informs
https://pubsonline.informs.org/journal/ijaa

INFORMS JOURNAL ON APPLIED ANALYTICS
Vol. 54, No. 3, May–June 2024, pp. 282–296
ISSN 2644-0865 (print), ISSN 2644-0873 (online)

Identifying Popular Products at an Early Stage of Sales Season for Apparel Industry

Jiayun Wang,[a] Shanshan Wu,[a,b,*] Qingwei Jin,[a] Yijun Wang,[b] Can Chen[b]

[a] School of Management, Zhejiang University, Hangzhou 310058, China; [b] LineZone Data Technology Co. Ltd., Hangzhou 310052, China
*Corresponding author

Contact: wangjiayun@zju.edu.cn, https://orcid.org/0009-0000-0400-0735 (JW); wushanshan1989533@gmail.com, https://orcid.org/0009-0001-9883-8914 (SW); qingweijin@zju.edu.cn, https://orcid.org/0000-0001-6599-8754 (QJ); wangyijun@linezonedata.com (YW); chencan@linezonedata.com (CC)

Received: April 4, 2023
Revised: August 7, 2023; October 14, 2023; October 26, 2023
Accepted: November 4, 2023
Published Online in Articles in Advance: December 29, 2023

https://doi.org/10.1287/inte.2023.0022

Abstract. The early phase of launching a new apparel product is critical for gaining insights of its performance and classifying it into different categories such as fast selling, average selling, and slow selling. This information is crucial for optimizing product management strategies and making decisions regarding inventory planning, pricing, and marketing. Many apparel companies rely on rule-based methods conducted by experienced sales managers, which consume significant time and energy from managers and often result in delayed information and low prediction accuracy. We propose a new ranking-based method to identify the product popularity that predicts regional and national rankings of products based on sales data at an early stage of a sales season. Our method enables companies to efficiently identify popular products within a remarkably short span of two to four weeks. To validate its efficacy, we compare the model's predictions with actual orders from a fashion company in 2021, showcasing a notable 5.9% increase in sales volume when using our approach to guide order decisions.

Funding: J. Wang's work is supported in part by the National Natural Science Foundation of China [Grants 71931009 and 72171212]. S. Wu's work is supported in part by the Key R&D Program of Zhejiang Bilateral Industry Joint R&D Plan Project [Grant 2023C04047]. Q. Jin's work is supported in part by the National Natural Science Foundation of China [Grants 72010107002, 71821002, and 72171212].

Keywords: applications • predictive analytics • ranking-based method

Introduction

The apparel industry is complicated and fast paced. The industry faces the challenge of a short product life cycle, in which new designs and collections are introduced frequently, necessitating a quick turnaround time for manufacturing and distribution to meet market demands. Considering the seasonality of the consumer demand, a year is usually divided into four sales seasons, which start in March, June, September, and December, respectively. A sales season usually lasts three to four months, and new products are released consecutively. Apparel companies offer a wide range of products during every sales season, and consumption pattern varies a lot in different regions. This requires apparel companies to maintain a deep understanding of consumer preference and market trends to effectively manage the supply chain to ensure timely production and distribution of products.

In the supply chain of an apparel company, a two-step supply strategy is typically adopted, consisting of long-term planning and short-term replenishment. Long-term planning involves ordering predetermined quantities of products before the start of the sales season, and short-term replenishment involves restocking fast-selling products during the sales season. Identifying popular products is particularly important for short-term replenishment, as managers often fail to accurately predict the number of units to order in the absence of product-level data before the sales season. By replenishing fast-selling products during the sales season, companies can respond quickly and effectively to customer demand and thereby maintain customer satisfaction and loyalty. To identify popular products, sales analysis is conducted during the introduction period of a newly released product, typically lasting two to three weeks, to categorize products into fast selling, average selling, and slow selling. Fast-selling products are those with fast selling speed and potential high revenue, which should be prioritized in inventory and marketing efforts to capitalize on their popularity. Average-selling products sell at normal speed with limited potential in revenue. Maintaining inventory levels for these products is important but not necessarily prioritizing them in marketing efforts. Slow-selling products refer to those that are not meeting expected sales goals within a certain time frame. To prevent losses and improve sales, it is crucial to closely monitor slow-selling products and take proactive measures, such as adopting price strategies or changing the product display position. This analysis provides valuable insights into customer preferences and enables managers to make informed decisions about which products to replenish during the sales season.

Our research team consists of LineZone Data Technology Co. Ltd. (LineZone Data) and the revenue management team in the School of Management at Zhejiang University. LineZone Data is an information technology (IT) company that provides supply chain management services for the apparel industry. We observed that many clients still use a rule-based method manually to identify product popularity, which can be both time-consuming and error-prone. With hundreds of products to manage in a sales season, analyzing popularity can be a tough task for sales managers. Additionally, the time delay due to manual operations often causes missed opportunities to maximize profits. These companies have expressed a strong need for an algorithm that can automatically determine the popularity of each product in an early phase from sales and inventory data, thus resolving the limitations posed by manual methods.

The main difficulties in our study on product popularity classification are as follows:

- **Describing how well a product sells:** A typical rule-based method uses the sell-through rate as an evaluation indicator, in which the supply condition is not taken into account. For instance, consider a product with a potential sales volume of 1,000 units and an initial order quantity of 10,000 units. In this scenario, even if the product is fast-selling, it may not be correctly identified as such when using the sell-through rate as the sole criterion. Similarly, sales volume is influenced by many factors, such as store traffic, number of stores with stock, inventory volume, and out-of-stock rate, which is a biased indicator. In addition, using few data to predict the whole sales volume accurately is usually unrealistic. Instead, knowing whether a product is popular is more important than knowing the inaccurate sales volume at the early stage. Therefore, we need to find an indicator that is affected by neither supply conditions nor the length of time a product has been on the market.
- **Lacking historical sales data:** Apparel products are usually short-life products. When a new product enters the market, it is crucial to determine its popularity as quickly as possible. However, there are often few sales data available at the beginning of the sales season, making it difficult to predict.
- **Coping with a large number of products:** Apparel companies typically introduce hundreds of new products during each sales season. We need to develop a robust and accurate method to identify the popularity of new products and ensure that the model is scalable to handle a large amount of data from companies with multiple stores.

In response to these challenges, we develop a model that learns sales characteristics of products from the same period in previous years and predicts the popularity of newly launched products. After a product is released, our model requires only two to four weeks to categorize it as fast-selling, average-selling, or slow-selling with a high precision rate. Precision rate is defined as the proportion of relevant instances among the retrieved instances (Ting 2010). This automated approach reduces classification errors and enables quick and efficient analysis of large data sets.

Our main contributions lie in two aspects. First, we propose a new indicator called *AW Sales* (average weekly sales in main sales period) to measure the popularity of a product. Unlike traditional measures that consider sales volume over the entire sales period, *AW Sales* only takes into account the periods during which the majority of sales occur. It also eliminates the differences in store traffic, number of stores with stock, promotions, and days on market, and it provides a more accurate representation of a product's popularity. Second, because we are more concerned with the relative relationship of sales among products rather than the exact quantity of sales volume, we use the ranking algorithm LambdaMART as the predicting algorithm in our product popularity classification model. This is the first time the ranking algorithm has been applied to the sales prediction field, and our numerical experiments demonstrate that it outperforms commonly used machine learning algorithms.

Organization of the Paper

The rest of the paper is organized as follows. The Related Works section reviews literature relevant to our problem. Then, the Predictive Analytics: Methods section outlines the framework of our product popularity classification model and details the use of the ranking model in sales ranking. The Model Evaluation and the Implementation and Results sections present the results in testing and deployment. Finally, we summarize the paper.

Related Works

Research on classifying the popularity of apparel products is limited, and demand forecasting in fashion retail is the closest related field to our problem. Demand forecasting is essential in fashion retail because it impacts all stages of the supply chain and directly affects inventory shortages and overstock rates, which can significantly impact retailers' profits. However, demand forecasting in fashion retail presents distinctive challenges, including rapidly changing consumer preferences, unpredictable demand, extremely short product life cycles, and a lack of historical data. As a result, demand forecasting in fashion retail is one of the most difficult forecasting problems in the industry (Nenni et al. 2013). Studies on sales forecasting in fashion retail can be divided into two groups based on whether they use data of new products.

Time series forecasting is the most commonly used method and includes many well-known models: exponential smoothing (Brown 1963), Holt-Winters model (Winters 1960), regression models (Papalexopoulos and Hesterberg 1990), and autoregressive integrated moving average (ARIMA) models (Box and Pierce 1970). In fact, these statistical methods are not well suited to the field of fashion retail because most time series methods require not only large historical data sets in parameter estimation but also experience of the user (Choi 2016). Moreover, time series can only extract linear relationships (Gutierrez et al. 2008), resulting in poor performance compared with more sophisticated methods such as machine learning. Extreme learning machine (ELM) has become a popular method for fashion sales forecasting since Sun et al. (2008) applied it to sales forecasting. ELM has faster processing time and higher generalization performance compared with gradient-based learning algorithms. To enhance prediction accuracy, hybrid models, like those proposed by Wong and Guo (2010) and Choi (2016), have emerged, combining ELM with other algorithms. However, these prediction methods only achieve acceptable accuracy levels when there are ample data available, typically after products have been on the market for more than one month.

Forecasting sales for new products without historical data are a challenging task, and few studies have been conducted in this area. One possible approach is to use sales data from similar products to predict the sales of new products. Most new product forecasting methods typically involve two steps: first, clustering old products based on their attributes and sales performance, and second, classifying new products to forecast their sales. Therefore, new product forecasting methods differ mainly in how to cluster products. The concept of new product forecasting was first introduced by Thomassey and Fiordaliso (2006), who used k-means to cluster the sales curves and then classified the new products according to the prototypes. Giri et al. (2019) used pictures of clothing and sales information to cluster the products. Baardman et al. (2018) proposed a cluster-while-regress approach based on the similarity of product features and sales patterns. However, Singh et al. (2019) claimed that there is a lack of similarity in the sales curves even for products with similar attributes.

As the fast-fashion industry continues to experience rapid growth, extensive research has been conducted to explore various aspects of this field. Notably, Smith and Côté (2022) introduced a test market protocol to enhance the accuracy of future demand predictions. Moreover, researchers have investigated other critical aspects, such as optimizing markdown prices, interstore transshipment, and inventory shipments (Mantrala and Rao 2001, Caro et al. 2010, Sung et al. 2017). Whereas each of these studies focuses on distinct supply chain stages, our paper addresses a pivotal aspect at an earlier stage of the process.

Existing prediction methods that do not incorporate new product data heavily depend on clustering techniques and require sufficient external information, such as image data, text data, and Google Trends, to achieve good performance. However, in our specific case, we only have access to sales and inventory data from historical and current season products. This limits our ability to use existing sales-forecasting methods. Our proposed model takes a different approach in the following ways:

- **Forecasting objective:** Unlike traditional sales-forecasting models that aim to minimize the gap between predicted and actual sales, our model's objective is to accurately categorize products into three groups—fast selling, average selling, or slow selling.
- **Utilization of sales data of old products**: Compared with the traditional approach of obtaining prototypes and categorizing new products by clustering historical sales data, our model takes a more direct approach. It learns the sales characteristics of products with different levels of popularity using historical sales data from the same period. Subsequently, it uses this model to classify products directly. Additionally, we predict the popularity levels of new products solely based on sales data, without relying on external attributes such as image information, text description information, and category.

Nenni et al. (2013) argued that fashion markets are complex open systems that frequently demonstrate high levels of "chaos," and it is unrealistic to forecast demand for fashion products accurately. However, we believe that, although sales volume is affected by various factors, the popularity of products is inherent and can be inferred from the sales data. Although we cannot provide a precise quantitative sales volume, we can provide qualitative results that enable sales managers to develop better replenishment, allocation, and marketing strategies.

Predictive Analytics: Methods

Table 1 depicts the rule-based method used by Company A, a client of LineZone Data, to categorize product popularity based on sell-through rate. The company is a well-known children's apparel company in China with an operating income of more than RMB 10 billion in 2022. Under this method, a product with a high sell-through rate in a short period is considered fast-selling, whereas a product with a low sell-through rate is

Table 1. Rule-Based Method

	Days on market					
Sell-through rate	<15 days	≥15 days	≥30 days	≥45 days	≥60 days	≥90 days
< 10%	–	Slow	Slow	Slow	Slow	Slow
≥ 10%	–	Slow	Slow	Slow	Slow	Slow
≥ 15%	–	Slow	Slow	Slow	Slow	Slow
≥ 20%	Average	Slow	Slow	Slow	Slow	Slow
≥ 25%	Average	Average	Slow	Slow	Slow	Slow
≥ 35%	Fast	Average	Average	Slow	Slow	Slow
≥ 45%	Fast	Fast	Average	Average	Slow	Slow
≥ 55%	Fast	Fast	Fast	Average	Slow	Slow
≥ 65%	Fast	Fast	Fast	Fast	Average	Slow
≥ 75%	Fast	Fast	Fast	Fast	Average	Average
≥ 80%	Fast	Fast	Fast	Fast	Fast	Fast

Note. "Days on market" refers to the number of days from the introduction of each product.

deemed slow-selling. However, this method has a flaw: The sell-through rate is closely related to the product's initial inventory. If a product has a large inventory at the beginning of the sales season, its sell-through rate may remain low even if it is actually popular. Therefore, in our study, we propose a new indicator, *AW Sales*, to describe the popularity of products more accurately, and we use it as the dependent variable in our ranking-based model. This method addresses the limitations of the previous rule-based approach and is more accurate and efficient.

Figure 1 shows the framework of the product popularity classification model, including data processing, model training, and predicting. In data processing, we

Figure 1. (Color online) Framework of the Popularity Classification Model

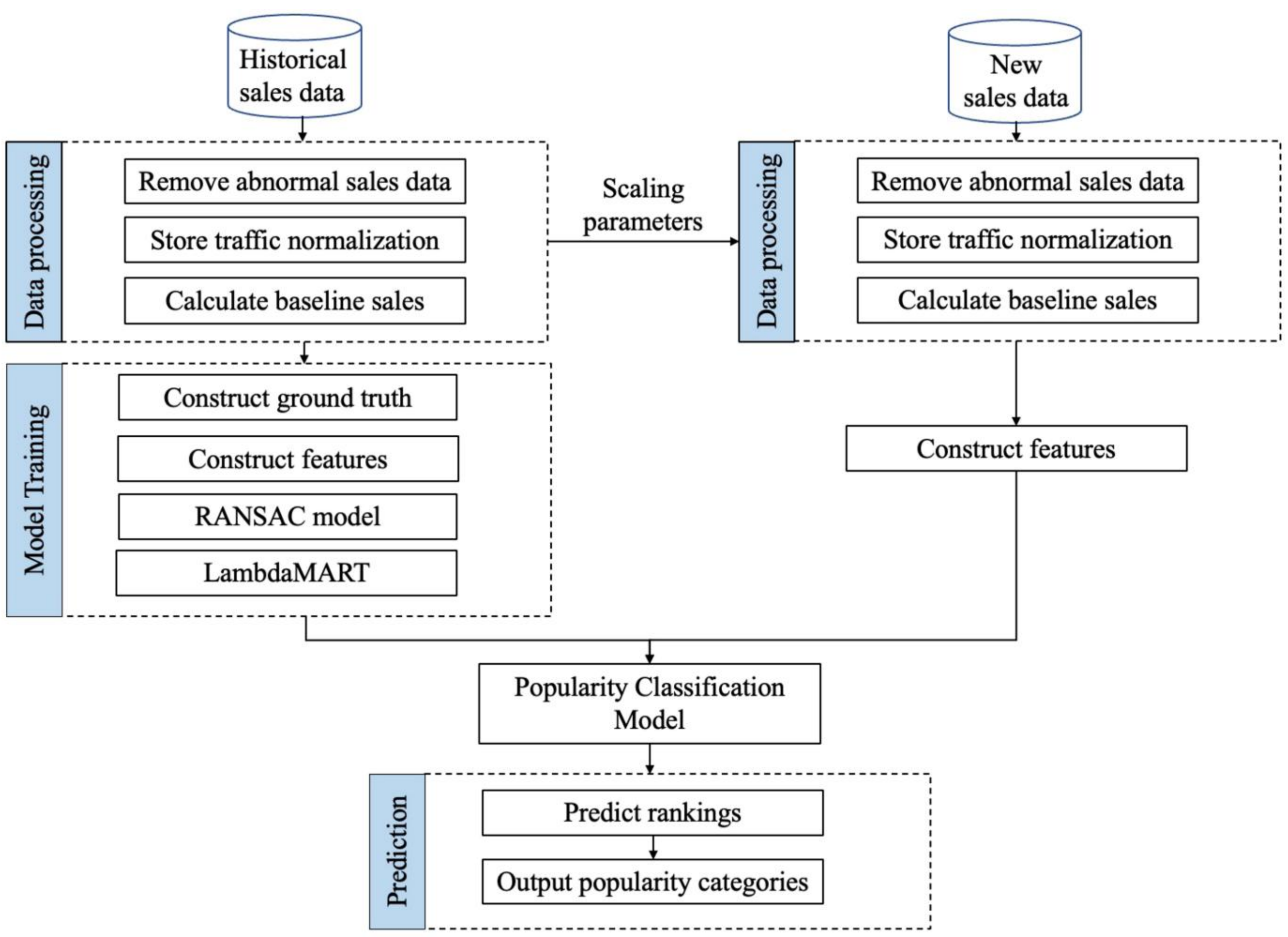


Table 2. Variables Used in the Popularity Classification Model

Variable	Comments
Mean sales	A measure of the regular performance of product.
Mean maximal sales	Captures unusual sales volume.
Active store ratio	Intended to evaluate the sales performance and market penetration of products.
First sale waiting days	Products that are sold soon after launching are considered as potential fast-selling products.
Mean waiting days	Products that are purchased multiple times over a period of time are considered potential fast-selling products.
Increasing subsequence length (*Length*)	A measure of potential fast-selling products; a longer subsequence indicates a stronger increasing trend.
Increasing subsequence range (*Range*)	A measure of the range of growth potential.
Increasing subsequence growth rate (*Growth Rate*)	A measure of potential sales trend.

remove sales that are not from individual buyers, normalize store traffic, and calculate baseline sales. In model training, we use the average weekly sales in the main sales period of each product at store level as the value to predict and construct features as shown in Table 2. We use the random sample consensus (RANSAC; Fischler and Bolles 1981) to estimate model interior points and remove outliers. Data of products from the same period in previous years are used as the training data so that we can get the model before sales season. After new products are launched, new sales data are obtained, and features are constructed in the same way after data processing. We feed features into our trained model to obtain regional ranking results. These regional rankings are then weighted to obtain national ranking, and products are classified as fast selling, average selling, and slow selling based on a preset threshold.

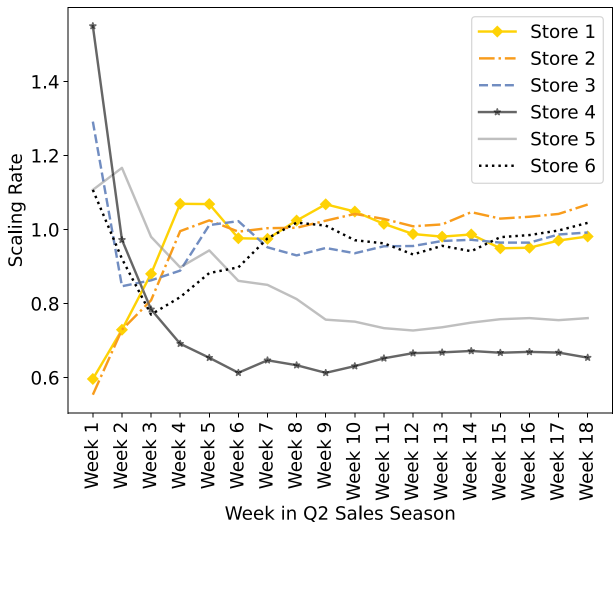
In the subsequent four sections, we will begin by providing insights into our data processing procedures, which encompass the normalization of store traffic and calculation of the baseline sales to determine sales unaffected by store traffic and promotions. Next, we will delve into the explanation of the dependent and independent variables used in our study. Last, we will elaborate on the rationale behind our selection of prediction algorithm. This discussion will shed light on the reasons and considerations that guide us in choosing the most appropriate algorithm for our model.

Data Processing: Store Traffic Normalization and Baseline Sales

In data processing, we remove sales that are not from individual buyers, including group sales data and returned sales data. To normalize store traffic, we adjust sales from different stores by normalizing the store traffic in each region to a standardized level. Last, we calculate the baseline sales volume using linear regression, which estimates the sales volume that would have occurred if no discounts were applied to the products.

In terms of normalization of total store traffic, the store scaling ratio is each store's total store traffic relative to the average total store traffic in the region. We then multiply each store's sales volume by its store scaling ratio to obtain normalized sales volume. Figure 2 illustrates the change of store scaling ratio in a sample region consisting of six stores during the Q2 sales season in 2021, showing that except for the first four weeks, the performance of stores remains relatively stable throughout the sales season. The observed instability in the store scaling ratio can be attributed to the gradual introduction of products over time, resulting in limited data availability during the early stages of the sales season, which hinders the accurate estimation of the store scaling ratio. To test the robustness of store scaling ratio in the first four weeks, a comparison test is conducted in which we utilize the fixed store scaling factor obtained at the 10th week of the sales season instead of daily calculations for both training and predicting. The results indicate that the precision rate exhibits only marginal variability, whereas the methodology using daily scaling factors exhibits higher recall rates compared with the use of fixed factors. Recall rate is

Figure 2. (Color online) Store Scaling Rate in a Sample Region



defined as fraction of relevant instances that are retrieved (Ting 2010). We posit that although the store scaling rate at the midpoint of the sales season provides a better representation of store traffic comparison, it may overlook the sales fluctuations of newly listed products, and it is better to use store scaling ratio calculated on a daily basis.

There is concern that the observed sales can be a censored measure of demand due to the occurrence of stockouts, wherein demand exceeds supply, resulting in unfulfilled transactions, leading to potential underestimation of market demand. Within the scope of our investigation, we scrutinize the occurrence of stockouts during the initial 28-day period and find them to be infrequent. This suggests that sales figures may not deviate much from true demand because of stockouts in our data sets. Nevertheless, in cases of frequent stockouts, there are methods available to mitigate their impact. One such approach involves using the average sales volume of the three days preceding and following the stockout period to replace the zero sales volume during the stockouts. This method helps reduce the influence of stockouts on the analysis and provides a more accurate representation of sales patterns. We emphasize the necessity of conducting a thorough assessment of stockout frequency before using our methodology to ensure appropriateness and validity.

Dependent Variable: Average Weekly Sales in Main Sales Period (*AW Sales*)

To accurately assess the popularity of products while accounting for the length of time a product has been launched, we introduce the concept of average weekly sales in main sales period (*AW Sales*). This metric provides a fairer and more robust measure of a product's popularity during the sales season.

Construction of *AW Sales*. *AW Sales* is calculated as the store's average weekly sales during its main sales period within a given region, using adjusted sales data. The adjusted sales data are obtained after data processing introduced in previous subsection. To identify the main sales period, which encompasses the time when the majority of sales occur, we propose a threshold-based method. The following steps are undertaken to determine the main sales period:

- Calculate the percentage of sales for each week.
- Sort the weekly sales percentages in descending order.
- Compute the cumulative percentage of weekly sales in descending order.
- Identify the weeks in which the cumulative percentage exceeds the threshold as the main sales period.
- If the number of weeks in the main sales period is less than the preset minimum number of weeks, additional weeks are included until the minimum is reached.

For instance, let us consider a product that has been sold for five weeks with sales percentages as follows: 15% (first week), 20% (second week), 25% (third week), 35% (fourth week), and 5% (fifth week). Assuming we set the threshold at 80%, the main sales period would consist of the second, third, and fourth weeks. We then use the average weekly sales during this main sales period as the indicator *AW Sales*.

Although the selected weeks may not be consecutive owing to fluctuations in sales volume, this does not pose an issue. Our primary concern is the sales performance during the period when most of the sales occur. Additionally, filtering out certain weeks using this method ensures a more accurate judgment of a product's popularity.

Independence from Influential Factors. *AW Sales* serves as a dependable indicator, remaining unaffected by various influential factors that could otherwise skew the assessment of a product's popularity. We illustrate its independence with respect to the following factors:

- **Differences in store traffic**: Store traffic can vary significantly across retail locations, leading to distorted sales volume figures. However, our normalization technique using the store scaling ratio eliminates the influence of store traffic fluctuations on the sales volume, making *AW Sales* independent of variations in store traffic.
- **Number of stores with initial stock**: Before the beginning of the sales season, products are divided into several assortments and different assortments are offered to different levels of stores. As a result, some products may be sold in a small proportion of stores, whereas others may be available in a larger number of stores, significantly influencing the total sales volume. *AW Sales* divides the total adjusted sales by the number of stores with initial stock and ensures that *AW Sales* is not affected by the number of stores with initial stock.
- **Discounts**: The application of discounts can significantly impact sales volume. We use linear regression to estimate baseline sales volume without any discounts to ensure that *AW Sales* is not affected by discounts.
- **Length of time the product has been launched (days on market)**: To avoid skewing average daily sales calculations because of the timing of product launches, our threshold-based method identifies the main sales period, encompassing the time when the majority of sales occur. Consequently, *AW Sales* is not influenced by the length of time a product has been launched.

By constructing *AW Sales* and demonstrating its independence from these relevant factors, we achieve a dependable indicator for accurately assessing the popularity of products during the sales season.

Independent Variable: Feature Description

The task of accurately forecasting sales volume at the beginning of a sales season is widely recognized to be a challenging task owing to various factors such as the seasonality of demand and the unpredictable nature of consumer behavior. However, analyzing sales data during the early period of the sales season can provide insights into the level of product popularity, which can guide marketing strategies to maximize profit. Our proposed model aims to capture sales characteristics of fast-selling, average-selling, and slow-selling products. The dependent variable in our model is the average weekly sales, termed the *AW Sales*, as mentioned earlier.

Table 2 displays variables employed in the popularity classification model. The independent variables are computed at the aggregate week-aggregate store level. The detailed calculation method of independent variables and the corresponding notations are provided in Tables A.1 and A.2 in Appendix A. The predictors in our model can be grouped into three categories. The first group comprises two variables designed to measure product popularity based on sales volume: *Mean sales*, which measures the regular sales volume of products, and *Mean maximal sales*, which captures unusual sales volume. The second group of predictors comprises the *Active store ratio* and *First sale waiting days*. The *Active store ratio* is the percentage of stores that generate sales, which relates to the range of stores that the product covers and the popularity of product. The *First sale waiting days* reflects the average time it takes for a new product to make its first sale in a set of stores. Shorter waiting time suggests higher demand and greater attractiveness to customers.

The third group of predictors are derived from the sequence of weekly adjusted sales volume. Given the inherent randomness in sales demand during this early phase, we propose a novel approach to capture trends by identifying the longest increasing subsequence in the sales data. Increasing subsequence is a sequence in which each number is greater than the preceding one and the longest increasing subsequence has the maximal length among all increasing subsequences of a given sequence. We use three measures to indicate the stability and the sales trend—that is, the length, range, and growth rate of the longest increasing subsequence. *Range* is the difference between the maximum and minimum values in the longest increasing subsequence. *Growth Rate* is the ratio of *Range* to *Length*. *Range* and *Growth Rate* capture explosive sales spikes, and *Length* indicates the stability of the rising sales trend. A higher value of *Length* indicates a more stable rising trend.

Prediction Model: Ranking-Based Algorithm

Whereas conventional machine learning methods focus on learning the relationship between features and sales to make accurate sales predictions, our primary interest lies in understanding product popularity and ranking rather than precise sales forecasts. Consequently, we have adopted a direct ranking approach that leverages the features to predict the rankings of products. This approach allows us to prioritize the relative popularity and positioning of products within our study, which aligns more closely with our objectives. By focusing on rankings, we can gain valuable insights into product performance and customer preferences, beyond traditional sales-forecasting methodologies.

The selection of the LambdaMART model is based on two key factors. First, this model places greater emphasis on the relative popularity of products rather than their exact sales values. Diverging from benchmark models that aim to minimize the discrepancy between predicted and actual sales quantities, the LambdaMART ranking model directly uses features to forecast product popularity rankings. Its gradient solely depends on the positions of products in the rankings, independent of the predicted values. Given our interest in discerning the relative sales relationships among products rather than their precise quantities, we believe this approach can yield more efficient and effective outcomes.

Second, the evaluation metric used by LambdaMART assigns higher weights to higher-ranked search results, thus emphasizing the importance of precisely identifying and ranking the most relevant and popular products. This prioritization is crucial, as it allows the model to accurately identify the most popular products, which can significantly impact profitability. By focusing on the accurate ranking of top items, the LambdaMART model optimizes its predictive performance in identifying fast-selling products, ultimately contributing to enhanced profit-maximizing strategies.

To select the most effective algorithm for our model, we test six popular algorithms: multiple linear regression (MLR; Olive 2017), support vector machine (SVM; Cortes and Vapnik 1995), eXtreme Gradient Boosting (XGB; Chen and Guestrin 2016), random forest (RF; Ho 1995), neural network (NN; Anderson 1995) with two layers, and the LambdaMART (Burges 2010) ranking model. The performances of the LambdaMART model and five benchmark models are evaluated by precision rate and recall rate to show results more intuitively. The precision rate represents the proportion of real fast-selling products among all the products predicted as fast-selling products, whereas the recall rate represents the proportion of real fast-selling products that are identified by the model. We describe more details of precision rate and recall rate in Appendix B.

We compare these models based on their precision rates and recall rates of fast-selling products, and the details are summarized in Table 3. Although SVM exhibits the highest prediction precision rates during the first month, its recall rate is only 26.7% after the new product launches for two weeks, indicating that it may

Table 3. Precision Rates and Recall Rates of Six Different Models in Identification of Fast-Selling Products

	Algorithm	14 days	21 days	1 month	2 months	3 months
Precision rate	LambdaMART	0.635	0.663	0.733	**0.841**	**0.895**
	MLR	0.632	0.569	0.604	0.759	0.770
	SVM	**0.720**	**0.729**	**0.762**	0.530	0.573
	XGB	0.594	0.680	0.717	0.747	0.706
	RF	0.557	0.576	0.604	0.632	0.661
	NN	0.611	0.648	0.724	0.766	0.732
Recall rate	LambdaMART	**0.794**	**0.794**	**0.809**	**0.779**	**0.784**
	MLR	0.629	0.634	0.673	0.733	0.776
	SVM	0.267	0.347	0.396	0.614	0.574
	XGB	0.564	0.515	0.351	0.733	0.713
	RF	0.728	0.564	0.619	0.535	0.589
	NN	0.708	0.520	0.559	0.728	0.721

Note. The bold numbers represent the largest value among six models, and the underlined numbers represent the second largest value.

miss many opportunities to identify fast-selling products. MLR demonstrates strong performance relative to machine learning methodologies like SVM, XGB, RF, and NN. However, it does not surpass the LambdaMART ranking model, which consistently achieves superior precision rates and recall rates on the test set. The LambdaMART ranking model has the best overall performance, achieving high precision rates and recall rates throughout the entire life cycle. Its recall rates are much higher than those of other models in the introduction stage, and its precision rates are significantly higher than those of other algorithms in the growth and maturity stage. Therefore, we select the LambdaMART ranking model as the prediction algorithm for our popularity classification model.

The relationship between features and the dependent variable in different regions may vary because of various factors such as cultural differences and economic conditions. As a result, it is necessary to train an independent model for each region. Nonetheless, this approach may result in a problem in which some models are trained with too few samples. To overcome this problem, we adopt the strategy used in document retrieval, in which a universal model is trained using samples with different search keywords to minimize the error in the ranking results for all queries, and this method has been realized in LambdaMART. Similarly, we train our model using samples from various regions. Just as in the realm of document retrieval, a typical sample consists of a query that is the search keyword and documents that need to be ranked. In our proposed model, we use region as the query and use data in each region as documents. By training our model in this way, we can develop a robust and accurate prediction model that can effectively forecast the popularity of products across various regions.

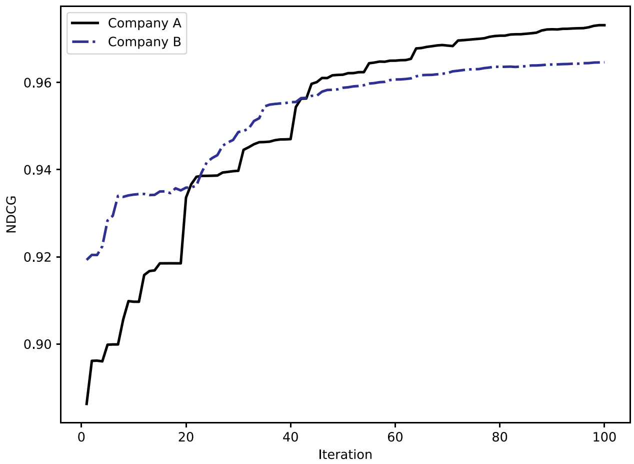
Model Evaluation

Our study uses data from two companies; owing to companies' concerns regarding disclosure of private information to competitors: we refer to them as Company A (the same as in the Predictive Analytics: Methods section) and Company B. Company A is a prominent apparel company in China, operating more than 2,000 stores across the mainland. During the Q2 sales season in 2021 and 2022, Company A released 1,485 and 1,568 products, respectively. Company B is a newly established apparel company in China specializing in young men's clothing. During the same sales season in 2021 and 2022, Company B released 240 and 547 products, respectively.

To evaluate the model performance on the training set, we use the normalized discounted cumulative gain (*NDCG*) metric, which is the evaluation metric employed by LambdaMART. A comprehensive explanation of *NDCG* is provided in Appendix C. *NDCG* is normalized to yield a score within the range of zero to one, with one denoting a perfect ranking. We use the sales data of the entire sales season to calculate the *AW Sales* of each product as ground truth and use the ranking as the ideal ranking. Our model converges after 100 iterations, and the *NDCG* score of the training set from Company A reaches 0.973, whereas the *NDCG* score of the training set from Company B reaches 0.965. This indicates that the ranking of products generated by our models on training set is highly aligned with the ideal ranking. The evolution of the *NDCG* scores for both Company A and Company B is illustrated in Figure 3.

To evaluate the performance of our model, illustrated in Figure 4, we implement a rolling evaluation approach. We select historical data matching the current duration to train the model and then utilize this trained model for predicting product performance. For example, when predicting the sales performance of products two weeks after the commencement of the sales season, we employ historical data from the corresponding two-week period after the start of sales seasons in previous quarters. To determine the predicted label for each product, we use a classification approach using a consistent threshold based on

Figure 3. (Color online) Normalized Discounted Cumulative Gain of Training Set



the rankings computed by our model. Following the general rules of the apparel industry, the top 20% of products are classified as the fast-selling products and the bottom 30% of products are classified as the slow-selling products. After calculating the *AW Sales*, we can assign each product a label as fast selling, average selling, or slow selling. These rankings are diligently calculated on a daily basis throughout the entire sales season, ensuring a robust and reliable evaluation process.

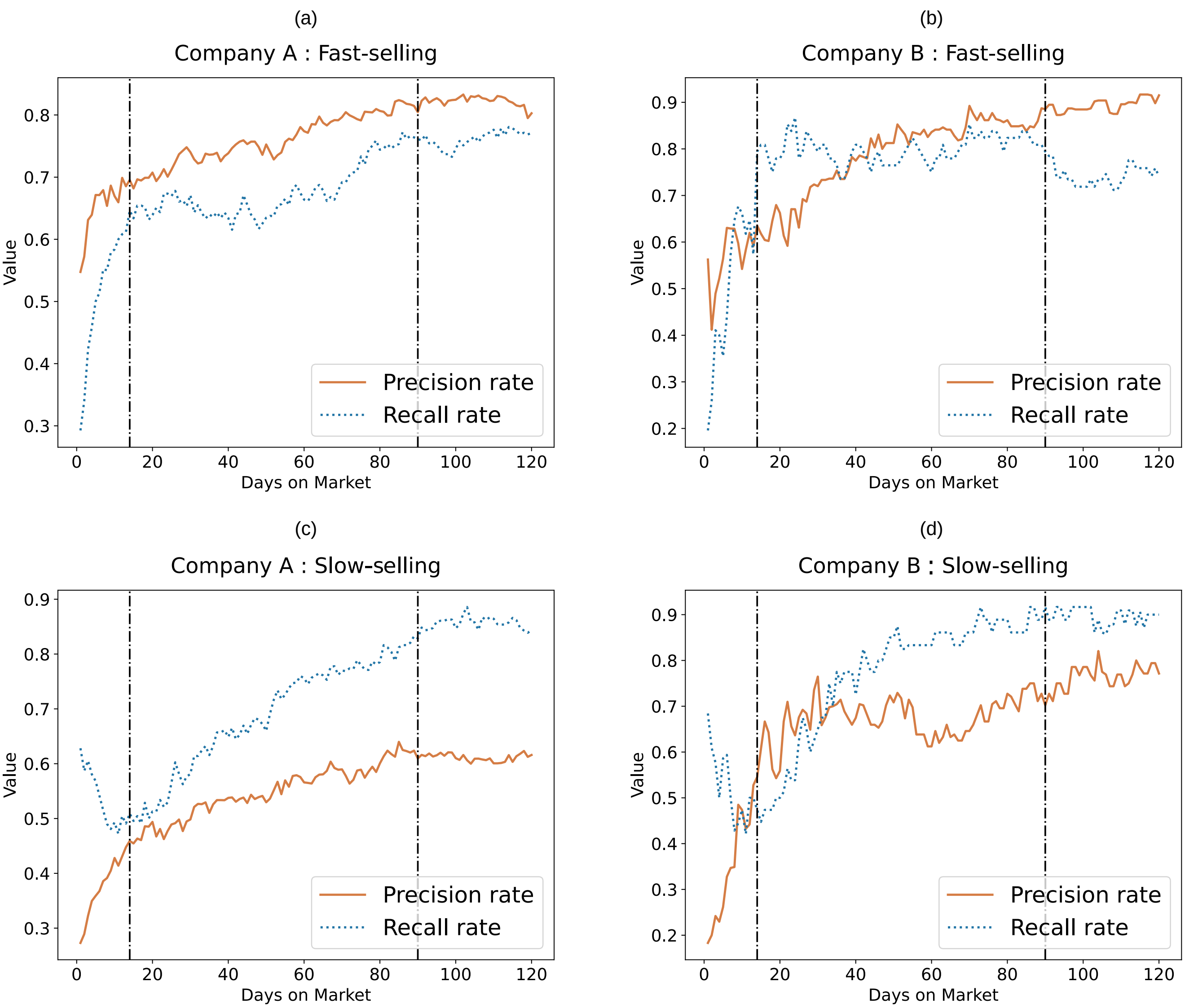
However, we acknowledge the potential for evaluation deviations due to the varied product launch dates within the sales season. We align the products by days on market and calculate the precision rates and recall rates. Figure 4 divides time into three periods: the introduction period (0–14 days), the growth and maturity period (14–90 days), and the decline period (greater than 90 days). Each period shows distinctive characteristics in performance. During the introduction period, the precision rate increases rapidly and achieves steady growth during the middle stage of the life cycle. After the middle stage, both the precision rate and recall rate are typically at a high level and show little fluctuation until the end of the life cycle. We observe that the model performs well in predicting fast-selling products, achieving high precision rates and recall rates after only two weeks of launch.

Specifically, when products are on the market for 14 days, products from Company A achieve a precision rate of 0.695 and a recall rate of 0.646, whereas products from Company B achieve a precision rate of 0.635 and a recall rate of 0.794. Additionally, the model maintains high precision rates throughout the middle and late stages of the life cycle. We then conduct a thorough examination on prediction errors.

Most classification errors occur at the boundaries of three categories—that is, predicting a product as fast selling or average selling or predicting a product as average selling or slow selling. Given the preset threshold, it is understandable that products near the boundaries are easy to wrongly classify, and actually, these errors may not be critical issues because the products at the boundaries can be classified as either category—their impact on the operation is not that significant. However, we focus on two other types of errors that can cause significant impact on the operation. The first type of error occurs when a fast-selling product is predicted as slow selling. The second type of error occurs when a slow-selling product is predicted as fast selling. We focus on these two errors because the former one can lead to profit loss, such as losing the opportunity to sell products in more stores because of failure to identify fast-selling products in time. The latter mistake misleads companies to place orders for slow-selling products, resulting in excess inventory at the end of sales season.

We use *Error* 1 to represent the proportion of products predicted as slow selling that are actually fast selling compared with the total number of real fast-selling products. Similarly, *Error* 2 represents the proportion of products predicted as fast selling that are actually slow selling compared with the total number of products predicted as fast-selling. We describe more details of error frequencies in Appendix B. Table 4 illustrates frequencies of both errors for Company A and Company B at different days on market. As the number of days a product has been on the market increases, the frequency of both errors decreases rapidly. Specifically, for Company A, 4.95% of the real fast-selling products are misclassified as slow-selling products, and 0.47% of the slow-selling products are misclassified as fast-selling products when products are launched for two weeks. After three months, frequencies of both errors are less than 0.30%. For Company B, no fast-selling products are misclassified as slow-selling products except within the first two weeks. Although the second type of error occurs at a slightly higher rate at the beginning of the sales season, no slow-selling products are classified as fast-selling products after two months. The infrequent occurrence of both errors indicates that our model is effective in predicting which products are likely to sell well and which products may not. Specifically, our model has a low likelihood of misclassifying slow-selling products as fast selling or vice versa. This suggests that our model is able to accurately distinguish between these two categories, reducing the risk of inefficient inventory management. Our model's performance in identifying slow-selling products is generally weaker than fast-selling products for two reasons. First, the preset threshold lacks clear boundaries among the three categories, leading to errors in classification. Second, our ranking model gives higher weight to correctly classify products with higher popularity rankings, making it more prone to misclassification of less popular products. Despite these limitations, our model

Figure 4. (Color online) Precision Rates and Recall Rates at Different Days on Market



Notes. The first vertical line on the graph represents the time point of 14 days on the market, and the second vertical line represents the time point of 90 days on the market. The leftmost portion of the graph corresponds to the introduction period, the middle portion corresponds to the growth and maturity period, and the rightmost portion corresponds to the decline period.

still provides valuable insights into product performance and can be used to optimize inventory management strategies.

Implementation and Results

We developed a software module after testing the proposed model. The module extracts input data from data warehouse using existing SQL scripts. Products that have been launched for more than seven days are added to the list for ranking, and the module ranks their popularity on a daily basis. In addition to outputting the popularity level of each product, the module conducts data analysis to provide sales managers with specific suggestions.

For fast-selling products, the module provides a comparison of its selling speed with the average selling speed of all products to support the forecasting results and generates a list of high-performing stores where the product is not available. This allows sales managers to identify sales trends and thus increase profits by selling the product in more stores. The module also considers related products that are often sold alongside fast-selling products and generates a list of stores where the related product is not available.

For slow-selling products, the module not only provides a comparison of their selling speed to the average selling speed of all products but also suggests discounts for slow-selling products. This helps ensure slow-selling

Table 4. Frequencies of Both Errors

	Company A		Company B	
Days on market	Error 1	Error 2	Error 1	Error 2
2 weeks	4.95%	0.50%	0	5.88%
1 month	3.33%	<0.30%	0	1.33%
2 months	2.37%	<0.30%	0	0
>2 Months	<0.30%	<0.30%	0	0

Notes. Error 1 refers to the proportion of products predicted as slow-selling that are actually fast-selling compared with the total number of real fast-selling products. Error 2 refers to the proportion of products predicted as fast-selling that are actually slow-selling compared with the total number of products predicted as fast-selling.

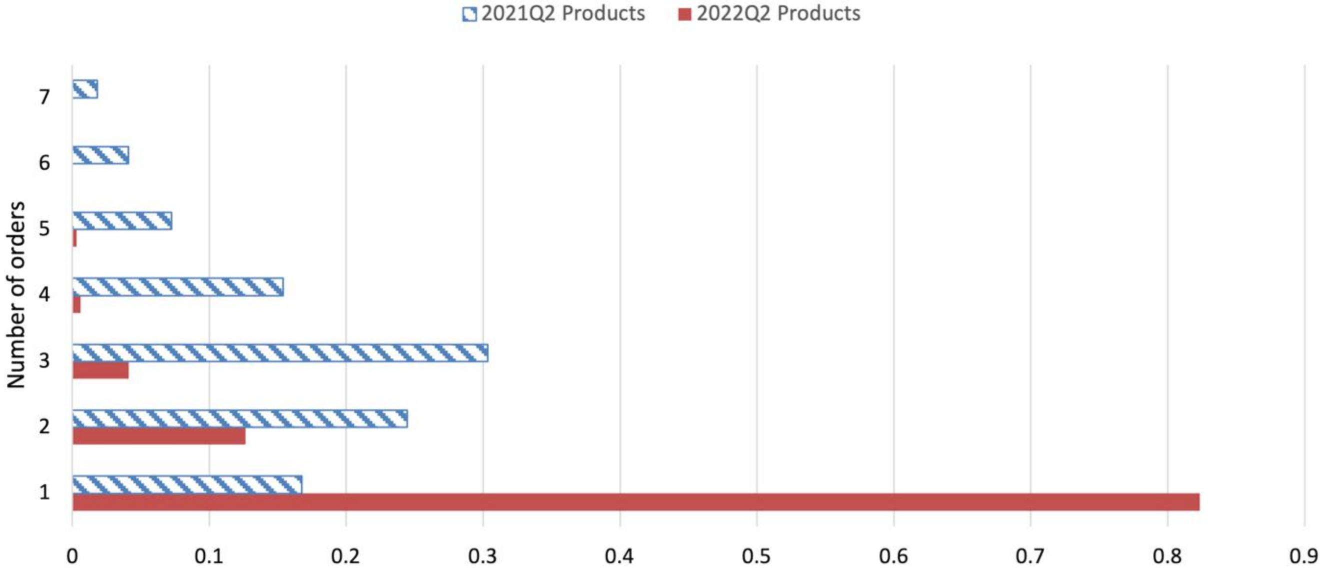
products reach high sell-through rates at the end of the sales season.

Company B, which subscribed to the module service in January 2022, observes an increase in quarterly revenue compared with the same period in 2021 and a decrease in out-of-stock rate compared with 2021. Company B's sales managers attribute this improvement to the use of the module. The remainder of this section quantifies the benefits generated by the module by comparing manual identification method with our module. The quantitative benefits of our module are as follows:

- **Decrease in the number of orders**: In 2021, before using the module, Company B adopted a strategy of placing multiple orders with small quantities to prevent stockouts. Once sales mangers observe growing sales trend for one product, they will place orders on that product. However, this strategy leads to high order cost and a high out-of-stock rate. As shown in Figure 5, the maximum number of orders placed on one product during the whole sales season is seven times, and 83% of products have had more than two orders placed, indicating a lack of accurate grasp of fast-selling products and confidence in prediction. This approach results in high order costs and inefficient inventory management. In 2022, sales managers in Company B used the module as a guide. As a result, the number of products with more than one order decreased by 66% compared with the same period in 2021. The maximum number of orders placed on one product dropped from seven to five. Although the number of orders decreased, the out-of-stock rate in the midstage of fast-selling products dropped from 15.3% to 4.6%, and at the end of the sales season, the sell-through rate had increased by 1.3% compared with the same period in 2021. Overall, the use of the module has improved Company B's inventory management efficiency and sales performance.

- **Increase in sell-through rate:** In practice, to ensure accuracy, we define the model identification date as the date when the model classifies the same popularity category for a product for three consecutive days. On average, a product receives a repeat order 26 days after it has been introduced to the market. Comparing the identification date of the module with the real order date, we discover that our module identifies those fast-

Figure 5. (Color online) Distribution of the Number of Orders



selling products with repeat orders 17 days earlier than rule-based methods. This early period of time allows the companies to adopt new order strategies before the sales season. Companies can order smaller quantities before the sales season and place repeat orders after fast-selling products have been identified by the module, which can significantly increase the sell-through rate at the end of the season and alleviate the company's financial burden during the stocking period. To measure the impact of this approach, we analyze Company B's sales and inventory data from 2021. Assuming a disposal cost of 10% of the tag price at the end of the season, we reduce Company B's order quantities by two weeks before the sales season. We observe an increase in sell-through rate from 50.88% to 85.53% at the end of the season in 2021 and a 74.4% reduction in disposal costs. This approach could potentially save Company B millions of yuan renminbi in disposal costs and significantly reduce storage and logistics costs.

Company A manages two to three times more products than Company B and introduces products gradually during the sales season. By identifying fast-selling products earlier, sales managers in Company A can order more of these products and introduce them to more stores earlier to maximize profits. We test 84 products from the Q2 sales season in 2021 that have records of decision date. We assume that, when inventory is sufficient, fast-selling, average-selling, and slow-selling products will be sold at their respective selling speed. In cases where inventory is insufficient, the selling speed will be limited by the inventory. Before the sales season, products are stocked under the assumption that all products are selling at average selling speed. This means that if a fast-selling product is not identified as such, it will continue to sell at average selling speed. Once a fast-selling product is identified, it will be sold at fast selling speed in more stores. If an average-selling or slow-selling product is mistakenly identified as a fast-selling product, it will be stocked under the assumption that it is a fast-selling product and sold in more stores. However, these products will still sell at their respective selling speed, resulting in lost profits due to the wrong decision.

Table 5 compares our model's predictions with the decisions made by managers during the Q2 sales season in 2021. In Scenario 1, the model identifies fast-selling products earlier than the rule-based method, allowing the products to be sold in more stores earlier, resulting in an increase in sales volume during the early period. In Scenario 3, the sales managers fail to identify a fast-selling product, but our model correctly identifies it. As a result, we can sell the fast-selling product at the fast selling speed in more stores, resulting in an increase in sales volume until the end of the sales season. In Scenario 8, the model misclassifies and thus orders average-selling or slow-selling products at quantities assuming they would be sold at the fast selling speed in more stores, resulting in the loss of excess units that exceed the selling speed. If the sales managers place orders completely according to the model's predictions, the sales volume can be improved by 39,558 units, accounting for 5.9% of actual total sale volume.

Conclusion

This study presents a novel approach to identify fast-selling apparel products at an early stage in the sales season using machine learning techniques. In our test, the model identifies fast-selling products 17 days earlier than the rule-based method, resulting in improved sell-through rates and increased profits. By comparing the model's predictions with real orders from a fashion company in 2021, we demonstrate a 5.9% increase in sales volume when using our approach to make order decisions.

Our popularity classification model frees managers from repetitive, complicated, and inefficient data analysis, allowing them to focus on the underlying factors that influence product performance. A software

Table 5. Comparisons in Sales Volume

Category	Scenario	Model	Manual	Comparison	Difference (units)
Fast-selling	1	✓	✓	Model identifies earlier than real, gains extra units during early period.	38,876
	2	✓	✓	Model identifies later than real, loses extra units during delay period.	−10,636
	3	✓	✗	Model gains extra units until end of sales season.	222,182
	4	✗	✓	Model loses extra units until end of sales season.	−180,078
	5	✗	✗	–	0
Average-selling or slow-selling	6	✓	✓	–	0
	7	✓	✗	Model gains extra until end of sales season.	13,738
	8	✗	✓	Model loses extra profit until end of sales season.	−43,277
	9	✗	✓	–	0
	10	✗	✗	Model identifies earlier than real, loses extra units in early period.	−1,247

Note. ✓, correct identification of products; ✗, incorrect identification.

module has been developed based on the popularity classification model, which helps companies reduce inventory levels before the sales season, identify fast-selling products at the beginning of the sales season, and suggest discounts for slow-selling products at the end of the sales season.

Using machine learning algorithms, sales managers in apparel companies can make informed decisions. Our model is estimated to generate millions of yuan renminbi in profit for LineZone Data's clients by reducing disposal costs and increasing potential sales volume. At present, the current module classifies products into different popularity levels and conducts basic data analysis. In light of the generality of our method, it is important to emphasize that our approach is not restricted solely to apparel items. As our methods focus on learning the sales patterns of products, we firmly believe that their application can be extended to various product types, encompassing mobile phones, fashion blind-box items, and other products characterized by unpredictable customer preferences and relatively short life spans. In the future, we hope to develop a more comprehensive module that integrates with inventory management systems to provide precise order quantities for companies in various domains.

Acknowledgments

The authors thank the chief editor Prof. Newman, anonymous associate editor, and anonymous reviewers for constructive comments that greatly improve this study.

Appendix A. Summary of Mathematical Formulations

Table A.1. Mathematical Notations

Notation	Definition
Indices and parameters	
m	The index of week.
n	The index of store.
k	The index of product.
K	Number of products to be ranked.
N_k	Number of stores with an initial stock of product k.
M_k	Number of weeks since the launch of product k.
s_{mnk}	The adjusted sales volume of product k in store n during week m.
x_{mnk}	1 if store n sells any quantity of product k in week m, and 0 otherwise.
D_{nk}	Length of the vector $\boldsymbol{d}_{nk}$.
L_k	Length of the vector $\boldsymbol{\ell}_k$.
Vectors	
$\boldsymbol{d}_{nk}$	$\boldsymbol{d}_{nk}$ is a sequence of days that represent the ordering times of orders of product k in store n. As an illustration, for product k in store n, suppose that we record sales on day 1 (one order), day 5 (three orders), and day 7 (two orders) after its initial launching. The vector $\boldsymbol{d}_{nk}$ would be denoted as (1, 5, 5, 5, 7, 7). Therefore, $\boldsymbol{d}_{nk}[0]$ represents the number of days before the first sale of product k in store n since its launching. Subsequently, the difference between $\boldsymbol{d}_{nk}[i+1]$ and $\boldsymbol{d}_{nk}[i]$ represents the number of days that elapsed between two consecutive orders of product k in store n.
$\boldsymbol{\ell}_k$	$\boldsymbol{\ell}_k$ refers to the longest increasing subsequence derived from the sequence of "weekly adjusted sales volume" of product k across all stores within a typical region. To illustrate, consider product k with a weekly adjusted sales volume sequence of (50, 162, 364, 285, 398, 488) in a region. In this case, $\boldsymbol{\ell}_k$ would be (50, 162, 364, 398, 488).

Note. Here, $\boldsymbol{d}_{nk}[i]$ and $\boldsymbol{\ell}_k[i]$ denote the $(i+1)$th elements of $\boldsymbol{d}_{nk}$ and $\boldsymbol{\ell}_k$, respectively.

Table A.2. Definitions of Variables Used in the Popularity Classification Model

Variable	Definition
Mean sales	$\sum_{n=1}^{N_k} \frac{\sum_{m=1}^{M_k} s_{mnk}}{M_k} \cdot \frac{1}{N_k}$
Mean maximal sales	$\sum_{n=1}^{N_k} \max_m(s_{mnk}) \cdot \frac{1}{N_k}$
Active store ratio	$\sum_{n=1}^{N_k} \frac{\sum_{m=1}^{M_k} x_{mnk}}{M_k} \cdot \frac{1}{N_k}$
First sale waiting days	$\sum_{n=1}^{N_k} \boldsymbol{d}_{nk}[0] \cdot \frac{1}{N_k}$
Mean waiting days	$\sum_{n=1}^{N_k} \sum_{i=0}^{D_{nk}-1} \frac{\boldsymbol{d}_{nk}[i+1] - \boldsymbol{d}_{nk}[i]}{D_{nk}-1} \cdot \frac{1}{N_k}$
Increasing subsequence length (*Length*)	L_k
Increasing subsequence range (*Range*)	$\boldsymbol{\ell}_k[L_k - 1] - \boldsymbol{\ell}_k[0]$
Increasing subsequence growth rate (*Growth Rate*)	$\frac{\boldsymbol{\ell}_k[L_k - 1] - \boldsymbol{\ell}_k[0]}{L_k}$

Appendix B. Forecast Performance Measures

This section describes performance measures used in this study: precision rate, recall rate, and error frequencies.

B.1. Precision Rate and Recall Rate

Precision rate and recall rate are commonly used in machine learning to evaluate the performance of classification models.

We compute the precision rate and recall rate as follows:

$$Precision\ rate = \frac{NFF}{NFF + NAF + NSF}, \tag{B.1}$$

$$Recall\ rate = \frac{NFF}{NFF + NFA + NFS}. \tag{B.2}$$

Here, *NFF* is the number of real fast-selling products that are predicted as fast-selling, *NAF* is the number of average-selling products that are predicted as fast-selling, *NSF* is the number of slow-selling products that are predicted as fast-selling, *NFA* is the number of fast-selling products that are predicted as average-selling, and *NFS* is the number of fast-selling products that are predicted as slow-selling.

B.2. Error Frequencies

The error frequencies are defined as follows:

$$Error\ 1 = \frac{NFS}{NF}, \tag{B.3}$$

$$Error\ 2 = \frac{NSF}{NFF + NAF + NSF}. \tag{B.4}$$

Here, *NFS* is the number of real fast-selling products that are predicted as slow-selling, *NF* is the number of real fast-selling products and *NSF* is the number of slow-selling products that are predicted as fast-selling.

Appendix C. Model Training Details

The evaluation metric used by LambdaMART is the normalized discounted cumulative gain (*NDCG*; Burges 2010). *NDCG* places significant emphasis on accurately ranking top items in a given query. For a specific query, *NDCG* is defined as the discounted cumulative gain (*DCG*) divided by the ideal discounted cumulative gain (*IDCG*).

$$DCG = \sum_{i=1}^{K} \frac{2^{aws_i} - 1}{log(1+i)} \tag{C.1}$$

$$NDCG = \frac{DCG}{IDCG} \tag{C.2}$$

In our context, K represents the number of products to be ranked, and aws_i denotes the value of *AW Sales* of the product at position i. The term *IDCG* refers to the ideal discounted cumulative gain, which represents the *DCG* value of the ideal order. The ideal order can be obtained from the *AW Sales* calculated based on the entire sales season. By comparing the model's *DCG* with the ideal order, *NDCG* evaluates the model's ability to accurately rank the most relevant results.

In our study, we perform grid search to fine-tune the parameters of LambdaMART. Among the various parameters, we find that *n_estimators* and *max_depth* have the most significant impact on the model's performance. The parameter *n_estimators* denotes the number of trees to use in the ensemble. Through systematic experimentation, we determine that setting *n_estimators* to 100 yields favorable results in terms of predictive accuracy and generalization capability. This choice strikes a balance between model complexity and performance, ensuring robust and efficient predictions. Similarly, the parameter *max_depth* represents the maximum depth of the trees in the ensemble. After rigorous analysis, we observe that setting *max_depth* to three leads to optimal model performance. By limiting the depth of the trees, we aim to prevent overfitting while capturing meaningful patterns and relationships in the data, thus enhancing the model's generality.

Furthermore, LambdaMART's feature of "hot-starting" greatly enhances the model's applicability in our context.

This capability allows the model to continue training on top of an already trained version after a sales season concludes. Consequently, the model can be updated or fine-tuned using data points from the previous sales season without losing the knowledge obtained during its initial training. This adaptability ensures that the model remains well suited for incremental learning, effectively capturing changes and trends in sales patterns over time.

References

Anderson JA (1995) *An Introduction to Neural Networks* (MIT Press, Cambridge, MA).

Baardman L, Levin I, Perakis G, Singhvi D, et al. (2018) Leveraging comparables for new product sales forecasting. *Production Oper. Management* 27(12):2340–2343.

Box GE, Pierce DA (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* 65(332):1509–1526.

Brown RG (1963) *Smoothing, Forecasting and Prediction of Discrete Time Series* (Prentice-Hall, Upper Saddle River, NJ).

Burges CJ (2010) From RankNet to LambdaRank to LambdaMart: An overview. *Learning* 11(23–581):81.

Caro F, Gallien J, Díaz M, García J, Corredoira JM, Montes M, Ramos JA, et al. (2010) Zara uses operations research to reengineer its global distribution process. *Interfaces* 40(1):71–84.

Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. Krishnapuram B, Shah M, eds. *Proc. 22nd ACM SIGKDD Internat. Conf. Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York), 785–794.

Choi TMJ (2016) *Information Systems for the Fashion and Apparel Industry* (Woodhead Publishing, Sawston, UK).

Cortes C, Vapnik V (1995) Support-vector networks. *Machine Learn.* 20(3):273–297.

Fischler MA, Bolles RC (1981) Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM* 24(6):381–395.

Giri C, Thomassey S, Balkow J, Zeng X (2019) Forecasting new apparel sales using deep learning and nonlinear neural network regression. Le Moullec Y, Shi J, eds. *Proc. Internat. Conf. Engrg. Sci. and Industrial Appl.* (IEEE, Piscataway, NJ), 1–6.

Gutierrez RS, Solis AO, Mukhopadhyay S (2008) Lumpy demand forecasting using neural networks. *Internat. J. Production Econom.* 111(2):409–420.

Ho TK (1995) Random decision forests. Abe K, Ebrahimi T, eds. *Proc. 3rd Internat. Conf. Document Analysis and Recognition*, vol. 1 (IEEE, New York), 278–282.

Mantrala MK, Rao S (2001) A decision-support system that helps retailers decide order quantities and markdowns for fashion goods. *Interfaces* 31(3 supplement):S146–S165.

Nenni ME, Giustiniano L, Pirolo L (2013) Demand forecasting in the fashion industry: A review. *Internat. J. Engrg. Business Management* 5(2013):5–36.

Olive DJ (2017) *Multiple Linear Regression* (Springer, Berlin).

Papalexopoulos AD, Hesterberg TC (1990) A regression-based approach to short-term system load forecasting. *IEEE Trans. Power Systems* 5(4):1535–1547.

Singh PK, Gupta Y, Jha N, Rajan A (2019) Fashion retail: Forecasting demand for new items. Preprint, submitted June 27, https://arxiv.org/abs/1907.01960.

Smith MA, Côté MJ (2022) Predictive analytics improves sales forecasts for a pop-up retailer. *INFORMS J. Appl. Analytics* 52(4): 379–389.

Sun ZL, Choi TM, Au KF, Yu Y (2008) Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems* 46(1):411–419.

Sung SW, Jang YJ, Kim JH, Lee J (2017) Business analytics for streamlined assort packing and distribution of fashion goods at kolon sport. *Interfaces* 47(6):555–573.

Thomassey S, Fiordaliso A (2006) A hybrid sales forecasting system based on clustering and decision trees. *Decision Support Systems* 42(1):408–421.

Ting KM (2010) Precision and recall. Sammut C, Webb GI, eds. *Encyclopedia of Machine Learning* (Springer US, Boston), 781–781.

Winters PR (1960) Forecasting sales by exponentially weighted moving averages. *Management Sci.* 6(3):324–342.

Wong WK, Guo Z (2010) A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm. *Internat. J. Production Econom.* 128(2):614–624.

Jiayun Wang is an MS student of management science and engineering at the School of Management, Zhejiang University. Her research is centered on the practical applications of data science and operations research in various real-world contexts such as inventory management and transportation.

Shanshan Wu is currently a director of customer success at Hangzhou LineZone Data Technology Co. Ltd. She received her PhD in computer science from Zhejiang University. In the past several years, her research has focused on the application of machine learning in supply chain management for apparel companies.

Qingwei Jin is an associate professor of management science in the School of Management, Zhejiang University. He received his PhD in industrial and systems engineering at North Carolina State University in 2011. His research interests include revenue management, retail operations management, and robust optimization methods.

Yijun Wang is currently the chief executive officer of Hangzhou LineZone Data Technology Co. Ltd. She received her PhD in supply chain management and marketing science from Rutgers University. She previously worked at Merrill Lynch. She has years of experience in modeling and optimization in e-commerce, finance, and the retailing industry.

Can Chen is currently the chairman of Hangzhou LineZone Data Technology Co. Ltd. He received his PhD in management science and information systems from Rutgers University. He previously worked at Merrill Lynch and assisted AT&T in data analytics. His current research focuses on data mining and modeling in the apparel industry.