

Panoramic Interests: Stylistic-Content Aware Personalized Headline Generation

Junhong Lian^{†¶}

Institute of Computing Technology,
Chinese Academy of Sciences
Beijing, China
lianjunhong23s@ict.ac.cn

Xiang Ao^{*†‡¶}

Institute of Computing Technology,
Chinese Academy of Sciences
Beijing, China
aoxiang@ict.ac.cn

Xinyu Liu^{*§¶}

Institute of Computing Technology,
Chinese Academy of Sciences
Beijing, China
liuxinyu@ict.ac.cn

Yang Liu^{†¶}

Institute of Computing Technology,
Chinese Academy of Sciences
Beijing, China
liuyang2023@ict.ac.cn

Qing He^{†¶}

Institute of Computing Technology,
Chinese Academy of Sciences
Beijing, China
heqing@ict.ac.cn

Abstract

Personalized news headline generation aims to provide users with attention-grabbing headlines that are tailored to their preferences. Prevailing methods focus on user-oriented content preferences, but most of them overlook the fact that diverse stylistic preferences are integral to users' panoramic interests, leading to suboptimal personalization. In view of this, we propose a novel Stylistic-Content Aware Personalized Headline Generation (SCAPE) framework. SCAPE extracts both content and stylistic features from headlines with the aid of large language model (LLM) collaboration. It further adaptively integrates users' long- and short-term interests through a contrastive learning-based hierarchical fusion network. By incorporating the panoramic interests into the headline generator, SCAPE reflects users' stylistic-content preferences during the generation process. Extensive experiments on the real-world dataset PENS demonstrate the superiority of SCAPE over baselines.

CCS Concepts

- **Computing methodologies** → **Natural language generation;**
- **Information systems** → **Personalization; Data mining.**

Keywords

Personalized Headline Generation, Stylistic-Content Awareness Fusion, User Preference Modeling, Large Language Models

ACM Reference Format:

Junhong Lian, Xiang Ao, Xinyu Liu, Yang Liu, and Qing He. 2025. Panoramic Interests: Stylistic-Content Aware Personalized Headline Generation. In *Companion Proceedings of the ACM Web Conference 2025 (WWW Companion '25)*, April 28-May 2, 2025, Sydney, NSW, Australia.

^{*}Corresponding authors.

[†]Key Lab of Intelligent Information Processing, ICT, CAS. Also with the Key Lab of AI Safety, CAS, Beijing, China.

[‡]Institute of Intelligent Computing Technology, CAS, Suzhou, China.

[§]High Performance Computer Research Center, ICT, CAS.

[¶]The authors are also with the University of Chinese Academy of Sciences, CAS.

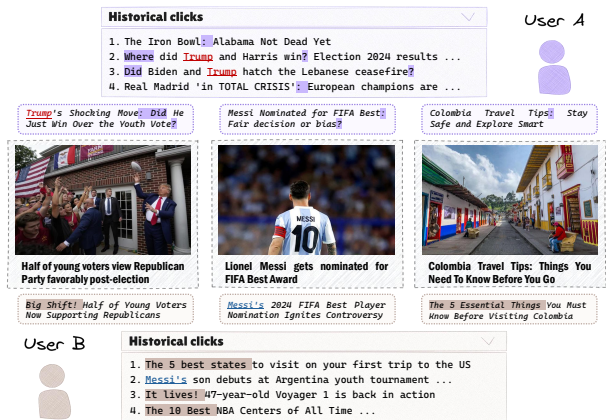


Figure 1: Illustration of the Joint Influence of Content Interests and Stylistic Preferences on Headline Personalization.

'25), April 28-May 2, 2025, Sydney, NSW, Australia. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3701716.3715539>

1 Introduction

Generating attractive news headlines has been firmly established as a special form of text summarization [7, 12]. Numerous efforts have focused on generating attention-grabbing headlines through personalized approaches, recognizing that readers with diverse preferences may find different focal points in the same news [1, 2]. These methods aimed to build engaging headlines tailored to individual users' reading interests by utilizing auxiliary information, such as user profiles and historical clicks [15, 16, 19, 21]. In spite of that, existing personalized approaches primarily emphasize user-oriented content-driven headlines, while largely overlooking the stylistic features of news headlines.

The stylistic features of news headlines are pivotal in journalism communication [3, 5]. Boosting headline engagement through text style transfer has become a widely-used approach [9, 20]. Style transfer-based methods incorporate specific stylistic elements or employ interrogative forms in news headlines to capture readers' attention [9, 14]. Nevertheless, these style transfer-based methods rely on a single, uniform stylistic strategy to boost engagement

risks crossing into clickbait [1, 19]. To improve content consistency, recent works [10, 20] have emphasized the style-content duality in headline generation, focusing on transferring to prototype headline styles while maintaining content decoupling. Despite these efforts, style transfer-based methods are unsatisfactory because they have yet to achieve user-oriented personalization.

In this work, we focus on encompassing both content interests and stylistic preferences to fully meet users' needs for personalized headline generation. The reason is that even readers with similar content preferences may exhibit distinct style preferences for headlines due to different personal reading habits. For example, consider the case illustrated in Figure 1, wherein users' content focus and stylistic preferences jointly affect the personalization of headlines. We mark their content interests and stylistic patterns with different colors and text highlighting. Based on User A's historical click patterns, we can infer that User A prefers the "subject-subordinate title" format and interrogative forms to increase engagement, while User B favors "numbered list" headlines and short exclamatory sentences. Hence, we contend that both content interests and stylistic preferences are interwoven to form users' panoramic interests in personalization and must be considered simultaneously. However, simultaneously addressing this issue is challenging. A primary obstacle is how to extract the inherent content and stylistic features encapsulated in headlines without ground-truth labels. Moreover, another challenge is to decouple implicit user panoramic interests and integrate them into headlines without explicit supervision.

To remedy the above challenges, we propose a novel framework named **Sty**listic-**C**ontent **A**ware **P**ersonalized Headline Generation (SCAPE). We first designed a headline inference module that uses LLM to infer content and stylistic attributes for each news headline individually. It subsequently produces an offline embedding table that encapsulates these attributes at the headline-level. Furthermore, a hierarchical gated fusion network is devised to adaptively integrate users' long- and short-term content interests and stylistic preferences. In addition, a self-supervised strategy is adopted to decouple the intertwined user preferences. Lastly, a personalized injection module leverages the fused user representation to guide a lightweight generator to produce final outputs. This ensures that the generated headlines reflect user's panoramic interests. Our main contributions are summarized as follows:

- To the best of our knowledge, this work represents the first attempt to incorporate both content and style preferences from user profiles for personalized news headline generation.
- We introduce a novel framework, SCAPE, which extracts the inherent features of headlines through LLM collaboration and facilitates personalized generation via a hierarchical fusion of both content interests and stylistic preferences.
- Extensive evaluations conducted on the real-world personalized news headline generation benchmark PENS [2] demonstrate the superior performance of our method.

2 Methodology

2.1 Problem Formulation

Consider a news database denoted as $\mathcal{D} = \{n_i = (t_i, b_i)\}_{i=1}^{|\mathcal{D}|}$, where t_i and b_i represent the original headline and the body of news n_i , and $|\mathcal{D}|$ is the total number of news. For a given user u , we denoted

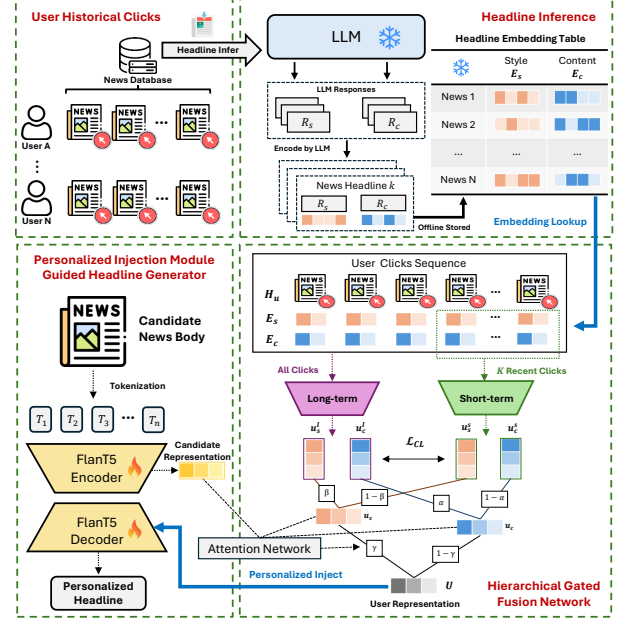


Figure 2: Architecture of the Proposed SCAPE Framework.

u 's click history as $H_u = [t_{h_1}, t_{h_2}, \dots, t_{h_L}]$, comprising L clicked headlines t_{h_j} ($j = 1, \dots, L$), where each t_{h_j} satisfies $t_{h_j} \in \{t_i \mid n_i = (t_i, b_i) \in \mathcal{D}\}$. Then, given a candidate news $n_c = (t_c, b_c) \in \mathcal{D}$, our target is to generate a personalized headline P_c^u for user u based on u 's click history H_u and the body b_c of the candidate news.

2.2 Our SCAPE Framework

In this section, we describe our proposed SCAPE framework, which is illustrated in Figure 2. It consists of a headline inference module, a hierarchical gated fusion network, and a personalized injection module for headline generator.

2.2.1 Headline Inference. Extracting the inherent features from headlines without ground-truth labels is challenging, but LLMs with extensive world knowledge offer a promising solution for inferring high-level latent concepts. Building on this, the headline inference module in SCAPE employs an instruction-tuned LLM, denoted as LLM_{inst} , to infer text style and content interests from headlines t_i in news database \mathcal{D} as follows:

$$R_s = LLM_{inst}(t_i, \mathcal{P}_{style}) \quad R_c = LLM_{inst}(t_i, \mathcal{P}_{content}) \quad (1)$$

where \mathcal{P}_{style} and $\mathcal{P}_{content}$ are the prompts designed to instruct the LLM to infer stylistic features and content interests, respectively, with R_s and R_c as corresponding responses.

Inspired by instruction-following text embedding [13], we design task-specific instructions I_s and I_c for embedding style and interest. These instructions are concatenated with news headlines and their corresponding responses, which are then encoded by an embedding-based LLM, denoted as LLM_{emb} . The offline stored headline embedding table E is constructed as follows:

$$E_s = LLM_{emb}([t_i, I_s, R_s]) \quad E_c = LLM_{emb}([t_i, I_c, R_c]) \quad (2)$$

$$E = \{(E_s(t_i), E_c(t_i)) \mid n_i = (t_i, b_i) \in \mathcal{D}\} \quad (3)$$

where $E_s(t_i)$ and $E_c(t_i)$ represent the style and content embeddings of headline t_i , respectively, and E is the set of style-content embedding pairs for all headlines in \mathcal{D} .

2.2.2 Hierarchical Stylistic-Content Awareness Fusion. Given that historical clicks reflect both stable personal traits and recent short-term characteristics, we employ an attention mechanism to aggregate all clicks and a GRU network for the K recent clicks. This yields the user’s long-term content interests representation u_c^l and short-term content interests representation u_c^s as follows:

$$\tilde{E}_c(t_{h_j}) = \text{MLP}(E_c(t_{h_j})) \quad \forall j \in \{1, \dots, L\} \quad (4)$$

$$u_c^l = \text{Attn}(\tilde{E}_c(t_{h_1}), \dots, \tilde{E}_c(t_{h_L})) \quad (5)$$

$$u_c^s = \text{GRU}([\tilde{E}_c(t_{h_{L-K+1}}), \dots, \tilde{E}_c(t_{h_L})]) \quad (6)$$

where $\tilde{E}_c(t_{h_j})$ is the projected content embedding for the j -th historical click with $L > K$. Similarly, we obtain the user’s long-term and short-term stylistic preferences, denoted as u_s^l and u_s^s .

We then obtain the fusion of long- and short-term representations corresponding to different types of interests, as follows:

$$\alpha = \sigma(W_{gc} \cdot [u_c^l, u_c^s] + b_{gc}) \quad \beta = \sigma(W_{gs} \cdot [u_s^l, u_s^s] + b_{gs}) \quad (7)$$

$$u_c = \alpha \cdot u_c^l + (1 - \alpha) \cdot u_c^s \quad u_s = \beta \cdot u_s^l + (1 - \beta) \cdot u_s^s \quad (8)$$

where σ denotes the sigmoid activation function, and α is the gating weight for content interests. W_{gc} and b_{gc} are learnable parameters. Similarly, we derive the gating weight β and representation u_s for style preferences through the same process.

Subsequently, we integrate the u_c and u_s based on the attention weights derived from the candidate article representation. This stylistic-content awareness fusion mechanism enables the stepwise integration of diverse representations of the user.

2.2.3 Personalized Injection Module Guided Generator. Once the user representation U is obtained, SCAPE integrates it into a decoder of the lightweight headline generator through the personalized injection module. Specifically, the user representation U is added to the input embeddings \mathbf{X} of each token in the decoder. This user-specific vector is subsequently propagated through the residual flow, thereby influencing the generated headlines. By integrating user preferences at the token level, the headline generator can better align its output with the user’s interests, producing stylistic-content-aware personalized headlines that effectively reflect their panoramic interests, formulated as follows:

$$\mathbf{X}' = \mathbf{X} + (\mathbf{1}_n \otimes U) \quad (9)$$

where $\mathbf{X} \in \mathbb{R}^{n \times d}$, $U \in \mathbb{R}^d$, n is the sequence length, d is the embedding dimension, and $\mathbf{1}_n \in \mathbb{R}^n$ is an all-ones vector.

2.3 Disentanglement Strategy

To prevent the u_c and u_s in Section 2.2.2 from collapsing into trivial forms, we adopt a mechanism inspired by [22] to promote the disentanglement of long- and short-term (LS-term) style and content interest representations. Specifically, we aggregate users’ historical LS-term clicks using mean aggregation to derive proxies for stylistic preferences ($p_{u_s}^l$ and $p_{u_s}^s$) and content interests ($p_{u_c}^l$ and $p_{u_c}^s$).

We employ contrastive learning between user LS-term representations and proxies, ensuring that the learned representations of LS-term style or content are more similar to their respective proxies than to opposite proxies. For content interests modeling, a

user’s long-term content interest u_c^l should be more similar to the long-term content proxy $p_{u_c}^l$ than to the short-term content proxy $p_{u_c}^s$ or the long-term style proxy $p_{u_s}^l$.

The triplet loss function is used to achieve contrastive learning, and the final contrastive learning loss can be computed as follows:

$$\mathcal{L}_{\text{CL}} = \frac{1}{|\mathcal{R}|} \sum_{R \in \mathcal{R}} \frac{1}{|\mathcal{N}_R|} \sum_{N \in \mathcal{N}_R} \max(0, d(R, P_R) - d(R, N) + m) \quad (10)$$

where $\mathcal{R} = \{u_s^l, u_s^s, u_c^l, u_c^s\}$ is user representations set, and \mathcal{N}_R is negative proxies set. P_R is the positive proxy corresponding to R . d denotes the Euclidean distance, and m is a positive margin value.

3 Experiments

3.1 Experimental Setup

3.1.1 Datasets and Baselines. We use the publicly available personalized news headline generation dataset PENS [2] as benchmark. The dataset comprises 500,000 anonymized impressions from over 445,000 users and 113,762 news articles, capturing detailed historical click data to reflect nuanced personalized preferences. The test set includes 3,940 news items annotated by 103 users, each providing 200 unique parallel headlines as the gold standard for personalized headline evaluation.

We compare SCAPE with SOTA personalized headline generation methods on the PENS benchmark, including pointer-network-based frameworks [1, 2, 21] and methods based on pre-trained language models [15, 19]. We also expanded our evaluation by employing a prompt-based method to comprehensively investigate the performance of LLMs against strong personalized baselines, evaluating both open-source LLMs [4, 18] of various sizes and closed-source LLMs [6, 8, 11] via API services.

3.1.2 Evaluation Metrics. We use ROUGE metrics to measure lexical similarity, with ROUGE-1 and ROUGE-2 for informativeness, and ROUGE-L for fluency. To evaluate the factual consistency of generated headlines, we follow previous work [19] that reported Fact Scores. As for personalization, due to the lack of a widely accepted evaluation method, we design a pairwise comparison task [17], where strong LLMs evaluate candidates against original headlines based on the user’s historical clicks. To minimize bias, we swap the contextual order of candidates and perform two independent assessments, marking inconsistencies as ties. Personalization performance is reported as "win/tie/lose" outcomes between the candidates and original headlines across baseline methods.

3.1.3 Implementation Details. We use FlanT5-base as the backbone for the headline generator. The model is pre-trained on general headline generation, as [15, 19], for 2 epochs with a peak learning rate (LR) of $1e - 4$ and cosine decay. Early stopping is applied within 5 epochs for subsequent steps, with peak LRs set to $1e - 3$, $1e - 6$, and $1e - 5$, respectively. Other details are consistent with prior work [19]. The collaborative LLMs are Qwen2.5-72B-Instruct and GTE-Qwen2-7B-Instruct. We employ the Huggingface ROUGE pipeline as prior work [15] and re-report results for FPG [19]. In the pairwise comparison evaluation of personalization performance, we use Qwen2.5-72B-Instruct as the judge to obtain the comprehensive assessment. We use NVIDIA A800 80GB GPU for our experiments.

Table 1: Performance of the Compared Baseline Methods.

Types	Methods	ROUGE-1	ROUGE-2	ROUGE-L	Fact Scores
Open-source LLMs	FlanT5-large	24.91	8.48	21.13	80.83
	FlanT5-XL	27.73	10.08	23.23	84.16
	Qwen2.5-1.5B	27.45	9.45	22.24	83.34
	Qwen2.5-7B	<u>29.54</u>	<u>10.47</u>	<u>24.04</u>	<u>88.67</u>
LLMs API services	GLM-4-Flash	26.91	8.40	21.47	85.88
	GLM-4-Air	27.45	8.89	22.03	88.02
	DeepSeek-V2.5	28.27	9.55	22.60	86.90
	GPT-4o	<u>29.58</u>	<u>11.03</u>	<u>24.26</u>	<u>90.11</u>
Personalized models	PENS-NRMS	26.15	9.37	21.03	50.73
	PENS-NAML	28.01	10.72	22.24	50.16
	PNG	28.78	11.27	22.39	51.23
	EUI-PENS	32.34	13.93	26.90	NA
	FPG	33.06	13.76	26.78	<u>89.55</u>
	GTP	<u>33.84</u>	<u>14.23</u>	<u>27.85</u>	NA
	SCAPE_{ours}	34.26*	14.79*	28.36*	92.36*

The symbol * denotes the significance level with $p \leq 0.05$. **Bold font** indicates the best-performing method. Underline indicates the second-best results in the group.

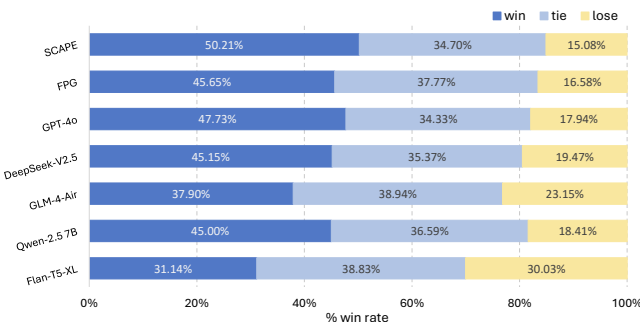


Figure 3: Win Rates in Personalization Evaluation.

3.2 Results and Analysis

Table 1 compares SCAPE with baseline methods, demonstrating its superior performance and setting a new benchmark for personalized headline generation in SOTA results. SCAPE significantly outperforms other methods in informativeness and fluency. This result highlights that SCAPE better improves content coverage through its stylistic-content-aware user modeling via LLM collaboration. Furthermore, SCAPE emphasizes the importance of balancing long- and short-term content interests and stylistic preferences to improve the factual consistency of personalized headlines.

Figure 3 shows the win rates of personalization evaluation using LLMs with customized prompts as judges. While GPT-4o shows higher win rates than the personalized model FPG, LLMs still struggle to generate personalized headlines based on prompt engineering. This underscores the challenge posed by the complexity and diversity of user’s clicks preferences. Our SCAPE framework takes a crucial step forward in personalized headlines generation by considering both user-oriented content interests and modeling linguistic style preferences in users’ historical clicks.

4 Conclusion

In this paper, we propose SCAPE, a novel framework to tackle the challenges of insufficiently user preference modeling in personalized headline generation. SCAPE introduces a headline inference module that extracts the inherent stylistic features and content interests from news headlines without explicit supervision. A hierarchical gated fusion mechanism is further introduced

to dynamically combine both long- and short-term interests for panoramic user modeling, which then guides the headline generator to produce stylistic-content aware personalized headlines. Extensive experiments on the PENS dataset show that SCAPE sets a new state-of-the-art benchmark for personalized headline generation.

Acknowledgments

The research work supported by National Key R&D Plan No.2022YFC3303303, the National Natural Science Foundation of China under Grant (No. U2436209, 62476263). Xiang Ao is also supported by the Project of Youth Innovation Promotion Association CAS, Beijing Nova Program 20230484430, the Innovation Funding of ICT, CAS under Grant No. E461060.

References

- [1] Xiang Ao, Ling Luo, Xiting Wang, et al. 2023. Put Your Voice on Stage: Personalized Headline Generation for News Articles. *ACM TKDD* (2023).
- [2] Xiang Ao, Xiting Wang, Ling Luo, Ying Qiao, Qing He, and Xing Xie. 2021. PENS: A Dataset and Generic Framework for Personalized News Headline Generation. In *ACL* 2021.
- [3] Allan Bell. 1991. *The language of news media*. Blackwell Oxford.
- [4] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, et al. 2024. Scaling instruction-finetuned language models. *JMLR* (2024).
- [5] John Fiske. 2010. *Introduction to communication studies*. Routledge.
- [6] Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Dan Zhang, Diego Rojas, Guanyu Feng, Hanlin Zhao, et al. 2024. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *arXiv preprint arXiv:2406.12793* (2024).
- [7] Xiaotao Gu, Yuning Mao, Jiawei Han, et al. 2020. Generating Representative Headlines for News Stories. In *The Web Conference* 2020.
- [8] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276* (2024).
- [9] Di Jin, Zhijing Jin, Joey Tianyi Zhou, et al. 2020. Hooks in the Headline: Learning to Generate Headlines with Controlled Styles. In *ACL* 2020.
- [10] Mingzhe Li, Xiuying Chen, Min Yang, et al. 2021. Learning to Write Eye-Catching Headlines via Disentanglement. In *AAAI* 2021.
- [11] Aixin Liu, Bei Feng, Bin Wang, Bingxuan Wang, Bo Liu, Chenggang Zhao, Chengqi Deng, Chong Ruan, Damai Dai, Daya Guo, et al. 2024. Deepseek-v2: A strong, economical, and efficient mixture-of-experts language model. *arXiv preprint arXiv:2405.04434* (2024).
- [12] Ling Luo, Xiang Ao, Yan Song, Feiyang Pan, Min Yang, and Qing He. 2019. Reading like HER: Human Reading Inspired Extractive Summarization. In *EMNLP-IJCNLP* 2019.
- [13] Letian Peng, Yuwei Zhang, Zilong Wang, Jayanth Srinivasa, Gaowen Liu, Zihan Wang, and Jingbo Shang. 2024. Answer is All You Need: Instruction-following Text Embedding via Answering the Question. In *ACL* 2024.
- [14] Kai Shu, Suhang Wang, Thai Le, et al. 2018. Deep Headline Generation for Clickbait Detection. In *ICDM* 2018.
- [15] Yun-Zhu Song, Yi-Syuan Chen, Lu Wang, and Hong-Han Shuai. 2023. General then Personal: Decoupling and Pre-training for Personalized Headline Generation. *TACL* (2023).
- [16] Xiaoyu Tan, Leijun Cheng, Xihe Qiu, et al. 2024. Enhancing Personalized Headline Generation via Offline Goal-Conditioned Reinforcement Learning with Large Language Models. In *KDD* 2024.
- [17] Peiyi Wang, Lei Li, Liang Chen, Zefan Cai, Dawei Zhu, Binghui Lin, Yunbo Cao, Qi Liu, Tianyu Liu, and Zhifang Sui. 2024. Large language models are not fair evaluators. In *ACL* 2024.
- [18] An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, et al. 2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671* (2024).
- [19] Zhao Yang, Junhong Lian, and Xiang Ao. 2023. Fact-Preserved Personalized News Headline Generation. In *ICDM* 2023.
- [20] Boning Zhang and Yang Yang. 2023. MediaHG: Rethinking Eye-catchy Features in Social Media Headline Generation. In *EMNLP* 2023.
- [21] Kui Zhang, Guangquan Lu, Guixian Zhang, Zhi Lei, and Lijuan Wu. 2022. Personalized Headline Generation with Enhanced User Interest Perception. In *ICANN* 2022.
- [22] Yu Zheng, Chen Gao, Jianxin Chang, Yanan Niu, Yang Song, Depeng Jin, and Yong Li. 2022. Disentangling long and short-term interests for recommendation. In *The Web Conference* 2022.