Beyond GLORY: Global Graph-Enhanced Personalized News Recommendations

Roye Katzav
royek@post.bgu.ac.il

Natalie Morad
moradna@post.bgu.ac.il

Dudi Biton
bitondud@post.bgu.ac.il

Abstract

The rapid evolution of online news platforms necessitates advanced recommendation systems to enhance user experience by providing personalized news content. The GLORY [1] method has established a significant framework by combining global and local news representations for personalized recommendations. Recognizing opportunities for further advancements, our research introduces modifications to the GLORY method by focusing on two main aspects: refining the loss function and optimizing the prediction mechanism. Specifically, we experiment with integrating a Pairwise ranking loss alongside or in place of the original loss function and adopt a Deep Neural Network (DNN) for prediction to potentially improve upon the method's accuracy in presenting relevant news articles to users. Conducted experiments on the MIND datasets reveal that our proposed modifications exhibit the potential to enhance the GLORY method's performance in some cases while achieving comparable results in others. This indicates a promising direction for refining news recommendation systems. Our work contributes to the ongoing exploration of optimizing recommendation algorithms, presenting a pathway for future enhancements in personalized news dissemination. The code for our enhancements to the GLORY method is documented in a publicly accessible repository[1].

1 Introduction

In the rapidly evolving landscape of online news consumption, personalized news recommendation systems play a crucial role in delivering content that aligns with individual user interests. The challenge lies in the dynamic nature of news content, characterized by its timeliness and the ever-changing preferences of users. Traditional content-based recommendation methods [2, 3, 4, 5, 6, 7], leveraging user-item interaction data, have made significant strides in addressing these challenges. However, they often rely on extracting semantic information from textual data, focusing predominantly on a user's local reading history. This approach, while effective, overlooks the global interaction patterns that offer a broader view of user interests and behaviors beyond the immediate content of news articles.

To address these gaps, Yang et al. presented GLORY [1] (Global-LOcal news Recommendation sYstem). The GLORY method marks a significant leap forward by leveraging a global news graph for personalized news recommendation. This global perspective enriches news representations and uncovers latent behavioral patterns, offering more relevant recommendations beyond the local view of a user's reading history.

While the GLORY method has made significant strides in addressing the limitations of traditional content-based recommendation systems by leveraging global and local news representations, it introduces its own set of challenges. Specifically, the GLORY framework employs a relatively simple prediction mechanism that relies on a SoftMax function and dot product between user embeddings and candidate news embeddings. This approach, while effective in certain contexts, does not facilitate

[1] https://github.com/dudi709/BeyondGLORY

complex learning that could further refine the personalization of news recommendations. Additionally, the use of Noise-Contrastive Estimation (NCE) loss within the GLORY method, though useful for distinguishing relevant from irrelevant articles, does not directly prioritize the ranking of articles by relevance. This aspect is critical in recommendation systems, where the goal is not only to identify relevant content but also to rank it according to the likelihood of user interest.

In this work, we propose enhancements to the GLORY model aimed at overcoming its limitations. First, we introduce a more sophisticated prediction mechanism that employs a Deep Neural Network (DNN), designed to enable smarter learning and a more nuanced understanding of user preferences and content relevance. Second, we incorporate a Pairwise ranking loss [8] alongside or in place of the existing NCE loss. This modification directly addresses the ranking by relevance, allowing for more effective prioritization of news articles that users are most likely to find engaging and relevant. Through these enhancements, our approach seeks to advance the personalization capabilities of the GLORY method, providing a pathway to more accurate news recommendation systems.

By integrating advanced loss functions and leveraging the predictive power of DNNs, we have developed a method that demonstrates comparable, and in some instances superior, performance to the original GLORY method. The main contributions of this article include:

- Demonstrating the efficacy of Pairwise Ranking loss [8] in enhancing the ranking accuracy of news recommendation systems.
- Exploring the potential of DNN-based prediction components to capture complex user-item interactions.

The remainder of this paper is organized as follows: Section 2 provides a detailed review of the background and related work, covering the evolution of personalized news recommendation systems. Section 3 introduces our proposed enhancements to the GLORY method, detailing both the theoretical underpinnings and the practical implementations of these modifications. In Section 4, we describe the experimental setup, including the datasets used, the computational environment, and the evaluation metrics. Section 5 presents the results of our experiments, comparing the performance of the original GLORY method with our enhanced version. Section 6 discusses these results, focusing on the implications of our findings and the limitations of our approach. Finally, Section 7 concludes the paper with a summary of our contributions and outlines directions for future research.

2 Background & Related Work

This section reviews the evolution of news recommendation systems, from feature-based to advanced neural and graph-based models. We highlight the GLORY method and discuss enhancements to improve its personalization capabilities.

2.1 Challenges in News Recommendation Systems

In the field of news recommendation systems, the primary goal is to provide users with personalized news content that aligns with their interests and preferences. These systems have evolved significantly with the advent of digital media, offering a solution to the information overload problem by filtering and prioritizing news articles for individual users. However, the domain of news recommendation presents several unique challenges that need to be addressed to enhance the effectiveness of these systems. In their comprehensive study, Chong Feng et al. [9] identified key categories of challenges in the literature, including:

1. The cold-start problem, arises when new users or articles have limited interaction data.
2. Scalability, ensuring the system can handle a large number of users and articles.
3. Recency, emphasizing the importance of providing timely and relevant news.
4. Content diversity, ensuring a varied selection of news topics.
5. User engagement, maintaining user interest and interaction with the system.
6. Evaluation metrics, developing appropriate measures to assess the performance of recommendation algorithms.

Addressing these challenges is crucial for the development of effective and user-centric news recommendation systems.

2.2 Feature-based News Recommendation

In news recommendation, feature-based methods leverage user and news features extracted from reading history, employing techniques such as LibFM [10] and DeepFM [11]. These methods are effective at modeling interactions between features, which is valuable for addressing the data sparsity prevalent in this field.

Factorization machines (FMs) are versatile in capturing complex interactions in high-dimensional spaces, while DeepFM [11] combines deep learning with FMs to learn high-level feature interactions. Despite their effectiveness, these methods face challenges such as computational complexity and the need for feature engineering.

Research in this area focuses on developing more efficient algorithms and automated feature engineering techniques to reduce complexity and improve recommendation accuracy. Feature-based news recommendation continues to be a promising approach for delivering personalized news content to users.

However, these methods face challenges in capturing the dynamic nature of user preferences and news content. To enhance feature representations, some approaches incorporate additional information like user demographics or news content features. Despite these enhancements, accurately capturing complex user-news interactions remains a challenge.

2.3 Neural NLP Techniques in News Recommendation

The abundance of textual information in news recommendation scenarios has prompted the adoption of content-based systems leveraging natural language processing (NLP) techniques. Models like NAML [12] employ convolutional neural networks (CNNs) [13] to extract features from news articles and gated recurrent units (GRUs) [14] to model user preferences. Similarly, NRMS [15] utilizes multi-head self-attention mechanisms to effectively model both news content and user interactions.

Furthermore, models like HieRec [4] introduce complex structures to enhance news and user modeling. HieRec uses a three-layer hierarchical structure to capture user preferences at different granularity levels. MINER [3] features a poly attention scheme with disagreement regularization and category-aware attention weighting, enabling the extraction of multiple user interest vectors for more precise recommendations. These advancements in neural network architectures and NLP techniques have significantly improved the accuracy and personalization of news recommendations.

2.4 Graph-based News Recommendation

In recent years, graph-based techniques have gained prominence in natural language processing (NLP), particularly in news recommendations [16, 17, 18]. These approaches leverage graph structures to capture more nuanced representations of news content and user interactions. For example, GERL [19] utilizes a dual approach, combining text-based methods to analyze news titles and topics with graph-based methods to extract insights from user-news interactions. Similarly, GNewsRec [20] employs a heterogeneous user-news-topic graph to uncover long-term user interests, while User-as-graph [21] introduces a novel graph pooling method to amalgamate a user's local news graph based on their reading history. DIGAT [22] further enhances content representation by exploiting a semantic augmentation graph for candidate news articles.

Additionally, the survey by Shoujin Wang et al. [23] provides a comprehensive overview of graph neural networks (GNNs) in NLP, highlighting their potential in various tasks, including news recommendation. This work emphasizes the growing importance of graph-based methodologies in understanding complex language structures and user interactions.

Despite these advancements, existing methods, including both content-based and graph-based approaches, often focus on modeling the current user, overlooking the global dynamics of news and the rich information available in global user-news interactions. Future research in graph-based news recommendation could explore ways to address this limitation by incorporating a more comprehensive understanding of global news dynamics.

2.5 GLORY

The GLORY [1] method represents a significant advancement in personalized news recommendation. It introduces a global perspective by leveraging a global news graph, which captures latent behavioral patterns across users and enriches news representations. This approach allows GLORY to provide more relevant and personalized recommendations by understanding user preferences and news content comprehensively. GLORY's unique strength lies in its ability to go beyond the local view of a user's reading history. It utilizes global information to enhance news representations, enabling the model to uncover hidden motivations and behaviors that are not evident from semantic information alone.

Despite the GLORY model's innovative approach to personalizing news recommendations through the use of global news graphs, there remains significant room for enhancement. Specifically, refining the model's ability to precisely rank and differentiate news articles based on user relevance, and understanding the complex interactions between user preferences and news content, presents a promising avenue for further research.

3 Method

In this section, we discuss the original GLORY method as presented by Yang et al. [1]. Subsequently, we propose a set of enhancements aimed at refining and advancing the original method. An overview of the original GLORY method is provided in Figure 1.

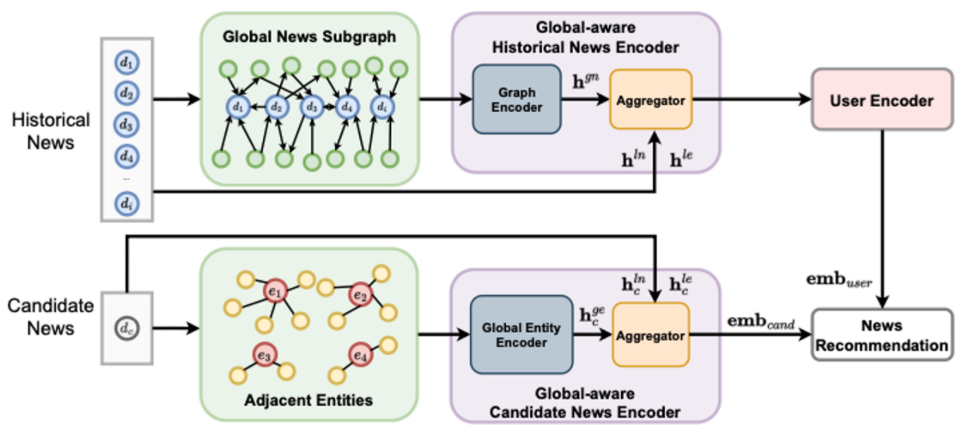

Figure 1: An Overview of GLORY

3.1 Original GLORY Method

The GLORY framework is designed to optimize news recommendation systems by improving the representation of both historical and candidate news. This improvement is achieved through a multi-layered approach that combines local and global information, employing advanced machine-learning techniques to create enriched data representations.

3.1.1 Local Representation Learning

GLORY starts by constructing local representations of news articles. This is done by initializing news article titles with pre-trained GloVe [24] embeddings, enabling the model to capture semantic nuances. Attention mechanisms, specifically multi-head self-attention, are then employed to aggregate these embeddings, resulting in a rich local news representation denoted as h_{ln}.

Similarly, local entity representations are derived using pre-trained TransE [25] embeddings sourced from WikiData [26]. These embeddings encapsulate relational data between entities, and the same attention framework is applied to distill this information into a local entity representation h_{le}.

3.1.2 Global Contextualization

To enhance the news recommendation system with a comprehensive understanding of user preferences and the wider news environment, the GLORY framework integrates two distinct graph structures: the Global News Graph and the Global Entity Graph. The **Global News Graph** aggregates users' reading histories into a graph $G_n = (V_n, E_n)$, where each node represents a news article, and the directed edges reflect the transition from one article to another by a user. Instead of using the entire global graph, GLORY extracts a subgraph relevant to the user's interests, allowing for a focused representation that is both efficient and personalized. The **Global Entity Graph** differs from the global news graph by not utilizing subgraphs but rather employing adjacent entities directly. This graph is undirected $G_e = (V_e, E_e)$ and captures the entities' interconnectedness. This direct use of adjacent entities avoids the complexity of subgraph extraction and provides a straightforward approach to incorporating global entity relationships into the recommendation process.

3.1.3 Encoding with Global Awareness

The GLORY method encodes both historical and candidate news with a global awareness that stems from the constructed graphs. Unlike the local encoding process that deals with articles and entities in isolation, global-aware encoding captures the interconnected nature of news consumption.

The **Global-Aware Historical News Encoder** employs a Gated Graph Neural Network (GGNN) to enhance the representation of historical news extracted from the user-specific subgraph. It then integrates global and local insights through an attention-pooling mechanism, ensuring a nuanced aggregation of information. The resulting output is subsequently processed by an attention mechanism, which generates the user's embedding. Conversely, the **Global-Aware Candidate News Encoder** processes the candidate news in a manner akin to the local entity encoding but enhances it by integrating global information from the entity graph. This method does not employ GGNNs but rather focuses on the adjacency of entities to provide a current and relevant representation of candidate news. Similar to the historical news encoder, it aggregates the global and local data using an attention-pooling network, resulting in candidates' news embedding.

3.1.4 Final Recommendation Prediction

The final recommendation of GLORY is achieved through the application of a dot product between the user embedding, emb_{user}, and the candidate news embedding, emb_{cand}. The resulting scores are then passed through a softmax function to obtain a probability distribution over the candidate news items, reflecting their relevance to the user's interests:

$$\hat{y}_i = softmax(emb_{user} \cdot emb_{cand})$$

3.2 Proposed Enhancements for GLORY

In this work, we propose a series of enhancements to the GLORY model to improve its performance in recommendation systems. Our focus is on: (1) refining the model's ability to differentiate and rank news articles based on their relevance to the user's interests; (2) learning more complex, non-linear interactions between user preferences and news candidate embeddings. The enhancements target two main components of the GLORY model: the loss function and the prediction component.

3.2.1 Adjustments to the Loss Function

The original GLORY model employed the Noise-Contrastive Estimation (NCE) loss function, which optimized the log-likelihood for positive samples in contrast to negative samples. The NCE loss function is defined as:

$$L_{NCE} = -\sum_{i=1}^{N} \log \frac{\exp(\hat{y}_i^+)}{\exp(\hat{y}_i^+) + \sum_{j=1}^{K_{neg}} \exp(\hat{y}_i^j)}$$

where $\hat{y}_i^+$ represents the model's output for the positive sample, K_{neg} is the number of negative samples, and $\hat{y}_i^j$ represents the model's output for the j-th negative sample.

While the NCE loss function was effective in distinguishing relevant from irrelevant candidates, it did not inherently rank candidates by their relevance, which is an important aspect of recommendation systems. To improve upon this, we introduce the Pairwise Ranking Loss [8].

3.2.2 Pairwise Ranking Loss

The Pairwise Ranking Loss [8] is introduced to ensure that relevant candidates are not only identified but also prioritized over irrelevant ones. This is achieved by penalizing the model when the score difference between relevant (positive) and less relevant (negative) samples is less than a predefined margin. The pairwise ranking loss function [8] is formulated as follows:

$$L_{PairwiseRanking} = \frac{1}{N \cdot K_{neg}} \sum_{i=1}^{N} \sum_{j=1}^{K_{neg}} \max(0, m - (s_{pos,i} - s_{neg,i,j}))$$

Where m represents the margin, N the number of positive samples, K_{neg} the number of negative samples, $s_{pos,i}$ the score of the i-th positive sample, and $s_{neg,i,j}$ the score of the j-th negative sample associated with the i-th positive sample.

3.2.3 Enhancement of the Prediction Component

To capture more complex, non-linear interactions between user preferences and news candidate embeddings, we replaced the prediction component (i.e., dot product and softmax) of the GLORY model with a more sophisticated neural network architecture. The new prediction component utilizes a neural network with multiple hidden layers, allowing for a deeper and more nuanced understanding of user preferences.

4 Evaluations

4.1 Datasets

For evaluating our method's enhancements, we used the MIND dataset [27], sourced from anonymized user interactions on the Microsoft News [2] website. This dataset comes in two versions: MIND-large, with data from one million users over six weeks, and MIND-small, a subset including 50,000 users. These datasets provide a robust basis for testing and comparing the performance of our improved recommendation system.

4.2 Experimental Setup

All experiments were carried out on a computing environment equipped with an RTX 6000 GPU and 60GB RAM, ensuring adequate computational resources. Following the GLORY [1] work, we adopted the same parameter settings for a fair comparison. In addition, regarding our proposed enhancement: (1) We implemented the pairwise ranking loss, setting the margin to 1 (m=1). We will evaluate the impact of integrating pairwise ranking loss both as a standalone feature and in conjunction with the NCE loss (i.e., $L_{\text{enhanced}} = L_{\text{NCE}} + L_{\text{PairwiseRanking}}$), to determine its efficacy in enhancing the ranking process. (2) Additionally, we implemented a DNN architecture for the prediction component, which processes concatenated user and candidate embeddings. The architecture is designed with a series of hidden layers. Specifically, the network starts with an input layer of size equal to twice the embedding dimension, reflecting the concatenated embeddings (user and candidates). It then progresses through hidden layers with the following neuron counts: 512, 256, 128, 64, 32, 16, and 8, each followed by a ReLU activation function to introduce non-linearity and facilitate effective learning. The final layer, a single neuron output, computes a relevance score for each candidate, quantifying the predicted engagement probability between a user and a news article.

4.3 Experimental Plan

To systematically evaluate the efficacy of our proposed improvements on the GLORY model, we adopted an ablation study methodology. This enabled us to meticulously analyze the performance

[2] `https://news.microsoft.com/`

impact of each individual modification by comparing it against the baseline performance of the original GLORY method. Our experimental setup consisted of several configurations, each designed to test different combinations of loss functions and prediction components:

- Combination of NCE loss with Pairwise Ranking loss, utilizing the original prediction mechanism, denoted as *NCE + Pairwise & Orig. Pred.*
- Integration of NCE loss with Pairwise Ranking loss, alongside our advanced DNN-based prediction component, denoted as *NCE + Pairwise & DNN.*
- Application of Pairwise Ranking loss with the original prediction framework, denoted as *Pairwise & Orig. Pred.*
- Implementation of Pairwise Ranking loss with our newly developed DNN-based prediction component, denoted as *Pairwise & DNN.*
- Employment of NCE loss with our DNN-based prediction component, exclusively, denoted as *NCE & DNN.*

This structured approach not only facilitated a granular understanding of how each proposed improvement impacts the GLORY model's overall performance but also allowed for the direct attribution of performance changes to specific modifications. Through this meticulous examination, we aimed to isolate the contributions of our improvements to the model's effectiveness, thereby providing clear insights into their value.

4.4 Evaluation Metrics

Our evaluation employs the GLORY [1] method's metrics, facilitating a direct comparison with the original GLORY results. The metrics used are as follows:

Area under the ROC Curve (AUC): AUC measures the area under the Receiver Operating Characteristic curve, which plots the true positive rate against the false positive rate at various threshold settings. It provides an aggregate measure of performance across all possible classification thresholds. A higher AUC value indicates better model performance. The AUC is defined as:

$$\text{AUC} = \int_0^1 TPR(d)\, dFPR$$

where TPR is the true positive rate and FPR is the false positive rate as functions of the decision threshold d.

Mean Reciprocal Rank (MRR): MRR is a statistic measure for evaluating any process that produces a list of possible responses to a sample of queries, ordered by probability of correctness. The reciprocal rank of a query response is the multiplicative inverse of the rank of the first correct answer. The MRR is the average of the reciprocal ranks of results for a sample of queries Q:

$$\text{MRR} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\text{rank}_i}$$

where rank_i is the rank of the first relevant item for the i-th query.

Normalized Discounted Cumulative Gain (nDCG@N): nDCG is a measure of ranking quality. It is used to evaluate the model based on the order of the recommended items and takes into account the position of the relevant items within the recommendation list. It is particularly useful when the relevance scores are not binary. For a set of N recommendations, it is defined as:

$$\text{nDCG@N} = \frac{DCG@N}{IDCG@N}$$

where $DCG@N$ (Discounted Cumulative Gain) is:

$$DCG@N = \sum_{i=1}^{N} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$

and $IDCG@N$ is the ideal $DCG@N$, computed by placing the most relevant items in the top positions. rel_i is the relevance score of item at position i. This was examined using N=5 and 10.

These metrics collectively provide a comprehensive view of the model's performance in terms of both ranking accuracy and the ability to prioritize highly relevant items.

5 Results

This section presents the outcomes of our experimental evaluation, aimed at assessing the impact of the proposed enhancements on the performance of the GLORY method within the context of the MIND datasets. The experiments were meticulously designed to ensure a fair comparison, strictly adhering to the parameter settings and computational environment as delineated in Section 4. It is noteworthy that the results of the original GLORY method reported in this paper slightly deviate from the original results published in the GLORY [1] paper. This variation is attributed to differences in the execution environments, underscoring the importance of uniform testing conditions to ensure comparability. The results, detailed in Table 1, were derived from a consistent environment, facilitating a reliable assessment of the enhancements' efficacy.

Table 1: Performance of the Original GLORY [1] and Our Proposed Enhancements Evaluated on MIND Datasets [27]

Method/Improvements	MIND-small				MIND-large			
	AUC	**MRR**	**nDCG@5**	**nDCG@10**	**AUC**	**MRR**	**nDCG@5**	**nDCG@10**
GLORY (Baseline)	**67.96**	32.75	**36.18**	42.48	69.02	33.80	37.44	43.81
NCE + Pairwise & Orig. Pred.	67.95	**32.78**	36.16	**42.52**	68.87	33.66	37.38	43.66
NCE + Pairwise & DNN	67.41	32.03	35.44	41.72	68.97	33.24	36.93	43.35
Pairwise & Orig. Pred.	67.92	32.43	35.94	42.18	**69.15**	**33.83**	**37.55**	**43.86**
Pairwise & DNN	67.04	31.77	35.10	41.48	68.83	33.24	36.63	43.33
NCE & DNN	66.80	31.48	34.96	41.24	68.83	33.24	36.63	43.33

Table 1 presents the results of the original GLORY [1] method and our proposed enhancements, which were evaluated using the AUC, MRR, nDCG@5, and nDCG@10 metrics on both MIND-small and MIND-large datasets [27]. As illustrated in Table 1, the GLORY method serves as our baseline. When examining the enhancements, particularly the integration of the NCE loss with Pairwise Ranking loss utilizing the original prediction mechanism (*NCE + Pairwise & Orig. Pred.*), we observe a marginal improvement in MRR and nDCG@10 on the MIND-small dataset, suggesting a slight enhancement in ranking accuracy.

Conversely, the application of the same combination but with the DNN-based prediction component (*NCE + Pairwise & DNN*) demonstrated a decrease in performance metrics across both datasets. This outcome suggests that while the DNN-based component introduces a more complex model architecture, it does not unequivocally translate to better performance, and may require additional refinement and adjustment to fully leverage its capabilities within the context of news recommendation systems.

Interestingly, the utilization of Pairwise Ranking loss with the original prediction framework (*Pairwise & Orig. Pred.*) resulted in the highest performance on the MIND-large dataset across all metrics, underscoring the potential efficacy of direct ranking optimizations in large-scale environments. However, the integration of Pairwise Ranking loss with the DNN-based prediction component (*Pairwise & DNN*) did not yield similar gains, indicating that the sophisticated model architecture might require additional refinements to fully leverage the benefits of Pairwise Ranking loss in this context.

Furthermore, the exclusive application of the NCE loss with the DNN-based prediction component (*NCE & DNN*) did not result in significant performance improvements, reiterating the necessity for a nuanced approach to integrating advanced machine learning techniques within the GLORY framework.

In summary, our experiments reveal the complexities of enhancing the GLORY framework. Some adjustments, like replacing the NCE to Pairwise Ranking loss with the original predictor, show an increase in the performance. Yet, incorporating a DNN-based predictor proves there's a fine line between complexity and effectiveness.

6 Discussion

The experimental findings from integrating various enhancements into the GLORY framework offer insightful perspectives on the nuanced interplay between model complexity, architectural adjustments, and performance outcomes in the domain of news recommendation systems. This discussion elaborates on the implications of these findings and potential limitations.

The modest performance gains observed from the integration of NCE with Pairwise Ranking loss using the original prediction mechanism on the MIND-small dataset underline the potential of combining loss functions to improve ranking accuracy. This suggests that even small adjustments within the existing framework can lead to improvements, which reinforces the importance of loss function selection in the model's ability to discern and prioritize relevant news articles.

In addition, the superior performance of Pairwise Ranking loss with the original prediction framework, particularly on the MIND-large dataset, showcases the effectiveness of direct ranking optimizations. This reinforces the notion that for large-scale datasets, focusing on optimizing the ranking process itself can be more beneficial than integrating more complex predictive components. The discrepancy in performance gains between the MIND-small and MIND-large datasets also highlights the scalability of certain enhancements and the variable impact of dataset size on the efficacy of applied methods.

Conversely, the application of the DNN-based prediction component across both datasets did not yield improvements over the baseline, questioning the assumption that more complex models automatically lead to better outcomes. Theoretically, these components are designed to capture intricate user-item interactions, but their effectiveness largely depends on the specifics of the architecture and parameter settings. This highlights an important consideration in deploying deep learning for recommendation systems: the balance between model complexity and achieving tangible performance enhancements requires careful tuning and validation to ensure a positive contribution to the system's overall performance.

The lack of improvement with DNN-based enhancements may point to a trade-off between model complexity and practical utility. As models grow more complex, they have the potential to identify deeper patterns but may also become more prone to overfitting and influenced by noise within the data. Moreover, the increased complexity can make these models harder to optimize and adjust, contributing to their failure to enhance performance beyond the established baseline. This necessitates a strategic approach in model development and highlights the need for meticulous model evaluation to truly leverage the capabilities of DNN-based components in recommendation systems.

7 Conclusions and Future Work

This study has proposed and evaluated several enhancements to the GLORY [1] method for personalized news recommendations, focusing on refining the loss function and optimizing the prediction mechanism. Our meticulous evaluation, leveraging the comprehensive MIND [27] datasets, has shed light on the nuanced impact of these enhancements, revealing that strategic modifications, such as the integration of Pairwise Ranking loss with the original prediction framework, can indeed enhance model performance. However, our exploration of DNN-based prediction components illustrated that complexity alone does not guarantee success, underscoring the intricate balance between model sophistication and practical efficacy in the dynamic landscape of personalized news recommendations.

Future work may include exploring alternative DNN architectures and integrating additional types of data to enrich the model's understanding of user preferences, such as incorporating multimodal data.

References

[1] B. Yang, D. Liu, T. Suzumura, R. Dong, and I. Li, "Going beyond local: Global graph-enhanced personalized news recommendations," in *Proceedings of the 17th ACM Conference on Recommender Systems*, 2023, pp. 24–34.

[2] M. An, F. Wu, C. Wu, K. Zhang, Z. Liu, and X. Xie, "Neural news recommendation with long-and short-term user representations," in *Proceedings of the 57th annual meeting of the association for computational linguistics*, 2019, pp. 336–345.

[3] J. Li, J. Zhu, Q. Bi, G. Cai, L. Shang, Z. Dong, X. Jiang, and Q. Liu, "Miner: Multi-interest matching network for news recommendation," in *Findings of the Association for Computational Linguistics: ACL 2022*, 2022, pp. 343–352.

[4] T. Qi, F. Wu, C. Wu, P. Yang, Y. Yu, X. Xie, and Y. Huang, "Hierec: Hierarchical user interest modeling for personalized news recommendation," *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 5446–5456, 2021.

[5] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Neural news recommendation with attentive multi-view learning," *arXiv preprint arXiv:1907.05576*, 2019.

[6] ——, "Npa: neural news recommendation with personalized attention," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, pp. 2576–2584.

[7] C. Wu, F. Wu, S. Ge, T. Qi, Y. Huang, and X. Xie, "Neural news recommendation with multi-head self-attention," in *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, 2019, pp. 6389–6394.

[8] S. Sidana, M. Trofimov, O. Horodnitskii, C. Laclau, Y. Maximov, and M.-R. Amini, "Representation learning and pairwise ranking for implicit feedback in recommendation systems," *arXiv preprint arXiv:1705.00105*, 2017.

[9] C. Feng, M. Khan, A. U. Rahman, and A. Ahmad, "News recommendation systems - accomplishments, challenges & future directions," *Journal Name*, vol. Volume Number, no. Issue Number, p. Page Range, Publication Year.

[10] S. Rendle, "Factorization machines with libfm," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 3, no. 3, pp. 1–22, 2012.

[11] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, "Deepfm: a factorization-machine based neural network for ctr prediction," *arXiv preprint arXiv:1703.04247*, 2017.

[12] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Neural news recommendation with attentive multi-view learning," *arXiv preprint arXiv:1907.05576*, 2019.

[13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[15] C. Wu, F. Wu, S. Ge, T. Qi, Y. Huang, and X. Xie, "Neural news recommendation with multi-head self-attention," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 6389–6394, 2019.

[16] I. Li, L. Song, K. Xu, and D. Yu, "Variational graph autoencoding as cheap supervision for amr coreference resolution," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2022, pp. 2790–2800.

[17] I. Li, V. Yan, T. Li, R. Qu, and D. Radev, "Unsupervised cross-domain prerequisite chain learning using variational graph autoencoders," *arXiv preprint arXiv:2105.03505*, 2021.

[18] L. Wu, Y. Chen, K. Shen, X. Guo, H. Gao, S. Li, J. Pei, B. Long *et al.*, "Graph neural networks for natural language processing: A survey," *Foundations and Trends® in Machine Learning*, vol. 16, no. 2, pp. 119–328, 2023.

[19] S. Ge, C. Wu, F. Wu, T. Qi, and Y. Huang, "Graph enhanced representation learning for news recommendation," pp. 2863–2869, 2020.

[20] L. Hu, C. Li, C. Shi, C. Yang, and C. Shao, "Graph neural news recommendation with long-term and short-term interest modeling," *Information Processing & Management*, vol. 57, no. 2, p. 102142, 2020.

[21] C. Wu, F. Wu, Y. Huang, and X. Xie, "User-as-graph: User modeling with heterogeneous graph pooling for news recommendation," in *IJCAI*, 2021, pp. 1624–1630.

[22] Z. Mao, J. Li, H. Wang, X. Zeng, and K.-F. Wong, "Digat: Modeling news recommendation with dual-graph interaction," *arXiv preprint arXiv:2210.05196*, 2022.

[23] S. Wang, L. Hu, Y. Wang, X. He, Q. Z. Sheng, M. A. Orgun, L. Cao, and F. Ricci, "Graph learning based recommender systems: A review," *arXiv preprint arXiv:2105.06339*, 2021.

[24] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[25] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in neural information processing systems*, vol. 26, 2013.

[26] D. Vrandečić and M. Krötzsch, "Wikidata: a free collaborative knowledgebase," *Communications of the ACM*, vol. 57, no. 10, pp. 78–85, 2014.

[27] F. Wu, Y. Qiao, J.-H. Chen, C. Wu, T. Qi, J. Lian, D. Liu, X. Xie, J. Gao, W. Wu *et al.*, "Mind: A large-scale dataset for news recommendation," in *Proceedings of the 58th annual meeting of the association for computational linguistics*, 2020, pp. 3597–3606.