Fact-Preserved Personalized News Headline Generation

Zhao Yang*†‡, **Junhong Lian***†‡, Xiang Ao†‡§✉

† Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, CAS, Beijing 100190, China.
‡ University of Chinese Academy of Sciences, Beijing 100049, China.
§ Institute of Intelligent Computing Technology, Suzhou, CAS.

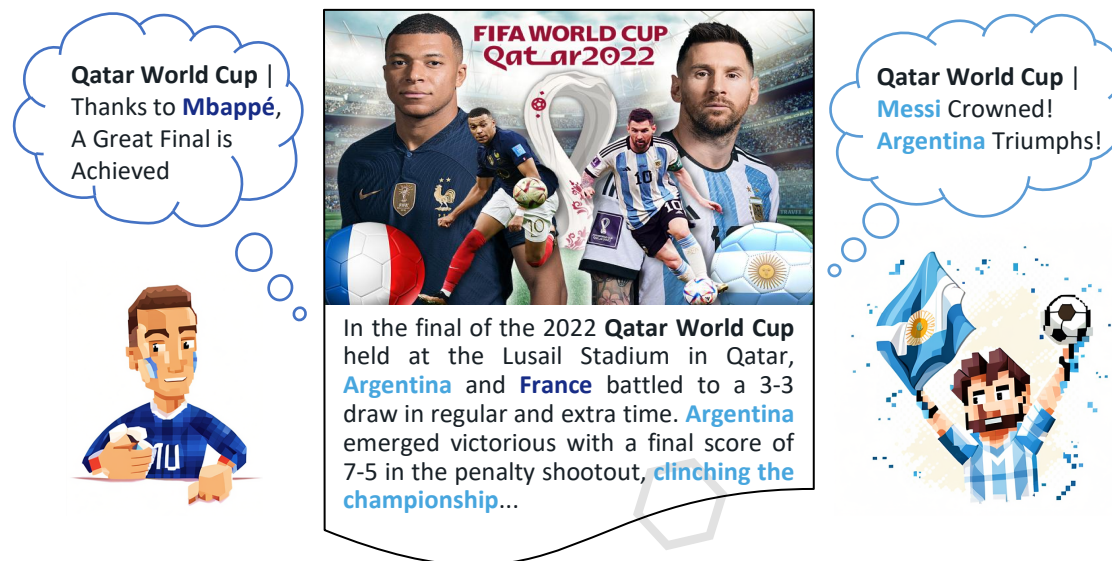{yangzhao20s, lianjunhong23s, aoxiang}@ict.ac.cn

IEEE ICDM 2023, Shanghai, China.
Dec, 2023

Introduction

Personalization in News Headlines

Personalized news headline generation involves generating a *user-specific headline* based on the *user's reading interest*.

The fundamental idea is that readers with different preferences can find their focal characters or aspects they are interested in, even within the same news.

Qatar World Cup | Thanks to **Mbappé**, A Great Final is Achieved

Qatar World Cup | **Messi** Crowned! **Argentina** Triumphs!

In the final of the 2022 **Qatar World Cup** held at the Lusail Stadium in Qatar, **Argentina** and **France** battled to a 3-3 draw in regular and extra time. **Argentina** emerged victorious with a final score of 7-5 in the penalty shootout, **clinching the championship**...

Motivation

Eye-attracting Headlines or Potential Clickbait

Excessive personalization may threaten the *factual consistency* of news headlines.

Original Headline

Bucks Rip Warriors 118-99

News Article

Giannis Antetokounmpo scored 30 points and had his third triple-double of the season to help Milwaukee rout the Golden State Warriors 118-99 on Thursday night, ...

Curry with 12 points, eight rebounds, four assists and two steals ...

Stephen Curry gave a Golden State Warriors fan a special moment ...

A's Historical Clicks

1. Giannis Antetokounmpo cleared from protocols
2. Bucks' Giannis Antetokounmpo(illness) available on Friday
3. Giannis Antetokounmpo expected to play on Christmas

Giannis Antetokounmpo✓ triple-double help Bucks rip
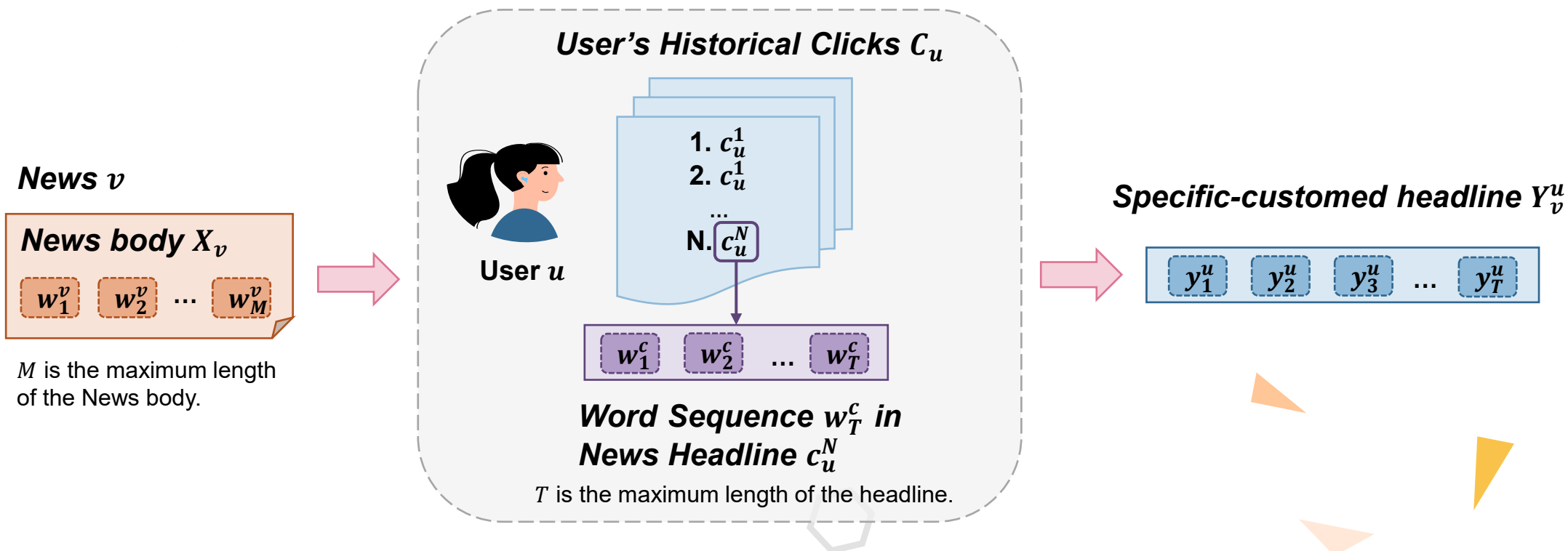
Eye-attracting Headlines✓

B's Historical Clicks

1. Curry and the Warriors take on the Rockets
2. Stephen eager to rectify Christmas Day struggles against Suns
3. Splash brothers Curry, Thompson lead Warriors past Pistons
4. Rockets look to take 2-0 regular-season series lead over the Spurs

Stephen Curry✓ bring young Rockets fan✗ special moment

Potential Clickbait✗

Problem Formulation

We perceived news headline generation as a *headline-specialized summarization* task.

News v

News body X_v

w_1^v w_2^v ... w_M^v

M is the maximum length of the News body.

User's Historical Clicks C_u

User u

1. c_u^1
2. c_u^1
...
N. c_u^N

w_1^c w_2^c ... w_T^c

Word Sequence w_T^c in News Headline c_u^N

T is the maximum length of the headline.

Specific-customed headline Y_v^u

y_1^u y_2^u y_3^u ... y_T^u

Method

The Framework of FPG

It has N layers of transformer blocks in both news encoder and decoder.

(a) is *history encoder*, (b) is *personalized news encoder*, and (c) is *user-guided decoder*.

ICDM 2023 – 23rd IEEE International Conference on Data Mining
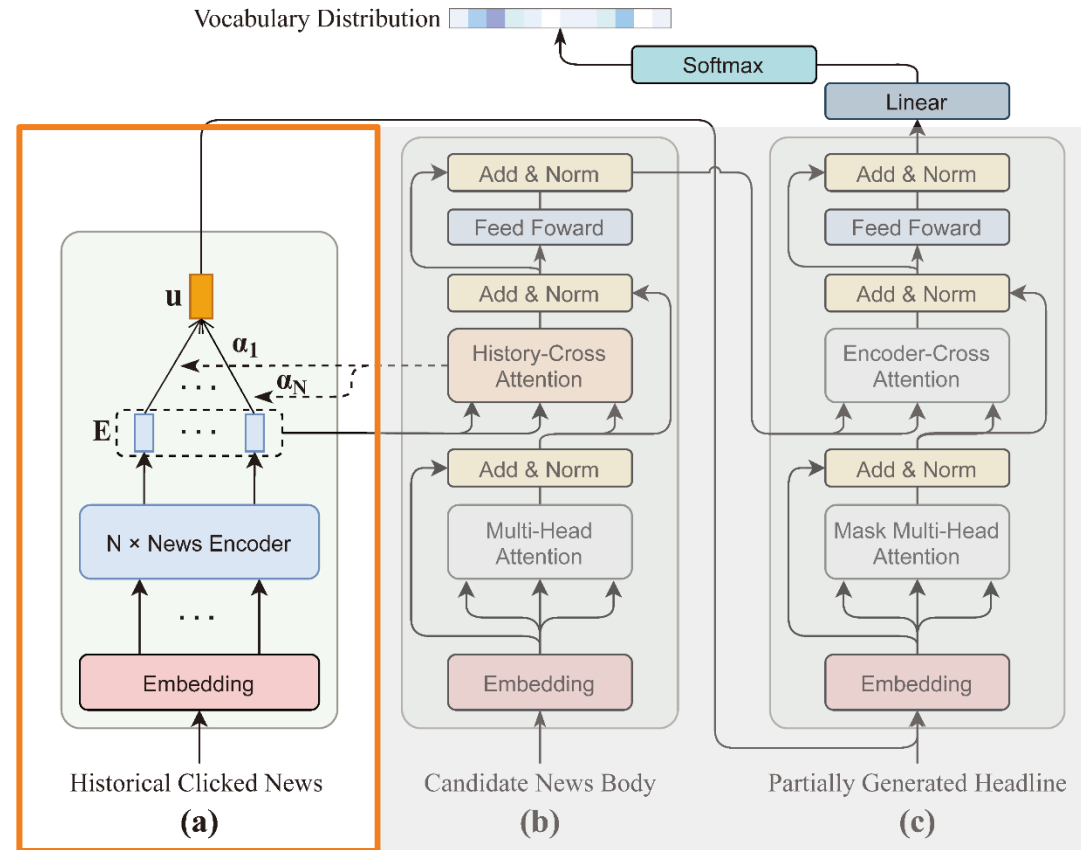
Method
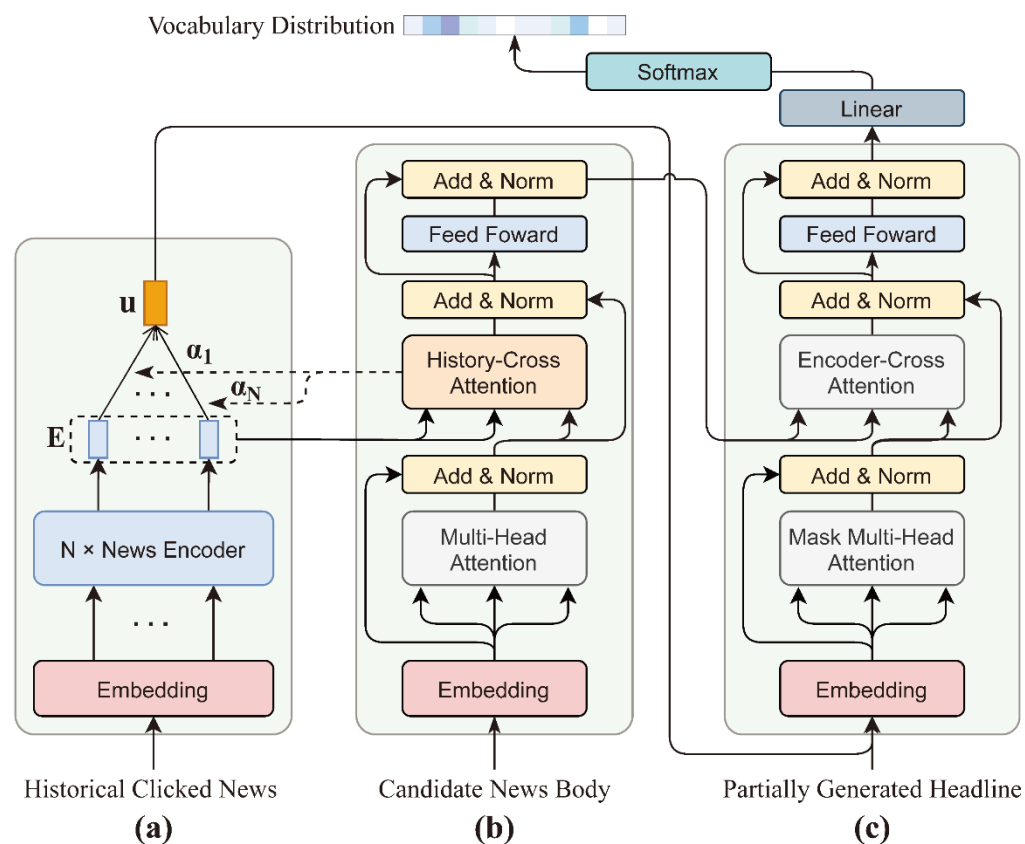
History Encoder

The history encoder aims to *learn users' interest representations* based on their *historical behaviors*. We denote E_u as *the news-level user interests* of u.

$$a_j = \text{Softmax}(\mathbf{h}_j^c \tanh(\mathbf{V}_a \mathbf{h}_j^{c\top} + \mathbf{b}_a))$$

$$\mathbf{e_c} = \sum_{j=1}^{T} a_j \mathbf{h}_j^c$$

Where $\mathbf{V}_a \in \mathbb{R}^{d_e \times d_e}, \mathbf{b}_a \in \mathbb{R}^{d_e \times 1}$.

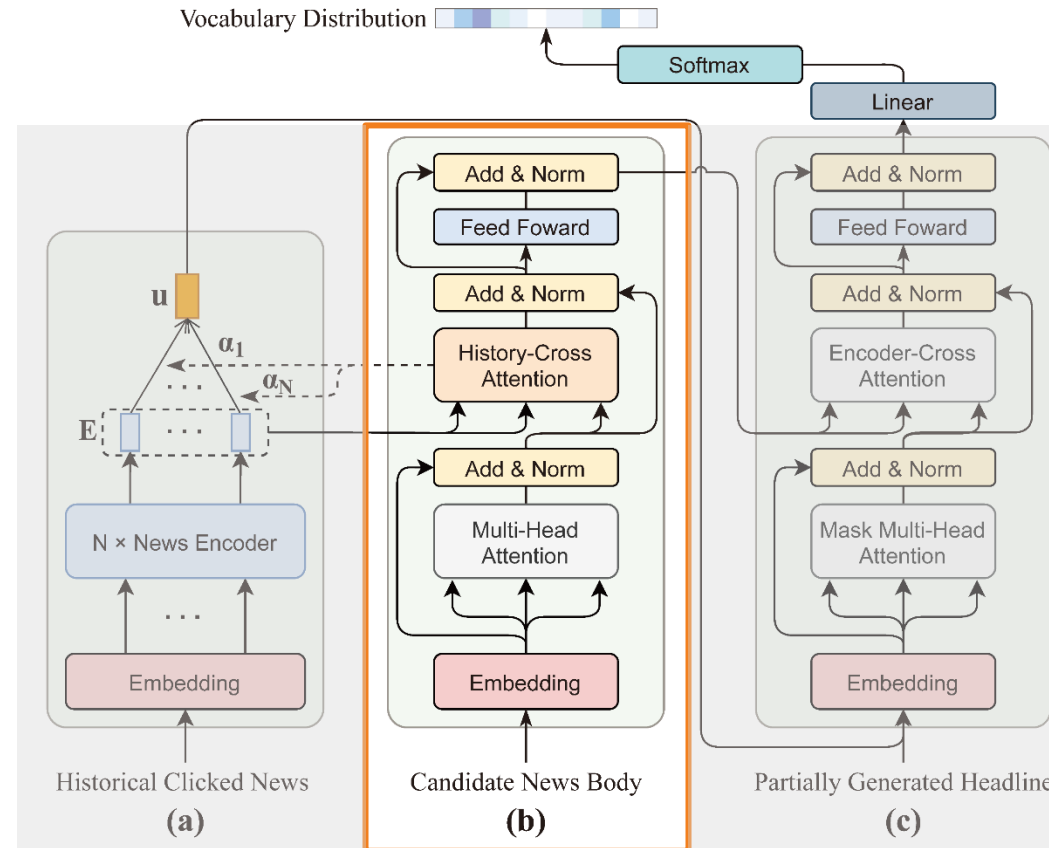ICDM 2023 – 23rd IEEE International Conference on Data Mining

Personalized News Encoder

The personalized news encoder intends to *encode a candidate news* by matching user interests, using a *history-cross attention* to integrate the user's historical behaviors.

$$\mathbf{Q}_h = \mathbf{X}^\top \mathbf{H}^Q, \ \mathbf{K}_h = \mathbf{E}_u^\top \mathbf{H}^K, \ \mathbf{V}_h = \mathbf{E}_u^\top \mathbf{H}^V$$

$$\mathbf{X}_p = \mathrm{Softmax}(\frac{\mathbf{Q}_h^\top \mathbf{K}_h}{\sqrt{d_e}})\mathbf{V}_h$$

Where $\mathbf{H}^Q, \mathbf{H}^K, \mathbf{H}^V \in \mathbb{R}^{d_e \times d_e}$ are learnable parameter

ICDM 2023 – 23rd IEEE International Conference on Data Mining

User-guided Decoder

The user-guided decoder generates a personalized headline based on a *fact-aware* user's *global interests representation* and the relevance of *news-level interests*.

$$\mathbf{u} = \sum_{j=1}^{N} \alpha_j \mathbf{e}_j$$

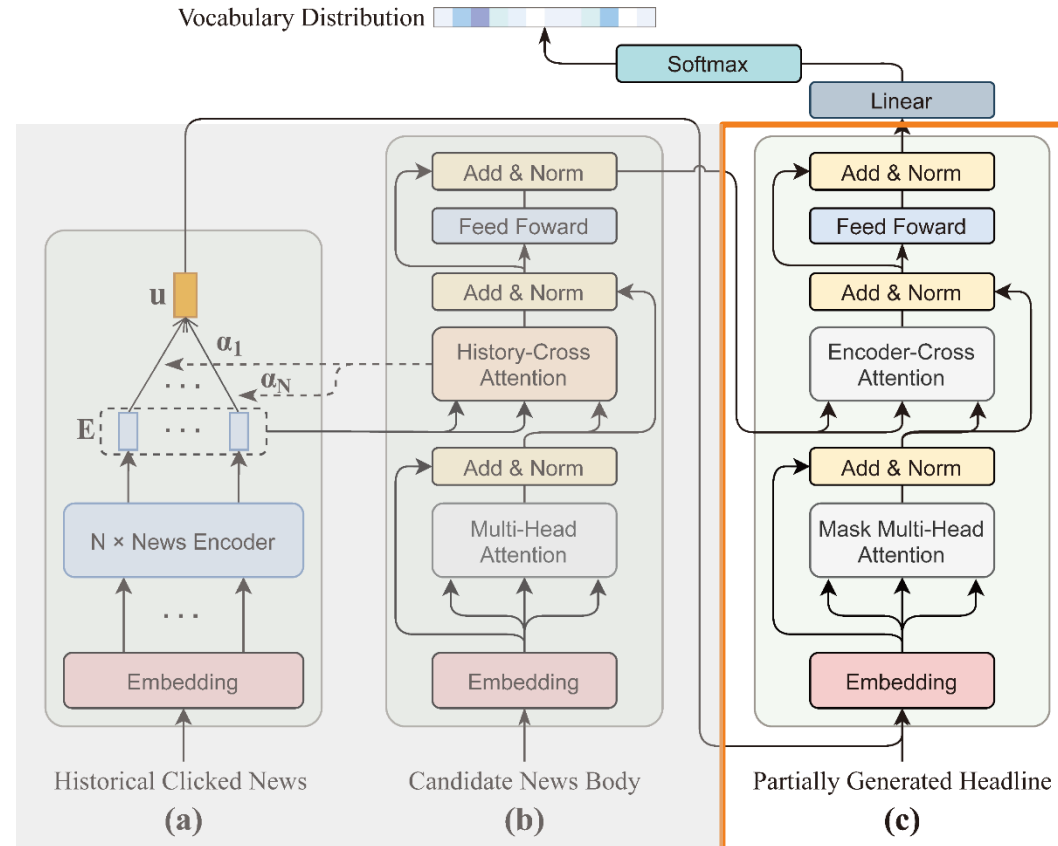Where $\{\alpha_{1,...,N}\}$ are attention scores of history-cross attention

$$P(\hat{y}_t) = \text{Softmax}(\mathbf{S}_t^{\mathsf{N}} \mathbf{W}_v + \mathbf{b}_v)$$

Where $\mathbf{W}_v \in \mathbb{R}^{d_e \times \|\mathbb{V}\|}$ and $\mathbf{b}_v \in \mathbb{R}^{1 \times \|\mathbb{V}\|}$ are learnable parameter

We use *the negative log-likelihood* as the loss function to train the headline generation model:

$$\mathcal{L}_{NLL} = -\sum_{i=1}^{T} \log P(y_i | y_1, \ldots, y_{i-1}; X, C)$$

Where T is the length of the generated headline.

ICDM 2023 – 23rd IEEE International Conference on Data Mining

Fact-enhanced Training

We apply a *multi-stage fact-enhanced training phase* to improve the factual consistency of generated headlines by minimizing a *contrastive learning loss L_{CLL}*.

$$\mathcal{L}_{CLL} = - \underbrace{\mathbb{E}_{x,c,y^+ \in \mathcal{D}*} \log P(y^+; x, c)}_{L_C^+}$$

$$- \underbrace{\mathbb{E}_{x,c,y^- \in \mathcal{D}*} \log(1 - P(y^-; x, c))}_{L_C^-}$$

$$\mathcal{D}* = \{X, C, Y^+, Y^-\}$$

C: **User's Historical Clicks**

X: **Candidate News**

Y: **Specific-customed Headline**

Y^+: **High Factual Score Headline**

Y^-: **Factual Errors Headline**

D^l: **News Clicked by a Maximum of l users**
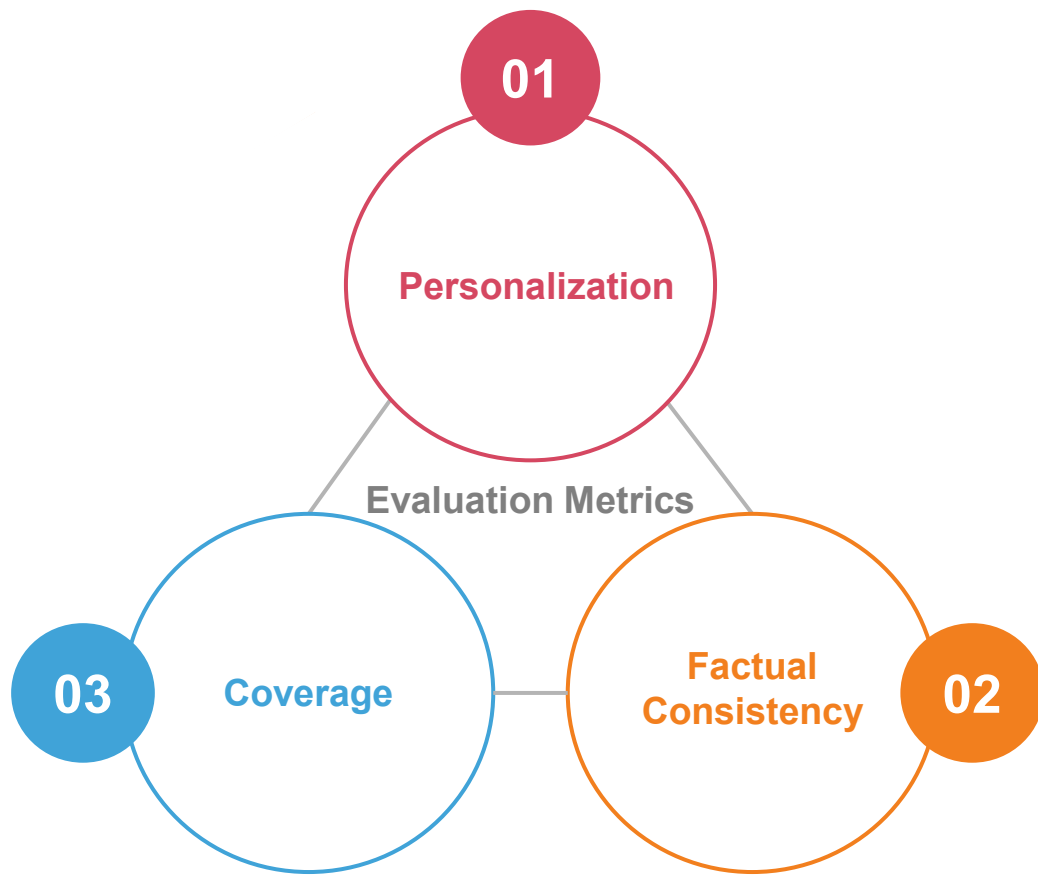

Algorithm 1: Training schedule of FPG

Input: $\mathcal{C} = \{X, Y\}$, $\mathcal{D}_l = \{X, C, Y\}$,
$\mathcal{D}* = \{X, C, Y^+, Y^-\}$;

Initialize Transformer parameters ξ with BART-base;

Other parameters θ are randomly initialized;

1. Pre-train the headline generator with MLE;

2. Froze ξ to train the history encoder;

for *epoch=1:epoch₂* **do**
    Sample $\{X_i, C_i, Y_i\}$ from $\mathcal{D}_l$;
    Update θ via minimizing Eq.(7);
end

3. Train all parameters of FPG;

for *epoch=1:epoch₃* **do**
    Sample $\{X_i, C_i, Y_i\}$ from $\mathcal{D}_l$;
    Update θ and ξ via minimizing Eq.(7);
end

4. Fact-enhanced training;

for *epoch=1:epoch₄* **do**
    Sample $\{X_i, C_i, Y_i^+, Y_i^-\}$ from $\mathcal{D}*$;
    Update ξ via minimizing Eq.(8);
end

Evaluation Metrics

We adopt a *three-pronged approach* to comprehensively evaluate headline quality.

01 Personalization

Evaluation Metrics

03 Coverage

Factual Consistency **02**

Personalization

$$\mathrm{P}_{sim}(\max/\mathrm{avg}) = \mathrm{Max/Mean} \underset{c \in C_u}{sim(c, y)}$$

Factual **Consistency**

FactCC Score

Coverage

The average F1 of ROUGE scores

ICDM 2023 – 23rd IEEE International Conference on Data Mining

Overall Performances on PENS benchmark

Our method outperforms others in *balancing personalization with factual consistency* in news headline generation, achieving the *highest coverage scores*, which suggests it creates more informative and fluent headlines that align well with users' interests.

TABLE II
THE OVERALL PERFORMANCES OF COMPARED METHODS.

Methods	Metrics					
	P_C(avg)	P_C(max)	FactCC	ROUGE-1	ROUGE-2	ROUGE-L
PGN	2.71	16.20	65.08	19.86	4.76	18.83
PG+Transformer	2.66	16.99	53.26	20.64	4.03	18.62
Transformer	2.70	16.36	61.61	19.54	4.72	16.36
BART	2.72	17.13	86.67	26.27	9.88	22.85
PENS-NPA	3.43	21.48	51.51	21.08	4.03	19.03
PENS-EBNR	3.31	20.04	52.74	20.88	3.87	18.66
PENS-NRMS	3.52	21.66	50.73	21.35	4.22	19.18
PENS-NAML	**3.93**	**22.73**	50.16	22.81	4.90	19.65
FPG-SA	2.89	17.20	89.20	27.13	10.39	23.09
FPG-CNN	2.88	17.20	**89.29**	27.20	10.41	23.15
FPG-GRU	2.88	17.27	89.26	**27.33**	**10.51**	**23.30**

ICDM 2023 – 23rd IEEE International Conference on Data Mining

Case Study

Our model successfully combined *user interests* with *factual consistency*, creating personalized and *informative* headlines that *avoid potential clickbait*.

TABLE VI

A CASE OF PERSONALIZED NEWS HEADLINES GENERATED FROM THREE DIFFERENT MODELS. THE UNDERLINED WORDS INDICATE SOME CRITICAL FACTS OF THE ORIGINAL NEWS ARTICLE. THE WORDS IN RED FONT REPRESENT CONSISTENT KEYWORDS OF PERSONALIZED INTERESTS, AND THE WORDS IN BLUE FONT DENOTE FACTUAL ERRORS AND CORRESPONDING FACTS IN THE NEWS ARTICLE.

News Article	**Justin Rose** didn't just dominate Thursday afternoon's marquee pairing at the 2019 U.S. Open, he **tied a record** set by his more famous playing partner. With **an opening 65**, Rose matched **Tiger Woods' first-round score in 2000** for the lowest-ever **U.S. Open round at Pebble Beach**. The Englishman also grabbed a one-shot lead after one round thanks to three consecutive birdies to cap his day …
Historical Clicks	**US Open: Tiger Woods** finishes strong at Pebble Beach **Tiger Woods** fought back Sunday and had his best **U.S. Open** score in 10 years 2019 **U.S. Open Tiger** Tracker: **Woods** shoots second-round 72
Manually-written Headline	Justin Rose tied **Woods'** score of 65 in **U.S. Open** round at Pebble Beach
PENS-NAML	**Tiger Woods first round score**�’ U.S. Open
BART	Rose take the lead, ties **Tiger' Pebble Beach record**✔
FPG-GRU	Justin Rose ties **Tiger Woods' U.S. Open record**✔ with opening 65 at Pebble Beach

Thanks.

Fact-Preserved Personalized News Headline Generation
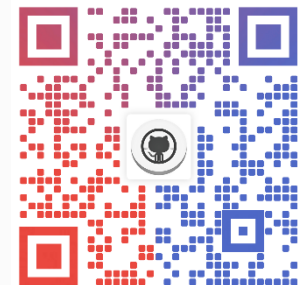
Zhao Yang*†‡, **Junhong Lian***†‡, Xiang Ao†‡§✉

† Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, CAS, Beijing 100190, China.
‡ University of Chinese Academy of Sciences, Beijing 100049, China.
§ Institute of Intelligent Computing Technology, Suzhou, CAS.

{yangzhao20s, lianjunhong23s, aoxiang}@ict.ac.cn

The source code will be ready soon at https://github.com/ictmldm/FPG

IEEE ICDM 2023, Shanghai, China.
Dec, 2023