UNIVERSITY OF AMSTERDAM
Communication Science

Is having a boyfriend embarrassing now?An automated content analysis of YouTube comments on The Vogue article

Name: Ada Shi
Course: Big Data and Automated Content Analysis 1
Lecturer: Dr. Anne Kroon
Date: December 17, 2025
Word count: 2042

Is having a boyfriend embarrassing now?An automated content analysis of YouTube comments on The Vogue article

Introduction

Is having a boyfriend embarrassing now? This provocative question, posed in a British Vogue article by Chanté Joseph (2025), immediately captured my attention, particularly as multiple friends shared this piece with me shortly after its publication. The article, with its catchy and somewhat controversial title, sparked widespread discussion in online public discourse. As Joseph noted in an interview with CNN, audiences' reactions to her article were polarized: "Some people love it, and some people hate it" (Radford, 2025).

Joseph's article examines contemporary attitudes toward heterosexual relationships and their public performance, arguing that overt displays of romantic partnerships, once markers of social status and maturity, have become less common and are increasingly perceived as "uncool" or even embarrassing. She observes that many women now post subtle or cropped representations of their partners on social media, navigating the benefits of partnership while avoiding being overly defined by it. Joseph further notes that audiences often mute or unfollow accounts that are excessively boyfriend-centric, and that some women, especially social media influencers, hesitate to publicly affirm their relationships due to superstition or social backlash. Overall, she contends that being partnered no longer automatically affirms womanhood and that selective visibility and curated independence reflect broader shifts in gender norms, relational dynamics, and digital identity disclosure in contemporary society.

These observations resonate with broader societal trends identified in the literature. For instance, Twenge (2023) highlights that younger generations prioritize autonomy and personal choice, and often reevaluate traditional markers of social success such as romantic partnerships. McRobbie (2020) emphasizes that contemporary cultural narratives, shaped by postfeminist and neoliberal ideologies, encourage women to exercise independence and curate their public personas, reflecting changing expectations about gender roles and relational identity.

Building on these insights, the present study explores how audiences engage with the statement in the article's title through computational online discourse analysis. While highly exploratory, this study is motivated by nuanced societal changes in which women appear increasingly independent and less constrained by traditional expectations of presenting themselves in relation to men. By analyzing public reactions, the study aims to illuminate how these shifts manifest in everyday online discussions, particularly among individuals who engage with the topic without taking an explicit stance. Given the exploratory nature of this research, no hypotheses were formulated. Instead, the study is guided by the following open-ended research questions:

RQ1*: To what extent do people agree, disagree, or neither agree nor disagree with the statement that it is embarrassing to have a boyfriend now?*

RQ2*: What are the major topics discussed by people who neither agree nor disagree?*

Importantly, the focus of this study is not on analyzing all of the nuanced arguments and perspectives presented in the article per se, which are rich, complex, and difficult to capture in large-scale analysis. Instead, the research specifically investigates how people react to the statement, “Is having a boyfriend embarrassing now?”, as expressed through public online comments.

Dateset Description

To address the above questions, this study draws on publicly available data collected via the official YouTube Data API v3 (Google, 2025). Among major social media platforms, YouTube offers the most accessible official API, allowing data collection without prior application approval. Although ethical considerations regarding informed consent remain, YouTube is deemed a suitable platform for the analysis due to its impactful scale as the world’s second most visited website (Arthurs et al., 2018). Given the time constraints of this study, YouTube’s API also provided a practical means of accessing large-scale online discourse with a relatively generous quota allocation of 10000 units per day (Google, 2025). Alternative platforms, including Reddit, Threads, TikTok, Instagram, and X (formerly Twitter), were considered but excluded due to restricted API access.

A stepwise approach was employed to collect YouTube comments related to the public debate surrounding the British Vogue article. First, relevant videos were identified using a keyword-based search query (“is having a boyfriend embarrassing now”) submitted to YouTube’s search endpoint. For each identified video, top-level comments were retrieved using the commentThreads endpoint, and replies were excluded to maintain analytical consistency.

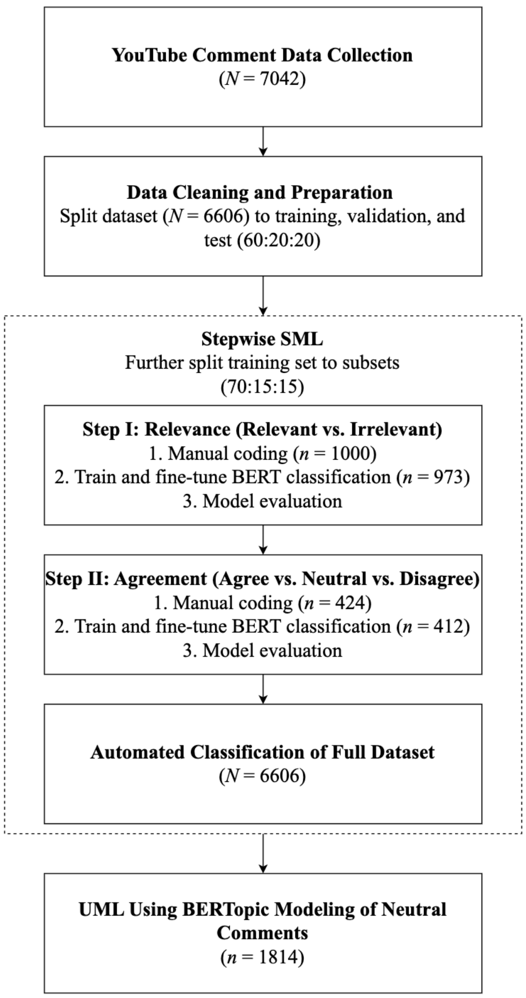
A total of 7042 comments were collected from 49 videos between October 29, 2025, and December 4, 2025. After removing non-English content and non-analyzable records, such as URLs, in-text timestamps, and decorative symbols (e.g., miscellaneous symbols and pictographs), the cleaned dataset comprised 6606 comments, including English text and emojis. The dataset was converted into a Pandas dataframe, where each observation included the comment text and its cleaned version, temporal metadata (i.e., comment and video upload dates), an engagement indicator (i.e., number of likes), and video-level information (i.e., video ID, title, URL, and description). This dataset is suitable for the study for three main reasons: YouTube comments offer genuine, unsolicited opinions; the dataset is large enough to support both supervised and unsupervised analyses; and using the official API ensures transparent and reproducible data collection.

Analytical Strategy

This study employs a stepwise mixed-method machine learning approach, combining manual content analysis, supervised machine learning (SML), and unsupervised machine learning (UML) to analyze large-scale YouTube data in Python 3.12.7 (see Figure 1 for the flowchart). First, SML was applied to classify the comments because it efficiently handles implicit variables in large datasets when a smaller subset can be manually coded (Boumans & Trilling, 2016). To ensure data integrity and prevent information leakage, the cleaned dataset was first split into training (60%), validation (20%), and test (20%) sets, a widely used ratio balancing model learning and evaluation (Zhou et al., 2024). Manual coding was performed exclusively on the training set, which was further split into training (70%), validation (15%), and test (15%) subsets for model training, hyperparameter tuning, and performance evaluation (Rimal et al., 2025). The final BERT classifiers were then applied to the full dataset.

Figure 1

Analytical Strategy Flowchart



Step 1: Supervised Machine Learning for Relevance Classification

In this study, because not all comments directly address the focal statement, a relevance filtering step was conducted before substantive analysis. A random sample of 1000 comments was manually coded by one researcher for relevance (*1 = relevant*, *0 = irrelevant*). After removing duplicates, 973 coded comments remained and were split into training (70%), validation (15%), and test (15%) sets. A BERT-based transformer model was fine-tuned for relevance classification. BERT was chosen because transformer models outperform traditional classifiers (e.g., bag-of-words or TF-IDF models) in sentiment analysis and capturing contextual meaning in short, informal texts such as social media comments (Bikku et al., 2023; Bilal & Almazroi, 2022; Talukder et al., 2025), an advantage also emphasized in the course. Model performance was evaluated on a validation set. With a moderate training loss of .47 over four epochs, the final model achieved an acceptable F1 score of .695, a recall of .661, a precision of .732, and an accuracy of .753 (see Table 1 for stepwise training evaluation). These metrics show that the model performs well and there was no overfitting (Saito & Rehmsmeier, 2015; Goodfellow et al., 2016).

Table 1

Stepwise Training Evaluation of the First BERT Model

Step	Training Loss	Validation Loss	F1	Recall	Precision	Accuracy
50	.656	.597	.529	.371	.920	.719
100	.431	.575	.563	.435	.794	.712
150	.350	.556	.695	.661	.732	.753

Step 2: Supervised Machine Learning for Agreement Classification

Relevant comments were further analyzed for stance toward the focal statement. 424 relevant comments from the sample were annotated for agreement into three categories: agree (1), neither agree nor disagree (0), and disagree (-1). After removing duplicates, 412 comments remained and were again split into training (70%), validation (15%), and test (15%) sets, using the same rationale as in Step 1 to balance model training and evaluation. A second BERT classifier was fine-tuned to predict agreement categories. Class weighting was applied during training to address class imbalance, improving performance on underrepresented categories such as agree and disagree (Chen et al., 2024). Over five epochs, the final model achieved an average F1 score of 0.616 (see Table 2 for stepwise training evaluation). The trained model was then applied to all comments identified as relevant in Step 1, yielding automated agreement classifications for 2974 comments.

Table 2

Stepwise Training Evaluation of the Second BERT Model

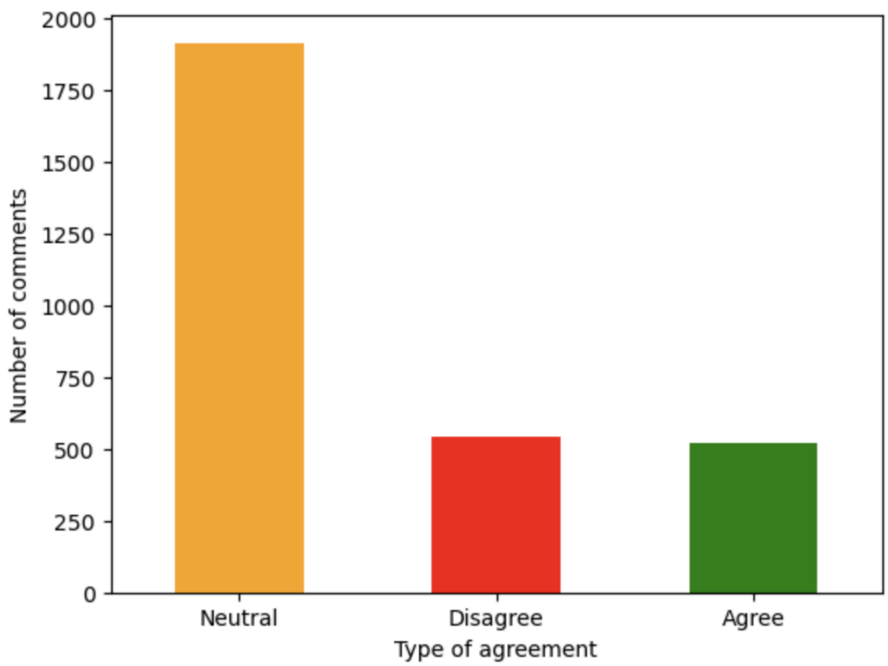
Step	Training Loss	Validation Loss	F1	Recall	Precision	Accuracy
50	.000	2.882	.661	.610	.639	.661
100	.000	3.113	.613	.533	.568	.613

150	.000	2.931	.661	.619	.630	.661

RQ1 examined how much people agree, disagree, or neither agree nor disagree with the statement "It's embarrassing to have a boyfriend now". Most relevant comments neither agreed nor disagreed (64.36%), whereas 17.45% expressed agreement and 18.19% expressed disagreement (see Figure 1). In terms of audience engagement, agreeing comments received the highest average number of likes (M = 40.38, SD = 156.03), followed by neutral (M = 19.82, SD = 83.72) and disagreeing comments (M = 18.17, SD = 95.71). Interestingly, although agreeing comments occurred the least frequently, they elicited the highest level of engagement.

Figure 1

Distribution of Comment Agreement With the Focal Statement



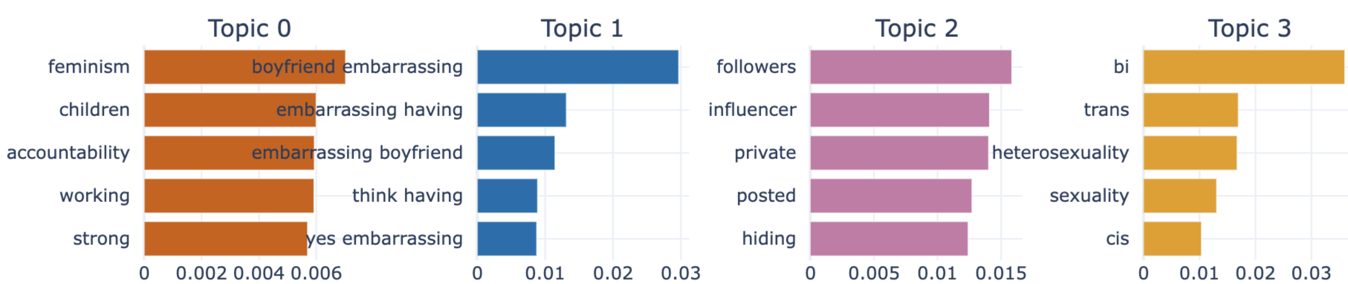
Step 3: Unsupervised Machine Learning for Topic Exploration

To explore the content of comments that neither agreed nor disagreed with the focal statement, BERTopic, an unsupervised topic modeling approach, was applied to this subset (n = 1914). BERTopic was chosen because it combines transformer-based embeddings with clustering and topic representation, generating more semantically coherent topics than traditional methods such as Latent Dirichlet Allocation (LDA), particularly for short texts (Grootendorst, 2022).

RQ2 examined the major topics discussed in comments that neither agreed nor disagreed with the focal statement. The BERTopic model identified five topics, including one outlier (Topic -1) with heterogeneous comments that could not be clustered (38.48% of neutral comments). The substantic topics (see Figure 2) were interpreted qualitatively: Topic 0, focused on gender roles and social expectations (24.09%), illustrated by "Every time you speak, it all makes more sense, I hate the stigma around balancing work and family…", which aligns with the article's discussion of shifting norms around how relationships and partnership roles are valued and displayed online; Topic 1, centered on mixed opinions about romantic relationships (16.36%), illustrated by "As a former woman, I do not think having a boyfriend is embarrassing, but social norms make it awkward…"; Topic 2, relating to influencer culture and privacy (13.99%), illustrated by "…as a heterosexual man, I feel like it's okay not to share everything online…", which reflects the article's emphasis on navigating privacy and selectively sharing relationship information; and Topic 3, concerning sexual identity and orientation (7.72%), illustrated by "The thing about being a bi girl with a boyfriend is navigating others' assumptions…".

Figure 2

Top Keywords for Each Topic



Conclusion

The findings of this study provide insight into public reactions to the statement, "Is having a boyfriend embarrassing now?" According to the results, in response to RQ1, most people neither explicitly agreed nor disagreed that it is embarrassing to have a boyfriend now. Interestingly, although comments expressing agreement were less frequent than those expressing disagreement or neutrality, they received the highest number of likes, suggesting that many people resonated with these perspectives even though they did not leave their own comments. Regarding RQ2, participants who remained neutral discussed multiple topics, including gender roles and social expectations, mixed opinions about romantic relationships, influencer culture and privacy, and sexual identity and orientation. This indicates that neutral comments are often multifaceted, connecting personal experiences with broader societal debates and highlighting avenues for future research. It also aligns with observations in the introduction about evolving cultural attitudes toward relationships and the curated performance of independence in digital spaces (Joseph, 2025; Twenge, 2023; McRobbie, 2020).

While the study employed state-of-the-art methods, several limitations should be noted. First, data were sourced solely from YouTube and included only English-language comments, limiting the generalizability of the findings. Moreover, model performance was constrained by single-researcher manual coding, class imbalance (i.e., some categories had far fewer examples for the model to learn from), and interpretive ambiguity inherent in automatically generated topics. As recommendations, this study encourages future studies to improve reliability and depth by incorporating multiple coders, collecting data from diverse platforms, and using longitudinal analyses to better capture evolving patterns of online engagement.

References

Arthurs, J., Drakopoulou, S., & Gandini, A. (2018). Researching YouTube. *Convergence: The International Journal of Research into New Media Technologies, 24*(1), 3–15. https://doi.org/10.1177/1354856517737222

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism, 4*(1), 8–23. https://doi.org/10.1080/21670811.2015.1096598

Bikku, B., Jarugula, J., Kongala, L., Tummala, N. D., & Donthiboina, N. V. (2023). *Exploring the effectiveness of BERT for sentiment analysis on large-scale social media data.* https://doi.org/10.1109/conit59222.2023.10205600

Bilal, M., & Almazroi, A. A. (2022). Effectiveness of fine-tuned BERT model in classification of helpful and unhelpful online customer reviews. *Electronic Commerce Research*. https://doi.org/10.1007/s10660-022-09560-w

Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, P. (2024). A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review, 57*(6). https://doi.org/10.1007/s10462-024-10759-6

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Google. (2025). *YouTube Data API v3* [Web API]. https://developers.google.com/youtube/v3

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Joseph, C. (2025, October 25). *Is having a boyfriend embarrassing now?* British Vogue. https://www.vogue.co.uk/article/is-having-a-boyfriend-embarrassing-now

McRobbie, A. (2020). *Feminism and the politics of resilience: Essays on gender, media, and the end of welfare*. Polity Press.

Radford, A. (2025, November 13). *Is having a boyfriend embarrassing now? Here's what the author says*. CNN. https://edition.cnn.com/2025/11/13/business/video/vogue-article-boyfriend-embarrassing-creators-digvid

Rimal, Y., Sharma, N., Siddhartha Paudel, Abeer Alsadoon, Madhav Parsad Koirala, & Gill, S. (2025). Comparative analysis of heart disease prediction using logistic regression, SVM, KNN, and random forest with cross-validation for improved accuracy. *Scientific Reports, 15*(1). https://doi.org/10.1038/s41598-025-93675-1

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), e0118432. https://doi.org/10.1371/journal.pone.0118432

Talukder, M. A., Uddin, M. A., Roy, S., Ghose, P., Sarker, S., Ansam Khraisat, Kazi, M., Rahman, M. M., & Hakimi, M. (2025). A hybrid deep learning model for sentiment analysis of COVID-19 tweets with class balancing. *Scientific Reports, 15*(1). https://doi.org/10.1038/s41598-025-97778-7

Twenge, J. M. (2023). *Generations: The real differences between Gen Z, Millennials, Gen X, Boomers, and Silents – and what they mean for America's future*. Atria Books.

Zhou, B., Liu, J., Cui, S., & Zhao, Y. (2024). A large-scale spatio-temporal multimodal fusion framework for traffic prediction. *Big Data Mining and Analytics, 7*(3), 621–636. https://doi.org/10.26599/bdma.2024.9020020

Appendix:
Supplementary Material

Figure 3

Word Cloud for Each Topic