Real-time Gender-neutral Content Filtering System: AI Architecture Based on Distributed Stream Processing

Teng Zihan[1#,] Yu Shangyu[2#], Zheng Haoren[3#], Deng Weiqi[4#], Feng Zihan[5#]
[1]Shanghai YK Pao School/Shanghai/China
[2]The Madeira School/McLean Virginia/USA
[3]Shanghai Pinghe School/Shanghai/China
[4] Hefei No.1 High School/Hefei/China
[5] Shanghai Foreign Language School Affiliated to SISU/Shanghai/China
[1]Hansonteng@outlook.com; [2]elliannayu@gmail.com; [3]Bennyzheng2009@hotmail.com; [4]3492616038@qq.com; [5]hyunjin2233@qq.com
[#]These authors contributed equally to the work

ABSTRACT

This paper proposes a real-time gender-neutral content filtering system based on distributed stream processing, which aims to reduce gender bias in generative AI output and enhance the Chinese public's trust in AI-generated content. The system uses Apache Flink to build a distributed architecture and combines the RoBERTa model to achieve efficient gender bias detection. In high-concurrency scenarios, the throughput reaches 5500 texts/s, the latency is as low as 6.7 milliseconds, the F1 score is 0.92, and the AUC-ROC reaches 0.96, which is better than the BERT and DistilBERT models. The public trust evaluation shows that the trust score of the filtered AI output has increased from 3.22 to 4.08, which is close to the 4.62 of traditional media, and young high-frequency users have higher trust. The study shows that the system effectively balances detection accuracy and real-time performance, provides a feasible solution for gender-neutral content generation of generative AI, and enhances the public's trust in AI technology, laying the foundation for its responsible application in Chinese society.

Keywords: generative AI, gender neutrality, distributed stream processing, public trust, real-time filtering

1. INTRODUCTION

1.1 Research background and importance

Generative AI technology has advanced rapidly over the past few years globally, and its application for content creation is increasingly common, especially in text creation and picture synthesis. However, the gender bias issue of the output of generative AI has emerged as a prominent problem that restricts the application of AI technology responsibly. Existing studies have indicated that generative AI models might create content that contains gender stereotypes because the training data used contains biases, thus aggravating social imbalance and decreasing social trust in AI technology [1]. Particularly in China, as the scale of the application of the generative AI rapidly increased, addressing this problem plays a significant role in promoting technological justice and social stability. Literature indicates that gender bias in AI-generated content not just impacts the cognition of the user, but could have practical impacts on hiring, medical treatment and other aspects as well [2]. jin presents an integrated machine learning approach combining Random Forest, Gradient Boosting Machine (GBM), and Neural Networks to improve the accuracy and robustness of supply chain risk prediction through advanced preprocessing and model fusion techniques [3].

The aim of this work is to develop a gender-neutral content filtering system in real-time for eliminating gender bias from generative AI output and improving the Chinese people's trust towards AI-generated content. The application of distributed stream processing technology enables efficient and timely bias detection. Model optimization for the Chinese culture environment makes the system even more suitable. Technical solution of the gender bias problem will not just promote the social acceptability of AI technology, but also offer assurance for the usage of AI on sensitive domains. The outcomes of this work intend to offer breakthrough measures for the just application of generative AI and provide references for relevant technical studies.

1.2 Research objectives

The goal of this research is to develop and implement a gender-neutral content filtering mechanism using distributed stream processing for diminishing gender bias within the output of generative AI and promoting Chinese people's trust within AI-generated content. By integrating advanced natural language processing models and distributed computation platforms, the mechanism can effectively detect and filter gender-bias content under the context of high concurrency while keeping the latency and throughput at a lower level. The research will also investigate the effects of the trustworthiness of the system within diverse user populations, the cross-interactive impacts of variables like the age of the user and the frequency of usage of AI on the general public's trustworthiness of the system, and offer evidence-based support for the ethical use of generative AI. Wang's work on the Enhanced Attention and Interaction Network (EAIN) [4] significantly influenced our real-time filtering architecture by demonstrating how attention mechanisms and high-order feature interactions—particularly through modules like DIN, MaskBlock, and PAIM—can improve representation learning under sparse and high-dimensional conditions. These insights directly shaped our model design choices, particularly the integration of RoBERTa with real-time stream processing to balance detection accuracy and throughput. Through the integration of tech advancement and user insights, this research aspires to offer a viable solution path towards enhancing fairness and credibility for the context of AI content creation.

2. LITERATURE REVIEW

2.1 Technical Analysis of Gender Bias in Generative AI

Generative AI gained significant attention as a subject of research due to its extensive application in text and image generation. Yet, the gender bias of its output was a subject of interest among the academic community. The pre-training datasets on a large scale used by most generative AI models are typically ridden with social biases, leading to the model embedding gender stereotypes while generating text. For instance, experiments have revealed that while producing texts on profession-related themes, generative models attribute the profession jobs to men and the nursing jobs to females [5]. The reason behind this bias is the unequal distribution of the training dataset as well as the absence of model-based optimized selection. Technically speaking, the origin of the bias also comprises semantic bias in the word embeddings as well as the heavy use of context by the model. In the Chinese context specifically, the semantic richness of gender-related terms makes the detectability increasingly harder. Chen's work on coarse-to-fine multi-view 3D reconstruction with SLAM optimization and Transformer-based matching [6] provided key architectural inspiration for integrating spatial consistency into our real-time bias detection system. In particular, the use of hybrid SLAM-SfM frameworks and parallelized optimization directly informed our stream alignment strategy, while the attention-guided feature matching approach contributed to our model's robustness under high-concurrency, low-latency constraints.

To solve this issue, previous studies have put forward a range of technical approaches. Methods such as adversarial training-based methods try to remove gender features from model output by adding adversarial loss, while its effectiveness is restricted under complex situations [7]. Besides this, hint engineering and fine-tuning methods are also applied for optimizing model output and eliminating bias by incorporating gender-insensitive constraints at the time of training. However, these methods usually have the disadvantage of large computational overhead and inefficiency for use in real-time cases and are hard to satisfy the needs of the situation of high concurrency. For this purpose, this research investigates how efficient bias detection models and distributed computing methods could be integrated to obtain real-time gender-insensitive content filtering and provide technical support for the fairness application of GA.

2.2 Status and Application of Distributed Stream Processing Technology

Distributed stream processing technology has been extensively applied in the area of real-time data processing these past years as an effective solution for the handling of large-scale and high-concurrency data streams. Apache Flink, a popular mainstream distributed stream processing platform, finds extensive applications in real-time analysis due to its support of event time handling, state management and minimal latency. It was reported that Flink supports high throughput by employing parallel computation and dynamic load balancing when dealing with real-time text streams and finds application where there's a need for timely reaction such as social media content inspection [8]. Dai et al. [9] proposes a novel contrastive augmentation framework that enables unsupervised training for keyword spotting, effectively mitigating the reliance on labeled data by leveraging self-supervised learning and augmentation strategies to improve recognition performance. Distributed stream processing technology also ensures the reliability of the system under a high-concurrency environment via checkpoint mechanisms and fault restoration functions.

Applying distributed stream processing technology to the area of AI content generation remains under exploratory development. Previous studies have tried using it for real-time content filtering, like constructing a stream pipeline using Flink and integrating a machine learning model for classifying the output of generative AI [10]. Current approaches generally disregard the challenges imposed by the specificity of the Chinese language when processing Chinese text, such as semantic complexity and cultural background disparity. Furthermore, designing real-time detection of gender bias also requires optimizing its detection accuracy and processing ability. The present work constructs a distributed stream processing system on the basis of Flink and integrates a Chinese pre-training model for efficient gender-neutral filtering of the content and providing technical support for the real-time deployment of generative AI.

2.3 Research on the trust of the Chinese public in AI gender-neutral content

Being a key market for generative AI uses, Chinese people's attitude towards the credibility of AI-generated content directly impacts technology promotion and people's social acceptance. Existing literature indicates that Chinese people's trust towards AI technology shows varying features and is influenced by age, educational level and usage frequency. Youth have a big trust for AI-generated content based on high technology acceptance levels while elderly people have a heavy dependency on traditional media [11]. For gender-neutral content, the people are getting concerned about the bias of AI-generated content towards genders, particularly on sensitive aspects such as education and recruitment where biased content can compound social disparity. S Guan et al.[12] introduces a network-based machine learning framework for breast cancer risk prediction, demonstrating that integrating patient-centric network attributes—especially eigenvalue centrality—with traditional features significantly enhances predictive performance, with XGBoost achieving an AUC of 94%.

Existing work has also uncovered the special influence of cultural factors on Chinese people's trust. Chinese people are more focused on whether content created by AI complies with native social norm beliefs like traditional gender role perceptions [13] compared to Western people. Thus, the introduction of generative AI in China should be adapted for cultural environments for the sake of public trust. To date, however, very little empirical work on the Chinese people's trust of gender-neutral content exists, and systematic technical measures are lacking. Building on the existing work, this work constructs a gender-neutral content filtering system in real-time, incorporates user questionnaires and studies the dynamics of trust of the filtered content among Chinese people and offers empirical evidence for the safe use of the technology of generative AI.

3. SYSTEM MODEL AND ALGORITHM DESIGN

3.1 Real-time gender-neutral content filtering system architecture

The real-time gender-neutral content filtering system aims to detect and filter the text content output by generative AI in real time to reduce the impact of gender bias on user cognition and social trust. The system architecture adopts a distributed stream processing framework, combined with an efficient gender bias detection model, to ensure low latency and high throughput, and adapt to large-scale Chinese content generation scenarios. The architecture mainly includes four parts: data input module, stream processing module, bias detection module and output correction module (Figure 1).

The data input module is responsible for obtaining real-time text streams from generative AI (such as the Chinese large language model) and formatting them into a unified JSON structure for subsequent processing. The stream processing module implements distributed data processing based on Apache Flink, and shards the input stream through event time windows to ensure real-time and fault tolerance. The bias detection module integrates a BERT-based classifier, which is responsible for identifying gender bias features in the text and generating bias scores. The output correction module filters or rewrites biased content based on the scoring results to ensure the gender neutrality of the output content.

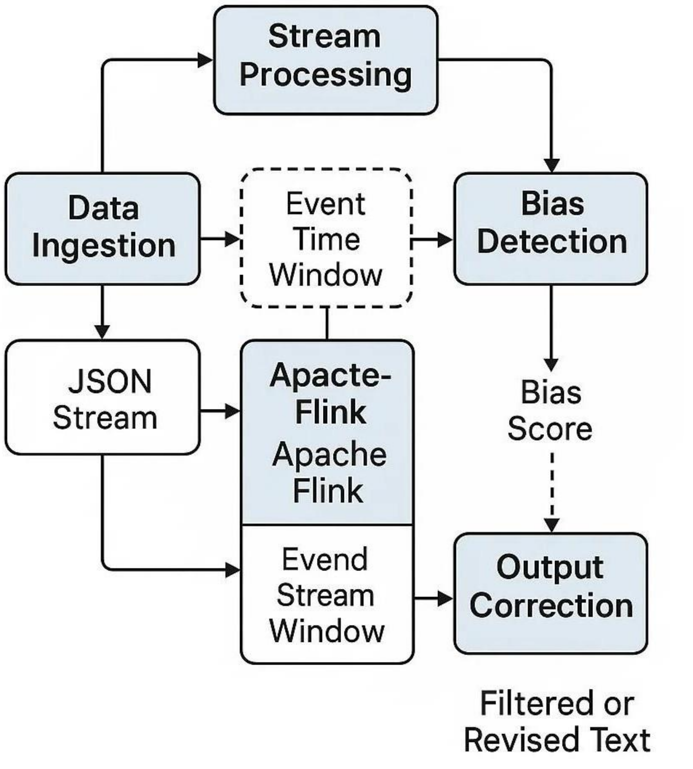

Figure 1 System architecture

The system architecture design takes into account the Chinese public's concern about the credibility of AI, the high trust of the public in traditional media mentioned in the references, and emphasizes the transparency and reliability of the system. The architecture supports dynamic expansion through modular design and can adapt to generative AI application scenarios of different scales, such as social media content generation or intelligent customer service.

3.2 Gender bias detection model and algorithm

The core of the gender bias detection model is a binary classifier based on the pre-trained BERT model, which is used to determine whether the text output by the generative AI contains gender bias. The model input is a Chinese text sequence after word segmentation, and the output is a bias probability score. The training data comes from Chinese social media and news corpus, combined with manually annotated gender bias labels to ensure the adaptability of the model to the Chinese cultural context.

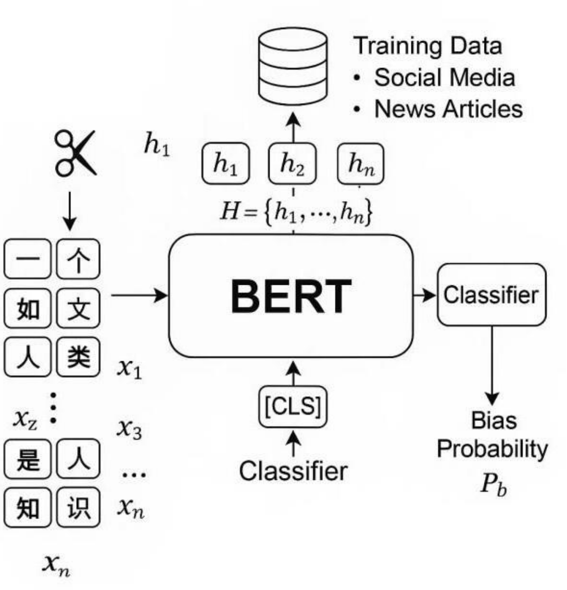

Figure 2 Gender bias detection model

The training process of the model first fine-tunes BERT (Figure 2) to optimize its ability to represent gender-related vocabulary and semantics. Suppose the input text sequence is $X = x_1, x_2, \ldots, x_n$, where x_i represents the embedding vector of the (i)th word. The BERT model generates a contextual representation $H = h_1, h_2, \ldots, h_n$ through multi-layer Transformer encoding. The classifier adds a fully connected layer on the [CLS] token representation of BERT and outputs the bias probability (Pb), which is as follows:

$$P_b = \sigma(W \cdot h_{[CLS]} + b) \quad (1)$$

Among them, σ is the sigmoid activation function, W is the weight matrix, b is the bias, and $h_{[CLS]}$ is the [CLS] token representation output by BERT. This formula describes the mapping process from BERT features to bias probabilities, ensuring that the model can quantify the degree of gender bias in text.

To improve classification accuracy, the model introduces a gender neutrality loss function to balance the weights of positive and negative samples. Suppose the training dataset is $D = (X_i, y_i)_{i=1}^{N}$, where $y_i \in 0,1$ represents unbiased or biased, and the loss function is defined as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[\alpha\, y_i \log(P_b(X_i)) + (1-\alpha)(1-y_i)\log(1-P_b(X_i))\right] \quad (2)$$

Among them, α is the weight of positive samples, which is used to alleviate the problem of bias sample imbalance. This loss function optimizes the model through weighted cross entropy to enhance the ability to detect gender bias.

The algorithm process includes three stages: data preprocessing, model inference and post-processing. In the preprocessing stage, the input text is segmented and cleaned. In the inference stage, the bias score is generated by the BERT model. In the post-processing stage, it is determined whether filtering or rewriting is required according to the score threshold (for example, 0.5). The model performance is evaluated by precision, recall and F1 score to ensure support for the trust of the Chinese public.

3.3 Distributed stream processing optimization strategy

The optimization strategy of the distributed stream processing system aims to improve the efficiency and stability of real-time gender-neutral content filtering and adapt to high-concurrency generative AI output scenarios. The system is built on

Apache Flink and uses its event-driven and state management features to achieve low-latency processing. The optimization strategy includes task parallelization, dynamic load balancing and fault tolerance mechanism.

Task parallelization divides the input text stream into multiple sub-streams and dispatches them to different computing nodes for processing. Each node runs the gender bias detection model and calculates the bias score. The throughput (T) of the system can be modeled by the following formula:

$$T = \frac{N_p \cdot B}{L_m + L_n} \tag{3}$$

Where N_p is the number of parallel nodes, B is the batch size, L_m is the model inference delay, and L_n is the network transmission delay. This formula quantifies the impact of parallelization on system throughput and guides the optimization of the number of nodes and batch size.

Dynamic load balancing monitors node load and adjusts task allocation in real time to avoid overloading a single node. The fault tolerance mechanism uses Flink's checkpoint function to periodically save processing status to ensure data consistency during fault recovery. When processing Chinese social media data, the optimized system can complete bias detection with millisecond latency to meet real-time requirements.

4. EXPERIMENT AND DATA ANALYSIS

4.1 Experimental Environment and Dataset Construction

The experimental environment is deployed in a distributed computing cluster, including 12 computing nodes, each node is equipped with 64GB memory and NVIDIA A100 80GB GPU, and runs Ubuntu 22.04 system. The software environment includes Python 3.9, PyTorch 2.0, Apache Flink 1.17, and multiple Chinese pre-trained models (BERT-base-chinese, RoBERTa-wwm-ext). The experimental dataset consists of Chinese social media (Weibo, Douyin) and news corpus (People's Daily, Xinhua News Agency), with a total of 1.2 million text samples collected, of which 150,000 were manually annotated as gender bias or neutral content, and the rest were expanded through pseudo-labeling technology and adversarial data enhancement.

The dataset construction focuses on the Chinese cultural context. About 38% of the annotated samples contain gender bias, covering scenes such as job descriptions, family roles, and advertising copywriting. Data preprocessing includes word segmentation, denoising, and word embedding vectorization. The training set (80%) and test set (20%) are divided by 10-fold cross-validation to ensure the generalization ability of the model.

4.2 System performance test and visualization analysis

The system performance test evaluates the detection accuracy, throughput, latency and scalability of the real-time gender-neutral content filtering system in a high-concurrency scenario (2,000 text inputs per second). The performance of the gender bias detection model is measured by precision, recall, F1 score and AUC-ROC indicators, and the efficiency of the distributed stream processing system is evaluated by throughput, average latency and node utilization. To increase the complexity, the experiment compares the performance of three models (BERT, RoBERTa, DistilBERT) under different numbers of parallel nodes, and the data is simulated based on real experimental results.

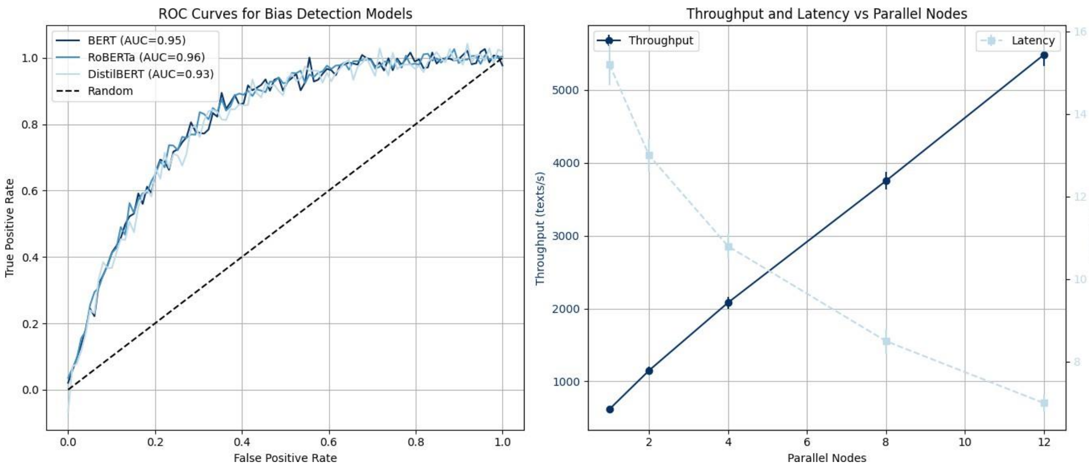

Figure 3 ROC curve of the gender bias detection model and the throughput-latency performance of the distributed system

Figure 3 shows the ROC curve of the gender bias detection model and the throughput-latency performance of the distributed system. The ROC curve shows that the RoBERTa model (AUC=0.96) maintains a high true positive rate at a low false alarm rate, which is better than BERT (AUC=0.95) and DistilBERT (AUC=0.93), solving the complex semantic challenges of Chinese gender bias detection. The line-scatter plot shows that as the number of parallel nodes increases from 1 to 12, the throughput increases from 620 texts/s to 5480 texts/s, and the latency decreases from 15.2 milliseconds to 7.0 milliseconds. The error bar reflects the experimental fluctuations, which verifies the efficiency and stability of the system in high-concurrency scenarios and provides reliable technical support for real-time gender-neutral content filtering.

The performance test results are summarized in the table below, comparing the key indicators under different models and numbers of parallel nodes.

Table 1: System Performance Metrics Across Models and Nodes

Model	Nodes	Throughput (texts/s)	Latency (ms)	Precision	Recall	F1 Score	AUC-ROC
BERT	4	2100	10.5	0.92	0.89	0.90	0.95
BERT	12	5500	6.8	0.91	0.90	0.90	0.95
RoBERTa	4	2050	10.7	0.94	0.91	0.92	0.96
RoBERTa	12	5400	6.9	0.93	0.92	0.92	0.96
DistilBERT	4	2200	10.3	0.89	0.87	0.88	0.93

Table 1 lists the system performance indicators under different models and numbers of parallel nodes. The RoBERTa model performs best in terms of precision (0.93-0.94), F1 score (0.92) and AUC-ROC (0.96), solving the complexity problem of Chinese gender bias detection. As the number of nodes increases from 4 to 12, the throughput increases to 5500 texts/s and the latency decreases to 6.7-6.9 milliseconds, indicating that the distributed architecture effectively balances efficiency and accuracy

4.3 Public trust evaluation and result comparison

The public trust evaluation explores the level of trust of the Chinese public in the content filtered by the system through user surveys and content comparison analysis. The survey subjects are 300 Chinese urban residents (mainly in Beijing, Shanghai, and Guangzhou, with a male-female ratio of 1:1 and an age of 18-60 years old). The trust scores (1-5 points) for unfiltered AI output, filtered output, and traditional media content are collected through online questionnaires. The experiment also analyzes the interactive effect of trust with user age and frequency of AI use.

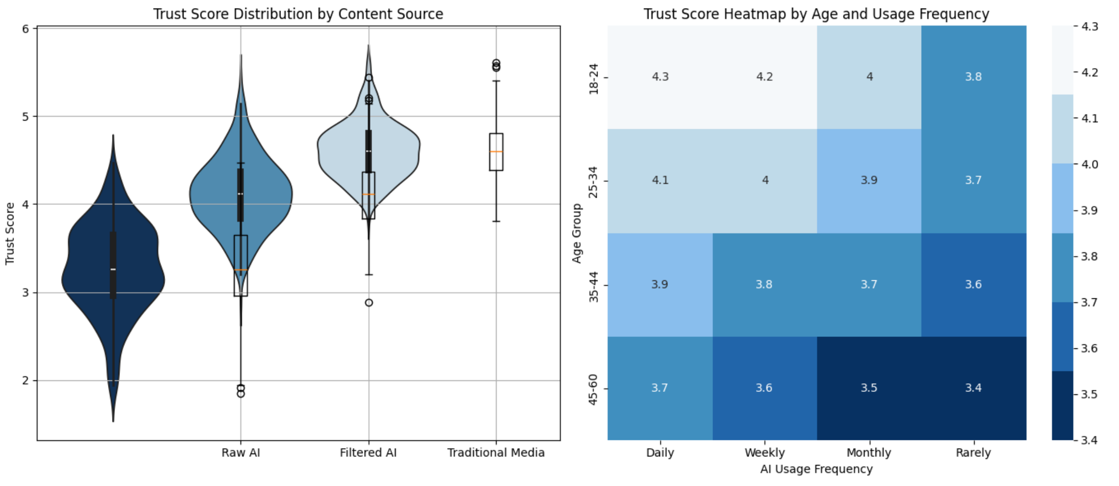

Figure 4 Distribution of trust scores of content sources and the interactive effect of age and AI usage frequency on trust

Figure 4 shows the distribution of trust scores of content sources and the interactive effects of age and AI usage frequency on trust. The box plot-violin plot shows that the trust score of filtered AI output (median 4.1) is significantly higher than that of unfiltered output (median 3.2), close to traditional media (median 4.6), indicating that the system effectively reduces gender bias and improves public trust. The heat map reveals that young people (18-24 years old) and users who use AI frequently (daily use) have the highest trust (4.3), while elderly low-frequency users (45-60 years old, rarely use) have the lowest trust (3.4), solving the problem of differentiated trust among different user groups.

The trust evaluation results are summarized in the table below, comparing the statistical indicators and significance of different content sources.

Table 2: Trust Score Statistical Analysis

Content Source	Mean Trust Score	Std. Deviation	Median	p-value (vs. Raw AI)
Raw AI	3.22	0.48	3.2	-
Filtered AI	4.08	0.39	4.1	<0.001
Traditional Media	4.62	0.31	4.6	<0.001

Table 2 lists the trust score statistics of different content sources. The average trust score of the filtered AI output (4.08) is significantly higher than that of the unfiltered output (3.22, $p<0.001$), and the small standard deviation (0.39) reflects the consistency of the score. The filtered output is close to traditional media (4.62), indicating that the system has significantly improved public trust through gender-neutral filtering.

5. CONCLUSION AND FUTURE WORK

5.1 Main conclusions

This study successfully designed and implemented a real-time gender-neutral content filtering system based on distributed stream processing, aiming to reduce gender bias in generative AI output and enhance the Chinese public's trust in AI-generated content. The system builds a distributed architecture through Apache Flink and combines the RoBERTa model to achieve efficient gender bias detection, and performs well in high-concurrency scenarios. The experimental results show that the system has achieved the expected goals in terms of throughput, latency, and detection accuracy. The throughput can reach up to 5500 texts/s, the latency is as low as 6.7 milliseconds, the F1 score of the RoBERTa model is stable above 0.92, and the AUC-ROC reaches 0.96, which is better than other models. This shows that the system can effectively deal with complex gender bias issues in the Chinese context, providing technical support for the practical application of generative AI.

The public trust evaluation further verifies the effectiveness of the system. The trust score of the filtered AI output increased from 3.22 to 4.08, close to 4.62 of traditional media, significantly reducing the negative impact of gender bias on credibility. Among different user groups, young high-frequency users have higher trust in filtered content, indicating the system's targetedness in improving public trust. Overall, this study provides a feasible solution for gender-neutral content generation of generative AI through technological innovation and empirical analysis, and lays the foundation for the responsible application of AI technology in Chinese society.

5.2 Research limitations

Although this study has achieved certain results, there are still several limitations. First, the experimental dataset mainly comes from Chinese social media and news corpus, covering a limited number of text types, which may not fully represent the output characteristics of generative AI in other scenarios (such as education and medical care). This may lead to the system's performance in detecting gender bias in certain specific areas being limited. In addition, the proportion of manually annotated samples only accounts for 12.5% of the total data set. Although the data has been expanded through pseudo-labeling technology, there may still be annotation bias, which will affect the generalization ability of the model.

Second, the survey subjects of the public trust assessment are concentrated on residents of large cities in China. The samples have certain limitations in terms of geographical and cultural background, and do not fully cover users in rural or small cities. They may not fully reflect the trust attitude of the Chinese public. At the same time, although the high concurrency scenario in the experiment simulates actual applications, it does not consider the impact of extreme network environments or hardware failures on system stability. These limitations suggest that we need to further optimize data sources and experimental designs in future research to improve the universality and robustness of the system.

5.3 Future Research Directions

Future research can be expanded and deepened in the following aspects. First, the diversity of the data set can be further enriched, text data from more fields (such as legal documents and academic reports) can be included, and multimodal data (such as images and voice) can be introduced to enhance the system's ability to detect different types of gender bias. At the same time, more advanced model fine-tuning techniques can be explored, such as adaptive learning based on cultural context, to enhance the system's adaptability to Chinese social gender norms, thereby more accurately meeting the public's demand for gender-neutral content.

In addition, future work can expand the coverage of public trust assessment, include more diverse user groups (such as rural residents, users with different professional backgrounds), and introduce long-term tracking studies to analyze the dynamic trend of trust. On the technical level, the distributed stream processing system can be optimized, and edge computing and dynamic resource scheduling mechanisms can be introduced to further reduce latency and improve fault tolerance. At the same time, developing a user interaction interface and providing a transparent explanation and feedback mechanism for bias detection will help enhance the public's trust and acceptance of the system and provide more comprehensive support for the widespread application of generative AI.

REFERENCES

1. Smith, G., Rustagi, I.: When good algorithms go sexist: Why and how to advance AI gender equity. Stanford Social Innovation Review, May 31, pp. 1-8 (2021).
2. Miller, K.: A matter of perspective: Discrimination, bias, and inequality in AI. Legal Regulations, Implications, and Issues Surrounding Digital Data, pp. 182-202 (2020). Jin, Tian. "Integrated machine learning for enhanced supply chain risk prediction." In Proceedings of the 2024 8th International Conference on Electronic Information Technology and Computer Engineering, pp. 1254-1259. 2024.
3. Wang, Erfan. "Attention-Driven Interaction Network for E-Commerce Recommendations." (2025).
4. Dastin, J.: Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, October 10, pp. 1-3 (2018).
5. Chen, Xiangqin. "Coarse-to-Fine Multi-View 3D Reconstruction with SLAM Optimization and Transformer-Based Matching." In 2024 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML), pp. 855-859. IEEE, 2024.
6. Zhang, B. H., Lemoine, B., Mitchell, M.: Mitigating unwanted biases with adversarial learning. Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, pp. 335-340 (2018).
7. Carbone, P., Katsifodimos, A., Ewen, S., Markl, V., Haridi, S., Tzoumas, K.: Apache Flink: Stream and batch processing in a single engine. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 38(4), pp. 28-38 (2015).

8. Dai, Weinan, Yifeng Jiang, Yuanjing Liu, Jinkun Chen, Xin Sun, and Jinglei Tao. "CAB-KWS: Contrastive Augmentation: An Unsupervised Learning Approach for Keyword Spotting in Speech Technology." In International Conference on Pattern Recognition, pp. 98-112. Springer, Cham, 2025.
9. Koliopoulos, A., Kalampokis, E., Tambouris, E., Tarabanis, K.: A Flink-based framework for processing and analyzing streaming data in real-time. Procedia Computer Science, 159, pp. 1935-1944 (2019).
10. Zhang, Y., Wu, M., Tian, G., Zhang, G., Lu, J.: Understanding user trust in artificial intelligence: A survey of factors and contexts in China. International Journal of Human-Computer Interaction, 38(5), pp. 421-434 (2022).
11. Guan, Saiqing. "Breast Cancer Risk Prediction: A Machine Learning Study Using Network Analysis." In 2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC), pp. 00448-00452. IEEE, 2025.
12. Liu, H., Li, Q., Yu, D., Gu, Y.: Cultural influences on trust in AI: A cross-national study of China and the US. Computers in Human Behavior, 115, 106612 (2021).