DIVAN: Deep-Interest Virality-Aware Network to Exploit Temporal Dynamics in News Recommendation

Antonio Ferrara
Politecnico di Bari
Bari, Italy
antonio.ferrara@poliba.it

Marco Valentini
Politecnico di Bari
Bari, Italy
m.valentini7@studenti.poliba.it

Paolo Masciullo
Politecnico di Bari
Bari, Italy
p.masciullo@studenti.poliba.it

Antonio De Candia
Politecnico di Bari
Bari, Italy
a.decandia6@studenti.poliba.it

Davide Abbattista
Politecnico di Bari
Bari, Italy
d.abbattista@studenti.poliba.it

Riccardo Fusco
Politecnico di Bari
Bari, Italy
r.fusco@studenti.poliba.it

Claudio Pomo
Politecnico di Bari
Bari, Italy
claudio.pomo@poliba.it

Vito Walter Anelli
Politecnico di Bari
Bari, Italy
vitowalter.anelli@poliba.it

Giovanni Maria Biancofiore
Politecnico di Bari
Bari, Italy
giovannimaria.biancofiore@poliba.it

Ludovico Boratto
Università di Cagliari
Cagliari, Italy
ludovico.boratto@unica.it

Fedelucio Narducci
Politecnico di Bari
Bari, Italy
fedelucio.narducci@poliba.it

ABSTRACT

In today's era of information overload, personalized news recommendation systems are crucial for connecting users with relevant content. The dynamic nature of user interests and the fleeting popularity of news articles pose significant challenges to accurate prediction. For this reason, the RecSys 2024 Challenge aims to inspire innovative solutions in this field. This study presents DIVAN (Deep-Interest Virality-Aware Network), our solution for the RecSys 2024 Challenge, combining a Deep Interest Network (DIN) for personalized user interest representation with a Virality-aware Click Predictor that utilizes temporal features to estimate click probability based on news popularity. A user-specific weight balances the influence of DIN and virality-based predictions, enhancing personalization and accuracy. Experiments on the Ekstra Bladet dataset from the Challenge demonstrate how promising DIVAN is in accuracy and beyond-accuracy performance.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

Recommender Systems, News Recommendation

RecSys Challenge '24, October 14–18, 2024, Bari, Italy

ACM ISBN 979-8-4007-1127-5/24/10
https://doi.org/10.1145/3687151.3687153

ACM Reference Format:
Antonio Ferrara, Marco Valentini, Paolo Masciullo, Antonio De Candia, Davide Abbattista, Riccardo Fusco, Claudio Pomo, Vito Walter Anelli, Giovanni Maria Biancofiore, Ludovico Boratto, and Fedelucio Narducci. 2024. DIVAN: Deep-Interest Virality-Aware Network to Exploit Temporal Dynamics in News Recommendation. In *ACM RecSys Challenge 2024 (RecSys Challenge '24), October 14–18, 2024, Bari, Italy*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3687151.3687153

1 INTRODUCTION

Recommendation systems are essential for enhancing user experiences on online platforms by providing personalized content suggestions [9]. News recommendation, in particular, faces unique challenges due to the dynamic nature of news content and user preferences [12]. Addressing these challenges is crucial for engaging users and delivering relevant content. The RecSys 2024 Challenge[1] aims to inspire innovative solutions in this field, focusing on predicting user clicks on news articles using the Ekstra Bladet News Recommendation Dataset (EB-NeRD)[2], a comprehensive dataset with millions of user interactions, article information, and temporal dynamics spanning six weeks. Researchers must develop models that leverage these data points to predict user clicks, considering factors like evolving user interests [12], news cycles [6], and the ephemeral nature of news content [2, 10]. To achieve this, participants are provided with a rich dataset, including a user's browsing history, details about the browsing session, personal metadata, and a list of candidate news articles.

This paper introduces **DIVAN**[3] (Deep-Interest Virality-Aware Network), a novel framework combining collaborative filtering with

[1] https://recsys.acm.org/recsys24/challenge/
[2] https://recsys.eb.dk/dataset/
[3] Implementation available at https://github.com/sisinflab/DIVAN.

virality-based predictions. DIVAN utilizes an Attention-based Click predictor based on Deep-Interest Network (DIN) [14] to model user interests through attention mechanisms, weighting past interactions depending on their relevance to candidate news articles. The framework incorporates contextual information (time of day, device type) and article features (text embeddings, category, sentiment). Recognizing that virality can affect users' click behaviours, DIVAN integrates a Virality-aware Click Predictor. This component uses temporal (time since publication, hour of day) and content features (sentiment, category) to estimate click probability based on news virality, effectively recommending news accounting for user interests and temporal dynamics. DIVAN dynamically balances the influence of collaborative filtering and virality-based predictions through a user-specific weight, catering to individual preferences for viral versus personalized content. The two key contributions of this paper are: i) the joint modeling of user interests and news virality, ii) a user-specific weighting mechanism to estimate a tailored trade-off between personalization and virality.

2 RELATED WORK

Personalized news recommendation seeks to tailor news selection to users' individual interests, which are typically modeled based on logs collected during user navigation. News articles, as well as user profiles, are often encoded in news recommender systems using explicit features or through deep-learning techniques extracting information from text, titles, and images [12]. News articles often have short popularity and virality lifecycles, characterized by rapid surges in views followed by declines in interest and click rates. This temporal aspect of popularity can be exploited to enhance news recommender systems. For instance, SCENE [5] employs a two-stage personalized news recommendation approach that uses content characteristics along with popularity and recency as adjustment factors in final news ranking. PENR [11] includes a popularity prediction task to enhance the news encoder's performance by counting and normalizing click behaviours within a multi-task learning framework. PPSR [1] leverages news popularity to address cold-start issues by retrieving semantically similar news for non-clicked articles, distinguishing between genuinely unpopular news and cold-start news, thereby improving prediction accuracy.

Our solution to the RecSys challenge builds on PP-Rec [7], which mixes a time-aware popularity predictor with user-news matching scores from separate user and item encoders. We enhance this approach by integrating concepts from the Deep Interest Network [14] into the PP-Rec, unifying user profiles with candidate news to better capture hidden relationships and improve recommendations.

3 DATA AND FEATURE ENGINEERING

EB-NeRD Dataset. The **E**kstra **B**ladet **Ne**ws **R**ecommendation **D**ataset (EB-NeRD) includes data from approximately 2 million user interactions with Ekstra Bladet's online platform over a six-week period (April 27th to June 8, 2023), chosen to avoid anomalies from major events like holidays or elections. The dataset comprises four components: *Articles* (with features for 100,000 news articles), *Artifacts* (semantic embeddings from titles, subtitles, and bodies using models like multilingual BERT and RoBERTa), *Behavior* (40 million logs of user clicks over a 7-day period), and *History* (articles clicked by users in the 21 days prior). Behavior and History data are split into training, validation, and testing sets, and no user profile embeddings are included.

Data Preprocessing.

The data preprocessing phase involved removing redundant data, handling missing values, reducing dimensionality, and preparing the data for model input. We collected content from training, validation, and test sets for news articles, encoding features like named entities, topics, and sentiment labels. Principal Component Analysis (PCA) was used to reduce embedding size, retaining 80%, 90%, and 95% of the original variance with 64, 128, and 256 components, respectively. RoBERTa embeddings, which retained the highest variance with 64 components, were used in the model input. We excluded impressions with multiple clicked articles, likely due to data collection errors, and trimmed user histories to the 50 most recent articles, focusing on short-term interests. Missing values for features such as read time and scroll percentage were imputed using mean values.

Preparing Data for the Model. The preprocessed data is engineered and prepared to construct a dataset representing user-article interactions. We expand the behavior data so that each news article in the inview corresponds to a positive or negative sample, indicating whether the user clicks the article. Each sample includes information to describe the user, the news article, the temporal aspects, and the context of their interaction. Sensitive features (i.e., gender, age, and postcode) as well as future features unavailable in the test set, like read time and scroll percentage of the clicked article, are discarded. Instead, new temporal features are introduced: i) the time elapsed between article publication and appearance in the user's impression, ii) the hour of the day, and iii) the day of the week the impression occurred. These features reveal time-sensitive patterns in news consumption, such as article freshness, user reading habits, and the news popularity lifecycle. The original features used for this computation (impression, publication, and last modification time) are dropped due to the high correlation with the new ones.

4 DIVAN

In this section, we elaborate on our solution, diving into the main aspects that guided our choices.

4.1 Model Architecture

Motivated by the peculiar dynamics of the popularity of news articles, we employ a manifold architecture, represented in Figure 1, to discover user interests as conditioned by the virality of the candidate news. On the one hand, according to Zhou et al. [14], we employ an Attention-based Click Predictor to adaptively model the user profile based on the previous interactions and the current candidate item. On the other hand, we incorporate a Virality-aware Click Predictor to capture the popularity dynamics of news articles.

4.1.1 Representational Layer. We feed our model with samples constructed as detailed in Section 3. We project the candidate article, user history articles, origin article, context characteristics, and user characteristics into a multidimensional embedding space. These projections are then concatenated to compose a comprehensive representation of each data component. This approach enables a

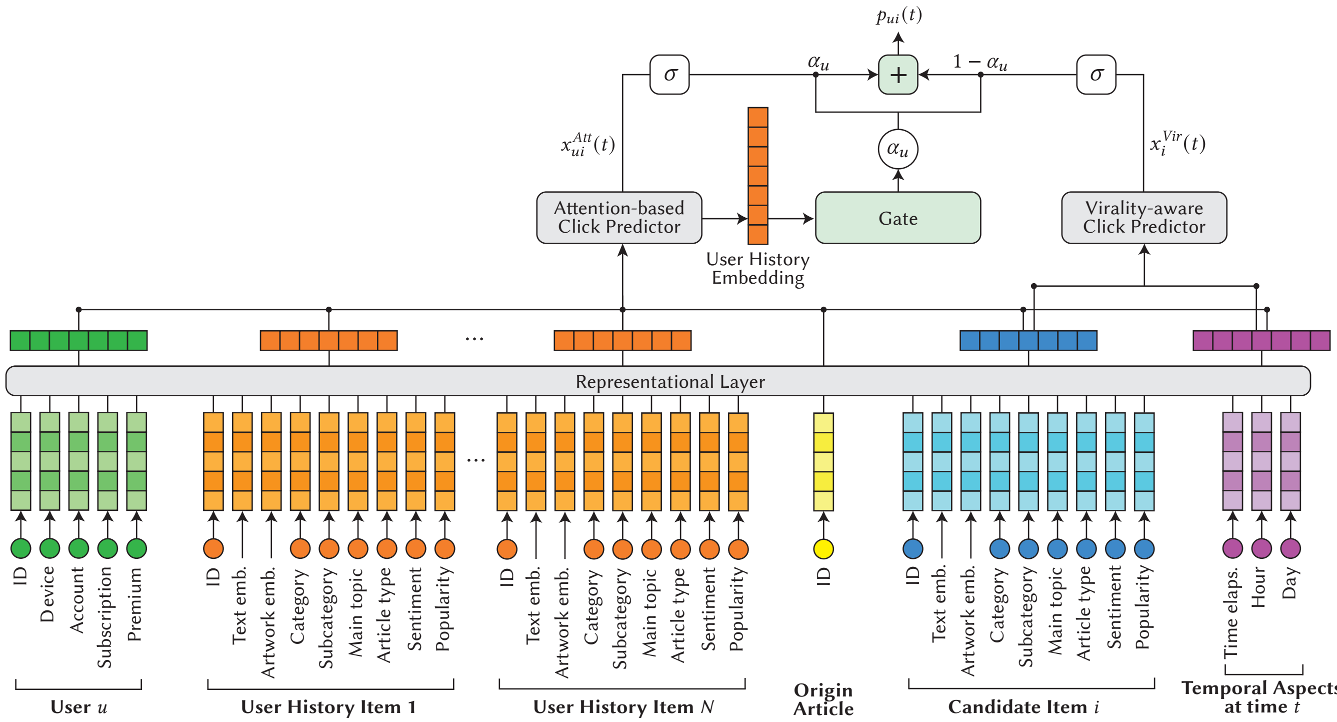

Figure 1: Representation of the architecture of DIVAN with its inputs and outputs.

unified representation that effectively captures the diverse aspects of the user, the news article, and the interaction context.

4.1.2 Attention-based Click Predictor. Our Attention-based Click Predictor utilizes the architecture of DIN[4] [14]. This model enhances user representation through a cross-attention mechanism that identifies user interests by weighting previously interacted items based on their relevance to the candidate article. This method effectively captures the user's diverse and context-specific interests. We then define a dynamic user representation that depends on both the user's interaction history and the candidate news article. This user profile is concatenated with representations of user characteristics, the candidate news, the origin item, and temporal aspects, yielding the input for a feed-forward neural network, which predicts the score $x_{ui}^{Att}(t)$ for the user clicking on the candidate item.

4.1.3 Virality-aware Click Predictor. News articles typically exhibit short virality lifecycles, marked by rapid surges in views followed by swift declines in interest and click rates. This pattern is especially pronounced in the news domain. Viral news often attracts a diverse audience with varying interests, making it particularly effective for engaging cold users with little to no historical data.

These observations motivate the introduction of a Virality-aware Click Predictor, which accounts for the ephemeral nature of news popularity and its unique dynamics. To encode these dynamics, we incorporate news recency features into the Virality-aware Click Predictor. This allows the predictor to estimate the probability of user-article interactions based on the content features of the candidate item, being aware of the temporal context of the interaction.

The architecture of the Virality-aware Click Predictor is a straightforward multilayer perceptron. Its input is derived from the representational layer, while its output is the score $x_i^{Vir}(t)$, indicating the likelihood of an item being clicked at time t based on its virality. This design enhances the learning of embeddings for both the candidate item and its temporal aspects. These embeddings are shared with the Attention-based Click Predictor, but here we focus solely on the temporal dynamics of the interaction.

4.1.4 User-based Click Prediction. We combine the attention-based and the virality-aware click predictions through a weighted linear combination that counterbalances their impact. The trade-off between collaborative and virality contributions varies on a user basis, since some users are more inclined to click on viral items, while others prefer news that better aligns with their specific interests.

Notably, given a user u and a news item i, a personalized trade-off parameter α_u is used to combine the click probability predictions $\sigma(x_{ui}^{Att}(t))$ and $\sigma(x_i^{Vir}(t))$, obtained by applying a sigmoid function to the outputs of the Attention-based and Virality-aware Click Predictors, respectively. The parameter α_u is the output of a Gate, i.e., a sigmoid-activated MLP, trained to determine the propensity of a user u towards viral news from the representation of such user computed by the Attention-based Click Predictor. Finally, the click probability $p_{ui}(t)$ is computed as follows:

$$p_{ui}(t) = \alpha_u \cdot \sigma(x_{ui}^{Att}(t)) + (1-\alpha_u) \cdot \sigma(x_i^{Vir}(t)). \tag{1}$$

[4]Available in the FuxiCTR library [15].

Table 1: Summary of the competition results in terms of accuracy and beyond-accuracy metrics

Model	AUC	MRR	nDCG@5	nDCG@10	Intralist Diversity	Novelty	Serendipity	Coverage
Most Popular	0.4171	0.2573	0.2770	0.3800	0.7328	12.5449	0.8031	0.020
Most Viral	0.4326	0.2641	0.2881	0.3873	0.7328	12.5449	0.8031	0.020
:D (1st)	0.8933	0.7350	0.7923	0.8002	0.7697	3.7017	0.7678	0.632
BlackPearl (2nd)	0.8825	0.7165	0.7762	0.7861	0.6877	4.9304	0.7538	0.320
Tom3TK (3rd)	0.8707	0.7029	0.7631	0.7751	0.7121	4.7701	0.7543	0.612
FeatureSalad (Academic 1st)	0.8494	0.6638	0.7296	0.7451	0.7328	12.5449	0.8031	0.020
DIN	0.6927	0.4724	0.5292	0.5771	0.7421	5.9069	0.7785	0.692
DIVAN	0.7091	0.4851	0.5435	0.5891	0.6580	8.7076	0.7745	0.872

4.2 Model Optimization

The training phase updates the embeddings of all features in the model, ensuring accurate user-article click predictions that reflect actual user interactions with news articles.

We train the model using a Bayesian Personalized Ranking (BPR) loss [8], which assumes that a user u prefers a consumed item over a non-consumed item and allows us to optimize the model by maximizing, for each pair of clicked news and non-clicked news, a function of their difference. Thus, for each clicked article, we sample 14 negative interactions to let the model learn the user interests.

The final loss function combines contributions from the Attention-based and Virality-aware Click Predictors. It uses Binary Cross-Entropy (BCE) Loss for the Virality-aware Click Predictor to distinguish clicked from non-clicked samples and BPR Loss for the Attention-based Click Predictor to prioritize ranking clicked news higher. Additionally, a penalty term is included to prevent the Virality-aware Click Predictor from being dominated by the Attention-based Click Predictor, avoiding model degeneration into DIN.

5 EXPERIMENTS

In this section, we present the experimental protocol and the results obtained by our model, followed by a discussion of the findings.

Experimental Setup. Throughout the competition, we extensively explored and validated various configurations for our architecture. However, here we describe only the details of the chosen model.

The Attention-based Click Predictor utilizes a deep neural network with hidden layers consisting of 1024, 512, and 256 units, all activated by ReLU functions. The network is trained using a dropout rate of 0.1. Additionally, the attention mechanism comprises hidden layers with 512 and 256 units, also activated by ReLU functions and trained with a dropout rate of 0.2. The embedding regularizer has been set to 1.e-4. On the other hand, the Virality-aware Click Predictor has been designed with hidden layers of 512, 256, and 128 units, activated by ReLU functions, and trained with a 0.3 dropout rate. The gate module consists of hidden layers with 128 and 64 units and is trained with a dropout rate of 0.2. The optimization has been performed using Adam [4] with a learning rate of 0.001, β_1 set to 0.9, β_2 set to 0.999, and without weight decay.

Discussion. The evaluation, considering accuracy and beyond-accuracy [3], has been performed using CodaBench[5] [13]. The results in Table 1 compare DIVAN with i) two unpersonalized models, i.e., Most Popular and a virality-based ranker (named Most Viral) where the item's popularity is discounted by a negative exponential factor based on the time elapsed between its publication and its appearance in the impression, ii) the top three teams and the best academic team of the RecSys 2024 Challenge, and iii) the DIN model, equivalent to DIVAN without the Virality-aware Click Predictor.

Notably, DIVAN performs better than DIN in terms of accuracy metrics and significantly outperforms the non-personalized Most Popular and Most Viral recommenders. The improvement over DIN, though not sufficient to reach the top of the Challenge leaderboard, empirically supports the introduction of the Virality-aware Click Predictor, and demonstrates its critical role in capturing the temporal aspects influencing user interests and news article trends.

Interestingly, DIVAN has a lower Intralist Diversity than DIN and the other models, but achieves higher novelty and coverage. This suggests that while DIVAN continues to capture user interests effectively thanks to its Attention-based Click Predictor, it trades DIN's ability to explore a broad range of interests for the opportunity to provide fresh and trending articles from a narrower set of interests, thanks to the Virality-aware Click Predictor. This aspect is crucial for long-tail content discovery, enhancing user satisfaction by not repetitively recommending only the most popular items.

6 CONCLUSION

This paper presented DIVAN, a solution to the RecSys 2024 Challenge. DIVAN addresses the challenges of dynamic user interests and ephemeral news popularity by jointly modeling user preferences and content virality. Experimental results on the provided EB-NeRD dataset demonstrated superior performance in accuracy metrics compared to baseline models while also improving novelty and coverage. These findings suggest that integrating virality signals on a user-specific basis enhances recommendation quality and content discovery. Future research directions may include refining the model architecture and evaluating its efficacy across diverse news domains and user demographics.

ACKNOWLEDGMENTS

The authors acknowledge the partial support by: CT_FINCONS_III, OVS Fashion Retail Reloaded, LUTECH DIGITALE 4.0, Secure Safe Apulia, IDENTITA, REACH-XY. The authors acknowledge the CINECA award under the ISCRA initiative for the availability of high-performance computing resources.

[5]https://www.codabench.org/

REFERENCES

[1] Yuhan Ding, Bang Wang, Xiangyang Cui, and Minghua Xu. 2024. Popularity prediction with semantic retrieval for news recommendation. *Expert Syst. Appl.* 247 (2024), 123308. https://doi.org/10.1016/J.ESWA.2024.123308

[2] Chong Feng, Muzammil Khan, Arif Ur Rahman, and Arshad Ahmad. 2020. News Recommendation Systems - Accomplishments, Challenges & Future Directions. *IEEE Access* 8 (2020), 16702–16725. https://doi.org/10.1109/ACCESS.2020.2967792

[3] Marius Kaminskas and Derek Bridge. 2017. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Trans. Interact. Intell. Syst.* 7, 1 (2017), 2:1–2:42.

[4] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.). http://arxiv.org/abs/1412.6980

[5] Lei Li, Dingding Wang, Tao Li, Daniel Knox, and Balaji Padmanabhan. 2011. SCENE: a scalable two-stage personalized news recommendation system. In *Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2011, Beijing, China, July 25-29, 2011*, Wei-Ying Ma, Jian-Yun Nie, Ricardo Baeza-Yates, Tat-Seng Chua, and W. Bruce Croft (Eds.). ACM, 125–134. https://doi.org/10.1145/2009916.2009937

[6] Nuno Moniz, Luís Torgo, Magdalini Eirinaki, and Paula Branco. 2017. A Framework for Recommendation of Highly Popular News Lacking Social Feedback. *New Gener. Comput.* 35, 4 (2017), 417–450. https://doi.org/10.1007/S00354-017-0019-X

[7] Tao Qi, Fangzhao Wu, Chuhan Wu, and Yongfeng Huang. 2021. PP-Rec: News Recommendation with Personalized User Interest and Time-aware News Popularity. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli (Eds.). Association for Computational Linguistics, 5457–5467. https://doi.org/10.18653/V1/2021.ACL-LONG.424

[8] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *UAI 2009, Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, June 18-21, 2009*, Jeff A. Bilmes and Andrew Y. Ng (Eds.). AUAI Press, 452–461. https://www.auai.org/uai2009/papers/UAI2009_0139_48141db02b9f0b02bc7158819ebfa2c7.pdf

[9] Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). 2022. *Recommender Systems Handbook*. Springer US. https://doi.org/10.1007/978-1-0716-2197-4

[10] David Vallet, Shlomo Berkovsky, Sebastien Ardon, Anirban Mahanti, and Mohamed Ali Kâafar. 2015. Characterizing and Predicting Viral-and-Popular Video Content. In *Proceedings of the 24th ACM International Conference on Information and Knowledge Management, CIKM 2015, Melbourne, VIC, Australia, October 19 - 23, 2015*, James Bailey, Alistair Moffat, Charu C. Aggarwal, Maarten de Rijke, Ravi Kumar, Vanessa Murdock, Timos K. Sellis, and Jeffrey Xu Yu (Eds.). ACM, 1591–1600. https://doi.org/10.1145/2806416.2806556

[11] Jingkun Wang, Yipu Chen, Zichun Wang, and Wen Zhao. 2021. Popularity-Enhanced News Recommendation with Multi-View Interest Representation. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (Eds.). ACM, 1949–1958. https://doi.org/10.1145/3459637.3482462

[12] Chuhan Wu, Fangzhao Wu, Yongfeng Huang, and Xing Xie. 2023. Personalized News Recommendation: Methods and Challenges. *ACM Trans. Inf. Syst.* 41, 1 (2023), 24:1–24:50. https://doi.org/10.1145/3530257

[13] Zhen Xu, Sergio Escalera, Adrien Pavão, Magali Richard, Wei-Wei Tu, Quanming Yao, Huan Zhao, and Isabelle Guyon. 2022. Codabench: Flexible, easy-to-use, and reproducible meta-benchmark platform. *Patterns* 3, 7 (2022), 100543.

[14] Guorui Zhou, Xiaoqiang Zhu, Chengru 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. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*, Yike Guo and Faisal Farooq (Eds.). ACM, 1059–1068. https://doi.org/10.1145/3219819.3219823

[15] Jieming Zhu, Jinyang Liu, Shuai Yang, Qi Zhang, and Xiuqiang He. 2021. Open Benchmarking for Click-Through Rate Prediction. In *CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1 - 5, 2021*, Gianluca Demartini, Guido Zuccon, J. Shane Culpepper, Zi Huang, and Hanghang Tong (Eds.). ACM, 2759–2769. https://doi.org/10.1145/3459637.3482486