M23 Attacks: Public Sentiments Alternation in Congolese Newspapers

Zeyu Han[1,*] John Quattrochi[2]

December 7, 2024

Abstract

The March 23 Movement (M23) rebel attack in November 2021 reignited disputation between the Democratic Republic of the Congo (DRC) and Rwanda, which was accused of sponsoring the hostile operations. Meanwhile, the Congolese public sentiments towards Rwanda changed rapidly as a manifestation of protest which was reflected in Congolese newspapers to a certain extent. This paper endeavors to leverage advanced large language models (LLMs) to perform sentiment analysis on various mainstream newspapers to investigate sentiments trends and emotion distributions to systematically understand sentiments shifts during the M23 rebel attack. We utilized BERT-based multilingual sentiment model to analyze the alternation of sentiment trends and degrees, and DistilRoBERTa sentiment model to identify specific emotions in journalistic articles. The sentiments shifts was corresponding to our initial hypothesis that state-run media would change sentiment the most to increase the support for the state. Additionally, we applied Term Frequency-Inverse Document Frequency (TF-IDF) and Latent Semantic Analysis (LSA) to discover significant terms and topics in Congolese newspapers. The prominence of "M23," "Goma," and "Rwanda" in the newspapers highlights the frequent association of the M23 attacks with Rwanda in public discourse. Our project page including code and data is available at here.

Keywords: public opinion; sentiment analysis; media expression; national inclination

1 Introduction

In 2012, M23, primarily composed of ethnic Tutsi people, launched an armed rebellion against the government of the DRC. The conflict triggered a humanitarian crisis such as slaughter and refugee problems until a peace agreement was eventually signed in 2013 with military intervention by the United Nations (UN) peacekeeping force [1, 2, 3]. However, M23 violated the agreement and reactivated the armed conflict with the DRC government in November 2021 [4], and UN reports [5, 6] suggested that Rwanda was potentially supporting M23 by providing weapons and supplies to illegally disseizin Congo's natural resources [2].

As neighboring countries, the relationships between the DRC and Rwanda have long been in a delicate state of balance. On the one hand, despite sharing a common cultural heritage and traditions, the Tutsi and Hutu populations in both countries have a complex and intertwined history marked by conflict and tension, culminating in the 1994 Rwandan Genocide [7]. This prolonged conflict has caused devastating consequences, fueling armed conflicts, economic decline, and refugee crisis that has inflicted serious national antagonism on both sides. On the other hand, Rwanda and the DRC maintain substantial import and export trades reaching $487.2M in 2022 [8]. The significant economic interactions between two countries promoted the progress of cultural exchange and strengthen the bonds between the two nations. Consequently, the 2021 M23 rebel attack severely strained relations between the two countries and profoundly impacted Congolese public opinion and sentiments towards Rwanda.

It is challenging to explicitly survey the public opinions of Congolese due to armed conflict, poor transportation infrastructure, low communication capacity, and its large geographic extent. Although Twitter(X) serves as an effective platform for public expression after some critical social events occurred [9], lack of equipment and internet access hinders its ability to fully capture public sentiments comprehensively. Conversely, newspapers could offer valuable insights into the shifts in public sentiments towards Rwanda following the M23 rebel attack due to its widespread and preciseness of information.

This project aims to leverage internet-based text and natural language processing to systematically analyze how sentiments expressed in Congolese newspapers towards Rwanda changed before and after the November 2021 M23 rebel attacks. Our contribution could be summarized as follows:

- Collect articles from various media sources for a comprehensive exploration of Congolese public sentiments.
- Utilize diverse pre-trained LLMs to perform sentiment analysis to acquire Congolese holistic sentiments trend and specific emotions towards Rwanda.
- Compare state-run media to UN-affiliated media to private media to explore sentiment changes in distribution and degrees.

2 Related Works

Recently, LLMs have emerged as effective tools in natural language processing (NLP), achieving state-of-the-art performance in tasks such as understanding context, sentiment analysis, and text generation. For instance, BERT series [10] developed by Google and LLaMA[11] series developed by Meta have exhibited extraordinary performance across various downstream tasks. Furthermore, fine-tuned versions of these models are proving valuable disciplines beyond NLP, such as sociology and political science, offering researchers innovative methods for text analysis with accurate results [12]. These advanced LLMs could be deployed effectively to analyze the sentiments expressed in newspapers, providing nuanced insights into public opinion and sentiments.

Public opinion often intertwines with public sentiment, and journalistic sentiment expressions during domestic conflicts are frequently associated with two key emotions: love and fear [13]. These expressions and appeals could provide crucial perspectives to the political situation surrounding significant events, enabling researchers to acquire better understanding of public attitudes and concerns. Furthermore, the armed conflicts tends to evoke widespread insecurity and patriotism which exert a profound impact on government policy making through the "rally 'round the flag" effect [14]. This effect often stimulates public demand for tough government to restore peace and stability, accompanied by corresponding expressions of anger. In addition, considering the tendency for individuals to seek information aligned with their political preferences [15], the high-follower newspapers on Twitter (X) could partially represent the Congolese public sentiments which avoid the inevitable challenges associated with conducting surveys locally.

3 Dataset

We collected our datasets in three steps: (i) Obtaining valid official sources of newspapers. (ii) Retrieving related articles and announcements according to key words search such as "M23". (iii) Extracting articles from newspaper websites.

3.1 Data Collection

Sources Acquisition. To obtain credible sources of newspapers, we initially utilized ”Republic of Congo media guide” [16] to identify the recognized Congolese government media outlets. Then we took United Nations official website as reference to figure out corresponding trustworthy UN affiliated medias of Congo. Finally, we included private media outlets with the most Twitter followers which could reflect Congolese public sentiments on certain topics.

Articles Retrieval. To compile relevant textual data on M23 rebel attacks and its impact, we conducted keyword searches by using ”M23 Rebels” and ”M23 Attacks in 2021”. Subsequently, We manually reviewed the resulting website links to ensure their relevance and suitability for our analysis.

Title: Est de la RDC: à l'ONU, Kinshasa dit attendre du Conseil de sécurité plus de "fermeté" à l'égard de tous les pays qui ont des combattants et supplétifs sur son sol | Actualite.cd

Article Text:

L'ambassadeur Zénon Mukongo, représentant permanent de la RDC auprès des Nations-Unies, a réaffirmé l'engagement de son pays à poursuivre la mise en œuvre du Programme de désarmement, démobilisation, relèvement communautaire et stabilisation (P-DDRCS).Dans son intervention lundi 30 septembre devant les membres du conseil de sécurité de l'ONU, il a indiqué que ce programme est la clé de la stratégie nationale pour désarmer, démobiliser et réintégrer les combattants en leur offrant des perspectives économiques viables et durables tout en stabilisant les zones touchées par les conflits. Occasion pour lui d'appeler l'ONU de demander aux Etats voisins de récupérer également leurs combattants en RDC à la base de l'insécurité dans sa partie orientale." Le PDDRCS est un levier crucial pour la paix en RDC et nous invitons nos partenaires et amis à le soutenir. C'est ici le lieu de féliciter l'initiative du groupe des pays des amis du DDR qui tiendra une réunion spéciale sur la RDC demain mardi 1er octobre 2024 au siège des Nations-Unies. Je ne saurais clore ce chapitre sans insister sur la nécessité absolue du rapatriement des combattants étrangers qui ont trop longtemps oecumé l'Est de mon pays. À ce propos, ma délégation attend du Conseil plus de fermeté à l'égard de tous les pays qui ont des combattants et supplétifs sur le territoire congolais pour qu'ils puissent initier des dialogues sincères avec leurs ressortissants combattants qui doivent retourner dans leurs pays d'origine ", a plaidé l'ambassadeur Zénon Mukongo.Le diplomate congolais a rappelé la nécessité pour les partenaires internationaux et amis de la RDC de soutenir et d'accompagner la mise en œuvre de ce programme." Mon gouvernement est conscient que le programme de désarmement, démobilisation, relèvement communautaire et stabilisation reste un outil essentiel pour la protection des civils et l'instauration d'une paix et d'une stabilité durable en RDC. Cependant, il reste d'avis qu'au tant que la participation active du gouvernement et sa vision claire de la voie à suivre reste essentielle pour garantir les soutiens durables des partenaires internationaux à ce programme crucial, autant l'accompagnement de la Monusco reste déterminant pour l'opérationnalisation du PDDRCS. À ce propos, mon gouvernement estime qu'il y a des zones où l'on peut avancer très rapidement comme en Ituri ou le processus de réconciliation est en cours. En gros, le programme peut être lancé là où c'est faisable ", a lancé le diplomate congolais.La fusion du programme de DDR et celui de Stabilisation et Reconstruction des zones sortant des conflits armés (STAREC), ont donné naissance au programme DDRCS. Cette fusion vise non seulement plus de cohérence, mais aussi l'efficacité dans la réinsertion des démobilisés « vers des activités économiques et d'intérêt public, loin du métier des armes. »Ce programme était attendu depuis longtemps pour permettre la prise en charge des combattants qui quittent le maquis. Depuis sa mise en place et le changement des animateurs à sa tête, ce programme éprouve d'énormes difficultés pour décoller et contribuer réellement au désarmement, démobilisation, relèvement communautaire des combattants issus des groupes armés.Clément MUAMBA

Figure 1: This article comes from Actualite which represents the attitude of Congolese private newspaper to rebel army and their appeals for other countries.

Data Extraction. To extract the articles from newspaper webpages, we utilized the readability[1] package to gather the main content of the journals and Beautifulsoup[2] package to remove unnecessary URLs, advertisements and tags. As demonstrated in Figure 1, this process effectively reduced data noise for further analysis. Since the DRC is a French-speaking region, we employed the Google Translate API[3] to English versions for its compatibility with the English-based LLMs. Then, we utilized the datetime[4] package to extract the publication year of each article, enabling time-based visualization of the data.

3.2 Dataset Description

Our datasets contains three sub-datasets depending on the newspaper types. The sub-datasets comprise 48 articles from and UN media, and 300 articles from private media which is related to the M23 rebel attacks from 2020 to 2024.

Datasets Composition. Our datasets of Congolese newspapers are composed by three parts: (1) State-run Media: Radio of Television National Congolais (RTNC) represents government attitude for M23 rebel attack and Rwanda. (2) UN-run Media: Radio Okapi represents international outlooks to frictions between the DRC and Rwanda. (3) Private newspapers with the largest twitter followers: Actualite and 7SUR7 partially reflect the

[1] Readability: https://pypi.org/project/readability/
[2] BeautifulSoup: https://pypi.org/project/beautifulsoup4/
[3] https://pypi.org/project/googletrans/
[4] https://docs.python.org/3/library/datetime.html

Congolese public sentiment tendency.

Timeline Division. In order to observe the alternation of public sentiments caused by 2021 M23 rebel attack, we divided the time of datasets into three distinct periods. The first period, covering 2020 to 2021, served as the baseline to allow us to observe the trends in Congolese public sentiments to Rwanda. The second period, from 2021 to 2022, was a critical interval which revealed the potential surges and radical alternations of different sentiments. The third period, from 2023 to 2024, represents the long-term effect of M23 rebel attack to Congolese public sentiments towards Rwanda.

Variables

- **Source:** Webpage links and references of articles.
- **Year:** Published year of articles.
- **Articles:** Textual content extracted from the source links, which is crucial to analyze Congolese public sentiments and attitudes towards the M23 and Rwanda.
- **sentiment:** The sentiment of the articles, determined through sentiment analysis conducted by LLMs.
- **degree:** The corresponding strength or intensity of the sentiment.
- **Translated Articles:** English articles translated from French to fit English-based LLM.

Our curated dataset is optimized to leverage LLMs by providing relevant and structured text for accurate sentiment analysis. Furthermore, variables such as "Year" allow us to examine differences in public sentiments across various media outlets and to visualize changes in sentiment distribution and intensity over time.

3.3 Potential Limitation

While our dataset is curated to optimize sentiment analysis with LLMs, it has certain limitations. The datasets may not capture the comprehensive public sentiments due to the limited number of media outlets included, and some articles may inevitably contain biases which potentially influenced the sentiment analysis results [17]. Additionally, the translated content might introduce translation inaccuracies, which could impact sentiment analysis results.

4 Methodology

In this section, we describe two NLP techniques applied in textual analysis: algorithm-based feature extraction and sentiment analysis through LLMs. In Section 4.1, we introduced *Term-Frequency Inverse Document Frequency* (TF-IDF) to extract and quantify significant words in newspaper texts and *Latent Semantic Analysis* (LSA) to identify potential topics on english-version datasets from Section 3.2. Then, we leveraged the MULTILINGUAL BERT[18] and DISTILROBERTA[19] sentiment models in Section 4.2 to complete emotional classification to observe the sentiments alternation of Congolese newspapers.

4.1 Feature Extraction

TF-IDF is an effective method for identifying significant terms by comparing their frequency in a specific document to their frequency across the entire corpus. It combines two metrics: Term Frequency (TF) and Inverse Document Frequency (IDF) to assign a TF-IDF score to each term. As shown in the following equations:

$$TF(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d} \quad (1)$$

$$IDF(t, D) = \log \frac{\text{Total number of documents in the corpus } |D|}{\text{Number of documents containing term } t} \quad (2)$$

$$TF\text{-}IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (3)$$

The Equation 1 calculates the frequency of a term within a single document relative to the total number of terms in that document. Subsequently, we used Equation 2 to compute the Inverse Document Frequency, which measures how unique or rare a term is across the entire corpus. Finally, the multiplication of TF and IDF provides the TF-IDF score for each term as shown in Equation 3. Suppose there exists two documents, "I like apple" and "I love orange", the terms "like" and "apple" are important in Document 1, while "love" and "orange" are significant in Document 2, as demonstrated in Table 1.

Term	TF (Doc 1)	TF (Doc 2)	IDF	TF-IDF (Doc 1)	TF-IDF (Doc 2)
I	$\frac{1}{3} = 0.33$	$\frac{1}{3} = 0.33$	$\log \frac{2}{2} = 0$	$0.33 \times 0 = 0$	$0.33 \times 0 = 0$
like	$\frac{1}{3} = 0.33$	$\frac{0}{3} = 0$	$\log \frac{2}{1} = 0.69$	$0.33 \times 0.69 = 0.23$	$0 \times 0.69 = 0$
apple	$\frac{1}{3} = 0.33$	$\frac{0}{3} = 0$	$\log \frac{2}{1} = 0.69$	$0.33 \times 0.69 = 0.23$	$0 \times 0.69 = 0$
love	$\frac{0}{3} = 0$	$\frac{1}{3} = 0.33$	$\log \frac{2}{1} = 0.69$	$0 \times 0.69 = 0$	$0.33 \times 0.69 = 0.23$
orange	$\frac{0}{3} = 0$	$\frac{1}{3} = 0.33$	$\log \frac{2}{1} = 0.69$	$0 \times 0.69 = 0$	$0.33 \times 0.69 = 0.23$

Table 1: TF-IDF calculation example, the outputs of each cell is in range {0,1}.

LSA is an effective topic modeling technique which utilizes Singular Value Decomposition (SVD) in Equation 4 to factorize the TF-IDF matrix into three submatrices: Term-topic matrix (U), Diagonal matrix of singular values (Σ) indicating importance of topics, and Document-topic matrix (V). On the one hand, it identifies clusters of terms corresponding to latent topics by reducing the dimensionality and noise in the text. On the other hand, it can handle synonyms and polysemy, making it more effective at identifying accurate topic terms.

$$A = U\Sigma V^T \quad (4)$$

These two efficient techniques helped us explore the public concerns and focal points of Congolese newspapers beyond the keyword 'M23.' Additionally, they provided insights into identifying the words commonly associated with specific topics in the newspapers to understand the Congolese public sentiments towards Rwanda and M23.

4.2 Sentiment Analysis

Sentiment Analysis is a NLP classification task that assigns different emotion labels and corresponding degrees to texts by analyzing the relationships between words, phrases, and sentence structures. And LLM is usually an effective tool for performing sentiment analysis on large-scale textual data due to its various pre-training data.

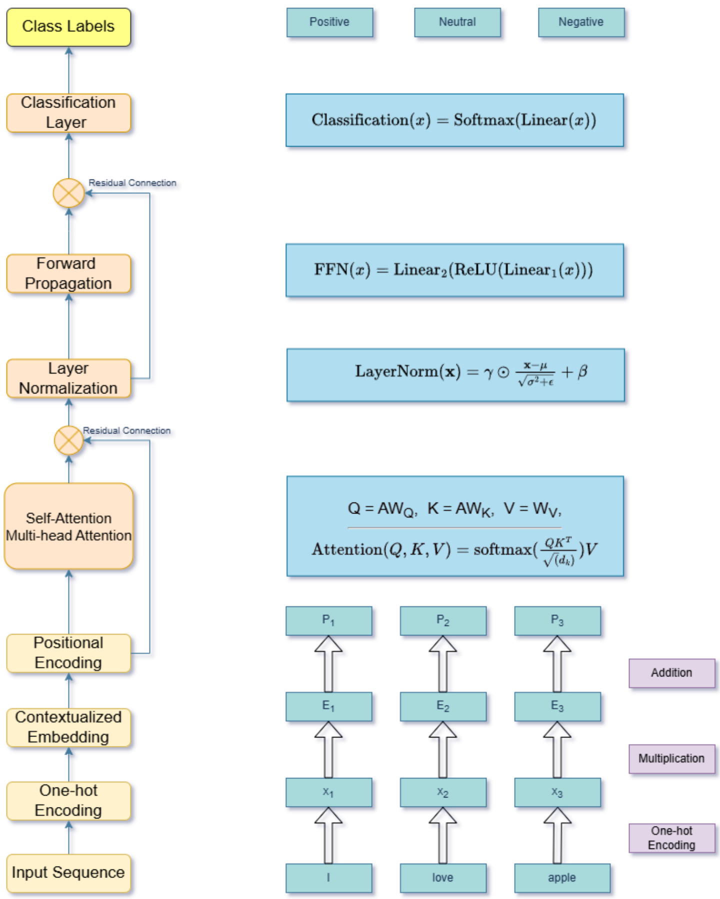


Figure 2: Mechanisms of Transformer based auto-regressive LLMs classification task. If the input sequence is "I love apple", then Transformer[20] would use the attention formula and activation functions to calculate its vector representation. Then it used $softmax$ function to assign the labels with highest probability.

4.2.1 Attention Mechanism

In order to capture the correlation of long sequences of text and improve classification performance, one of the most efficient techniques used in LLMs is Transformers[20] structure which leveraged attention mechanism such as self-attention and multi-head attention to transfer the text into vector representations. And the core formula of of the attention mechanism is Scaled Dot-Product Attention as below:

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

where **Q** represents Query matrix, **K** represents Key matrix, and **V** represents Value matrix. For any input sequence, each word is associated with a unique set of Q, K, and V matrices incorporating positional encoding to capture word order as we showed in Figure 2. And these comprehensive vector representations of text significantly enhance the accuracy of classification tasks such as sentiment analysis.

4.2.2 Models Deployment

We deployed the Multilingual Bert[18] and DistilRoberta[19] models on Google Colab, leveraging an A100 GPU with 40GB of RAM to perform sentiment analysis on four Congolese newspaper articles from 2020 to 2024.

- Multilingual BERT [18] is a transformer-based model pre-trained on 104 languages using the masked language modeling technique [10], enabling it to effectively handle classification tasks for French text. As a

BERT-based model, it captures contextual relationships between words within a sentence, allowing for more accurate sentiment analysis. Its outputs consist of five sentiment labels: very negative, negative, neutral, positive, and very positive.

- DistilRoBERTa [19] is an English-based distilled version of the RoBERTa [21] model through knowledge distillation technique[22], which retains most of the performance of the original complex model while using significantly fewer parameters for efficient deployment. It provides comprehensive understanding of the text with seven specific emotion labels: anger, fear, disgust, sadness, neutral, surprise, and joy.

Subsequently, we adjusted the hyperparameters, "max_length" to 512 and "truncation" to true, to ensure the models accommodated our datasets effectively. Furthermore, as the sequence length of articles in our datasets was too long, we divided the text into chunks to facilitate efficient Google translation.

4.2.3 Evaluation

After obtaining the sentiment analysis results, we structured the sentiment labels as columns with their corresponding degrees as values in the final output for each dataset. In order to explore the alternation in Congolese public sentiments towards Rwanda, we computed the mean of sentiment degrees across different time periods to identify trends in Congolese sentiments.

As an evaluation metric, we hypothesized that a change in the mean sentiment degree greater than 0.3 indicates a substantial shift in public sentiments, given that the overall sentiment degree range spans from 0 to 1. Then, we visualized the sentiment distribution of each newspaper by year using bar plots and the variation in public sentiment degrees using box plots to observe patterns in Congolese public sentiment shifts towards Rwanda and M23 as reflected in the newspapers.

5 Results and Analysis

In this section, we focused on the Radio Okapi and Actualite datasets to compare significant terms, topic words, sentiment distribution, and sentiment degrees to provide insights from both a public-interest media and a private media perspective. Furthermore, both datasets contained a large number of articles that could offer comprehensive understanding about the alternation in Congolese public sentiment.

Additional tables for other datasets can be found in Appendix A, while corresponding figures are presented in Appendix B for further reference. In Appendix C, it contains the statistical results and mean degree alternation following time of different newspapers after sentiment analysis of LLMs.

5.1 Terms and Topics

5.1.1 TF-IDF

After applying the TF-IDF algorithm to the newspaper datasets, we identified the ten most significant terms based on their respective TF-IDF scores. And these terms potentially reflect public sentiments and attitudes as conveyed through the newspapers.

As shown in Table 2a, the words "m23," "drc," and "rwanda" emerged as the top three most significant terms in the Actualite newspaper, suggesting that these terms frequently appear together in the context of news articles. This reflects the strong association of M23 with Rwanda which corresponds to widely perception among Congolese audiences. Furthermore, the high frequency of the words "force" and "army" highlights the significant emphasis on military involvement in the coverage. This suggests that discussions surrounding M23 in the Actualite newspaper often revolve around military actions and security dynamics in the region.

Word	TF-IDF Score
m23	57.51
drc	42.07
rwanda	35.14
congolese	34.59
force	31.26
army	27.92
kivu	27.70
armed	27.59
goma	27.17
north	26.97

(a) Top 10 words in the Actualite.

Word	TF-IDF Score
modified	54.72
m23	38.20
posted	36.87
drc	31.45
mar	31.05
le	29.56
published	27.69
rebel	27.09
congolese	26.19
kivu	25.15

(b) Top 10 words in the Radio Okapi.

Table 2: TF-IDF scores of the top 10 words in the Actualite and Radio Okapi datasets.

Conversely, according to Figure 2b, the words "modified," "m23," and "posted" are top three significant terms in the Radio Okapi newspaper. Compared to the terms in Actualite, Radio Okapi did not associate M23 with Rwanda to publicly express suspicion or critical sentiments. It primarily illustrated posted announcements and reported on rebel activities based on observations in the North Kivu region.

5.1.2 LSA

The latent semantic analysis revealed potential terms associated with underlying topics. The topic terms in Actualite primarily focused on "m23," "rwanda," "army," "defense," and "government" as shown in Table 3, which positioned Rwanda as the primary aggressor in M23 rebel attack and expressed anger towards it. Some articles even suggested military involvements to restore the peace.

Topic	Top Terms
1	m23, drc, rwanda, congolese, rebels, army, armed, goma, security, kivu
2	goma, sake, rutshuru, clashes, city, army, rebels, wazalendo, fardc, displaced
3	rwanda, rwandan, congolese, army, kigali, kinshasa, accused, tutsi, parc, afp
4	force, regional, eac, sadc, rebels, summit, withdrawal, m23, african, troops
5	siege, national, state, elected, government, parc, des, bunagana, defense, minister

Table 3: Top 10 Terms for the 5 Most Prominent Topics in Actualite Identified by LSA.

Topic	Top Terms
1	m23, drc, rebel, congolese, territory, kivu, north, group, fardc, armed
2	game, enemy, call, property, governor, vindictive, play, compatriot, act, religious
3	drc, tshisekedi, president, state, rwanda, congolese, country, republic, félix, peace
4	people, displaced, humanitarian, monusco, population, young, site, kivu, beni, goma
5	monusco, keita, bintou, un, group, armed, council, united, drc, humanitarian

Table 4: Top 10 Terms for the 5 Most Prominent Topics in Radio Okapi Identified by LSA.

According to Table 4, the potential topic terms, such as "MONUSCO", "humanitarian", "peace", performed an effort of Radio Okapi to downplay the relationship between M23 and Rwanda. It emphasized the role of the United Nations Organization Stabilization Mission in the DRC (MONUSCO), which is a UN peacekeeping operation dedicated to maintaining peace and security. Unlike Actualite, Radio Okapi refrained from using inciting or emotional language to express its sentiments towards Rwanda to avoid the misunderstanding and tention.

5.2 Sentiment Distribution

5.2.1 Multilingual Bert Distribution

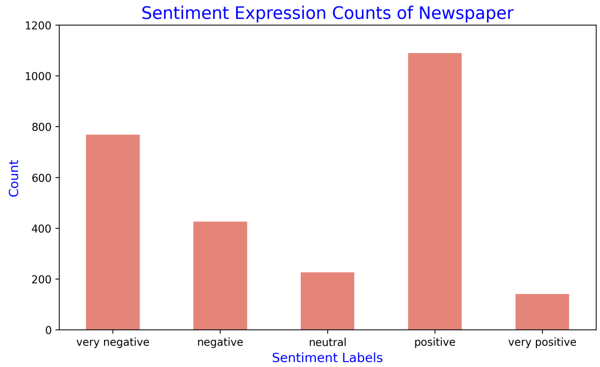
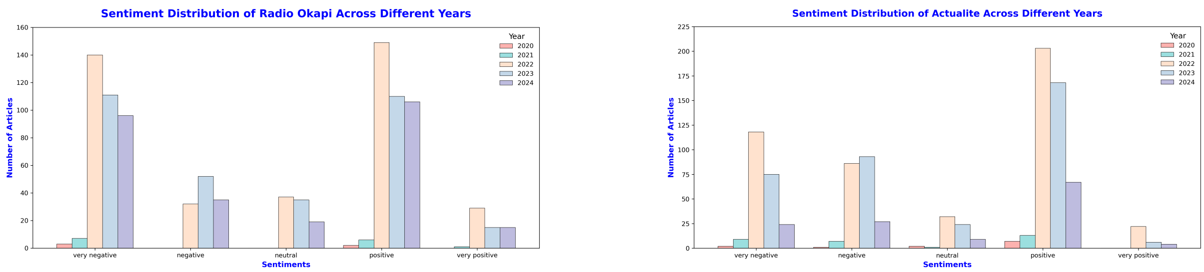


Figure 3: Overall Sentiment Distribution among Congolese Newspapers

According to explore sentiment distributions on our datasets, the Congolese newspapers holistically demonstrated a polarized distribution of public sentiments towards Rwanda after M23 rebel attack as Figure 3 shown. On the one hand, the number of articles expressing very negative sentiment towards Rwanda reached almost 800, while those with very positive sentiment numbered only 180. On the other hand, despite approximately 400 articles expressing negative attitudes towards Rwanda, around 1100 articles expressed positive attitudes towards Rwanda, representing the Congolese expectation of a peaceful relationship with Rwanda.



(a) UN-affiliated media Radio Okapi

(b) Congolese private media Actualite

Figure 4: Sentiment Distributions in Radio Okapi and Actualite through MULTILINGUAL BERT.

As a UN-affiliated media outlet, the number of articles in Radio Okapi expressing both very negative and positive sentiments experienced similar growth rates in 2022, with an approximate 15-fold increase compared to 2021 as shown in Figure 4a. The surge in articles expressing very negative sentiment reflected calls from Congolese officials for international intervention and condemnation of the M23 rebel attacks, allegedly supported by Rwanda, aligning with the perspectives presented by Alan et al. [14]. Similarly, there was a notable increase in articles expressing positive sentiment, advocating for a peaceful resolution to the conflict through negotiation and dialogue.

According to Figure 6b, as a private media outlet, Actualite favored publishing articles with positive sentiment over those with very negative sentiment. It exhibited an upward trend, with articles expressing positive sentiment increasing approximately tenfold, compared to a fivefold rise in those with very negative attitudes. However, articles with negative sentiment also experienced a surge in 2022, which reflected growing concerns over regional instability and conflict dynamics. This trend was also evident in 7SUR7, as shown in Figure 12 in Appendix B, indicating that private newspapers inevitably expressed negative sentiments following rebel attacks, despite initially attempting to convey positive sentiments.

5.2.2 DistilRoberta Distribution

The sentiment analysis results from DistilRoBERTa [19] allowed us to indetify specific emotions of newspapers beyond basic sentiment categorization. As Figure 5 shown, we observed that anger, fear, and neutral sentiments were the most frequently expressed categories in Congolese newspapers, indicating a concentration on the negative attitudes towards Rwanda.

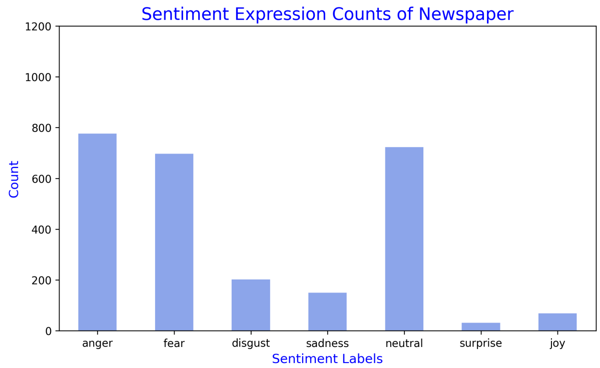
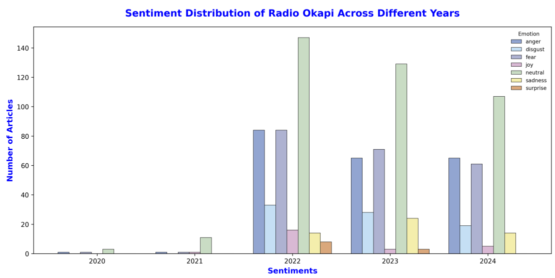
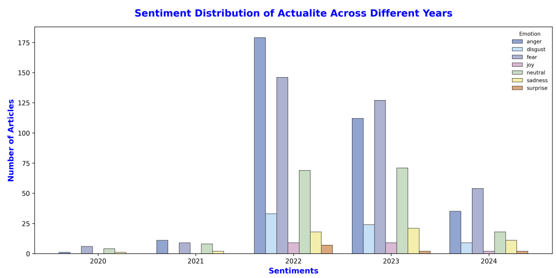


Figure 5: Overall Sentiment Distribution among Congolese Newspapers

Figure 7a demonstrated that while Radio Okapi experienced an surge in negative sentiments like anger and fear following the 2021 M23 rebel attacks, it predominantly maintained a neutral stance to avoid inciting hate. Furthermore, the anger and fear sentiment in Radio Okapi may focus on humanitarian crisis such as slaughter and disease according to topic terms through LSA in Section 5.1.2.



(a) UN-affiliated media Radio Okapi



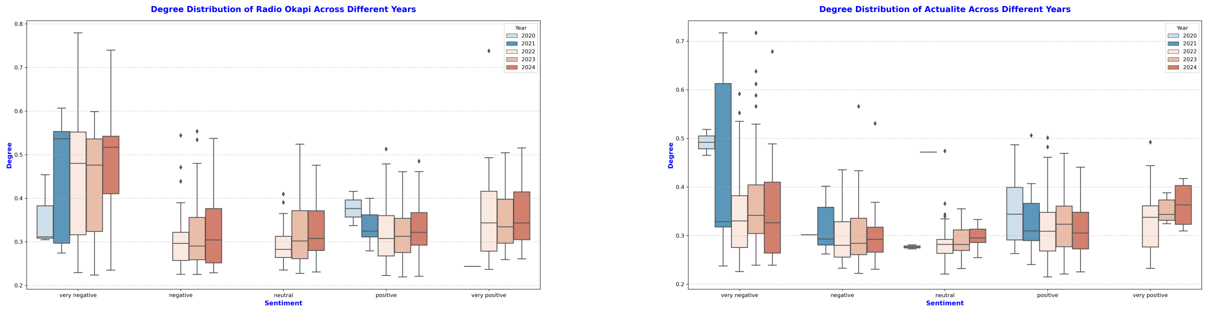
(b) Congolese private media Actualite

Figure 6: Sentiment Distributions in Radio Okapi and Actualite through DistilRoBERTa.

However, Actualite exhibited a different trend as shown in Figure 6b compared to Radio Okapi. It exhibited a significant surge in articles expressing anger and fear in 2022 following the 2021 M23 rebel attack, and these negative sentiments became dominant in reports related to M23 and Rwanda. Although the number of negative articles decreased in subsequent years, this trend persisted into 2024 which reflected the prolonged negative attitudes of Actualite towards Rwanda due to armed conflicts. Additionally, the increase in anger may correlate with the demand for military intervention, as highlighted by the significant terms discussed in Section 5.1.1.

5.3 Sentiment Degree

Due to the limited number of articles available from RTNC, we excluded this dataset from our sentiment degree analysis. The scarcity of data prevented us from identifying insightful trends in public sentiment reflected in its reporting.



(a) Radio Okapi Sentiments Degree

(b) Actualite Sentiments Degree

Figure 7: Corresponding degrees of different sentiments in various Congolese newspapers from 2020 to 2024.

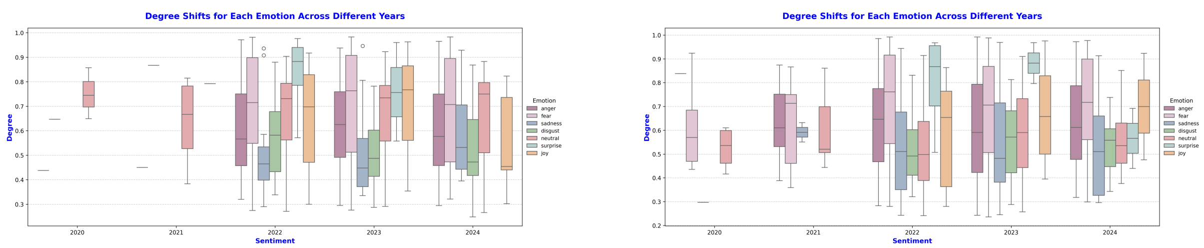
5.3.1 Multilingual Bert Degree

As illustrated in Figure 7a, Radio Okapi predominantly exhibited a moderate level of negative sentiment in 2020. The very negative sentiment degree of majority of articles fell within the 0.3 to 0.4 range, with a median score of 0.31, indicating a generally restrained and less extreme negative sentiment in their reporting. However, the higher median (0.55) and wider interquartile range (IQR) from 0.3 to 0.57 indicated a remarkable increase in negative articles following the M23 rebel attack. Moreover, the upward trend in the minimum negativity scores from 2021 to 2024 suggests a broad-based shift in Congolese public opinion towards a more intensely negative view of Rwanda which aligned with [15] claimed that rebel attacks could reshaped the geopolitics and social sentiments.

According to Figure 7b, the initial IQR of the degree of very negative sentiment ranged from 0.48 to 0.51, with a median of 0.49, indicating that Actualite's overall attitude towards Rwanda was rather negative in 2020 since their former armed operations. After 2021 M23 rebel attack, the range and variability of very negative sentiment expanded to 0.33–0.61, while the median decreased to 0.34 and remained relatively stable through 2024. This trend suggests that, although Actualite initially maintained a negative stance towards Rwanda, its highest degree of very negative sentiment gradually decreased from 2022 to 2024. This shift indicated that Actualite was deliberately moderating its language to avoid using inciting content.

5.3.2 DistilRoberta Degree

According to Figure 6, we observed a significant increase in the emotions of anger and fear. Subsequently, we analyzed the box plots of sentiment degree variations to acquire a deeper understanding of these changes. As shown in Figure 8, both Radio Okapi and Actualite exhibited a high IQR for fear (0.55–0.9) and anger (0.48–0.75). Furthermore, the degrees of fear and anger remained stable from 2022 to 2024, indicating long-term negative

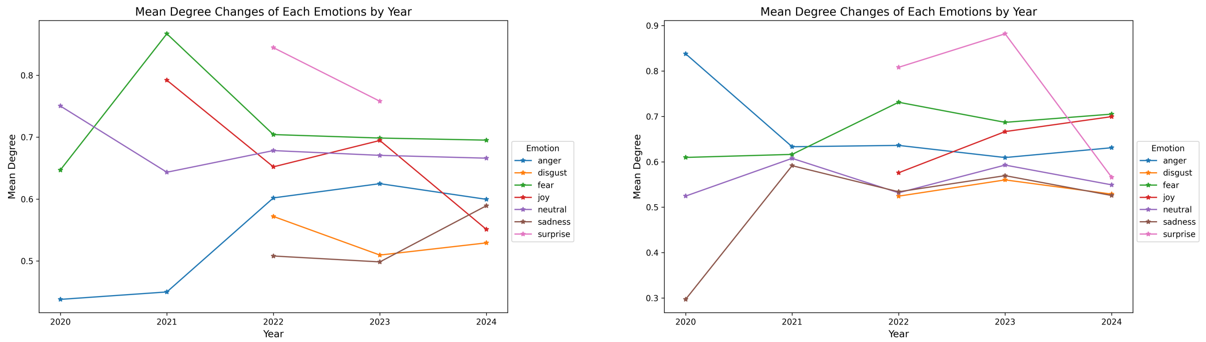


(a) Radio Okapi Sentiments Degree

(b) Actualite Sentiments Degree

Figure 8: Corresponding degrees of different sentiments in various Congolese newspapers from 2020 to 2024.

sentiments in Congolese newspapers toward Rwanda. Additionally, the higher degree of neutral sentiment in Radio Okapi compared to Actualite suggests that public-interest media in Congo tends to adopt a more balanced and less emotionally charged reporting style.



(a) DistilRoberta Okapi Mean Degree Alternation Following Time

(b) DistilRoberta Actualite Mean Degree Alternation Following Time

Figure 9: Comparison of DistilRoberta Mean Degree Alternation for Radio Okapi and Actualite Over Time.

Figure 9a reveals an interesting trend of Radio Okapi journals. While the average level of anger expressed increased significantly in 2022, the average level of fear expressed showed a sudden decrease. This contrasting shift in sentiments suggested a potential change in reporting focus to condemn the humanitarian crisis caused by the m23 to attract international attentions. The negative sentiments represented in Actualite converged to a mean degree of 0.6 in 2021, while the mean degree of anger remained consistent and the mean degree of fear increased from 0.61 to 0.75. This trend demonstrated that the Actualite newspaper may maintain long-term concerns about national security and the well-being of Congolese society.

6 Conclusion

Our study employed sentiment analysis and feature extraction techniques, such as TF-IDF and latent semantic analysis, to investigate public concerns and sentiment shifts reflected in Congolese newspapers during the 2021 M23 rebel attack period. On the one hand, the sentiments expressed in newspapers partially represented public sentiment trends, with Congolese media consistently portraying negative emotions, such as fear and anger, toward Rwanda and M23. On the other hand, topic words like ”m23,” ”rwanda,” and ”army” highlighted underlying public sentiments, reflecting a collective demand for a decisive government response, including military action, as particularly emphasized in private media outlets.

7 Limitation

The limited budget restricted our ability to employ more sophisticated large language models (LLMs) for sentiment analysis. Future research could benefit from utilizing advanced models such as GPT and LLaMA, which offer enhanced capabilities in sentiment analysis to provide better accuracy and in-context understanding. Leveraging these models would enable a more nuanced and comprehensive exploration of the sentiments alternation of Congolese newspapers.

References

[1] Al Jazeera. *DRC government and M23 sign peace deal*. Accessed: 2024-11-27. Dec. 2013. URL: https://www.aljazeera.com/news/2013/12/12/drc-government-and-m23-sign-peace-deal.

[2] ACLED Data. *Actor Profile: M23 (DRC)*. Accessed: 2024-11-27. Mar. 2023. URL: https://acleddata.com/2023/03/23/actor-profile-m23-drc/.

[3] International Centre for Counter-Terrorism (ICCT). *M23: The Forgotten War and an Overshadowed Ceasefire*. Accessed: 2024-11-27. 2024. URL: https://www.icct.nl/publication/m23-forgotten-war-and-overshadowed-ceasefire.

[4] Xinhua News. *Conflict in DRC continues to escalate*. Accessed: 2024-11-27. Nov. 2021. URL: http://www.news.cn/english/africa/2021-11/09/c_1310299193.htm.

[5] Refworld. *Country Report by RVI*. Accessed: 2024-11-27. 2012. URL: https://www.refworld.org/reference/countryrep/rvi/2012/en/97677.

[6] United Nations. *United Nations Report*. Accessed: 2024-11-27. 2022. URL: https://documents.un.org/doc/undoc/gen/n22/757/86/pdf/n2275786.pdf.

[7] The Conversation. *Rwanda and DRC's Turbulent Past Continues to Fuel Their Torrid Relationship*. Accessed: 2024-11-27. 2022. URL: https://theconversation.com/rwanda-and-drcs-turbulent-past-continues-to-fuel-their-torrid-relationship-188405.

[8] Observatory of Economic Complexity (OEC). *Rwanda and DRC Bilateral Trade Profile*. Accessed: 2024-11-27. 2024. URL: https://oec.world/en/profile/bilateral-country/rwa/partner/cod.

[9] Juan S. Morales. "Legislating during war: Conflict and politics in Colombia". In: *Journal of Public Economics* 193 (2021), p. 104325. ISSN: 0047-2727. DOI: https://doi.org/10.1016/j.jpubeco.2020.104325. URL: https://www.sciencedirect.com/science/article/pii/S0047272720301894.

[10] Jacob Devlin et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2019. arXiv: 1810.04805 [cs.CL]. URL: https://arxiv.org/abs/1810.04805.

[11] Hugo Touvron et al. *Llama 2: Open Foundation and Fine-Tuned Chat Models*. 2023. arXiv: 2307.09288 [cs.CL]. URL: https://arxiv.org/abs/2307.09288.

[12] Patrick Y. Wu et al. *Large Language Models Can Be Used to Estimate the Latent Positions of Politicians*. 2023. arXiv: 2303.12057 [cs.CY]. URL: https://arxiv.org/abs/2303.12057.

[13] Marta Milena Barrios. "Colombian cries: Internal armed conflict and emotions in letters to the editor". In: *Journalism* 18.2 (2017), pp. 159–175. DOI: 10.1177/1464884915605030. eprint: https://doi.org/10.1177/1464884915605030. URL: https://doi.org/10.1177/1464884915605030.

[14] Alan J. Lambert, J. P. Schott, and Laura Scherer. "Threat, Politics, and Attitudes: Toward a Greater Understanding of Rally-'Round-the-Flag Effects". In: *Current Directions in Psychological Science* 20.6 (2011), pp. 343–348. DOI: 10.1177/0963721411422060. eprint: https://doi.org/10.1177/0963721411422060. URL: https://doi.org/10.1177/0963721411422060.

[15] Thomas Zeitzoff. "How Social Media Is Changing Conflict". In: *Journal of Conflict Resolution* 61.9 (2017), pp. 1970–1991. DOI: 10.1177/0022002717721392. eprint: https://doi.org/10.1177/0022002717721392. URL: https://doi.org/10.1177/0022002717721392.

[16] BBC News. *Republic of Congo Media Guide*. Accessed: 2024-10-28. Apr. 2023. URL: https://www.bbc.com/news/world-africa-14121194.

[17] Leila Demarest and Arnim Langer. "Peace journalism on a shoestring? Conflict reporting in Nigeria's national news media". In: *Journalism* 22.3 (2021), pp. 671–688. DOI: 10.1177/1464884918797611. eprint: https://doi.org/10.1177/1464884918797611. URL: https://doi.org/10.1177/1464884918797611.

[18] Jacob Devlin et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *CoRR* abs/1810.04805 (2018). arXiv: 1810.04805. URL: http://arxiv.org/abs/1810.04805.

[19] Victor Sanh et al. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter". In: *ArXiv* abs/1910.01108 (2019).

[20] Ashish Vaswani et al. "Attention is All you Need". In: *Neural Information Processing Systems*. 2017. URL: https://api.semanticscholar.org/CorpusID:13756489.

[21] Yinhan Liu et al. *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. 2019. arXiv: 1907.11692 [cs.CL]. URL: https://arxiv.org/abs/1907.11692.

[22] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. *Distilling the Knowledge in a Neural Network*. 2015. arXiv: 1503.02531 [stat.ML]. URL: https://arxiv.org/abs/1503.02531.

A Term Rankings and Topics

This section provides additional tables showcasing the term rankings and topics for other datasets analyzed in this study.

Word	TF-IDF Score
m23	2.47
drc	2.22
rwanda	2.15
state	1.92
president	1.91
rwandan	1.87
goma	1.78
congolese	1.69
force	1.64
press	1.54

Table 5: Top 10 words in the RTNC.

Word	TF-IDF Score
m23	24.19
congolese	23.95
rwanda	21.39
drc	21.20
country	18.92
kivu	18.62
republic	18.06
congo	17.63
president	17.32
north	16.79

Table 6: Top 10 words in the 7SUR7.

Topic	Top Terms
1	drc, m23, rwanda, president, rwandan, congolese, state, republic, united, security
2	national, provincial, electoral, elections, january, results, independent, 14, policy, commission
3	goma, city, fardc, terrorists, alleged, virunga, cut, electricity, mayor, army
4	heads, luanda, mini, process, force, monusco, regional, summit, eac, nairobi
5	minister, defense, meeting, materials, council, government, restoring, part, terrorists, security

Table 7: Top 10 Terms for the 5 Most Prominent Topics in RTNC Identified by LSA.

Topic	Top Terms
1	congolese, m23, rwanda, drc, kivu, country, republic, army, congo, force
2	kivu, goma, north, city, army, rebel, rutshuru, territory, bunagana, clash
3	regional, african, luanda, eac, sadc, nairobi, community, summit, deployment, ceasefire
4	rwanda, rwandan, army, fardc, kigali, kagame, congolese, soldier, congo, m23
5	eac, south, bukavu, regional, fayulu, martin, sadc, force, city, african

Table 8: Top 10 Terms for the 5 Most Prominent Topics in 7sur7 Identified by LSA.

B Sentiment Results

This section provides additional figures related to the sentiment analysis of other datasets conducted in this study.

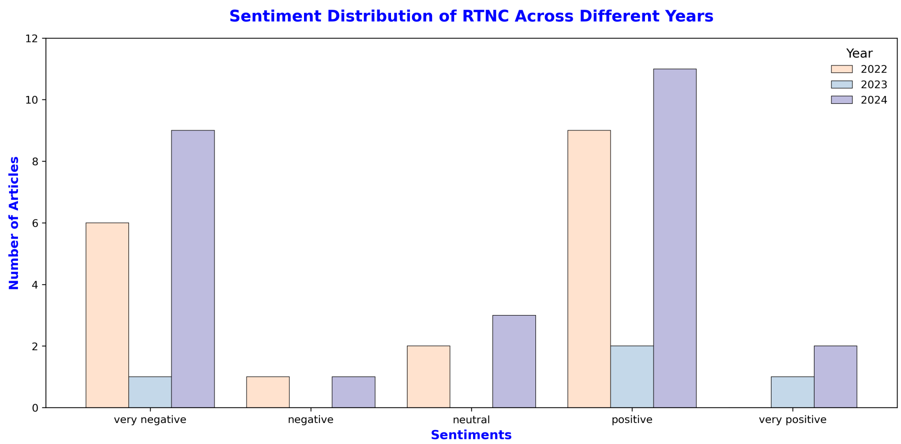


Figure 10: Distribution Congolese State-Run media RTNC

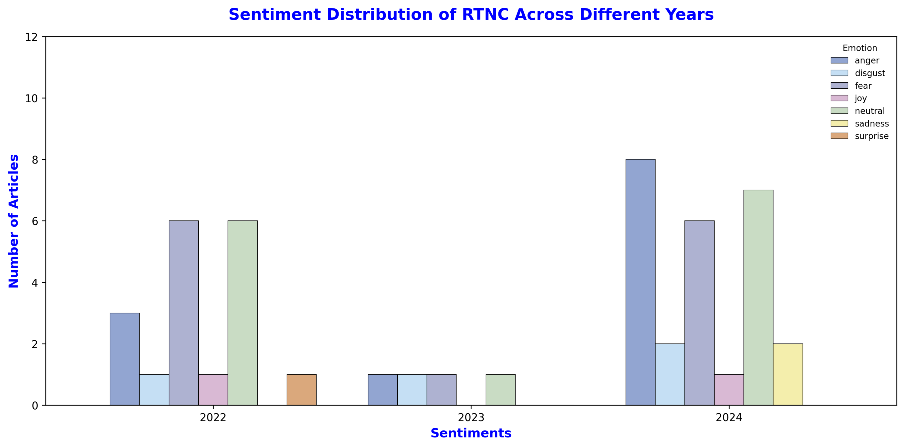


Figure 11: Distribution of DistilRoberta Congolese State-Run media RTNC

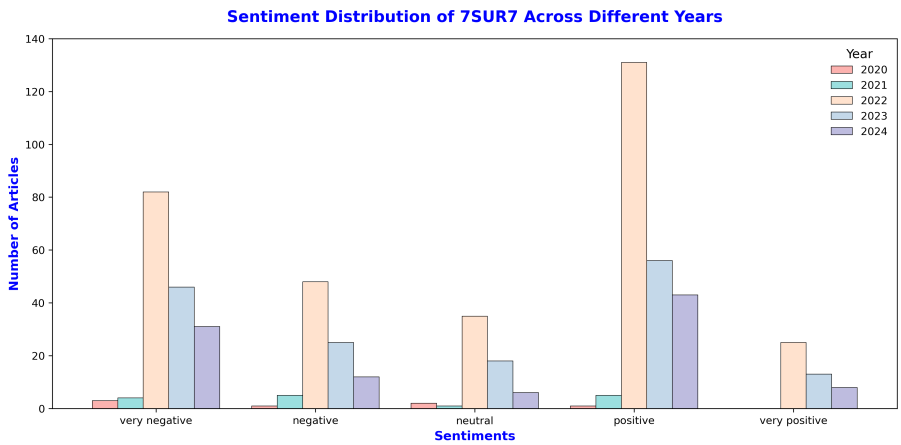


Figure 12: Distribution of Congolese private media 7SUR7

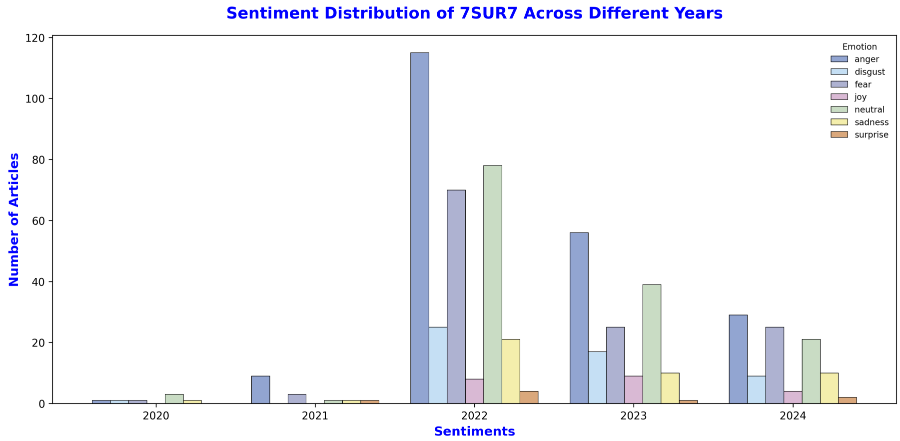


Figure 13: Distribution of DistilRoberta Congolese private media 7SUR7

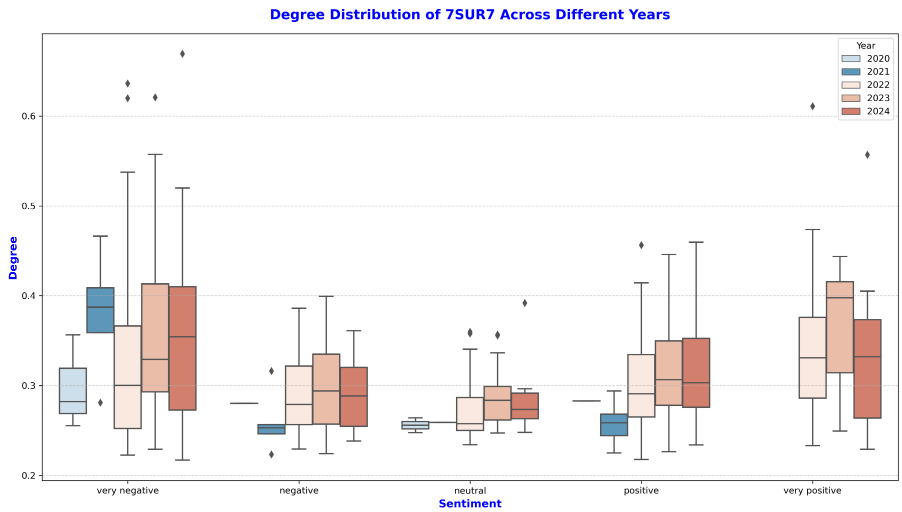


Figure 14: Degree of Congolese private media 7SUR7

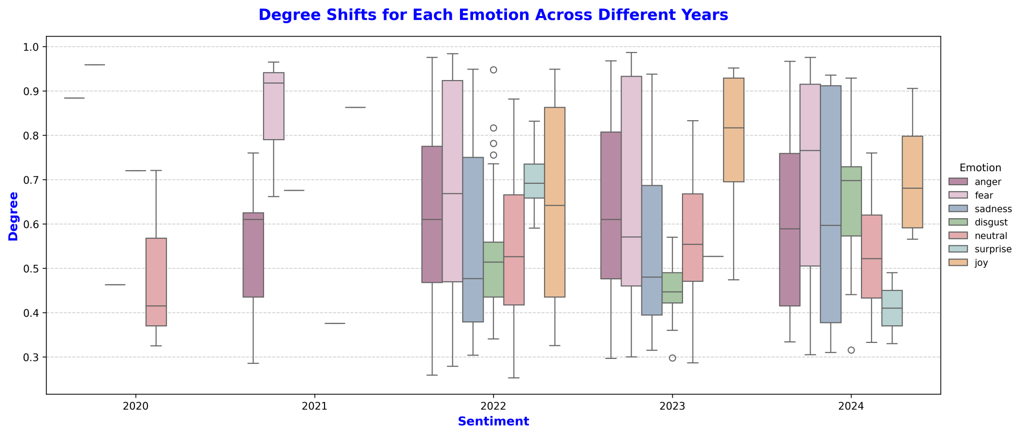


Figure 15: Degree of DistilRoberta Congolese private media 7SUR7

C Statistical Results

This section provides statistics of sentiment degrees acquired from both LLMs in this study.

Year	sentiment	count	mean	std	min	25%	50%	75%	max
2022	negative	1.0	0.227165	NaN	0.227165	0.227165	0.227165	0.227165	0.227165
	neutral	2.0	0.302910	0.031179	0.280863	0.291887	0.302910	0.313934	0.324957
	positive	9.0	0.311933	0.047497	0.268750	0.278376	0.291441	0.343369	0.412900
	very negative	6.0	0.304078	0.040103	0.269265	0.276008	0.291121	0.318496	0.374370
2023	positive	2.0	0.321895	0.058747	0.280355	0.301125	0.321895	0.342666	0.363436
	very negative	1.0	0.523857	NaN	0.523857	0.523857	0.523857	0.523857	0.523857
	very positive	1.0	0.321814	NaN	0.321814	0.321814	0.321814	0.321814	0.321814
2024	negative	1.0	0.352944	NaN	0.352944	0.352944	0.352944	0.352944	0.352944
	neutral	3.0	0.298124	0.016756	0.288449	0.288449	0.288449	0.302961	0.317472
	positive	11.0	0.342360	0.081644	0.229060	0.278606	0.326008	0.427552	0.443906
	very negative	9.0	0.341198	0.123009	0.227896	0.253818	0.305009	0.356044	0.619036
	very positive	2.0	0.466176	0.019591	0.452323	0.459249	0.466176	0.473102	0.480028

Figure 16: Multilingual RTNC Statistics

Year	sentiment	count	mean	std	min	25%	50%	75%	max
2020	positive	2.0	0.376374	0.055603	0.337056	0.356715	0.376374	0.396032	0.415691
	very negative	3.0	0.356751	0.084137	0.305000	0.308210	0.311421	0.382627	0.453834
2021	positive	6.0	0.334881	0.044082	0.279071	0.311061	0.324713	0.362001	0.399849
	very negative	7.0	0.445354	0.148197	0.273999	0.296913	0.536389	0.553192	0.606881
	very positive	1.0	0.243587	NaN	0.243587	0.243587	0.243587	0.243587	0.243587
2022	negative	32.0	0.307447	0.072532	0.225315	0.257234	0.296357	0.321926	0.544051
	neutral	37.0	0.291225	0.039076	0.235322	0.264054	0.282752	0.312705	0.409437
	positive	149.0	0.318471	0.063121	0.222852	0.267480	0.307459	0.360246	0.512791
	very negative	140.0	0.440023	0.126433	0.229345	0.316365	0.480058	0.551993	0.779392
	very positive	29.0	0.356528	0.099714	0.236899	0.278584	0.343286	0.416112	0.737916
2023	negative	52.0	0.311439	0.073172	0.225229	0.258737	0.290325	0.355874	0.553427
	neutral	35.0	0.324622	0.079670	0.227452	0.261126	0.301948	0.371427	0.524095
	positive	110.0	0.315521	0.052847	0.219328	0.275422	0.312691	0.353741	0.460711
	very negative	111.0	0.433783	0.112732	0.223523	0.323722	0.475970	0.536086	0.598744
	very positive	15.0	0.354455	0.071289	0.259066	0.296705	0.334326	0.397744	0.504320
2024	negative	35.0	0.319311	0.076483	0.229058	0.251697	0.304150	0.376149	0.537254
	neutral	19.0	0.325482	0.066627	0.230741	0.280109	0.307619	0.371043	0.475642
	positive	106.0	0.328176	0.055488	0.221062	0.292290	0.321821	0.367039	0.484906
	very negative	96.0	0.472264	0.111238	0.235105	0.410337	0.517002	0.542268	0.739683
	very positive	15.0	0.367402	0.079310	0.260925	0.304960	0.343088	0.414548	0.515511

Figure 17: Multilingual Okapi Statistics

Year	sentiment	count	mean	std	min	25%	50%	75%	max
2020	negative	1.0	0.301689	NaN	0.301689	0.301689	0.301689	0.301689	0.301689
	neutral	2.0	0.276749	0.005987	0.272515	0.274632	0.276749	0.278866	0.280983
	positive	7.0	0.353501	0.080269	0.262962	0.291091	0.344058	0.399260	0.486786
	very negative	2.0	0.491786	0.037703	0.465125	0.478456	0.491786	0.505116	0.518446
2021	negative	7.0	0.319251	0.054929	0.262282	0.280756	0.292936	0.358206	0.401614
	neutral	1.0	0.471916	NaN	0.471916	0.471916	0.471916	0.471916	0.471916
	positive	13.0	0.332663	0.073226	0.240420	0.289869	0.309694	0.366408	0.505981
	very negative	9.0	0.437424	0.189246	0.237453	0.317821	0.328628	0.612805	0.716912
2022	negative	86.0	0.293995	0.050860	0.233053	0.255640	0.280023	0.328465	0.435430
	neutral	32.0	0.287687	0.047260	0.221015	0.263537	0.282113	0.292241	0.473911
	positive	203.0	0.314210	0.057526	0.214952	0.268424	0.309041	0.347743	0.501288
	very negative	118.0	0.340068	0.077083	0.226062	0.275591	0.330242	0.382198	0.591537
	very positive	22.0	0.336423	0.067288	0.232738	0.276723	0.337726	0.361197	0.492045
2023	negative	93.0	0.300786	0.056869	0.222386	0.260932	0.284160	0.335896	0.565712
	neutral	24.0	0.286368	0.030867	0.232180	0.269320	0.281769	0.311663	0.355130
	positive	168.0	0.324667	0.057406	0.220827	0.277129	0.323515	0.360732	0.469208
	very negative	75.0	0.374241	0.110051	0.239271	0.304224	0.341654	0.404292	0.716912
	very positive	6.0	0.351717	0.026928	0.324501	0.331408	0.343577	0.373293	0.388098
2024	negative	27.0	0.299392	0.058846	0.230815	0.265966	0.292455	0.317270	0.530693
	neutral	9.0	0.297617	0.024025	0.254703	0.285688	0.295403	0.313322	0.333133
	positive	67.0	0.312954	0.052023	0.225425	0.272719	0.305292	0.348079	0.440782
	very negative	24.0	0.344269	0.107126	0.239188	0.264179	0.326338	0.409880	0.678490
	very positive	4.0	0.363240	0.052580	0.309599	0.323203	0.362996	0.403034	0.417369

Figure 18: Multilingual Actualite Statistics

Year	sentiment	count	mean	std	min	25%	50%	75%	max
2020	negative	1.0	0.280116	NaN	0.280116	0.280116	0.280116	0.280116	0.280116
	neutral	2.0	0.255991	0.011595	0.247792	0.251891	0.255991	0.260090	0.264190
	positive	1.0	0.282840	NaN	0.282840	0.282840	0.282840	0.282840	0.282840
	very negative	3.0	0.298042	0.052369	0.255466	0.268803	0.282141	0.319329	0.356518
2021	negative	5.0	0.259010	0.034430	0.223238	0.246159	0.252895	0.256705	0.316056
	neutral	1.0	0.258980	NaN	0.258980	0.258980	0.258980	0.258980	0.258980
	positive	5.0	0.257978	0.025801	0.225094	0.244203	0.258624	0.268091	0.293880
	very negative	4.0	0.380429	0.076149	0.280854	0.358845	0.387277	0.408861	0.466308
2022	negative	48.0	0.289249	0.041377	0.229262	0.256534	0.278908	0.321791	0.386219
	neutral	35.0	0.272253	0.032430	0.234112	0.250035	0.257677	0.286738	0.359776
	positive	131.0	0.300116	0.047751	0.217869	0.265141	0.290820	0.334593	0.456397
	very negative	82.0	0.323052	0.087348	0.222545	0.252156	0.300185	0.366398	0.636225
	very positive	25.0	0.343745	0.085245	0.233118	0.286022	0.330900	0.375883	0.610905
2023	negative	25.0	0.301172	0.049369	0.224359	0.257017	0.293950	0.335090	0.399415
	neutral	18.0	0.289223	0.035009	0.247289	0.261605	0.283724	0.298977	0.356754
	positive	56.0	0.312756	0.049027	0.226391	0.277982	0.306599	0.349803	0.445992
	very negative	46.0	0.357109	0.100750	0.229119	0.293106	0.329200	0.413173	0.620844
	very positive	13.0	0.367996	0.065529	0.249270	0.314275	0.397636	0.415469	0.443783
2024	negative	12.0	0.290118	0.042911	0.238202	0.254631	0.288408	0.320211	0.361122
	neutral	6.0	0.290656	0.052147	0.247867	0.263029	0.273515	0.291670	0.391783
	positive	43.0	0.315377	0.055001	0.233885	0.275992	0.303194	0.352665	0.459648
	very negative	31.0	0.356117	0.104134	0.216946	0.272804	0.354207	0.410070	0.669130
	very positive	8.0	0.341654	0.105954	0.229092	0.263843	0.331971	0.373365	0.556834

Figure 19: Multilingual 7SUR7 Statistics

Year	Emotion	count	mean	std	min	25%	50%	75%	max
2022	anger	3.0	0.508333	0.010017	0.498	0.50350	0.5090	0.51350	0.518
	disgust	1.0	0.575000	NaN	0.575	0.57500	0.5750	0.57500	0.575
	fear	6.0	0.606167	0.257994	0.370	0.39100	0.5315	0.80325	0.966
	joy	1.0	0.914000	NaN	0.914	0.91400	0.9140	0.91400	0.914
	neutral	6.0	0.433000	0.164072	0.238	0.34875	0.4145	0.46600	0.721
	surprise	1.0	0.749000	NaN	0.749	0.74900	0.7490	0.74900	0.749
2023	anger	1.0	0.939000	NaN	0.939	0.93900	0.9390	0.93900	0.939
	disgust	1.0	0.465000	NaN	0.465	0.46500	0.4650	0.46500	0.465
	fear	1.0	0.953000	NaN	0.953	0.95300	0.9530	0.95300	0.953
	neutral	1.0	0.752000	NaN	0.752	0.75200	0.7520	0.75200	0.752
2024	anger	8.0	0.629625	0.236203	0.303	0.42875	0.6525	0.75125	0.977
	disgust	2.0	0.693500	0.392444	0.416	0.55475	0.6935	0.83225	0.971
	fear	6.0	0.695667	0.165793	0.448	0.60075	0.7130	0.80875	0.896
	joy	1.0	0.515000	NaN	0.515	0.51500	0.5150	0.51500	0.515
	neutral	7.0	0.706429	0.172776	0.401	0.63850	0.7520	0.78150	0.952
	sadness	2.0	0.702500	0.275065	0.508	0.60525	0.7025	0.79975	0.897

Figure 20: Roberta rtnc Statistics

Year	Emotion	count	mean	std	min	25%	50%	75%	max
2020	**anger**	1.0	0.438000	NaN	0.438	0.43800	0.4380	0.43800	0.438
	fear	1.0	0.647000	NaN	0.647	0.64700	0.6470	0.64700	0.647
	neutral	3.0	0.750333	0.104103	0.649	0.69700	0.7450	0.80100	0.857
2021	**anger**	1.0	0.450000	NaN	0.450	0.45000	0.4500	0.45000	0.450
	fear	1.0	0.867000	NaN	0.867	0.86700	0.8670	0.86700	0.867
	joy	1.0	0.792000	NaN	0.792	0.79200	0.7920	0.79200	0.792
	neutral	11.0	0.643364	0.161179	0.383	0.52700	0.6670	0.78250	0.815
2022	**anger**	84.0	0.601917	0.192185	0.320	0.45700	0.5665	0.75050	0.971
	disgust	33.0	0.572000	0.166368	0.338	0.43300	0.5820	0.67800	0.880
	fear	84.0	0.704190	0.208917	0.274	0.54850	0.7145	0.89850	0.982
	joy	16.0	0.652062	0.205562	0.300	0.47125	0.6975	0.82900	0.917
	neutral	147.0	0.678333	0.150549	0.271	0.56250	0.7310	0.79350	0.904
	sadness	14.0	0.508071	0.192712	0.290	0.39850	0.4645	0.53375	0.936
	surprise	8.0	0.844625	0.134061	0.571	0.78525	0.8825	0.94000	0.976
2023	**anger**	65.0	0.624923	0.172088	0.295	0.49100	0.6250	0.75900	0.938
	disgust	28.0	0.509571	0.131696	0.287	0.41400	0.4875	0.60250	0.782
	fear	71.0	0.698592	0.218846	0.276	0.51200	0.7630	0.90800	0.983
	joy	3.0	0.694667	0.310877	0.354	0.56050	0.7670	0.86500	0.963
	neutral	129.0	0.670636	0.159249	0.291	0.55800	0.7340	0.78500	0.923
	sadness	24.0	0.498542	0.151883	0.335	0.37150	0.4475	0.56850	0.946
	surprise	3.0	0.758000	0.201007	0.558	0.65700	0.7560	0.85800	0.960
2024	**anger**	65.0	0.599600	0.184357	0.294	0.45800	0.5770	0.75000	0.965
	disgust	19.0	0.529263	0.170344	0.248	0.41700	0.4720	0.64550	0.868
	fear	61.0	0.695197	0.217540	0.321	0.47300	0.7070	0.89500	0.983
	joy	5.0	0.551000	0.219055	0.302	0.44000	0.4540	0.73600	0.823
	neutral	107.0	0.666084	0.167752	0.266	0.51000	0.7500	0.79650	0.883
	sadness	14.0	0.589286	0.173295	0.395	0.44250	0.5320	0.70550	0.929

Figure 21: Roberta okapi Statistics

Year	Emotion	count	mean	std	min	25%	50%	75%	max
2020	anger	1.0	0.838000	NaN	0.838	0.83800	0.8380	0.83800	0.838
	fear	6.0	0.609667	0.187850	0.436	0.46950	0.5695	0.68450	0.924
	neutral	4.0	0.524500	0.093646	0.416	0.46175	0.5360	0.59875	0.610
	sadness	1.0	0.297000	NaN	0.297	0.29700	0.2970	0.29700	0.297
2021	anger	11.0	0.633182	0.173115	0.388	0.53150	0.6100	0.75150	0.874
	fear	9.0	0.616333	0.191609	0.359	0.46100	0.7130	0.75000	0.866
	neutral	8.0	0.607625	0.166101	0.444	0.50625	0.5205	0.69900	0.861
	sadness	2.0	0.591500	0.057276	0.551	0.57125	0.5915	0.61175	0.632
2022	anger	179.0	0.636140	0.188722	0.283	0.46750	0.6460	0.77500	0.985
	disgust	33.0	0.524333	0.155042	0.321	0.41100	0.4920	0.60200	0.831
	fear	146.0	0.731199	0.207008	0.280	0.54400	0.7615	0.91600	0.992
	joy	9.0	0.575889	0.229395	0.280	0.36300	0.6540	0.76400	0.863
	neutral	69.0	0.531783	0.175465	0.241	0.38900	0.4980	0.63700	0.914
	sadness	18.0	0.534000	0.215963	0.243	0.35050	0.5105	0.67650	0.944
	surprise	7.0	0.808143	0.186382	0.507	0.70250	0.8670	0.95550	0.967
2023	anger	112.0	0.609402	0.202199	0.243	0.42275	0.5910	0.79350	0.992
	disgust	24.0	0.560167	0.165496	0.288	0.42150	0.5715	0.68175	0.813
	fear	127.0	0.686913	0.201704	0.236	0.50600	0.7060	0.86800	0.988
	joy	9.0	0.666556	0.196917	0.395	0.50000	0.6580	0.82900	0.975
	neutral	71.0	0.592817	0.181848	0.257	0.44300	0.5900	0.73250	0.910
	sadness	21.0	0.569524	0.233636	0.245	0.38200	0.4820	0.71500	0.969
	surprise	2.0	0.882000	0.121622	0.796	0.83900	0.8820	0.92500	0.968
2024	anger	35.0	0.631000	0.186245	0.318	0.47750	0.6120	0.78700	0.972
	disgust	9.0	0.528889	0.120335	0.343	0.44700	0.5580	0.60600	0.738
	fear	54.0	0.705037	0.198279	0.299	0.56025	0.7170	0.89900	0.977
	joy	2.0	0.699500	0.316077	0.476	0.58775	0.6995	0.81125	0.923
	neutral	18.0	0.549278	0.125084	0.376	0.46150	0.5350	0.63000	0.851
	sadness	11.0	0.525818	0.235001	0.296	0.32650	0.5100	0.66000	0.913
	surprise	2.0	0.566000	0.178191	0.440	0.50300	0.5660	0.62900	0.692

Figure 22: Roberta actualite Statistics

Year	Emotion	count	mean	std	min	25%	50%	75%	max
2020	anger	1.0	0.884000	NaN	0.884	0.88400	0.8840	0.88400	0.884
	disgust	1.0	0.720000	NaN	0.720	0.72000	0.7200	0.72000	0.720
	fear	1.0	0.959000	NaN	0.959	0.95900	0.9590	0.95900	0.959
	neutral	3.0	0.487000	0.207586	0.325	0.37000	0.4150	0.56800	0.721
	sadness	1.0	0.463000	NaN	0.463	0.46300	0.4630	0.46300	0.463
2021	anger	9.0	0.549000	0.141678	0.286	0.43500	0.6100	0.62500	0.760
	fear	3.0	0.848333	0.163072	0.662	0.79000	0.9180	0.94150	0.965
	neutral	1.0	0.376000	NaN	0.376	0.37600	0.3760	0.37600	0.376
	sadness	1.0	0.676000	NaN	0.676	0.67600	0.6760	0.67600	0.676
	surprise	1.0	0.863000	NaN	0.863	0.86300	0.8630	0.86300	0.863
2022	anger	115.0	0.626296	0.189493	0.259	0.46800	0.6100	0.77500	0.976
	disgust	25.0	0.541600	0.153425	0.341	0.43500	0.5140	0.55900	0.948
	fear	70.0	0.676271	0.223605	0.279	0.46975	0.6685	0.92325	0.984
	joy	8.0	0.636250	0.246844	0.326	0.43550	0.6420	0.86325	0.949
	neutral	78.0	0.547833	0.178788	0.253	0.41775	0.5265	0.66575	0.882
	sadness	21.0	0.564810	0.223129	0.304	0.37900	0.4770	0.75000	0.949
	surprise	4.0	0.701750	0.099436	0.591	0.65850	0.6920	0.73525	0.832
2023	anger	56.0	0.628839	0.191101	0.297	0.47625	0.6100	0.80750	0.968
	disgust	17.0	0.447824	0.069897	0.298	0.42200	0.4470	0.49000	0.570
	fear	25.0	0.670800	0.243355	0.300	0.46000	0.5710	0.93300	0.987
	joy	9.0	0.775556	0.187450	0.474	0.69500	0.8170	0.92900	0.952
	neutral	39.0	0.564897	0.142059	0.287	0.47100	0.5540	0.66800	0.833
	sadness	10.0	0.548300	0.215485	0.315	0.39475	0.4800	0.68675	0.938
	surprise	1.0	0.527000	NaN	0.527	0.52700	0.5270	0.52700	0.527
2024	anger	29.0	0.598034	0.192462	0.334	0.41500	0.5890	0.75900	0.967
	disgust	9.0	0.649333	0.188223	0.316	0.57300	0.6980	0.72900	0.929
	fear	25.0	0.694600	0.235937	0.305	0.50500	0.7660	0.91500	0.976
	joy	4.0	0.708500	0.157000	0.566	0.59150	0.6810	0.79800	0.906
	neutral	21.0	0.517571	0.122662	0.333	0.43300	0.5220	0.62000	0.760
	sadness	10.0	0.627200	0.268619	0.310	0.37725	0.5970	0.91200	0.936

Figure 23: Roberta 7sur7 Statistics

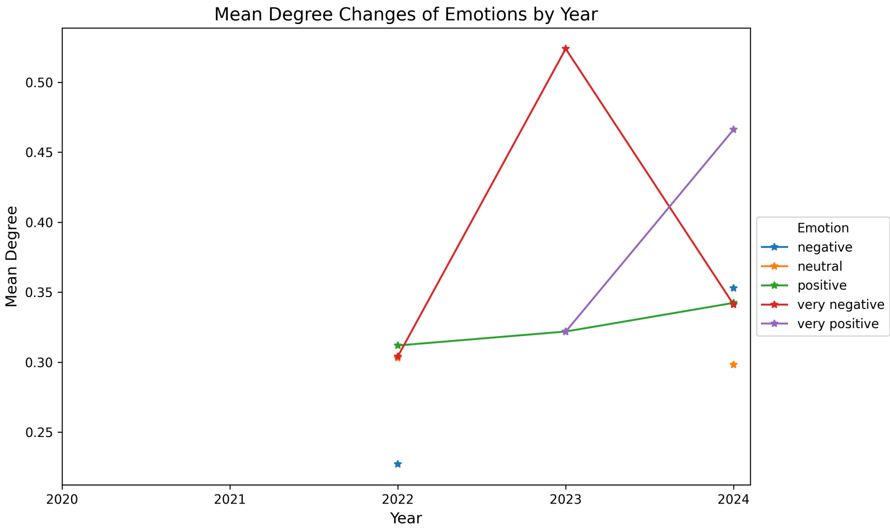


Figure 24: Bert RTNC Mean Degree Alternation Following Time

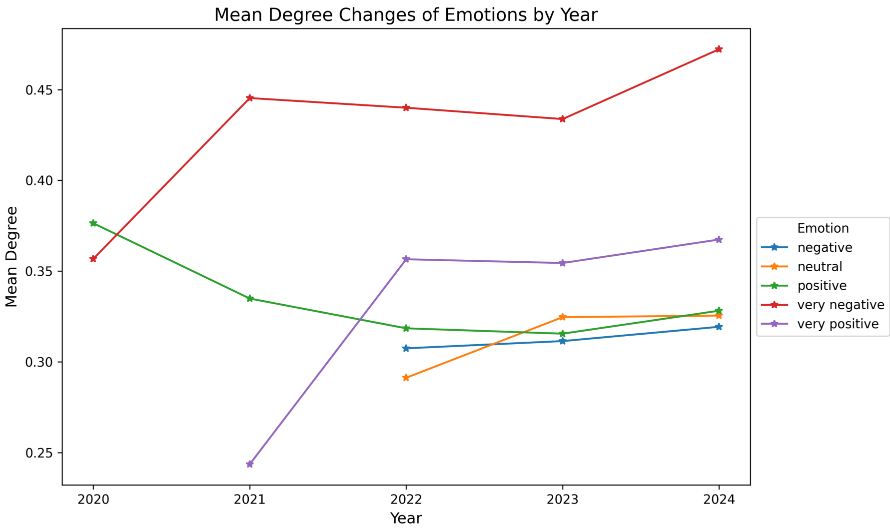


Figure 25: Bert Okapi Mean Degree Alternation Following Time

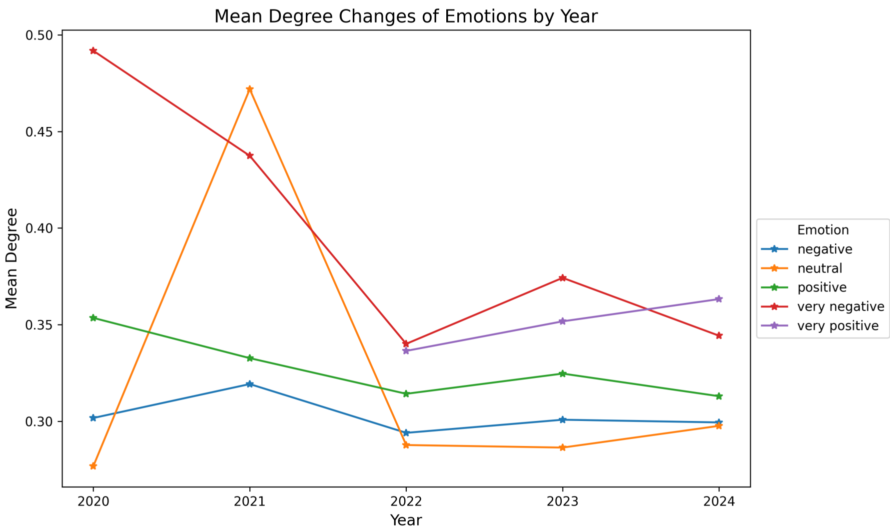


Figure 26: Bert Actualite Mean Degree Alternation Following Time

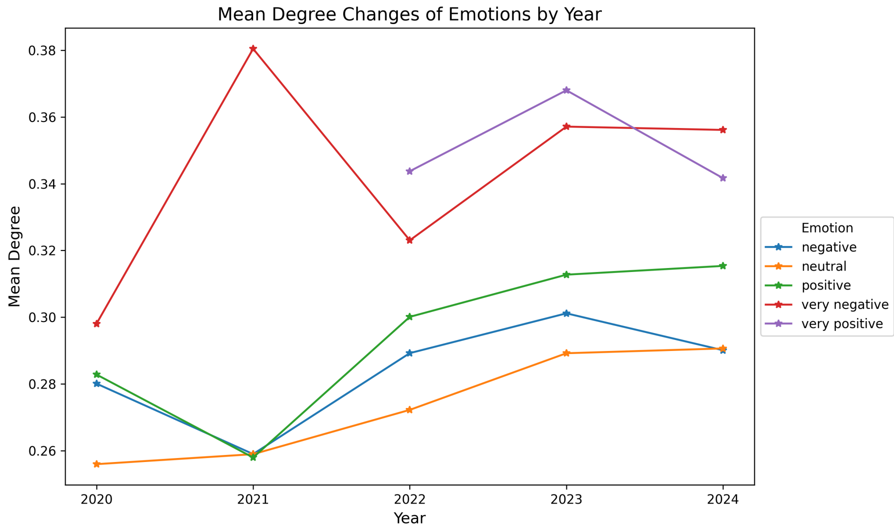


Figure 27: Bert 7SUR7 Mean Degree Alternation Following Time

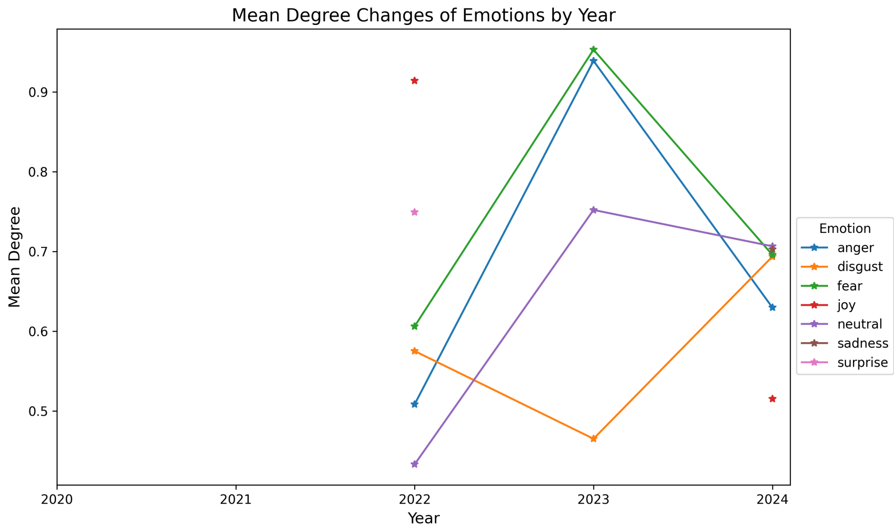


Figure 28: DistilRoberta RTNC Mean Degree Alternation Following Time

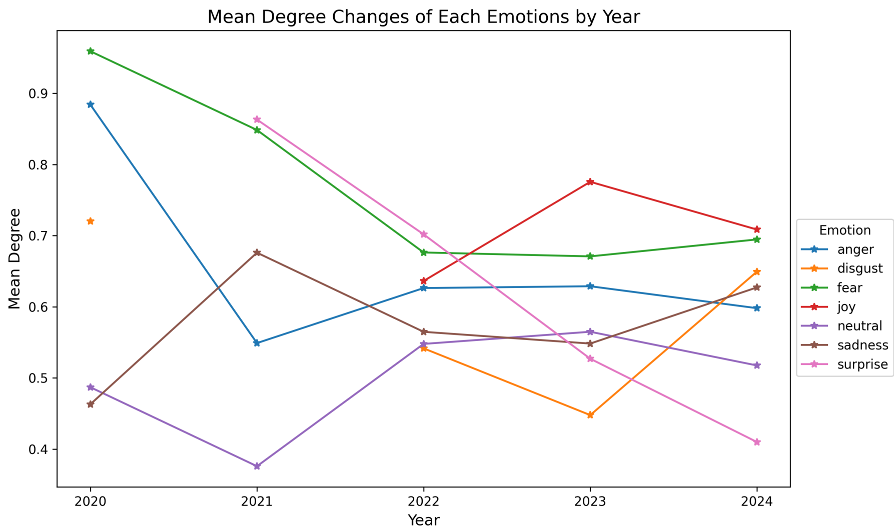


Figure 29: DistilRoberta 7SUR7 Mean Degree Alternation Following Time