Improving the Text Convolution Mechanism with Large Language Model for Review-Based Recommendation

Yoonhyuk Choi
School of Computing and Augmented Intelligence
Arizona State University
Tempe, USA
ychoi139@asu.edu

Fahim Tasneema Azad
School of Computing and Augmented Intelligence
Arizona State University
Tempe, USA
fazad@asu.edu

2024 IEEE International Conference on Big Data (BigData) | 979-8-3503-6248-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/BigData62323.2024.10825014

***Abstract*—Recent studies in recommender systems focus on addressing data sparsity and cold-start problems by utilizing side information, such as tags, images, and testimonials. Among these, user-written testimonials (purchase reviews) are precious for analyzing personal preferences, and many methods have been developed based on this context. Generally, existing methods apply 2D text convolution followed by selecting important words using the attention mechanism. However, the text convolution scheme inevitably suffers from information loss since the number of words in reviews commonly exceeds hundreds. To address this limitation, we focus on the Large Language Model (LLM), which has shown promising results in various fields, including search engines, natural language processing, and healthcare. In particular, LLM has demonstrated excellent performance in text summarization and QA tasks, leading to the development of text-based recommender systems. Nevertheless, LLM alone struggles to perform collaborative filtering, which is essential in a recommender system. Thus, we propose LLM-based text summarization before applying 2D convolution, followed by the widely used collaborative filtering mechanism. This approach can improve recommendation quality by removing unnecessary words in advance, reducing the smoothing effect while capturing the rich user-item interactions. Our method is integrated with recent text-based recommendation algorithms, which have proven to improve the quality of all baselines by about 16.9 % on average. We conduct experiments and ablation studies using benchmark datasets, demonstrating that our method is scalable and efficient.**

***Index Terms*—Recommender system, large language model, review-based recommendation, convolutional neural network, information theory**

I. Introduction

Recommender systems have become fundamental tools across real-world platforms such as Netflix, Amazon, and Yelp [1], [2]. Although they have achieved remarkable commercial success, the quality of recommendations is quite sensitive to scenarios with scarce data. To address this limitation, recent studies have employed various forms of side information, including social relationships [3] and item images [4]. In particular, the textual data (e.g., user reviews) has become the most widely used information, where several recent studies [5]–[8] have achieved state-of-the-art performance. However, these methods simply utilize the review texts as input for text convolution without pre-processing. This approach presents an issue as it is difficult to distinguish between important and less important words because the weights of the convolution filter are applied equally to all words [9]. This raises the question: Would there be a performance improvement if the important words relevant to the domain were appropriately selected from the reviews and used as input for the convolution layer?

To answer the above question, we focus on information theory to determine which kinds of words are necessary in a sentence. Entropy, which is relevant to information gain (IG), has been a criterion for judging the importance of inputs in a prediction [10]. Based on this idea, prior methods have suggested entropy-based parsing for language models [11], feature selection using IG and divergence for text categorization [12], and maximum entropy models for mobile text classification [13]. A recent study [14] also showed that uncertainty (entropy) is directly related to few-shot text classification. More recently, [15] revealed that words with low information gain (IG) but high occurrence frequency need to be filtered out, highlighting the necessity of text summarization before text convolution [16]. To this end, we perform text summarization based on the Large Language Model (LLM), which has recently gained significant attention. Since LLM serves as a pre-processing step at the very beginning of the model, our method has the advantage of reducing entropy while maintaining collaborative filtering.

Large Language Models (LLMs) have shown prominent results in solving language-based problems [17]–[19]. As is well known, LLMs can solve multiple tasks, including question-answering [20], sequential modeling [21], text completion [22], and text summarization [23]. There are various types of LLMs, such as GPT, BARD, BERT, PaLM, and LLaMA, each differing slightly depending on their learning methods and the amount of data they are trained on [24]. Specifically, we focus on the most widely used ChatGPT [25], [26], which is known for its low cost and high efficiency. However, an important aspect besides the model itself is the appropriate use of prompts [27], [28]. A prompt guides the LLM to perform the target task effectively, and we designed ours to capture

979-8-3503-6248-0/24/$31.00 ©2024 IEEE

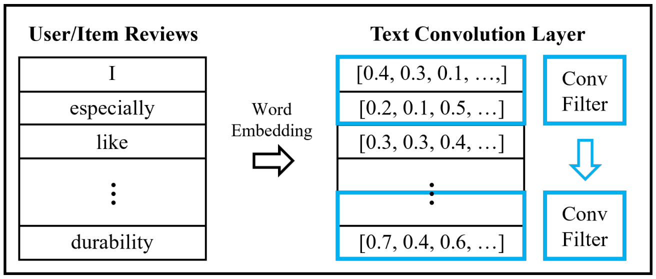

Fig. 1: Mechanism of text convolution layer: Each word is embedded using the pre-trained embedding functions like word2vec [31] and GloVe [32]. Then, the convolution filter slides down and smooths the adjacent words

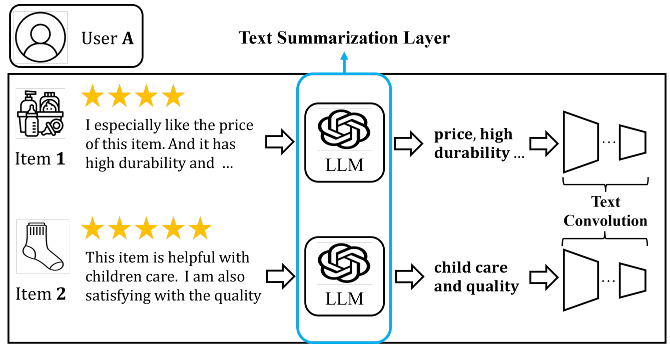

Fig. 2: Mechanism of incorporating LLM with text convolution. For each review, we apply text summarization using LLM, followed by the text convolution layer

words suitable for the recommender system [29], [30].

From this perspective, we analyze whether the LLM can mitigate the information loss inherent in the text convolution algorithm. Our contributions can be summarized as follows:

- We propose LLM-based text summarization before the text convolution layer. Our method is scalable, as LLMs can be integrated with any text-processing algorithms. Additionally, it is robust in addressing collaborative filtering problems that are challenging when using only LLMs.
- Based on information theory, we theoretically prove that LLMs can mitigate the information loss of the plain text processing mechanism.
- We conduct extensive experiments using state-of-the-art review-based methods. The results demonstrate that LLMs significantly improve recommendation quality.

II. Method

We start with a brief introduction to the text convolution mechanism. As illustrated in Figure 1, the text convolution layer consists of the following two steps: (1) For each user or item, we concatenate their reviews and apply pre-trained word embedding functions (word2vec[1] [31] or GloVe[2] [32]). Then, (2) the two-dimensional convolution filters slide down and perform matrix multiplication for each window. It is important to note that the parameters of the convolution filter are applied equally to all words, which may result in smoothing both important and less important words.

To address this issue, we suggest applying the LLM (e.g., ChatGPT or LLaMA) to capture useful words during preprocessing. In Figure 2, we describe the overall mechanism of incorporating the LLM before the text convolution layer. For example, assume that *user A* has evaluated two *items (1 and 2)* after purchasing them. Each review may contain less useful words: *"I, especially, this item"* for *item 1*, and *"This item is, I am also satisfying"* for *item 2*, which are unnecessary for modeling this user's interest. To remove such phrases, we set the prompt for the LLM as shown in Table III. Returning to Figure 2, as shown in the middle, the texts can be summarized to contain keyword-like information, mitigating information loss in the subsequent text convolution layer. Finally, the extracted word features can be utilized as the latent representations of users and items for collaborative filtering.

[1]https://code.google.com/archive/p/word2vec
[2]https://nlp.stanford.edu/projects/glove

III. Time Complexity Analysis

We begin with the computational cost of the plain text convolution algorithm, DeepCoNN [5]. For a brief explanation, let us assume the number of parameters in this model is $A + B$, which comprises a single feature extractor (FE) (A) and a prediction layer (B). In addition, we need to consider the cost of the LLM, where we employ GPT-4o for text summarization. In detail, the time complexity of GPT-4o is known to be 320ms per token. Assuming that a review has k tokens on average and there are N_t reviews in a target domain, we can approximate the running time as $\mathcal{O}((A + B) \cdot N_t \cdot k) \times 320ms$. In the following section, we theoretically analyze the effectiveness of our method based on information theory.

IV. Theoretical Analysis

Before delving into the information gain, we introduce Shannon's entropy [10] for the d-sized vector sets below:

$$H(T) = -\sum_{i=1}^{d} T_i log_d T_i. \quad (1)$$

Using the above equation, the information gain $IG(T, a)$ under separation condition (attribute) a is given by:

$$IG(T, a) = H(T) - H(T|a). \quad (2)$$

If the LLM can filter out the subsets of a with low $IG(T, a)$, it is evident that the information gain will increase since $H(T|a)$ decreases [33]. Additionally, please note that low uncertainty $H(\cdot)$ is vital for confident prediction and parameter convergence [34]. Finally, a recent study [35] found that entropy minimization is related to the improvement of text summarization scores. From this observation, we can infer that an LLM with the highest summarization score (e.g., ROUGE-L) can be more helpful when integrated with the text convolution layer [36].

TABLE I: (RQ1) The experimental results are measured using the NDCG@10 and HR@10 scores on four Amazon datasets. Bold methods with underlines indicate that a large language model is applied before text convolution. For each dataset, symbol ¶ demonstrates the method with the highest HR@10 improvement, and bold indicates the best HR@10 score

Recommendation Type	Method	Metric	Gift Card			All Beauty			Amazon Fashion		
			Cloth	CDs	Toys	Cloth	CDs	Toys	Cloth	CDs	Toys
Single-Domain	DeepCoNN [5]	NDCD@10		0.105			0.089			0.099	
		HR@10		0.182			0.175			0.178	
	DeepCoNN	NDCD@10		0.130			0.104			0.125	
		HR@10		0.227¶			0.193			0.212	
	NARRE [37]	NDCD@10		0.116			0.111			0.106	
		HR@10		0.208			0.202			0.194	
	NARRE	NDCD@10		0.132			0.142			0.135	
		HR@10		0.236			0.239¶			0.233	
	AHN [38]	NDCD@10		0.141			0.142			0.102	
		HR@10		0.248			0.254			0.187	
	AHN	NDCD@10		0.155			0.163			0.130	
		HR@10		0.279			0.284			0.226	
Cross-Domain	RC-DFM [39]	NDCD@10	0.146	0.122	0.130	0.135	0.132	0.128	0.098	0.112	0.105
		HR@10	0.269	0.225	0.246	0.260	0.254	0.249	0.181	0.212	0.215
	RC-DFM	NDCD@10	0.168	0.141	0.152	0.154	0.152	0.144	0.131	0.144	0.138
		HR@10	0.302	0.249	0.271	0.292	0.277	0.270	0.230	0.239	0.242
	CATN [40]	NDCD@10	0.152	0.130	0.137	0.140	0.133	0.131	0.097	0.110	0.113
		HR@10	0.291	0.236	0.254	0.258	0.259	0.251	0.188	0.202	0.217
	CATN	NDCD@10	0.171	0.149	0.160	0.155	0.157	0.146	0.123	0.137	0.141
		HR@10	0.311	0.265	0.279	0.291	0.284	0.275	0.230	0.248	0.261¶
	SER [8]	NDCD@10	0.160	0.146	0.159	0.150	0.143	0.146	0.141	0.144	0.138
		HR@10	0.300	0.285	0.299	0.288	0.272	0.279	0.250	0.280	0.276
	SER	NDCD@10	0.188	0.170	0.179	0.165	0.162	0.164	0.170	0.171	0.169
		HR@10	0.324	0.313	0.315	0.296	0.291	0.302	0.292	0.331	0.329
	HEAD [41]	NDCD@10	0.184	0.177	0.180	0.161	0.158	0.157	0.153	0.172	0.165
		HR@10	0.339	0.332	0.340	0.309	0.302	0.305	0.291	0.322	0.318
	HEAD	NDCD@10	0.215	0.206	0.211	0.187	0.185	0.170	0.181	0.200	0.189
		HR@10	**0.358**	**0.343**	**0.370**	**0.322**	**0.321**	**0.314**	**0.341**	**0.363**	**0.369**

TABLE II: Details of the benchmark datasets

Domain	Dataset	# users	# items	# reviews
Source	Clothing (Cloth)	1,219,520	376,858	11,285,464
	CDs and Vinyl (CDs)	112,391	73,713	1,443,755
	Toys and Games (Toys)	208,143	78,772	1,828,971
Target	Gift Card	457	148	2,972
	All Beauty	990	85	5,269
	Amazon Fashion	406	31	3,176

V. Experiments

To demonstrate the superiority of our proposed method, we formulated the following Research Questions (RQs):

- **RQ1:** Does LLM improve the performance of text-based recommendation algorithms?
- **RQ2:** What information is preserved in the summarized reviews? Can only the important information be extracted from the original reviews?
- **RQ3:** Does the LLM-based text summarization reduce entropy compared to traditional reviews?

A. Baselines and Experimental Setup

Baselines. We found that the single-domain recommendation algorithms are quite outdated. Thus, we further employ recently proposed cross-domain techniques as baselines. (1) For single-domain methods, we use DeepCoNN [5], NARRE [37], and AHN [38]. (2) For cross-domain schemes, we use RC-DFM [39], CATN [40], SER [8], and HEAD [41]. Please refer to these articles for more details. The implementation can be found in the anonymous *GitHub*[3]

Experimental setup. As described in Table II, we utilize the *Amazon*[4] 5-core review datasets following [8], [42], where users and items have at least five interactions. We split the dataset into 80%/10%/10% for training, validation, and testing, without considering temporal sequence [43], [44]. Due to the limitation on the size of data that GPT-4 can process at one time, smaller datasets were selected as the target. [45]. We employ an early stopping technique to terminate the training process if the best validation score is not updated for 300

[3]https://github.com/ChoiYoonhyuk/LLM_short

[4]https://cseweb.ucsd.edu/ jmcauley/datasets/amazon_v2/

TABLE III: (RQ2) For target datasets, we describe the original and the summarized texts, including its rating (rat) and vote

Prompt		Summarize the **reviewText** of each row by capturing important words related to the [category]
Gift Card (rat: 1.0, vote: 9)	Original	I WAS SO EMBARRASSED TO HAVE THE CARDS NOT WORK. BE CAREFUL. MAKE SURE YOU CHECK YOUR BALANCE FIRST BEFORE ORDERING OR YOU MAY BE SORRY...
	Summary	**EMBARRASSED** HAVE CARDS WORK. **CAREFUL**. MAKE SURE CHECK YOUR BALANCE...
All Beauty (rat: 5.0, vote: 25)	Original	HEY!! I am an Aqua Velva Man and absolutely love this stuff, been using it for over 50 years. This is a true after shave lotion classic. Not quite sure how many women that have been attracted ...
	Summary	Aqua Velva **absolutely love** this stuff, been using over...
Amazon Fashion (rat: 4.0, vote: 35)	Original	great price for the product, though the sizes tend to be bigger (based on mens size i think). there wasn't a size chart to refer to when i was ordering, so i ended up buying two, each at a difference size
	Summary	**Great price** product, sizes tend bigger. No size chart ordering, bought two different sizes

iterations. Lastly, we report the test accuracy using the best validation score.

B. Performance Analysis (RQ1)

In Table I, we evaluate the models based on two metrics: the Normalized Discounted Cumulative Gain (NDCG@10) and Hit Ratio (HR@10). Methods with underlines indicate that LLM-based summarization is applied for text summarization on a plain algorithm. Firstly, we observe that LLM improves the recommendation quality of all methods. We believe this is because LLM filters out less relevant information, which can better reflect the user preferences inherent in the reviews. Additionally, the performance improvement was more significant for cross-domain algorithms. Due to the relatively large size of the source domain, it can be inferred that the more refined the reviews are, the more significantly the performance of the target domain will improve.

C. Case Study (RQ2)

In Table III, we show illustrative examples of abbreviated texts compared to the original ones. Here, we take three target domains (*Gift Card, All Beauty, and Amazon Fashion*) for analysis. As shown in this table, the prompt is set as *summarize the reviewText of each row by capturing important words related to the [category]*, where we manually set this to reflect the category of a target domain. As we can see, the LLMs well preserves meaningful words that are related to the domain, user preference, and semantic meaning of words while disregarding generic text. This confirms that LLMs can improve text convolution by filtering out less important words.

D. Information Gain Analysis (RQ3)

In Figure 3, we compare the entropy of the original files and the summarized files to measure the performance of the LLM. In each file, we randomly select 100 reviews with more than 20 words. Then, we concatenate them to generate a single document to measure Shannon's entropy [10]. For a fair comparison, we describe the entropy of all datasets in Table II. Firstly, we found that the entropy of the datasets is influenced by the average length of the sampled sentences. In addition, we observe that the summarized texts ($*$) generally have lower entropy compared to the original ones ($/$). This supports our assumption in Eq. 2, where reduced uncertainty can improve recommendation quality.

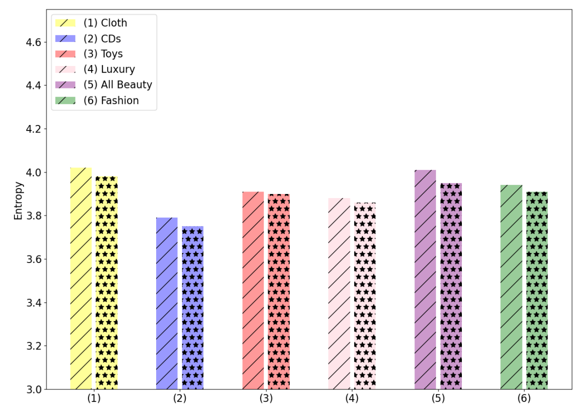

Fig. 3: (RQ3) For each dataset in Table II, we concatenate all reviews and measure the Shannon's entropy [10] before ($/$) and after text summarization ($*$)

VI. CONCLUSION

Recent studies on recommender systems have tackled the issue of data sparsity by incorporating user reviews. However, these methods merely utilize 2D convolution, potentially failing to capture important words by over-smoothing abundant texts. To address this limitation, we propose to remove less important parts from reviews through LLM-based summarization in advance by focusing on the relevant parts of the reviews. Additionally, we provide theoretical proof showing that our technique can alleviate information loss compared to the plain text processing algorithm. Our empirical analysis demonstrates that applying our method is effective for both single and cross-domain text recommendation methods, showing performance improvement across various datasets by reducing the uncertainty of textual information.

REFERENCES

[1] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: a survey," *Decision support systems*, vol. 74, pp. 12–32, 2015.

[2] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM computing surveys (CSUR)*, vol. 52, no. 1, pp. 1–38, 2019.

[3] P. Kazienko, K. Musial, and T. Kajdanowicz, "Multidimensional social network in the social recommender system," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 4, pp. 746–759, 2011.

[4] S. K. Addagarla and A. Amalanathan, "Probabilistic unsupervised machine learning approach for a similar image recommender system for e-commerce," *Symmetry*, vol. 12, no. 11, p. 1783, 2020.
[5] L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for recommendation," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 2017, pp. 425–434.
[6] S. M. Al-Ghuribi and S. A. M. Noah, "Multi-criteria review-based recommender system–the state of the art," *IEEE Access*, vol. 7, pp. 169 446–169 468, 2019.
[7] M. Srifi, A. Oussous, A. Ait Lahcen, and S. Mouline, "Recommender systems based on collaborative filtering using review texts—a survey," *Information*, vol. 11, no. 6, p. 317, 2020.
[8] Y. Choi, J. Choi, T. Ko, H. Byun, and C.-K. Kim, "Based domain disentanglement without duplicate users or contexts for cross-domain recommendation," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 2022, pp. 293–303.
[9] Z. Yang, Z. Hu, R. Salakhutdinov, and T. Berg-Kirkpatrick, "Improved variational autoencoders for text modeling using dilated convolutions," in *International conference on machine learning*. PMLR, 2017, pp. 3881–3890.
[10] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.
[11] A. Ratnaparkhi, "Learning to parse natural language with maximum entropy models," *Machine learning*, vol. 34, pp. 151–175, 1999.
[12] C. Lee and G. G. Lee, "Information gain and divergence-based feature selection for machine learning-based text categorization," *Information processing & management*, vol. 42, no. 1, pp. 155–165, 2006.
[13] C. Yin and J. Xi, "Maximum entropy model for mobile text classification in cloud computing using improved information gain algorithm," *Multimedia Tools and Applications*, vol. 76, pp. 16 875–16 891, 2017.
[14] S. Mukherjee and A. Awadallah, "Uncertainty-aware self-training for few-shot text classification," *Advances in Neural Information Processing Systems*, vol. 33, pp. 21 199–21 212, 2020.
[15] M. Dong, H. Xu, and Q. Xu, "Text classification based on improved information gain algorithm and convolutional neural network," in *Testbeds and Research Infrastructures for the Development of Networks and Communications: 14th EAI International Conference, TridentCom 2019, Changsha, China, December 7-8, 2019, Proceedings 14*. Springer, 2020, pp. 184–198.
[16] T. Goyal, J. J. Li, and G. Durrett, "News summarization and evaluation in the era of gpt-3," *arXiv preprint arXiv:2209.12356*, 2022.
[17] J. Liu, C. Liu, P. Zhou, R. Lv, K. Zhou, and Y. Zhang, "Is chatgpt a good recommender? a preliminary study," *arXiv preprint arXiv:2304.10149*, 2023.
[18] J. Shuai, L. Wu, K. Zhang, P. Sun, R. Hong, and M. Wang, "Topic-enhanced graph neural networks for extraction-based explainable recommendation," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023, pp. 1188–1197.
[19] R. Schaeffer, B. Miranda, and S. Koyejo, "Are emergent abilities of large language models a mirage?" *Advances in Neural Information Processing Systems*, vol. 36, 2024.
[20] Y. Zhuang, Y. Yu, K. Wang, H. Sun, and C. Zhang, "Toolqa: A dataset for llm question answering with external tools," *Advances in Neural Information Processing Systems*, vol. 36, 2024.
[21] J. Lin, R. Shan, C. Zhu, K. Du, B. Chen, S. Quan, R. Tang, Y. Yu, and W. Zhang, "Rella: Retrieval-enhanced large language models for lifelong sequential behavior comprehension in recommendation," in *Proceedings of the ACM on Web Conference 2024*, 2024, pp. 3497–3508.
[22] T. Wu, S. He, J. Liu, S. Sun, K. Liu, Q.-L. Han, and Y. Tang, "A brief overview of chatgpt: The history, status quo and potential future development," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 5, pp. 1122–1136, 2023.
[23] H. Xu, H. Liu, Z. Lv, Q. Yang, and W. Wang, "Sentiment-aware review summarization with personalized multi-task fine-tuning," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 2826–2835.
[24] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang *et al.*, "A survey on evaluation of large language models," *ACM Transactions on Intelligent Systems and Technology*, vol. 15, no. 3, pp. 1–45, 2024.
[25] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier *et al.*, "Chatgpt for good? on opportunities and challenges of large language models for education," *Learning and individual differences*, vol. 103, p. 102274, 2023.
[26] H. Zhang, X. Liu, and J. Zhang, "Summit: Iterative text summarization via chatgpt," *arXiv preprint arXiv:2305.14835*, 2023.
[27] T. Shin, Y. Razeghi, R. L. Logan IV, E. Wallace, and S. Singh, "Autoprompt: Eliciting knowledge from language models with automatically generated prompts," *arXiv preprint arXiv:2010.15980*, 2020.
[28] L. Reynolds and K. McDonell, "Prompt programming for large language models: Beyond the few-shot paradigm," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–7.
[29] P. Liu, L. Zhang, and J. A. Gulla, "Pre-train, prompt, and recommendation: A comprehensive survey of language modeling paradigm adaptations in recommender systems," *Transactions of the Association for Computational Linguistics*, vol. 11, pp. 1553–1571, 2023.
[30] X. Ren, W. Wei, L. Xia, L. Su, S. Cheng, J. Wang, D. Yin, and C. Huang, "Representation learning with large language models for recommendation," in *Proceedings of the ACM on Web Conference 2024*, 2024, pp. 3464–3475.
[31] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in neural information processing systems*, vol. 26, 2013.
[32] 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.
[33] D. Roobaert, G. Karakoulas, and N. V. Chawla, "Information gain, correlation and support vector machines," in *Feature extraction: Foundations and applications*. Springer, 2006, pp. 463–470.
[34] J. Moon, J. Kim, Y. Shin, and S. Hwang, "Confidence-aware learning for deep neural networks," in *international conference on machine learning*. PMLR, 2020, pp. 7034–7044.
[35] A. Khurana and V. Bhatnagar, "Investigating entropy for extractive document summarization," *Expert Systems with Applications*, vol. 187, p. 115820, 2022.
[36] J. Wang, Y. Liang, F. Meng, Z. Sun, H. Shi, Z. Li, J. Xu, J. Qu, and J. Zhou, "Is chatgpt a good nlg evaluator? a preliminary study," *arXiv preprint arXiv:2303.04048*, 2023.
[37] C. Chen, M. Zhang, Y. Liu, and S. Ma, "Neural attentional rating regression with review-level explanations," in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 1583–1592.
[38] X. Dong, J. Ni, W. Cheng, Z. Chen, B. Zong, D. Song, Y. Liu, H. Chen, and G. De Melo, "Asymmetrical hierarchical networks with attentive interactions for interpretable review-based recommendation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 7667–7674.
[39] W. Fu, Z. Peng, S. Wang, Y. Xu, and J. Li, "Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 94–101.
[40] C. Zhao, C. Li, R. Xiao, H. Deng, and A. Sun, "Catn: Cross-domain recommendation for cold-start users via aspect transfer network," *arXiv preprint arXiv:2005.10549*, 2020.
[41] Y. Choi, "Based cross-domain recommendation via hyperbolic embedding and hierarchy-aware domain disentanglement," *arXiv preprint arXiv:2403.20298*, 2024.
[42] C. Zhao, H. Zhao, M. He, J. Zhang, and J. Fan, "Cross-domain recommendation via user interest alignment," in *Proceedings of the ACM Web Conference 2023*, 2023, pp. 887–896.
[43] J. Xu and Y. Cai, "Decoupled hyperbolic graph attention network for cross-domain named entity recognition," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023, pp. 591–600.
[44] S. Wang, S. Guo, L. Wang, T. Liu, and H. Xu, "Hdnr: A hyperbolic-based debiased approach for personalized news recommendation," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023, pp. 259–268.
[45] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, "Gpt-4 technical report," *arXiv preprint arXiv:2303.08774*, 2023.