Towards Joint Information Retrieval and Recommender Systems

Simone Merlo[1]

[1]University of Padua, Padua, Italy

Abstract

Everyone has some information need related to work tasks, entertainment or other fields. The technological components that are used to answer them usually are Information Retrieval (IR) systems and Recommender Systems (RS). Despite these two types of systems are traditionally developed in isolation, since the nineties it was clear that there were common aspects between IR and RS. Indeed, they are both concerned with retrieving the most relevant documents or items in a collection according to a user request. Only recently some efforts have been directed towards the development of joint IR and RS systems. Nonetheless, most of the created systems focus on gaining the knowledge to carry out one of the two tasks based on the data of the other. A few relevant results really addressed the issue of joint IR and RS but they present several limitations: most of existing models are jointly optimized by aggregating data from both tasks without considering that users' intents in IR and RS sometimes may be different; current models focus on personalization without considering cold-start users; lack of appropriate, public datasets suitable for training and evaluating such models. This paper outlines the author's PhD research objectives in designing new models and resources that allow to overcome the discussed limitations.

Keywords

Information Retrieval, Recommender Systems, Conversational Search, Conversational Recommendation

1. Motivation

The information access scenario is increasingly expanding due to the growing need of people to seek for information. In this field the two major components used to satisfy the users' information needs are: Information Retrieval (IR) systems and Recommender Systems (RS). The former provides the most relevant documents –or items, depending on the application– given a textual query (the information need), the latter suggests to the user some items based on the user's historical interactions (e.g., clicked or purchased items). Despite these two categories of systems are often considered as independent, there exist several points of connection [1]. Indeed, both IR and RS are mainly concerned with providing the users with the piece of information that is most suitable for their information need. However, differences are still present: while IR systems are mainly concerned with retrieving the most relevant documents in a collection, which usually are the "most similar" ones, to the request, RS may be also required to recommend items that are complementary, and not only similar, to the users' preferences.

Nowadays the results of IR systems and RS are often merged together to provide the users with a more comprehensive answer (e.g., the suggested products in the search engines results page). Thus, IR and RS are perceived as joint tasks and even though under the hood there are some variations, the end user does not notice differences between those two technologies. Recently the research community started to develop systems performing both the IR and RS tasks jointly, noticing a promising increase in performance [2, 3, 4]. However, this novel research frontier still presents limitations:

- **(L1)** most of existing models are jointly optimized by simply aggregating data from both tasks without considering that user intents in IR and RecSys sometimes may be different.
- **(L2)** Current models deeply focus on user history without considering that the search task (IR) could be performed by cold-start users or by external people/agents.

SEBD 2025: 33rd Symposium On Advanced Database Systems, June 16–19, 2025, Ischia, Italy
simone.merlo@phd.unipd.it (S. Merlo)
https://dei.unipd.it/~merlosimon (S. Merlo)
0009-0003-8003-4795 (S. Merlo)

- **(L3)** Lack of appropriate, public datasets suitable for training and evaluating models performing both IR and RS tasks jointly.

Moreover, along with the traditional ones, conversational information access systems started to be extensively developed and employed. This is mainly due to the naturalness and ease of interaction that this type of systems enable. Indeed, conversations represent the most natural interaction interface for humans. Conversational Information access systems include both Conversational Search (CS) [5] and Conversational Recommendation (CR) [6] systems. However, in the conversational context CS and CR are still thought as completely independent, even if their integration could lead to significant advantages. Indeed, a user seeking for a recommendation may realize to be in need for additional, external information, or, a user with a "search style" information need may benefit from some recommendation (*e.g.*, when looking for products).

This work summarizes the author's Ph.D. studies [7] in the field of joint IR and RS and outlines his future research directions, aiming to overcome the current limitations of this novel research field, also in the conversational context. The main objective is to develop new publicly available resources including datasets and models that allow to exploit the advantages derived from the joint modeling of IR and RS.

2. Background and Related Work

2.1. Joint Information Retrieval and Recommendation

Traditionally, for historical and industrial reasons, IR systems and RS are developed in isolation. Indeed, when users express an information need that requires both retrieval and recommendation, two separate systems are often employed, one for each task, and the results are then merged together. This is evident when the information need concerns a product, in fact, in that case, the user might be asking both for recommendations or information (retrieval) related to it. Nonetheless, since the nineties it was clear that there were common aspects between IR and RS, in fact, Belkin and Croft [1] consider them as "two sides of the same coin". However, this research area has never been explored until recent times. Indeed, recently, the research community started to develop systems carrying out both IR and RS tasks jointly, noticing a promising increase in performance. In particular, Zamani and Croft [3, 2] have shown that developing such models allows to improve the performance thanks to the sharing of knowledge between IR and RS. However, most of the developed systems focus on refining RS capabilities by exploiting the search data [8] or on gaining the knowledge to carry out one of the two tasks based on the data of the other [2]. Only a few relevant results really addressed the issue of joint IR and RS: a first proposed model based on graph neural networks was SRJGraph [9] while the current state-of the art is represented by the Unified Information Access (UIA) framework [4]. UIA fine-tunes some pre-trained BERT models in order to learn a dense representation in a latent space of both the input data (request) and the items to be retrieved, in such a way that the most relevant matches are the ones whose dense representation is the closest in space with respect to the one of the request. Furthermore, this framework employs an Attentive Personalization Network that is used to grasp knowledge from the previous user interactions.

However, despite some models performing both retrieval and recommendation jointly have been developed there is still no publicly available dataset specifically designed to train and evaluate this kind of systems. Indeed, the majority of existing approaches employ either or private datasets or datasets designed for a single task which are processed to fit for both task (*i.e.*, generating the queries for RS datasets or the user interaction histories for IR datasets). In this context popular datasets are: the Amazon Reviews recommendation dataset [10] and the Amazon ESCI IR dataset [11].

2.2. Conversational Information Retrieval and Recommendation

In the conversational context, several CS [5] and CR [6] systems have been developed. Along with systems also many datasets have been published [12, 13, 14, 15] and for CS dedicated conference tracks have been created, including TREC CAsT [16, 17, 18, 19] and TREC iKAT [20]. However, joint Conversational Search and Recommendation (CSR) has not been explored yet. We ascribe this absence of joint CSR systems to the lack of appropriate CSR dataset. Nonetheless, also in the conversational context, end users would benefit from the joint modeling of IR and RS which would naturally reflect real-world behaviour. Indeed, users seeking for recommendations may need additional information to adjust their query (or the other way around).

2.3. Evaluation in Information Retrieval and Recommendation

The classical IR evaluation measures [21] usually take into account the rank of the retrieved documents for a set of test queries. The most popular measure, in this sense, is Discounted Cumulative Gain (DCG) [22]. Nonetheless, even evaluation measures which do not take into account the rank are widely used: Mean Average Precision (MAP) and recall. However, independently of the employed measures, in IR, the top-k retrieved documents are usually considered to compute them.

In recommendation instead, the scenario is slightly different. The results are usually considered per user rather than per query. Indeed, usually a few user interactions (often only the most recent in time) are retained as test data and are used to evaluate the system's recommendations, while the remaining interactions are employed for training purposes. Nonetheless, DCG, precision and recall represent evaluation measures that are employed also in recommendation. Moreover, other measures that are often used in recommendation are Mean Absolute Error (MAE) and Hit Ratio (HR).

3. Research Goals

The proposed research focuses on the development of innovative joint IR and RS models, overcoming the limitations discussed in Section 1. Therefore we define the following research objectives:

- **(O1) models**: develop new models to carry out IR and RS tasks jointly that are able to capture the differences (and not only the similarities) between the user intents (overcoming L1). Moreover, we aim at improving the current systems user management, in order to effectively exploit the user historical interactions while allowing cold-start and external users to take advantage of the benefits of joint IR and RS (overcoming L2). Finally, we focus on realizing all the models with particular attention to their computational burden.
- **(O2) evaluation**: evaluate the performance of the new models both from the efficiency and effectiveness point of view. Moreover, we aim to generate some datasets that are be specifically built for the evaluation of models performing IR and RS tasks jointly (overcoming L3).

Furthermore, other important objectives are: diffusing the knowledge about joint IR and RS to the research community and providing further and more accurate evidence of the effectiveness of the frameworks based on this concept. We pursue these goals not only in the traditional information access scenario but also in the conversational context.

4. Proposed Approaches and Current Work

In this section we propose some approaches to achieve the objectives and overcome the limitations discussed in Sections 3 and 1, respectively. Moreover, we report the current state of our work and how we applied the proposed approaches.

4.1. Proposed Approaches

In the following we describe the main approaches that we plan to apply for each of the core goals defined in Section 3.

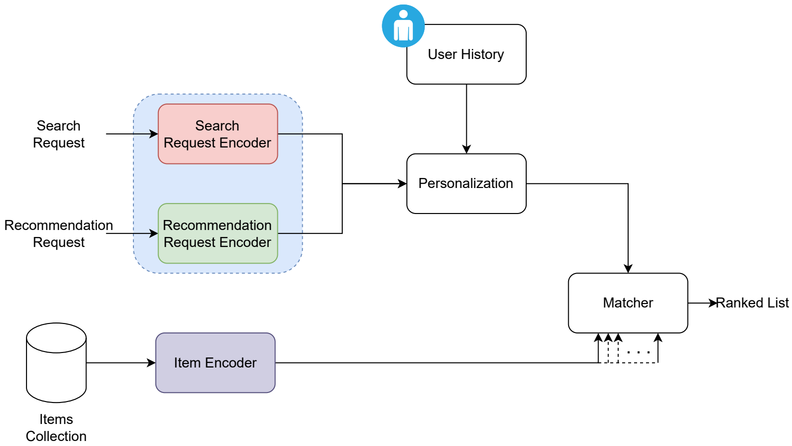

Figure 1: Proposed architecture for new models in the context of joint IR and RS.

4.1.1. Models

In Figure 1 we report a possible general scheme that the new models will adopt in order to fulfill objective O1. In particular, to capture the differences between IR and RS, joint models could encode the retrieval requests (red block) using a different encoder than the one used to for the recommendation requests (green block). Meanwhile, the collection items and documents could be encoded using a single encoder (purple block) that will be shared among the two tasks. This would allow to learn a way of representing the items and documents that is more general and task-independent. Thus, this model architecture would enable the sharing of knowledge between the tasks, while having the ability to obtain representations of the retrieval or recommendation requests that are task specific. In the majority of systems present in literature it is common to learn how to encode the requests (independently from the task) using a single encoder (light blue box in Figure 1), eventually slightly modifying them before the encoding. However, we claim that this is not sufficient. Moreover, both in IR and RS, the most common (modern) practice is to encode the documents or items into latent spaces by means of dense vector representations [23, 24, 25, 26]. Thus, to implement the encoders of this new architecture, the same techniques could be exploited. In particular, it would be possible to both use pre-trained models (e.g., BERT [27]) and fine-tune them or design ad hoc neural networks needing a full training.

The personalization module in Figure 1 is used to adapt the request encodings to take into account also the user preferences. The most popular techniques used to do this exploit attention or graphs [28].

To allow cold-start or external users to exploit these systems there are different possible paths: (1) removing the personalization block when the user is not authenticated or new; (2) creating some "dummy" base profiles to represent the historical interactions of the cold-start or external users. These profiles could be empty, randomly created or contain the most popular interactions.

4.1.2. Evaluation

Using appropriate datasets is crucial to accurately evaluate new methods and models. To create a new dataset specifically built for "joint IR and RS" possible paths to follow are: applying some processing to an already existing dataset or merging datasets that were built for a specific task (i.e. only for IR or only for RS) exploiting some common features (e.g., the items ids for product related datasets). The former has already been widely adopted [3, 8]. Indeed, many product recommendation datasets are enriched with search style queries, which correspond to user reviews or from product categories. The latter , instead has not been explored yet. We ascribe this to the complexity in finding dataset that share the same features. However, this approach could help to improve the quality of the obtained datasets, since both the recommendation and search data would represent real-world data. Finally, given the recent advancements in the deep learning field, Large Language Models (LLMs) could be exploited for creating new joint IR and RS datasets. For example, they could be employed to generate the search style queries that are missing from RS dataset starting from some reviews of the users and the associated products (which usually are both included in RS datasets).

An open challenge in evaluation concerns how to handle the different evaluation methodologies used in IR and RS. This reflects also on the ground truths distributed with the datasets. The ideal scenario would involve devising a uniform evaluation methodology.

4.2. Current Work

Our initial work was directed towards studying and analysing the state-of-the-art in the IR, RS and joint IR and RS fields. For this purpose, we reproduced, replicated and generalized the UIA framework [7] which represents the joint IR and RS state-of-the-art model. For replicability and generalizability, we modified both the training and the data processing pipelines. This reproducibility work revealed that the datasets employed to train this kind of systems and the way in which they are processed play a fundamental role. Moreover, we discovered that the stability of UIA and, in general, of joint IR and RS models, may strongly depend on the task considered (*i.e.*, UIA is much more stable in IR-related tasks than RS-related ones). We plan to exploit the knowledge gathered through this reproducibility work to create new datasets and systems in the joint IR and RS field, taking care of their weaknesses.

To develop new models appropriate datasets are needed. For this purpose, after our initial reproducibility work, we concentrated on the development of a new resource to allow proper training and evaluation of joint IR and RS models. In particular, we focused on the conversational domain and we created a new joint CSR dataset. First we formalized the requirements of the ideal joint CSR collection and then we created a dataset trying to satisfy all the requirements. Our dataset includes all the elements required by the Cranfield paradigm [29] adapted to fit in the conversational domain: (i) **information needs**, which include conversations and user profiles; (ii) **corpora**, Amazon Reviews [10] and MS-MARCO v2.1 [30] which are two well known RS and IR corpora, respectively; (iii) **human annotations**, which include quality assessments of the conversations, intent labels and relevance judgments both non-personalized (for CS) and personalized (for CR). The main challenge that we encountered in the creation of such dataset was related to the differences in evaluation between IR and RS. In particular, we decided to generate an IR style ground truth also for recommendation but we needed to take care of personalization which plays a fundamental role in RS. This work is currently under review. We plan to exploit this dataset for the development of new joint CSR systems.

5. Research Issues for the Doctoral Consortium

In the context of the Doctoral Consortium, the main questions (DC1-3) that require discussion with experienced researchers are:

DC1:Which could be possible architectures and components of joint IR and RS models which allow to consider both the differences and similarities between these two fields (also in the conversational context)?

DC2: The evaluation of joint IR and RS systems is still performed with independent data and methodologies for the two tasks. Which could be possible strategies to uniform the evaluation methodologies considering the importance of personalization in RS?

DC3: The most natural domain of application for joint IR and RS systems is the product domain. Which could be other suitable domains for such systems?

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] N. J. Belkin, W. B. Croft, Information filtering and information retrieval: Two sides of the same coin?, Commun. ACM 35 (1992) 29–38. URL: https://doi.org/10.1145/138859.138861. doi:10.1145/138859.138861.

[2] H. Zamani, W. B. Croft, Learning a joint search and recommendation model from user-item interactions, in: J. Caverlee, X. B. Hu, M. Lalmas, W. Wang (Eds.), WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020, ACM, 2020, pp. 717–725. URL: https://doi.org/10.1145/3336191.3371818. doi:10.1145/3336191.3371818.

[3] H. Zamani, W. B. Croft, Joint modeling and optimization of search and recommendation, in: O. Alonso, G. Silvello (Eds.), Proceedings of the First Biennial Conference on Design of Experimental Search & Information Retrieval Systems, Bertinoro, Italy, August 28-31, 2018, volume 2167 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2018, pp. 36–41. URL: https://ceur-ws.org/Vol-2167/paper2.pdf.

[4] H. Zeng, S. Kallumadi, Z. Alibadi, R. F. Nogueira, H. Zamani, A personalized dense retrieval framework for unified information access, in: H. Chen, W. E. Duh, H. Huang, M. P. Kato, J. Mothe, B. Poblete (Eds.), Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023, ACM, 2023, pp. 121–130. URL: https://doi.org/10.1145/3539618.3591626. doi:10.1145/3539618.3591626.

[5] F. Mo, K. Mao, Z. Zhao, H. Qian, H. Chen, Y. Cheng, X. Li, Y. Zhu, Z. Dou, J. Nie, A survey of conversational search, CoRR abs/2410.15576 (2024). URL: https://doi.org/10.48550/arXiv.2410.15576. doi:10.48550/ARXIV.2410.15576. arXiv:2410.15576.

[6] D. Jannach, A. Manzoor, W. Cai, L. Chen, A survey on conversational recommender systems, ACM Comput. Surv. 54 (2022) 105:1–105:36. URL: https://doi.org/10.1145/3453154. doi:10.1145/3453154.

[7] S. Merlo, G. Faggioli, N. Ferro, A reproducibility study for joint information retrieval and recommendation in product search, in: Advances in Information Retrieval - 47th European Conference on Information Retrieval, ECIR 2025, Lucca, Italy, April 06-10, 2025, Proceedings, Part IV, 2025. URL: https://www.dei.unipd.it/~merlosimon/papers/paper-ECIR-MFF.pdf.

[8] Z. Si, Z. Sun, X. Zhang, J. Xu, X. Zang, Y. Song, K. Gai, J. Wen, When search meets recommendation: Learning disentangled search representation for recommendation, in: H. Chen, W. E. Duh, H. Huang, M. P. Kato, J. Mothe, B. Poblete (Eds.), Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023, Taipei, Taiwan, July 23-27, 2023, ACM, 2023, pp. 1313–1323. URL: https://doi.org/10.1145/3539618.3591786. doi:10.1145/3539618.3591786.

[9] K. Zhao, Y. Zheng, T. Zhuang, X. Li, X. Zeng, Joint learning of e-commerce search and recommendation with a unified graph neural network, in: K. S. Candan, H. Liu, L. Akoglu, X. L. Dong, J. Tang (Eds.), WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022, ACM, 2022, pp. 1461–1469. URL: https://doi.org/10.1145/3488560.3498414. doi:10.1145/3488560.3498414.

[10] Y. Hou, J. Li, Z. He, A. Yan, X. Chen, J. J. McAuley, Bridging language and items for retrieval and recommendation, CoRR abs/2403.03952 (2024). URL: https://doi.org/10.48550/arXiv.2403.03952. doi:10.48550/ARXIV.2403.03952. arXiv:2403.03952.

[11] C. K. Reddy, L. Màrquez, F. Valero, N. Rao, H. Zaragoza, S. Bandyopadhyay, A. Biswas, A. Xing, K. Subbian, Shopping queries dataset: A large-scale ESCI benchmark for improving product search, CoRR abs/2206.06588 (2022). URL: https://doi.org/10.48550/arXiv.2206.06588. doi:10.48550/ARXIV.2206.06588. arXiv:2206.06588.

[12] S. A. Hayati, D. Kang, Q. Zhu, W. Shi, Z. Yu, INSPIRED: toward sociable recommendation dialog systems, in: B. Webber, T. Cohn, Y. He, Y. Liu (Eds.), Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020, Association for Computational Linguistics, 2020, pp. 8142–8152. URL: https://doi.org/10.18653/v1/

2020.emnlp-main.654. doi:10.18653/V1/2020.EMNLP-MAIN.654.
[13] R. Li, S. E. Kahou, H. Schulz, V. Michalski, L. Charlin, C. Pal, Towards deep conversational recommendations, in: S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett (Eds.), Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada, 2018, pp. 9748–9758. URL: https://proceedings.neurips.cc/paper/2018/hash/800de15c79c8d840f4e78d3af937d4d4-Abstract.html.
[14] R. Anantha, S. Vakulenko, Z. Tu, S. Longpre, S. Pulman, S. Chappidi, Open-domain question answering goes conversational via question rewriting, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tür, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, Association for Computational Linguistics, 2021, pp. 520–534. URL: https://doi.org/10.18653/v1/2021.naacl-main.44. doi:10.18653/V1/2021.NAACL-MAIN.44.
[15] E. Choi, H. He, M. Iyyer, M. Yatskar, W. Yih, Y. Choi, P. Liang, L. Zettlemoyer, Quac: Question answering in context, in: E. Riloff, D. Chiang, J. Hockenmaier, J. Tsujii (Eds.), Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, Association for Computational Linguistics, 2018, pp. 2174–2184. URL: https://doi.org/10.18653/v1/d18-1241. doi:10.18653/V1/D18-1241.
[16] P. Owoicho, J. Dalton, M. Aliannejadi, L. Azzopardi, J. R. Trippas, S. Vakulenko, TREC cast 2022: Going beyond user ask and system retrieve with initiative and response generation, in: I. Soboroff, A. Ellis (Eds.), Proceedings of the Thirty-First Text REtrieval Conference, TREC 2022, online, November 15-19, 2022, volume 500-338 of *NIST Special Publication*, National Institute of Standards and Technology (NIST), 2022. URL: https://trec.nist.gov/pubs/trec31/papers/Overview_cast.pdf.
[17] J. Dalton, C. Xiong, J. Callan, TREC cast 2021: The conversational assistance track overview, in: I. Soboroff, A. Ellis (Eds.), Proceedings of the Thirtieth Text REtrieval Conference, TREC 2021, online, November 15-19, 2021, volume 500-335 of *NIST Special Publication*, National Institute of Standards and Technology (NIST), 2021. URL: https://trec.nist.gov/pubs/trec30/papers/Overview-CAsT.pdf.
[18] J. Dalton, C. Xiong, J. Callan, Cast 2020: The conversational assistance track overview, in: E. M. Voorhees, A. Ellis (Eds.), Proceedings of the Twenty-Ninth Text REtrieval Conference, TREC 2020, Virtual Event [Gaithersburg, Maryland, USA], November 16-20, 2020, volume 1266 of *NIST Special Publication*, National Institute of Standards and Technology (NIST), 2020. URL: https://trec.nist.gov/pubs/trec29/papers/OVERVIEW.C.pdf.
[19] J. Dalton, C. Xiong, J. Callan, TREC cast 2019: The conversational assistance track overview, CoRR abs/2003.13624 (2020). URL: https://arxiv.org/abs/2003.13624. arXiv:2003.13624.
[20] M. Aliannejadi, Z. Abbasiantaeb, S. Chatterjee, J. Dalton, L. Azzopardi, TREC ikat 2023: The interactive knowledge assistance track overview, in: I. Soboroff, A. Ellis (Eds.), The Thirty-Second Text REtrieval Conference Proceedings (TREC 2023), Gaithersburg, MD, USA, November 14-17, 2023, volume 500-xxx of *NIST Special Publication*, National Institute of Standards and Technology (NIST), 2023. URL: https://trec.nist.gov/pubs/trec32/papers/Overview_ikat.pdf.
[21] N. Ferro, M. Maistro, Evaluation of IR systems, in: O. Alonso, R. Baeza-Yates (Eds.), Information Retrieval: Advanced Topics and Techniques, volume 60 of *ACM Books*, ACM, 2024, pp. 111–191. URL: https://doi.org/10.1145/3674127.3674132. doi:10.1145/3674127.3674132.
[22] K. Järvelin, J. Kekäläinen, Cumulated gain-based evaluation of IR techniques, ACM Trans. Inf. Syst. 20 (2002) 422–446. URL: http://doi.acm.org/10.1145/582415.582418. doi:10.1145/582415.582418.
[23] S. Lin, J. Yang, J. Lin, Distilling dense representations for ranking using tightly-coupled teachers, CoRR abs/2010.11386 (2020). URL: https://arxiv.org/abs/2010.11386. arXiv:2010.11386.
[24] G. Izacard, M. Caron, L. Hosseini, S. Riedel, P. Bojanowski, A. Joulin, E. Grave, Unsupervised dense information retrieval with contrastive learning, Trans. Mach. Learn. Res. 2022 (2022). URL: https://openreview.net/forum?id=jKN1pXi7b0.

[25] J. Wang, P. Rathi, H. Sundaram, A pre-trained zero-shot sequential recommendation framework via popularity dynamics, in: T. D. Noia, P. Lops, T. Joachims, K. Verbert, P. Castells, Z. Dong, B. London (Eds.), Proceedings of the 18th ACM Conference on Recommender Systems, RecSys 2024, Bari, Italy, October 14-18, 2024, ACM, 2024, pp. 433–443. URL: https://doi.org/10.1145/3640457.3688145. doi:10.1145/3640457.3688145.

[26] K. Balasubramanian, A. Alshabanah, E. Markowitz, G. V. Steeg, M. Annavaram, Biased user history synthesis for personalized long-tail item recommendation, in: T. D. Noia, P. Lops, T. Joachims, K. Verbert, P. Castells, Z. Dong, B. London (Eds.), Proceedings of the 18th ACM Conference on Recommender Systems, RecSys 2024, Bari, Italy, October 14-18, 2024, ACM, 2024, pp. 189–199. URL: https://doi.org/10.1145/3640457.3688141. doi:10.1145/3640457.3688141.

[27] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers), Association for Computational Linguistics, 2019, pp. 4171–4186. URL: https://doi.org/10.18653/v1/n19-1423. doi:10.18653/V1/N19-1423.

[28] M. F. Aljunid, M. D. Huchaiah, M. K. Hooshmand, W. A. Ali, A. M. Shetty, S. Q. Alzoubah, A collaborative filtering recommender systems: Survey, Neurocomputing 617 (2025) 128718. URL: https://doi.org/10.1016/j.neucom.2024.128718. doi:10.1016/J.NEUCOM.2024.128718.

[29] C. Cleverdon, The Cranfield tests on index language devices, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997, p. 47–59.

[30] T. Nguyen, M. Rosenberg, X. Song, J. Gao, S. Tiwary, R. Majumder, L. Deng, MS MARCO: A human generated machine reading comprehension dataset, in: T. R. Besold, A. Bordes, A. S. d'Avila Garcez, G. Wayne (Eds.), Proceedings of the Workshop on Cognitive Computation: Integrating neural and symbolic approaches 2016 co-located with the 30th Annual Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, December 9, 2016, volume 1773 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2016. URL: https://ceur-ws.org/Vol-1773/CoCoNIPS_2016_paper9.pdf.