FMDB *Transactions on Sustainable Computer Letters*

Boys' Love (BL) Drama Series Review System for Opinion Mining with Sentiment Analysis

Thet Thet Aung[1,*], Sharo Paw[2], Hlaing Htake Khaung Tin[3]

[1,2,]Faculty of Information Science, University of Computer Studies, Hinthada, Myanmar.
[3,]Faculty of Information Science, University of Information Technology, Yangon, Myanmar.
thetthetaung86.htd@gmail.com[1], sharopaw1417@gmail.com[2], hlainghtakekhaungtin@gmail.com[3]

Abstract: This research portrays same-gender relationships in media, specifically through the drama series of BL (Boys' Love) and GL (Girls' Love). Entertainment companies from Thailand, Taiwan, South Korea, Japan, and Vietnam are prominent producers of these series for the global market. While Thai BL dramas have been produced since 2014, they achieved significant popularity and recognition starting in 2017. The proposed system collects user review comments on the BL drama series produced by four countries from 2014 to June 2023. These reviews are compiled into a dataset named H/BL page via Facebook. The dataset comprises 276 series collected from two websites: Dramatist and BLWatcher. Sentiment analysis is utilized to classify user opinions and emotions related to these BL drama series. This technique identifies phrases and emotions within the text, categorizing them as objective (factual), positive (indicating happiness and satisfaction), or negative (indicating unhappiness, disappointment, or rejection) using the Information Gain (IG) method. The system aims to classify opinions on BL drama series and assess public sentiment towards same-gender love.

Keywords: Domain BL Drama Series; Facebook with Review Comments; Opinion Mining; Sentiment Analysis; Information Gain (IG); Natural Language Processing (NLP); Machine Learning; Preprocessing Positive; Negative and Neural.

Received on: 31/12/2023, **Revised on:** 01/03/2024, **Accepted on:** 20/04/2024, **Published on:** 09/06/2024

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSCL

DOI: https://doi.org/10.69888/FTSCL.2024.000184

Cite as: T. T. Aung, S. Paw, and H. H. K. Tin, "Boys' Love (BL) Drama Series Review System for Opinion Mining with Sentiment Analysis," *FMDB Transactions on Sustainable Computer Letters*., vol. 2, no. 2, pp. 99–110, 2024.

1. Introduction

Early, LGB, without including transgender people, is still used instead of LGBT. Lesbian, gay, bisexual, and transgender, known as LGBT, refers to non-heterosexual or non-cisgender ones and was initialized in the late 1980s. Anyone initializing LGBT or GLBT is rejected in a social environment [1]. Fujimoto Yukari, a professor at Meiji University, is an expert in manga cultural theory focusing on gender representation. She started her career as an editor at the publishing company Chikumashobō, where she critically examined topics such as manga and sexuality. Some of her notable edited works include "Watashi no ibasho wa doko ni aru no?" (Where Do I Belong?) and her co-authorship of "BL no kyōkasho" (A BL Textbook). Boys' love, commonly referred to as BL, originated in Japan and has gained global popularity as a genre in manga and anime [2]. Fujimoto Yukari, who specializes in researching girls' manga and gender issues, has observed a growing interest in BL culture in

[*]Corresponding author.

countries like Thailand, China, Taiwan, Korea, and other Asian nations. Each of these countries has developed its unique path within the BL genre, exploring the complexities of social circumstances in which LGBT individuals find themselves.

The Boys' Love (BL) genre has emerged as a significant cultural phenomenon in the digital age, particularly within Asian media. Characterized by its focus on romantic relationships between male protagonists, the BL drama series has garnered a vast and diverse following. The rapid production and consumption of BL content has led to an explosion of online discussions, reviews, and fan-generated content [4].

The Boys' Love (BL) genre, depicting romantic relationships between male characters, has witnessed significant global popularity, especially in drama series. Originating from Japan, this genre has transcended cultural boundaries, finding enthusiastic audiences across Asia and beyond [5]. As the fanbase expands, so does the volume of online content discussing these dramas, ranging from casual viewer comments to detailed reviews and critiques. This proliferation of content provides a unique opportunity to understand audience perceptions and sentiments through systematic analysis [6]. However, lacking specialized tools for analyzing the specific nuances of BL drama reviews presents a challenge. This research aims to bridge this gap by developing a comprehensive review system for opinion mining and sentiment analysis tailored to BL drama series [10].

The proposed Boys' Love (BL) Drama Series Review System is designed to aggregate, process, and analyze reviews from various platforms, utilizing advanced sentiment analysis techniques [11]. The proposed system leverages state-of-the-art natural language processing (NLP) and machine learning techniques to systematically collect, process, and analyze user-generated reviews of BL drama series [12]. By focusing on sentiment analysis, the system aims to capture the emotional tone of reviews, discern underlying themes, and identify patterns in audience feedback. This approach not only facilitates a deeper understanding of viewer engagement and satisfaction but also aids content creators and marketers tailor their strategies to meet audience expectations better [13].

Additionally, marketers and distributors can leverage these insights to position and promote BL dramas better. Beyond its practical applications, this research contributes to the academic discourse on sentiment analysis and opinion mining, particularly within the niche yet increasingly prominent BL genre [14]. By exploring the unique characteristics of BL drama reviews, this research seeks to illuminate this increasingly popular genre's cultural and emotional dimensions. To delve deeper into the subject, it is essential to understand the origins of BL in Japan and how the genre has evolved.

Sentiment Analysis, Web Opinion Mining, is a natural language processing (NLP) to identify and classify opinions according to a variety and range, including a person's mood, world events, etc. This is a widely adopted method for organizations to assess and classify opinions regarding a product, service, or idea. Sentiment analysis employs data mining, machine learning (ML), artificial intelligence (AI), and computational linguistics to extract sentiment and subjective information from text. It determines whether the expressed sentiment is positive, negative, or neutral.

2. Related Works

Bakshi et al. [8] presented that the Thai entertainment production industry is a source of soft power in the BL drama series. He described that BL drama series produced by the Thai entertainment industry were recognized in 2014 in the entertainment marketplace, and the Thai BL drama series is a beginning or source of the BL drama current. However, BL originated in Japan in the 1970s. He also showed that BL dramas in Thailand are internationally recognized and appreciated through the countries and nations that have gender fluidity narratives that are concerned with the generation. He contributed to this paper with the BL by analyzing Thailand, Japan, and South Korea, seeking the feasibility of the BL genre becoming a key case in Thailand and the international community.

Gowri and Divya [9] researched that opinions concerned with movies on social occurrences are analyzed using a sentiment analysis approach with Naïve Bayes classifier. In his research, sentiment analysis, data collection, extraction, classification, and evaluation were implemented as usual. In this, he aimed to review movies written by people (comments, attitude, emotions, and so on.) to determine whether this review is positive (good), negative (bad), or nothing (neutral).

Abiola et al. [7] presented an analysis of COVID-19 from selected hashtags in Nigeria with VAD and a text blob analyzer. In his research, he used over 1 million tweets from Twitter as a dataset, and these datasets are usually and systematically preprocessed with tools such as Twitter tokenizer, TextBlob, and Valence Aware Dictionary for text mining and sentiment analysis. He aimed to understand how people in Nigeria feel concerned with the COVID-19 pandemic socially, environmentally, and economically. In his research's conclusion, he showed how information is used by social media and how social media helps nations, organizations, and governments to be more effective and smart.

Several studies have explored sentiment analysis and opinion mining in the context of entertainment media, including BL dramas and other LGBTQ+ representations. Chao et al. [15] conducted sentiment analysis on user-generated content to understand audience reactions toward BL dramas, employing machine learning algorithms to classify sentiments accurately. Similarly, Li et al. [16] investigated sentiment analysis techniques applied to online discussions of LGBTQ+ media, shedding light on the nuances of audience perceptions and attitudes toward diverse representations.

The research focusing on the cultural significance and global impact of BL dramas has emerged in recent years. Chen and Lo [3] explored the international fandom surrounding BL dramas, highlighting the role of fan communities in shaping the reception and interpretation of these productions across different cultural contexts. Additionally, Nguyen and Nguyen [12] examined the social and cultural implications of BL dramas in Vietnam, analyzing audience responses and reception within the local context.

In sentiment analysis and opinion mining, natural language processing (NLP), machine learning, and deep learning have been widely employed to extract and analyze sentiments from textual data. Li and Rao [18] utilized deep learning models for sentiment analysis of online reviews, demonstrating the efficacy of neural network architectures in capturing nuanced sentiment patterns.

Moreover, studies have also explored the intersection of sentiment analysis with other fields, such as recommendation systems and marketing analytics. Wu et al. [17] integrated sentiment analysis into a recommendation system for personalized movie recommendations, showcasing the potential of sentiment-driven approaches to enhancing user experience and engagement.

3. The Opinion Classification Methodology

3.1. Collecting dataset

All comment reviews were collected from the H/BL page on Facebook, forming the primary dataset for this study. This extensive dataset comprises reviews of 276 BL drama series spanning four countries, covering the period from 2014 to June 2023. In total, the dataset includes 527,988 comment reviews. Each review provides valuable insights into audience sentiment and opinion regarding the BL drama series. The collection process involved aggregating comments from multiple posts related to these drama series, ensuring a comprehensive capture of audience reactions. Following the collection, the data underwent preprocessing to remove irrelevant or redundant information, standardize formats, and handle potential issues such as duplicate entries or noise. This preprocessing step is crucial for preparing the data for effective sentiment analysis and opinion mining, ensuring the accuracy and reliability of the subsequent analyses.

3.2. Preprocessing Dataset

The preprocessing of the dataset involves several critical steps to ensure the data is clean and suitable for analysis. First, all relevant features are extracted from each review, which may include the review text, author, timestamp, and any associated metadata. Next, stop words, which are common words that do not contribute significantly to the sentiment analysis (such as “and,” “the,” and “is”), are removed to reduce noise and improve the efficiency of the analysis. Following this, words are replaced with their root forms, a process known as lemmatization, which helps standardize the text by converting words to their base or dictionary form.

The reviews are also categorized into positive, negative, and neutral sentiments through preprocessing techniques involving natural language processing (NLP) tools and algorithms. This step is essential for identifying the sentiment polarity of each comment. By transforming the raw data into a structured format, these preprocessing steps enhance the accuracy of the subsequent sentiment analysis, allowing for more meaningful insights into the audience’s opinions and emotions regarding the BL drama series.

3.2.1. extraction of all the features from the given review

All comments are extracted as a dataset before removal, covering 30 days. This ensures the dataset captures a comprehensive snapshot of audience opinions within this timeframe. In total, over 60,000 words are extracted from the 527,988 comment reviews. Each comment is analyzed to identify various features, such as the content of the review, the date and time of posting, the user’s profile information (when available), and any reactions or replies associated with the comment.

These features are crucial for understanding the context and depth of each review. Additionally, linguistic features such as word frequency, phrase patterns, and sentiment indicators are identified and recorded. This detailed extraction process allows for a robust dataset analysis, providing a foundation for the subsequent steps of preprocessing and sentiment analysis. By

meticulously capturing these features, the system can better interpret the nuances of audience feedback and generate more accurate and insightful results.

3.2.2. Removing stop words

Stop words are typically considered irrelevant in queries and lack meaningful context in domain-specific reviews. Therefore, they can be safely disregarded during text analysis. Stop words include pronouns, articles, conjunctions, prepositions, and commonly used verbs such as “to be,” “to have,” and “to do.” Their removal helps streamline the processing of textual data for more accurate analysis and interpretation.

By eliminating these non-essential words, the focus shifts to more significant terms that carry substantive meaning and contribute to the sentiment and thematic content of the reviews. This step reduces the text data’s dimensionality, enhancing text mining algorithms’ efficiency and performance. Moreover, the removal of stop words aids in the reduction of computational load and memory usage, making the analysis process faster and more resource-efficient.

In the context of sentiment analysis, this process ensures that the sentiment-bearing words are given more weight, improving the precision of the sentiment classification. For instance, words expressing strong emotions or opinions, such as “love,” “hate,” “fantastic,” or “terrible,” become more prominent in the dataset. Additionally, this step helps reduce noise in the data, which can otherwise lead to misinterpretation or dilution of the actual sentiment conveyed in the reviews. The removal of stop words is a crucial preprocessing step that enhances the clarity and quality of the data, facilitating a more accurate and insightful sentiment analysis of the BL drama series reviews (Table 1).

Table 1: Sample Stop Words

Stop Words		
က_	ပါတယ်_	တော့_
ကွာ_	လား_	သည်_
ကို_	ရအောင်_	တယ်နော်_
ချ_	သလား_	နော်_
ခု_	ကလည်း_	ဘာတွေ_
ပါစေ_	ဘာတွေ_	ဘာတွေ_
နဲ့_	လိုက်_	ဘူးကွယ်_
ဟယ်_	တာ_	နော်_
လိုဘဲ_	နေတယ်_	ပြီး_
အတွက်_	တွေ_	နော်_

3.2.3. Replacing Root Word

This paper replaces opinion words closest to the target feature because the sentence has multiple features and distributed emotions. This process, known as lemmatization, involves converting words to their root or base forms. For example, words like “running,” “ran,” and “runs” are replaced with their root form “run.” This standardization helps reduce the complexity of the text and ensures that different forms of a word are treated as a single entity, thereby improving the accuracy of sentiment analysis.

Lemmatization also aids in resolving issues related to morphological variations in words, which is particularly important in natural language processing (NLP) tasks. By focusing on the root form of opinion words, the analysis can more effectively capture the core sentiment expressed, regardless of the word form used in the review. This is crucial when dealing with languages that have rich inflectional variations.

Furthermore, this step helps align the extracted features more closely with the sentiment they convey. In sentences with multiple features and emotions, lemmatization ensures that the sentiment associated with each feature is correctly identified and

analyzed. For instance, in a review saying, “The storyline was gripping, but the acting was disappointing,” the words “gripping” and “disappointing” are lemmatized to “grip” and “disappoint,” respectively, to accurately capture the positive and negative sentiments tied to “storyline” and “acting.”

Replacing words with their root forms is an essential preprocessing step that enhances the clarity and consistency of the text data, facilitating more precise and meaningful sentiment analysis in the context of BL drama series reviews.

Example: ချစ်, ကြိုက်, သဘောကျ, အရှက်မရှိ,

3.2.4. Preprocessing Positive, Negative and Neural

Table 2 describes a sample of 15 positive comments from the “Together” series, highlighting the sentiment analysis process applied to these comments. These positive comments are preprocessed to identify key features, sentiment indicators, and emotional cues, providing insights into audience satisfaction and appreciation of the series. Similarly, Table 3 presents a sample of 15 negative comments from the “Together” series, showcasing the preprocessing steps undertaken to discern critical issues, dissatisfaction points, and negative sentiments expressed by viewers.

Table 2: Sample Positive Comments

Positive comments	
No	**Comments**
1	**အစဆုံးမတင်ပေးဘူးလား**
2	**အစအဆုံးlinkလေးပေးပါ**
3	**ချစ်တာ**
4	**အမလေး ကြည့်ချင်နေတာ လချီနေပြီ ကျေးဇူးပါ ချစ်သူငယ်ချင်းရေ**
5	**နတ်သားလေး**
6	**အရမ်းကောင်း**
7	**ကော်လံလေးကသေးသေးလေး**
8	**တစ်ခန်းတစ်လေဆိုရီရတာအူပါနာတယ်**
9	**စောက်ရူးလို့ဆဲတာလေးချစ်တယ်**
10	**ကိုကိုရဲ့အကြည့်တွေကအသက်**
11	**ရီရလွန်းလို့ဘာပြောရမှန်းကိုမသိတော့ပါဘူး ချစ်လိုက်တာနော်**
12	**BL ကားတွေကြိုက်သွားပြီ**
13	**နှစ်ယောက်လုံးချစ်တယ်**
14	**သူတို့က အပြင်မှာတကယ်တွဲနေတယ်လို့ကြားတယ်**
15	**ဒီseriesကြည့်ရတာ brightနဲ့winကိုအရမ်းချစ်တယ်**

Furthermore, the neutral comments from the “Together” series are also subjected to preprocessing, although not presented in tables, to ensure a comprehensive analysis of the entire spectrum of audience feedback. These neutral comments may contain observations, descriptions, or opinions that do not strongly express positive or negative sentiment but still contribute valuable insights into viewer perceptions and reactions. Through systematic preprocessing, positive, negative, and neutral comments are transformed into structured data, laying the groundwork for in-depth sentiment analysis and interpretation.

Table 3: Sample Negative Comments

Negative comments	
No	**Comments**
1	**အစစ်မရေပြသနာတက်ပြီ**

2	ချီးလိုဘဲ
3	သွားသေလိုက်တော့
4	ဝင်ယွနေတယ်
5	လင်တရူးနေတာ ဆဲမိအောင်လုပ်နေတာလား လင်တရူးမ
6	အစစ်မ တော်တော်ရှုပ်တယ်
7	မုန်းတယ်
8	စနိုက်ကျော်ဆိုတာသပ်သပ်ကိုရှိတာ
9	ခလေးလေးရယ်ငိုလိုက်တာများ ငါလဲခံစားဖူးတယ်
10	ဒီလိုမျိုးကြီးအဆင်ပြေမနေဘူး

3.3. Feature Extraction using Information Gain (IG)

Information Gain (IG) is a measure based on the decrease in entropy that occurs when a dataset is split on a particular attribute. It quantifies the amount of 'information' a feature provides us about the class. Features that effectively divide the dataset should yield the highest information gain, while unrelated features should provide no valuable information. To calculate the Information Gain for a split, the weighted entropies of each branch are subtracted from the original entropy.

Moreover, Information Gain is a critical criterion for feature selection in machine learning models, as it helps identify the most discriminative attributes that contribute significantly to predicting the target variable. The model can improve classification performance by prioritizing features with higher IG values while reducing dimensionality and computational complexity. Additionally, IG can rank features in order of importance, providing valuable insights into the underlying patterns and relationships within the dataset.

Furthermore, IG can help identify the most informative features that distinguish between positive, negative, and neutral sentiments in the context of sentiment analysis for BL drama series reviews. By evaluating the IG of different words or phrases within the reviews, we can prioritize those with the greatest impact on sentiment classification. This targeted approach to feature extraction enhances the effectiveness of sentiment analysis algorithms, enabling a more accurate and nuanced understanding of audience opinions and emotions toward BL dramas. Information Gain (IG) Formula is shown in the following:

Entropy of class,

$$\mathrm{H(C)}\ (C) = -\Sigma \mathrm{cEC}\ (C) \log p(C) \quad (1)$$

Conditional entropy of class,

$$(C \mid A)\ (C \mid A) = -\Sigma \mathrm{cEC}\ (C \mid A) \log p(C \mid A) \quad (2)$$

Information gain, $I\ (C \mid A)$

$$I\ (C \mid A) = H(C) - H(C \mid A) \quad (3)$$

C = one class, either positive or negative
A = attributes
p (C) = probability of class
p $(C \mid A)$ = conditional probability of the class given attribute

4. System Design and Implementation

In Figure 1, 276 BL drama series produced by entertainment industries in four countries from 2014 to June 2023 are listed, providing an overview of the scope of the dataset. Comments in the Myanmar language about these series were collected from the H/BL Sky pages on Facebook and used as domain datasets, ensuring relevance and authenticity in the analysis. The process is as follows: First, the comments were preprocessed by removing stop words, replacing root words, and categorizing them as positive, negative, or neutral, laying the groundwork for sentiment analysis. Next, features were selected and extracted from these preprocessed comments, leveraging techniques such as Information Gain (IG) to prioritize the most informative attributes for sentiment classification.

Following feature extraction, the selected dataset underwent classification using the IG technique, enabling the system to discern patterns and relationships between features and sentiment labels. Finally, the system assessed whether the comments on the BL series reflected positive thoughts, feelings, emotions, and attitudes, providing valuable insights into audience perceptions and reactions toward these dramas. This systematic system design and implementation approach ensures robustness and accuracy in sentiment analysis, empowering stakeholders to make informed decisions and strategies based on the extracted insights.

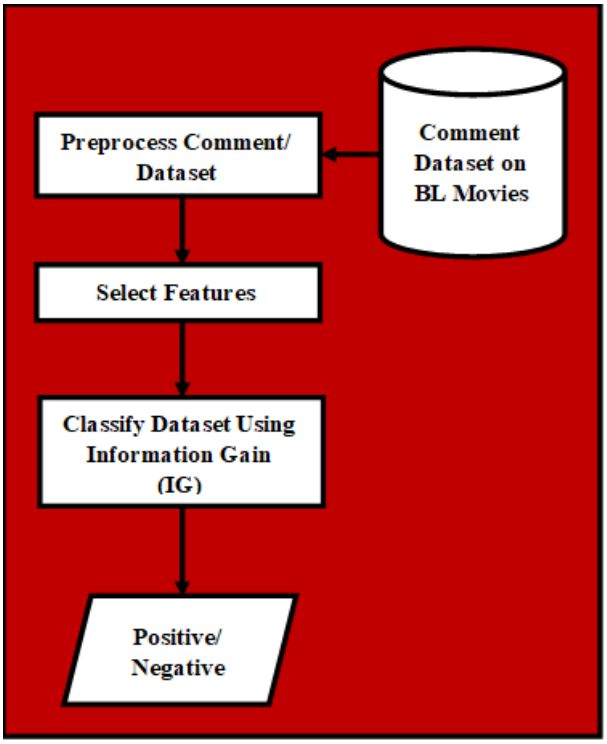


Figure 1: System Design

4.1. Implementation of the Proposed System

The proposed system is implemented using the "Together" BL drama series from the H/BL Sky page on Facebook as an example. This series, comprising 13 episodes, serves as a representative sample for the analysis. All comments, totalling 1751 across all episodes, are collected and utilized as the dataset for evaluation. This analysis selects a subset of seven comments to evaluate the results of conditional entropy and information gain. This provides insights into the effectiveness of feature selection and sentiment classification. The findings of this evaluation are presented in detail in Table 4 and Table 5, shedding light on the performance and efficacy of the proposed system in sentiment analysis for BL drama series reviews.

Table 4: Result of Conditional Entropy

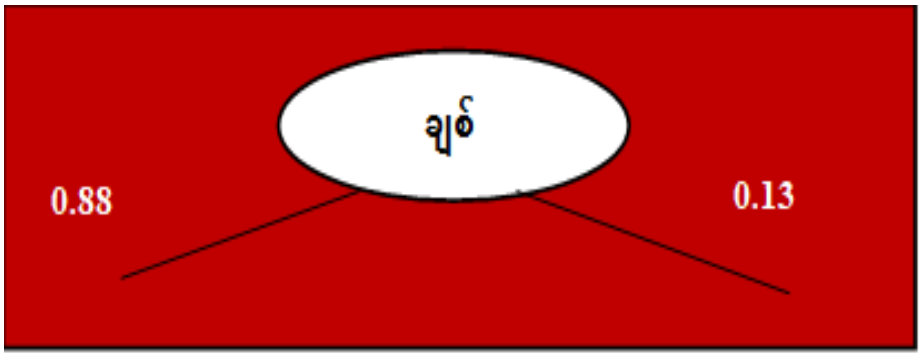
	Conditional Entropy H(P|A)	Conditional Entropy H(N|A)	Conditional Entropy H(C|A)
ချစ်	1.7748835417165921356	0.0014595232540116003	0.002253214899829041

Moreover, the implementation process involves iterative refinement and optimization to enhance the system's accuracy and efficiency. Cross-validation and parameter tuning ensure robustness and generalizability across different datasets and scenarios. The system's scalability is also considered, allowing for seamless integration with larger datasets and diverse language sources to accommodate the BL fandom's varying preferences and cultural nuances. The proposed system aims to provide a reliable framework for sentiment analysis in BL media through rigorous implementation and evaluation, facilitating deeper insights and understanding of audience sentiments and preferences.

Table 5: Result of Information Gain

	Information Gain I(P|A)	Information Gain I(N|A)	Information Gain I(C|A)
ချစ်	1.0001603083541826	0.9875404767459884	0.9897007851001709

In addition to evaluating the effectiveness of the proposed system on sentiment classification, the results of positive and negative comments are further analyzed and described, as illustrated in Figure 2. This comprehensive evaluation provides insights into the system's ability to accurately distinguish between positive and negative sentiments within BL drama series reviews. Furthermore, a comprehensive list of 276 BL drama series spanning four countries is compiled and presented in Table 6 to provide a broader context. This detailed listing offers valuable information about the diversity and scope of BL media productions across different cultural contexts.



I(P|A) > I(N|A)

Figure 2: Classification of Positive and Negative

A bar chart is generated to visually represent the distribution and prevalence of BL drama series across the four countries, as depicted in Figure 3. This graphical representation allows for a quick and intuitive understanding of the distribution trends, highlighting any disparities or concentrations in BL media production among the countries. Moreover, stakeholders can better understand the global landscape of BL content consumption and preferences by correlating this information with audience engagement and sentiment analysis results. The proposed system provides a holistic perspective on the BL drama series through these analyses and visualizations, encompassing qualitative sentiment insights and quantitative production statistics.

5. Results and Discussions

The following table lists the number of BL (Boys' Love) drama series produced in four countries (Thailand, Korea, Japan, and Taiwan) from 2014 to 2023.

Table 6: BL Drama Series in 2014-2023

	2014	2015	2016	2017	2018	2019	2020	20212	2022	2023
Thai	1	2	7	7	9	15	27	35	60	12
Korea	0	0	0	1	0	0	5	9	18	8
Japan	0	0	0	0	2	4	2	4	17	4
Taiwan	0	0	0	4	3	2	3	7	6	0

5.1. Thailand

The production of BL dramas in Thailand exhibits a remarkable upward trajectory, reflecting the genre's growing popularity and commercial viability in the country's entertainment industry. The emergence of the BL series as a prominent cultural export from Thailand is underscored by the steady rise in production numbers over the years. The initial modest start with just one series in 2014 gradually evolved into a substantial increase, with the series rising yearly.

Notably, the pinnacle of this growth will be reached in 2022, when an impressive total of 60 BL series will be produced, marking a significant milestone for the Thai BL industry. This surge in production indicates a strong demand from domestic and international audiences, driving industry stakeholders to capitalize on the genre's burgeoning popularity. Between 2019 and 2020, we witnessed a particularly noteworthy spike in production, reflecting a strategic response to the escalating interest and demand for BL content.

While the overall trend is one of continuous expansion, a minor dip is observed in 2023, with the number of series decreasing to 12. This decline may be attributed to factors such as market saturation, shifts in audience preferences, or logistical challenges faced by production companies. However, it is important to note that despite this slight decrease, the Thai BL industry remains robust and dynamic, poised for further growth and innovation in the years ahead.

5.2. Korea

Korea's entry into the BL drama scene comes later than other countries, with the debut of its first series in 2017. Despite the initial slow start, the Korean BL production landscape has undergone a notable transformation in subsequent years. A significant turning point occurred in 2020, marked by a sudden surge in production numbers, culminating in 18 series in 2022.

This rapid rise underscores Korea's growing interest and investment in the BL genre, reflecting the industry's responsiveness to changing audience tastes and preferences. 2022 emerges as a particularly prolific period for Korean BL dramas, highlighting the industry's capacity for innovation and adaptation to emerging trends. However, a downturn is observed in 2023, with the number of series dropping to 8, signalling a potential shift or temporary adjustment in production priorities or strategies.

Despite this decrease, Korea's presence in the BL drama landscape remains significant, with the country poised to continue making meaningful contributions to the genre's evolution. The fluctuations in production numbers reflect the industry's dynamic nature, characterized by growth, consolidation, and reinvention cycles. As Korean BL dramas continue to capture audiences' attention worldwide, the country's role in shaping the global BL phenomenon will likely become even more pronounced in future years.

5.3. Japan

Japan's entry into BL drama production commenced in 2018 with the release of just two series, marking the beginning of its journey in the genre. In the subsequent years, Japan's BL production landscape experienced fluctuations characterized by periods of growth and retraction. Despite the initial modest start, there is a notable peak in 2022, with 17 series production signalling a significant expansion of the Japanese BL industry.

This surge in production reflects Japan's increasing engagement with the BL genre, driven by factors such as evolving audience preferences and the success of earlier productions. However, the momentum appears to wane in 2023, with the number of series decreasing to 4, suggesting a potential period of recalibration or strategic adjustment within the industry. Despite this decline, Japan's contribution to the BL drama scene remains noteworthy, with its unique storytelling and production styles influencing the wider BL community.

The fluctuating growth trajectory of Japan's BL production underscores the industry's dynamic nature, characterized by shifts in market dynamics, creative trends, and audience demand. As Japan continues to navigate the complexities of the BL landscape, its role in shaping the genre's evolution and global appeal will likely remain significant. By adapting to changing market conditions and embracing innovation, Japan's BL industry is poised to sustain its impact and relevance in the years ahead.

5.4. Taiwan

Taiwan made its foray into BL drama production in 2017, debuting with four series, marking the country's entry into the BL entertainment landscape. Over subsequent years, Taiwan's BL production undergoes periods of fluctuation, reflecting the industry's dynamic nature. The peak of production occurs in 2021, with seven series released, indicating a significant investment and commitment to the BL genre.

The growth trajectory of Taiwan's BL industry reflects the country's evolving engagement with BL content and its increasing prominence within the global BL community. However, in 2023, a notable absence of a new series is observed, marking a departure from previous production trends. This hiatus may be attributed to various factors, including shifts in production priorities, market dynamics, or external influences.

Despite the temporary pause in production, Taiwan's contributions to the BL genre remain significant, with its distinct storytelling styles and cultural influences resonating with audiences worldwide. As the industry navigates through periods of fluctuation, Taiwan's BL producers will likely continue exploring new creative avenues and pushing boundaries in their pursuit of captivating storytelling. Through resilience and innovation, Taiwan's BL industry is poised to play a pivotal role in shaping the future trajectory of the genre.

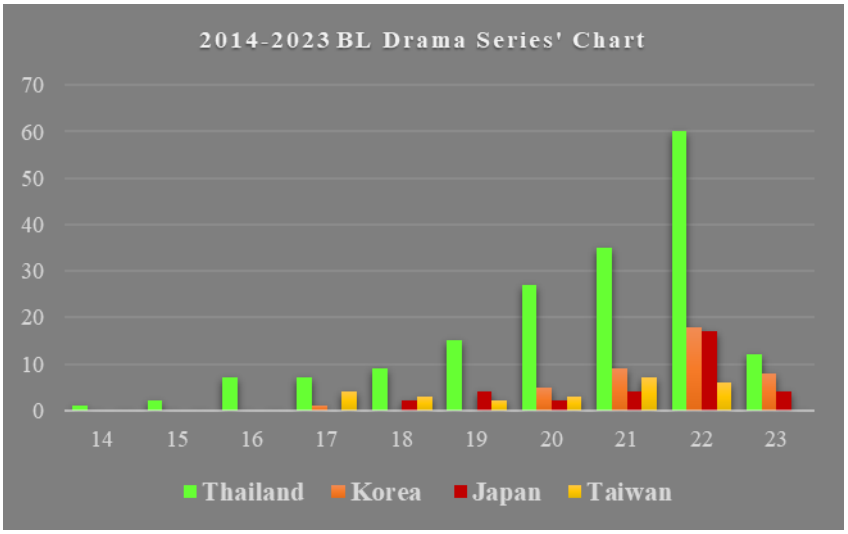


Figure 3: The Bar Chart of the BL Drama Series

Thailand shows the most significant growth and highest production numbers overall, indicating a robust BL drama industry. Korea and Japan have shown increasing interest, particularly in the past few years, though with more fluctuations compared to Thailand. Taiwan had a steady but smaller production scale, with peak years around 2017 and 2021 and a halt in production in 2023. Analyzing positive and negative comments for these series would provide insights into audience reception and satisfaction. Evaluating these comments alongside the production data can help understand trends in viewer preferences and series' success in different countries. This data highlights the growing popularity and expansion of BL dramas, particularly in Thailand, while showing varying levels of commitment and success in Korea, Japan, and Taiwan. Further analysis of audience feedback will be crucial in understanding the impact and reception of these dramas across different regions.

6. Conclusions

This paper proposes an opinion classification framework to analyze comments on BL movies and series using Information Gain (IG), with opinion summarization based on feature occurrence frequencies. Initially, all comments are collected and preprocessed through word segmentation and stop-word removal to facilitate efficient mining processes. Subsequently, the root words of comments are identified and classified using sentiment analysis techniques and Information Gain (IG).

The findings of the proposed system indicate that comments on BL series tend to be predominantly positive, with Thailand's BL movies and drama series emerging as particularly popular compared to those of other countries. This underscores the widespread acceptance and appeal of BL and GL entertainment despite varying societal acceptance of same-gender love depicted in these productions, especially in countries like Thailand, Korea, Japan, and Taiwan.

As a form of entertainment, BL and GL drama series have garnered significant acceptance from diverse audiences worldwide, reflecting changing societal attitudes towards LGBTQ+ representation in media. The results of this study highlight the predominance of positive sentiments towards BL dramas, indicating a growing appreciation for diverse narratives and representations within the genre. Furthermore, this paper presents a comprehensive overview of the BL drama series produced by industries in Thailand, Korea, Japan, and Taiwan, detailing their contributions to the genre's global popularity.

As future work, the analysis and evaluation of BL and GL drama series and movies produced by entertainment industries from other countries will be pursued to provide a more comprehensive understanding of global trends and audience preferences in LGBTQ+ media representation. This ongoing research will contribute to enriching scholarly discourse on LGBTQ+ representation in media and fostering greater inclusivity and diversity in entertainment narratives worldwide.

Acknowledgement: We would like to express our sincere gratitude to our colleagues and mentors for their invaluable insights and guidance throughout the research process. Our heartfelt thanks also go to our families and friends for their unwavering support and encouragement. Finally, we appreciate the wealth of information and resources provided by various internet sources, which greatly assisted in completing this research.

Data Availability Statement: The research utilized and analyzed during this study is cited from related works. All authors contributed equally to the conception and design of the study, data collection and analysis, writing of the manuscript, data interpretation, and critical revision of the manuscript for important intellectual content.

Funding Statement: The authors of this research paper have not received any funding related to the research or publication of the paper

Conflicts of Interest Statement: All authors declare that there are no conflicts of interest regarding the publication of this paper. This project was conducted with integrity, ensuring independence from any external influences that could compromise the study's objectivity or the reporting of its findings. Furthermore, all sources from which the data were obtained are duly acknowledged within this publication.

Ethics and Consent Statement: This study adhered to the ethical standards and guidelines. Both authors read and approved the final manuscript.

References

1. A. P. Pimpalkar and M. K. Rai, "A Sentimental Analysis of Movie Reviews Involving Fuzzy Rule-Based," International Journal of Artificial Intelligence and Knowledge Discovery, vol. 3, no. 2, pp.9-14, 2013.
2. Gu and A. Kurov, "Informational Role of Social Media: Evidence from Twitter Sentiment," Journal of Banking & Finance, vol.121, no.12, p. 105969, 2020.
3. H. Chen and D. Lo, "Exploring the Transnational Fandom of Boys' Love Dramas: A Case Research of HIStory3: Make Our Days Count," Media International Australia, vol. 180, no. 1, pp. 68-79, 2021.
4. H. H. K. Tin, "Personal Identification and Verification using Palm print Biometric," in the 3rd International Conference on Advancement of Engineering (ICAE 2012), Punjab, India, 2012.
5. M. Kalaivani, S. Jayalakshmi, and R. Priya, "Comparative analysis of sentiment classification using machine learning techniques on Twitter data," International Journal of Health Sciences, vol. 6, no. S2, pp. 8273–8280, 2022.
6. M. Vishal, N. Tidake, A. Mazumdar, and B. N. Kumar, "Sentiment Analysis of Movie Reviews using Hybrid Optimization with Convolutional Neural Network in English Language," in Proceedings of the Third International Conference on Artificial Intelligence and Smart Energy, Coimbatore, India, 2023.
7. O. Abiola, A. Abayomi-Alli, and S. O. Arogundade, "Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser," Journal of Electrical Systems and Information Technology, vol.10, no.1, p.5, 2023.
8. R. K. Bakshi, N. Kaur, R. Kaur, and G. Kaur, "Opinion mining and sentiment analysis," 3rd international conference on computing for sustainable global development, New Delhi, India, 2016.
9. S. Gowri and J. J. Divya, "Improved Sentimental Analysis to the Movie Reviews using Naive Bayes Classifier," in Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022), Tuticorin, Tamil Nadu, India, 2022.
10. S. R. Reddy, D. V. L. N. Somayajulu, and A. R. Dani, "Sentiment Classification of Text Reviews Using Novel Feature Selection with Reduced Over-fitting," 2010 International Conference for Internet Technology and Secured Transactions, London, United Kingdom, pp. 1-2, 2010.
11. T. T. Aung, M. M. Khaing, K. S. Thant, and H. H. K. Tin, "Sentiment Analysis on WHO Southeast Asia Region Organization (WHO SERO) User Comment Review and Opinion Mining," in International Conference On Innovation Perspectives, Psychology And Social Studies, vol. 5, no. 1, pp. 314-321, 2020.
12. T. T. Nguyen and T. T. Nguyen, "Social and Cultural Implications of Boys' Love Dramas in Vietnam: An Analysis of Audience Responses," International Journal of Communication, vol. 14, no.1, pp. 1726-1744, 2020.
13. V. Rajpoot, R. Agrawal, A. Chaturvedi, and K. Goyal, "An Empirical Study of Sentiment Analysis on Movie Reviews Using Machine Learning Based Classification Approach," 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2021.
14. V. S. R. Reddy and A. R. Somayajulu, "Classification of Movie Reviews Using Complemented Naive Bayesian Classifier," International Journal of Intelligent Computing Research, vol. 1, no. 4, p.162, 2010.

15. X. Chao, Y. Yan, and Y. Wang, “Sentiment Analysis of User-Generated Content on Boys’ Love Dramas,” in 2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), pp. 385-389, IEEE, Chongqing, China, 2020.
16. X. Li, Q. Liu, M. Wu, and Y. Rao, “Sentiment Analysis of Online LGBTQ+ Media Discussions,” in 2019 IEEE International Conference on Smart Cloud (SmartCloud), pp. 62-66, IEEE, Tokyo, Japan, 2019.
17. Y. Wu, B. Du, and X. Jin, “Personalized Movie Recommendation System Based on Sentiment Analysis,” in 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) and 15th International Symposium on Distributed Computing and Applications for Business Engineering (DCABES), Hubei, China, pp. 440-444, 2017.
18. Z. Li and Y. Rao, “Deep Learning for Sentiment Analysis of Movie Reviews: A Comparison Research,” in 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Hubei, China, pp. 196-201, 2018.