International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN - 2395-1990
Online ISSN : 2394-4099
OPEN ACCESS
Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET2310661

Enhancing Image Recommendations in Social Contexts using Hierarchical Attention

Gelasam Usha*, Dr. D. Shobha Rani
Department of CSE, Chadalawada Ramanamma Engineering College, Tirupati, Andhra Pradesh, India

ARTICLEINFO

Article History :

Accepted: 01 Dec 2023
Published: 26 Dec 2023

Publication Issue :
Volume 10, Issue 6
November-December-2023
Page Number :
343-351

ABSTRACT

The paper introduces a novel approach to address the computational challenges faced by recommender systems operating on large-scale datasets, particularly in the context of social contextual image recommendation. Recognizing the need for a more efficient means of comparing numerous items to identify users' preferences, the proposed hierarchical consideration model delves into three crucial factors: transfer history, social impact, and owner adoration. These factors encapsulate nuanced aspects of user preferences, deriving from intricate relationships between users and images. To operationalize this, a hierarchical attention network is designed, explicitly reflecting the hierarchical nature of users' latent interests within the identified key aspects. Leveraging embeddings from state-of-the-art deep learning models tailored for different data types, the hierarchical attention network dynamically adjusts its focus on varying content levels. Extensive experimentation on real-world datasets underscores the model's superiority, with compelling results demonstrating its effectiveness and adaptability, particularly in contrast to existing approaches. The study culminates in highlighting the model's prowess in navigating diverse data landscapes, establishing its potential as an advanced solution for large-scale recommender systems.

Keywords : Image Recommendation, social network, intention modeling, General Recommendation, Hierarchical Attentive Social Contextual recommendation.

I. INTRODUCTION

In the era of mobile Internet, the pervasive practice of capturing and sharing daily life events through user-generated images on social media platforms has surged. The prevalence of image tweets, exemplified by platforms like Twitter and Sina Weibo, underscores the growing significance of understanding images for downstream applications such as event detection, image tweet retrieval, and recommendation [1]. Despite the challenges posed by the technical complexities of mining semantics from images, the primary hurdle lies in the unique social media context. While image recognition algorithms can provide visual tags, the ever-evolving nature of social media amplifies the importance of contextual information, encompassing events and intent, for a comprehensive understanding of image tweets [2-5]. For instance, a seemingly innocent image may carry significant socio-political implications, emphasizing the need to go beyond visual tags. The paper focuses on addressing the intricate task of comprehending users' preferences for images and recommending images within social image-based platforms [6]. It highlights the dual role of users as both consumers and creators of images, emphasizing the importance of understanding upload behavior. Users also form social networks, sharing their image preferences, and the rich heterogeneous contextual information in these networks offers valuable insights into interpreting users' preferences for highly subjective content [7]. However, the challenge remains in delineating the heterogeneous social contextual aspects influencing users' preferences. Moreover, in the preference selection process, different users prioritize various social contextual perspectives for personalized image preferences. The complexity is further compounded by unique user preferences with individuals like Lily favoring images similar to her uploads, while others, like Bob, are influenced by social neighbors to align their preferences with their social circle. This intricate interplay of user preferences and social contextual factors poses a significant challenge in the realm of image recommendation within social media platforms.

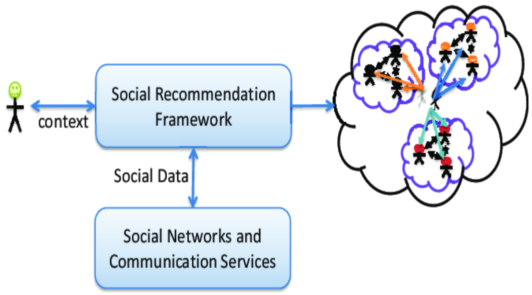

Figure 1: An overall framework of social contextual image recommendation.

In the proposed hierarchical structure, the approach involves the initial formation of auxiliary angle representations for each user, followed by the aggregation of these three viewpoint representations into an auxiliary user interest vector. Fig.1 shows this two-step process enhances the modeling of user preferences in the recommendation system. First, individual client profiles are constructed by considering three key perspectives—transfer history, social impact, and proprietor adoration—resulting in specialized auxiliary angle representations. Subsequently, these representations are combined into a comprehensive auxiliary user interest vector, capturing the multifaceted nature of user preferences [8]. Second, recognizing that different components within each perspective and varying importance of different perspectives contribute differentially to each user's recommendation process, the hierarchical attention network incorporates two levels of attention mechanisms. At the element level, attention is applied to specific features within each perspective, allowing the model to discern the significance of diverse elements. Simultaneously, at the angle level, attention is directed towards the overarching perspectives, accommodating the unique preferences of each user. This hierarchical attention architecture ensures a nuanced understanding of users' latent interests, contributing to a more refined and personalized recommendation system [9].

II. RELATED WORKS

Image Recommendation: In many image based social networks, images are associated with rich context in-formation, e.g., the text in the image, the hash tags. Re-searchers proposed to apply factorization machines for image recommendation by considering the rich context information [6]. Recently, deep Convolution Neural Net- works (CNNs) have been successfully applied to analyzing visual imagery by automatic image representation in the modeling process [27]. Thus, it is a natural idea to leverage visual features of

CNNs to enhance image recommendation performance [18], [28], [17], [5]. E.g., VBPR is an extension of BPR for image recommendation, on top of which it learned an additional visual dimension from CNN that modeled users' visual preferences [18]. There are some other image recommendation models that tackled the temporal dynamics of users' preferences to images over time [17], or users' location preferences for image recommendation [15], [25], [17]. As well studied in the computer vision community, in parallel to the visual content information from deep CNNs, images convey rich style information. Researchers showed that many brands post images that show the philosophy and lifestyle of a brand [14], images posted by users also reflect users' personality [13]. Recently, Gatys et al. proposed a new model of extracting image styles based on the feature maps of convolution neural networks [10]. The proposed model showed high perceptual quality for extracting image style, and has been successfully applied to related tasks, such as image style transfer [11], and high-resolution image stylisation [12]. We argue that the visual image style also plays a vital role for evaluating users' visual experience in recommender systems. Thus, we leverage both the image content and the image style for recommendation.

Social Contextual Recommendation: Social scientists have long converged that a user's preference is similar to or influenced by her social connections, with the social theories of homophile and social influence [3]. With the prevalence of social networks, a popular research direction is to leverage the social data to improve recommendation performance [33], [23], [24], [11]. E.g., Ma et al. proposed a latent factor based model with social regularization terms for recommendation [23]. Since most of these social recommendation tasks are formulated as non- convex optimizing problems, researchers have designed an unsupervised deep learning model to initialize model parameters for better performance [9]. Besides, ContextMF is proposed to fuse the individual preference and interpersonal influence with auxiliary text content information from social networks [24]. As the implicit influence of trusts and ratings are valuable for recommendation, TrustSVD is proposed to incorporate the influence of trusted users on the prediction of items for an active user [16]. The proposed technique extended the SVD++ with social trust information. Social recommendation has also been considered with social circle [28], online social recommendation [19], social network evolution [20], and so on.

Besides, as the social network could be seen as a graph, the recent surge of network embedding is also closely related to our work [8]. Network embedding models encode the graph structural information into a low latent space, such that each node is represented as an embedding in this latent space. Many network embedding models have been proposed [37], [44], [48], [47]. The network embedding could be used for the attention networks. We distinguish from these works as the focus of this paper is not to advance the sophisticated network embedding models. We put emphasis on how to enhance recommendation performance by leveraging various data embeddings.

Attention Mechanism: Neural science studies have shown that people focus on specific parts of the input rather than using all available information [22]. Attention mechanism is such intuitive ideas that automatically models and selects the most pertinent piece of information, which learns to assign attentive weights for a set of inputs, with higher (lower) weights indicate that, the corresponding in- puts are more informative to generate the output. Attention mechanism is widely used in many neural network based tasks, such as machine translation [4] and image captioning [13]. Recently, the attention mechanism is also widely used for recommender systems [19], [12], [23], [21]. Given the classical collaborative filtering scenario with user-item interaction behavior, NAIS extended the classical item based recommendation models by distinguishing the importance of different historical items in a user profile [19]. With users' temporal behavior, the attention networks were proposed to learn which historical behavior is more important for the user's current temporal decision [31], [32]. A lot of attention based recommendation models have been developed to better exploit the auxiliary information to improve recommendation performance. E.g., ANSR is proposed with a social attention module to learn adaptive social influence strength for social recommendation [23]. Given the review or the text of an item, attention networks were developed to learn informative sentences or words for recommendation [15], [11]. While the above models perform the standard vanilla attention to learn to attend on a specific piece of information, the co-attention mechanism is concerned to learn attention weights from two sequences [21], [26], [36]. E.g., in the hash tag recommendation with text and image information, the co-attention network is designed to learn which part of the text is distinctive for images, and simultaneously the important visual features for the text [56]. Besides, researchers have made a comprehensive survey the attention based recommendation models [37]. In some real- world applications, there exists hierarchical structure among the data, several pioneering works have been proposed to deal with this kind of relationship [14], [29]. E.g., a hierarchical attention model is proposed to model the hierarchical relationships of word, sentence and document for document classification [54]. Our work borrows ideas from the atten- tion mechanism, and we extend this idea by designing a hierarchical structure to model the complex social contextual aspects that influence users' preferences. Nevertheless, different from the natural hierarchical structure of words, sentences and documents in natural language processing, the hierarchical structure that influences a user's decision from complex heterogeneous data sources is summarized by our proposed model. Specifically, our proposed model has a two-layered hierarchical structure with the bottom layer attention network that summarizes each aspect from the various elements of this aspect. By taking the output of each aspect from the bottom layer, the top-layer attention network learns the importance of the three aspects.

The work that is most similar to ours is the Attentive Collaborative Filtering (ACF) for image and video recommendation [5]. By assuming there exists item level and component level implicitness that underlines a user's preference, an attention based recommendation model is pro- posed with the component level attention and the item level attention. Our work borrows the idea of applying attention mechanism for recommendation, and it differs from ACF and previous works from both the research perspective and the application point. From the technical perspective, we model the complex social contextual aspects of users' interests from heterogeneous data sources in a unified recommendation model. In contrast, ACF only leverages the image (video) content information. From the application view, our proposed model could benefit researchers and engineers in related areas when heterogeneous data are available.

III.PROPOSED METHODOLOGY

The proposed model builds upon the well-established foundation of dormant factor-based models, which operate on the premise that clients and items can be predicted within a low latent space. In contrast to traditional models, our approach extends beyond the fundamental dormant client interest vector. For each user, we identify and incorporate three key aspects—transfer history, social impact, and proprietor profound respect—that collectively influence their preferences. Each of these perspectives captures a contextual factor derived from the intricate relationships among users and images, providing a more comprehensive understanding of user behavior [18], [19]. Specifically, the transfer history perspective focuses on condensing each user's uploaded images to effectively characterize their interests. This nuanced consideration of transfer history enhances the model's ability to capture and interpret user preferences, contributing to a more sophisticated and context-aware recommendation system.

The Hierarchical Attentive Social Contextual recommendation (HASC) model, as illustrated in Fig. 2, constitutes a novel hierarchical neural network designed to capture users' preferences for unfamiliar images through two levels of attention with integrated social contextual modeling. At the top layer of the attention network, the model delineates the significance of three contextual viewpoints—transfer history, social impact, and maker deference—for users' decision-making. This top-layer attention is derived from the lower-layer attention networks that aggregate intricate elements within each perspective [19]. For a given user and image I, featuring the three identified social contextual viewpoints, attentive degrees (denoted as α[l], where l = 1, 2, 3) indicate the importance of each perspective at the top layer (depicted as the perspective significance attention with the orange component in the figure).

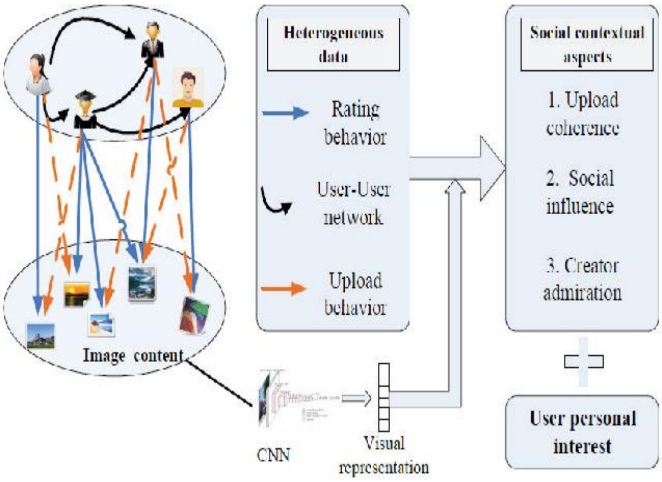

Figure 2: Architecture of Proposed Methodology for Social Contextual Image recommendation

Algorithm 1 encapsulates the model's focus on capturing two specific attributes. Firstly, recognizing the hierarchical structure inherent in social contextual recommendation, where diverse elements contribute to each viewpoint and users exhibit three distinct aspects, the model constructs a hierarchical client interest representation. In this hierarchical structure [20], auxiliary angle representations for each user are initially formed, and subsequently, the three viewpoint representations are aggregated into an auxiliary client interest vector. Secondly, acknowledging that different elements within each perspective and various perspectives hold varying levels of in formativeness for each user in the recommendation process, the hierarchical attention network incorporates two levels of attention mechanisms. At the element level, attention is directed towards specific features within each perspective, capturing their individual significance. Simultaneously, at the viewpoint level, attention is applied to the overarching perspectives, accommodating the diverse preferences of each user. This dual-level attention architecture ensures a nuanced understanding of user preferences, fostering a recommendation system that is sensitive to both individual elements and overarching contextual factors.

The algorithm steps for the Hierarchical Attentive Social Contextual recommendation (HASC) model:

Algorithm 1: the learning algorithm of HASC

Input:

- (U): Set of users
- (I): Set of items (images)
- (C): Set of social contextual features
- (Y): Set of user-item interactions or preferences

Parameters:

- (Theta): Model parameters, including weights and biases
- (alpha): Attention weights
- (beta): Attention weights at the viewpoint level

Algorithm Steps:

1. Initialization:

- Initialize model parameters (Theta) randomly or using pre-trained embeddings.
- Initialize attention weights (alpha) and (beta) with small random values.

2. Forward Pass:

- For each user (u) and item (i), perform the following steps:
- Embed user and item features using neural networks tailored for each type of data.

- Compute element-level attention weights (alpha[l]) for each social contextual feature using the attention mechanism.
- Combine the element-level attention weights with the features to obtain weighted representations.
- Compute viewpoint-level attention weights (beta) using the attention mechanism.
- Combine the weighted representations with viewpoint-level attention weights to obtain the final user interest vector.

3. Prediction:
- Use the obtained user interest vector to predict the user's preference for the given item.
- Compute the predicted preference (hat{y}_{ui}).

4. Loss Calculation:
- Compute the loss between the predicted preference (hat{y}_{ui}) and the actual user feedback (y_{ui}) using an appropriate loss function.

5. Backward Pass (Backpropagation):
- Compute the gradients of the loss with respect to the model parameters (Theta), attention weights (alpha) and (beta).
- Update the model parameters using an optimization algorithm such as stochastic gradient descent (SGD) or its variants.

6. Repeat:
- Repeat steps 2-5 for a predefined number of epochs or until convergence.

7. Evaluation:
- Periodically evaluate the model's performance on a validation set to monitor generalization.

8. Fine-tuning (Optional):
- Adjust the model architecture or hyper parameters based on the evaluation results for potential improvement.

9. Testing:
- Evaluate the final trained model on a separate test set to assess its performance in a real-world scenario.

This algorithm outlines the training and prediction steps for the HASC model, emphasizing the hierarchical attention mechanism's role in capturing the intricate relationships among users, items, and social contextual features for improved recommendation accuracy.

IV.RESULT ANALYSIS

The evolution of cloud computing has brought about numerous benefits, with cloud storage being a particularly convenient technology that enables users to expand their storage capacity. However, the use of cloud storage also introduces a series of security challenges. Users relinquish direct control over the physical storage of their data, leading to a separation of ownership and management of information. Fig 3 to address the security concerns associated with cloud storage, we propose a Transport Layer Security (TLS) framework based on a fog computing model, complemented by the design of a Hash-Solomon algorithm. Through theoretical security analysis, the proposed design has been proven to be feasible. By intelligently distributing the proportion of data blocks stored across different servers, we ensure the security of data on each server. Moreover, the theoretical impossibility of breaking the encoding matrix reinforces the robustness of our security approach. This TLS framework and Hash-Solomon algorithm offer a promising solution to enhance the security of data in cloud storage environments.

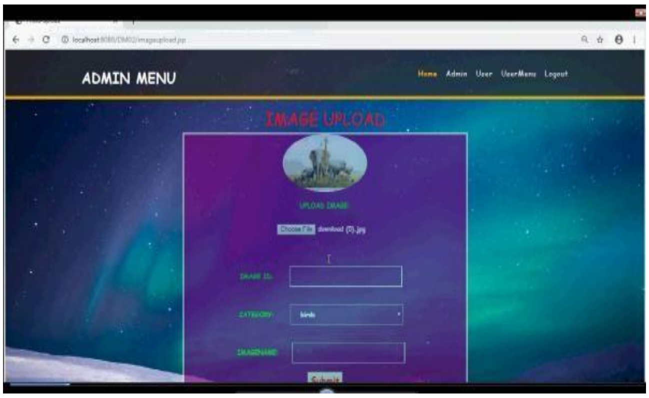

Figure 3(a): Results for Image Recommendations.

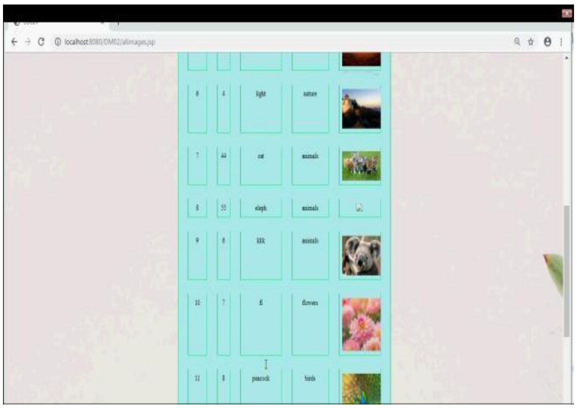

Figure 3(b): Results for Image Recommendations.

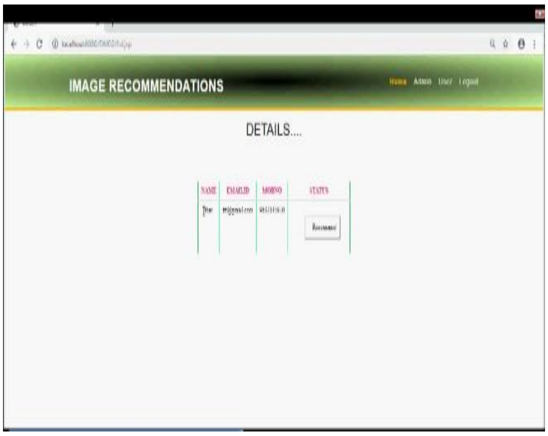

Figure 3(c): Results for Image Recommendations.

V. CONCLUSION

In this paper, we have introduced a novel approach to social contextual image recommendation through the development of a Hierarchical Attentive Social Contextual (HASC) model. Going beyond conventional user interest modeling, our model identifies and incorporates three critical social contextual aspects that shape a user's preference for images from diverse data sources: the upload history aspect, the social influence aspect, and the owner admiration aspect. The cornerstone of our model is a carefully designed hierarchical attention network, adept at naturally capturing the hierarchical relationship of users' interests given these three identified aspects. Importantly, we leverage data embeddings from rich and heterogeneous data sources, enabling the hierarchical attention networks to dynamically adapt their focus to different levels of content importance. Through extensive experiments conducted on real-world datasets, our proposed HASC model consistently surpasses various state-of-the-art baselines for image recommendation. This demonstrates the efficacy of our model in effectively capturing and utilizing social contextual information, offering a superior solution for personalized and context-aware image recommendations.

VI.REFERENCES

[1] flickr statistics. https://expandedramblings.com/index.php/ flickr- stats/, 2017. [online; accessed 20-jan-2018].

[2] P. Krishna Kishore, S. Ramamoorthy, V.N. Rajavarman, "Mitigation of HTTP Flood DDoS Attack in Application Layer Using Machine Learning and Isolation Forest," SSRG International Journal of Electrical and Electronics Engineering, vol. 10, no. 10, pp. 6-19, 2023. Crossref, https://doi.org/10.14445/23488379 /IJEEE-V10I10P102

[3] Krishna Kishore, P., Prathima, K., Eswari, D.S., Goud, K.S. (2023). Bidirectional LSTM-Based Sentiment Analysis of Context-Sensitive Lexicon for Imbalanced Text. In: Bhateja, V., Sunitha, K.V.N., Chen, YW., Zhang, YD. (eds) Intelligent System Design. Lecture Notes in Networks and Systems, vol 494. Springer, Singapore. https://doi.org/10.1007/978-981-19-4863-3_27

[4] d. bahdanau, k. cho, and y. bengio. neural machine translation by jointly learning to align and translate. in iclr, 2015.

[5] j. chen, h. zhang, x. he, l. nie, w. liu, and t.-s. chua. attentive collaborative filtering: multimedia recommendation with itemand component-level attention. in sigir, pages 335–344. acm, 2017.

[6] t. chen, x. he, and m.-y. kan. context-aware image tweet modelling and recommendation. in mm, pages 1018–1027. acm, 2016.

[7] P. Krishna Kishore, S. Ramamoorthy, V.N. Rajavarman, ARTP: Anomaly based real time prevention of Distributed Denial of Service attacks on the web using machine learning approach, International Journal of Intelligent Networks, Volume 4, 2023, Pages 38-45, ISSN 2666-6030, https://doi.org/10.1016/j.ijin.2022.12.001.

[8] p. cui, x. wang, j. pei, and w. zhu. a survey on network embedding. tkde, 2018.

[9] s. deng, l. huang, g. xu, x. wu, and z. wu. on deep learning for trust- aware recommendations in social networks. tnnlss, 28(5):1164–1177, 2017.

[10] gatys, a. s. ecker, and m. bethge. texture synthesis using convolutional neural networks. in nips, pages 262–270, 2015.

[11] a. gatys, a. s. ecker, and m. bethge. image style transfer using convolutional neural networks. in cvpr, pages 2414–2423, 2016.

[12] a. gatys, a. s. ecker, m. bethge, a. hertzmann, and e. shechtman. controlling perceptual factors in neural style transfer. in cvpr, pages 3985–3993, 2017.

[13] F. Gelli, X. He, T. Chen, and T.-S. Chua. How personality affects our likes: Towards a better understanding of actionable images. In MM, pages 1828–1837. ACM, 2017.

[14] F. Gelli, T. Uricchio, X. He, A. Del Bimbo, and T.-S. Chua. Beyond the product: Discovering image posts for brands in social media. In MM. ACM, 2018.

[15] Y. Gong and Q. Zhang. Hashtag recommendation using attentionbased convolutional neural network. In IJCAI, pages 2782– 2788, 2016.

[16] G. Guo, J. Zhang, and N. Yorke-Smith. A novel recommendation model regularized with user trust and item ratings. TKDE, 28(7):1607–1620, 2016.

[17] R. He, C. Fang, Z. Wang, and J. McAuley. Vista: a visually, socially, and temporally-aware model for artistic recommendation. In Recsys, pages 309–316. ACM, 2016.

[18] R. He and J. McAuley. Vbpr: Visual bayesian personalized ranking from implicit feedback. In AAAI, pages 144–150, 2016.

[19] X. He, Z. He, J. Song, Z. Liu, Y.-G. Jiang, and T.-S. Chua. Nais: Neural attentive item similarity model for recommendation. TKDE, 2018.

[20] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua. Neural collaborative filtering. In WWW, pages 173–182, 2017.

[21] B. Hu, C. Shi, W. X. Zhao, and P. S. Yu. Leveraging meta- path based context for top- n recommendation with a neural co- attention model. In SIGKDD, pages 1531–1540. ACM, 2018.

[22] Munivara Prasad, K., Samba Siva, V., Krishna Kishore, P., Sreenivasulu, M. (2019). DITFEC: Drift Identification in Traffic-Flow Streams for DDoS Attack Defense Through Ensemble Classifier. In: Peng, SL., Dey, N., Bundele, M. (eds) Computing and Network Sustainability. Lecture Notes in Networks and Systems, vol 75. Springer, Singapore. https://doi.org/10.1007/978-981-13-7150-9_32

[23] M. Jamali and M. Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In RecSys, pages 135–142. ACM, 2010.

[24] M. Jiang, P. Cui, F. Wang, W. Zhu, and S. Yang. Scalable recom- mendation with social contextual information. TKDE, 26(11):2789– 2802, 2014.

[25] T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. In ICLR, 2017.

[26] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In KDD, pages 426–434. ACM, 2008.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classifi- cation with deep convolutional neural networks. In NIPS, pages 1097–1105, 2012.

[28] C. Lei, D. Liu, W. Li, Z.-J. Zha, and H. Li. Comparative deep learning of hybrid representations for image recommendations. In CVPR, pages 2545–2553, 2016.
[29] J. Li, M.-T. Luong, and D. Jurafsky. A hierarchical neural autoen- coder for paragraphs and documents. arXiv:1506.01057, 2015.
[30] D. C. Liu, S. Rogers, R. Shiau, D. Kislyuk, K. C. Ma, Z. Zhong,J. Liu, and Y. Jing. Related pins at pinterest: The evolution of a real-world recommender system. In WWW, pages 583–592, 2017.
[31] Q. Liu, Y. Zeng, R. Mokhosi, and H. Zhang. Stamp: short-term attention/memory priority model for session-based recommenda- tion. In SIGKDD, pages 1831–1839. ACM, 2018.
[32] P. Loyola, C. Liu, and Y. Hirate. Modeling user session and intent with an attention-based encoder-decoder architecture. In RecSys, pages 147–151. ACM, 2017.
[33] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King. Recommender systems with social regularization. In WSDM, pages 287–296. ACM, 2011.
[34] A. Mnih and R. R. Salakhutdinov. Probabilistic matrix factoriza- tion. In NIPS, pages 1257– 1264, 2008.

Cite this article as :

Gelasam Usha, Dr. D. Shobha Rani, "Enhancing Image Recommendations in Social Contexts using Hierarchical Attention", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 6, pp. 343-351, November-December 2023. Available at doi : https://doi.org/10.32628/IJSRSET2310661
Journal URL : https://ijsrset.com/IJSRSET2310661