IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 1, March 2024, pp. 786~797
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp786-797

A recommender system-using novel deep network collaborative filtering

Shruthi Nagaraj, Blessed Prince Palayyan
Department of Computer Science and Engineering, Presidency University, Bangalore, India

Article Info

Article history:

Received May 4, 2023
Revised Oct 25, 2023
Accepted Nov 7, 2023

Keywords:

Deep network collaborative filtering
Dual network
Graph neural network
Network collaborative filtering
Recommender system

ABSTRACT

The recommendation model aims to predict the user's preferred items among million through analyzing the user-item relations; furthermore, Collaborative Filtering has been utilized as one of the successful recommendation approaches in last few years; however, it has the issue of sparsity. This research work develops a deep network collaborative filtering (DeepNCF), which incorporates graph neural network (GNN), and novel network collaborative filtering (NCF) for performance enhancement. At first user-item dual network is constructed, thereafter-custom weighted dual mode modularity is developed for edge clustering. Furthermore, GNN is utilized for capturing the complex relation between user and item. DeepNCF is evaluated considering the two distinctive. The experimental analysis is carried out on two datasets for Amazon and movielens dataset for recall@20 and recall@50 and the normalized discounted cumulative gain (NDCG) metric is evaluated for Amazon Dataset for NDCG@20 and NDCG@50. The proposed method outperforms the most relevant research and is accurate enough to give personalized recommendations and diversity.

Corresponding Author:

Shruthi Nagaraj
Department of Computer Science and Engineering, Presidency University
Bangalore, India
Email: shruthin_12@rediffmail.com

1. INTRODUCTION

Nowadays, it is customary to make purchases online. Consumers sometimes find it difficult to select the most interesting product from among the various ones offered by E-platforms. A substantial amount of internet information may now be available rapidly with advances in recent technological developments and the popularity of online services. For a diversity of online services and goods, users can leave reviews, comments, and ratings. Because of recent advancements in global computing, the problem of online data overload has arisen. Finding meaningful and helpful information online is becoming more and more challenging because of the data flood. Yet, several contemporary approaches with reduced computing demands may now more effectively and efficiently direct users to the relevant material. As a result, generating recommender systems has gained a lot of traction lately. Generally considering, recommender systems serve as information filtration tools, providing customers with pertinent and practical information [1]–[3]. For each page of the website, we are on, a recommendation algorithm is used. For instance, e-commerce companies typically provide "guess your favorite" on the front page. There are many more websites, including those for movies, videos, news, books, restaurants, maps, and other content, that have built-in recommendation engines, amongst which a few of the ones are already in use.

Collaborative, content-based, and hybrid systems are the three primary categories of recommender systems (RSs). Without taking into account any information about other users, content-based filtering (CBF) predicts a user's preferences based on his or her information (gender, age and activities on social media). Given that it employs several strategies to provide the user with the information they need, CBF might be seen as an information-filtering task. Instead of looking for particular data within an incoming stream, filtering is frequently considered of as the removal of undesirable data (viewed as noise) from that stream. The most often used strategy is based on the semantic content of an item. In this scenario, a number of the guiding principles from the information retrieval discipline, upon which it is built, are applied: products are suggested based on comparisons between their content and the user profile. This profile is shown as a table with the user-specified key elements and weights. This method of information retrieval is straightforward, quick, and effective [4]–[7]. The majority of current research has been on recommendation systems based on deep learning. Deep learning model topologies may be easily changed to accommodate different recommendation situations and specific application conditions. By mining feature combinations and fitting data patterns, deep learning recommendation models outperform traditional recommendation algorithms. Yet, the majority of these current recommendation algorithms ignore the potential worth of supplementary data, user ratings, and user reviews. Auxiliary information, a type of implicit feedback, contains user characteristics and item characteristics. Hence, it is possible to mine auxiliary data for implicit attributes of objects. Yet, it is important to remember that different users may have different rating habits. Ratings from users may seem to indicate a user's preference for a certain item. As an illustration, some people constantly offer great ratings, whilst others may have high expectations and frequently offer negative ones. In this situation, it is challenging to evaluate an item's quality just based on customer reviews. In addition, just because two people give a product a high rating does not always mean that they have the same opinion on it. Because of his rating habits, one individual could just give everything that is not invasive, a good review while the other might truly appreciate the product. We will incorporate user feedback into the recommendation system to address these issues efficiently. Both user ratings and reviews are explicit observations, but since reviews are more customized than ratings, they are more likely to represent people's emotional tendencies [8]–[11].

The internet, books, e-learning, travel, movies, music, e-commerce, news, specialized research resources and television shows. are just a few of the applications in a variety of recommender systems. To offer its customers personalized recommendations across a range of applications, it is critical to create better and more efficient recommender systems. For a wider range of applications, the current generation of recommender systems must be improved to offer better useful and suitable recommendations. It is necessary to conduct further research on recent studies on recommender systems that focus on different applications. Further, research contribution is given as follows,

- Deep network collaborative filtering (DeepNCF) comprises a novel network collaborative filtering and graph neural network (GNN) for performance enhancement; user-item based dual mode network is constructed through rating matrix; later custom weighted dual mode modularity is developed for edge clustering.
- GNN is adopted for optimization of edge clustering and understand the more complex relationship between user as well as item. Integrated approach of network collaborative filtering (NCF) and GNN. Furthermore, Edge reduction approach is used for selection of non-redundant edges.
- DeepNCF is evaluated considering the two distinctive dataset of Amazon Books and MovieLens considering the metrics normalized discounted cumulative gain (NDCG) and recall with different variants.

This Particular research is organized as follows: First section of the research work starts with the background of recommendation system, importance of collaborative filtering and integration of deep learning towards recommendation system. Second section discusses the existing work for recommendation system. DeepNCF is designed and develop in third section along with mathematical modelling and architecture; DeepNCF is evaluated considering the different dataset along with comparative analysis.

2. RELATED WORK

Graph neural network (GNN) has been one of the key research areas in recommendation system since development of earlier one; this section discuss little recent existing mechanism that adopted GNN for exploiting the optimal relationship among user and items. To develop precise session embedding and train item representations from session graphs, gated graph neural networks were employed [12]. Identified the local and long-range interdependence of session-based recommendations using GNN and self-attention approaches [13]. To locate and create item and user choice embedding, several weighted graph attention networks were used [14]. Removed nonlinear activations and feature transformations from the network architecture to address graph convolutional networks [15]. They uses GNN to handle a variety of characteristics and created a component to look at the relationship between potential neighbor nodes [16].

Object interactions were captured in using a self-attentive graph neural network and a soft attention method [17]. Yet, statistics on user-item interaction are frequently scarce. By including item attribute data in interaction characteristics, some researchers have started to combine multimodal data to alleviate the data shortage. To find user and object embedding for the goal of making recommendations, consider persons and objects as knowledge networks, including their traits and interactions; Proposes linking user-item interaction graphs with user-user social networks to improve social recommendation [18], [19]. A technique for capturing the intricate relationship between explicit user decisions and edge data was developed by Inx [20].

To get over these limitations and provide more precise item recommendations, the review text and node attributes for graph neural network recommendation (RTN-GNNR) model combines the review text and node attributes for graph neural network recommendation. There are four parts to the RTN-GNNR. A bidirectional, gated network is suggested by the module for obtaining review text attributes. Bidirectional encoder representation from transformers (BERT) and an attention mechanism are used in the recurrent unit (Bi-GRU) text analysis method to assist the model choose the most important reviews [21], [22]. Construct a context-aware recency-based attention network (CARAN) that uses the attention mechanism to give recent visitation spots the highest priority based on the temporal and spatial context and the weather. Using linear interpolation and spatiotemporal matrices to characterize spatial distance smoothly, enables interaction between non-adjacent check-ins.

3. PROPOSED METHODOLOGY

This research develops DeepNCF, which integrates the graph neural network (GNN) and network collaborative filtering for performance enhancement in terms of diversity and recommendation accuracy. Figure 1 shows the DeepNCF workflow; at first novel dual mode graph is designed, later novel NCF with edge clustering and edge reduction is carried out; further we adopt GNN architecture and design Deep NCF train DeepNCF to make the recommendation. Figure 2 shows the proposed workflow. The first step here is to develop a bipartite graph, which performs link segmentation that generates a dual mode graph community. On this basis, a link segmentation algorithm divides the so that it has a robust structure. The second step is the execution of a deep reinforcement learning (DRL) into a specific single community such as the neighborhood chosen among the group that generates a varied type of recommendations. The in-depth discussion of the proposed model is discussed here.

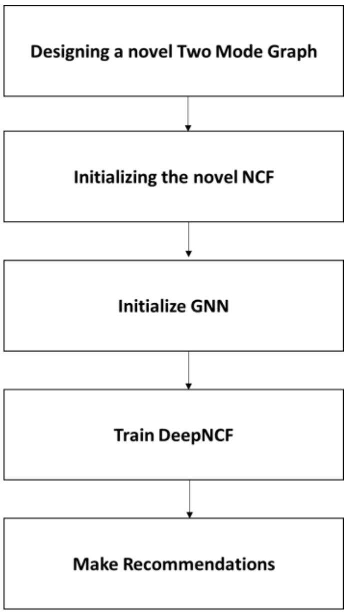

Figure 1. DeepNCF workflow

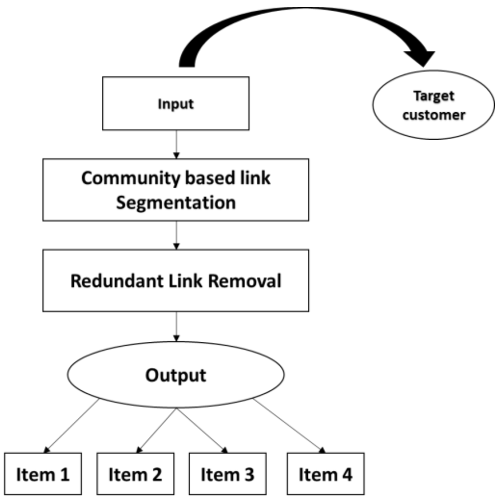

Figure 2. Proposed workflow

3.1. Problem statement

The main goal of this work is to draw attention to the problem of recommendation overfitting while ensuring recommendation accuracy and maintaining variation without the need for additional data. The traditional user-based collaborative filtering method wherein a neighborhood consists of the user's relevant ratings that focus on specific items. This will allow the users that are present in the neighborhood to retain the similarity with the target user, which consists of various choices. Because the links here are connected to various items that do not efficiently reflect a user's choice. Upon considering an example of users rating to meet the personalized requirements, henceforth ensuring that the candidate neighbor ensures the same rating features irrespective of the correlated dense items with the target user. To enhance the recommendation system, the users recommend a specific category of items. The existing techniques are based on the mechanism of user-based collaborative filtering which enhances the performance of the recommendation system recently. Solving the overfitting in user-based collaborative filtering is further applicable to various techniques used in recent Recommendation systems. The proposed system (PS) approach is shown below.

3.2. Edge clustering

Simultaneously, the dual mode network has gained widespread use for various events such as customer-item relationships. Here there exist two categories of the non-overlapping category of nodes for uppermost and bottommost nodes. A specific single link here collaborates with a set of nodes that belongs to various categories. The two-mode graph is an essential measure to enhance the community structure of the two-mode graph. To attain an effective community identification model that proposes a dual mode graph which states that there exists top k top nodes and e below nodes in the dual mode graph. Henceforth the two-mode graph here is considered as $k * e$ incidence matrix B that focuses on the communication among the top and bottom nodes whereas $B_{m,n} = 1$ if there exists a link between uppermost node m and bottommost node n and $B_{m,n} = 0$. Hence $C_{m,n}$ shows the probabilities associated with any random model through a link that is present between the uppermost node m and bottommost node n. Accordingly, to the connection matrix B and C, for the two-mode graph, D is shown as,

$$D = \sum_{m=1}^{k} \sum_{n=1}^{e} \left(\frac{B_{m,n}}{K} - C_{m,n}\right)\varphi(H(m), H(n)) \tag{1}$$

$$= \frac{1}{K}\sum_{m=1}^{k} \sum_{n=1}^{e} \left(\frac{B_{m,n}}{} - \frac{P(m)P(n)}{K}\right) \varphi(H(m), H(n)) \tag{2}$$

Here K denotes the number of links in B. $H(m)$ In addition, $H(n)$ denotes the degree of nodes m and n irrespectively. The φ function here shows the $\varphi(H(m), H(n))$ to equalize the nodes m and n are segmented into a similar community else 0. The RS consists of customer-item relation; it shows the dual mode graph. An RS is transformed into $IBG = H(T, J, S)$, which shows the $|T|$ denotes the user nodes, $|J|$ consists of item nodes and $|S|$ denotes the number of links. $S_{t,j}$ shows that a customer t has rated an item$j$. Henceforth IBG is denoted as $|T|$*$|J|$ adjacent matrix B. The adjacency matrix B in (3) shows the correlation degree between the node t and item$j$, which consists of two links that accommodate the transformation. The links consisting of degree $\boldsymbol{\alpha} \le l_{t,j} \le 1$ are denoted as strong links and $0 \le g_{t,j} \le \boldsymbol{\alpha}$ are denoted as weak links.

$$G_{t,j}^{v} = s(t,j) = \ 0 \ \ else \frac{c_{t,j}}{Maxj\in J\{} \ 0 < c_{t,j} \ < \mu \ 1.0 \ \ \mu \le c_{t,j} \ \le 1 \tag{3}$$

Here each link denotes the parameter for each community, to analyze a co-efficient $Prob(l_f | S_{t,j})$ that denotes the link $S_{t,j}$ that denotes the community l_f. In this step, each link is associated with many communities at a specific time and the main advantage levied on the dual mode graph is weighted by the co-efficient. The dual mode graph is reframed upon substitution by $\varphi(S(t), S(j))$ with different parameters as $\psi_{t,j,f}$ and $\omega_{t,j,f}$ irrespectively.

$$X^{v} = \frac{1}{|S|}\sum_{f=1}^{F} \sum_{t=1}^{|T|} \sum_{j=1}^{|J|} \left(\psi_{t,j,f}, G_{t,j}^{v} - \omega_{t,j,f} \frac{P(t)P(j)}{|S|}\right) \tag{4}$$

$$\psi_{t,j,f} = pred(S_f | S_{t,j}) \tag{5}$$

The two co-efficient $\psi_{t,j,f}$ and $\omega_{t,j,f}$ are constructed. For $\psi_{t,j,f}$ the probability that is associated with $S_{t,j}$ to the model belonging to a communityS_f. $\psi_{t,j,f}$ Is equivalent to 1 if the customer node t and item node j are segmented into similar community S_f if $S(T) = S(J) = S_f$ this is equivalent to 0. Henceforth the probability of two nodes associated with a similar community is high in comparison if the nodes considered are neighbours. We can state that,

$$\omega_{t,j,f} = pred^{J}(S_f | S_{t,j}) pred^{T}(S_f | S_{t,j}) \tag{6}$$

$$pred^J\left(S_{t,j}\right) = \frac{\sum_{j\in J} pred(S_f|S_{t,j})}{|J|} \tag{7}$$

$$pred^T\left(S_{t,j}\right) = \frac{\sum_{t\in T} pred(S_f|S_{t,j})}{|T|} \tag{8}$$

Accordingly, to (4) -(8) a dual mode weighted graph is,

$$X^v = \frac{1}{|S|}\sum_{f=1}^{F} \sum_{t=1}^{|T|} \sum_{j=1}^{|J|} \left(pred(S_f|S_{t,j})G_{t,j}^v - pred^J\left(S_{t,j}\right)pred^T\left(S_{t,j}\right)\frac{P(t)P(j)}{|S|}\right) \tag{9}$$

A high-end value X^v denotes a robust community structure for , the X^v shows the difference amidst a dual mode graph after the community is identified and an unstructured network doesn't contain a community structure. The wider the change the clearer the community structure is Assumption $X^v = 0$ then all the links belong to a similar community, if so then $F = 1$ we get, $\forall j \in J, t \in T \;\; pred\left(S_{t,j}\right) = 1$. Simultaneous to (7) and (8) we can find the equation,

$$pred^J\left(S_{t,j}\right)=\frac{\sum_{j\in J} pred(S_1|S_{t,j})}{|J|} = \frac{|J|}{|J|} = 1 \tag{10}$$

$$pred^T\left(S_{t,j}\right) = \frac{\sum_{t\in T} pred\left(S_{t,j}\right)}{|T|} = \frac{|T|}{|T|} = 1 \tag{11}$$

The $9 - th$ equation is simply reduced to,

$$X^v = \frac{1}{|S|}\sum_{f=1}^{F} \sum_{t=1}^{|T|} \sum_{j=1}^{|J|} \left(pred\left(S_{t,j}\right)G_{t,j}^v - pred^J\left(S_{t,j}\right)pred^T\left(S_{t,j}\right)\frac{P(t)P(j)}{|S|}\right) \tag{12}$$

$$= \frac{1}{|S|}\sum_{t=1}^{|T|} \sum_{j=1}^{|J|} \left(G_{t,j}^v - \frac{P(t)P(j)}{|S|}\right)$$

Henceforth if all inks are focused on the same community, then we have$X^v = 0$. The X^v uses local data the details of the link associated with each community. Community identification comes under the category of minima or maxima. A global value is introduced here as X^v, the Max_{deg} is related to global data, and integrate Max_{deg} into X^v. The Max_{deg} is denoted as,

$$Max_{deg} = \frac{1}{|U|}\frac{1}{|J|}\sum_{t=1}^{|T|} \sum_{j=1}^{|J|} \; 1 \le f \le Fmaxpred(S_f|S_{t,j}) \tag{13}$$

Here the dual mode graph is denoted by,

$$X_{Globloc}^v = \frac{2Max_{deg}+X^v}{3} \tag{14}$$

Consequently, by (14) we can show that $X_{Globloc}^v$ that focuses on local as well as global information whereas the variation exists in the [0, 1] range. Increasing the value $X_{Globloc}^v$, the effective result is generated in terms of community detection. Parallely, a novel algorithm is developed for a dual mode weighted graph for each link here $S_{t,j}$ the initialization is done as $pred(S_f|S_{t,j})$ for $\sum_{f=1}^{F} \quad pred(S_f|S_{t,j})$=1.

$$pred\left(S_{t,j}\right) = \frac{[pred(t)+\tau]*[pred(t)+\vartheta))}{\sum_{f'=1}^{F} \; [pred\left(t)+\tau]*\sum_{f'=1}^{F} \; [pred(t)+\vartheta)\right)} \tag{15}$$

Here τ and ϑ are considered as the hyper-parameter so that the denominator is not 0. Based on the values of$pred(S_f|S_{t,j})$ shown by (16) we try to compute the equation $pred(S_f|t\,)$ and $pred(S_f|j\,)$.

$$pred(t) = \frac{\sum_{j\in J(t)} pred(S_f|S_{t,j})}{\sum_{f'=1}^{F} \sum_{j\in J(t)} pred(S_{f'}|S_{t,j})} \tag{16}$$

$$pred(t) = \frac{\sum_{t\in T(j)} pred(S_f|S_{t,j})}{\sum_{f'=1}^{F} \sum_{t\in T(j)} pred(S_{f'}|S_{t,j})} \tag{17}$$

Based on the results considered $pred(t)$ and $pred(t)$ the value of $pred(S_f|S_{t,j})$ is re-evaluated by (15). The evaluation is carried out above until the $pred(S_f|S_{t,j})$ that reaches the community to achieve the large value for $pred(S_f|S_{t,j})$ as the final community that links the $S_{t,j}$. The custom dual mode weighted graph $X^v_{Globloc}$ is evaluated at each level from F_{min} to F_{max}. The higher the value for $X^v_{Globloc}$, the robust the community is, the maximum value obtained by $X^v_{Globloc}$ by considering an optimal number of F communities.the value F_{min} is estimated to be 2, whereas the F_{max} value is given by $F_{max} = (|T| * |J|)^{0.5}$. Edge clustering algorithm as shown in Algorithm 1.

Algorithm 1. Edge clustering algorithm

Input	Dual Mode Network $IBG = H(T, J, S)$, the number of communities F.
Step 1	Partition the $H(T, J, L)$ into F-number of communities $\{l_1, l_2, \dots \dots, l_f\}$.
Step 2	for each link $S_{t,j}$ do
Step 3	Initialize probability $Prob(l_f|S_{t,j})$, for $\sum_{f=1}^{F} \quad Prob(S_{t,j}) = 1$
Step 4	end for
Step 5	for each link $S_{t,j}$ do
Step 6	repeat
Step 7	Evaluate $Prob(l_f|t)$ and $Prob(l_f|j)$
Step 8	Use the values of $Prob(l_f|t)$ and $Prob(l_f|j)$ to evaluate $Prob(S_{t,j})$ accordingly.
Step 9	Till it converges $Prob(S_{t,j})$
Step 10	Choose a community that has the largest $Prob(S_{t,j})$ that is the final community that connects $S_{t,j}$
Step 11	Update the value in $\{l_1, l_2, \dots \dots, l_f\}$
Step 12	end for
Step 13	Return k number of communities $\{l_1, l_2, \dots \dots, l_f\}$
output	F number of communities $\{l_1, l_2, \dots \dots, l_f\}$

3.3. Edge reduction

Once the edge clustering is done there exist many redundant links that may reduce the performance of the system with the IBG denotes as redundant links. Consider customers c_1, c_2 *and* c_3 by considering the items I_1, I_2, I_3, I_4, I_5 and I_6 is reduced into a similar community. Additionally, c_1, c_2 *and* c_3 generate comparisons with a target customer t_c and the comparisons denoted as ranks are depicted in the table. It is essential to ensure that the community includes this customer, hence c_1 *and* c_2 is selected. In the next phase, the information of customersc_1, c_2 *and* c_3. Via Figure 3 to evaluate a $S(c_1)$= $\{s_{1,1}, s_{1,2}, s_{1,3}\}$, $S(c_2)$= $\{s_{2,1}, s_{2,2}, s_{2,3}, s_{2,4}, s_{2,5}\}$ and $S(c_3)$= $\{s_{3,1}, s_{3,2}, s_{3,3}, s_{3,6}\}$. the links are connected to a similar item that is equal. $s_{1,2}$ is equivalent to $s_{2,2}$ upon computation with c_2 that has additional links $\{s_{2,4}, s_{2,5}\}$ in comparison with c_1, indicates c_2 that recommends items IM_4 and IM_5 additionally along with c_1. Edge reduction as shown in Algorithm 2.

The suggestions from the united c_1 and c_2 communities will be identical to those from c_2 alone. Thus, establishing connections with c_1 is irrational. Used to foster community The connections to IM_1, IM_2, and IM_3 are the same for c_2 and c_3 but c_3 also has a connection to IM_6. This suggests that while c_2 cannot approve IM_6, c_3 can. Therefore, to assess similarity links leading to c_3 must be preserved. So, c_2, c_3, rather than c_3, c_2, will be a superior decision if we chose the top two users as neighbours. As c_3's link set includes c_1, it is a stronger contender for acceptance into the community than c_1. Duplicate IBG connections are covered in this section. Letc_a stand in for a single user node, and let C represent the group of user nodes in the IBG. Connections in S(c_a)are judged redundant if another user node already exists such that S(c_a) S(c_b). The remaining linkages in a IBG are known as nonredundant links once all redundant connections have been eliminated, at which point the IBG ceases to be reducible. Duplicate links in the user's link list will be removed. Duplicate links prevent the algorithm from determining how similar a user is to a target user. As a result, this user is removed from the target user's list of prospective neighbors. In this study, we postulate that removing redundant connections produced by user nodes like $c1$ may broaden the variety of ideas. Comprehensive edge reduction, data are provided by the second technique. It is essential to stress that we conduct edge reduction is in each local community rather than the entire IBG for two reasons.

a) It is possible to lessen the computational difficulty of edge reduction.

b) Any number of unnecessary links may be deleted.

First, following edge clustering, each community's size is significantly reduced to the size of the IBG as a whole, representing the edge reduction that is performed in a lesser region. In addition, edge reduction is performed concurrently in many communities, hence decreasing computing time by a factor of T compared to processing the whole IBG. The second argument is because it is simple to derive that redundant links in the IBG are also considered redundant links in a single community; yet, a redundant link in a single community, redundant connections may be found throughoutIBG. The whole IBG features an enormous amount of user nodes and item nodes. Despite the number of connections that exist between each user node and the link set, due to its small size, this user node, which only appears within other user nodes, is challenging to locate. Finding out if a user node's link set is present in the link sets of other user nodes will be easier after the domain has been limited from the whole IBG to each distinct community. Novel NCF algorithm as shown in Algorithm 3. Moreover, these three-algorithm lead development of novel NCF; further, it is integrated with GNN for enhancement of performance which is carried out in the next section of the research.

Algorithm 2. Edge reduction

Input	G community l_f
Step 1	Fix the value $l_f^c = l_f$
Step 2	for $a = 1\ to\ \lvert T(l_f) \rvert do$
Step 3	for b =1 to $\lvert T(l_f) \rvert do$
Step 4	if S(l_a) ⊆ S(l_b) then
Step 5	if S(l_a) ⊆l_f^c then
Step 6	$l_f^c = l_f^c -$ S(l_a)
Step 7	end if
Step 8	end if
Step 9	end for
Step 10	end for
Step 11	return G community to l_f^c
output	G community to l_f^c

Algorithm 3. Novel NCF algorithm

Input	Item matrix IM and a target customer tc, K recommendations for the target customer tc K: the total items recommended to the target customer tc $X_{Globloc}^v$:the custom dual mode model J_{tc}: The item set which the target customer has not rated yet. $pred_{tc,j}$: predicted target score for the target customer tc for item j
Step 1	Map a IM to an item Dual Mode graph $IBG = H(T, J, S)$
Step 2	$X_{Globloc}^v = 0; F = 0$
Step 3	for $F' = f_{min}$ to f_{max} do
Step 4	Use edge clustering algorithm (alg 1) to get the communities $\{l_1, l_2, \dots \dots, l_f\} = ps(H(T, J, S), F')$.
Step 5	Evaluate the $X_{Globloc}^v$ value for $\{l_1, l_2, \dots \dots, l_f\}$
Step 6	If $X_{Globloc}^v \leq X_{Globloc}^v$, then
Step 7	$X_{Globloc}^v \leq X_{Globloc}^v$, $F = F'$,
Step 8	$H(T, J, S)$ is clustering into $\{l_1, l_2, \dots \dots, l_f\}$
Step 9	end if
Step 10	end for
Step 11	Compute $H(T, J, S)$ with community F and peak $X_{Globloc}^v$
Step 12	for $f = 1$ to F do
Step 13	Utilize (alg 2) to remove redundant links in l_f , such that $l_f^c = ps(l_f)$
Step 14	end for
Step 15	Fetch a set of communities $\{l_1^c, l_2^c, \dots., l_f^c\}$
Step 16	for each item $j \in J_{tc}$ do
Step 17	Extract the community l_f^c that consists of the target customer tc and item j
Step 18	Evaluate the similarity between the target customer tc for each customer $c \in l_f^c$
Step 19	Choose customers in l_f^c having maximum similarity for the neighbourhood Ni_{tc} of the target customer tc
Step 20	Predicting the score $pred_{tc,j}$ for item j for Ni_{tc}
Step 21	end for
Step 22	The K items with $pred_{tc,j}$ recommended to the target customer tc
output	K recommendations for the target customer tc, $X_{Globloc}^v$, J_{tc}, $pred_{tc,j}$

4. RESULT SECTION

In this section, we compare the effectiveness and robustness of the proposed PS approach comparison with the existing state-of-art methods on two datasets i.e. Amazon Dataset and MovieLens Dataset, which are the benchmark datasets. We evaluate recall for Amazon dataset and MovieLens dataset for recall @20 and recall@ 50, whereas the NDCG metric is evaluated for the amazon database for metrics NDCG@20 and NDCG@50. The various state-of-art techniques used for comparison are MostPopular, YouTube DNN, GRU4Rec, MIND, ComiRec-DR, ComiRec-SA and Fat technique.

4.1. Methods considered for evaluation

The methods considered for evaluations are: i) MostPopular is a traditional recommendation method, which will recommend the most popular items to users, ii) YouTube DNN [23] is one of the most successful deep learning models of industry recommender system, iii) GRU4Rec [24] is the first time to use a gated recurrent neural network to model the sequence relationship between items in the sequence, iv) MIND [25] is a recommendation model of multi-interest extraction using a capsule network, v) ComiRec-DR [26] is the capsule network model based on the original dynamic routing method that has better performance than MIND, vi) ComiRec-SA [26] is a multi-interest extraction model based on self-attention, vii) Future aware diverse trends (FAT) [27] is a recent state-of-the-art model, which takes into consideration of users' feature sequence. FAT uses ComiRec-DR methods to get the future sequence multi-interest, and it computes correlations among different users by using a collaborative filtering method.

Moreover, the above all method has been compared in the existing model [28], thus proposed DeepNCF is compared with existing model.

4.1.1. Dataset details

The dataset details are: i) Amazon Dataset: This consists of product reviews and metadata from Amazon. In our experimental analysis, we use the book category of the Amazon dataset. The length of each training sample is truncated to 20, and ii) MovieLens Dataset: The benchmark data set in the field of recommendations is the MovieLens dataset. These facts consist of user demographics (age, gender), movie ratings, and movie information (style kind, age and employment). In this case, we use the Movielens-1M dataset, which contains 1 million score records, to examine how the PS framework works.

4.2. Recall

We use the better interpretable average per user rather than the global average. The recall is evaluated here for two databases, one is amazon Dataset and the other is the MovieLens dataset. For both of these datasets, the recall is evaluated at metrics@20 and metrics@50.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

– Recall@20 for Amazon dataset

In Figure 3 and Table 1 we can see that the recall@20 is evaluated for Amazon Dataset, the method MostPopular method gives a value of 1.368 which is the least value, whereas we can see that ComiRec –SA denotes an average value of 5.103 and FAT technique shows a value of 5.177. However, the existing approach EA attains a value of 6.761, means–ends analysis (MEA) technique denotes a higher value than the existing system as 7.084 and in comparison, with the existing techniques, and our proposed model ensures better performance and gives a value of 13.89.

– Recall@50 for Amazon dataset

In Figure 4 and Table 2 the recall@20 is evaluated for Amazon Dataset, the method MostPopular method gives a 1.368 recall value, which is the least value; whereas the method self efficacy academic (SEA) method generates an average value of 8.172 and future aware diverse trends (FAT) technique gives a value of 8.161. However, the existing approach existing approach (EA) attains a value of 10.625, gives a higher value than the existing system as 11.054, in comparison with the existing techniques our proposed model ensures better performance and gives a value of 21.84.

– Recall@20 for MovieLens Dataset

In Figure 5 and Table 3 we can see that the recall@20 is evaluated for the MovieLens dataset, the method MostPopular method gives a value of 4.636, which is the least value, whereas the ComiRec -SA denotes an average value of 11.205 and the MIND technique shows a value of 11.155. However, the existing approach (EA) attains a value of 10.958, YouTube-DNN technique denotes a higher value than the existing system as 13.558, in comparison with the existing techniques our proposed model ensures better performance and gives a value of 20.41.

In Figure 6 and Table 4 we can see that the recall@50 is evaluated for MovieLens Dataset, the method MostPopular method gives a value of 10.829 which is the least value, whereas we can see that MIND denotes an average value of 23.889 and the SEA technique shows a value of 24.325. However, the existing approach EA attains a value of 23.478, YouTube-DNN technique denotes a higher value than the existing system as 26.596 and in comparison, with the existing techniques, and our proposed model ensures better performance and gives a value of 47.05.

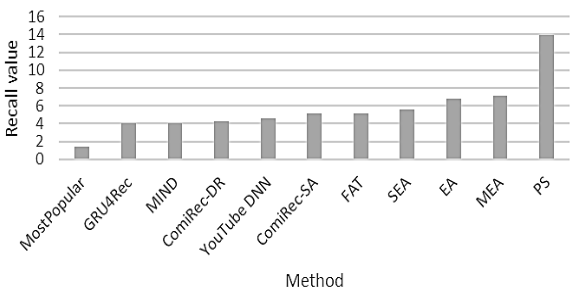

Figure 3. Comparison recall@20 for Amazon dataset

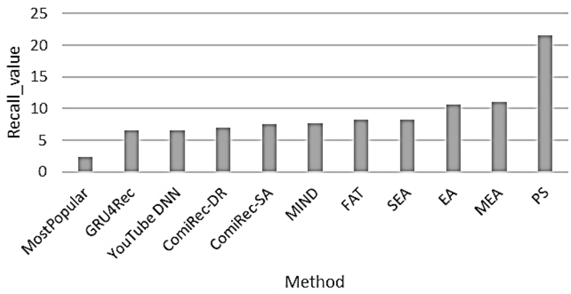

Figure 4. Comparison of recall@50 for various techniques

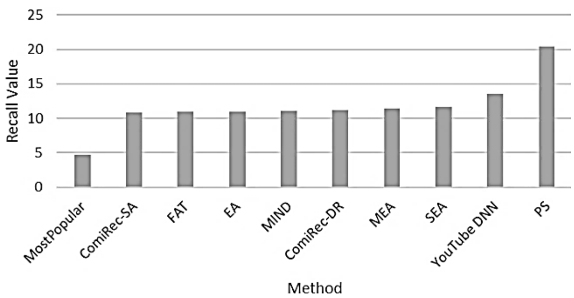

Figure 5. Comparison of recall@20 for MovieLens dataset

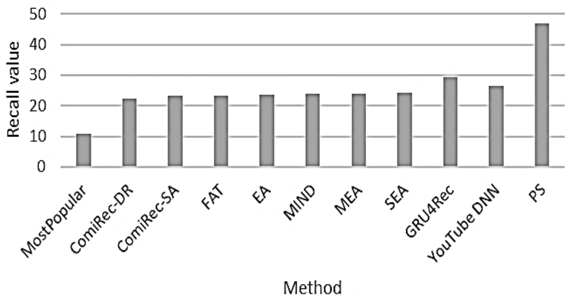

Figure 6. Comparison of recall@50 for MovieLens dataset

Table 1. Comparison recalls@20 for various technique

Method	Recall@20	Method	Recall@20
MostPopular	1.368	FAT	5.177
GRU4Rec	4.057	SEA	5.576
MIND	4.035	EA	6.761
ComiRec-DR	4.299	MEA	7.084
YouTube DNN	4.573	PS	13.89
ComiRec-SA	5.103		

Table 2. Comparison of recall@50 for various techniques

Method	Recall@50	Method	Recall@50
MostPopular	2.4	FAT	8.161
GRU4Rec	6.501	SEA	8.172
YouTube DNN	6.544	EA	10.625
ComiRec-DR	6.957	MEA	11.054
ComiRec-SA	7.458	PS	21.48
MIND	7.638		

Table 3. Comparison of recall@20 for MovieLens dataset

Method	Recall@20	Method	Recall@20
MostPopular	4.636	ComiRec-DR	11.205
ComiRec-SA	10.834	MEA	11.366
FAT	10.892	SEA	11.645
EA	10.985	YouTube DNN	13.558
MIND	11.115	PS	20.41

Table 4. Comparison of recall@50 for MovieLens dataset

Method	Recall@50	Method	Recall@50
MostPopular	10.829	MEA	23.962
ComiRec-DR	22.391	SEA	24.325
ComiRec-SA	23.137	GRU4Rec	29.467
FAT	23.364	YouTube DNN	26.596
EA	23.478	PS	47.05
MIND	23.889		

4.3. NDCG

Normalized Discounted Cumulative Gain (NDCG) considers the location of the correct project. In Figure 7 and Table 5 the NDCG@20 is evaluated for Amazon Dataset, the method MostPopular method gives a value of 2.259, which is the least value, whereas we can see that ComiRec-DR gives an average value of 4.508. The FAT technique shows a value of 4.07, however, the existing approach EA attains a value of 5.613, MIND technique denotes a higher value than the existing system as 7.933, in comparison with the existing techniques our proposed model ensures better performance and gives a value of 7.95.

In Figure 8 and Table 6 the NDCG@50 is evaluated for Amazon dataset, the method MostPopular method gives a value of 3.936, which is the least value, whereas YouTube DNN denotes an average value of 5.039 and ComiRec-DR technique shows a value of 5.591. However, the existing approach EA attains a value of 7.184, MEA technique denotes a higher value than the existing system as 7.238, in comparison with the existing techniques our proposed model ensures better performance and gives a value of 10.02.

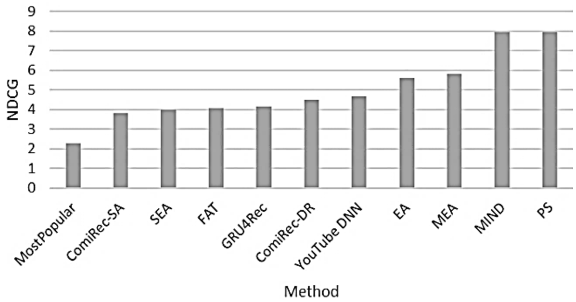

Figure 7. Comparison of NDCG@20 for MovieLens dataset

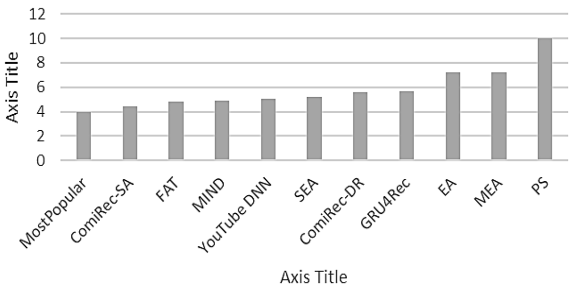

Figure 8. Comparison of NDCG@50 for MovieLens dataset

Table 5. Comparison of NDCG@20 for MovieLens dataset

Method	NDCG@20	Method	NDCG@20
MostPopular	2.259	YouTube DNN	4.673
ComiRec-SA	3.807	EA	5.613
SEA	3.99	MEA	5.816
FAT	4.07	MIND	7.933
GRU4Rec	4.163	PS	7.95
ComiRec-DR	4.508		

Table 6. Comparison of NDCG@50 for MovieLens dataset

Method	NDCG@50	Method	NDCG@50
MostPopular	3.936	ComiRec-DR	5.591
ComiRec-SA	4.41	GRU4Rec	5.678
FAT	4.824	EA	7.184
MIND	4.907	MEA	7.238
YouTube DNN	5.039	PS	10.02
SEA	5.231		

4.4. Comparative analysis

The comparative analysis carried out here by the proposed system with the existing system, for Amazon Dataset the metric Recall@20 the existing system gives a value of 6.761. Table 7 shows the comparative analysis. Our proposed model denotes a value of 13.89 the improvisation done is 69.0427%, for Recall@50 the existing system gives a value of 10.625 and our proposed model denotes a value of 21.48 the improvisation done is 67.62%. For MovieLens Dataset the metric Recall@20 the existing system gives a value of 10.895 and our proposed model gives a value of 20.41 the improvisation done is 60.0414%, MovieLens Dataset the metric Recall@50 the existing system gives a value of 23.478 and our proposed model denotes a value of 47.05 the improvisation done is 66.8444%. , for Amazon Dataset the metric NDCG@20 the existing system gives a value of 5.613 and our proposed model denotes a value of 7.95 the improvisation done is 34.4614%. , for Amazon Dataset the metric NDCG@50 the existing system gives a value of 7.184 and our proposed model denotes a value of 10.02 the improvisation done is 32.9691%, upon the comparative analysis of the proposed system with the existing system it can be concluded that our proposed model outperforms the existing model.

Table 7. Comparative analysis

Dataset Details	ES	DeepNCF	Improvisation
Amazon Dataset Recall@20	6.761	13.89	69.0427%
Amazon Dataset Recall@50	10.625	21.48	67.62%
MovieLens Dataset Recall@20	10.985	20.41	60.0414%
MovieLens Dataset Recall@50	23.478	47.05	66.8444%
Amazon Dataset NDCG@ 20	5.613	7.95	34.4614%
Amazon Dataset NDCG@ 50	7.184	10.02	32.9691%

5. CONCLUSION

In this paper, a DeepNCF model is developed that addresses the several issues such as overfitting, sparsity and performance metrics; DeepNCF incorporated the dual mode network into the collaborative filtering; DeepNCF designs user-item dual mode network from rating matrix; further considering the edges in network, it carries out edge clustering through custom weighted dual model modularity for higher densing. Furthermore, GNN is utilized for learning the complex relation among the user and items. The performance is better of the DeepNCF model in comparison with other existing approaches on two datasets. The edge-clustering algorithm is applicable here for developing a dual mode graph and implementing a edge reduction algorithm for every single community. The DeepNCF attains higher metrics in comparison with the existing methods in aspects of recommendation accuracy and variation in attaining the item from the customer for rating data and discarding irrelevant information.

REFERENCES

[1] A. Gatzioura, J. Vinagre, A. M. Jorge, and M. Sanchez-Marre, "A Hybrid Recommender System for Improving Automatic Playlist Continuation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 5, pp. 1819–1830, 2021, doi: 10.1109/TKDE.2019.2952099.

[2] K. Al Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, "Hybrid recommender system for tourism based on big data and AI: A conceptual framework," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 47–55, 2021, doi: 10.26599/BDMA.2020.9020015.

[3] Z. Yang and M. Zhang, "TextOG: A Recommendation Model for Rating Prediction Based on Heterogeneous Fusion of Review Data," *IEEE Access*, vol. 8, pp. 159566–159573, 2020, doi: 10.1109/ACCESS.2020.3020942.

[4] X. Yang, Y. Guo, Y. Liu, and H. Steck, "A survey of collaborative filtering based social recommender systems," *Computer Communications*, vol. 41, pp. 1–10, 2014, doi: 10.1016/j.comcom.2013.06.009.

[5] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. S. Chua, "Neural collaborative filtering," in *26th International World Wide Web Conference, WWW 2017*, in WWW '17. International World Wide Web Conferences Steering Committee, Apr. 2017, pp. 173–182. doi: 10.1145/3038912.3052569.

[6] M. Iqbal *et al.*, "Kernel Context Recommender System (KCR): A Scalable Context-Aware Recommender System Algorithm," *IEEE Access*, vol. 7, pp. 24719–24737, 2019, doi: 10.1109/ACCESS.2019.2897003.

[7] L. Wu, L. Chen, R. Hong, Y. Fu, X. Xie, and M. Wang, "A Hierarchical Attention Model for Social Contextual Image Recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 1854–1867, 2020, doi: 10.1109/TKDE.2019.2913394.

[8] J. Zhang, X. Shi, S. Zhao, and I. King, "STAR-GCN: Stacked and reconstructed graph convolutional networks for recommender systems," in *IJCAI International Joint Conference on Artificial Intelligence*, in IJCAI-2019, vol. 2019- August. International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 4264–4270. doi: 10.24963/ijcai.2019/592.

[9] W. Wang, G. Zhang, and J. Lu, "Hierarchy Visualization for Group Recommender Systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 6, pp. 1152–1163, 2019, doi: 10.1109/TSMC.2017.2699690.

[10] B. Xiao, X. Xie, C. Yang, and Y. Wang, "RTN-GNNR: Fusing Review Text Features and Node Features for Graph Neural Network Recommendation," *IEEE Access*, vol. 10, pp. 114165–114177, 2022, doi: 10.1109/ACCESS.2022.3218882.

[11] A. Breitfuss, K. Errou, A. Kurteva, and A. Fensel, "Representing emotions with knowledge graphs for movie recommendations," *Future Generation Computer Systems*, vol. 125, pp. 715–725, 2021, doi: 10.1016/j.future.2021.06.001.

[12] S. Wu, Y. Tang, Y. Zhu, L. Wang, X. Xie, and T. Tan, "Session-based recommendation with graph neural networks," *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, vol. 33, no. 01, pp. 346–353, 2019, doi: 10.1609/aaai.v33i01.3301346.

[13] C. Xu *et al.*, "Graph contextualized self-attention network for session-based recommendation," in *IJCAI International Joint Conference on Artificial Intelligence*, in IJCAI-2019, vol. 2019- August. International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 3940–3946. doi: 10.24963/ijcai.2019/547.

[14] R. Qiu, Z. Huang, J. Li, and H. Yin, "Rethinking the item order in session-based recommendation with graph neural networks," in *International Conference on Information and Knowledge Management, Proceedings*, in CIKM '19. ACM, Nov. 2019, pp. 579–588. doi: 10.1145/3357384.3358010.

[15] X. He, K. Deng, X. Wang, Y. Li, Y. D. Zhang, and M. Wang, "LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation," in *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, in SIGIR '20. ACM, 2020, pp. 639–648. doi: 10.1145/3397271.3401063.

[16] Z. Duan, H. Xu, Y. Huang, J. Feng, and Y. Wang, "Multivariate Time Series Forecasting with Transfer Entropy Graph," *Tsinghua Science and Technology*, vol. 28, no. 1, pp. 141–149, 2023, doi: 10.26599/TST.2021.9010081.

[17] S. M. Al-Ghuribi and S. A. Mohd Noah, "Multi-Criteria Review-Based Recommender System-The State of the Art," *IEEE Access*, vol. 7, pp. 169446–169468, 2019, doi: 10.1109/ACCESS.2019.2954861.

[18] S. Fan *et al.*, "Metapath-guided heterogeneous graph neural network for intent recommendation," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '19. ACM, 2019, pp. 2478–2486. doi: 10.1145/3292500.3330673.

[19] W. Fan *et al.*, "Graph neural networks for social recommendation," in *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, in WWW '19. ACM, 2019, pp. 417–426. doi: 10.1145/3308558.3313488.

[20] J. Zhao *et al.*, "IntentGC: A scalable graph convolution framework fusing heterogeneous information for recommendation," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '19. ACM, 2019, pp. 2347–2357. doi: 10.1145/3292500.3330686.

[21] F. Luo, G. Ranzi, X. Wang, and Z. Y. Dong, "Social Information Filtering-Based Electricity Retail Plan Recommender System for Smart Grid End Users," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 95–104, 2019, doi: 10.1109/TSG.2017.2732346.

[22] M. B. Hossain, M. S. Arefin, I. H. Sarker, M. Kowsher, P. K. Dhar, and T. Koshiba, "CARAN: A Context-Aware Recency-Based Attention Network for Point-of-Interest Recommendation," *IEEE Access*, vol. 10, pp. 36299–36310, 2022, doi:

10.1109/ACCESS.2022.3161941.

[23] P. Covington, J. Adams, and E. Sargin, “Deep neural networks for youtube recommendations,” in *RecSys 2016 - Proceedings of the 10th ACM Conference on Recommender Systems*, in RecSys ’16. ACM, Sep. 2016, pp. 191–198. doi: 10.1145/2959100.2959190.

[24] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, “Session-based recommendations with recurrent neural networks,” *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 2016.

[25] G. Zhou *et al.*, “Deep interest network for click-through rate prediction,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’18. ACM, 2018, pp. 1059–1068. doi: 10.1145/3219819.3219823.

[26] Y. Cen, J. Zhang, X. Zou, C. Zhou, H. Yang, and J. Tang, “Controllable Multi-Interest Framework for Recommendation,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’20. ACM, 2020, pp. 2942–2951. doi: 10.1145/3394486.3403344.

[27] Y. Lu *et al.*, “Future-aware diverse trends framework for recommendation,” in *The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW 2021*, in WWW ’21. ACM, Apr. 2021, pp. 2992–3001. doi: 10.1145/3442381.3449791.

[28] D. Yin and S. Feng, “Enhanced Attention Framework for Multi-Interest Sequential Recommendation,” *IEEE Access*, vol. 10, pp. 67703–67712, 2022, doi: 10.1109/ACCESS.2022.3185063.

BIOGRAPHIES OF AUTHOR

Shruthi Nagaraj is a Research Scholar in the school of Computer Science and engineering at Presidency University, Bengaluru. She has a Bachelor's degree in Computer Science Engineering from VTU, Master's degree in Software Engineering from Ramaiah Institute of Technology, VTU. She has two years of experience working in the IT industry in Communication and Network Security. Her area of interests in research are Recommendation Systems, Deep Learning and Machine Learning Technologies, Artificial Intelligence and Intelligent Information Retrieval. She can be contacted at: shru.nag95@gmail.com.

Blessed Prince Palayyan is an Associate Professor with the school of computer science and engineering at Presidency University. He has served in the education sector for more than 17 years at college level and among the top university in India and has served with university abroad. He has received his bachelor’s degree from Manonmaniam Sundaranar University, Master’s degree from Annamalai University and Doctorate from Anna University, Chennai. His research interest includes major areas of computer science, particularly security in computing, Recommendation System, Image Processing, Artificial Intelligence, Robotics and IoT. He can be contacted at this email: blessedprince@presidencyuniversity.in.