of 20 percent.[8]

"Their fans are clearly embracing the new recommendations. Leveraging Amazon Personalize, we will be able to further push the limits in building data driven 1-to-1 personalised experiences for sports fans everywhere," says Wyndham Richardson, Pulselive's Managing Director & Co-Founder.

No doubt football supporters like Oliver are now more engaged than ever with their favorite players and clubs.

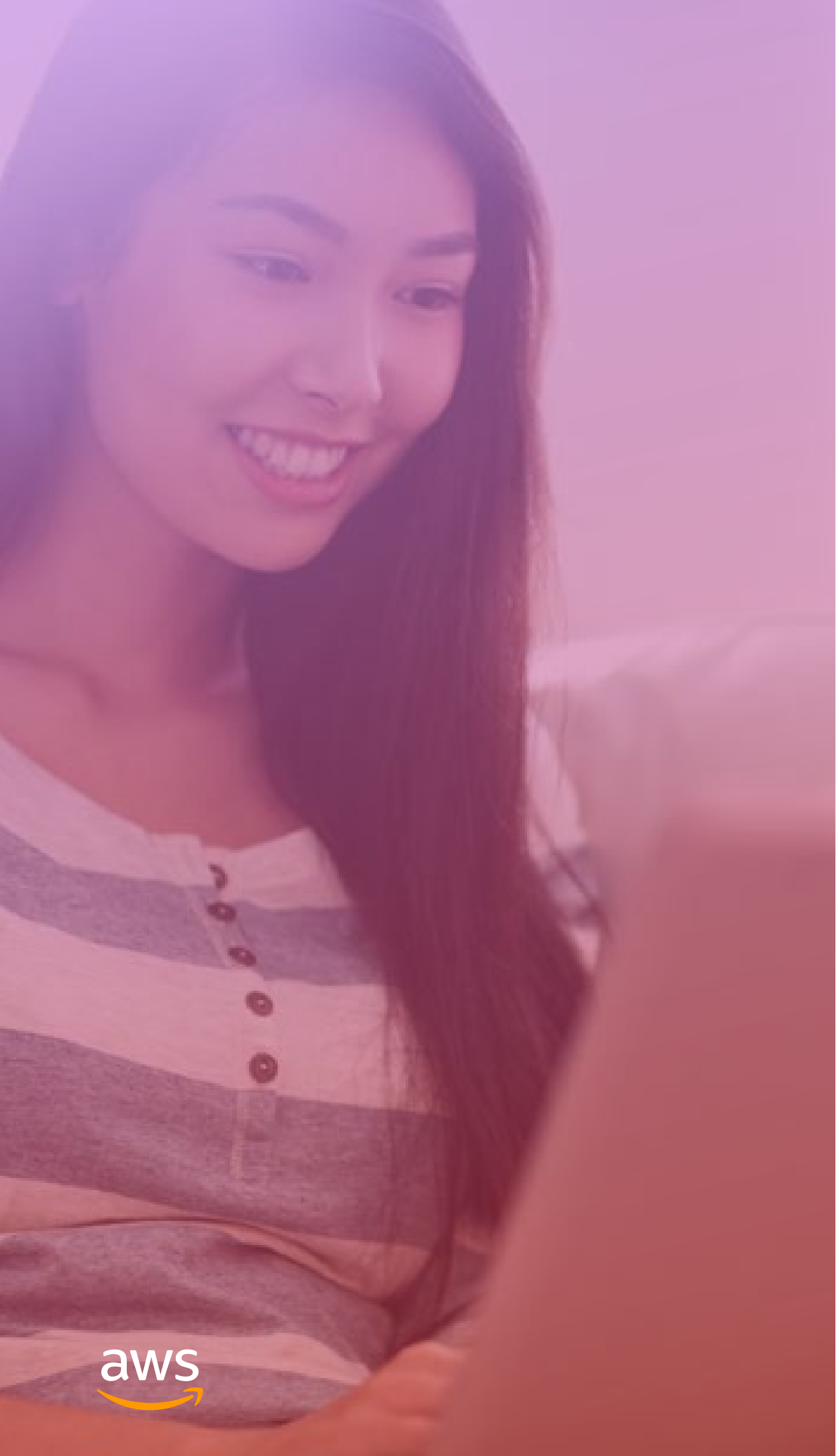


Pulselive personalization results

8. [Mark Wood, Increasing engagement with personalized online sports content, Amazon Web Services, July 22, 2020]

Chapter 5

Get started with Amazon Personalize



Create better user experience with real-time personalization

As Pulselive discovered, Amazon Personalize makes it easy for developers to build applications capable of delivering a wide array of personalization experiences, including specific video recommendations and re-ranking.

It is a fully managed ML service that goes beyond rigid and static rule-based recommendation systems by training, tuning, and deploying custom ML to deliver highly personalized recommendations to customers across industries such as retail, media, and entertainment.

Amazon Personalize provisions the necessary infrastructure and manages every aspect of the ML pipeline, from processing the data and identifying features to using the best algorithms and training, optimizing, and hosting the models.

Learn more about Amazon Personalize, the same ML technology used by Amazon.com for real-time personalized recommendations.