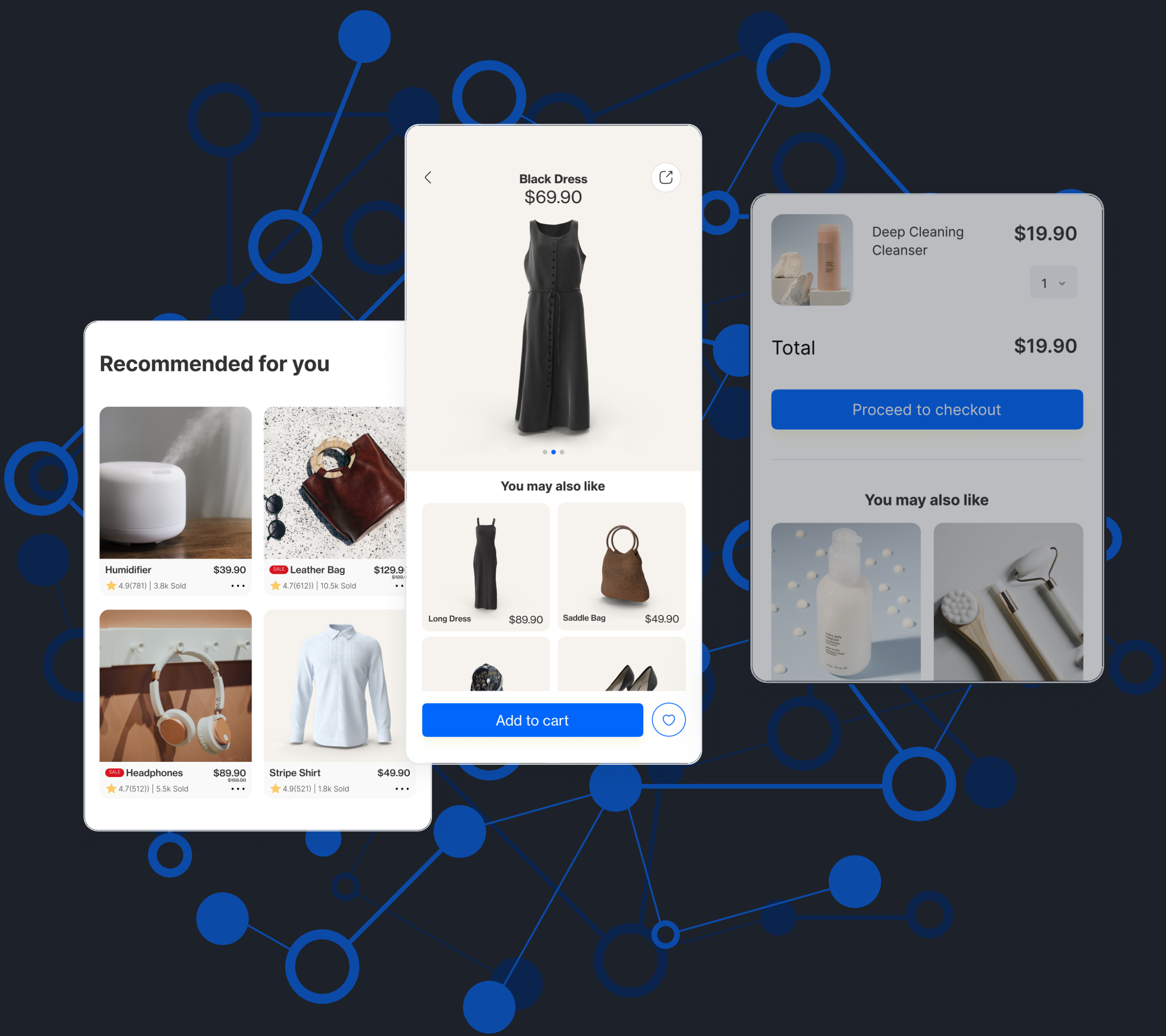
BytePlus

Deep-Learning based Recommendation System and BytePlus's Approach



Abstract

With the ever-growing volume of online data, recommendation systems have been an effective way to solve the prominent problem of information overload. Over the past few years, we have witnessed the success of deep learning in various domains, such as computer vision and natural language processing (NLP). This allows for tremendous advancements to the recommendation systems we know today.

This whitepaper delves into the development and evolution of recommendation systems, and how they have become highly integrated and essential in today's digital economy. We will also shine the spotlight on a leading recommendation system, BytePlus Recommend, to overview how it is applied in various uses today. The content within is specially tailored for engineers, system architects and data scientists who want to understand how an industry-leading recommendation system works in practice, as well as the key considerations they need to make when deploying a recommendation system.

Table of contents

Introduction

Accelerating Personalized Experiences with Recommendation Systems

With the extensive amount of products and content available in today's digital economy, personalization plays an essential role in forging an enriched customer experience. Digital businesses often store enormous volumes of data, including user data, web traffic, and transactional data that can help them make crucial business decisions regarding product roadmaps, marketing, and strategy planning.

And when consumers scour the Internet for information about a product or content, they are no longer seeking a list of unfiltered results — but a curation of recommendations that match their interest. That is why many digital businesses are increasingly using recommendation systems to keep pace with rising customer expectations for a personalized customer experience. Personalization through recommendation systems is pivotal to improving user engagement, which in turn, increases conversion rates and business revenues.

Content and e-commerce platform providers are amongst the first to find success with recommendation systems across a variety of use cases.

Netflix and YouTube, for example, analyze enormous amounts of data to provide viewers with a more personalized user experience. Recommendation systems not only help address the issue of 'choice overload' — they also improve the user experience to help raise customer retention.

On the e-commerce front, Amazon has successfully optimized online shopping experiences on its platform with recommendations across every step of the purchase funnel, including product landing pages and search results. When the simplicity of the purchase process is augmented by personalized product recommendations, the resulting experience can increase conversion rate and revenue while retaining customer loyalty.

Extent of Influence

Market statistics have shown that recommendation systems influence:

80%	35%	60%
of the content streamed on Netflix[1]	of products purchased on Amazon[2]	of video clicks on YouTube homepage[3]

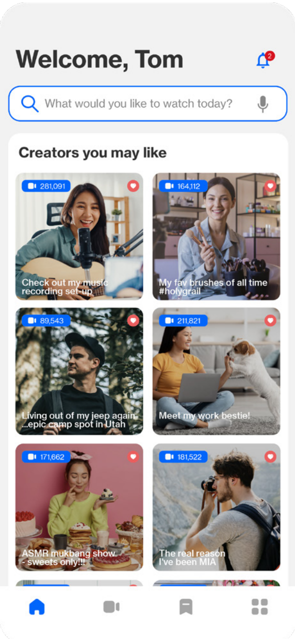
What these companies have in common is the commitment to high-quality recommendations in their personalization strategy. Hence, any digital enterprise seeking personalization as a competitive advantage must be prepared to adopt the same mandate.

Out-of-the-box recommendation solutions in the market that use basic or rule-based algorithms are unlikely to provide the flexibility needed by these businesses to tackle unique challenges. Instead, they must bolster their technology stack with best-in-class recommendation solutions that are:

- Highly customizable, yet easy to set up over an inclusive platform
- Driven by leading machine learning (ML) technology to deliver powerful performance
- Capable of real time updates to keep recommendations ultra-relevant
- Augmented with additional guidance to input data optimally and maximize its value

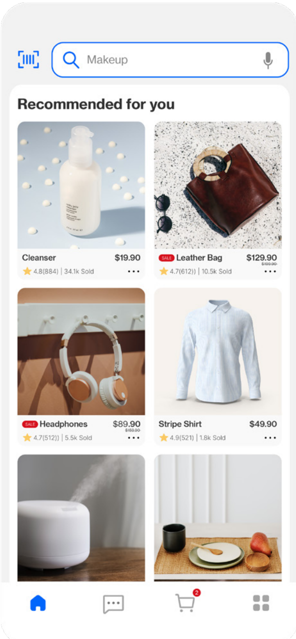
Rising popularity across industries

Digital companies in any industry can adopt recommendation systems for their personalization strategy to strengthen user engagement, increase conversion rates and grow business revenues.



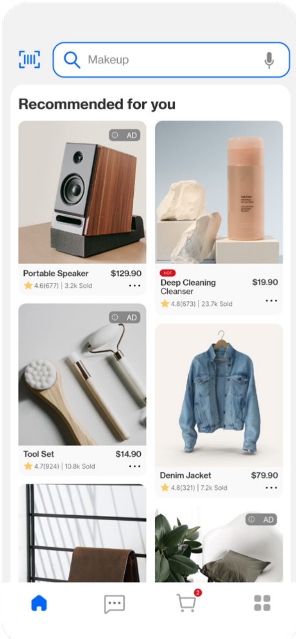
Video/Livestreaming

OTT video platforms and streaming service providers are tapping on recommendation systems to deliver relevant and accurate video recommendations to viewers, helping to improve user engagement and reduce customer churn.



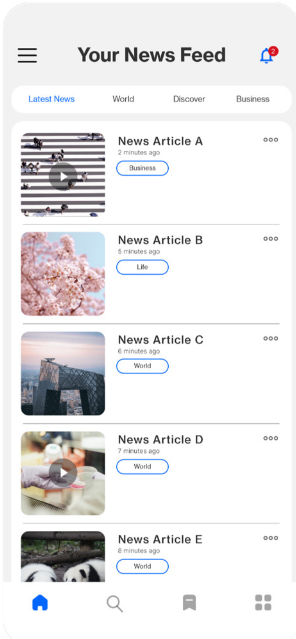
E-Commerce

Recommendation systems can help businesses provide more accurate product recommendations based on user, product and user behavior factors throughout the shopping journey.



Advertising

Recommendation systems are used to generate in-depth consumer insights based on specific user interaction and preferences, which can be used to improve the relevance of advertising efforts and digital marketing campaigns in real time.



Content

Content providers and media companies can leverage recommendation systems to gain deeper insights into content consumption patterns, make accurate recommendations, and deliver engaging, personalized experiences.

The Evolution of Recommendation Systems

Recommendation systems are a sub-class of information filtering systems that suggest relevant items such as movies to watch, products to buy, or anything else depending on industries to users[4]. The recommendation system has evolved drastically in the past decades, starting from traditional approaches such as popularity-based and rule-based models, to more personalised content-based filtering and collaborative filtering models like matrix factorisation or factorisation machine.

In recent years, deep learning models have seen a significant development and revolutionized the architecture of recommendation systems. They solve the problems of conventional models and significantly boost the performance of recommendation systems.

This whitepaper focuses on deep learning models and how they are implemented in recommendation systems.

Deep Learning Models

Deep learning[5] has brought essential improvements to the performance of recommendation systems, with models like deep neural networks able to capture non-linearity in data, thus helping to pick out complex patterns among user item interactions. They are also highly effective in learning deep representation, which retrieves data from text, images and videos while reducing the effort in feature design.

An example is the classic two-stage funnel process for information retrieval that was first proposed by YouTube (see Fig 16).

Terminology

Deep learning models	Part of a broader family of machine learning methods based on artificial neural networks with representation learning. Deep learning models can be supervised, semi-supervised or unsupervised.
Neural network	A network composed of artificial neurons or nodes, used for solving AI problems
Candidate generation	A key stage in the recommendation process, where the system generates a set of relevant candidates based on a specific filtering approach
Embedding	A way of representing entities using learned vector. It involves a mapping from a discrete set of queries and items to a vector space called the embedding space.

Here, every stage involves a neural network, so the network at the candidate generation stage will select a subset of candidates from a large pool of videos. Hundreds of candidates bearing general relevance to the user are generated at this stage, before going through a more thorough step to select the final list of videos for the user.

The process then reaches the ranking stage, where another network handles the assigning and ranking of each video with a preference score, so that the recommendations are delivered by order of importance.

This two-stage approach allows for effective and engaging recommendations from large pools of content or products.

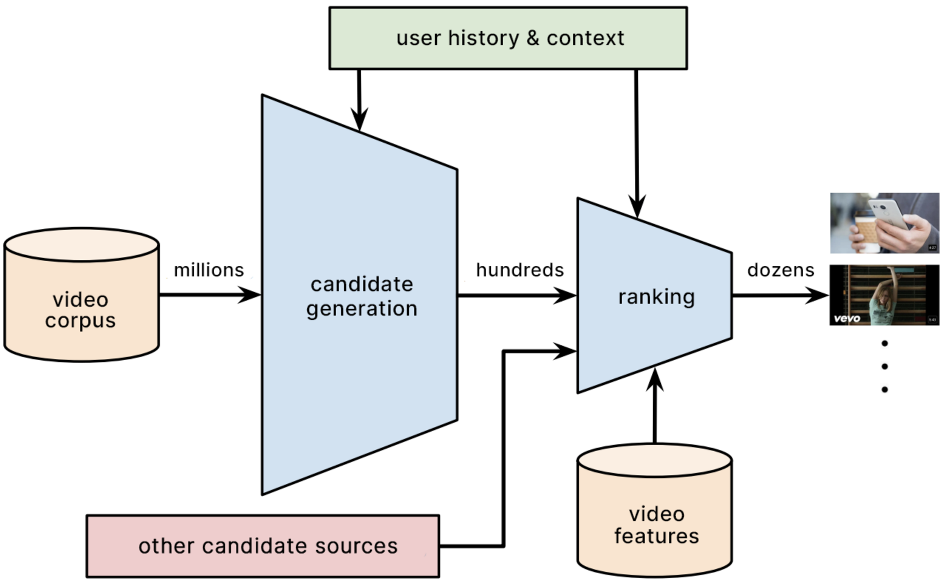


Fig. 1. Two-stage recommendation system structure

Recommendation Systems: The Breakdown

The recommendation systems widely adopted in the e-commerce and media industries are large scale versions of the system structure discussed previously. These typically consist of five main components:

Data Synchronisation	Collecting and synchronising raw data
Data Processing and Feature Engineering	Processing and preparing raw data for model training
Distributed Model Training Platform	A dedicated machine learning platform for model training
Candidate Generation, Ranking, and Re-ranking	Sub-steps of generating recommendations a. **Candidate Generation:** Retrieving a subset of relevant candidates based on different logics for every user b. **Ranking:** Further refinement of relevant candidates using a formula or metric like click-through rates c. **Re-ranking:** Applying additional rules like filtering, de-duplication
Evaluation:	Online and offline approaches, based on suitable metrics, to evaluate the performance of the results generated by the recommendation system.

The diagram below (Fig 2) is an overview of the overall architecture and workflow of a modern recommendation system.

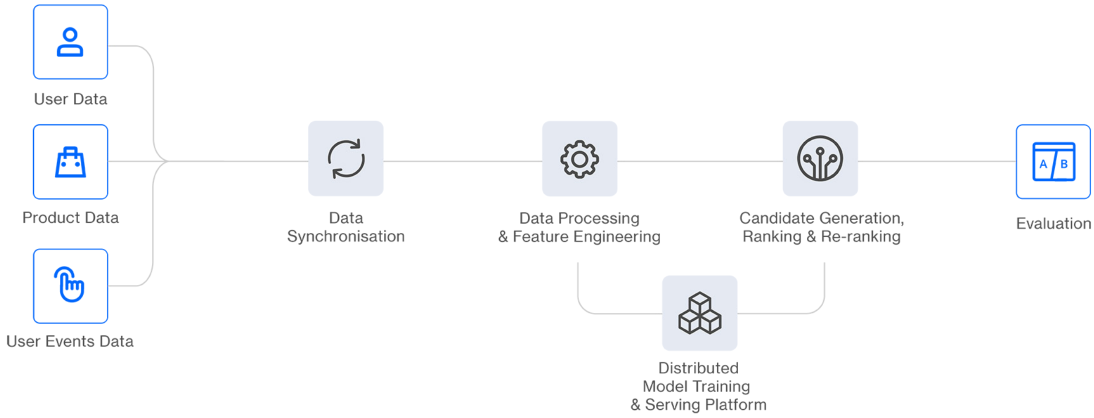


Fig. 2. Recommendation System Workflow

The 5 Components of a Recommendation System

Data Synchronisation

Data is a critical component of the system that ensures recommendation success, so the raw data collected is the fundamental building blocks. Model performance is directly related to how accurate, timely, and comprehensive the collected data is.

The more data there is on a user, the more efficient and personalized the recommendations are in optimizing a user's experience, whether it is to pique interest in an item or to aid a buying decision.

For example

To enable recommendations for e-commerce, three types of data are required:

User	ID, location, tags, ratings
Product	ID, title, price, reviews, ratings, price
UserEvent	Impression, click rate, 'add to cart', purchase rate, conversion rate

To enable recommendations based on user's behavior, two types of data must be collected:

Historical Data	Historical behaviours, user, product, and events data
Incremental Data	New behaviours, new changes or updates to the user or product

Data Processing and Feature Engineering

After collecting the relevant data, the next step is to process and extract meaningful features to serve as key inputs for various ML models.

Generally, the main outputs of the recommendation system platform are:

- Generating sample data for recommendation system training and evaluating
- Generating user and product profiles for model serving

With high-speed processing of large amounts of data, models can analyze and capture short-term changes in user interests.

A well-structured recommendation system can support accurate and optimal processing for incremental batch and real-time features. When done right, it increases the predictive power of the models, which defines the huge difference between a good and bad recommender.

Distributed Model Training and Serving Platform

The collection, processing and feature engineering of data is done in preparation for model training and serving on a well-developed distributed ML platform.

Over the past few years, recommendation systems have seen the most developments in this area. With the introduction of deep learning models, there has been an increase in the number of feature types that can be used as the inputs in recommendation systems.

The burst of innovation that advanced model architectures bring provides a great deal of uplift to recommendation performance. However, that also means that modern model training platforms need to support a variety of embedding types for candidate generation[7]. The more embedding types that the system supports, the more versatile it becomes in building machine learning models, thus generating better performance. But this makes businesses harder to scale up their recommendation capabilities.

For example, they must be flexible enough to support different model architectures to ensure optimized ranking through the system. When there is an increase in the number of features, the size of the model can easily go beyond the typical physical storage of a single device. To resolve this, a common approach is to shard the embedding into several parameter servers so that each can be updated more efficiently and quickly. The trained models can then be used to perform candidate generation and ranking to produce recommendations.

Model serving is then used to address challenges in low latency, high availability, and high throughput, which involves a high degree of complexity in architecture design and engineering optimization.

Candidate Generation, Ranking, and Re-ranking

Candidate generation, ranking, and re-ranking are key steps for generating recommendations[8]. When faced with many items and users, it is important to split the process into smaller parts for better result refinement.

The "funnel" approach is most used in recommendation systems as part of the model training and serving[9].

Candidate Generation

It is common to adopt multiple algorithms to generate different sets of candidates. These models are often known as candidate generators.

In the e-commerce industry, there could be tens of thousands of products available in the marketplace, so multiple candidate generators are used to generate hundreds of product candidates.

For example, one generator can be used to generate 100 candidates from the most popular products category, while another can general 100 candidates from products that are most like the target product. Likewise, another 100 or 200 candidates can be generated based on models like collaborative filtering or deep learning, respectively. These candidates will then be merged before being ranked.

The Candidate Generation Process

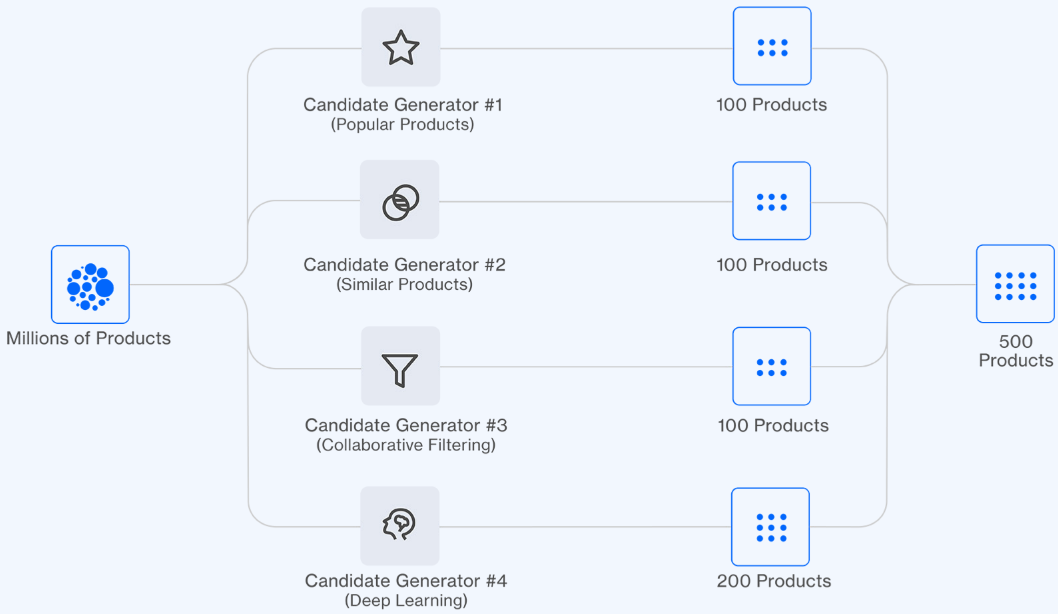


Fig. 3. The Candidate Generation Process

Ranking

After candidate generation, ranking is used to score the candidates so that the final set of items can be selected for display to the user. With a smaller pool of candidates generated from last stage, more features and more complex models are used to yield better results in scoring, such as approaches like multi-task learning[10], DeepFM[11], xDeepFM[12].

The table below compares the main differences between candidate generation and ranking.

	Pool of candidate	Output	Model Complexity	# Features	Processing speed	Accuracy	Focus
Candidate Generation	Millions	Hundreds	Low	Less	Very fast	Lower	Relevancy, diversity
Ranking	Hundreds	Dozens	High	More	Slower	High	Specific goals such as CTR

Re-ranking

Re-ranking is the final step which allows the business to rank the recommendation list again based on additional criteria to ensure the results are fresh, diversified, fair and so on.

Some of the commonly used rules include:

- **De-duplication**
  Identifying and removing duplicated items from the recommendation results
- **Diversification**
  Shuffling items based on rule sets such as limiting the appearance frequency of certain product categories
- **Filtering**
  Removing items based on rules sets such as being out-of-stock or having low ratings
- **Boosting**
  Shifting certain items up or down the recommendation list for purposes like promoting flagship store items

Re-ranking ensures that the final results are the most relevant, qualified, and personalized for the users.

Evaluation

Every model used by the recommendation system in learning would have gone through extensive experiments during its evaluation, both online and offline.

For offline evaluations, the commonly used metrics include precision, recall, AUC, mAP, P-R curve, and so on. However, most digital companies prefer to conduct online experiments such as live A/B testing to determine the effectiveness of a model, for various reasons.

Benefits of Online Experiments:

- **Accurate** – because online experiments are more accurate, they help to reduce data latency and data loss issues

- **Holistic** – online experiments bridge the gap for metrics that cannot be retrieved, such as the click-through rate, duration of each stay

- **Minimal bias** – online experiments usually remove training data biases as well, as training data is often derived based on existing recommendation systems. Therefore, a live A/B experiment is critical as it keeps all factors impacting the 2 groups of users the same, except for recommendation model, so we can fairly understand what is the impact of recommendation

As effective as online evaluation can be, it is equally important to also note the inherent challenges of A/B testing.

With limited user traffic, a significant amount of time is required for the other tests to complete. To resolve this, many companies have built or adopted a scalable A/B testing platform to support more efficient experiment rollout and measurements for parallel experiments.

The selection of suitable metrics to evaluate the performance of recommendation systems can be a challenge as well. At times, businesses might be confused on what are the right metrics to select for performance evaluation. They must bear in mind that selection is often impacted by the overall goal of the company and stakeholder teams; whether it is short term or long-term metrics; or sometimes comparing engagement metrics to conversion metrics.

5 Key Considerations for Selecting The Right Recommendation System

With the different types of recommendation systems available, it is worth considering the key characteristics and solution features that will best address the specific needs of different use cases. Here is a checklist of important considerations for online businesses to think through when deploying a recommendation system.

Data capability

It can be challenging to know what data to collect (which data fields, what frequency etc) for recommendation systems and even more challenging to understand what data issues you have that may impact recommendation performance. Outstanding recommendation systems not only guide you through which data to collect but also validate your data quality.

Model capability

Traditional rule-based model pales in comparison with the latest advances in ML based models. Choosing a recommendation system that encompasses ML based models will greatly enhance your user experience and business benefits.

Investment

Building a recommendation system in-house, especially one with robust ML functionalities can be a complex, time-intensive and costly process. It requires substantial investment in manpower as well as infrastructure.

Versatility

Whether a recommendation system provides flexibility to address your evolving business needs is also critical. Some recommendation providers only support out-of-the-box models that cannot be customized for your business goals.

BytePlus Recommend: The Approach

Many recommendation vendors in the market are rule-based or use basic AI algorithms to enhance their performance. By today's customer expectation standards, this basic level of personalization is underwhelming, and may not deliver the desired outcomes, despite the time and money invested to boost business performance.

ByteDance has been a pioneer in recommendation systems since its formation in 2012. A number of ByteDance products have adopted the use recommendation systems to improve their user experience, including TikTok, Douyin, Toutiao, and more apps.

BytePlus Recommend aims to help businesses quickly and easily create highly personalized recommendations across the entire user journey. The solution is backed by the company's industry-leading recommendation technology, robust support, and highly customizable models in an easy-to-use platform.

30%
Increase on average to conversion rate

50%
Increase on average to engagement time

30%
Increase on average to gross merchandise value

*Disclaimer: uplifts are calculated using average performance from our existing clients

The detailed dashboard on the platform provides visibility of the learning models in real time, to deliver insights into the recommendation performance and how much uplift has been delivered.

Key components of the BytePlus Recommendation approach include:

- Market leading ML algorithms
- ML platform dedicated to recommendation systems
- Deep expertise for highly customizable solutions

Market Leading Machine Learning Algorithms

BytePlus Recommend is known for using advanced ML algorithms for both candidate generation and ranking. These include proprietary algorithms such as Deep Reinforcement Learning and Deep Retrieval.

The Deep Retrieval model performs end-to-end candidate retrieval with close to linear computational speed, despite the complexity associated with large-scale recommendations[13]. Being an end-to-end model, it is not limited by the two tower[14] constraint most models struggle with and can support the large-scale candidate generation increasingly demanded by the Internet industry.

Deep Retrieval end-to-end training aligns the learning of the user embeddings and the retrievable structure under the same objective function, direct from user-item interactions. Because embeddings make ML models more efficient and easier to work with, BytePlus Recommend supports the embedding of large numbers of real-time feature dimensions, Natural Language Processing (NLP) and Computer Vision (CV) features, and graph-based features.

BytePlus Recommend can process high data volume and efficiently trains models with a wide variety of parameters because of the unique ML algorithms used.

With different online businesses having vastly different recommendation needs, the BytePlus Recommend model architecture is customizable so that the solution is tailored to the host of recommendation scenarios unique to each business. This ensures that metrics like the click-through rates or conversion rates are optimized.

Understanding Natural Language Processing (NLP) and Computer Vision (CV)

NLP refers to how we understand the contents of text, including the contextual nuances of the language within them. The contextual information like product entity or user sentiment are extracted from various user and product information to better serve model training.

CV refers to how we understand multimedia content like product videos or images. This includes various methods to process, analyze, and extract relevant information like numerical or symbolic data from digital multimedia. It transforms visual data into useful features that can be used to improve the performance of the recommendations.

Machine Learning Platform Dedicated to Recommendation Systems

A strong machine learning platform provides faster, more versatile performance when training and serving models. The BytePlus Recommend platform supports advanced ML capabilities such as distributed training of data at any scale, as well as real-time feature engineering and serving.

Distributed training helps to ensure that the recommender can handle real-time training of up to a hundred million parameters and data. Despite the large amount of data present, fast training and prediction is possible on the platform's high functioning graphics processing unit (GPU) and tensor processing unit (TPU).

Features are extracted and computed from the collected data using BytePlus's proprietary feature-learning systems, Instance Profile Service (IPS) and Graph-based embedding to handle the complex computation logic and latency requirements that have traditionally challenged conventional extraction process. With the powerful platform solution, these features can be processed quickly in real-time, so that their value does not gradually diminish as the user's interest change over time to render the recommendation irrelevant.

Given the volatility of the e-commerce industry, it is important to have a reliable system that can sustain ever-changing profiles of a large user base[15]. Having IPS integrated in the ML platform optimizes BytePlus Recommend for high availability and performance, allowing it to become a one-stop service for storing and serving profile data.

Additionally, when paired with an up-scale optimizing compiler, BytePlus Recommend uses fewer resources and executes faster. This enables the fast serving of complex models which ensures that the recommendations provided to the users are accurate and effective. Model complexity in this case, is no longer a challenge but an opportunity.

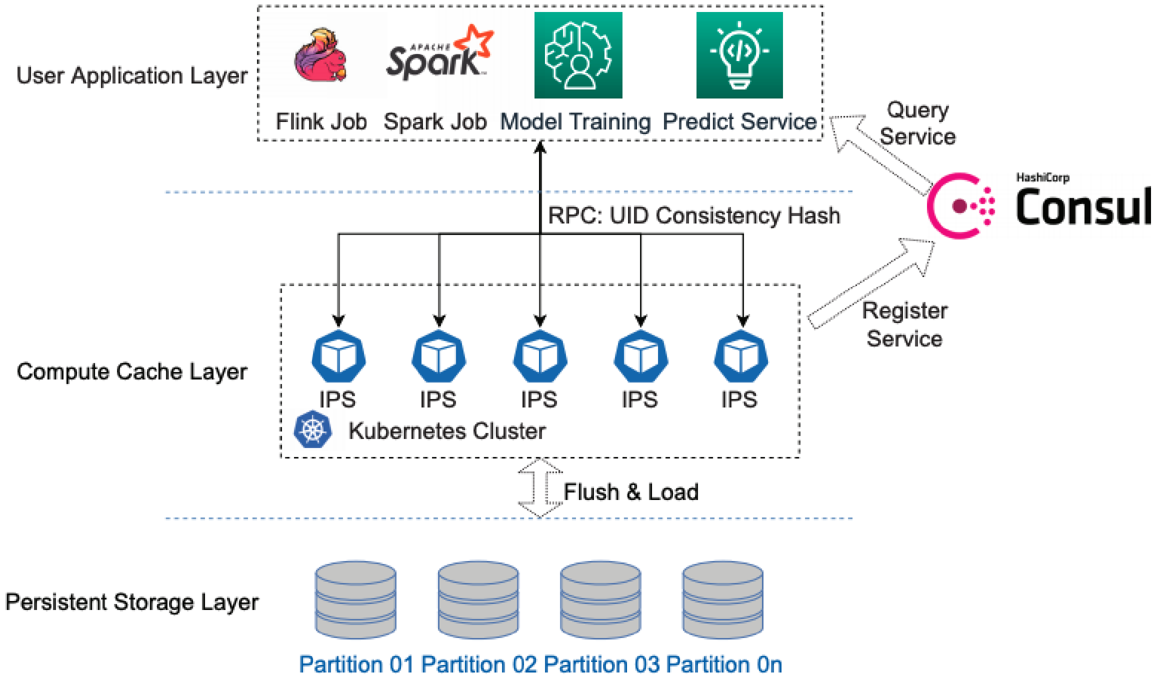


Fig. 4. High Level Architecture of IPS

Understanding Instance Profile Service (IPS)

IPS is a large-scale distributed system, developed in-house to manage unstructured profile data and perform real-time engineering.

Modelled after time, category, action and feature stats, this system can effectively detect and characterize a user's interests over time. All user profile data are stored as hashed literals along with strict privacy and access controls.

Using Instance data (impression, action, and features), the in-memory structures employed in IPS allow for flexible time window query and embedded multi-layer hash maps which support fast feature querying.

Deep Expertise for Highly Customizable Solutions

BytePlus Recommend is designed to provide a hyper- personalized experience, powered by large-scale ML technology. With a robust infrastructure and highly experienced ML experts across the world, BytePlus has both the network options to support extreme scalability—and the expertise to create recommendation solutions with lower knowledge barrier to benefit businesses.

The BytePlus Recommend team provides an integrative collaboration process for each client to ensure optimized model performance and to minimize unnecessary integration. By customizing each model to meet specific client needs, BytePlus Recommend supports all possible scenarios and an ultra-large-scale modelling.

Expertise in action

Data Collection

The BytePlus Recommend team provides clients with guidance on the types of data or specific fields that best suit their needs. This facilitates the running of the best performing model to achieve optimal results, and includes specific User Event data like impressions, scene behavior or product data like image URLs.

Backed by the rich experience in different use cases demanded by the market, the team can provide step-by-step guidance for myriad businesses to accurately and efficiently prepare their data for synchronization.

Data Validation

Checking the quality of data is one of the most overlooked areas when it comes to recommendation systems. Hence, BytePlus Recommend provides a streamlined approach to ensure that the data quality is up to standard after it has been synched. This includes checks on whether the data matches the schema, the data quantity and quality (empty rate or duplication rate) or if the data is satisfactory.

Advanced data validation techniques like verifying joining impressions and click data are used to ensure that accurate and sufficient input data are present for model training.

Performance Evaluation

BytePlus Recommend ensures that live A/B experiments are conducted with statistical rigor to verify things like the clickthrough rate (CTR) and conversion rate (CVR) metrics.

The in-house built A/B testing platform is optimized to maximize statistical significance and evaluation of algorithm performances. BytePlus conducts tests like the Z-test for conversion metrics, Welch's T-test for per-user metrics, and the Delta Method for click-through rate metrics. This allows for more comprehensive and accurate results.

Businesses can choose to split user traffic to compare BytePlus Recommend to a control group, whether in-house, on any existing model in the market, or by combining several at the same time. This ensures that the business is supplied with all the information and data needed for informed decision making, prior to investing in a recommendation system.

Common Use Cases

Explore how recommendation systems can be effectively deployed to drive business benefits. BytePlus Recommend is used as subject reference in the following use cases:

Use case 1: Improving User Engagement with Personalized Livestreaming Recommendations

A leading livestreaming platform in Japan wanted to improve user loyalty and decided to deploy BytePlus Recommend to significantly raise clickthrough rate (CTR) by 40% and stay time by 50%, as compared to another third-party solution previously used by the company.

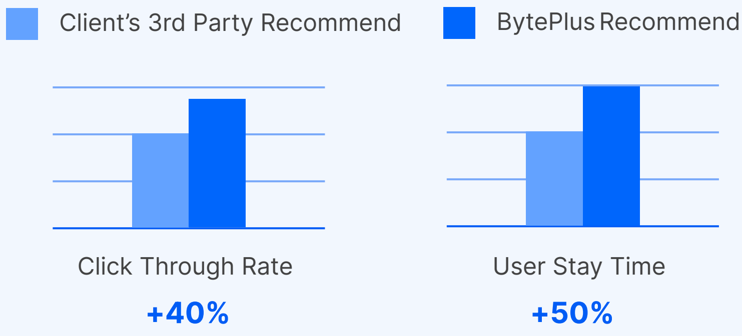
The Implementation

BytePlus Recommend identified 3 key areas to make recommendations more effective and to elevate user loyalty:

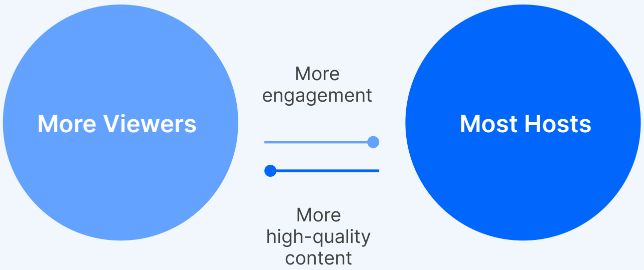
- **Data** – Collect meaningful data on users, videos, and their interactions in real-time — including the capture of real-time data changes for processing on the fly to prepare for model training.
- **ML Models** – Deploy models built for greater efficiency of livestream video recommendations.
- **Customization** – Enable fine-tuned recommendation results based on specific business logic such as de-duplication, filtering, and diversification. These are built on top of previous ML models to complete the final solution.

A/B Testing

To evaluate the BytePlus Recommend model, an A/B test was conducted over a period of 2 months. We observed a 40% CTR uplift and 50% stay time uplift, respectively.



This significant uplift also triggered a positive feedback loop —where more active users incentivize more hosts to provide more high-quality content—to result in even more user engagement.



Following the evaluation, BytePlus Recommend was rolled out to 100% user traffic.

Next steps

To ensure recommendation quality is maintained, the model has been designed for continuous improvement while in progress. This is supplemented by monthly review sessions between the two parties to enhance the monitoring process and collaboration efforts. Moving forward, talks about further collaboration on a new recommendation scenario has begun.

Use case 2: Delivering a Personalized E-Commerce User Experience

A leading full category e-commerce player in Indonesia aimed to provide exceptional user experience to drive increased user engagement and achieve long-term business success. The company's primary objective is to improve the relevancy of products recommended to users to increase total conversions.

The BytePlus team was able to provide best practices and enable high-quality data to be collected through a series of data synchronization and validation, which are then fed into advanced machine learning models. The collaboration resulted in an uplift of 50% for clicks, 30% for conversions and 25% for Gross Merchandise Value (GMV) contributed by recommendations.

The Implementation

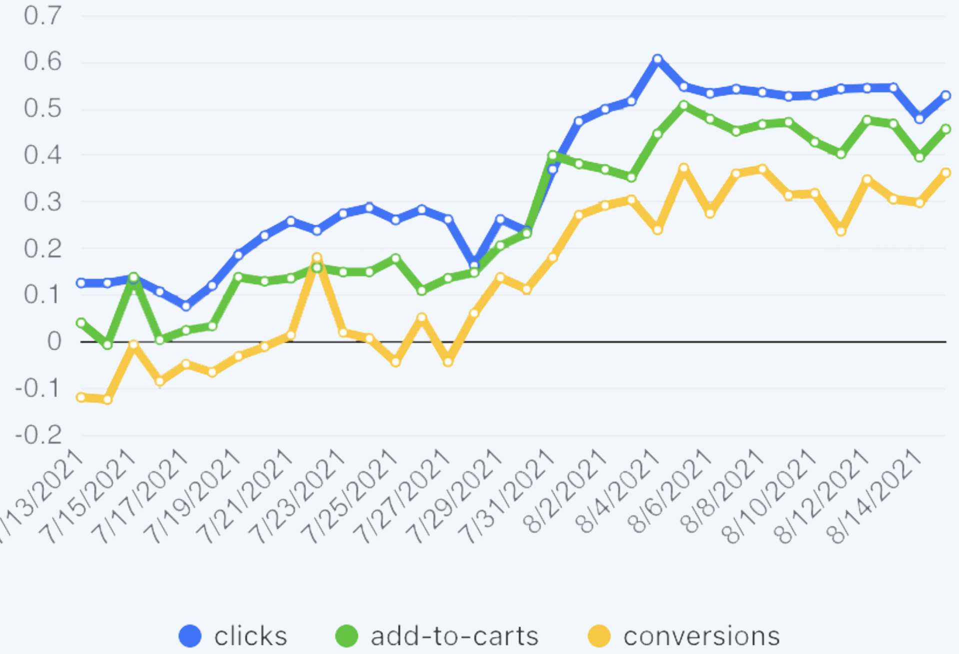
1. The need to improve the quality of impression data was one of the issues identified during the early stages of the collaboration. To ensure this demand was met, the BytePlus team offered streamlined data synchronization and validation services to guide data collection and quality verification. Given that fixing impression data requires additional resources and effort from the client, BytePlus team provided an alternative solution to collect server-side impression using AckServer Impressions API.

2. To address the lack of information available in product descriptions, the BytePlus team explored and leveraged text and image embedding techniques to achieve a positive uplift on product information availability.

3. The large amount of product and user behavior data was efficiently processed by the multiple candidate generators and models implemented by BytePlus. And different strategies were launched using A/B testing to yield the best performance.

A/B Testing

With personalized real-time recommendation, the BytePlus team witnessed remarkable success in product detail page recommendation. During the 6-week live A/B experiment period, the BytePlus team leveraged online feedback and iterated the model accordingly — the performance continues to.



As a result, BytePlus Recommend helped the e-commerce firm achieve a 50% improvement in clicks, 30% for conversions and 25% for GMV from product detail page recommendations. The company also successfully achieved their topline GMV with minimal resources.

Conclusion

In this whitepaper, we sought to explain how recommendation systems are being revolutionized by the implementation of deep learning-based models to increase their effectiveness. We looked at how recommendation systems work and offered insight into how the recommendation process has been stepped up by BytePlus's recommendation system.

Our deep dive into the holistic approach by BytePlus Recommend shed light on how best to implement a modern recommendation system, including how to tap BytePlus Recommend's expertise to adopt our recommender and subsequently, improve core business metrics.

If you are keen on applying BytePlus Recommend to your business, please reach out to our team through the following:

- Fill in the contact form here
- Follow/comment on the BytePlus LinkedIn page

About

ByteDance has been a pioneer in recommendation systems since its formation in 2012. A host of ByteDance products have adopted recommendations to improve their user experience, including TikTok, Douyin, Toutiao, and other apps.

BytePlus, as the enterprise tech brand of ByteDance, offers the recommendation solution to digital businesses to provide them with efficient solutions to build better user experience and grow their business.

BytePlus is committed to maximizing the impact and value of our recommendation systems to improve business performance.

References

1. Gutierrez, D. (2022). Netflix-Uses-Big-Data-to-Drive-Success - insideBIGDATA. Retrieved 17 January 2022, from *https://insidebigdata.com/2018/01/20/netflix-uses-big-data-drive-success/netflix-uses-big-data-to-drive-success/*

2. (2022). Retrieved 17 January 2022, from *https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers*

3. Keenan, T. (2014). Technocreep: The Surrender of Privacy and the Capitalization of Intimacy.

4. Ricci, F., Rokach, L., & Shapira, B. (2010). Introduction to recommender systems handbook. Recommender Systems Handbook, 1–35. *https://doi.org/10.1007/978-0-387-85820-3_1*

5. Zhang, S., Yao, L., Sun, A., & Tay, Y. (2020). Deep Learning based Recommender System. ACM Computing Surveys, 52(1), 1–38. *https://doi.org/10.1145/3285029*

6. Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube recommendations. Proceedings of the 10th ACM Conference on Recommender Systems. *https://doi.org/10.1145/2959100.2959190*

7. Google. (n.d.). Embeddings | machine learning crash course | google developers. Google. Retrieved January 17, 2022, from *https://developers.google.com/machine-learning/crash-course/embeddings/video-lecture*

8. Sergios Karagiannakos. (2021, May 20). An introduction to recommendation systems: An overview of machine and Deep Learning Architectures. AI Summer. Retrieved January 17, 2022, from *https://theaisummer.com/recommendation-systems/*

9. Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube recommendations. Proceedings of the 10th ACM Conference on Recommender Systems. *https://doi.org/10.1145/2959100.2959190*

10. Ruder, S. (2017, June 15). An overview of multi-task learning in Deep Neural Networks. arXiv.org. Retrieved January 17, 2022, from *https://arxiv.org/abs/1706.05098*

11. Guo, H., Tang, R., Ye, Y., Li, Z., & He, X. (2017, March 13). DeepFM: A factorization-machine based neural network for CTR prediction. arXiv.org. Retrieved January 17, 2022, from *https://arxiv.org/abs/1703.04247*

12. Lian, J., Zhou, X., Zhang, F., Chen, Z., Xie, X., & Sun, G. (2018, May 30). XDeepFM: Combining explicit and implicit feature interactions for recommender systems. arXiv.org. Retrieved January 17, 2022, from *https://arxiv.org/abs/1803.05170*

13. Gao, W., Fan, X., Wang, C., Sun, J., Jia, K., Xiao, W., Ding, R., Bin, X., Yang, H., & Liu, X. (n.d.). Deep retrieval: Learning a retrievable structure ... - arxiv. Retrieved January 17, 2022, from *https://arxiv.org/pdf/2007.07203.pdf*

14. Introducing tensorflow recommenders. The TensorFlow Blog. (n.d.). Retrieved January 17, 2022, from *https://blog.tensorflow.org/2020/09/introducing-tensorflow-recommenders.html*

15. Shi, R., Liu, Y., Chen, J., Zou, X., Chen, Y., Fan, M., Cai, Z., Zhang, G., Li, Z., & Liang, Y. (2021). IPS: Unified Profile Management for Ubiquitous Online recommendations. 2021 IEEE 37th International Conference on Data Engineering (ICDE). *https://doi.org/10.1109/icde51399.2021.00288*