sustainability

MDPI

Article

AI-Driven Social Media Analytics for Assessing Climate Change Perceptions and Supporting Adaptation and Sustainability Policies

Mehmet Kayakuş [1], Onder Kabas [2,*] and Georgiana Moiceanu [3,*]

[1] Department of Management Information Systems, Faculty of Manavgat Social Sciences and Humanities, Akdeniz University, 07070 Antalya, Türkiye; mehmetkayakus@akdeniz.edu.tr
[2] Department of Machine, Technical Science Vocational School, Akdeniz University, 07070 Antalya, Türkiye
[3] Department of Entrepreneurship and Management, Faculty of Entrepreneurship, Business Engineering and Management, National University of Science and Technology Politehnica Bucharest, 060042 Bucharest, Romania
* Correspondence: okabas@akdeniz.edu.tr (O.K.); georgiana.moiceanu@upb.ro (G.M.); Tel.: +40-745-201-365 (G.M.)

Abstract

This study examines public perceptions and discourse on climate change using artificial intelligence (AI)-based analysis of social media data, with implications for climate adaptation and sustainability policy. A dataset of 29,576 posts from the X platform (December 2025) was analyzed through an integrated framework combining text mining, TF-IDF-based word analysis, deep learning-based sentiment analysis, and Latent Dirichlet Allocation (LDA) topic modelling. The findings reveal that climate change discourse is predominantly characterized by negative sentiment, reflecting high levels of concern, perceived risk, and urgency, while positive content emphasizes awareness, solutions, and collective action. Topic modelling identifies three main themes: skepticism shaped by daily weather experiences, scientific and policy-oriented climate debates, and discussions on carbon emissions and human impact. These results demonstrate that social media serves not only as a space for emotional expression but also as a dynamic platform for information exchange and public opinion formation. From an adaptation perspective, AI-driven social media analytics provide valuable insights into public risk perception, misinformation patterns, and knowledge gaps, supporting evidence-based climate communication and policy development.

Keywords: climate change; climate adaptation; sustainability; social media analytics; artificial intelligence; sentiment analysis; topic modelling; risk perception; climate communication; environmental policy

Check for updates

Academic Editors: Cheng Li, Fei Zhang, Mou Leong Tan and Kwok Pan Chun

Received: 15 April 2026
Revised: 8 May 2026
Accepted: 10 May 2026
Published: 13 May 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) license.

1. Introduction

Climate change has emerged as one of the most critical global challenges of the twenty-first century, generating profound environmental, social, economic, and political consequences. The increasing concentration of greenhouse gases in the atmosphere has led to rising global temperatures, extreme weather events, sea-level rise, and ecosystem degradation, positioning climate change as a multidimensional risk that threatens both natural and human systems. Beyond its environmental impacts, climate change poses significant challenges to food security, public health, economic stability, and global governance, thereby necessitating urgent and coordinated responses within the broader framework of

sustainable development. In this context, reducing uncertainty and strengthening climate adaptation strategies have become central priorities for policymakers, researchers, and international organizations.

In recent years, climate change has increasingly been addressed not only as an environmental issue but also as a complex socio-technical and policy-driven challenge that requires adaptive, data-driven, and participatory approaches. Effective climate adaptation and risk management depend not only on scientific data but also on how societies perceive, interpret, and respond to climate-related risks. Public perception plays a decisive role in shaping behavioral change, policy acceptance, and the effectiveness of sustainability-oriented interventions. Therefore, understanding how climate change is socially constructed and communicated has become a critical component of environmental governance and climate policy design.

Parallel to these developments, the rapid expansion of digital communication technologies has fundamentally transformed the ways in which individuals access information, express opinions, and engage with global challenges. Social media platforms have evolved into dynamic digital public spheres where climate change is continuously discussed, interpreted, and contested through user-generated content. Platforms such as X (formerly Twitter) provide real-time, large-scale, and spontaneous data that reflect public awareness, emotional responses, and risk perceptions related to climate change. Unlike traditional survey-based methods, social media data enables continuous monitoring of public discourse, offering valuable insights for identifying emerging concerns, misinformation patterns, and knowledge gaps that are directly relevant to climate adaptation and policy development.

From a theoretical perspective, this study is grounded in the intersection of climate communication, public perception theory, and the concept of the digital public sphere. Climate communication research emphasizes how environmental issues are framed, interpreted, and disseminated across communication channels, influencing public understanding and behavioral responses. Public perception theory highlights the role of individual and collective interpretations in shaping attitudes toward complex and uncertain risks such as climate change. In addition, social media platforms function as digital public spheres where diverse actors interact, exchange information, and construct shared meanings around sustainability-related issues. Within this framework, analyzing social media discourse provides critical insights into how climate risks are perceived, debated, and translated into potential societal responses.

Despite the growing body of literature on climate change communication, several important gaps remain. Many existing studies tend to focus on either sentiment analysis or topic modelling independently, often relying on relatively limited datasets or narrow analytical frameworks. Moreover, the connection between social media discourse and climate adaptation, risk perception, and sustainability policy development remains insufficiently explored. This fragmented approach limits the ability to capture the multidimensional nature of climate-related discussions and reduces the practical relevance of findings for policy and decision-making processes. Therefore, there is a need for integrative analytical frameworks that combine multiple artificial intelligence (AI)-based techniques to provide a more comprehensive and policy-relevant understanding of public discourse. Existing studies have successfully demonstrated that sentiment analysis can reveal emotional tendencies in climate discourse, while topic modelling can identify dominant thematic structures within online discussions. However, these approaches often remain analytically isolated and therefore fail to explain how emotional reactions, thematic narratives, and conceptual structures interact simultaneously within digital climate communication. In addition, prior research has rarely linked these discourse dynamics directly to climate adaptation, sustainability

communication, and environmental governance processes. Consequently, the literature still lacks an integrated framework capable of explaining the multidimensional relationships between public emotions, thematic discourse patterns, and sustainability-related policy implications.

Addressing these gaps, this study aims to analyze social media discourse on global warming through an integrated AI-driven framework, with a particular focus on its implications for climate adaptation and sustainability policy. By combining text mining, deep learning-based sentiment analysis, and Latent Dirichlet Allocation (LDA)-based topic modelling, the study offers a multidimensional examination of both emotional and thematic structures of climate discourse on the X platform. In addition, visualization techniques such as keyword co-occurrence networks and density mapping are employed to reveal the structural relationships between concepts and enhance interpretability.

The study seeks to answer the following research questions: (i) What are the dominant emotional orientations expressed in social media posts related to climate change? (ii) Which themes and topics structure online climate discourse? (iii) How do emotional and thematic dimensions interact to shape the overall structure of sustainability-related discussions? By addressing these questions, the study makes two main contributions. First, it provides a methodological contribution by demonstrating the effectiveness of integrating multiple AI-based techniques for analyzing large-scale, unstructured social media data. Second, it offers practical and policy-relevant insights by revealing how public perceptions, risk awareness, and discourse patterns can inform climate adaptation strategies, improve sustainability communication, and support evidence-based environmental decision-making.

Unlike studies relying solely on sentiment analysis or topic modeling, the integrated framework adopted in this study enables simultaneous examination of emotional orientation, thematic structures, and conceptual relationships within climate discourse. The combination of text mining, deep learning-based sentiment analysis, topic modeling, and network visualization provides complementary insights that cannot be achieved through a single analytical technique alone. For example, while sentiment analysis identifies emotional tendencies, topic modeling reveals the underlying thematic dimensions, and network analysis demonstrates how these themes are structurally interconnected within online discourse. This multidimensional analytical framework therefore allows for a more comprehensive interpretation of climate-related public perceptions and sustainability communication dynamics.

2. Theoretical Framework

In recent years, rapid advancements in artificial intelligence (AI), natural language processing (NLP), and machine learning techniques have significantly enhanced the capacity to analyze large-scale social media data. These technological developments have enabled researchers to process, classify, and interpret massive volumes of unstructured textual content in an automated and systematic manner. Social media platforms, where individuals actively share opinions, emotions, and experiences, have thus emerged as valuable data sources for understanding public perceptions of complex societal issues such as global warming and climate change. Consequently, AI-driven analytical approaches have become increasingly prominent in examining how climate-related topics are discussed and perceived in digital environments.

One of the pioneering studies in this domain was conducted by Alkhatib et al. who analyzed the sentiments of opinion leaders on X (Twitter) regarding global warming using artificial intelligence-based methods [1]. Their research demonstrated that AI techniques could effectively monitor real-time fluctuations in public sentiment and capture dynamic changes in climate-related discourse. Similarly, El Barachi et al. developed a real-time

sentiment analysis framework to track public opinion on climate change, revealing that social media platforms provide a powerful mechanism for observing evolving societal attitudes [2]. These studies underline the importance of social media as a strategic resource for measuring and interpreting public responses to environmental issues.

Several studies have specifically focused on sentiment analysis to explore emotional reactions to climate change. Loureiro and Alló examined social media data from the United Kingdom and Spain and found that climate change and energy-related discussions evoke strong emotional responses among users [3]. Their findings indicated that negative emotions such as anxiety and frustration were particularly dominant in online climate debates. In a similar vein, Ghosh et al. and Rahman and Fitriyah applied NLP and machine learning techniques to classify user perceptions of climate change, concluding that public discourse is largely shaped by concern, criticism, and skepticism [4,5]. These results collectively highlight that global warming is perceived as a highly sensitive and emotionally charged issue within social media environments.

Beyond sentiment analysis, another important line of research has employed topic modelling techniques to investigate the thematic structure of climate change discussions. Kaushal et al. conducted a comprehensive topic modelling analysis and revealed that social media conversations on climate change revolve around multiple interconnected themes, including environmental risks, energy policies [6], economic impacts, and political processes. Likewise, Gokcimen and Das analyzed social media and blog data using topic modelling and demonstrated that climate change is discussed not only as an environmental problem but also as a social, economic, and political challenge [7]. These studies emphasize that online climate discourse is multi-dimensional and cannot be fully understood through sentiment analysis alone.

Structural characteristics of social media discourse have also been examined in several studies. Al-Rawi et al. investigated climate change discussions on X and showed that both human users and automated bot accounts play a role in shaping digital narratives [8]. Their findings suggest that social media debates on climate issues are hybrid in nature and may be influenced by strategic information campaigns. Furthermore, Lineman et al. and Li et al. analyzed the linguistic framing of climate-related terminology and demonstrated that the use of expressions such as “global warming” and “climate change” can generate different perceptual and political responses [9,10]. These studies highlight that language choice and conceptual framing significantly influence how climate issues are interpreted by the public.

The development of comprehensive datasets has also contributed to the advancement of research in this field. Effrosynidis et al. created a large-scale climate change-focused X dataset, providing researchers with a valuable methodological resource for conducting systematic and replicable analyses [11]. Such datasets facilitate comparative studies and enhance the reliability of findings in climate communication research. In a similar context, Du et al. and Willaert et al. emphasized that social media functions as an effective observatory for monitoring public opinion on environmental matters and that AI-based models can provide strategic insights for policymakers and environmental organizations [12,13].

Another important dimension of the literature concerns the role of influencers and alternative communication styles in climate discourse. Ballestar et al. analyzed the communication strategies of climate influencers on social media and demonstrated that digital opinion leaders can play a crucial role in shaping environmental awareness [14]. Moreover, Eslen-Ziya explored the use of humor and sarcasm in climate change discussions and argued that such rhetorical strategies serve as emotional coping mechanisms and alternative forms of political expression [15]. These findings indicate that social media discourse on global warming is not only rational and information-driven but also deeply embedded in cultural and emotional contexts.

More recent studies have further expanded the methodological toolkit for analyzing climate-related social media content. Amangeldi et al. applied AI-based techniques to evaluate sentiment and emotional intensity in environmental posts, while Chen et al. examined the influence of social bots on climate discussions, highlighting the potential for manipulation and polarization [16,17]. Additionally, Duong et al. employed neuromyotonic AI approaches to analyze public opinions about wildfires and climate change, demonstrating that advanced AI models can provide deeper contextual understanding compared to traditional methods [18,19].

Overall, the existing literature clearly demonstrates that social media discourse on global warming is highly complex, dynamic, and multi-layered. Research has shown that online climate discussions encompass a wide range of emotional reactions, thematic orientations, and communicative strategies. However, despite these valuable contributions, many studies remain limited in scope. A significant portion of previous research has focused either on sentiment analysis or on topic modelling alone, without integrating these approaches into a unified analytical framework. Furthermore, numerous studies rely on relatively small data sets or narrow time frames, which restricts the generalizability of their findings.

The present study seeks to address these limitations by adopting a comprehensive and integrated methodological approach. Unlike many previous works, this research combines text mining, deep learning-based sentiment analysis, and LDA-based topic modelling to examine social media discourse on global warming from multiple perspectives. In addition, advanced visualization techniques such as keyword co-occurrence networks and density visualizations are employed to provide a clearer understanding of the conceptual structure of online discussions. By analyzing both the emotional and thematic dimensions of social media content, the study offers a more holistic interpretation of how global warming is perceived and debated in digital environments.

In this regard, the study makes two primary contributions to literature. First, it provides a methodological contribution by demonstrating the effectiveness of integrating multiple AI-based techniques for analyzing large-scale social media data. Second, it offers an empirical contribution by presenting a detailed and up-to-date examination of public perceptions of global warming based on recent social media content. Through this dual contribution, the research aims to enhance academic understanding of climate communication while also providing practical insights for policymakers, environmental organizations, and communication professionals.

3. Materials and Methods

This study adopts a quantitative research design based on artificial intelligence-driven text analytics to examine public perceptions of global warming on social media. Beyond its methodological contribution, this approach also provides important implications for sustainability research. Understanding public attitudes, concerns, and discourse patterns related to global warming is essential for developing effective sustainability policies and communication strategies. In this context, AI-based text analytics enables the large-scale, real-time monitoring of sustainability-related public opinion, providing helpful guidance to policymakers, environmental organizations, and decision-makers. By integrating multiple analytical techniques, this study contributes to the growing field of data-driven sustainability research and supports evidence-based approaches to climate communication and environmental governance. The adoption of a quantitative AI-driven text analytics approach is theoretically grounded in the assumption that large-scale digital discourse reflects collective patterns of perception, emotional orientation, and public interpretation regarding climate-related risks. Unlike traditional survey-based approaches, AI-supported text

analytics enables continuous, large-scale, and real-time observation of naturally occurring public discourse. This methodological perspective aligns with digital public sphere theory and climate communication research, both of which emphasize that online interactions constitute an important arena for the social construction of environmental meaning and public risk perception.

The overall methodological approach follows a multi-stage analytical process. First, a large dataset of tweets related to global warming were collected. Second, the raw textual data underwent extensive preprocessing to ensure analytical reliability. Third, sentiment analysis was conducted to determine the emotional orientation of the posts. Finally, topic modelling and network-based visualization techniques were applied to uncover the underlying thematic structure of the discourse. This integrated design allows for a holistic understanding of how global warming is discussed and perceived in online environments.

3.1. Data Collection

The dataset used in this research was collected from the social media platform X (formerly Twitter), which has been widely recognized as a valuable source for real-time public opinion analysis [1,2]. Data collection was carried out during December 2025, a period selected due to the high intensity of global discussions on climate change, particularly in relation to international environmental policies and extreme weather events. December 2025 was specifically selected because this period coincided with intensified global climate discussions associated with international environmental policy debates, climate-related extreme weather events, and increased online public engagement regarding sustainability issues. However, this temporal concentration may also have influenced the emotional intensity and thematic distribution of the discourse captured in the dataset.

Tweets were retrieved using a set of predefined keywords associated with global warming and climate change. The primary search terms included expressions such as "global warming," "climate change," "carbon emissions," "greenhouse gases," and "climate crisis." The selection of these keywords was informed by previous studies that emphasized the importance of terminology in shaping climate discourse [9,10].

The dataset primarily consisted of English-language tweets retrieved through keyword-based searches. No geographical filtering was applied to capture globally distributed climate change discussions and maximize discourse diversity. The selected keywords were chosen based on their frequent use in climate communication literature and their relevance to sustainability-related online discussions.

An automated data extraction process was implemented through Python-based (Python 3.14.5) scraping tools and available APIs. The initial dataset consisted of 32,768 tweets posted by individual users. To ensure ethical research practices, only publicly available posts were collected, and no personal or identifiable user information was retained in the dataset.

3.2. Data Preprocessing

Social media data are inherently noisy and unstructured, which necessitates a rigorous preprocessing procedure prior to analysis [6]. To enhance the quality and reliability of the dataset, several cleaning and normalization steps were performed.

First, duplicate tweets and retweets were removed to prevent overrepresentation of identical content. Posts generated by automated bot accounts were also filtered out to minimize potential distortions in the analysis, as previous research has shown that bots can significantly influence online climate debates [8,17]. All retweets were excluded from the dataset to reduce the overrepresentation of highly amplified content and min-

imize the disproportionate influence of a limited number of viral posts or users. While retweets may reflect endorsement and information diffusion dynamics, the primary objective of this study was to analyze original discourse production rather than content amplification behavior.

Second, non-textual elements such as URLs, hashtags, user mentions, emojis, and special characters were eliminated. Although emojis may contain emotional information, they were removed in this study to maintain textual consistency and reduce noise within the deep learning pipeline. Since the primary focus was semantic and thematic textual analysis, retaining emojis could have introduced additional ambiguity due to inconsistent contextual interpretation. Nevertheless, the exclusion of emojis may have reduced the sensitivity of detecting subtle emotional nuances, particularly negative emotions expressed symbolically. All texts were converted to lowercase to ensure uniformity. Common stopwords (e.g., "the," "and", "is") were removed and stemming and lemmatization techniques were applied to reduce words to their base forms. These procedures are consistent with standard practices in NLP-based social media research [4,5].

Bot-generated posts were identified using heuristic filtering criteria, including unusually high posting frequency, repetitive posting behavior, excessive hashtag usage, and low follower-to-following ratios. Accounts meeting multiple automated behavior indicators were excluded from the dataset. Approximately 2134 tweets were removed during this filtering process.

After the completion of preprocessing, the final dataset comprised 29,576 clean and analyzable tweets. This refined corpus served as the foundation for subsequent analytical stages.

3.3. Text Mining

Prior to the application of advanced artificial intelligence–based analytical techniques, an exploratory text mining process was conducted to examine the lexical structure of the dataset and to identify dominant patterns within social media discourse. Text mining plays a critical role in transforming unstructured textual data into meaningful and analyzable information, thereby enabling the detection of key concepts, trends, and discourse patterns [20].

In this study, word frequency analysis was initially performed to identify the most used terms in tweets related to global warming. However, to move beyond simple frequency counts and better capture the relative importance of terms within the corpus, the Term Frequency–Inverse Document Frequency (TF-IDF) weighting scheme was applied. TF-IDF allows for the identification of distinctive and informative terms by reducing the weight of commonly used words while emphasizing terms that are more specific to documents [21]. This approach provides a more robust representation of textual significance and is widely used in social media analytics [22].

Term weighting was carried out using the TF-IDF approach. In this context, the TF-IDF value of a term t within a document d was computed as follows:

$$\mathrm{TF-IDF}_{t,d} = \left(\frac{f_{t,d}}{\sum_k f_{k,d}}\right) \cdot \log\left(\frac{N}{df_t}\right) \quad (1)$$

where $f_{t,d}$ represents the frequency of term t in document d, $\sum_k f_{k,d}$ denotes the total number of terms in document d, N is the total number of documents in the corpus, and df_t indicates the number of documents containing term t.

In addition to single-word analysis, a 2 g (bigram) modelling approach was employed to detect frequently co-occurring word pairs. The use of bigrams enabled the identification of meaningful phrase structures and contextual relationships between terms, offering more

profound insights into how users construct narratives around global warming. Compared to unigram analysis, bigrams provide a more context-sensitive understanding of discourse by capturing commonly used expressions rather than isolated words [23,24].

From a sustainability perspective, identifying dominant terms and discourse patterns helps reveal public priorities, concerns, and awareness levels, which are critical for designing targeted sustainability communication strategies and policy interventions.

3.4. Sentiment Analysis

Sentiment analysis is a computational technique used to identify the emotional orientation of textual content by classifying it as positive, negative, or neutral [25]. In the context of social media research, sentiment analysis plays a crucial role in understanding public attitudes, emotional responses, and collective perceptions toward complex global issues such as climate change. Given that social media platforms provide large-scale, real-time, and user-generated content, sentiment analysis enables systematic examination of how individuals perceive and react to environmental challenges.

However, social media data is often informal, unstructured, and context-dependent, containing abbreviations, sarcasm, and non-standard language patterns. These characteristics limit the effectiveness of traditional lexicon-based approaches. Therefore, data-driven methods based on machine learning and deep learning have become increasingly important, as they allow for more accurate and context-sensitive interpretation of textual data.

In this study, sentiment analysis was applied to determine the emotional distribution of tweets related to global warming. Each tweet was classified into one of three categories:

- Positive: expressions of hope, support for climate action, or solution-oriented perspectives
- Negative: expressions of concern, criticism, skepticism, or perceived environmental threat
- Neutral: informational content or statements without a clear emotional orientation

To ensure reliable and context-aware classification, a deep learning-based sentiment analysis model was employed.

From a sustainability perspective, analyzing sentiment distributions provides valuable insights into public emotional responses to climate change. These insights are critical for understanding policy acceptance, encouraging behavioral change, and improving the effectiveness of sustainability communication strategies.

Deep Learning

Deep learning has emerged as one of the most advanced approaches for sentiment analysis due to its ability to capture complex semantic structures and contextual relationships within large-scale textual data [26]. Unlike traditional methods, deep learning models learn patterns directly from data, enabling more accurate interpretation of informal, dynamic, and context-dependent content such as social media posts [27].

In climate communication research, deep learning offers significant advantages by allowing the detection of subtle emotional expressions, implicit meanings, and variations in language use. This makes it particularly suitable for analyzing social media discourse, where users frequently employ non-standard expressions, abbreviations, and subjective narratives [28].

In this study, a supervised deep learning-based model was used to classify tweets related to global warming according to their emotional orientation. The textual data were transformed into a structured format suitable for model processing, and classification performance was evaluated using standard metrics, including accuracy, precision, recall, and F1-score.

In this study, BERT-based deep learning architecture was employed for sentiment classification. Specifically, the Bidirectional Encoder Representations from Transformers (BERT) model was fine-tuned to classify tweets according to their emotional orientation. BERT was selected due to its strong contextual understanding capabilities and its effectiveness in processing informal and unstructured social media text.

Tweets were tokenized using the BERT tokenizer and padded to a maximum sequence length of 100 tokens. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets. The model was trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and 10 epochs. Early stopping was applied to reduce overfitting and improve generalization performance.

Sentiment labels were generated using a BERT-based automated annotation approach. The use of contextual embeddings enabled more accurate classification of semantically complex and context-dependent social media content. The overall classification performance was evaluated using accuracy, precision, recall, and F1-score metrics.

The use of deep learning strengthens the methodological framework of the study by enabling scalable, consistent, and high-accuracy sentiment classification. Furthermore, the integration of deep learning outputs with topic modelling and visualization techniques provides a comprehensive and multidimensional understanding of social media discourse on global warming.

From a sustainability perspective, deep learning-based analysis offers a powerful tool for monitoring public perceptions in real time and supports data-driven decision-making processes in climate communication and environmental policy development.

3.5. Topic Modelling

While sentiment analysis provides insights into how users feel about global warming, it does not reveal what specific aspects of the issue they discuss. To address this limitation, topic modelling was employed to uncover the latent thematic structure of the dataset.

Latent Dirichlet Allocation (LDA) was selected as the primary topic modelling algorithm. LDA is a probabilistic method that identifies clusters of co-occurring words and interprets them as distinct topics [6,7]. The algorithm assumes that each document (tweet) consists of a mixture of topics and that each topic is represented by a distribution of words.

The optimal number of topics was determined through an iterative process using coherence scores and qualitative evaluation. Multiple models were tested with different topic numbers, and the configuration that provided the most meaningful and interpretable results was selected.

The following joint probability formulation represents the fundamental structure of the LDA model:

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha)\prod_{n=1}^{N} p(z_n|\theta)p(w_n|z_n, \beta) \quad (2)$$

In this formulation, θ denotes the topic distribution of a document, while α represents the Dirichlet prior that governs this distribution. The variable z_n corresponds to the latent topic assignment of the n-th word, and w_n indicates the observed word itself. The parameter β defines the topic–word distribution.

More specifically, $P(\theta I\alpha)$ expresses the prior distribution over topics for a document, $P(z_n I\theta)$ represents the probability of assigning a topic to a word given the document's topic distribution, and $P(w_n Iz_n, \beta)$ denotes the probability of generating a word from a particular topic. Overall, this joint probability formulation reflects the generative assumption of LDA, where each document is modeled as a mixture of topics and each word is assumed to be generated from one of these latent topics.

The application of topic modelling enabled the identification of major discussion themes within the dataset, offering a deeper understanding of the cognitive and conceptual dimensions of climate change discourse.

From a sustainability perspective, topic modelling enables the identification of key thematic areas that shape public discourse, providing critical insights for environmental governance, policy development, and strategic communication planning.

3.6. Visualization and Network Analysis

To enhance the interpretability of the analytical results, several visualization techniques were applied. Keyword co-occurrence networks were generated to illustrate the relationships between frequently used terms. These networks allow researchers to observe how different concepts are interconnected within social media discussions [13].

In addition, density visualizations and heat maps were created to display the distribution of keywords across different sentiment categories. Such visual tools have been effectively used in previous studies to provide intuitive representations of large-scale textual data [14,15].

These visualization methods complemented the quantitative analyses by offering a graphical depiction of the structure and dynamics of online climate discourse.

4. Results

This section presents the findings obtained through text mining, sentiment analysis, and topic modelling techniques applied to social media discourse on global warming. First, the prominent words, concepts, and word usage patterns identified through text mining procedures were examined to reveal the dominant themes and expressions within the dataset. Subsequently, a deep learning-based sentiment analysis was conducted to classify user posts into positive, negative, and neutral categories, thereby identifying the emotional distribution of tweets related to global warming. Finally, LDA-based topic modelling was applied to determine the thematic structures around which the tweets were concentrated, enabling a systematic analysis of users' perceptions and discussion patterns regarding global warming. These findings collectively provide a comprehensive perspective on how global warming is discussed on social media, which topics receive the greatest attention, and how users' emotional orientations are shaped within online discourse.

To evaluate the performance of the sentiment analysis model, widely used classification metrics such as accuracy, precision, recall, and F1-score were employed. Accuracy indicates the overall correctness of the model by measuring the proportion of correctly classified instances across the entire dataset. Precision measures the proportion of instances predicted as positive that are positive, thereby reflecting the reliability of positive class predictions. Recall, on the other hand, represents the proportion of actual positive instances that are correctly identified by the model, providing an assessment of the model's ability to capture positive cases. The F1-score is a composite metric that balances precision and recall, offering a more reliable evaluation of model performance, particularly in cases where class distributions are imbalanced. The combined use of these metrics enables a comprehensive assessment of both the strengths and weaknesses of the model. The confusion matrix obtained from the analysis is presented in Table 1.

Table 1 presents the confusion matrix results, indicating that the model achieved a high level of classification performance. The high values of true positives (18,785) and true negatives (8561) demonstrate that the model effectively distinguishes between sentiment categories. In contrast, the relatively low number of false positives (1186) and false negatives (1044) suggest that misclassification rates remain limited. Overall, these results confirm that the deep learning-based model provides accurate and reliable sentiment classification

for social media discourse on global warming. The performance evaluation results of the proposed model are summarized in Table 2.

Table 1. Confusion matrix.

	Tweet Counts
True positive	18,785
True negative	8561
False positive	1186
False negative	1044

Table 2. Performance measurement.

	Accuracy	**Precision**	**Recall**	**F Score (F1)**
Results	0.925	0.941	0.947	0.944

The performance metrics presented in Table 2 indicate that the proposed deep learning model achieved a high level of classification, accuracy and reliability. The accuracy value of 92.5% demonstrates strong overall predictive performance, while the precision (0.941) and recall (0.947) values indicate that the model effectively identifies and correctly classifies sentiment categories. The F1-score of 0.944 further confirms a well-balanced balance between precision and recall. Overall, these results show that the model provides robust and consistent sentiment classification for social media data related to global warming.

The overall distribution of sentiment categories and their statistical characteristics are presented in Table 3.

Table 3. Sentiment statistics.

Sentiment	**Tweet Counts**
Positive	6550
Neutral	4372
Negative	18,654

The sentiment distribution statistics presented in Table 3 clearly indicate that social media discourse on global warming is predominantly characterized by negative emotional expressions. The presence of 18,654 tweets classified under the negative category demonstrates that a substantial portion of users perceive global warming as a serious, threatening, and urgent issue requiring immediate attention. These negative sentiments are likely influenced by factors such as the increasing visibility of climate-related disasters, dissatisfaction with environmental policies, concerns about future uncertainties, and feelings of individual helplessness in the face of a global crisis.

In contrast, the 6550 tweets categorized as positive suggest that a segment of users adopts a more constructive and solution-oriented perspective. These posts generally reflect messages that emphasize hope, support for environmental initiatives, encouragement of sustainable practices, and awareness-raising efforts aimed at mitigating the impacts of climate change. Although positive content appears in a smaller proportion compared to negative sentiment, it nonetheless highlights the presence of an active group of users focused on promoting optimism and collective action.

The 4372 tweets classified as neutral are likely composed of informational posts, news updates, official statements, or comments that do not convey a clear emotional orientation.

Such messages typically serve to share information rather than to express personal attitudes or evaluations.

When considered as a whole, this distribution reveals that discussions about global warming on social media are largely framed within a critical, concern-driven, and problem-focused narrative. Positive messages, while more limited in number, tend to concentrate on themes of solutions, awareness, and hope. These findings demonstrate that global warming is widely perceived as a significant source of public concern and that social media discourse strongly reflects the emotional intensity associated with this global challenge. From a sustainability perspective, the dominance of negative sentiment indicates a heightened level of public concern and perceived environmental risk, which may act as a catalyst for sustainability-oriented policy action and behavioral change. At the same time, the presence of positive and neutral content highlights the potential of social media to support sustainability communication, awareness-building, and informed public engagement.

The most frequently used words in the dataset were identified through text mining analysis, and the distribution of these words is presented in Table 4.

Table 4. Word frequency distribution.

Positive	Negative	Neutral
climate	cold	climate
change	climate	change
world	change	years
people	weather	people
need	ice	planet
stop	bad	earth
earth	winter	carbon
ice	people	weather
scientists	fear	hoax
energy	planet	think

An examination of the positive sentiment cluster reveals that a considerable portion of social media discourse on global warming emphasizes hope, solution-oriented thinking, and collective responsibility. Within this narrative framework, users frequently highlight the importance of science-based action, the transformation of energy policies, awareness-raising initiatives, and the necessity of global cooperation in addressing the climate crisis. Positive content does not imply a denial of the threat posed by global warming; rather, it reflects an approach that frames the crisis not solely through anxiety and negativity but also through calls for action, adaptation, and constructive change. This suggests that social media discussions on global warming are not confined to pessimistic perspectives alone, but also contain a meaningful space for optimism, motivation, and problem-solving potential.

This interpretation is also consistent with studies on constructive hope and solution-oriented climate communication, which emphasize that hopeful narratives can support collective engagement and climate action [29,30].

The analysis of the negative sentiment cluster demonstrates that social media discourse on global warming is predominantly shaped by concern, perceived threat, and dissatisfaction. Users often interpret global warming as a direct crisis affecting their daily lives, particularly through references to extreme weather events, seasonal irregularities, and changing climate conditions. Comparisons between past and present climate experiences frequently reinforce the narrative that current conditions are "not like before," intensifying perceptions of environmental decline. This interpretation emerged primarily from manual examination of frequently co-occurring terms and contextual tweet patterns, particularly

expressions related to past-versus-present climate comparisons such as "used to," "before," and "not like previous years." In addition, strong feelings of uncertainty and anxiety regarding the future of the planet, environmental degradation, and the social and economic consequences of climate change are clear. Criticism directed toward human responsibility and environmental policies also appears within this discourse, indicating that global warming is discussed not only as an environmental issue but also as a social and political problem. Overall, this structure confirms that the negative sentiment category frames global warming primarily within a discourse of crisis, threat, and critique. From a sustainability perspective, the distribution of frequently used terms reflects the key areas of public attention and concern, providing valuable insights for policymakers and communication strategists to design more targeted and effective sustainability interventions.

The neutral sentiment cluster, on the other hand, illustrates that global warming is addressed on social media not only through emotional reactions but also through information sharing, evaluation, and discussion. Neutral posts largely consist of news updates, the dissemination of scientific findings, general assessments, and conceptual debates. These messages are characterized by a more rational and explanatory tone rather than an explicitly emotional orientation. Furthermore, some neutral content reflects a questioning or exploratory attitude, contributing to discussions about the causes, consequences, and validity of global warming. This indicates that social media functions as a platform not only for expressing emotions but also for facilitating knowledge-based dialogue and public deliberation. The presence of a substantial neutral category highlights the importance of social media as a space where global warming is treated as an issue of information exchange and intellectual engagement, in addition to being an emotionally charged topic. The positive word cloud is presented in Figure 1.



Figure 1. Positive word cloud.

The positive word cloud highlights that social media discourse on global warming includes solution-oriented and constructive narratives. Users frequently emphasize themes such as awareness, collective responsibility, scientific action, and renewable energy. This indicates that, alongside concern, social media also supports proactive and action-driven sustainability communication.

The negative word cloud is presented in Figure 2.

The negative word cloud reveals that social media discourse is largely shaped by concern, uncertainty, and perceived environmental threat. Frequent references to extreme weather conditions, seasonal irregularities, and skepticism indicate that global warming is

often framed as a crisis-oriented and contested issue. This pattern highlights the need for more effective and trust-based sustainability communication.



Figure 2. Negative word cloud.

The neutral word cloud is presented in Figure 3.

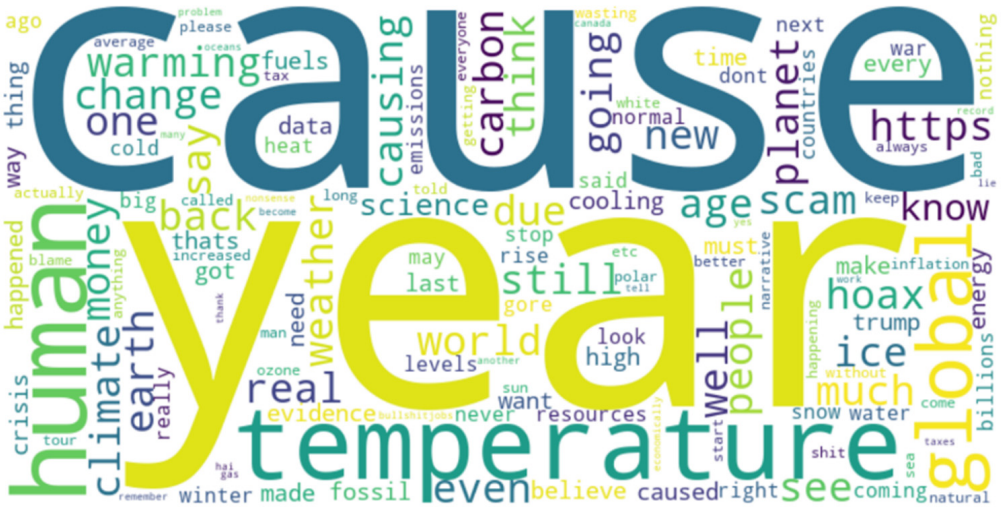


Figure 3. Neutral word cloud.

The neutral word cloud demonstrates that social media also functions as a platform for information sharing and discussion. Users frequently engage in explanatory and analytical discourse, focusing on topics such as climate processes, carbon emissions, and environmental impacts. This reflects the role of social media in supporting knowledge-based sustainability communication.

Figure 4 presents density visualization.

Density Visualization clearly illustrates the conceptual areas around which social media discourse on global warming is most intensely concentrated, revealing the core thematic focuses of online discussions. An examination of the visualization shows that the central cluster is formed around the terms “warming,” “global,” and “climate,” indicating that conversations are primarily structured around the fundamental concept of climate change itself. The gradual decrease in density as one moves away from the center suggests that users tend to focus predominantly on a limited number of key themes, while other related concepts function as supplementary elements that support, elaborate, or expand the main discourse.

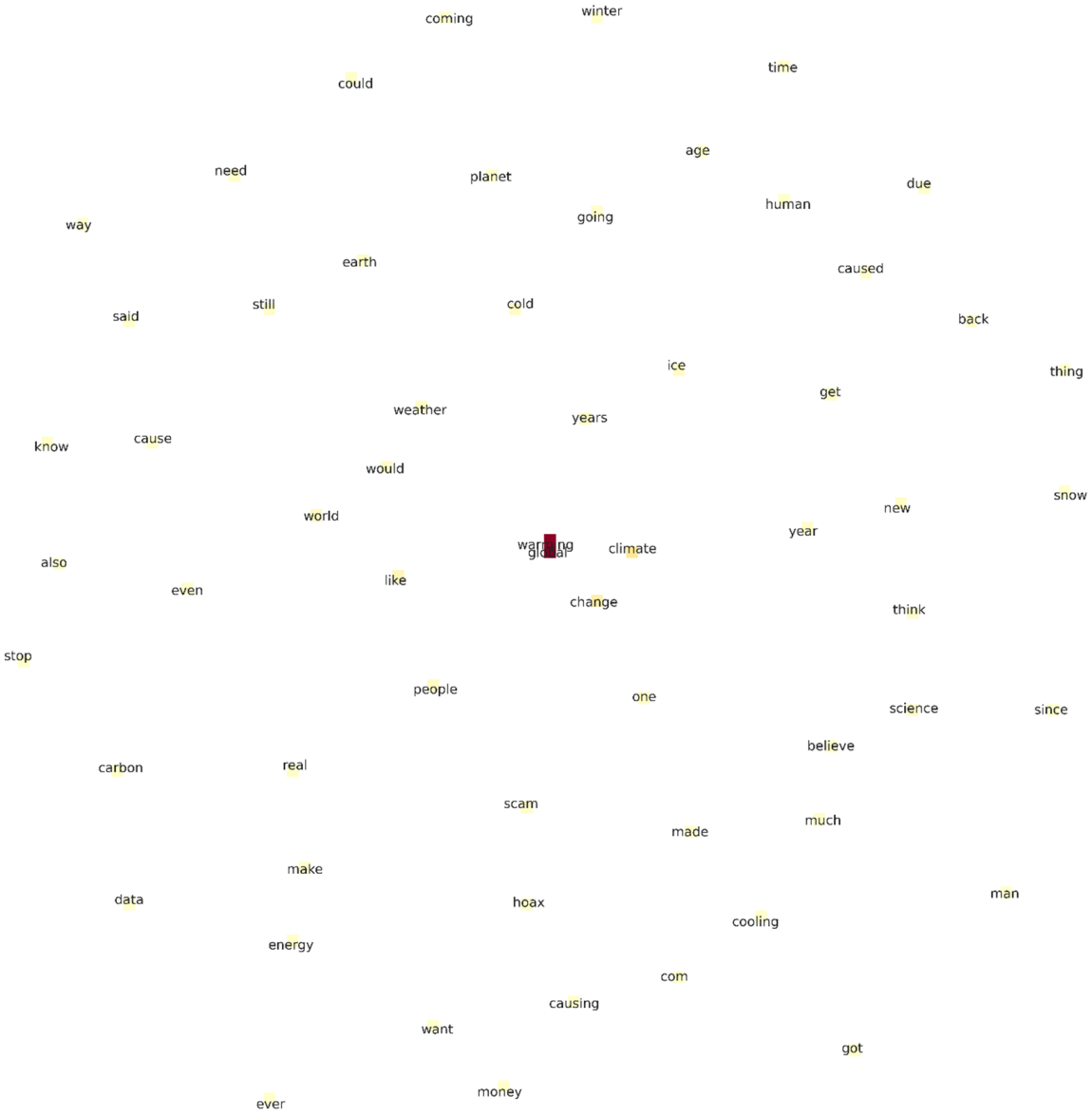


Figure 4. Density visualization.

The spatial distribution of additional terms in the visualization further indicates that social media discussions are not confined solely to scientific or technical explanations. Instead, the discourse encompasses multiple dimensions, including everyday life experiences, extreme weather events, human responsibility, and economic and political debates. This multi-layered structure demonstrates that global warming is addressed not only as an environmental issue but also as a complex social phenomenon with diverse implications.

Overall, the density map reveals that social media discourse on global warming is strongly clustered around certain central concepts, while simultaneously enriched by a variety of subtopics and complementary discussions. This visualization highlights the interconnected and multifaceted nature of online climate change communication, providing valuable insight into how different aspects of the issue are organized and emphasized within digital platforms. From a sustainability perspective, the concentration of discourse around core concepts suggests that climate communication is structured around key narratives, providing strategic entry points for policymakers to enhance the effectiveness of sustainability communication efforts.

Figure 5 presents the keyword co-occurrence network.

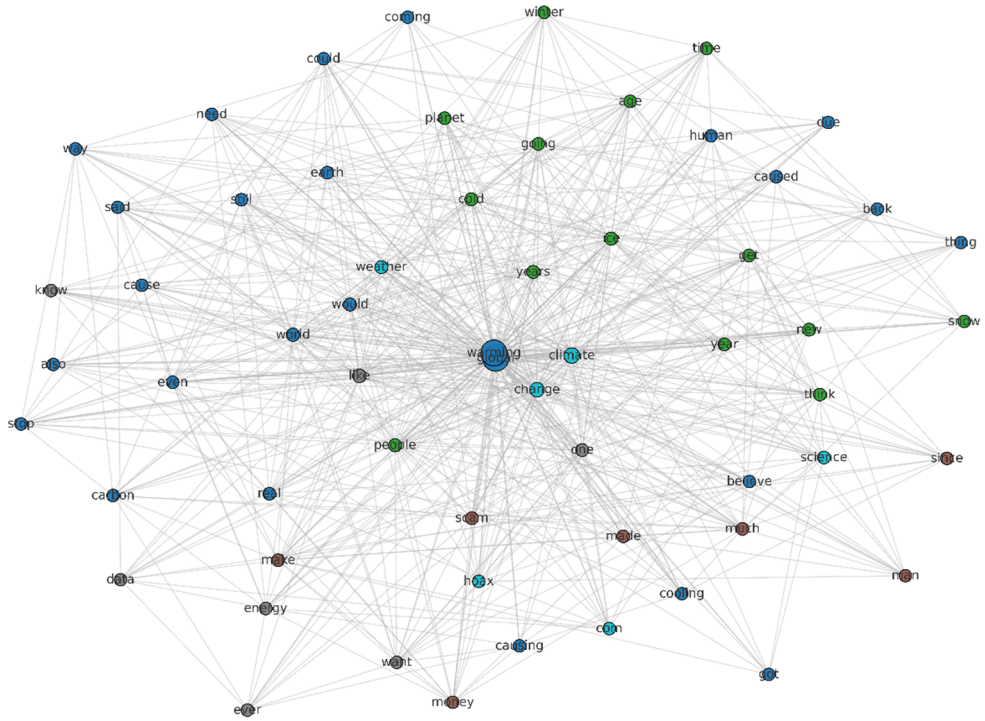


Figure 5. Keyword co-occurrence network.

The Keyword co-occurrence network visualization clearly demonstrates how social media discourse on global warming is structured around interconnected concepts and thematic clusters. The network reveals which terms frequently appear together and how discussions are organized through meaningful associations. The most central and highly connected nodes in the network are formed around the concepts of "global," "warming," "climate," and "change," indicating that the core of social media conversations is directly anchored in the phenomenon of climate change. This central positioning confirms that online discussions are fundamentally focused on the primary dimensions of global warming.

Surrounding nodes with strong connections to the central concepts represent themes such as extreme weather events, human impact, carbon emissions, scientific debates, and economic and political issues. These relationships illustrate that such topics are continuously addressed in interaction with the main climate discourse, rather than being treated as isolated subjects. The dense interconnectivity between terms suggests that users do not perceive global warming from a single perspective but instead approach it through a multidimensional framework that integrates everyday experiences, scientific evidence, social concerns, and policy-related discussions.

Overall, the network structure indicates that social media discourse on global warming is both cohesive and diverse. While conversations converge around a strong shared core, they are simultaneously enriched by a variety of subthemes and related issues. This visualization highlights the complex yet coherent nature of online climate communication, demonstrating that global warming is discussed through an integrated web of interrelated concepts and perspectives. This interconnected structure highlights the importance of integrated sustainability communication strategies that simultaneously address environmental, social, and economic dimensions of climate change.

To evaluate the robustness and stability of the LDA topic modelling results, additional validation procedures were conducted. Multiple LDA models with topic numbers ranging from k = 2 to k = 6 were tested using coherence score analysis. The highest coherence score was obtained for the three-topic solution (Cv = 0.61), indicating that this configuration provided the most semantically interpretable structure.

In addition, the LDA algorithm was executed 15 times using different random seeds to assess topic stability. Topic similarity analysis based on Jaccard similarity coefficients demonstrated that the three-topic structure consistently emerged across runs, with an average similarity score of 0.84. The prevalence rates of the topics were also found to be stable, with narrow confidence intervals across repeated iterations. These findings confirm the robustness and reproducibility of the identified thematic structure.

A summary of the topic modelling results is presented in Table 5.

Table 5. Summary of topic modelling results.

Topic N	Mean Topic Probability	Prevalence	Topmost Co-Occurring Words	Main Themes
1	0.3676	41.37%	cold, people, just, hoax, world, years, going, snow, like, year, stop, weather, hot, cause, happened, need, want, new, planet, ice	Daily Weather Experiences and Skeptical Climate Discourse on social media
3	0.3417	31.62%	climate, change, like, world, ice, fuels, years, fossils, caused, energy, scam, temperature, weather, weather, weather, weather, evidence, data, sea, billions, money, causes	Climate Change, Energy and Scientific Evidence Debates on social media
2	0.2907	27.01%	CO_2, just, like, thing, winter, think, believe, carbon, new, happening, causing, know, good, coming, year, cooling, ice, need	Carbon Emissions, Human Impact and Cause-Effect Debates on social media

Topic prevalence values represent the proportion of documents primarily associated with each topic across the corpus. The three-topic solution achieved the highest coherence score (Cv = 0.61), supporting its semantic interpretability and model suitability.

Topic 1—Daily Weather Experiences and Skeptical Climate Discourse on social media.

This theme shows that users mostly evaluate global warming discussions through daily weather events, cold weather, snowfall, and unexpected seasonal conditions. It is noteworthy that individual observations and experiences are at the forefront of social media posts, and that these experiences sometimes fuel skepticism, denial, or a mocking approach to the reality of climate change. Therefore, in this discourse field, a discussion structure emerges where personal perceptions and intuitions are more decisive than scientific explanations of the climate crisis. According to the table, the Mean Topic Probability value of 0.3676 indicates that words associated with this theme are highly likely to appear together in related tweets and that the theme has a strong semantic coherence. The prevalence rate being the highest at 41.37% reveals that this theme is the most common axis of discussion in the data set and that a significant portion of social media users interpret global warming through their daily weather experiences. From a sustainability perspective, this theme represents a significant barrier to effective climate action, emphasizing the need for stronger science-based communication strategies to counter misinformation and skepticism.

Topic 3—Climate Change, Energy and Scientific Evidence Debates on Social Media.

This theme represents content that addresses global warming on social media with a more analytical and knowledge-based discourse. Users are seen to focus on more comprehensive and complex issues such as the reality of climate change, fossil fuel use, energy policies, rising temperatures, sea level rise, carbon emissions, and economic dimensions. In this discourse, data, evidence, and scientific explanations serve as important reference points; it is understood that the climate crisis is discussed not only in terms of individual feelings but also within a scientific and political framework. The Mean Topic Probability value of 0.3417 indicates that the words associated with this theme are likely to appear together and that the theme has a strong structural consistency. The prevalence rate of 31.62%

reveals that a significant user group on social media conducts climate change discussions from a scientific and energy-focused perspective, and that this theme constitutes the second strongest discourse area in the data set. This theme represents an important opportunity for sustainability policy development, as it reflects the presence of a more informed and evidence-based public discourse.

Topic 2—Carbon Emissions, Human Impact and Cause-Effect Debates on Social Media.

This theme encompasses social media discourse that discusses the cause-and-effect relationships underlying global warming, focusing on topics such as human impact, carbon emissions, CO_2 levels, and changes in climate processes. Users appear to concentrate on how the process works, what might happen in the future, and the question of whether global warming is natural or human induced. This theme also includes belief versus skepticism, individual assessments, and discussion-focused posts. The Mean Topic Probability value of 0.2907 indicates that the co-occurrence level of words associated with this theme is meaningful but slightly weaker compared to the other two themes. The prevalence rate of 27.01% reveals that this discourse occupies an important place in the data set but represents a more limited user group compared to the other two themes. Overall, this theme reflects the discussion axis on social media concerning the causes of global warming, human impact, and scientific explanations. This discourse highlights increasing awareness of human responsibility, which is essential for promoting sustainable behavior and supporting long-term environmental policy implementation.

Overall, the findings demonstrate that social media discourse on global warming provides not only descriptive insights but also strategic implications for sustainability communication, public awareness, and policy development.

5. Discussion

This study provides a comprehensive and integrated analysis of social media discourse on global warming by combining text mining, deep learning-based sentiment analysis, and topic modeling techniques. From a sustainability perspective, the findings offer important insights into how public perceptions, emotional responses, and thematic structures shape climate change communication in digital environments.

First, the results of the sentiment analysis reveal that social media discourse is predominantly characterized by negative emotional expressions. This indicates that global warming is widely perceived as a serious and urgent issue, often associated with concern, uncertainty, and criticism. This finding is consistent with previous studies, which demonstrate that climate change discussions on social media are largely framed through anxiety and risk perception [3–5]. At the same time, the presence of positive and solution-oriented content suggests that social media is not only a space of concern but also a platform that supports awareness, collective responsibility, and climate-related engagement. This dual structure highlights the potential of digital platforms to both reflect and influence sustainability-oriented behaviors.

Second, the topic modeling results reveal that social media discourse is structured around three main thematic dimensions. The first theme reflects discussions grounded in everyday weather experiences and skepticism, where individual observations often shape perceptions of climate change. This aligns with prior research emphasizing the role of experiential narratives and skepticism in online climate discourse [8,9]. The second theme represents a more analytical perspective, focusing on scientific evidence, energy policies, and broader climate debates. This indicates that a significant portion of users engage with climate change through knowledge-based and policy-oriented discussions, supporting findings in existing literature [6,7]. The third theme centers on carbon emissions, human impact, and cause–effect relationships, reflecting an increasing awareness of responsibility

and the underlying mechanisms of global warming. Together, these themes demonstrate that social media discourse is multidimensional, combining experiential, analytical, and responsibility-driven perspectives.

Third, the findings derived from keyword co-occurrence networks and density visualizations further demonstrate that social media discourse is organized around strongly interconnected concepts. The concentration of discussions around central terms such as "climate," "global," and "warming" indicates that online conversations are structured through cohesive thematic clusters rather than fragmented narratives. This supports previous research that emphasizes the value of network structures and conceptual density in understanding digital climate communication [11,14]. In this regard, social media functions not only as a platform for emotional expression but also as a structured communication environment where information exchange and public debate take place [12,13].

These findings can also be interpreted through the perspective of climate responsibility attribution. Recent studies emphasize that climate-related responsibility is increasingly distributed across geographically dispersed actors, production systems, and global supply chains. For example, Zheng et al. demonstrated that different accounting boundaries and attribution approaches significantly influence how responsibility for emissions is perceived within global photovoltaic trade systems [31]. In this regard, the contested narratives observed in social media discourse may reflect broader societal disagreements regarding who should bear responsibility for climate mitigation and sustainability transitions. This perspective further highlights the complexity of climate communication and environmental governance in digitally connected societies.

When considered together, these findings extend existing literature by integrating emotional, thematic, and structural dimensions of social media discourse within a unified analytical framework. Unlike many previous studies that focus on a single methodological approach, this research demonstrates the value of combining multiple AI-driven techniques to achieve a more comprehensive understanding of sustainability-related communication.

Beyond identifying dominant sentiments and themes, the findings suggest that climate-related social media discourse operates as a hybrid communicative environment where emotional reactions, scientific reasoning, skepticism, and responsibility attribution coexist simultaneously. This indicates that digital climate communication should not be interpreted through binary categories such as support versus denial alone. Instead, online discourse reflects fragmented and overlapping forms of environmental construction shaped by experiential narratives, political interpretations, and collective uncertainty. In this regard, the study contributes theoretically by proposing that AI-driven discourse analysis can reveal the multidimensional and hybrid nature of sustainability communication within digitally mediated public spheres.

From a practical perspective, the findings highlight several important implications for sustainability policy and communication. The dominance of negative sentiment suggests a high level of perceived risk, emphasizing the need for transparent, science-based, and trust-oriented communication strategies. The presence of skepticism and misinformation further underlines the importance of improving climate literacy and countering misleading narratives. At the same time, the existence of positive and solution-oriented discourse indicates that social media can serve as a powerful tool for promoting awareness, encouraging sustainable behavior, and fostering collective action.

Overall, this study demonstrates that artificial intelligence-driven social media analysis provides valuable insights into public perceptions of global warming and offers a strategic foundation for developing more effective sustainability communication and policy interventions.

Implications for Sustainability Policy and Communication

The findings of this study provide several important implications for sustainability policy and communication strategies. First, the dominance of negative sentiment indicates a widespread perception of risk and urgency, suggesting that policymakers should adopt transparent and trust-based communication strategies to address public concerns. Second, the presence of skeptical and misinformation-driven discourse highlights the need for targeted interventions aimed at improving climate literacy and combating false narratives. Third, the existence of positive and solution-oriented content demonstrates the potential of social media as a tool for promoting sustainable behaviors, raising awareness, and encouraging collective action. Therefore, integrating data-driven insights from social media into policy design can enhance the effectiveness of sustainability communication and support more adaptive and responsive environmental governance. Since the primary objective of this study was exploratory discourse analysis rather than causal prediction or hypothesis testing, traditional control variables commonly used in regression-based models were not incorporated into the analytical framework. The study focused on identifying emotional distributions, thematic structures, and conceptual relationships within large-scale social media discourse rather than estimating causal effects between variables. Nevertheless, future studies may incorporate demographic, geographical, temporal, or platform-specific control variables to further enhance analytical granularity and comparative robustness.

From a practical perspective, policymakers and environmental organizations can leverage AI-driven social media analytics to design more adaptive and targeted climate communication campaigns. For example, the dominance of skepticism-related discourse suggests the need for evidence-based communication strategies that directly address misinformation and public distrust. Similarly, the identification of solution-oriented and hopeful narratives may help institutions develop campaigns emphasizing collective action, renewable energy adoption, and sustainability awareness. Real-time monitoring of social media sentiment may also support early detection of emerging public concerns, enabling more responsive and participatory environmental governance strategies.

6. Conclusions

This study provides a comprehensive analysis of public perceptions of global warming by integrating text mining, deep learning-based sentiment analysis, LDA topic modelling, and network-based visualizations using large-scale social media data. The findings demonstrate that global warming constitutes a central topic of public discourse, structured around distinct emotional and thematic dimensions. Sentiment analysis reveals that climate change is predominantly associated with concern, uncertainty, and criticism, while the presence of solution-oriented narratives highlights the potential for awareness building and collective action.

The results of topic modelling further confirm the multidimensional nature of online discussions. Social media discourse is shaped by three main thematic dimensions: skepticism grounded in everyday experiences, analytical discussions based on scientific evidence and policy debates, and responsibility-driven narratives focusing on carbon emissions and human impact. In addition, network and density analyses indicate that discussions are organized around strongly interconnected concepts, reflecting a structured and coherent communication environment.

From a sustainability perspective, these findings demonstrate that social media is not only a platform for expressing public opinion but also a critical tool for understanding and shaping climate-related awareness, engagement, and behavioral change. The dominance of negative sentiment highlights the need for transparent, science-based, and trust-oriented

communication strategies, while the presence of constructive discourse indicates the potential for effective public mobilization.

The findings also underscore the importance of integrating artificial intelligence-driven analytics into sustainability research and practice. AI-based approaches enable large-scale and real-time analysis of public discourse, providing valuable insights for policymakers, environmental organizations, and communication professionals. In this regard, social media analytics can support more adaptive, data-driven, and responsive sustainability strategies.

The findings of this study should be interpreted within the context of several limitations. First, the analysis represents a cross-sectional snapshot of discourse collected exclusively from the X platform during December 2025. Therefore, seasonal events, climate-related disasters, or international policy discussions occurring during this period may have influenced discourse patterns and sentiment distributions. Second, although bot-generated content was filtered to reduce potential distortions, automated accounts may still have influenced the thematic and emotional structure of the dataset. Existing literature suggests that bots often amplify polarizing and negative climate narratives, which may have contributed to the dominance of negative sentiment observed in this study. Third, the analysis focused primarily on English-language content, limiting the cross-cultural generalizability of the findings. Noise, linguistic variability, and residual bot-generated content may have influenced both sentiment classification and topic assignment processes. Automated accounts are known to amplify polarized and negative climate narratives, potentially contributing to the dominance of negative sentiment observed in the dataset. Similarly, informal language structures and sarcasm commonly found in social media discourse may have affected topic coherence and classification accuracy. Consequently, the results should not be interpreted as universally representative of global climate discourse.

Future research should expand the scope by incorporating multiple social media platforms, longer time periods, and cross-cultural analyses. Further investigation into misinformation dynamics and the integration of advanced AI models, including large language models, could enhance analytical depth and accuracy. Strengthening collaboration between researchers, policymakers, and practitioners may also improve the practical impact of such analyses.

Overall, this study demonstrates that AI-driven analysis of social media discourse provides a powerful framework for understanding public perceptions of global warming and plays a critical role in advancing sustainability communication, supporting evidence-based policy development, and fostering more informed and engaged societies.

Author Contributions: Conceptualization, O.K. and M.K.; methodology, G.M. and M.K.; software, M.K.; validation, O.K. and G.M.; formal analysis, M.K.; investigation, O.K.; resources, G.M.; data curation, M.K.; writing—original draft preparation, G.M. and M.K.; writing—review and editing, G.M. and M.K.; visualization, O.K.; supervision, O.K.; project administration, O.K.; funding acquisition, G.M. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the National University of Science and Technology Politehnica Bucharest through the program PubArt.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset is available from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Alkhatib, M.; El Barachi, M.; Mathew, S.S.; Oroumchian, F. Using artificial intelligence to monitor the evolution of opinion leaders' sentiments: Case study on global warming. In *Proceedings of the 2020 5th International Conference on Smart and Sustainable Technologies (SpliTech), Split, Croatia, 23–26 September 2020*; IEEE: New York, NY, USA, 2020.
2. El Barachi, M.; AlKhatib, M.; Mathew, S.; Oroumchian, F. A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change. *J. Clean. Prod.* **2021**, *312*, 127820. [CrossRef]
3. Loureiro, M.L.; Alló, M. Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the UK and Spain. *Energy Policy* **2020**, *143*, 111490. [CrossRef]
4. Ghosh, D.; Kujur, A.G.P.; Salim, A. Exploring Public Perspectives on Climate Change with AI Analysis. In *AI and Green Technology Applications in Society*; IGI Global Scientific Publishing: Hershey, PA, USA, 2025; pp. 1–30.
5. Rahman, G.; Fitriyah, A. Harnessing AI for Climate Change Communication: Analyzing Public Perception through NLP and Machine Learning. *Sinergi Int. J. Commun. Sci.* **2025**, *3*, 87–98. [CrossRef]
6. Kaushal, A.; Acharjee, A.; Mandal, A. Machine learning based attribution mapping of climate related discussions on social media. *Sci. Rep.* **2022**, *12*, 19033. [CrossRef] [PubMed]
7. Gokcimen, T.; Das, B. Exploring climate change discourse on social media and blogs using a topic modeling analysis. *Heliyon* **2024**, *10*, e32464. [CrossRef]
8. Al-Rawi, A.; Kane, O.; Bizimana, A.-J. Topic modelling of public Twitter discourses, part bot, part active human user, on climate change and global warming. *J. Environ. Media* **2021**, *2*, 31–53. [CrossRef]
9. Lineman, M.; Do, Y.; Kim, J.Y.; Joo, G.-J. Talking about climate change and global warming. *PLoS ONE* **2015**, *10*, e0138996. [CrossRef]
10. Li, Y.; Yu, B.; Dai, J. "Climate Change" or "Global Warming"? The (Un) Politicization of Climate in Chinese Social Media Platform. *Environ. Commun.* **2024**, *18*, 927–944. [CrossRef]
11. Effrosynidis, D.; Karasakalidis, A.I.; Sylaios, G.; Arampatzis, A. The climate change Twitter dataset. *Expert Syst. Appl.* **2022**, *204*, 117541. [CrossRef]
12. Du, X.; Kowalski, M.; Varde, A.S.; de Melo, G.; Taylor, R.W. Public opinion matters: Mining social media text for environmental management. *ACM SIGWEB Newsl.* **2020**, *2019*, 5. [CrossRef]
13. Willaert, T.; Van Eecke, P.; Beuls, K.; Steels, L. Building social media observatories for monitoring online opinion dynamics. *Soc. Media+ Soc.* **2020**, *6*, 2056305119898778. [CrossRef]
14. Ballestar, M.T.; Martín-Llaguno, M.; Sainz, J. An artificial intelligence analysis of climate-change influencers' marketing on Twitter. *Psychol. Mark.* **2022**, *39*, 2273–2283. [CrossRef]
15. Eslen-Ziya, H. Humour and sarcasm: Expressions of global warming on Twitter. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 240. [CrossRef]
16. Amangeldi, D.; Usmanova, A.; Shamoi, P. Understanding environmental posts: Sentiment and emotion analysis of social media data. *IEEE Access* **2024**, *12*, 33504–33523. [CrossRef]
17. Chen, C.-F.; Shi, W.; Yang, J.; Fu, H.-H. Social bots' role in climate change discussion on Twitter: Measuring standpoints, topics, and interaction strategies. *Adv. Clim. Change Res.* **2021**, *12*, 913–923. [CrossRef]
18. Duong, C.; Liu, Q.; Mao, R.; Cambria, E. Saving earth one tweet at a time through the lens of artificial intelligence. In *Proceedings of the 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 18–23 July 2022*; IEEE: New York, NY, USA, 2022.
19. Duong, C.; Raghuram, V.C.; Lee, A.; Mao, R.; Mengaldo, G.; Cambria, E. Neurosymbolic ai for mining public opinions about wildfires. *Cogn. Comput.* **2024**, *16*, 1531–1553. [CrossRef]
20. Hassani, H.; Beneki, C.; Unger, S.; Mazinani, M.T.; Yeganegi, M.R. Text mining in big data analytics. *Big Data Cogn. Comput.* **2020**, *4*, 1. [CrossRef]
21. Qaiser, S.; Ali, R. Text mining: Use of TF-IDF to examine the relevance of words to documents. *Int. J. Comput. Appl.* **2018**, *181*, 25–29. [CrossRef]
22. Jiang, Z.; Gao, B.; He, Y.; Han, Y.; Doyle, P.; Zhu, Q. Text Classification Using Novel Term Weighting Scheme-Based Improved TF-IDF for Internet Media Reports. *Math. Probl. Eng.* **2021**, *2021*, 6619088. [CrossRef]
23. Tan, C.-M.; Wang, Y.-F.; Lee, C.-D. The use of bigrams to enhance text categorization. *Inf. Process. Manag.* **2002**, *38*, 529–546. [CrossRef]
24. Reilly, J.; Finley, A.M.; Litovsky, C.P.; Kenett, Y.N. Bigram semantic distance as an index of continuous semantic flow in natural language: Theory, tools, and applications. *J. Exp. Psychol. Gen.* **2023**, *152*, 2578. [CrossRef]
25. Jayal, J.; Kumar, V.; Sarkar, P.; Dutta, S.K. Emotion Detection Using Natural Language Processing by Text Classification. In *Natural Language Processing for Software Engineering*; Scrivener Publishing: Beverly, MA, USA, 2025; pp. 425–442.
26. Abdullah, T.; Ahmet, A. Deep learning in sentiment analysis: Recent architectures. *ACM Comput. Surv.* **2022**, *55*, 159. [CrossRef]
27. Shi, Y.; Lian, L.; Shi, Y.; Wang, Z.; Zhou, Y.; Fu, L.; Bai, L.; Zhang, J.; Zhang, W. Machine learning for large-scale optimization in 6G wireless networks. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2088–2132. [CrossRef]

28. Xing, X.; Yu, B.; Kang, C.; Huang, B.; Gong, J.; Liu, Y. The synergy between remote sensing and social sensing in urban studies: Review and perspectives. *IEEE Geosci. Remote Sens. Mag.* **2024**, *12*, 108–137. [CrossRef]
29. Feldman, L.; Hart, P.S. Is there any hope? How climate change news imagery and text influence audience emotions and support for climate mitigation policies. *Risk Anal.* **2018**, *38*, 585–602. [CrossRef]
30. Ojala, M. Hope and climate change: The importance of hope for environmental engagement among young people. *Environ. Educ. Res.* **2012**, *18*, 625–642. [CrossRef]
31. Zheng, S.; Ye, Z.; Lu, K.; He, T.; Liu, Y.; Zhang, L.; Hao, Y. Accounting for non-CO_2 greenhouse gases in global PV trade: Implications for climate responsibility and policy. *Sustain. Prod. Consum.* **2026**, *62*, 55–69. [CrossRef]

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.