

Article

InfluEmo: Influence of Emotions on Instagram Influencers' Success

Chiara Felicia Schettini¹ and Giovanna Maria Dimitri^{2,*} 

¹ Department of Information Engineering and Mathematics, University of Siena, 53100 Siena, Italy; chiara.schettini@student.unisi.it

² Department of Social and Political Sciences, University of Milan, 20122 Milano, Italy

* Correspondence: giovanna.dimitri@unimi.it

Abstract

The use of social networks has been shown as a powerful tool for driving populations opinions towards specific trends and interests. Yet what actually makes the success of a profile? Are emotions responsible for driving the public opinion and the opinion of the followers? We present a study on the influence of emotions in their success. To do so, we first created a novel dataset called InfluEmo, crawled from Instagram, in which we designed and analyzed the impact of emotions in influencers' success. The dataset InfluEmo is novel and freely available. Automatic emotion extraction yielded promising results, supporting our hypothesis that specific emotional profiles in influencers' posted content are associated with measurable indicators of success measured as number of followers. These findings suggest that emotions might play a systematic and quantifiable role in shaping public opinion and influencing users' interactions on Instagram. Using the novel InfluEmo dataset ($\approx 38,000$ posts, ≈ 970 profiles, 4 domains: fashion, climate, AI, and journalism), the paper shows, in fact, that more positive emotional language is consistently associated with higher engagement, with fashion influencers achieving the highest average likes ($\approx 138,885$ /post) and lowest emotional entropy, while AI, climate, and journalism content—characterized by more neutral or mixed emotions—exhibits lower likes (≈ 6761 – $19,544$ /post), weaker sentiment–likes correlations, and higher entropy, indicating that positivity and emotional predictability outperform informational complexity in driving Instagram success.

Keywords: sentiment analysis; emotion recognition; machine learning; instagram



Academic Editor: Hovhannes Harutyunyan

Received: 29 November 2025

Revised: 26 January 2026

Accepted: 3 February 2026

Published: 10 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Social media's influence on economic growth is profound, largely due to its impact on consumer behaviour. In this context, social media influencers play a strategic role in shaping public perception and driving it towards certain behaviours and products. Yet what drives the success of an influencer over others? Can emotions, shared through their posts and publications, make a difference in their success as influencers? In the last few years, social media has played a strategic role, not just from an economic and strategic point of view, but also because it represents a tool that can significantly impact users' minds, driving and the way choices and acts are made [1,2]. In this context, an influencer can be defined as anyone who has the power to influence others' decisions, due to their relationship and interaction established with their personal audience, more specifically, their followers. Typically, an influencer is actively engaged in a specific niche, and is a phenomenon which has seen a sharp rise with the diffusion of social networks. Influencers

often collaborate with companies, helping them to sell their products or services in the context of what is known as “Influencer Marketing” [3]. The phenomenon of influencer marketing has, in fact, rapidly emerged as a leading segment in the marketing field, as evidenced by significant research interest [4–10]. To pursue their goals, influencers on social media platforms continuously share perspectives on a range of products, services, and brands, positioning them as opinion leaders who engage with extensive follower networks [11]. Brands, in fact, are highly interested in social media influencers (SMIs) because they are perceived by audiences as authentic, faithful, and relatable sources, which extends the brand’s reach to larger audiences [11–13]. Influencers are usually categorized by follower counts: micro-influencers with up to 10,000 followers, meso-influencers up to one million, and macro-influencers exceeding a million followers. Reflecting the effectiveness of this approach, U.S. marketers rank influencer marketing as the second most effective strategy, with 94% reporting positive outcomes from influencer marketing campaigns [14]. Yet what is the role of emotions in the final success which an influencer reaches? Does it impact the final success and can we quantify the role of emotions in the final success which influencers hold? The analysis of emotions has been, in fact, profoundly analyzed in several contexts [15–18]. However, how much emotions affect the role of social network influencers is still an open research questions, also due to the lack of publicly available datasets in which this research aspect could be analyzed. In this context, we introduced several novelties in our article. First of all, we performed a novel comprehensive analysis to understand the impact that emotions have on success, analysing it through an affective computing perspective on the texts shared by influencers through captions. To do so, we designed and web-scraped a novel dataset, which we named InfluEmo, where we investigated the role of emotions in understanding influencers’ success on Instagram. The paper is structured as follows. In Section 2, we present an overview of related works with respect to our study. In Section 3, we describe the methods used, both for scraping and designing the new dataset, and for performing emotion recognition. In Section 4, we thoroughly describe the dataset and its characteristics. In Section 5, we describe the emotional recognition experiments performed and the analysis which correlated the success of such profiles together with the number of influencers. In Section 7, we draw conclusions and sketch future developments of our work.

2. Related Work

In this Section, we will discuss previous related works, showing the differences and novelties of our proposed work. In [19], Oliveira et al. analyzed the behaviour on Instagram of users, focusing on how positive sentiments and concise captions boost engagement for fashion influencers on Instagram. However, this study focused the analysis on the specific sector of fashion influencers and did not release their dataset publicly. Similarly, Zhan et al. [20] performed Sentiment Analysis of Instagram Captions and explored user preferences through sentiment analysis of Instagram captions, without addressing the success of influencers. In other cases, present in the literature, only specific or limited studies are analyzed. For instance, in [21], Mackay et al. examined how positive messages from Canadian influencers positively influenced engagement during the pandemic but the dataset was limited by only Canadian influencers. Moreover, in [22] Bashari et al. proposed a machine learning method to distinguish influential posts on Instagram by analyzing only user-generated content, with a dataset containing influencers and non influencers, but without explaining the relationship between the success expressed by the content and their emotions evoked. Other related works provide overviews and reviews on the topic, as, for instance, in [23]. Here, Nandwani et al. provide an overview of sentiment analysis techniques and emotion detection in text, highlighting specific applications and

challenges. As a further example, in [24], Ranjan et al.'s approach is strictly based on comments in social media using an NLP approach, without focusing on the influencer and the emotion perceived by their content. In addition, in [25], the authors examined how text and images on social media convey emotions, using neural networks to enhance sentiment interpretation. However, the sentiment analysis is limited, as it describes emotions only in terms of polarities without exploring the impact of influencers. Furthermore, other works concentrated on specific domains such as [26,27], where the authors explore specific areas, such as financial headlines to improve market forecasting, and user preferences across social platforms, without addressing influencer success metrics or how emotions are evoked by the content. In the article [28], Gu et al. investigate how the emotional expressions that social media influencers display in their posts affect their popularity by shaping how followers perceive them. They conducted one correlational study and three experiments to examine how emotional style (positive vs. negative) influences perceptions of warmth and competence, online engagement, and intentions to spread electronic word-of-mouth. Rather than simple emotional valence, the study finds that perceived authenticity and perceived appropriateness of emotional expressions are key processes linking influencers' emotions to their popularity: posts judged as both appropriate and authentic tend to enhance popularity indicators. Positive emotions are generally seen as more appropriate, but their effects on perceived authenticity are variable and context-dependent. Overall, the results suggest that successful influencers must balance emotional norms (appropriateness) with genuineness (authenticity) to build followership. In this study, however, data are retrieved from user studies, and not automatically from post and social media information. Moreover, some further more recent studies focus on specific countries and do not actually expand the analysis performed to a wider larger number of different cultural situations [29,30]. These studies employ various analyses across different social media platforms. However, not all of them focus specifically on Instagram, nor do they pay particular attention to the emotions evoked by captions and their impact on success on this platform. Furthermore, none of these studies provided a dataset that compares different types of influencers based on their areas of interest.

3. Methods

3.1. Sentiment Score Computation

All of the analyses were performed in RStudio, version 4.3.1 [31], and the dataset used was scraped from the social network Instagram ([instagram.com](https://www.instagram.com)) [32]. The sentiment analysis was performed to evaluate the sentiment of the text, using the *Syuzhet* package [33]. The predefined method, *syuzhet*, is an R package version 4.4.2 developed by Matthew Jockers, which uses existing sentiment dictionaries such as the NRC Emotion Lexicon [34]. In our analysis, the words could be belonging to three emotion classes: positive, negative, and neutral. The calculation of the overall score works as follows:

- **Positive Sentiment:** if the score associated with the word is greater than 0.
- **Negative Sentiment:** if the score associated with the word is less than 0.
- **Neutral Sentiment:** if the score is equal to 0.

To assign a word to one of the classes, the NRC Emotion Lexicon and the NRC Emotion Intensity Lexicon were used [35]. The lexicon consists of words associated with 8 emotions (joy, sadness, anger, fear, disgust, surprise, anticipation, and disgust) and 2 polarities (positive and negative). The NRC Emotion Lexicon is based on associating a specific emotion with each word, while the NRC Emotion Intensity Lexicon focuses on the intensity associated with each word for its corresponding emotion. For the sentiment association of each word, the emotion-sentiment transfer method is applied, which assigns a sentiment score to each word based on the associated emotion [35]. To compute

the sentiment score, we used the following equation (with Pos being positive and Neg being negative):

$$S(w_i) = \sum_{j=1}^8 [E(w_{ij}) \times P(e_j \rightarrow \text{Pos}) - E(w_{ij}) \times P(e_j \rightarrow \text{Neg})] \quad (1)$$

where

- $S(w_i)$ is the sentiment score associated to the i -th word w_i .
- $E(w_{ij})$ is the score related to the intensity of the emotion associated to each word w_i in respect to the emotion e_j .
- $P(e_j \rightarrow \text{Pos})$ and $P(e_j \rightarrow \text{Neg})$ are the probabilities that the emotion e_j is associated with a positive or negative polarity.

This method classifies each word as positive or negative and provides the corresponding emotional intensity. In our work, we proceeded in the following way: we first considered the caption of each scraped influencers post. We then tokenized the sentence, and we computed a summative score for each post, considering the average scores of all of the words belonging to it.

In Table 1, we report an example of a sentence. If the text contains the following sentence: 'I hate war because it leads to destruction', we first obtained the tokens: 'I', 'hate', 'war', 'because', 'it', 'leads', 'to', 'destruction'. We then computed the sentiment score for each of the token, to obtain an average sentiment score for the post, which corresponds to -0.2 in this case.

Table 1. Example of the sentiment score computed for an instagram post in InfluEmo.

Token	Word	Sentiment Score
1	I	0
2	hate	-0.75
3	war	-0.50
4	because	0.00
5	it	0.00
6	leads	0.40
7	to	0
8	destruction	-0.75

3.2. Impact of Emotions on the Sentiment Score

We further expanded our analysis to also investigate how each of the eight emotions contributed to the sentiment score detected in each token. As previously mentioned, we classified emotions into 8 classes as described in [36,37]:

1. **Anger:** Associated with aggression, hostility, and frustration.
2. **Fear:** Linked to uncertainty, threat, and anxiety.
3. **Anticipation:** Involves expectations and predictions.
4. **Trust:** Refers to confidence and reliability.
5. **Joy:** Represents feelings of pleasure and happiness.
6. **Sadness:** Reflects emotions of pain or sorrow.
7. **Disgust:** Expresses repulsion towards someone or certain situations.
8. **Surprise:** Reaction to shock or unexpected events.

In our code, the NRC Emotion Lexicon [34] calculates the association of these words with the eight emotions and the two polarities.

In order to understand the influence of each of the 8 emotions in the sentiment score, we implemented the following procedure:

1. **Step 1:** The NRC dictionary [35] was used for classifying emotions. This dictionary contains eight emotion categories and two polarities (positive and negative).
2. **Step 2:** We considered texts taken from the influencers datasets, and tokenized the captions.
3. **Step 3:** Stop words, such as prepositions and other commonly used words, were removed.
4. **Step 4:** The resulting dataset, now containing only relevant words, was merged with the NRC dictionary. Each word contained in the captions was then assigned a corresponding sentiment label from the NRC dictionary.
5. **Step 5:** Finally, the frequency of words associated with each sentiment was calculated, and the results were sorted by frequency. This allowed each sentiment to be represented by its occurrence count, as described in Section 5.

3.3. Shannon Emotional Entropy

Shannon entropy was applied, after the sentiment analysis scores, to offer a comprehensive view of the language structure within each group of influencers of our dataset (see Section 4 for description on InfluEmo design). Shannon entropy, in fact, was used as a measure of the uncertainty or randomness in a set of data. We could assess that for a given probability distribution, it quantifies the average amount of information or “surprise” associated with each possible outcome. The formula for Shannon entropy H is

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

where

- $H(X)$: Represents the Shannon entropy of a random variable X .
- n : The number of unique outcomes or symbols in the distribution of X .
- $p(x_i)$: The probability of each outcome x_i .
- $\log_2 p(x_i)$: The binary logarithm (base 2) of $p(x_i)$, representing the amount of information in bits associated with outcome x_i .

More specifically, in our work, we used Shannon entropy in order to compute the so-called *emotional entropy* of a sentence. The concept of “emotional entropy” in a string refers to the degree of valence variability or conflict within the emotional tones of the words it contains. Essentially, emotional entropy serves as an indicator of the unpredictability or level of surprise arising from the coherence—or lack thereof—in the emotional expressions throughout the sentence. High emotional entropy suggests greater inconsistency in emotional tone, indicating a mix of positive and negative sentiments, while low entropy would indicate a more uniform emotional expression. This metric thus provides insights into the emotional complexity of a given string by assessing the fluctuation in sentiment across its language. We further computed the Pearson correlation coefficient in order to understand both the correlation between emotions and between the presence of a certain emotion and success of an influencer.

3.4. Engagement Calculation

Moving forward, we used also the so-called engagement score in order to understand relationships between emotions and engagements of the users. The engagement rate measures how much interaction there is of the public users to a specific content, and so how they behave and the feelings that are generated by a specific content [38].

The calculation of the engagement score is the following:

$$\text{Engagement Rate} = \frac{\text{interactions}}{\text{number of followers}} \times 100$$

Since the number of followers was not available, as an information, we modified the formula considering the number of likes and comments as the denominator of the ratio. From our analysis of interactions on the posts, it was observed that engagement plays a vital role in determining the success of a content strategy. The engagement was calculated using the variables of “likes” and “comments” to assess user interactions. More specifically, we computed two types of engagement rates:

$$\text{EngagementRateLikes} = \frac{\text{number of likes}}{\text{number of posts}} \times 100 \quad (2)$$

$$\text{EngagementRateComments} = \frac{\text{number of comments}}{\text{number of posts}} \times 100 \quad (3)$$

3.5. Definition of Influencer Success

This paper defines influencer success as a set of engagement-based, quantitative proxies derived from Instagram interactions, measured using likes and comments. For each post, the raw number of likes and comments is used as a primary indicator of audience response, e.g., average likes per post vary strongly by group (fashion \approx 138,885, climate \approx 19,544, AI \approx 6761, and journalists \approx 8538). Engagement rates (normalized by activity): Since follower counts were unavailable, success is quantified through post-normalized engagement metrics defined above, that is EngagementRateLikes and EngagementRateComments, capturing how much interaction each post generates on average. Moreover, success is further measured by computing Pearson correlations between the sentiment score of caption words and the number of likes, where positive correlations (strongest in fashion and journalism) indicate that more positive emotional language yields more engagement. As a further metric, this paper uses Shannon emotional entropy of captions and lower entropy (more emotionally coherent and predictable language) is associated with higher engagement, while higher entropy corresponds to reduced success (notably in climate and AI content).

4. InfluEmo Dataset

4.1. Dataset Scraping and Design

To download and extract our dataset, we used a web scraping application, capable of extracting a users profile post, called ‘Apify’ [39]. The results of Apify are dynamic, meaning that the extraction process can always change over time. Apify adheres to Instagram’s ethical guidelines, as the extraction does not include private user data such as email addresses, location, or gender. The extraction only involves what the users choose to publish and only public profiles. In our case, we decided to design the novel dataset, considering four areas where influencers are highly present on Instagram: artificial intelligence (AI), climate, journalists, and fashion. The four datasets are structured to contain key variables such as profiles, posts, comments, likes, and captions. All of these elements are publicly released and can all be critical in the definition of emotion recognition. To obtain the data for each group, an automatic research was conducted through the web to identify a list of the most influential public Instagram profiles within each group of interest. The profiles were then collected along with their profiles’ URLs. In the scraping procedure, we proceeded by collecting over 100 of URLs for each influencer group, which were then fed to Apify in order to perform the scraping itself. Due to the extensive nature of the data extraction, it took more than a month to complete the process. The overall design process and the manual/topic-based categorization grouping process were carried out through the following steps:

1. **Topic Identification:** An initial research phase identified key topics relevant to the study. These topics were derived from analyzing influencers' posts, captions, and engagement metrics, with common themes emerging in areas such as climate activism, artificial intelligence, journalism, and fashion.
2. **Data Collection:** Data was gathered by focusing on influencers who demonstrated a strong interest in the identified topics. This involved examining their posts, likes, comments, and interactions with followers.
3. **Manual/topic-based categorization methodology:** Influencers were grouped into based on the primary topics they discussed and the overall similarity of their content.
4. **Manual/topic-based categorization criteria:** The criteria for forming groups included
 - **Content Analysis:** Assessing the frequency and relevance of keywords within captions and content related to each topic featured in influencers' posts.
 - **Engagement Metrics:** Evaluating likes and comments associated with specific topics.

The dataset collected is freely available and downloadable at the following link: https://drive.google.com/drive/folders/19M04M7DtLzP_cE-6KIik8Q4Ze1pFuPyN?usp=sharing (accessed on 15 January 2026).

4.2. Dataset Description

For each influencers' post, we also retrieved the corresponding caption, as well as additional information such as likes and comments. In Table 2, we report the main statistics and content information for each influencers group included in InFluEmo.

Table 2. Summary of influencer datasets.

Information	Climate	AI	Journalist	Fashion
Posts	3883	13,528	3447	17,085
Images	1285	5841	1276	4146
Sidecars	1480	3053	674	8524
Videos	1118	4634	1497	4415
Comments	1,497,042	1,251,386	1,268,510	12,278,121
Likes	77,024,098	91,609,122	29,730,855	2,583,396,940
Profiles	193	304	195	280
Captions	3825	12,981	3331	15,928
Range of Time	31 December 2011 to 25 May 2024	30 December 2013 to 23 May 2024	2 May 2011 to 2 May 2024	22 May 2012 to 23 May 2024

4.3. Data Cleaning

In the final collected dataset, we also had the two features of 'sponsored' and 'paid-partnership', which represent the collaborations between users and organizations. These features were not considered in the calculation of engagement due to insufficient data in respect to the total number of posts of each group as shown in Table 3.

Table 3. Data of paid partnerships and sponsors for each influencer group.

Feature	Fashion	Climate	AI	Journalist
paidPartnership	442	62	117	4
sponsor/0/id	313	37	77	3
Number of Posts	17,085	3883	13,528	3447

Regarding the size of the dataset, it is important to mention that the total number of posts does not match the total number of URLs in the "url" column, as some URLs correspond to non-existent data. Specifically, in the fashion influencer dataset, there are 17,092 URLs but only 17,085 actual posts. For climate influencers, there are 3903 URLs compared to 3883 real posts. In the AI influencer dataset, there are 13,553 URLs and

13,528 actual posts. Finally, the journalist dataset contains 3455 URLs but only 3447 real posts. Because of this, for the calculation of the total number of posts was considered the column “type” which contains the type of the published content that exist. This is important to avoid compromising the integrity of the dataset, ensuring that the analyses are based on true values. All the columns related to comments, such as first comment and latest comments, were used for the calculation in Section 3.4 in order to provide a picture of user interactions. In addition, these columns were also taken into consideration to describe and understand the sentiment emerging from user comments by analyzing the text.

5. Experiments and Results

After conducting a preliminary analysis of the captions, which outlines the content of all posted items, the following steps of analysis were undertaken.

Step 1: Identify the most frequent words associated with each caption.

Step 2: Filter out words deemed “not relevant” for the analysis. “Not relevant” refers to words with a neutral sentiment score, and so equal to 0, using the package [33].

Step 3: Analyze the number of likes associated with captions that contain non-neutral sentiment words.

Step 4: Order the results in descending order based on the number of likes.

Step 5: Perform a sentiment evaluation using the package [33] to determine the sentiment generated by the most liked words.

Step 6: Analyze the correlation between the sentiment associated with the words and their respective likes.

Step 7: Plot the results, demonstrating that words with higher likes are associated with a positive sentiment score.

A detailed explanation follows. Firstly, we extracted the more popular words adopted in the captions with the higher number of likes as shown, for example by Figure 11, for each word was associated with their specific number of likes. Figure 1 shows an example.

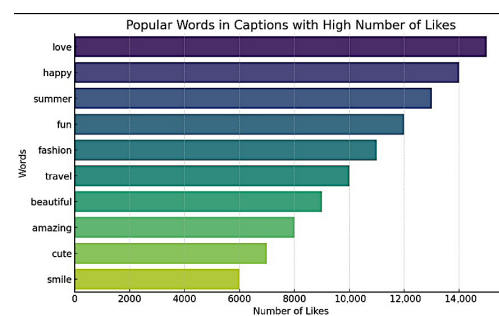


Figure 1. Example of correlation between words with the most likes and the sentiment they express for the fashion influencers.

Then, we considered the words and relative number of likes, for which we computed the corresponding sentiment scores. Finally, the correlation between the ‘likes’ and the ‘sentiment_score’ was computed in order to understand the relationship between the number of likes and the sentiment expressed. For example, if we find a positive correlation coefficient (e.g., 0.80), this suggests that as the sentiment score increases (more positive sentiments), and the number of likes also tends to increase. Conversely, a negative correlation coefficient (e.g., -0.40) would indicate that higher sentiment scores are associated with fewer likes.

5.1. Description of the Experiments

Firstly, as mentioned before, an analysis was made of the text and of the words of caption. We furthermore identified the words that were more frequently used. All of these words were appropriately filtered by removing those that were not useful for the sentiment analysis, such as prepositions. Additionally, less frequent words were excluded to focus on the most relevant terms. The remaining words were then sorted based on their frequency. Subsequently, another filtering was performed to conduct an analysis that relates them to the sentiment score. Neutral sentiment results have a score of 0, so the objective was to analyze words that can have positive or negative impact on the users. Consequently, when all the relevant words were identified, we performed a further analysis for the number of “likes”. It is possible to see the words which have more likes and if the sentiment they generate is positive or negative.

This part of the analysis was interesting not only to gain insight into what users can convey, but also to understand the impact that followers can have through posts. Furthermore, this allowed us to determine whether the sentiments shared by influencers on social media are reflected in the perceptions of their followers. Eventually, a box plot was plotted that showed the sentiment perceived from the comments of the followers, as showed.

To further relate social media engagement with users’ emotions, we relied on the so-called Jakobson’s theory, which provides a useful framework for analyzing social media engagement [38]. This theory helps to understand how communication strategies impact the interactions between influencers and followers. Consequently, engagement interactions are defined based on the following points:

- Emotional: The type of communication that is adopted is designed to make followers attached to what is being promoted.
- Cognitive: A communication strategy that creates positive–cognitive associations and reinforces brand awareness.
- Conative: A strategy that induces customers to react, such as through engagement in communication.

5.2. Results

In this section, we will describe the results obtained, revealing the results for each of the influencers groups considered.

5.2.1. Fashion Influencer

In Figure 2, we show a box plot for understanding how the caption made by the influencers can impact the users, looking at the comments of the followers.

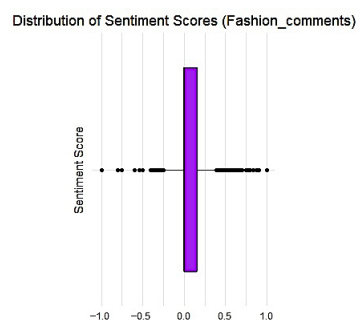


Figure 2. Sentiment score distribution of the comments of the followers for the fashion influencers group.

As can be seen in Figure 2, the box appears to be more skewed towards the right side, suggesting that the distribution of positive comments is more varied compared to the negative ones. This suggests that there are more positive scores, reflecting a positive sentiment, in comparison to negative ones. Positive comments are distributed across a wider range, up to 0.5, whereas negative comments tend to be closer to 0, meaning they are only slightly negative. Regarding the outliers, which are present on both sides of the axis, they indicate that there is no clear predominance of either negative or positive scores in the extreme comments. The engagement calculation takes into account interactions such as “likes” and “comments” per post, calculating the average. Engagement is calculated by summing the total number of likes and comments for each post and then dividing by the number of posts. The number of average likes is 138,884.6 and the number of average comments is 660.0785. The entropy scores distribution is reported in Figure 3.

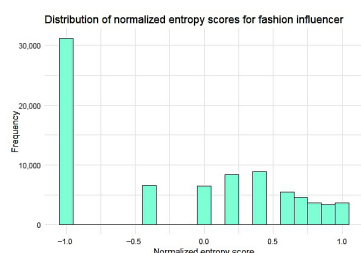


Figure 3. Fashion entropy scores distribution.

The same steps were made for all the other groups. For the fashion influencers, what was obtained from the first filtering of the most used words is shown in Table 4 together with their sentiment score.

Table 4. Top fashion influencers’ words, with relative counts and sentiment scores.

Word	Total_Likes	Sentiment_Score
love	40,3071,965	0.75
life	123,996,708	0
happy	113,063,991	0.75
years	91,582,606	0
work	80,947,489	0.25
always	80,856,713	0
beautiful	76,229,893	0.75
fun	66,097,736	0.75
special	63,789,201	0.8
summer	59,489,500	0
collection	56,858,174	0
ever	53,125,829	0
good	51,236,265	0.75
friends	43,571,541	0
incredible	39,978,179	0.5
hair	39,174,293	0
season	36,660,558	0
favorite	34,766,112	0.75
set	32,714,894	0
ad	30,113,009	0
creative	28,220,260	0.75
paris	27,386,331	0
need	26,077,077	0
great	25,713,584	0.5
spring	24,994,950	0
ready	23,613,361	0.8
director	18,786,516	0.25
shop	15,904,853	0
photographer	10,470,899	0
bag	4,065,846	0

The correlation analysis from the likes and score sentiment is reported in Figure 4.

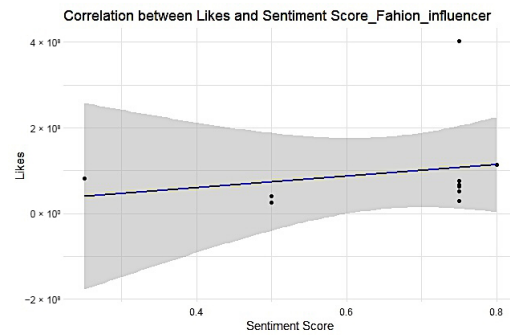


Figure 4. Correlation between the sentiment score of the words and likes.

The plot shown in Figure 4 with a regression line above, that approximates data with a linear relationship, reveals the relationship between likes and sentiment scores of a fashion influencer. The x-axis is labelled as the sentiment score, which varies from around 0.3 to 0.8. The y-axis represents the number of likes (ranging from 0 to around 400 million). We can further see the blue regression line which indicates the relationship that exists between the sentiment score and number of likes. In other words, the sentiment score and number of likes appear to increase together. The confidence interval is represented by the shaded grey areas around the regression line. This shows the potential band in which the actual regression line could be with some degree of certainty.

Further more in Figure 5, the emotions are shown on the y-axis and their frequency is reported on the x-axis. Each emotion is represented by a different colour. The fashion influencers show a high frequency of positive emotion, followed by joy. The negative emotions, such as anger and surprise, have a lower frequency.

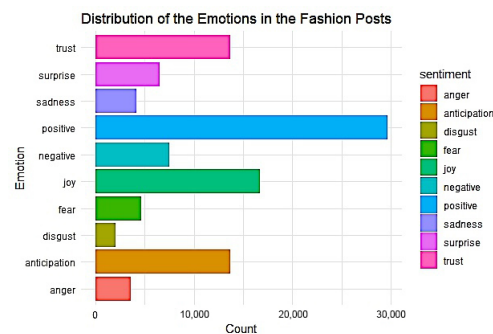


Figure 5. Distributions of the emotions in the Fashion group.

5.2.2. Climate Influencers

Regarding climate influencers, the box plot in Figure 6 shows the level of engagement of their followers in response to the influencers’ posts.

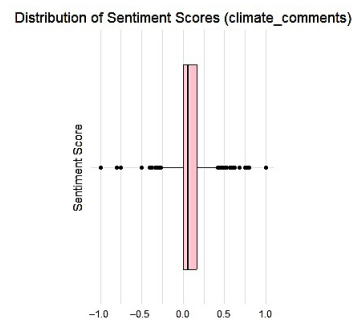


Figure 6. Sentiment score of the comments of the followers for the climate influencers.

It can be seen in Figure 6 that the box plot appears to be slightly extended to the right. This suggests a slight distribution of negative comments on the left side of the median, represented by the vertical black line. While the majority of comments are positive, a slight portion reflects negative sentiment as well. In fact, in this specific case, the negative comments are more than those shown in the previously mentioned Section 5.2.1. Regarding the outliers present on both sides of the axis, they indicate that there is no clear dominance of either negative or positive sentiment in the extreme comments.

The engagement calculation takes into account interactions such as “likes” and “comments” per post mentioned before in Section 5.2.1. In the case of the climate influencers groups, we obtained a number of average likes equal to 19,544.12 and a number of average comments equal to 379.8635.

In Table 5, we report the statistics of words, number of likes, and the sentiment score for the climate influencers groups.

Table 5. Climate words, number of likes, and sentiment scores.

Word	Total_Likes	Sentiment_Score
climate	24,062,497	0
people	17,010,825	0
week	14,837,788	0
world	12,315,440	0
thank	8,700,186	0.5
bio	8,244,309	0
love	8,135,223	0.75
time	7,033,452	0
support	6,871,694	0.5
share	5,996,525	0.5
crisis	5,797,730	-0.75
justice	5,407,179	1
planet	5,361,310	0
book	5,101,078	0
years	5,002,144	0
help	4,973,898	0
even	4,892,008	0
change	4,645,568	0
need	4,536,778	0
community	3,695,966	0.6
work	3,685,095	0.25
action	3,676,430	0.25
water	3,496,459	0
earth	3,407,339	0
made	3,141,681	0
together	2,871,161	0
food	2,483,302	0.4
always	1,574,595	0
sustainability	1,320,829	1
groupsy	0	0

After filtering the data, it is evident that not all of the most frequently used words have a positive sentiment score. For example, the word crisis has a negative score of -0.75 . We further evaluated the correlation between the number of likes and the word sentiment scores. The results are shown in Figure 7.

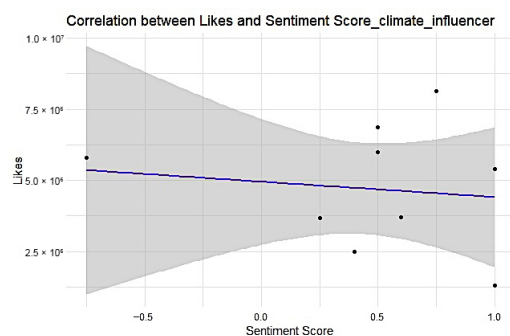


Figure 7. Correlation in the climate influencers group between the sentiment score of the words and likes.

The plot shown in Figure 7, with a regression line above, reveals the relationship between likes and sentiment scores of a climate influencer. On the x-axis, we represented the labelled sentiment score, which varies from around -0.5 to 1 . The y-axis represents the number of likes (from 0 to around 10 million), so a very lower number in respect to the group of fashion influencers shown in Figure 4. We further overlaid a purple regression line which indicates the relationship that exists between the sentiment score and number of likes. This means that the sentiment score does not rise up as the number of likes increases and this could be due to the fact that the negative sentiment from some types of word are relatively high. The shaded grey regions around the regression line are the confidence interval. Then, the result for the entropy is shown in Figure 8.

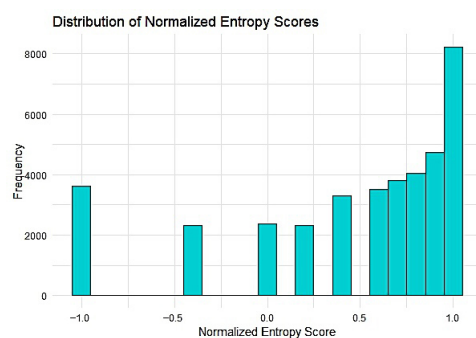


Figure 8. Climate entropy distribution score.

The plot of entropy showed in Figure 8 shows that a value from 0.6 to 1 seems to be more uniformly distributed, and a large spike exists, indicating that a greater proportion of content is unpredictable.

Regarding the emotions in the posts related to this group, the structure is the same as the one shown in Figure 5, where positive emotions are dominant, but at the same time, negative emotions, such as fear and anger, are also relevant. We see this results in Figure 9. In this case, anticipation is notable, but it is not as significant, as reported in the plot of entropy shown in Figure 8.

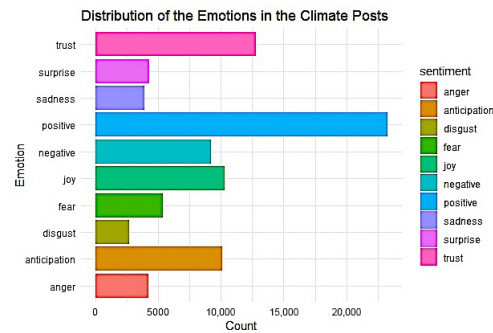


Figure 9. Distributions of the emotions in the climate group.

5.2.3. AI Influencers

In Figure 10, we plot the distribution of the sentiment scores for the AI influencers group.

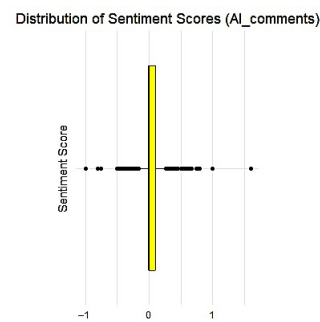


Figure 10. Sentiment score of comments by followers for the AI influencers group.

As shown in Figure 10, the distribution of sentiment scores is concentrated around 0, indicating that the majority of comments have a neutral sentiment. There are some outliers on both the negative (left) and positive (right) sides, but they are relatively few compared to the central concentration.

Before obtaining a sentiment score of these words, several operations of filtering were necessary. This because the language that frequently occurs is based on strictly technical language, which has a neutral sentiment score. From these results, in Table 6, we can understand that filtering the text before the analysis is extremely important. This is because many of these “educational posts” are based on technical language, which would otherwise turn out to be mainly characterized by a neutral sentiment. In Figure 11, we can observe the correlation plots illustrating the relationship between the sentiment scores of the words and the number of likes, indicating that a slightly higher number of likes is associated with a more positive sentiment score.

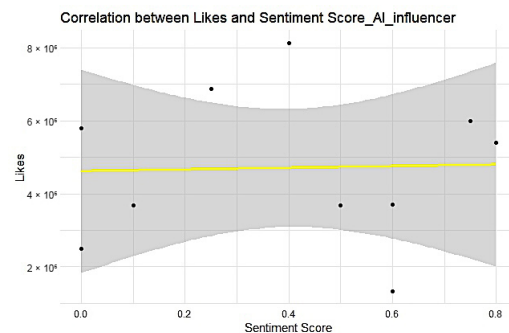


Figure 11. Correlation between the sentiment score of the words and likes.

Table 6. Number of likes and sentiment scores for the AI influencers.

Word	Total_Likes	Sentiment_Score
ai	16,262,047	0
bio	14,337,158	0
world	8,662,443	0
video	8,467,508	0
science	7,789,124	0
people	7,536,562	0
tech	7,058,360	0
image	5,660,375	0
love	5,294,934	0.75
learn	4,822,838	0.8
technology	4,555,970	0.1
chatgpt	4,360,019	0
work	4,337,798	0.25
light	3,703,840	0
space	3,353,432	0
need	3,258,700	0
human	3,110,915	0
openai	2,870,174	0
research	2,385,034	0
experience	2,379,570	0
artificialintelligence	1,917,340	0
data	1,422,167	0
episode	1,383,397	0
google	1,216,349	0
nvidia	867,356	0
learning	697,344	0.8
app	647,389	0
startup	640,531	0
art	603,972	0.6
business	592,498	0
digital	456,956	0
coding	308,073	0
machinelearning	251,693	0
kdnuggets	32,434	0

The plot in the Figure 11 shows a weak correlation between “Likes” and “Sentiment Score”. The yellow regression line is almost flat, likely due to the fact that most of the terms used have a neutral sentiment, resulting in a sentiment score close to 0. This demonstrates that the remaining words with either a positive or negative sentiment score are relatively few in comparison, and thus do not significantly affect the overall distribution compared to the neutral terms. Some points, like the one with 8 million likes, are outliers, but they do not have a clear effect on the general trend. The calculation of engagement explained in Section 3.4 follows the same procedure as in Sections 5.2.1 and 5.2.2, and was performed. In this AI influencers case, we counted an average number of likes of 6760.751 and an average number of comments of 92.35321. The engagement is calculated by summing the total number of likes and comments for each post and then dividing by the number of posts. In Figure 12, we can see the entropy distribution plot obtained.

The bar chart illustrates that the distribution of normalized entropy scores, in which the x-axis represents the normalized scores, have a range that rises from -1.0 to 1.0 , while the y-axis represents the frequency, indicating the frequency of occurrences for each bin of normalized entropy scores. There is a notable peak at -1.0 , suggesting a very high frequency of samples with normalized entropy scores at that extreme showing unpredictability in the posted content, but the remainder of the distribution (from -0.5 to 1.0) displays a more balanced spread, with slightly higher frequencies grouped around 0.5 , showing a slight predictability. In this group, positive emotions are more frequent,

as shown in Figure 13, but in contrast, fear also has a significant presence. The count for anticipation is around 30,000, which corresponds to what we can see in the entropy plot in Figure 12, while anger and surprise are the least relevant.

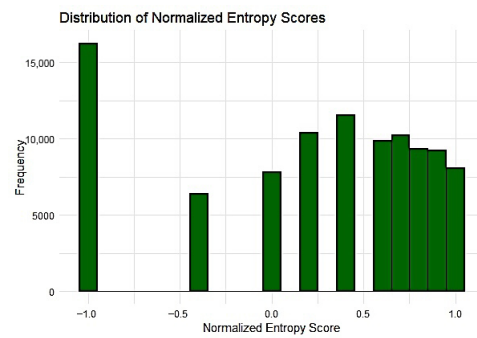


Figure 12. AI entropy distribution.

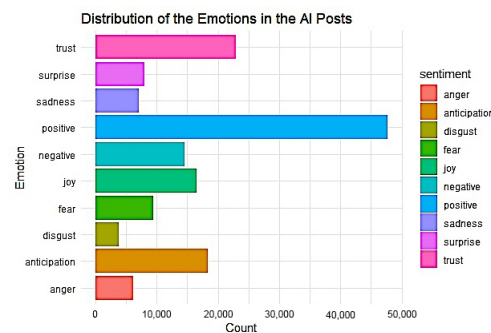


Figure 13. Distribution of the emotions in the AI posts.

5.2.4. Journalist Influencers

We report here the analysis performed on the journalists influencers group.

As shown in Figure 14, the median, represented by the vertical black line, is not exactly in the center of the box but slightly shifted to the left. This indicates that 50% of the sentiment scores have a tendency towards more positive values compared to negative scores. In contrast, in the box plot of the “Sentiment score of the comments of the followers fashion” in Figure 2, the negative scores were significantly lower. There are some outlier values on both the negative and positive sides, but they appear to be few compared to the central concentration. For the engagement calculation, the process is the same explained in Section 5.2.1. The results are reported with an average of likes of 8538.407 and an average number of comments of 364.305. From Table 7, we can see the words, their corresponding likes, and sentiment scores.

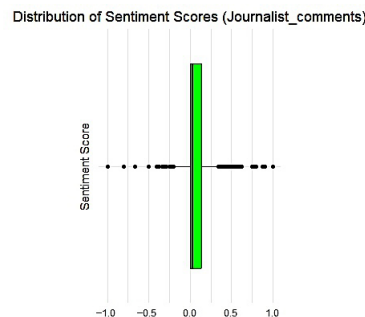


Figure 14. Sentiment score of the comments of the followers for the journalist influencers.

Table 7. Journalists’ most used words, number of likes, and relative sentiment scores.

Word	Total_Likes	Sentiment_Score
night	3,996,085	0
family	3,080,330	0
bio	2,616,037	0
today	2,518,991	0
happy	2,124,104	0.75
friend	2,077,665	0.8
people	2,018,108	0
year	1,981,505	0
watch	1,839,997	0
love	1,789,909	0.75
life	1,529,240	0
home	1,516,266	0
morning	1,410,616	0
team	1,337,031	0
great	1,268,229	0.5
every	1,087,248	0
special	967,102	0.8
women	966,229	0
back	961,378	0
world	908,839	0
story	827,618	0
news	810,657	0
join	742,187	0.25
thanks	664,167	0.6
conversation	66,0761	0
good	660,710	0.75
book	586,423	0
down	527,751	0
president	518,171	0
work	471,233	0.25

The plot for the correlation is shown in Figure 15.

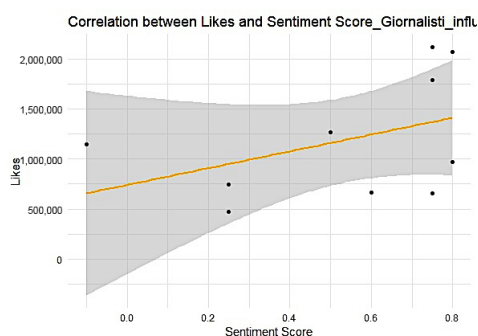


Figure 15. Correlation between the sentiment score of the words and likes.

The plot in Figure 15 explores how the number of “likes” correlates with the “sentiment scores” of the content made by journalist influencers. The x-axis represents the sentiment scores and varies from 0 to about 0.8, showing that all analyzed contents have a neutral to positive sentiment. The y-axis represents the likes range from 0 to over 2000.000.

Each black dot represents a data point, linking a specific comment’s sentiment score to the number of likes it received. The orange regression line illustrates the trend in the relationship between sentiment scores and likes. The line slopes upward, indicating that as sentiment scores rise, the number of likes tends to increase as well. The gray shaded area represents the confidence interval surrounding the regression line. There is a positive correlation between sentiment scores and likes. Generally, higher sentiment scores (more

positive terms) are linked to a greater number of likes. Some comments with similar sentiment scores received vastly different numbers of likes. Subsequently, a graphic of the entropy is illustrated in Figure 14.

Figure 16 shows that the distribution indicates that many comments are either very predictable (low entropy) or exhibit moderate to high unpredictability (higher entropy scores), while fewer comments fall into the mid-range entropy category.

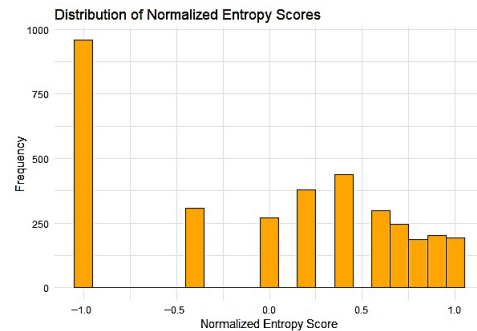


Figure 16. Journalist entropy.

In the bar plot shown in Figure 17, it is possible to see that positive emotions are frequent, with less frequency for sadness and disgust. Anticipation is also noteworthy, but it is not as frequent as in the groups of fashion influencers and AI influencers, as shown in their respective Sections 5.2.1 and 5.2.3.

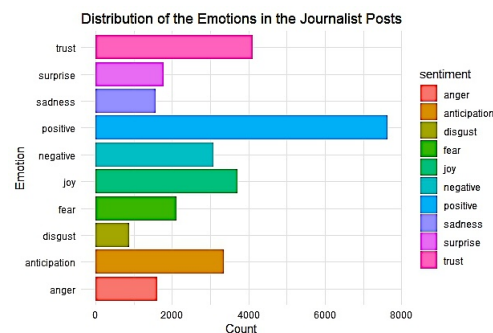


Figure 17. Distribution of the emotions in the journalist posts.

6. Discussion

Although a common thought might associate people's interest with simple content, what emerges from the plots is that higher engagement seems more closely related to the positivity of the language used. Additionally, the topics discussed using that type of language also play a role. This is demonstrated by the sentiment scores associated with the language, and this result explains why there is a significant concentration of followers in some groups compared to others. Thinking about the group of the fashion influencers, after filtration operations what was clear that there was no need to do several operations to obtain the words where the sentiment score was non-neutral, and all the most common words that are used have a positive sentiment score. Compared to the other datasets, it is straightforward that the datasets of climate and journalists adopt common words that have a negative impact on the followers, while for the AI influencer, what happens is that the filtration operations required more steps since the language used was of a technical nature and so the language that is used is purely neutral. So this might explain why there are not many followers, because there is no use of a strong collection of words with a positive sentiment score capable of attracting many people.

Analyzing the correlation of the graphs, it was possible to see that starting from a sentiment score (discussed in Section 5.2) of 0.4, the likes are greater in respect to the words that have a sentiment score less than 0.4.

From the results of the engagement, it is possible to see that greater engagement obtained from the fourth group is from the group of fashion influencers, explaining the correlation of the words with a positive sentiment score. The groups of journalists and climate show less likes in respect to the fashion influencer but with a great average of comments and this might be for the sentiment that is generated from the text.

Regarding the comments, it emerges that the reaction to influencers' captions reflects the same sentiment as the users' comments, showing that more positive captions tend to generate more positive comments. Vice versa, the more negative captions tend to generate more negative comments.

In conclusion, the dataset of AI has an average of comments and likes that is not really high and this could be a consequence of the huge number of neutral words adopted, not capable of capturing a great attention of followers.

In relation to the entropy, it is evident that less entropy is concentrated in the fashion dataset, suggesting that the text and arguments treated are more coherent, predictable, and not conflicting. In fact, this is also detectable from the distribution of emotions in the fashion dataset, which shows high values of anticipation and positivity. The other datasets exhibit higher entropy, leading to a loss of curiosity in the reader. It is a paradox, but considering the climate situation, the unpredictability of the news can make readers anxious about future scenarios. This result indicates that people prefer to follow clear and predictable content. Although the AI group also shows higher entropy compared to the fashion group, it has a significant value of entropy around -1 , as shown in Figure 12, indicating higher predictability. On the other hand, negative sentiments such as fear are also higher in the AI group compared to the fashion influencer. This underscores that predictability and positivity are key factors in gaining as much attention as possible.

7. Conclusions

The research question addressed is related to the significant influence of social media on consumer behaviour and how the opinions and language used by influencers shape the minds of their followers. The research explores why certain groups of influencers are more attractive than others, identifying the key factors behind this distinction. In conducting these analyses, some challenges arose due to changes in Instagram's policies in 2018. These changes made it more difficult to access information from Instagram profiles, complicating the management of interactions. The study focused on four different groups of Instagram influencers, referred to as groups, as they grouped influencers by similar interests. The groups considered were fashion influencers, journalist influencers, AI influencers, and climate influencers. The analysis examined the language used by these influencers in their posted content and the emotions evoked. The novel dataset collected, InFluEmo, included data such as the number of posts, comments, likes, profiles, and other types of observations. The analysis focused on understanding which language strategies are the most successful in attracting more people and increasing interaction between influencers and their followers. To uncover these patterns, sentiment analysis and entropy calculations were performed. Based on the results in Section 5.2, the study provides valuable insights on how to improve communication strategies to reach a larger audience, highlighting the importance of language in fostering trust and authenticity. Similar analyses performed were aimed at explaining the success of influencers and understanding the effects of social media on users. However, no attention was given to why most users were attracted to specific topics and less to others. The datasets provided in this study were properly elaborated to

perform the analysis. To the best of our knowledge, no public datasets exist of this type, based on specific research to clearly distinguish between each dataset according to the topics they addressed and taking into account the most influential people as influencers for each of these datasets.

In addition, the study was conducted by analyzing the main factors behind the success of influencers, interpreting the structure of the text used in the content posted by influencers, and assessing the impact they have on users by examining users' comments and likes. There is one important limitation to be highlighted. The possible deleted comments and modification to posts are not available in the dataset retrieved. Some negative comments could have been deleted, and not present in our analysis, introducing some sort of post-processing bias in the results obtained which could be taken into account. In further studies and data collection, if this information becomes available, it would be extremely interesting to be analyzed. On top of the above limitations, some further improvements on the number of hashtags and choice of the influencers could be made, in order to further improve the dataset and so on. The results showed a positive correlation between the number of likes and the sentiment they provoke. The more positive the sentiment generated, the more likes are associated with the content. It is also interesting to note that as entropy increases, engagement decreases. This is due to the fact that users tend to prefer predictable content with less complexity, which explains why, for example, the dataset of fashion influencers has more engagement compared to AI influencers, climate influencers, and journalist influencers, who spread a higher level of knowledge. This may be surprising, as one might expect people to be more attracted to richer content rather than to simpler content. Additionally, these results can be related to the study in [40], which discusses the poor results in academic performance that can be associated to a lack of attention that social media may promote. In fact, the most attractive content seems to be the one that requires less attention. Future work might include the extension of the dataset to further groups, as well as the investigation of possible cross-effects of success between Instagram and other social networks where the influencers might have connected accounts.

Author Contributions: Conceptualization, G.M.D. and C.F.S.; methodology, G.M.D.; software, C.F.S.; validation, G.M.D. and C.F.S.; formal analysis G.M.D. and C.F.S. investigation, C.F.S.; resources, C.F.S.; writing—original draft preparation, G.M.D. and C.F.S.; writing—review and editing, G.M.D. and C.F.S.; visualization, G.M.D. and C.F.S.; supervision, G.M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: No ethical approval was required and only public posts were retrieved.

Data Availability Statement: The dataset of InfluEmo is freely available at https://drive.google.com/drive/folders/19M04M7DtLzP_cE-6KlIk8Q4Ze1pFuPyN?usp=sharing (accessed on 1 February 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cornell, S.; Brander, R.; Peden, A. Selfie-related incidents: Narrative review and media content analysis. *J. Med. Internet Res.* **2023**, *25*, e47202.
2. Salazar-Fernández, C.; Baeza-Rivera, M.J.; Manríquez-Robles, D.; Salinas-Oñate, N.; Sallam, M. From conspiracy to hesitancy: The longitudinal impact of COVID-19 vaccine conspiracy theories on perceived vaccine effectiveness. *Vaccines* **2023**, *11*, 1150.
3. Casali, A. Social Media Marketing: Cos'è, Come si fa e Quali Vantaggi Porta alla tua Azienda. Available online: <https://www.digital4.biz/marketing/social-media-marketing-cos-e-come-farlo-vantaggi/> (accessed on 3 May 2023).
4. Audrezet, A.; de Kerviler, G.; Moulard, J. Authenticity under threat: When social media influencers need to go beyond self-presentation. *J. Bus. Res.* **2018**, *77*, 154–170.

5. Boerman, S. The effects of influencer marketing on online engagement. *Comput. Hum. Behav.* **2020**, *103*, 1–13.
6. Ki, C.; Kim, Y. The mechanism by which social media influencers persuade consumers: The role of influencer credibility and self-disclosure. *J. Retail. Consum. Serv.* **2019**, *49*, 562–570.
7. Ki, C.; Kim, Y. Influencer marketing on Instagram: How sponsorship disclosure affects consumer’s trust and purchase intentions. *J. Bus. Res.* **2019**, *101*, 525–536.
8. Lou, C.; Yuan, S. Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *J. Interact. Advert.* **2019**, *19*, 58–73.
9. Lou, C.; Yuan, S. Influencer marketing on social media: The impact of opinion leadership, parasocial interaction, and self-disclosure on consumer’s trust and purchase intention. *J. Soc. Media Soc.* **2019**, *8*, 1–19.
10. Mediakix. *The Influencer Marketing Report*; Mediakix: Santa Monica, CA, USA, 2019.
11. De Veirman, M.; Cauberghe, V.; Hudders, L. Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude. *Int. J. Advert.* **2017**, *36*, 798–828.
12. Djafarova, E.; Rushworth, C. Exploring the credibility of online celebrities’ Instagram profiles in influencing the purchase decisions of young female users. *Comput. Hum. Behav.* **2017**, *68*, 1–7.
13. Schouten, A.; Janssen, L.; Verspaget, M. Celebrity vs. Influencer endorsements in advertising: The role of identification, credibility, and Product-Endorser fit. *Int. J. Advert.* **2020**, *39*, 258–281.
14. Harrigan, P.; Daly, T.M.; Coussement, K.; Lee, J.A.; Soutar, G.N.; Evers, U. Identifying influencers on social media. *Int. J. Inf. Manag.* **2021**, *56*, 102246. <https://doi.org/10.1016/j.ijinfomgt.2020.102246>.
15. Pucci, F.; Fedele, P.; Dimitri, G.M. Speech emotion recognition with artificial intelligence for contact tracing in the COVID-19 pandemic. *Cogn. Comput. Syst.* **2023**, *5*, 71–85.
16. Florindi, F.; Fedele, P.; Dimitri, G.M. A novel solution for the development of a sentimental analysis chatbot integrating ChatGPT. *Pers. Ubiquitous Comput.* **2024**, *28*, 947–960.
17. Koolagudi, S.G.; Rao, K.S. Emotion recognition from speech: A review. *Int. J. Speech Technol.* **2012**, *15*, 99–117.
18. Khare, S.K.; Blanes-Vidal, V.; Nadimi, E.S.; Acharya, U.R. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Inf. Fusion* **2024**, *102*, 102019.
19. Cruz, J.P.O.; Passaro, L.C.; Almeida, J.M.; Marques-Neto, H.T. Characterizing fashion influencers’ behavior on Instagram. *Soc. Netw. Anal. Min.* **2024**, *14*, 147.
20. Zhan, M.; Tu, R.; Yu, Q. Understanding Readers: Conducting Sentiment Analysis of Instagram Captions. In Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence, Shenzhen, China, 8–10 December 2018.
21. MacKay, M.; Ford, C.; Colangeli, T.; Gillis, D.; McWhirter, J.E.; Papadopoulos, A. A content analysis of Canadian influencer crisis messages on Instagram and the public’s response during COVID-19. *BMC Public Health* **2022**, *22*, 673.
22. Bashari, B.; Fazl-Ersi, E. Influential post identification on Instagram through caption and hashtag analysis. *Meas. Control* **2020**, *53*, 409–415.
23. Nandwani, P.; Verma, R. A review on sentiment analysis and emotion detection from text. *Soc. Netw. Anal. Min.* **2021**, *11*, 81.
24. Ranjan, M.; Tiwari, S.; Sattar, A.M.; Tatkar, N.S. A New Approach for Carrying Out Sentiment Analysis of Social Media Comments Using Natural Language Processing. *Eng. Proc.* **2023**, *59*, 181.
25. Israa Khalaf Salman Al-Tameemi, Mohammad-Reza Feizi-Derakhshi, S.P.M.A. A Comprehensive Review of Visual-Textual Sentiment Analysis from Social Media Networks. *J. Comput. Soc. Sci.* **2023**, *7*, 2767–2838.
26. Fatouros, G.; Soldatos, J.; Kouroumalis, K.; Makridakis, G.; Kyriazis, D. Transforming sentiment analysis in the financial domain with ChatGPT. *Mach. Learn. Appl.* **2023**, *14*, 100508.
27. Sande, N.; Adeniyi, I.; Akinkunmi, A. Social Media Sentiment Analysis: A Comprehensive Analysis. 2024. Available online: https://www.researchgate.net/publication/378150391_Social_Media_Sentiment_Analysis_A_Comprehensive_Analysis/link/65ca4fc47900745497815c7a/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFuZSI6InB1YmxpY2F0aW9uIn19 (accessed on 15 January 2026).
28. Gu, S.; Heerdink, M.W.; van Kleef, G.A. How social media influencers’ emotional expressions shape their popularity: The role of authenticity and appropriateness. *Cogn. Emot.* **2025**, 1–23.
29. Azhar, K.A.; Wel, C.A.C.; Ab Hamid, S.N. “They Post, I Scroll, I Envy, I Buy”—How Social Media Influencers Shape Materialistic Values and Consumer Behavior Among Young Adults in Malaysia. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 172.
30. Alsalloum, A.A.; Gainous, J. Female social media influencers in the Arab Gulf: shaping marketing and advertising on Instagram. *Front. Commun.* **2025**, *10*, 1518386.
31. Rstudio. Index of /bin/windows/base/old/4.3.1. Available online: <https://posit.co/downloads/> (accessed on 16 June 2023).
32. Wikipedia. Instagram. Available online: <https://it.wikipedia.org/wiki/Instagram> (accessed on 14 September 2024).
33. Matthew Jockers. Introduction to the Syuzhet Package. Available online: <https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html> (accessed on 12 August 2023).

34. Jockers, M.L. `get_nrc_sentiment`: Get Emotions and Valence from NRC Dictionary. Available online: <https://www.rdocumentation.org/packages/syuzhet/> (accessed on 11 August 2023).
35. Wang, Y.; Huang, G.; Li, M. Automatically Constructing a Fine-Grained Sentiment Lexicon for Sentiment Analysis. *Cogn. Comput.* **2023**, *15*, 254–271.
36. Mohammad, S.M.; Turney, P.D. Crowdsourcing a Word–Emotion Association Lexicon. *Comput. Intell.* **2012**, *29*, 436–465.
37. Wikipedia. Robert Plutchik. 2024.
38. Ilaria Cazziol. Engagement: Cos'è, Perché è Importante, Come Calcolarlo e Migliorarlo. Available online: <https://www.shopify.com/it/blog/engagement> (accessed on 8 December 2022).
39. Apify. Instagram Scraper. Available online: <https://apify.com/> (accessed on November 2021).
40. Luca Braghieri, R.L.; Makarin, A. Social Media and Mental Health. *Am. Econ. Rev.* **2022**, *112*, 3660–3693.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.