Dual-blend insight recommendation system for e-commerce recommendations and enhance personalization

Sinzy Silvester[1], Shaji Kurain[2]
[1]Department of Computer Applications, IFIM College, Bengaluru, India
[2]Department of Computer Applications, Jagdish Sheth School of Management, Bengaluru, India

Article Info

Article history:

Received Nov 23, 2023
Revised Jan 22, 2024
Accepted Feb 4, 2024

Keywords:

E-commerce
Hit Rate
Mean reciprocal rate
Recommendation system
TMall

ABSTRACT

E-commerce, short for electronic commerce, refers to the buying and selling of goods and services over the internet. This digital transaction model has revolutionized the way businesses operate and consumers shop. In response to the burgeoning complexity of e-commerce datasets, this work addresses the need for advanced recommendation systems. This work introduces the dual-blend insight recommendation system (DIRS) model for personalized e-commerce recommendation system. The DIRS model involves dataset loading, preprocessing, and feature extraction, enabling training with recurrent neural network (RNN) and Bayesian personalized ranking (BPR) models. Recommendations are generated based on user-defined functions, i.e., location and session, and evaluation metrics such as hit rate (HR) and mean reciprocal rate (MRR) highlight DIRS's superior performance. The model is evaluated using the Tmall dataset. Results reveal DIRS consistently outperforms alternative algorithms, showcasing its effectiveness in 10k and 20k recommendation sets. This study provides valuable insights into optimizing e-commerce recommendations, emphasizing DIRS as a powerful model for enhancing user experience and engagement.

Corresponding Author:

Sinzy Silvester
Department of Computer Applications, IFIM College
Bengaluru, India
Email: sinzy1983@gmail.com

1. INTRODUCTION

E-commerce, the electronic buying and selling of goods and services over the internet, has witnessed a significant surge in popularity and usage in recent years [1]. The convenience of online shopping, coupled with a vast array of products available at one's fingertips, has led to a paradigm shift in consumer behaviour. The growth of e-commerce is propelled by various factors, including the widespread availability of high-speed internet, increased smartphone penetration, and evolving consumer preferences [2]. People now rely on e-commerce platforms not only for traditional retail purchases but also for services ranging from food delivery to digital subscriptions. One key aspect that has contributed to the success of e-commerce is the implementation of recommendation systems (RSs) [3]. These systems leverage algorithms [4], [5] to analyse user behaviour and preferences, offering personalized product suggestions. By studying a user's past interactions, purchases, and even the behaviour of similar users, RSs enhance the overall shopping experience [6].

However, challenges arise when RSs attempt to cater to users based on their session and location. User preferences can vary depending on the time of day or the context in which they are shopping [7]. Additionally, regional preferences and availability of certain products may influence recommendations. Providing a balance between real-time personalization and broader user trends is a complex task [8].

The cold start problem is another issue in e-commerce RSs [9]. This occurs when a system struggles to provide accurate suggestions for new users or items with limited historical data. Without sufficient information on a user's preferences, the system faces challenges in delivering relevant recommendations, potentially impacting the user's initial experience on the platform [10]. In conclusion, the rise of e-commerce is closely intertwined with the implementation of RSs, which plays a pivotal role in enhancing user engagement and satisfaction. While these systems bring personalized recommendations to users, they also bring challenges related to session-based and location-based preferences, as well as the cold start problem. Overcoming these hurdles is essential for e-commerce platforms to continually improve the accuracy and effectiveness of their RSs, ensuring a seamless and tailored shopping experience for users. Hence, the contribution of this work is as follows:

– This work involves a comprehensive exploration of the factors that impact the buying intention of consumer durables among Indian e-shoppers, with a specific focus on the location-based and session-based recommendation.
– This work entails the development of a robust model called dual-blend insight recommendation system (DIRS) that vividly illustrates how the identified factors intricately influence customers preferences in the context of purchasing consumer durables.
– This work involves addressing the cold-start issue in the RS using the time-centric transitional-matrix.
– This work involves evaluating the model using the Tmall dataset. The evaluation of the proposed work has been done on the basis of hit rate (HR) and mean reciprocal rate (MRR) and compared with existing works.

The manuscript is organized in the following way. In section 2, the literature survey has been discussed to understand factors that impact the buying intention of consumer durables among Indian E-shoppers. Further, in section 3, the DIRS model is discussed. In section 4, the proposed work is evaluated using Tmall dataset where the evaluation is done in terms of HR and MRR and compared with existing works. Finally, in section 5, the conclusion and future work of the work is presented.

2. LITERATURE SURVEY

Maher *et al.* [11], graph neural network (GNN), recurrent-neural-network (RNN), attention-based networks (ABN), pattern-matching (PM) and k-nearest-neighbors (KNN) are some of the deep learning (DL) methods that have been studied in this study. According to the findings, RNN, ABN and KNN methods perform better than other methods when it comes to making recommendations. Liu [12], presented two algorithms, first a short-term memory (STM) and short-term attention-memory priority (STAMP) algorithm for recommending products. They have used the YooChoose and Diginetica benchmarked datasets to test the presented algorithms. Through an analysis of both dataset's outcomes, this research's experimental findings show that the proposed algorithmic method helps to disclose preferences between different users, hence boosting the RSs overall accuracy. Maneeroj and Sritrakool [13] proposed a unified sequential RS which they have termed as personalized-preference drift-aware (PPD+) [14]. Utilizing soft labels, this PPD+ technique determines the optimal number of variables to employ in its classifications. They employed the UserBehaviour, GoodReads, and MovieLens datasets to test their technique. The findings demonstrate that, in comparison to the state-of-the-art methodologies, the proposed approach successfully clusters related items based on the user's past activities, hence yielding more reliable recommendations. Salampasis *et al.* [15], this work evaluated different RS approaches such as recomendation systems utilizing embedding, hybrid methods, item-to-vecor (Item2Vec) embedding, document-to-vector (Doc2Vec) embedding, long-short term-memory (LSTM), node-to-vector (Node2Vec), purchase intent task, and graphs. The RNN-LSTM approach outperformed the state-of-the-art techniques, highlighting its ability to produce more useful suggestions. Esmeli *et al.* [16], they suggested an architecture according to which a user's interests are first identified by a particular kind of ranking structure, followed by a RS algorithm is presented. They have employed YooChoose and RecSys to test the effectiveness associated with the presented work. The findings demonstrate that the performance of SBRS is enhanced by taking into account user actions including contextual elements during different sessions.

Gao and Li [17], presented a parallel-rule extraction approach that is built upon the Apriori method. For evaluation, they had sampled 1,000 individuals to test the effectiveness of the suggested method. It was clear from the findings that both of the original and updated semantics emotion recommendation approaches are useful for improving the quality of recommendations. Dang *et al.* [18], they employed two different types of hybrid sentiment classification algorithms to determine how reviewers feel about certain items and products. This study evaluated their approach with two Amazon datasets: movie reviews and food reviews. On both datasets, the offered sentiment-based technique outperformed previously available rating-based techniques in terms of the reliability of its recommendations. Elahi *et al.* [19], used the bidirectional encoder representations from transformers (BERT) and principal component analysis (PCA) in their proposed approach. The BERT helped to extract embeddings for the contextual sentence and PCA helped to reduce the dimensionality. They utilized two Amazon datasets-the music dataset and the games dataset-to test their approach. The results of the tests showed that there was connection among the ratings users gave and the emotions they expressed. Karthik and Ganapathy [20], they have introduced a product RS that employs fuzzy-recommendation logic as well as fuzzy-rules in ontological matching to provide accurate results. Experimental findings show that the suggested RS is more accurate at anticipating customer preferences including product preferences than the state-of-the-art approaches. As proposed in Liu and Ding [21], a syntactic-data-inquiring scheme (SDIS) is proposed for e-commerce sentiment-based RS. The suggested system improves the user's experience through the e-commerce website by putting the needs of the user first. The Amazon dataset was used for the evaluation of this approach. Experiment results show that the suggested SDIS improves suggestion efficacy by 15.1%, data analysis ratio by 9.41%, and modification rate by 17% as session frequency rises.

Sreenivasa and Nirmala [22] proposed hybrid location-centric prediction (HLCP) hybrid RS is meant to take into account both immediate and future needs. They utilized the Tmall dataset to test how well their RS works. Experimental findings demonstrate that the utilized HLCP RS has a higher MRR and HR than the state-of-the-art hybrid RS (RNN and KNN). Xu and Sang [23], offered a RS for an e-commerce purchasing platform which combines several different kinds of individualized recommendation algorithms. Random forest (RF), gradient boosting (GB), extreme gradient boosting (XGBoost), and decision tree (DT) are some of the algorithms used. The Tmall dataset was used to evaluate this work. Findings reveal that the model that incorporates XGBoost, GBDT, and LR proves to be the most successful, as it achieves the highest overall score in the testing. Yin *et al.* [24] presents the location-based CNN system for smartphone advertisement RS in e-commerce, that leverages information about user's whereabouts as input. This work's experimentation makes use of data collected during the 2015 Alibaba Group Smartphone Recommendation Algorithm Competition. The overall recall performance of location-based CNN was 8.14%, which is more than two percentage points greater than the standard approach. Xu and Wang [25], this research proposes the mobile computational working method. An online store administration system, complete with its 50,000 products and 2,000 customers, was selected as the evaluation's test subject. The testing findings show that the suggested intelligent RS for e-commerce products works well and yields accurate recommendations.

Lu and Zhang [26], mentioned seven attributes which will attract a given customer in the e-commerce. In the seven attributes, one of the most important attributes is the gender. They have found that the different genders tend to buy different products and the rate of online orders is more by the female genders. Ahouf and Lu [27], they tested the different web designs on five hundred thirty-two online shoppers. The results were finally evaluated using the structural equation modeling (SEM). After the analysis, they have come to a conclusion that the web design plays an important role for attracting the different genders. Sidlauskiene *et al.* [28], the gender and age were taken into consideration for all the analysis. This work analyzed three studies, one using a pre-test where 135 individuals participated and two experimentation studies where 180 and 237 participants participated respectively. The findings show that utilizing the chatbot, the individuals were able to get their related products. Hirt *et al.* [29], they built and released a tool that can accurately determine the gender of a sizable subset of German-speaking individuals who use Twitter from information such as their usernames, tweeting, and avatars. The research results have the possibility to increase the monetary value of consumer analytics and aid in the creation of more effective RSs by expanding their knowledge of customer demographics. Melchiorre *et al.* [30], they evaluated Bayesian personalized ranking (BPR) and alternative least square (ALS), two methods of matrix factorizing, and contrast their respective outcomes. When comparing the two methods RecGap, it was clear that BPR gives the most equitable outcomes compared to the customized models, whereas ALS's outcomes are much extremely biased. The results of these tests show that the matrix factoring approach has no noticeable impact on the impartiality of RSs. This study sheds light on the problems of unfairness and bias in music recommendation algorithms.

3. PROPOSED METHOD

3.1. Block diagram

In Figure 1, the block diagram outlines a comprehensive process of the DIRS model for leveraging the Tmall e-commerce dataset. The initial step involves loading the dataset, followed by a crucial phase of data preprocessing. During this phase, consumer, product, session, location, and behavior information are extracted, utilizing timestamps, and subsequently split into training and testing datasets. The next stage focuses on extracting consumer multi-behavior characteristics based on location and session-centric transitional matrices. This information is then used for training using both RNN and BPR models. The combination of these models enhances the system's ability to capture nuanced user behaviors. Following model training, the system proceeds to perform recommendations based on a user-defined function, such as generating top-10k and top-20k recommendation sets. Finally, the recommendation performance is evaluated using metrics like HR and MRR, providing valuable insights into the effectiveness of the implemented RS.

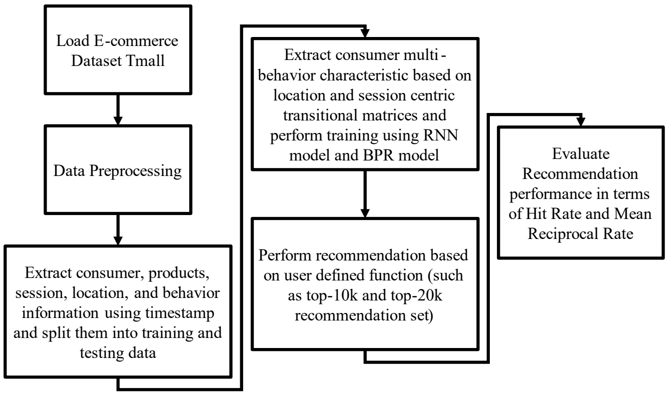

Figure 1. Block diagram of proposed work

3.2. System model

This work presents a DIRS for e-commerce which considers session and location factors of the consumer as the main factors for recommending a product. The DIRS was modelled to recommend products from the different sessions a consumer uses and location of the consumers for the given current session. In order to establish an association among the consumers present-session, present-location, past-session and past-location for the purpose of product recommendations, it is crucial to develop a comprehensive approach that takes into account both long-term and short-term session circumstances. The approach, known as the time-centric graph-positional attention-network (GPAN) approach [31], allowed more accurate understanding of the consumer’s preferences along with sessions (patterns) over time. Let us take into account a set of products, such as mobiles, and consumer durables, for the purpose of our analysis. This can be denoted using in (1). The variable $\mathcal{u}_n$ represents the overall products under consideration. Further, the set of consumers encompassing various demographic categories such as males, females, and children, and others can be represented using in (2). The variable $\mathcal{v}_n$ represents the overall count of consumers under consideration. This study examines the e-commerce application, which encompasses different sessions (actions) which is represented using in (3).

$$\mathcal{U} = \{\mathcal{u}_1, \mathcal{u}_2, \dots, \mathcal{u}_n\} \tag{1}$$

$$\mathcal{V} = \{\mathcal{v}_1, \mathcal{v}_2, \dots, \mathcal{v}_n\}. \tag{2}$$

$$\mathcal{C} = \{c_1, c_2, c_3, c_4\}. \tag{3}$$

The variable c_1 represents the click session, which refers to the sequence of user interactions with a website or application. It captures the specific pages or elements that a consumer has clicked on during their browsing session. On the other hand, c_2 represents the list of products that have been added to the cart. This variable tracks the products or services that the user has selected for potential purchase but has not yet completed the transaction. Moving on, c_3 represents the consumer's preferred list, which records the products that the consumer has marked or saved as their personal favorites. This variable allows users to easily access and revisit their preferred products in the future. Lastly, c_4 represents the purchased products. This variable indicates the products or services that the consumer has successfully bought or completed the transaction for. It serves as a record of the consumer’s past purchases. The objective at hand entails the anticipation of consumers purchasing behavior within the ongoing session, employing the proposed DIRS model.

3.3. Modelling of short-term and long-term sessions

To model the short-term and long-term session, the DIRS model utilizes the RNN structure. The structure is presented in Figure 2. The RNN structure is characterized by its composition of numerous hidden-layers. The RNN structure consists of various interconnected layers, including an input, output and various hidden-layers. Additionally, the RNN structure incorporates internal weight matrix structures, which play a crucial role throughout the network's functioning. The activation variable associated with the hidden-layers is acquired through the following equation

$$i_\ell^{v} = f\left(\mathcal{X} i_\ell^{v} + \mathcal{D} s_{w_\ell^{v}}\right), \tag{4}$$

The variable i_ℓ^{v}, which belongs to the set $\mathbb{S}^e$, represents the hidden-layers of consumer v at location l within a particular characteristic. Similarly, the variable $s_{w_\ell^{v}}$, also belonging to the set $\mathbb{S}^e$, represents the layer of the ℓ^{th} input product of consumer v. The activation variable, denoted as $f(i)$, is a mathematical representation used to model the behavior of a neuron or node in a neural network. It defines how the input signal, denoted as i, is transformed into an output signal. The transitional matrix, on the other hand, is a mathematical matrix that represents the relationships between the elements of a given system. In the context of the current product, it describes the transitions or transformations that occur within the system. The $\mathcal{D}$ in (5) has the capability to acquire and analyze data pertaining to the current session of consumers, while $\mathcal{X}$ possesses the ability to share signals indicating different characteristics. The previous session of consumer is depicted using in (6). In this study, the term adaptable-size represented as o is employed to refer to the overall input products considered or utilized in every layer of DIRS. The (7) depicts the change in probabilities based on location-centric data.

$$\mathcal{D} \in \mathbb{S}^e \tag{5}$$

$$\mathcal{W} \in \mathbb{S}^{e.} \tag{6}$$

$$\mathcal{D}_j \in \mathbb{S}^{e*e} \tag{7}$$

Where, $\mathcal{D}_j$ is the matrix which contains the location-centric data and belongs to the set of $\mathbb{S}^{e*e}$. In (4) is repeatedly implemented in order to calculate the state of every consumer location with respect to the session characteristics. The hidden-layer that constitutes the RNN exhibits dynamic characteristics in relation to session characteristics, wherein the fundamental structure demonstrates repetitiveness. The RNN encounters challenges when it comes to acquiring knowledge about immediate contextual information in session characteristics. In order to successfully acquire short-term session characteristic, the DIRS utilizes a feedforward-neural-network (FNN) that incorporates a single linear hidden-layer. The architectural depiction of the FNN is presented in Figure 3.

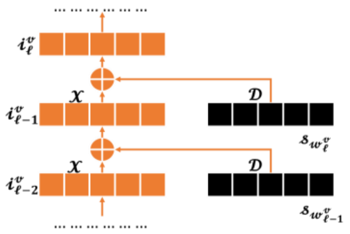

Figure 2. Modelling of short-term and long-term session

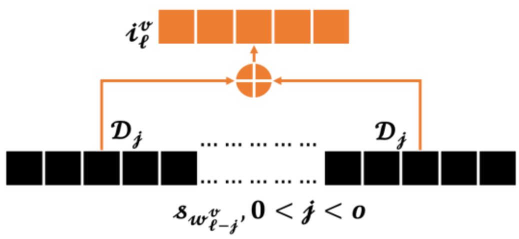

Figure 3. FNN Architecture used in DIRS

The extraction of session characteristic in a continuous session-window involves the utilization of clicked products and their corresponding transitional matrix structures within every session. The subsequent location is possible to be approximated through the utilization of a linear forecasting approach, as expressed by (8).

$$i_\ell^{v} = \sum_{j=0}^{o-1} \mathcal{D}_j s_{w_{\ell-j}^{v}} \tag{8}$$

Where, $\mathcal{D}_j$ is the location-centric transitional-matrix which contains the location-centric data and belongs to the set of $\mathbb{S}^{e*e}$ having different session characteristic. o refers to the overall input products considered or utilized in every layer of DIRS having the location and session characteristic. In order to effectively capture the dynamic session of consumers, the DIRS model employs session-based matrix structures as a means of extracting characteristics pertaining to various types of sessions. The estimation of the consumer v at location ℓ is conducted through the utilization of the subsequent in (9).

$$i_\ell^v = \mathcal{X} i_\ell^v + \sum_{j=0}^{o-1} \mathcal{D}_j \mathcal{N}_{c_{\ell-j}^v r_{\ell-j}^v} \quad (9)$$

The $\mathcal{N}_{c_{\ell-j}^v} \in \mathbb{S}^{e*e}$ represents a session-based transitional matrix approach that pertains to the session observed with respect to the j^{th} product of consumer v. The cold-start issue is easily and effectively mitigated by taking into account the initial condition where the value of v is equal to v_0. One crucial aspect to consider in this context pertains to the presence or absence of multiple types of sessions. If there exist just a singular form of session, it follows that the utilization of session-based matrix which becomes redundant and can therefore be removed. The DIRS model has the capability to extract the intrinsic characteristics pertaining to various session patterns exhibited in historical characteristics. Moreover, the determination of whether consumer v would participate in session c on product w during session sequence $\ell + 1$ can be achieved through computational analysis using in (10). The variable t_ℓ^v in (10) represents a visual illustration denoting the present location of the consumer v at the location ℓ. The fixed hidden-layer v_v belongs to the set $\mathbb{S}^e$ and encompasses the dynamic illustration i_ℓ^v.

$$z_{v,\ell+1,c,w} = (t_\ell^v)^{\mathcal{U}} \mathcal{N}_c s_w = (i_\ell^v + v_v)^{\mathcal{U}} \mathcal{N}_c s_w \quad (10)$$

3.4. DIRS

The current state of sequence-based models often fails to adequately account for the inherent variability in never ending session characteristics, commonly referred to as session-variation (SV), that exists between input characteristic sets. The utilization of the SV characteristic in predicting is highly advantageous, as it allows for the assessment of short-term session-window variations, which typically exerts a more determined influence on future purchasing session compared to long-term session variations. Let us examine two products, namely w_a and w_b, that have been recorded in a consumer’s purchasing cart. The consumer has recently acquired product w_a during midnight, while product w_b, was obtained several weeks ago. There exists a significant probability that upcoming purchasing decisions of consumers can be influenced by the presence of product w_a. Alternatively, if product w_b is bought during the preceding day, it is highly likely that both product w_a and w_b exerts a similar influence on the consumer's decision-making, owing to their shared appeal in the immediate context. Moreover, it is worth noting that the buying habits for particular products, like soaps and shampoo, exhibits regular patterns, typically occurring on a monthly basis. Consequently, the influence of session variations in these circumstances leads to a heightened level of dynamism. In light of the aforementioned conditions, the GPAN approach [31] is enhanced by integrating session variations information and using the DIRS architecture which has been presented in the Figure 4.

The DIRS architecture incorporates session and location characteristics for predicting what a consumer will choose. This architecture aims to provide accurate and personalized recommendations by considering the temporal aspect of consumer preferences and product popularity. The DIRS can forecast future consumer sessions and tailor recommendations accordingly. The integration of location and session characteristic enhances the system’s ability to adapt for recommendation. This study examines the phenomenon of consumer preference alteration in response to relocation to various regions or locations. To understand the customer preferences, the DIRS model utilizes the HLCP model [22]. Nevertheless, it remains crucial to take into account the inclusion of information pertaining to session variations. Hence, the DIRS model replaces location-centric transitional-matrix with session-centric transitional matrix and location-centric transitional matrix. The DIRS model is depicted in Figure 4. According to the findings presented in Figure 4, the determination of the location ℓ for a specific consumer v is carried out using in the (11) computational procedure.

$$i_\ell^v = \mathcal{X} i_{\ell-o}^v + \sum_{j=0}^{o-1} \mathcal{U}_{u_\ell^v - u_{\ell-1}^v} \mathcal{S}_{w_{\ell-j}^v} \quad (11)$$

The variable u_ℓ^v represents the current time (session) in the context of our study. On the other hand, $u_{\ell-1}^v$ represents the time at which every product for each layer of the DIRS is observed. Additionally, $\mathcal{U}_{u_\ell^v - u_{\ell-1}^v}$ represents the time-centric transitional-matrix with respect to the time variation $u_{\ell-1}^v$. The relationship between the variables $u_{\ell-1}^v$ and u_ℓ^v has been a subject of interest in various studies [22], [32]. Researchers have explored the dynamics and patterns exhibited by these variables over time. The utilization of the time-centric transitional-matrix facilitates the acquisition of session-specific characteristics which impacts majority of current activity log. Furthermore, in (12) is reformulated using the same approach as the HLCP model [22], yielding the following expression.

$$i_\ell^v = \mathcal{X} i_{\ell-o}^v + \sum_{j=0}^{\ell-1} \mathcal{U}_{u_\ell^v - u_{\ell-1}^v} \mathcal{S}_{w_{\ell-j}^v} \quad (12)$$

In the context of consumer sessions, the variable $i_l^v = v_0$ represents the initial state of consumers. In the context of modeling dynamic session characteristics, the DIRS model employs session-centric tranisitional-matrix using in (13). The estimation of the likelihood that a consumer $\mathcal{v}$ will engage in a specific session c towards a specific product $\mathcal{w}$ at the alternating location $\ell + 1$ is conducted in a way that is comparable to the HLCP approach [22]. This estimation is achieved by employing in (14).

$$i_\ell^v = \mathcal{X} i_{\ell-o}^v + \sum_{j=0}^{\ell-1} \mathcal{U}_{u_\ell^v - u_{\ell-1}^v} \mathcal{N}_{c_{\ell-m}^v} \mathcal{S}_{w_{\ell-j}^v} \tag{13}$$

$$z_{v,\ell+1,c,w} = (t_\ell^v)^{\mathcal{U}} \mathcal{N}_c \mathcal{S}_w = (i_\ell^v + v_v)^{\mathcal{U}} \mathcal{N}_c \mathcal{S}_w \tag{14}$$

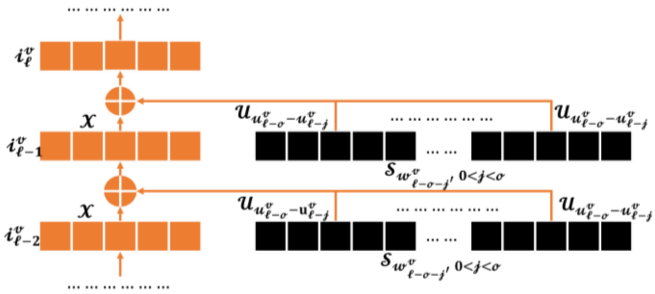

Figure 4. Architecture of DIRS model

3.5. Dynamic matrix adaptation

When faced with the task of continually acquiring distinctive possible session-variations, it becomes necessary to calculate a larger number of session-centric transitional-matrices. However, this approach may lead to issues with overfitting. In order to successfully tackle the issue at hand, it is crucial to evenly divide all potential session variation parameters into separate windows. The DIRS model focuses on the estimation of the transitional-matrices pertaining exclusively towards the upper and lower boundaries of the current session-window. The computation of transitional-matrices across every session variation in session-window is performed employing linear-interpolations (LI). The mathematical expression describing the time-centric transitional-matrices $\mathcal{U}_{ue}$, which represents the session-variation variable ue, can be formulated as (15).

$$\mathcal{U}_{ue} = \frac{\left[\mathcal{U}_{\mathcal{M}(u_e)}(\mathcal{V}(u_e) - u_e) + \mathcal{U}_{\mathcal{V}(u_e)}(u_e - \mathcal{M}(u_e))\right]}{\left[(\mathcal{V}(u_e) - u_e) + (u_e - \mathcal{M}(u_e))\right]} \tag{15}$$

The variables $\mathcal{V}(u_e)$ and $\mathcal{M}(u_e)$ represent the upper and lower boundaries of session variation u_e. Meanwhile, $\mathcal{U}_{\mathcal{V}(u_e)}$ and $\mathcal{U}_{\mathcal{M}(u_e)}$ represent the session-centric transitional-matrices for $\mathcal{V}(u_e)$ and $\mathcal{M}(u_e)$ respectively. The utilization of LI, presents a promising approach to tackle the challenge of understanding time-centric transitional-matrices in the presence of never-ending session variations. It is important to take into account that the optimization of time-centric transitional-matrices within each individual session-window follows a linear pattern. The global optimizing encompassing the total number of potential session-variations exhibits a non-linear behavior. The HLCP approach [22] has been found to be highly effective in accurately representing and analyzing the sequential behavioral traits exhibited by users. In the context of modeling session-centric information, it has been observed that the DIRS model exhibits greater efficiency compared to the HLCP model. According to our research findings, it has been observed that the DIRS model demonstrates optimal performance in scenarios where time-centric data is readily accessible. However, in cases where such data is lacking, the HLCP model [22] has been found to exhibit greater efficiency.

3.6. Bayesian personalized ranking model

The BPR model [33] is a widely employed pairwise ranking approach that is specifically designed for handling implicit feedback data. The utilization of BPR has emerged as a prominent objective parameter in the field of machine-learning (ML). In the context of BPR, it is generally assumed that customers have a preference for a selected collection of products over an unfavorable set. The objective of this study is to optimize the likelihood by utilizing in (16). The variable w' in (16) is utilized to represent negative-log probability in the context of this study. Additionally, the function $h(y)$ is a non-linear-function that has been chosen based on (17). Therefore, through the consideration of the negative-log probability, the forthcoming function objective is able to be reduced in a comparable manner using in (18).

$$\mathcal{p}(v, \ell+1, c, w \succ w') = \hbar\left(z_{v,\ell+1,c,w} - z_{v,\ell+1,c,w'}\right) \quad (16)$$

$$\hbar(y) = \frac{1}{1+e^{-y}} \quad (17)$$

$$\mathcal{K}_1 = \sum \log\left(1 + e^{-\left(z_{v,\ell+1,c,w} - z_{v,\ell+1,c,w'}\right)}\right) + \frac{\mu}{2}\|\ominus_1\|^2 \quad (18)$$

The variable μ is used to represent the regularization power control variable, which is a variable that determines the strength of regularization in a given model. On the other hand, the symbol $\ominus_1$ represents a set of parameters to be computed, namely $\mathcal{V}$, $\mathcal{S}$, $\mathcal{X}$, $\mathcal{U}$, and $\mathcal{N}$. The algorithm 1 for the BPR.

Algorithm 1. Enhanced BPR

Input	Overall Consumers $\mathcal{V}$, Products $\mathcal{U}$, Customer-Product Interaction Information $\mathcal{T}$ (session, location and time-centric transitional-matrices), Ratio-Reduction γ
Output	$\ominus$
Step 1.	For $v \leftarrow 1$ to N do
Step 2.	Generate Negative-log probability for various consumers
Step 3.	$S_v \leftarrow rand\_choose(\mathcal{V}, \mathcal{U}, \gamma)$
Step 4.	End
Step 5.	While convergence is not met do
Step 6.	Use Random-Sampling
Step 7.	$v \leftarrow$ Random selection of consumers using set $\mathcal{V}$
Step 8.	$j \leftarrow$ Random selection of purchased products using set $\mathcal{T}_v$
Step 9.	$j \leftarrow$ Random selection of negative products using set S_v
Step 10.	Compute gradient of $\ominus$ in accordance with BPR
Step 11.	Update the above variables
Step 12.	End

The purpose of this study is to assess the impact of DIRS on performance, as demonstrated in the subsequent section through empirical evidence related to HR and MRR performance metrics.

4. RESULTS AND DISCUSSION

In the results and discussion section, we initiate the analysis by elucidating the system requirements, outlining the necessary system requirements for our study. Following this, a detailed discussion on the dataset ensues, delineating its characteristics, sources, and relevance to our research context. Subsequently, we compare our obtained results with those presented in [22], [31], [32], employing the metrics of HR and MRR for a comprehensive evaluation.

4.1. System requirements

For the system configurations, the environment used for working with the Tmall dataset involves an Intel Core i7 processor, 16 GB RAM, and a Windows 11 operating system. The Python within the Annaconda framework was considered for the development of this research. In the next section, the Tmall dataset is discussed.

4.2. Dataset

The Tmall dataset, sourced from [34] is a comprehensive dataset widely utilized in the field of e-commerce for research and development purposes. This dataset contains a vast array of information related to user behaviours, item details, and transaction records within the Tmall online marketplace. The attributes in the Tmall dataset are given in Table 1.

Table 1. Attributes in the dataset

Field	Description
User_id	Unique user ID
Seller_id	Unique online seller ID
Item_id	Unique item ID
Category_id	Unique category ID
Online_Action_id	“0” denotes “click” while “1” for “buy”
Time_Stamp	Date of the format “yyyymmdd”
Merchant_id	Unique merchant ID
Location_id	Unique location ID
Budget	Budget constraints imposed on the merchant
Location_id_list	Available location list, e. g. 1:356:89
Merchant_id_list	You may recommend at most 10 merchants here, separated by “:”, e. g. 1:5:69

4.3. Hit-rate performance

The HR performance is evaluated using the following equation is evaluated using in the (19).

$$HR = \frac{TP}{TP+FN} \tag{19}$$

The Figure 5 represents the HR performance for different recommendation algorithms when considering a 10k recommendation set. The HR is a metric that measures the proportion of correct recommendations within a given set. In this context, the HR performance is as follows: HLCP (0.085), HTCP (0.457), GPAN (0.23), and DIRS (0.557). The values suggest that DIRS has the highest HR among the algorithms, indicating that nearly 53.80% of the recommendations made by this algorithm are correct within the 10k recommendation set. On the other hand, HTCP, HLCP, and GPAN have lower HR of 0.0457, 0.085, and 0.23 respectively. The Figure 6 represents the HR performance for different recommendation algorithms when considering a 20k recommendation set. The HR is a metric that measures the proportion of correct recommendations within a given set. In this context, the HR performance is as follows: HLCP (0.11), HTCP (0.534), GPAN (0.28), and DIRS (0.635). The values suggest that DIRS has the highest HR among the algorithms, indicating that nearly 51.46% of the recommendations made by this algorithm are correct within the 20k recommendation set. On the other hand, HTCP, HLCP, and GPAN have lower HR of 0.534, 0.11 and 0.28, respectively. Overall, the HR performance provides valuable insights into the accuracy of the recommendation algorithms, with DIRS exhibiting the highest performance.

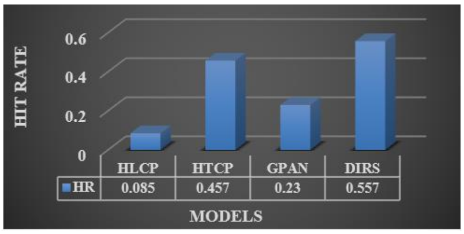

Figure 5. HR considering recommendation set of 10,000

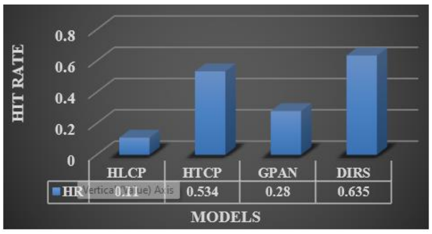

Figure 6. HR considering recommendation set of 20,000

4.4. Mean reciprocal rate

The MRR can be defined as the arithmetic mean of the reciprocal values associated with the ranking position of the initial N recommended products [31].

$$MRR = \frac{1}{|Q|}\sum_{i=1}^{|Q|}\frac{1}{rank_i} \tag{20}$$

The Figure 7 presents the MRR performance for various recommendation algorithms when considering a 10k recommendation set. MRR is a metric that evaluates the effectiveness of a recommendation system by measuring the quality of the top-ranked recommendation. In this context, the MRR values are as follows: HLCP (0.129), HTCP (0.197), GPAN (0.135), and DIRS (0.295). These values indicate the relative performance of each algorithm in terms of how well they place correct recommendations at the top of the list. DIRS stands out with the highest MRR of 47.90%, suggesting that, on average, the correct recommendations from this algorithm are placed higher in the ranking compared to HTCP (0.197), HLCP (0.129), and GPAN (0.135). The Figure 8 presents the MRR performance for various recommendation algorithms when considering a 20k recommendation set. MRR is a metric that evaluates the effectiveness of a recommendation system by measuring the quality of the top-ranked recommendation. In this context, the MRR values are as follows: HLCP (0.041), HTCP (0.298), GPAN (0.138), and DIRS (0.372). These values indicate the relative performance of each algorithm in terms of how well they place correct recommendations at the top of the list. DIRS stands out with the highest MRR of 57.25%, suggesting that, on average, the correct recommendations from this algorithm are placed higher in the ranking compared to HTCP (0.298), HLCP (0.041), and GPAN (0.138). Overall, the MRR performance indicates that DIRS is the most effective algorithm among the ones listed in terms of ranking correct recommendations higher within the 20k recommendation.

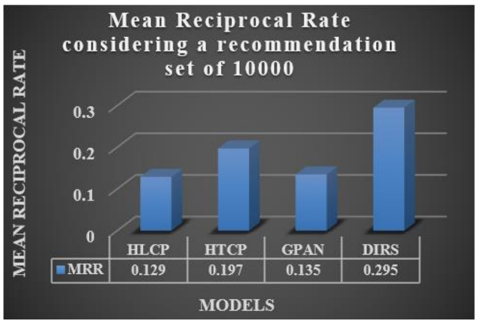

Figure 7. MRR considering recommendation of 10,000

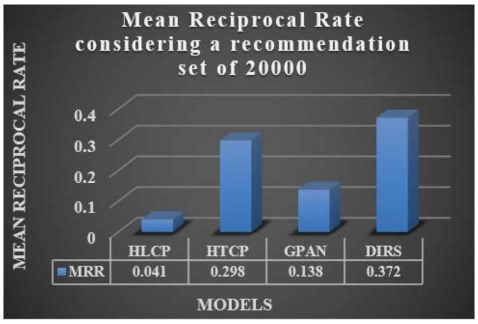

Figure 8. MRR considering recommendation of 20,000

5. CONCLUSION

In conclusion, this work presents a comprehensive and structured approach of leveraging the Tmall e-commerce dataset for personalized and optimized recommendations using the DIRS model. This work includes key steps, including dataset loading, data preprocessing, extraction of multi-behaviour characteristics, and training using both RNN and BPR models. The integration of these models enhances the system's capability to capture important consumer behaviours using location and session. The proposed DIRS model is evaluated using metrics such as HR and MRR. Analysing the HR performance across different recommendation sets reveals that DIRS consistently outperforms other algorithms, indicating a higher proportion of correct recommendations. This is evident for both 10k and 20k recommendation sets, emphasizing the effectiveness of the DIRS model. Similarly, the MRR results reinforce the superiority of DIRS, showcasing its ability to place correct recommendations higher in the ranking compared to alternative algorithms. Overall, the findings suggest that the DIRS model excels in providing accurate and high-quality recommendations within the Tmall e-commerce context. These results contribute valuable insights to the field of RSs, demonstrating the effectiveness of the proposed approach for enhancing user experience and engagement in e-commerce platforms. For future work, the proposed model can be fine tuned for gender and sentiment-based recommendations. Also, the performance can be evaluated using different datasets such as Amazon, Diginetica, and NowPlaying.

REFERENCES

[1] J. Gao, A. B. Siddik, S. K. Abbas, M. Hamayun, M. Masukujjaman, and S. S. Alam, "Impact of e-commerce and digital marketing adoption on the financial and sustainability performance of MSMEs during the COVID-19 pandemic: an empirical study," *Sustainability*, vol. 15, no. 2, p. 1594, Jan. 2023, doi: 10.3390/su15021594.

[2] M. Ciupac-Ulici, D.-G. Beju, V. P. Bresfelean, and G. Zanellato, "Which factors contribute to the global expansion of m-commerce?," *Electronics*, vol. 12, no. 1, p. 197, Jan. 2023, doi: 10.3390/electronics12010197.

[3] Y. Feng, "Enhancing e-commerce recommendation systems through approach of buyer's self-construal: necessity, theoretical ground, synthesis of a six-step model, and research agenda," *Frontiers Artificial Intelligence*, vol. 6, May 2023, doi: 10.3389/frai.2023.1167735.

[4] L. Liu, "E-commerce personalized recommendation based on machine learning technology," *Mobile Information Systems*, vol. 2022, pp. 1–11, Apr. 2022, doi: 10.1155/2022/1761579.

[5] F. Liu, "Design of personalized recommendation algorithms based on big data framework," *Journal of Physics: Conference Series*, vol. 2138, no. 1, p. 012025, Dec. 2021, doi: 10.1088/1742-6596/2138/1/012025.

[6] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: algorithms, challenges, metrics, and business opportunities," *Applied Sciences*, vol. 10, no. 21, p. 7748, Nov. 2020, doi: 10.3390/app10217748.

[7] M. B. Gulfraz, M. Sufyan, M. Mustak, J. Salminen, and D. K. Srivastava, "Understanding the impact of online customers' shopping experience on online impulsive buying: a study on two leading e-commerce platforms," *Journal of Retailing and Consumer Services*, vol. 68, no. 103000, p. 103000, Sep. 2022, doi: 10.1016/j.jretconser.2022.103000.

[8] A. Haleem, M. Javaid, M. A. Qadri, R. P. Singh, and R. Suman, "Artificial intelligence (AI) applications for marketing: a literature-based study," *International Journal of Intelligent Networks*, vol. 3, no. 3, pp. 119–132, 2022, doi: 10.1016/j.ijin.2022.08.005.

[9] A. Panteli and B. Boutsinas, "Addressing the cold-start problem in recommender systems based on frequent patterns," *Algorithms*, vol. 16, no. 4, p. 182, Mar. 2023, doi: 10.3390/a16040182.

[10] S. Milano, M. Taddeo, and L. Floridi, "Recommender systems and their ethical challenges," *AI & SOCIETY*, vol. 35, Feb. 2020, doi: 10.1007/s00146-020-00950-y.

[11] M. Maher *et al.*, "Comprehensive empirical evaluation of deep learning approaches for session-based recommendation in e-commerce," *Entropy*, vol. 24, no. 11, p. 1575, Oct. 2022, doi: 10.3390/e24111575.

[12] D. Liu, "Intelligent recommendation system based on the infusion algorithms with deep learning, attention network and clustering," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, May 2023, doi: 10.1007/s44196-023-00264-z.

[13] S. Maneeroj and N. Sritrakool, "An end-to-end personalized preference drift aware sequential recommender system with optimal item utilization," *IEEE Access*, vol. 10, pp. 62932-62952, 2022, doi: 10.1109/ACCESS.2022.3182390.
[14] N. Sritrakool and S. Maneeroj, "Personalized preference drift aware sequential recommender system," *IEEE Access*, vol. 9, pp. 155491-155506, 2021, doi: 10.1109/ACCESS.2021.3128769.
[15] M. Salampasis *et al.*, "A flexible session-based recommender system for e-commerce," *Applied Sciences*, vol. 13, no. 5, p. 3347, Mar. 2023, doi: 10.3390/app13053347.
[16] R. Esmeli, M. Bader-El-Den, H. Abdullahi, and D. Henderson, "Implicit feedback awareness for session-based recommendation in e-commerce," SN Computer Science, vol. 4, no. 3, Apr. 2023, doi: 10.1007/s42979-023-01752-x.
[17] L. Gao and J. Li, "E-commerce personalized recommendation model based on semantic sentiment," *Mobile Information Systems*, vol. 2022, pp. 1–10, Aug. 2022, doi: 10.1155/2022/7246802.
[18] C. N. Dang, M. N. Moreno-García, and F. D. la Prieta, "An approach to integrating sentiment analysis into recommender systems," Sensors, vol. 21, no. 16, p. 5666, Aug. 2021, doi: 10.3390/s21165666.
[19] M. Elahi, D. Khosh Kholgh, M. S. Kiarostami, M. Oussalah, and S. Saghari, "Hybrid recommendation by incorporating the sentiment of product reviews," *Information Sciences*, vol. 625, pp. 738–756, May 2023, doi: 10.1016/j.ins.2023.01.051.
[20] R. V. Karthik and S. Ganapathy, "A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce," *Applied Soft Computing*, vol. 108, p. 107396, Sep. 2021, doi: 10.1016/j.asoc.2021.107396.
[21] Y. Liu and Z. Ding, "Personalized recommendation model of electronic commerce in new media era based on semantic emotion analysis," *Frontiers in Psychology*, vol. 13, Jul. 2022, doi: 10.3389/fpsyg.2022.952622.
[22] B. R. Sreenivasa and C. R. Nirmala, "Hybrid location-centric e-commerce recommendation model using dynamic behavioural traits of customer," *Iran Journal of Computer Science*, vol. 2, no. 3, pp. 179–188, Jun. 2019, doi: 10.1007/s42044-019-00040-3.
[23] L. Xu and X. Sang, "E-commerce online shopping platform recommendation model based on integrated personalized recommendation," *Scientific Programming*, vol. 2022, pp. 1–9, Apr. 2022, doi: 10.1155/2022/4823828.
[24] C. Yin, S. Ding, and J. Wang, "Mobile marketing recommendation method based on user location feedback," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, May 2019, doi: 10.1186/s13673-019-0177-6.
[25] Q. Xu and J. Wang, "A social-aware and mobile computing-based e-commerce product recommendation system," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–8, Mar. 2022, doi: 10.1155/2022/9501246.
[26] B. Lu and S. Zhang, "A conjoint approach to understanding online buyers' decisions towards online marketplaces," *Journal of theoretical and applied electronic commerce research*, vol. 15, no. 3, pp. 69–83, 2020, doi: 10.4067/s0718-18762020000300106.
[27] A. A. Ahouf, and K. Lu, "Establishing trust in e-commerce through website design elements," *International Journal of Technology and Human Interaction*, vol. 18, no. 1, Jan. 2022, doi: 10.4018/ijthi.297615.
[28] J. Sidlauskiene, Y. Joye, and V. Auruskeviciene, "AI-based chatbots in conversational commerce and their effects on product and price perceptions," *Electronic Markets*, vol. 33, no. 1, May 2023, doi: 10.1007/s12525-023-00633-8.
[29] R. Hirt, N. Kühl, and G. Satzger, "Cognitive computing for customer profiling: meta classification for gender prediction," *Electronic Markets*, vol. 29, no. 1, pp. 93–106, Feb. 2019, doi: 10.1007/s12525-019-00336-z.
[30] A. B. Melchiorre, N. Rekabsaz, E. Parada-Cabaleiro, S. Brandl, O. Lesota, and M. Schedl, "Investigating gender fairness of recommendation algorithms in the music domain," *Information Processing & Management*, vol. 58, no. 5, p. 102666, Sep. 2021, doi: 10.1016/j.ipm.2021.102666.
[31] L. Dong, G. Zhu, Y. Wang, Y. Li, J. Duan and M. Sun, "A graph positional attention network for session-based recommendation," *IEEE Access*, vol. 11, pp. 7564-7573, 2023, doi: 10.1109/ACCESS.2023.3235353.
[32] B. R. Sreenivasa and C. R. Nirmala, "Hybrid time centric recommendation model for e-commerce applications using behavioral traits of user," *Information Technology and Management*, Mar. 2022, doi: 10.1007/s10799-022-00358-8.
[33] D. Xiang and Z. Zhang, "Cross-border e-commerce personalized recommendation based on fuzzy association specifications combined with complex preference model," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–9, Oct. 2020, doi: 10.1155/2020/8871126.
[34] "IJCAI-16 brick-and-mortar store recommendation Dataset_Tianchi datasets," tianchi.aliyun.com. https://tianchi.aliyun.com/dataset/53 (accessed Nov. 16, 2023).

BIOGRAPHIES OF AUTHORS

Sinzy Silvester received Master in Computer Applications (MCA) degree from Kerala University, India, in 2006 and the M.E. (CSE) from Anna University Coimbatore Tamil Nadu, in 2010 and pursuing Ph.D. from Visvesvaraya Technological University (VTU) Karnataka. Currently, she is an Assistant Professor at the Department of Computer Applications at IFIM College Bangalore. Her research interests include machine learning, deep learning, artificial intelligence, big data analytics, design, and analysis of algorithms. She can be contacted at email: sinzy1983@gmail.com.

Shaji Kurain holds a Ph.D. in Management Studies from Visvesvaraya Technological University (VTU), Karnataka. He is currently Head of HR Department at JACSOM. His thesis is on 'strategic HRM practices in small and medium enterprises in light of global competition'. His current research interests include future of work, new-gen workforce. His areas of interest are behavioral science, managerial psychology, leadership, human capital management. In addition, he is an empaneled expert in Times of India Ascent Series and an invited speaker at various national and international forums. He is a common face in NHRDN panel discussions. He is also an invited Guest Faculty at Uniglobe School of Business, Kathmandu. He can be contacted at email: shaji.kurian@jagsom.edu.in.