International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 10 | Oct 2023 www.irjet.net p-ISSN: 2395-0072

Prompt-Based Techniques for Addressing the Initial Data Scarcity in Personalized Recommendations

Aayush Devgan[1]

[1]*Student, Computer Science Department, VIT University, Vellore, Tamil Nadu*

Abstract - *Tapping into the vast reservoir of insights from pre-established language models, my study introduces a novel strategy for recommendation engines. The primary challenge tackled is the initial data scarcity in recommendation scenarios, especially prevalent in nascent enterprises or platforms lacking substantial user activity logs. Instead of relying on past user-item interactions, my approach morphs the recommendation mechanism into a sentiment analysis of languages reflecting user characteristics and item features. Sentiment tendencies are deduced via prompt education. While recommendation tools are pivotal in guiding users towards content resonating with their preferences, formulating tailor-made tools becomes arduous without prior interactions. Earlier research predominantly revolves around initial scenarios for either users or items and these methods, relying on past interactions in the same field, do not address my concern. I also introduce a standard to evaluate my method in the context of initial data scarcity, with outcomes underscoring its potency. To my understanding, this research stands as a pioneering effort in confronting the challenges of recommendations without prior data.*

Key Words: Recommendation System, Deep Learning, Prompt Based

1. INTRODUCTION

Recommender systems play a pivotal role in guiding users towards items that align with their preferences. For e-commerce platforms, having a precise recommendation system is crucial, as it enhances user involvement and drives sales. Classic recommendation approaches, like collaborative filtering [18, 19, 40] and content-oriented techniques [10], primarily harness data from past user-item interactions like clicks and buys to deduce user likings and recommend suitable items. Yet, these methodologies might not be effective where there's a lack of user-item interaction data. This issue is termed the initial-scarcity problem in recommendation systems, a challenge often faced by emerging businesses. Some previous works [22, 28, 31] have delved into the initial-scarcity dilemma, but as shown in Figure 1, they still bank on some extent of historical user-item data either for training or predicting. A basic way to circumvent the initial-scarcity issue might be to set manual guidelines, for instance, suggesting trending or in-season items [7, 31]. However, this might lead to generic recommendations, potentially diminishing the user experience.

Recent advancements in prompt-based learning using pre-existing language models (PLMs) in the field of natural language processing [3, 26, 39] appear to offer potential solutions to the initial-scarcity issue in recommendation systems. The central idea is that if user preferences and item attributes can be represented in natural language found in public datasets, then PLMs might be able to leverage this for making recommendations. Some experts [4, 13, 36, 49, 50] have begun exploring the feasibility of making tailored recommendations using established PLMs. The usual approach involves translating user interaction data into natural language and then reshaping the recommendation task to fit a language modeling structure with carefully curated templates. For instance, to recommend a movie, a template might look like "The viewer has seen movies A, B, and C. They might be interested in movie [D] next", where A to C are the most recent movies viewed by a user, and [D] is a potential movie recommendation [36, 50]. The movie with the highest predicted probability for [D] is then suggested. However, while these methods don't need a model trained on past user-item interactions, they still utilize such data during the prediction phase, which doesn't entirely address the primary initial-scarcity challenge I am focusing on.

In my research, I have put forth a straightforward yet potent method, termed Prompt Recognition, aimed at addressing the intricate issue of system initial-scarcity recommendations. With Prompt Recognition, I utilize both user and item profile characteristics, transform these into structured sentences, and employ PLMs to generate recommendations. To be more specific, I begin by establishing a verbalizer that translates profile attributes into natural language narratives. Next, I use a template to represent the recommendation challenge as a Masked Language Modeling (MLM) task [5], and then employ a PLM to execute this MLM task, leading to the final recommendations. Additionally, I introduce a benchmark for a more quantifiable assessment of my proposed solution. Here's a summary of my key takeaways:

I define the system initial-scarcity recommendation dilemma and present the community with its inaugural benchmark. I unveiled a prompt-based learning method, Prompt Recognition, explicitly designed for system initial-scarcity recommendation challenges. I carry out

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 10 | Oct 2023 www.irjet.net p-ISSN: 2395-0072

experimental analyses across diverse datasets and PLMs to showcase the potency of Prompt Recognition and contemplate the next steps and possibilities in this area.

I chose the click-through rate (CTR) prediction task [33] to set up my recommendation scenario. That is, each record $r_{u,i} \in \{0,1\}$ is a binary value, which $r_{u,i} = 1$ means the user u clicked the item i. The recommendation system f takes a user-item pair (c_u, c_i) as input to predict the probability $\hat{r}_{u,i} \in [0,1]$ that the user will click on the item as output based on the model $\hat{r}_{u,i} = P(r_{u,i} = 1 \mid c_u, c_i; \theta_f)$, where θ_f is the parameters of the recommendation system f. The goal of CTR prediction is to minimize the difference between the predicted probability $\hat{r}_{u,i}$ and the real user-item interaction $r_{u,i}$

In traditional supervised scenarios, a training set containing observed interaction triplets $\{(c_u, c_i, r_{u,i})\}$ between users and items is collected to optimize the model parameter θ_f. However, under the system initial-scarcity setting, I cannot obtain any interaction records, which is a common situation for start-ups or companies that have just launched their new business [32]. Thus, in the system initial-scarcity recommendation, I define a target dataset as $D_{tgt} = (U_{tgt}, I_{tgt}, R_{tgt})$ with an empty interaction matrix $R_{tgt} = \emptyset$. My goal is to recommend items in I_{tgt} to users in U_{tgt} by using their profile features $\{c_u\}$ and $\{c_i\}$. Using recorded interactions is not allowed, no matter in training or inference, but recommender system developers can explore available resources (e.g., account sign-up surveys) to build user and item profiles.

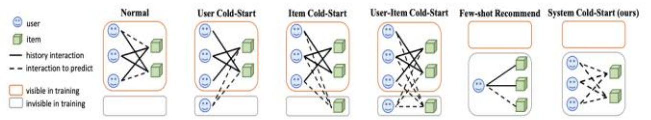

Fig. 1. Illustration of different cold-start recommendation scenarios, including user cold-start [31], item cold-start [28], user-item cold-start [22], few-shot recommendation [4, 50], and ours.

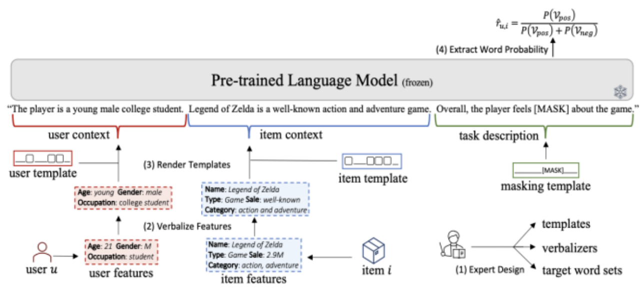

Fig. 2. PromptRec: Prompting PLMs to make personalized recommendations under the system cold-start setting.

2. METHODOLOGY

In the modern landscape of recommendation systems, the concept of "initial data scarcity" presents a daunting challenge. To address this, my study presents Prompt Recognition, a method that harnesses the power of pre-trained language models in a prompt learning framework. Conventional supervised learning methodologies, notably "Fully Supervised Learning" and the "Pre-train and Finetune" paradigm [25], often hit a wall in such situations due to the absence of foundational training data crucial for optimizing model parameters θ_f.

Delving deeper, while the allure of pre-trained language models (PLMs) in few-shot settings has been explored by some research [3,35], it's the innovative "prompt" learning [25] approach that transforms the downstream task to align with a pre-training task of language models. Contemporary recommendation strategies [4,36,50] that employ this prompt-centric paradigm seek to synchronize the recommendation task with masked language modeling. In this context, PLMs serve to gauge the likelihood of an item's descriptor emerging within a user-item narrative. Consider the instance: "A user clicking hiking shoes, will also click trekking poles". Here, the probability associated with "trekking poles" within this narrative reflects the user's inclination toward the item in question.

Yet, a mere prediction based on item descriptors doesn't always hit the mark, more so in initial data scarcity scenarios. The rationale is straightforward: the probability of an item's descriptor is inherently influenced by the vocabulary it comprises. Items named with frequently occurring terms might be deemed more probable, irrespective of their actual relevance to user inclinations or the contextual narrative. To illustrate, in general corpora, "League of Legends" is naturally poised to have a superior likelihood compared to "Legend of Zelda", given the higher frequency of the term "League" over "Zelda".

To address this problem, instead of simply predicting the item names, I propose a new approach called Prompt Recognition which: (1) provides common-sense descriptions to users and items in prompts; (2) predicts the probability of some chosen sentiment words (e.g., "good", "bad") based on the user-item context. Formally, a prompting function

f_{prompt} maps the user-item pair (c_u, c_i) into a L_{ctx}-word context $X_{u,i} \in V^{L_{ctx}}$ with a verbalizer $g_{verb}$and a template filler g_{fill}. Here, V is a predefined vocabulary set, g_{verb}

describes profile features with natural language words, and g_{fill} generates the context combining the user and item verbalized features with a manual template T, having a special word [MASK]. Given a PLM f_{plm} with a d-dimensional embedding space, the preference $\hat{r}_{u,i}$ from user u to item i is estimated by:

$$\hat{r}_{u,i} = \frac{P(V_{pos})}{P(V_{pos}) + P(V_{neg})},$$

where

$P(V) = \prod\_(w \in V) P([MASK] = w \mid E\_(u,i); f\_plm)$.

Here, $E_{u,i} \in R^{L_{ctx}} \times d$ is the word embedding matrix of the context $X_{u,i}$; $V_{pos}, V_{neg} \subset V$

are the predefined positive and negative sentiment vocabulary sets, respectively; $P\big([MASK] = w \mid E_{u,i}; f_{plm}\big)$ is the predicted probability that word w appears at the position [MASK] within the context $X_{u,i}$ according to the model f_{plm}.Figure 2 shows the overall framework of Prompt Recognition. Taking game recommendation as an example, I first describe the user u 's and item i 's profile features with natural language as $X_u = g_{verb}(c_u)$

and $X_i = g_{verb}(c_i)$ respectively, where g_{verb} is a human-designed function verbalizing a feature vector with words. Meanwhile, I design template T as "The player is a

age gender occupation. The name is categorized as a category video game created by a producer. Overall, the player feels [MASK] about the game.", where each underlined word is a slot. Then, the context $X_{u,i} = g_{fill}(X_u, X_i; T)^1$ is generated by filler g_{fill}, which fills the slots with the corresponding verbalized features. If let $V_{pos} = \{$ "good" $\}$ and $V_{neg} = \{$ "bad" $\}$, the predicted preference $r\hat{}_{u,i}$ is computed by normalizing the probabilities of observing "good" and "bad" at position [MASK].

3. EXPERIMENTS

Outlined in this experiment is the Initial Data Scarcity Recommendation Benchmark. This benchmark amalgamates three publicly accessible datasets, along with a dataset segmentation technique, crafted to mirror real-world initial-scarcity challenges. Additionally, it encompasses reference strategies, spanning both rule-driven and PLM-driven approaches. Notably, GAUC serves as the evaluation metric for all these methodologies.

Data Collections: The primary challenge posed by the initial-scarcity recommendation paradigm stems from the absence of past user-item interactions during the learning phase. To navigate this, systems lean heavily on profile attributes to tailor recommendations. From my research, three datasets aligned with these conditions were identified: the: In-Vehicle Coupon Recommendation (Coupon) [42], Mexico Restaurant Recommendation (Restaurant) [37], and MovieLens-100K (ML-100K) [16]. While the Coupon set aims to evaluate the accuracy of discount offerings to drivers, the Restaurant set focuses on gauging system efficiency in aligning with diner preferences. Concurrently, the ML-100K dataset delves into the efficacy of film suggestions for viewers. A concise overview of these datasets can be found in Table 1.

Table 1. Statistics of datasets.

Dataset	#User/#Feature	#Item/#Feature	#Interaction	Density
ML-100K	943/4	1682/22	100,000	3.15%
Coupon	8312/12	6924/13	12,684	0.01%
Restaurant	138/20	939/25	1,161	0.45%

Data Segmentation: My approach involves partitioning each dataset into training, validation, and testing subsets. The training portion contains 250 instances, the validation segment comprises 50, and the remainder forms the testing set. It's worth noting that the limited size of the training set facilitates the assessment of recommendation capabilities in a few-shot context, where minimal interaction data aids model refinement. To ward off hyper-parameter over-optimization, a modestly sized validation set is incorporated. The ultimate system evaluations hinge on performance metrics derived from the test subset. Detailed exploration of the few-shot scenario remains a subject for subsequent studies.

Data Transformation: This study conceptualizes the recommendation task as the Click-Through Rate (CTR) prediction challenge. Given that the foundational labels of my datasets reflect user-item preference intensities, a transformation process converts these into binary outcomes {0,1}, steered by specific thresholds [29, 51]. To elaborate, the chosen thresholds for ML-100K, Coupon, and Restaurant datasets stand at 4.0, 1.0, and 2.0, respectively.

Evaluation Metrics: Although the ROC-AUC score is a standard metric for binary classification tasks like CTR prediction, its application in personalized recommendations can be problematic. The inherent methodology of AUC, which involves evaluating all user-item interactions, might inadvertently introduce biases stemming from user-to-user variations [52]. To counteract this, the Group-AUC (GAUC) metric [17, 52] has been proposed. This method computes the AUC individually for each user and subsequently aggregates the results through a weighted average mechanism:

$$GAUC = \sum\_(u \in U)(\# 〖 history 〗\_u \times AUC(u))/(\sum\_(u \in U) \# \, history \, y\_u),$$

where history y_u is the number of records for user u, and $AUC(u)$ is the AUC over all interaction's records for user u.

Comparison Approaches: My analysis integrates two primary benchmark methodologies. The initial type encompasses strategies governed by manually curated rules. For instance, the Random approach entails an indiscriminate item recommendation to users. On the other hand, the subsequent category leans on unsupervised techniques grounded in Pre-trained

Language Models (PLMs). These methodologies leverage verbalized attributes of users and items, drawing insights directly from PLM outputs without the necessity for fine-tuning. A case in point is EmbOLism, which crafts embeddings for both user and item verbalized attributes, and then discerns user-item affinities by computing the dot product of these embeddings. The PairNSP model harnesses the next sentence prediction paradigm [5], fusing verbalized user and item features to evaluate their contextual compatibility within a PLM. Conversely, ItemLM [4] melds these features and gauges preferences based on the predicted likelihood of an item descriptor's presence in the given context. For a detailed overview, readers can refer to Table 2, which focuses on prompt learning in the realm of initial-scarcity recommendations.

Table 2. Prompt learning for cold-start recommendation.

Strategy	PLM	ML-100K	Coupon	Restaurant	avg.
Baselines					
Random	-	$50.02_{\pm 0.19}$	$49.76_{\pm 0.79}$	$50.44_{\pm 1.84}$	50.07
EmbSim	BERT	$50.07_{\pm 0.01}$	$51.40_{\pm 0.16^*}$	$46.46_{\pm 1.87}$	49.31
PairNSP	BERT	$49.71_{\pm 0.01}$	$50.06_{\pm 0.02}$	$52.67_{\pm 1.42^*}$	50.81
ItemLM	BERT	$50.10_{\pm 0.01}$	$30.97_{\pm 0.28}$	$49.08_{\pm 1.44}$	43.38
Ours					
	BERT	$\mathbf{52.53}_{\pm 0.01^*}$	$\mathbf{63.77}_{\pm 0.11^*}$	$\underline{54.79}_{\pm 1.40^*}$	57.03
PromptRec	GPT-Neo	$\underline{52.17}_{\pm 0.01^*}$	$\underline{58.88}_{\pm 0.17^*}$	$53.28_{\pm 1.89^*}$	54.78
	T5	$50.51_{\pm 0.01^*}$	$58.49_{\pm 0.13^*}$	$\mathbf{55.93}_{\pm 1.39^*}$	54.98

4. CONCLUSIONS

Delving into the realm of system initial data scarcity recommendation challenges, this manuscript offers both a comprehensive definition and a pioneering benchmark for the field. In addressing this issue, I've introduced a novel method, termed "Prompt Recognition", which leverages the strengths of PLMs to craft tailored recommendations without the need for supervised fine-tuning. I am optimistic about the potential of further academic endeavors in the initial-scarcity recommendation sphere, envisioning their applications in real-world business scenarios. Several avenues beckon deeper exploration in the future. These include refining the design of templates to maximize the efficacy of PLMs, discerning the optimal pre-trained language model from a plethora of choices, understanding the equilibrium between model dimensions and the performance of prompt learning, and probing into possible bias or privacy concerns inherent in PLM-driven recommendation systems, along with strategies to address them.

REFERENCES

[1] Sid Black, Leo Gao, Phil Wang, Connor Leahy, and Stella Biderman. 2021. GPT-Neo: Large Scale Autoregressive Language Modeling with Mesh-Tensorflow. https://doi.org/10.5281/zenodo.5297715 If you use this software, please cite it using these metadata.

[2] JesúS Bobadilla, Fernando Ortega, Antonio Hernando, and Jesús Bernal. 2012. A collaborative filtering approach to mitigate the new user cold start problem. Knowledge-based systems 26 (2012), 225–238.

[3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. Advances in neural information processing systems 33 (2020), 1877–1901.

[4] Zeyu Cui, Jianxin Ma, Chang Zhou, Jingren Zhou, and Hongxia Yang. 2022. M6-Rec: Generative Pretrained Language Models are Open-Ended Recommender Systems. arXiv preprint arXiv:2205.08084 (2022).

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018).

[6] Hao Ding, Yifei Ma, Anoop Deoras, Yuyang Wang, and Hao Wang. 2021. Zero-shot recommender systems. arXiv preprint arXiv:2105.08318 (2021).

[7] Sun Dong-ting, He Tao, and Zhang Fu-hai. 2012. Survey of cold-start problem in collaborative filtering recommender system. Computer and modernization 1, 201 (2012), 59.

[8] Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In International conference on machine learning. PMLR, 1126–1135.

[9] Yunfan Gao, Tao Sheng, Youlin Xiang, Yun Xiong, Haofen Wang, and Jiawei Zhang. 2023. Chat-REC: Towards Interactive and Explainable LLMs-Augmented Recommender System. arXiv preprint arXiv:2303.14524 (2023).

[10] Ruth Garcia and Xavier Amatriain. 2010. Weighted Content Based Methods for Recommending Connections in Online Social Networks. Recommender Systems and the Social Web (2010), 68.

[11] Shijie Geng, Shuchang Liu, Zuohui Fu, Yingqiang Ge, and Yongfeng Zhang. 2022. Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). arXiv preprint arXiv:2203.13366 (2022).

[12] Zhabiz Gharibshah, Xingquan Zhu, Arthur Hainline, and Michael Conway. 2020. Deep learning for user interest and response prediction in online display advertising. Data Science and Engineering 5, 1 (2020), 12–26.

[13] Lei Guo, Chunxiao Wang, Xinhua Wang, Lei Zhu, and Hongzhi Yin. 2023. Automated Prompting for Non-overlapping Cross-domain Sequential

Recommendation. arXiv preprint arXiv:2304.04218 (2023).

[14] James A Hanley and Barbara J McNeil. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 1 (1982), 29–36.

[15] Bowen Hao, Jing Zhang, Hongzhi Yin, Cuiping Li, and Hong Chen. 2021. Pre-training graph neural networks for cold-start users and items representation. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining. 265–273.

[16] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. Acm transactions on interactive intelligent systems (tiis) 5, 4 (2015), 1–19.

[17] Ruining He and Julian McAuley. 2016. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In proceedings of the 25th international conference on the world wide web. 507–517.

[18] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval. 639–648.

[19] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In Proceedings of the 26th international conference on world wide web. 173–182.

[20] Xiaoqi Jiao, Yichun Yin, Lifeng Shang, Xin Jiang, Xiao Chen, Linlin Li, Fang Wang, and Qun Liu. 2019. Tinybert: Distilling bert for natural language understanding. arXiv preprint arXiv:1909.10351 (2019).

[21] Betty R Kirkwood and Jonathan AC Sterne. 2010. Essential medical statistics. John Wiley & Sons.

[22] Hoyeop Lee, Jinbae Im, Seongwon Jang, Hyunsouk Cho, and Sehee Chung. 2019. Melu: Meta-learned user preference estimator for cold-start recommendation. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 1073–1082.

[23] Lei Li, Yongfeng Zhang, and Li Chen. 2022. Personalized prompt learning for explainable recommendation. arXiv preprint arXiv:2202.07371 (2022).

[24] Blerina Lika, Kostas Kolomvatsos, and Stathes Hadjiefthymiades. 2014. Facing the cold start problem in recommender systems. Expert systems with applications 41, 4 (2014), 2065–2073.

[25] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. 2023. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. Comput. Surveys 55, 9 (2023), 1–35.

[26] Xiao Liu, Yanan Zheng, Zhengxiao Du, Ming Ding, Yujie Qian, Zhilin Yang, and Jie Tang. 2021. GPT understands, too. arXiv preprint arXiv:2103.10385 (2021).

[27] Marwa Hussien Mohamed, Mohamed Helmy Khafagy, and Mohamed Hasan Ibrahim. 2019. Recommender systems challenges and solutions survey. In 2019 International Conference on Innovative Trends in Computer Engineering (ITCE). IEEE, 149–155.

[28] Wentao Ouyang, Xiuwu Zhang, Shukui Ren, Li Li, Kun Zhang, Jinmei Luo, Zhaojie Liu, and Yanlong Du. 2021. Learning graph meta embeddings for cold-start ads in click-through rate prediction. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in 9 Conference acronym ’XX, June 03–05, 2018, Woodstock, NY Xuansheng Wu et al. Information Retrieval. 1157–1166.

[29] Yanru Qu, Ting Bai, Weinan Zhang, Jianyun Nie, and Jian Tang. 2019. An end-to-end neighborhood-based interaction model for knowledge-enhanced recommendation. In Proceedings of the 1st international workshop on deep learning practice for high-dimensional sparse data. 1–9.

[30] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. The Journal of Machine Learning Research 21, 1 (2020), 5485–5551.

[31] Al Mamunur Rashid, Istvan Albert, Dan Cosley, Shyong K Lam, Sean M McNee, Joseph A Konstan, and John Riedl. 2002. Getting to know you: learning new user preferences in recommender systems. In Proceedings of the 7th international conference on Intelligent user interfaces. 127–134.

[32] Al Mamunur Rashid, George Karypis, and John Riedl. 2008. Learning preferences of new users in recommender systems: an information theoretic approach. Acm Sigkdd Explorations Newsletter 10, 2 (2008), 90–100.

[33] Matthew Richardson, Ewa Dominowska, and Robert Ragno. 2007. Predicting clicks: estimating the click-through rate for new ads. In Proceedings of the 16th international conference on the World Wide Web. 521–530.

[34] Rachna Sethi and Monica Mehrotra. 2021. Cold start in recommender systems—a survey from a domain perspective. In Intelligent Data Communication

Technologies and Internet of Things. Springer, 223–232.

[35] Taylor Shin, Yasaman Razeghi, Robert L Logan IV, Eric Wallace, and Sameer Singh. 2020. Autoprompt: Eliciting knowledge from language models with automatically generated prompts. arXiv preprint arXiv:2010.15980 (2020).

[36] Damien Sileo, Wout Vossen, and Robbe Raymaekers. 2022. Zero-Shot Recommendation as Language Modeling. In Advances in Information Retrieval: 44th European Conference on IR Research, ECIR 2022, Stavanger, Norway, April 10–14, 2022, Proceedings, Part II. Springer, 223–230.

[37] Blanca Vargas-Govea, Gabriel González-Serna, and Rafael Ponce-Medellın. 2011. Effects of relevant contextual features in the performance of a restaurant recommender system. ACM RecSys 11, 592 (2011), 56.

[38] Manasi Vartak, Arvind Thiagarajan, Conrado Miranda, Jeshua Bratman, and Hugo Larochelle. 2017. A meta-learning perspective on cold-start recommendations for items. Advances in neural information processing systems 30 (2017).

[39] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. Advances in neural information processing systems 30 (2017).

[40] Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2018. Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In Proceedings of the 27th ACM international conference on information and knowledge management. 417–426.

[41] Hongwei Wang, Fuzheng Zhang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2019. Multi-task feature learning for knowledge graph enhanced recommendation. In The world wide web conference. 2000–2010.

[42] Tong Wang, Cynthia Rudin, Finale Doshi-Velez, Yimin Liu, Erica Klampfl, and Perry MacNeille. 2017. A bayesian framework for learning rule sets for interpretable classification. The Journal of Machine Learning Research 18, 1 (2017), 2357–2393.

[43] Wenjie Wang, Xinyu Lin, Fuli Feng, Xiangnan He, and Tat-Seng Chua. 2023. Generative Recommendation: Towards Next-generation Recommender Paradigm. arXiv preprint arXiv:2304.03516 (2023).

[44] Xiaolei Wang, Kun Zhou, Ji-Rong Wen, and Wayne Xin Zhao. 2022. Towards Unified Conversational Recommender Systems via Knowledge-Enhanced Prompt Learning. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 1929–1937.

[45] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. 2022. Emergent abilities of large language models. arXiv preprint arXiv:2206.07682 (2022).

[46] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface's transformers: State-of-the-art natural language processing. arXiv preprint arXiv:1910.03771 (2019).

[47] Hongzhi Yin, Weiqing Wang, Hao Wang, Ling Chen, and Xiaofang Zhou. 2017. Spatial-aware hierarchical collaborative deep learning for POI recommendation. IEEE Transactions on Knowledge and Data Engineering 29, 11 (2017), 2537–2551.

[48] Junliang Yu, Hongzhi Yin, Xin Xia, Tong Chen, Jundong Li, and Zi Huang. 2022. Self-Supervised Learning for Recommender Systems: A Survey. arXiv preprint arXiv:2203.15876 (2022).

[49] Jiawei Zhang. 2023. Graph-ToolFormer: To Empower LLMs with Graph Reasoning Ability via Prompt Dataset Augmented by ChatGPT. (2023).

[50] Yuhui Zhang, Hao Ding, Zeren Shui, Yifei Ma, James Zou, Anoop Deoras, and Hao Wang. 2021. Language Models as Recommender Systems: Evaluations and Limitations. In I (Still) Can't Believe It's Not Better! NeurIPS 2021 Workshop.

[51] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining 1059–1068.

[52] Han Zhu, Junqi Jin, Chang Tan, Fei Pan, Yifan Zeng, Han Li, and Kun Gai. 2017. Optimized cost per click in taobao display advertising. Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 2191–2200.