

Analysis of satisfaction perceptions through online comments: The case of Artvin

Çevrimiçi yorumlar üzerinden memnuniyet algısının analizi: Artvin ili örneği

Muhammed Çağrı Aksu¹ 
İşıl Çiftçi⁴ 

Sılanur Soyar² 
Berra Uludere⁵ 

Emine Nisa İlkutlu³ 

¹ Assist. Prof. Dr., Artvin Coruh University, Faculty of Engineering, Department of Computer Engineering, Artvin, Türkiye



² Undergraduate Student, Artvin Coruh University, Faculty of Engineering, Department of Computer Engineering, Artvin, Türkiye



³ Undergraduate Student, Hacettepe University, Faculty of Engineering, Department of Geomatics Engineering, Ankara, Türkiye



⁴ Undergraduate Student, Hacettepe University, Faculty of Engineering, Department of Geomatics Engineering, Ankara, Türkiye



⁵ Undergraduate Student, Hacettepe University, Faculty of Engineering, Department of Geomatics Engineering, Ankara, Türkiye



Corresponding Author:
Muhammed Çağrı Aksu

Citation:

Aksu, M.Ç., Soyar, S., İlkutlu, E.N., Çiftçi, İ. & Uludere, B. (2026). Analysis of satisfaction perceptions through online comments: The case of Artvin. *OPUS- Journal of Society Research*, 23, e1933468. <https://doi.org/10.26466/opusjsr.1933468>

Open Access Statement:

This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). This license permits unrestricted use, distribution, and reproduction in any medium, provided that the original author(s) and the source are properly credited. The original publication in *OPUS Journal of Society Research* must be cited in accordance with accepted academic practice.

Review Note:

Evaluated by Double-Blind Peer Review

Ethics Reporting:

To report potential ethical concerns, contact: editorialoffice@opusjournal.net

Similarity screening was conducted via intihal.net.



Abstract

Social media and online review platforms provide valuable data for monitoring public perceptions of local environments, services, and everyday experiences. This study examines the sentiment orientation, temporal variation, and thematic structure of online comments related to Artvin through a multi-platform dataset collected from Facebook, Twitter/X, TikTok, Google Maps, Instagram, and YouTube. Artvin was selected as the research context because its natural and cultural attractions, mountainous geography, transportation and infrastructure conditions, environmental transformation processes, and everyday service experiences generate a multidimensional local digital discourse. After meaningless and duplicate content was removed, the comments were classified as negative, neutral, and positive using a BERT-based Turkish sentiment analysis model, and very short comments were additionally filtered at the topic modeling stage. Monthly sentiment scores were analyzed through time series and change point analyses, while BERTopic outputs were interpreted by grouping representative terms and comments under higher-order themes. The findings indicate that negative comments were more dominant in the digital discourse related to Artvin and that a statistically significant structural break occurred in February 2023. Topic modeling showed that negative discourse mainly focused on artificial or automatically generated content, transportation and infrastructure problems, local administration and political discussions, dam/expropriation processes, environmental impacts, and economic difficulties. Positive discourse clustered around gratitude, good wishes, natural beauty, scenery appreciation, and cultural values. The study shows that multi-platform social media analysis can support local public perception monitoring and contribute to data-driven regional decision-making processes.

Keywords: Social media analysis, sentiment analysis, time series analysis, topic modeling, BERT, Artvin

Öz

Sosyal medya ve çevrimiçi yorum platformları, bireylerin yerel çevreye, hizmetlere ve gündelik yaşam deneyimlerine ilişkin algılarını izlemek için önemli veri kaynakları sunmaktadır. Bu çalışma Artvin'e ilişkin çevrimiçi yorumların duygu yönelimini, zamansal değişimini ve tematik yapısını Facebook, Twitter/X, TikTok, Google Maps, Instagram ve YouTube'dan elde edilen çok platformlu bir veri kümesi üzerinden incelemektedir. Artvin, doğal ve kültürel çekicilikleri, dağlık coğrafi yapısı, ulaşım ve altyapı koşulları, çevresel dönüşüm süreçleri ve gündelik hizmet deneyimleri nedeniyle çok boyutlu yerel dijital söylem üreten bir araştırma bağlamı olarak seçilmiştir. Anlamsız ve yinelenen içerikler temizlendikten sonra yorumlar, Türkçe metinler için uygun BERT tabanlı bir duygu analizi modeliyle negatif, nötr ve pozitif olarak sınıflandırılmış; çok kısa yorumlar ise konu modelleme aşamasında ek filtrelemeye tabi tutulmuştur. Aylık duygu skorları zaman serisi ve kırılma noktası analizleriyle değerlendirilmiş; BERTopic çıktıları ise temsili terimler ve yorumlar dikkate alınarak üst temalar altında yorumlanmıştır. Bulgular, Artvin'e ilişkin dijital söylemde negatif yorumların daha baskın olduğunu ve Şubat 2023'te istatistiksel olarak anlamlı bir yapısal kırılma meydana geldiğini göstermektedir. Konu modelleme sonuçlarına göre negatif söylem daha çok yapay veya otomatik üretilmiş içeriklere yönelik tepkiler, ulaşım ve altyapı sorunları, yerel yönetim ve siyasal tartışmalar, baraj/kamulaştırma süreçleri, çevresel etkiler ve ekonomik zorluklar etrafında yoğunlaşmıştır. Pozitif söylem ise teşekkür, iyi dilek, doğal güzellikler, manzara beğenisi ve kültürel değerler çevresinde kümelenmiştir. Çalışma çok platformlu sosyal medya analizinin yerel kamu algısının izlenmesine ve veri temelli bölgesel karar destek süreçlerine katkı sağlayabileceğini göstermektedir.

Anahtar Kelimeler: Sosyal medya analizi, duygu analizi, zaman serisi analizi, konu modelleme, BERT, Artvin

Introduction

Monitoring public perception, lived experience, and service satisfaction at the local level is important for resource prioritization, service design, and public communication. Traditional tools such as surveys provide valuable evidence, but they may reflect rapid changes in the public agenda with delay due to cost, access, sampling, and periodic implementation constraints. Social media and online review platforms, in contrast, generate continuous and high-volume digital traces that reveal users' spontaneously expressed experiences and attitudes toward places, services, and events. These traces offer a useful basis for monitoring the multi-thematic agenda of a local context over time.

Previous studies have shown that user-generated content can affect perception and preference behavior and that satisfaction and image components can be examined through online reviews (Aydın, 2016; Belber & Eker, 2022; Eren & Kuvvetli, 2017). This is especially visible in tourism and service experience studies, where online reviews provide insights into destination image and perceived service quality. In the context of Artvin, social media data have been evaluated through touristic locations, and artificial intelligence-based approaches have been discussed for tourism decision support (Aksu & Karaman, 2022). Similarly, sentiment analysis and topic extraction on restaurant reviews in Artvin have shown that online feedback can reveal local service experience components (Durmuş, 2022). However, the local agenda is not limited to tourism or service experience. Online platforms also reflect transportation and infrastructure, local governance, economic conditions, environmental debates, and everyday life experiences. Therefore, a multi-platform and multi-year analysis can provide a broader and more continuous understanding of local public perception.

Sentiment analysis is one of the main components of this type of monitoring. It aims to automatically determine subjective orientation in texts, yet both lexicon-based and supervised classification-based approaches may face errors caused by context dependency, irony, and domain-specific

vocabulary (Pang & Lee, 2008). With the development of Transformer architectures, pre-trained language models have become widely used in text classification. BERT, in particular, has become a strong backbone for classification tasks because of its bidirectional contextual representations and domain adaptability (Devlin et al., 2019). In Turkish social media texts, however, morphological richness, spelling variation, abbreviations, slang, emojis, and platform-specific tags create challenges for data preparation and model generalizability. Studies on Turkish tweet sentiment analysis show that BERT-like models can produce applicable results for short social media texts (Yürütücü & Demir, 2023), while also indicating that model performance may vary according to dataset characteristics and training strategies (Arzu & Aydoğan, 2023). Moreover, preprocessing decisions may affect Turkish sentiment analysis results differently across datasets, making methodological transparency important (Görmez et al., 2024). Recent work also suggests that general-domain models may be limited for noisy and short social media texts and that social media-specific pretraining may offer performance and inference-cost advantages (Najafi & Varol, 2024).

Although sentiment analysis summarizes whether comments are positive, negative, or neutral, it does not fully explain the content areas through which these orientations are formed. Topic modeling is therefore needed to identify the thematic structure underlying sentiment. LDA, one of the classical approaches, models latent themes through document-word distributions (Blei et al., 2003). However, for short and heterogeneous social media texts, contextual embedding-based methods may improve topic coherence and interpretability (Bianchi et al., 2021). In this framework, BERTopic has become a useful approach for producing more interpretable topic clusters in short texts. Turkish examples show that BERTopic can support the dynamic tracking of customer complaint themes (Koruyan, 2022), and comparative work on Twitter data related to the February 6, 2023 earthquake indicates that BERTopic may

provide stronger topic elaboration and differentiation than LDA in certain contexts (Karadaş et al., 2025). Thus, in local online discourse monitoring, it is important to examine not only how positive or negative the discourse is, but also through which themes it is constructed.

The temporal dimension of sentiment is also important for identifying when local online discourse changes. Social media sentiment series often fluctuate due to events, agenda shifts, and platform dynamics. For this reason, possible regime changes should be tested statistically rather than interpreted only through visual inspection. Change point detection is a well-established approach for identifying abrupt changes in time-series characteristics such as mean or distribution (Aminikhanghahi & Cook, 2017). The PELT method is particularly useful for detecting multiple breakpoints efficiently in large datasets because of its near-linear computational cost (Killick et al., 2012). In social media research, change point analysis has been used to identify turning points in sentiment time series, showing that sentiment-time series integration can quantitatively track event-related shifts (Theocharopoulos et al., 2023). This approach can therefore support the systematic identification of periods in which local digital discourse changes.

Artvin was selected as the research setting because it combines multiple dimensions of local discussion, including natural and cultural attractions, tourism, transportation, infrastructure, environmental transformation, public services, and economic conditions. Its mountainous geography, natural beauty, tourism potential, transformations around the Çoruh Basin, and everyday service experiences make Artvin a meaningful case for monitoring online public perception. In this respect, Artvin generates digital discourse not only through tourist appreciation and destination image, but also through infrastructure, transportation, health, environment, economic life, and local governance. Multi-platform comments related to Artvin therefore provide an appropriate empirical basis for examining sentiment orientations, temporal changes, and thematic patterns at the local level.

The literature indicates three main gaps in the context of Artvin. First, integrated studies that analyze multi-platform user comments about a local geographical unit through a sentiment-time-theme framework remain limited. Existing studies generally focus on a single platform or a specific sub-context. Second, since many Turkish sentiment analysis studies rely on short time intervals, structural break analysis in multi-year sentiment series is relatively rare. Third, reproducible methodological designs that combine sentiment classification, change point detection, and topic modeling within a local-scale data science framework are still limited. This study addresses these gaps through a long-term, multi-platform dataset and an integrated analytical design.

Accordingly, this study aims to examine together (i) the long-term trajectory of sentiment polarity, (ii) structural breaks in monthly sentiment scores, and (iii) the thematic patterns in which positive and negative comments are concentrated, based on multi-platform social media comments related to Artvin. Within this scope, the following research questions were addressed:

RQ1: What pattern does the long-term sentiment orientation of multi-platform social media comments related to Artvin exhibit in terms of positive, neutral, and negative classes and continuous sentiment scores?

RQ2: Are there statistically significant structural breaks in the monthly sentiment time series, and how do the pre-break and post-break periods differ?

RQ3: Under which main themes do positive and negative comments cluster, and which dimensions of the local context do these themes reflect?

In summary, this study applies Transformer-based sentiment classification to long-term, multi-platform Turkish user comments, converts sentiment scores into a monthly time series, tests regime differentiation through change point analysis, and combines sentiment classes with topic modeling. In this way, the study contributes to local-level digital discourse monitoring by examining online

discourse related to Artvin not only in terms of sentiment direction, but also in terms of thematic content and temporal change.

Materials and Methods

This study adopted a data-driven research design combining quantitative content analysis, machine learning-based sentiment classification, time series analysis, change point detection, and topic modeling to examine online comments related to Artvin. The analytical process was structured according to the CRISP-DM methodology, which is widely used in data mining studies (Wirth & Hipp, 2000). The workflow consisted of data collection, preprocessing, sentiment classification, temporal analysis, change point detection, topic modeling, and integrated interpretation of the findings.

The dataset consisted of user comments collected from multiple social media and online review platforms between 10.04.2013 and 10.02.2026. The multi-platform structure enabled the analysis of comments produced in different interaction contexts within a common framework. After cleaning the raw data, a BERT-based Turkish sentiment analysis model was used to assign sentiment labels and scores to each comment. Monthly sentiment scores were then calculated for time series analysis, and change point detection was applied to identify structural shifts. Finally, BERTopic was used separately on positive and negative comments to reveal the thematic structure of the discourse. The general workflow of the study is presented in Figure 1.

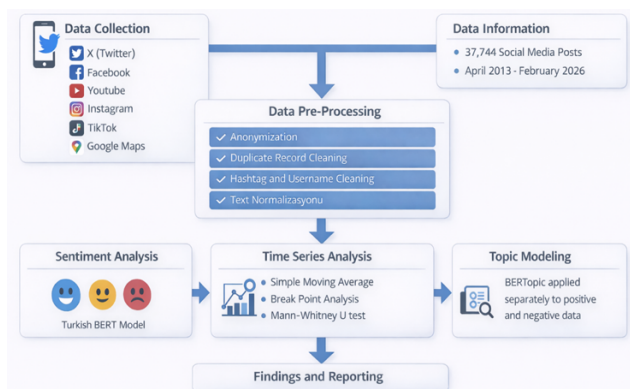


Figure 1. Research workflow and analysis process

Data Collection Process and Ethics

The dataset was created from publicly accessible comments related to Artvin. For each record, only comment text, posting date, and platform information were included in the analytical dataset. These variables formed the basis for sentiment analysis, time series analysis, and topic modeling.

Data collection was conducted through Apify, a third-party API-based data collection platform (Apify, n.d.). Only textual content available in public posts, videos, location pages, and comment fields was considered. Closed accounts, private groups, personal messages, and content requiring special access permission were excluded. Usernames, profile links, user IDs, contact information, device information, and similar personally identifiable data were not transferred to the analytical dataset.

The raw dataset covered the period from 10.04.2013 to 10.02.2026 and included 37,744 comments. The platform distribution was as follows: Facebook (11,234), Twitter/X (10,010), TikTok (7,428), Google Maps (5,076), Instagram (2,794), and YouTube (1,202). This structure indicates that the study was not limited to a single platform, but reflected a heterogeneous set of online commenting practices.

Since the dataset covers a long period, changes in platform access policies, API restrictions, search and visibility algorithms, and historical data availability were considered as methodological limitations. Therefore, the dataset should not be interpreted as the complete universe of online discourse related to Artvin. Rather, it represents publicly accessible multi-platform comments retrievable through a third-party API-based tool within the specified date range.

Ethically, the study was based on aggregated and anonymized analysis. No attempt was made to identify individual users, and the results were reported only through sentiment distributions, monthly averages, change point findings, and theme-level patterns. In line with the principle of data minimization, personal or technical information not required for the research questions was excluded.

Data Preprocessing

Before analysis, the raw comments were subjected to a multi-stage preprocessing procedure. First, date fields were checked and standardized for use in time series and change point analyses. No invalid-date records were identified.

Second, text fields were examined in terms of analytical relevance. Empty, meaningless, or functionally uninformative comments were removed, resulting in the exclusion of 783 records. Very short comments, especially those containing two words or fewer, were also flagged because of their limited value for topic modeling. These comments accounted for approximately 19.01% of the remaining dataset after excluding meaningless records.

Third, duplicate records were identified and removed to prevent repeated content from biasing the results. The duplication rate was approximately 20.81% of the remaining dataset after excluding meaningless records, and 7,692 duplicate records were deleted. After preprocessing, the dataset was reduced from 37,744 to 29,269 comments for sentiment and time series analyses. For topic modeling, additional filtering was applied to improve thematic separation, and 21,129 comments were used.

Descriptive indicators also showed that the dataset was dominated by short comments. The average comment length was 76.73 characters and 10.97 words, while the median values were 43 characters and 6 words. This right-skewed structure confirms the importance of duplicate removal and low-information text filtering before sentiment and topic analyses.

Sentiment Analysis

A deep learning-based sentiment classification approach was used to determine the sentiment orientation of the comments. The BERT-based model *incidelen/bert-base-turkish-sentiment-analysis-cased*, developed for Turkish texts, was used for classification (İncidelen & Aydoğan, 2025). The

model classified comments into three categories: negative, neutral, and positive.

For each comment, class probabilities were obtained and normalized through the Softmax function. The final sentiment label was assigned according to the highest probability. In addition to class labels, a continuous sentiment score was calculated by subtracting the negative-class probability from the positive-class probability. Thus, scores ranged from -1 to +1, where values closer to -1 indicate negative sentiment, values closer to +1 indicate positive sentiment, and values around zero indicate neutral or mixed sentiment.

The model outputs consisted of three variables for each comment: sentiment label, highest class probability, and continuous sentiment score. These outputs were used for overall sentiment distribution, monthly aggregation, change point analysis, and sentiment-specific topic modeling.

To evaluate the applicability of the model to this dataset, a manual validation procedure was conducted. A stratified random validation subset of 300 comments was created by selecting 100 comments from each model-predicted class. These comments were re-annotated by the researcher according to the same three-class coding scheme. The researcher-assigned labels were then compared with the model outputs using accuracy, precision, recall, F1-score, and a confusion matrix.

After sentiment classification, 15,148 comments were classified as negative, 7,880 as neutral, and 6,241 as positive. These class labels and sentiment scores formed the basis for the subsequent analyses.

Time Series Analysis and Change Point Detection

The continuous sentiment score of each comment was combined with its posting date and aggregated by month. Monthly average sentiment scores were calculated to reduce the effect of daily or weekly fluctuations and to make the underlying trend more interpretable (Basseville & Nikiforov, 1993). A three-month moving average was also calculated for visual interpretation, but statistical

tests were conducted on the original monthly averages rather than the smoothed series.

Change point analysis was applied to determine whether the monthly sentiment series contained a structural shift. The analysis identified February 2023 as the main breakpoint. This date was then used to divide the series into pre-break and post-break periods.

Because the distributional assumptions of monthly sentiment scores may not always be met and the number of observations differed across periods, the Mann–Whitney U test was used to compare the pre-break and post-break periods (Mann & Whitney, 1947). The significance level was set at $\alpha = 0.05$, and Bonferroni correction was considered when multiple comparisons were required (Dunn, 1961). In addition, Cliff's delta was calculated to evaluate the practical magnitude of the difference between periods (Cliff, 1993).

The pre-break mean sentiment score was -0.0027 , while the post-break mean was -0.2865 . The difference was statistically significant (Mann–Whitney $U = 2752.0$, $p < 0.001$), and the effect size was strong (Cliff's delta = 0.699). Since the pre-break mean was very close to zero, percentage change was interpreted cautiously. Therefore, the interpretation mainly relied on period means, Mann–Whitney U results, and Cliff's delta. Data organization, aggregation, visualization, and statistical analyses were conducted using Python-based data analysis tools.

Topic Modeling Procedure

Topic modeling was conducted to identify the content areas around which positive and negative comments clustered. Based on the sentiment labels obtained from the BERT-based classification stage, comments classified as negative and positive were analyzed as two separate subsets. This approach made it possible to examine the thematic foundations of negative and positive discourse comparatively.

The BERTopic approach was preferred because it can produce semantically grounded and inter-

pretable topic clusters in short and context-sensitive texts (Grootendorst, 2022). Since BERTopic is sensitive to embedding quality, vectorization choices, and clustering parameters, the parameter settings were determined by considering both the size of the dataset and the interpretability of the resulting clusters. Text embeddings were generated using the Sentence-Transformer-based paraphrase-multilingual-MiniLM-L12-v2 model because of its multilingual capability and suitability for Turkish texts (Reimers & Gurevych, 2019). Thus, comments were grouped not only according to lexical similarity, but also according to semantic proximity.

Before topic modeling, an additional thematic cleaning process was applied. Links, user mentions, unnecessary symbols, general Turkish stop words, and frequent but weakly discriminative social media expressions were removed. In the vectorization stage, CountVectorizer was used with a unigram + bigram structure ($ngram_range = (1,2)$). To reduce noise, $min_df = 2$ was used for very rare terms, while $max_df = 0.95$ was used for overly common terms.

Separate BERTopic models were run for negative and positive comments. In both models, the $nr_topics = "auto"$ setting was used so that the number of topics could be determined according to the structure of each subset. In addition, BERTopic's outlier-reduction mechanism and c-TF-IDF-based reassignment strategy were applied to reduce noisy assignments and improve topic representations.

The minimum topic size was determined separately for each sentiment-specific subset because the negative and positive datasets differed substantially in size. The negative topic-modeling subset included 12,018 comments, whereas the positive subset included 5,383 comments. Accordingly, $min_topic_size = 100$ corresponded to approximately 0.83% of the negative subset, while $min_topic_size = 30$ corresponded to approximately 0.56% of the positive subset. Thus, the thresholds were not selected arbitrarily, but were defined proportionally by considering subset size, topic stability, and interpretability.

This parameter choice was also necessary to avoid two methodological problems. If the same absolute threshold used for negative comments had been applied to the positive subset, smaller but semantically meaningful positive topics could have been excluded. Conversely, if a very low threshold had been applied to the larger negative subset, the model could have produced many small and fragmented clusters, reducing thematic interpretability. Therefore, the thresholds were used to obtain sufficiently large, semantically coherent, and interpretable topics within each sentiment class, rather than to compare the raw number of topics between negative and positive comments.

BERTopic outputs were evaluated using topic size, representative words, and representative texts. The label -1 was treated as outlier/noise and was not included in thematic interpretation. To increase interpretability, model-generated topics were grouped under higher-order themes. In this process, each topic was first examined in terms of its c-TF-IDF-based representative terms, topic size, and representative comments. Then, topics with conceptual similarity and common relevance to the local context of Artvin were grouped under broader themes.

Theme names were not generated automatically by the model. Instead, they were determined through researcher interpretation by considering representative terms, example comments, semantic proximity among topics, and relevance to the local context. Therefore, word frequency alone was not used as the basis for theme naming. This procedure made it possible to move beyond general sentiment distribution and identify the specific content areas toward which positive and negative evaluations were directed.

Results and Discussion

In this section, the findings are presented in four stages: descriptive characteristics of the dataset, sentiment analysis results, temporal sentiment trends and breakpoint findings, and topic modeling outputs. This structure allows the results to be interpreted from the general data profile toward

sentiment distribution, temporal change, and thematic content.

Descriptive Findings of the Dataset

The raw dataset consisted of 37,744 social media comments covering the period from 10.04.2013 to 10.02.2026. As shown in Table 1, the comments were collected from six platforms: Facebook, Twitter/X, TikTok, Google Maps, Instagram, and YouTube. The highest number of comments came from Facebook and Twitter/X, followed by TikTok and Google Maps. This distribution indicates that the analysis was not limited to a single platform, but reflected different online interaction environments.

Table 1. General characteristics of the dataset and platform distribution

Variable	Value
Date range	10.04.2013 – 10.02.2026
Raw data (number of comments)	37,744
Number of platforms	6
Facebook	11,234
Twitter/X	10,010
TikTok	7,428
Google Maps	5,076
Instagram	2,794
YouTube	1,202

Text-length statistics are presented in Table 2. The mean comment length was 76.73 characters and 10.97 words, while the median values were 43 characters and 6 words. This shows that the dataset was dominated by short comments, which is typical of social media data. In the preprocessing stage, after excluding meaningless records, 19.01% of the remaining comments contained two words or fewer, indicating that a considerable part of the dataset had limited contextual content.

Table 2. Descriptive statistics on text length in the raw dataset

Statistic	Character length	Word length
Number of observations (n)	37,744	37,744
Mean	76.73	10.97
Standard deviation	122.70	16.80
Minimum	1.00	1.00
1st percentile	3.00	1.00
5th percentile	3.00	1.00
Median (50%)	43.00	6.00
95th percentile	258.00	36.00
99th percentile	508.57	71.00
Maximum	4109.00	509.00

The preprocessing results are summarized in Table 3. No invalid-date records were identified. However, 783 meaningless or functionally uninformative records and 7,692 duplicate records were removed. After preprocessing, 29,269 comments remained for sentiment and time series analyses. For topic modeling, additional filtering was applied, and 21,129 comments were used. These steps were important for reducing noise and improving the reliability of the subsequent analyses.

Table 3. Data cleaning (preprocessing) steps and record counts

Cleaning step / indicator	Record count
Initial number of records	37,744
Removed due to meaningless text	783
Very short texts (flagged for topic modeling)	7,028
Removed due to duplicate records	7,692
Final dataset (for sentiment/time series analysis)	29,269
Number of records used for topic modeling	21,129
Proportion of very short comments (≤ 2 words)	19.01%
Duplication rate	20.81%

Sentiment Analysis Findings

Before interpreting the sentiment distribution of the final dataset, the performance of the BERT-based sentiment analysis model was evaluated through a manual validation procedure. For this purpose, a stratified random validation subset consisting of 300 comments was created by selecting 100 comments from each of the model-predicted negative, neutral, and positive classes. These comments were then re-annotated by the researcher according to the same three-class coding scheme, and the researcher-assigned labels were compared with the model predictions.

The overall validation results are presented in Table 4. The model achieved an accuracy of 0.887. The macro precision, macro F1-score, and weighted F1-score values were also 0.887, while macro recall was 0.893. These values indicate that the model produced generally consistent and balanced results across the three sentiment classes.

Table 4. Manual validation results of the sentiment analysis model

Indicator	Value
Validation sample	300 comments
Accuracy	0.887
Macro Precision	0.887
Macro Recall	0.893
Macro F1-score	0.887
Weighted F1-score	0.887

Class-based performance values are presented in Table 5. The negative class had a high recall value of 0.99, indicating that the model correctly captured most comments coded as negative by the researcher. The positive class had a high precision value of 0.98, showing that comments predicted as positive by the model largely overlapped with researcher coding. The neutral class produced relatively lower but still acceptable values. This is expected because neutral social media comments often include ambiguous, short, descriptive, or weakly evaluative expressions.

Table 5. Class-based performance results of the sentiment analysis model

Class	Precision	Recall	F1-score	Support
Negative	0.87	0.99	0.93	88
Neutral	0.81	0.86	0.84	94
Positive	0.98	0.83	0.90	118
Accuracy			0.89	300
Macro avg	0.89	0.89	0.89	300
Weighted avg	0.89	0.89	0.89	300

The confusion matrix in Table 6 provides a more detailed view of the agreement between the researcher-assigned labels and the model predictions. The rows represent the labels assigned by the researcher, while the columns represent the labels predicted by the model.

Table 6. Confusion matrix of the sentiment analysis model

Researcher label / Model label	Negative	Neutral	Positive
Negative	87	1	0
Neutral	11	81	2
Positive	2	18	98

As shown in Table 6, the model correctly classified most negative and positive comments. Of the 88 comments coded as negative by the researcher, 87 were also predicted as negative by

the model. Similarly, 98 of the 118 positive comments were correctly classified. More transitions were observed in the neutral class, where 11 comments were classified as negative and 2 as positive. In addition, 18 positive comments were classified as neutral. These results suggest that the model performed strongly in detecting clear positive and negative sentiment, while neutral expressions were more difficult to distinguish because of their ambiguous and context-dependent structure.

Overall, the manual validation findings support the applicability of the BERT-based model to the Artvin dataset. The accuracy, F1-score values, and confusion matrix indicate that the model outputs are sufficiently reliable for use in the subsequent time series, breakpoint, and topic modeling analyses.

The sentiment distribution of the final dataset is presented in Figure 2. Of the 29,269 comments, 15,148 were classified as negative, 7,880 as neutral, and 6,241 as positive. This corresponds to 51.76% negative, 26.92% neutral, and 21.33% positive comments. The dominance of negative sentiment suggests that online discourse related to Artvin was largely shaped by criticism, complaints, and dissatisfaction. This finding is consistent with the literature indicating that social media is frequently used for expressing complaints and negative experiences (Sigurdsson et al., 2021), while also reflecting broader evaluative content (Pang & Lee, 2008).

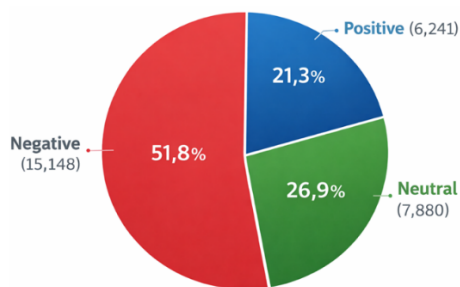


Figure 2. Distribution of sentiment classes in the final dataset

The relatively high share of neutral comments also deserves attention. It shows that the dataset included not only clear evaluations but also informational, descriptive, questioning, or weakly emo-

tional expressions. This is consistent with the heterogeneous and short-text nature of social media data. Although Transformer-based models provide strong contextual representations, class boundaries may still be affected by dataset characteristics, training strategies, and preprocessing decisions (Arzu & Aydoğan, 2023; Devlin et al., 2019; Görmez et al., 2024). Therefore, the sentiment distribution should not be interpreted as a direct measurement of the overall public attitude toward Artvin, but as an indicator of how Artvin-related discourse appears across publicly accessible online platforms.

Temporal Sentiment Trends and Breakpoint Findings

The continuous sentiment scores calculated for each comment were aggregated at the monthly level in order to examine the temporal trajectory of online discourse related to Artvin. The change in monthly average sentiment scores over time, together with the moving-average curve and the segment means derived from breakpoint analysis, is presented in Figure 3. Figure 3 shows that the series exhibited a fluctuating structure throughout the study period, but that after a certain date the average level shifted toward a more negative regime. Such short-term fluctuations are expected in social-media-based sentiment series because online sentiment production is often sensitive to agenda topics, events, and platform dynamics. Therefore, rather than evaluating the trend only through visual inspection, statistically testing possible regime changes provides a stronger analytical approach (Aminikhanghahi & Cook, 2017).

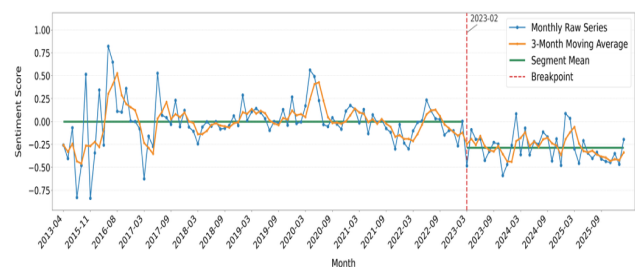


Figure 3. Time series and breakpoint of the monthly average sentiment scores of social media comments.

The breakpoint analysis identified the main structural breakpoint in the series as February 2023 (2023-02). Change point methods provide an effective framework for detecting abrupt changes in time-series characteristics such as mean or distribution (Aminikhanghahi & Cook, 2017). Among these approaches, PELT is frequently used because it enables efficient breakpoint detection with near-linear computational cost (Killick et al., 2012). In this study, breakpoint analysis was used to test whether the visually observed increase in negativity corresponded to a statistically supported regime change.

that sentiment-time series integration can make event-related shifts quantitatively visible (Theodoropoulos et al., 2023).

To enable a more detailed visual evaluation of the change point, the period around February 2023 was also examined separately. In this context, the monthly average sentiment scores and three-month moving average values for the period between January 2022 and December 2023 are presented in Figure 4. This focused interval makes it possible to observe the pattern immediately before and after the detected breakpoint more clearly.

Table 7. Comparison of monthly sentiment scores before and after the breakpoint.

Period	Time range	n (months)	Mean sentiment score	Mann-Whitney U	p	Cliff's delta (δ)
Pre-break	2013-04 – 2023-01	90	-0.0027			
Post-break	2023-02 – 2026-02	36	-0.2865			
Pre-post comparison				2752.0	< 0.001	0.699

The comparative statistics for the pre-break and post-break periods are presented in Table 7. The average sentiment score was -0.0027 in the pre-break period (2013-04 to 2023-01; n = 90 months), whereas it was -0.2865 in the post-break period (2023-02 to 2026-02; n = 36 months). The difference between the two periods was statistically significant (Mann-Whitney U = 2752.0; $p < 0.001$). Cliff's delta was calculated as 0.699, indicating that the observed difference was not only statistically significant but also practically strong. This finding supports the interpretation that comments related to Artvin shifted into a more negative sentiment regime after February 2023.

Since the pre-break average was very close to zero, percentage-based change indicators had limited interpretive value. Therefore, the magnitude and direction of the change were evaluated mainly through segment means, the Mann-Whitney U test, and Cliff's delta. This result indicates that the predominance of negative sentiment was not merely a static distributional feature, but also reflected a temporal shift after a specific point. This interpretation is consistent with studies that generate sentiment predictions from social media texts, derive time series, and identify turning points through change point analysis; such studies show

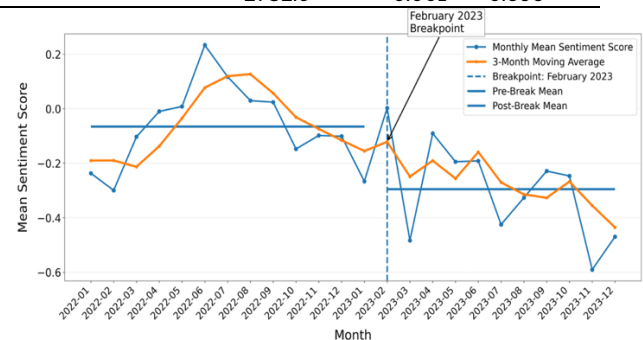


Figure 4. Monthly average sentiment scores and 3-month moving average around the February 2023 change point

As shown in Figure 4, monthly average sentiment scores began to cluster within a more negative range after February 2023. The downward tendency of the three-month moving average curve after the change point also indicates that the change was not merely the result of isolated monthly fluctuations, but pointed to a more persistent trend toward negativity. Taken together, Figure 3, Figure 4, and Table 7 show that there was a statistically significant structural break in the sentiment pattern of social media comments related to Artvin as of 2023-02 and that the discourse remained at a more negative level thereafter.

Topic Modeling Findings

In this section, the findings of the BERTopic-based topic modeling analysis conducted on comments classified as negative and positive following the sentiment analysis are considered together. In short and heterogeneous social media texts, pure sentiment distribution answers the question of “how negative/how positive?”, whereas topic modeling makes visible the question of “through which content axes are these sentiments produced?” Although classical generative topic modeling approaches such as LDA provide a strong framework for extracting latent themes from texts (Blei et al., 2003), evidence suggesting that contextual embeddings can strengthen thematic separation and increase topic coherence in short texts such as social media content supports the preference for approaches such as BERTopic (Bianchi et al., 2021). Applications conducted in the Turkish context also report that BERTopic can produce more detailed and interpretable thematic clusters in short and noisy texts (Karadaş et al., 2025; Koruyan, 2022). In this study as well, the outputs of topic modeling were evaluated separately for each sentiment polarity, and the topics were interpreted on the basis of representative terms and representative comments before being merged under broader themes. The theme distributions and topic correspondences in negative and positive comments are presented in Table 8 and Table 9, respectively.

Thematic Structure in Negative Comments

Topic modeling was conducted separately for negative and positive comments in order to identify the thematic foundations of sentiment polarity. The main themes obtained from negative comments are presented in Table 8. The largest theme was “Reactions to artificial/automatically generated content” (25.96%), followed by “Transportation and Infrastructure Problems” (20.08%) and “Local Administration and Political Discourses” (13.27%). Together, these three themes accounted for a large part of the negative discourse.

This result shows that negative sentiment was not produced only through direct dissatisfaction with local services or places. It was also associated with content credibility, platform-specific reactions, infrastructure problems, and governance-related evaluations. Other themes, such as dam/expropriation and environmental impacts, economic hardship, environmental degradation, street animals, food-service dissatisfaction, and health-service complaints, indicate that negative discourse extended across multiple dimensions of local life.

Table 8. Main themes and representative terms in negative comments

Theme	c-TF-IDF-based representative terms	Total Share (n) (%)
Reactions to artificial/automatically generated content	hiç, ren, böyle, bile, be, şimdi, bak, geyiği, ren geyiği, bunlar	3120 25.96
Transportation and Infrastructure Problems	yol, yolu, yer, güzel, kötü, cam, park, yollar, teras, her, ev, bina, sağlam, beyaz, evler, evi, tünelden, önceki beyaz, tünelden önceki, evleri	2413 20.08
Local administration and political discourses	oy, chp, devlet, belediye, akp, ülke, başkan, olsun, her, başka, sayın, yazıklar, yazıklar olsun, hangi, zaman, bilgi, kaç, bende, hanım	1595 13.27
Dam/expropriation and environmental impacts	su, barajı, barajı, suyu, çoruh, yusu-feli, nehri, nehir, altında, sular	812 6.76
Critical discourses directed at a specific academic figure	öğretim, üyesi, öğretim üyesi, üniversite, keşke, eğitim, kyk, okul, meslek, vah	673 5.60
Post-event expressions of consolation and good wishes	allah, olsun, geçmiş, geçmiş olsun, versin, olsun allah, korusun, beterrinden, inşallah, beterrinden korusun	576 4.79
Economic hardship, cost of living, and price increase discourses	para, zam, bin, tl, iş, lira, vergi, emekli, sonra, her	574 4.78
Environmental degradation and nature conservation discourses	plastik, toprağa, imha, tarım, doğaya, zarar, kasaları, plastik kasaları, bakanlığı, plastikleri, orman, ağaç, ağacı, ormanları, ağaçlar, hocam, ağacın, yangın, ağaçları, yeşil	512 4.26
National/ethnic identity discourses	türk, türkiye, gürcü, gürcüler, ahıska, müslüman, laz, türkiye ye	415 3.45
Street animals and animal welfare	köpek, hayvan, kedi, sokak, hayvanları, hayvanlar, köpekleri, hayvanların, başıboş, tüm	413 3.44
Food-service experience and dissatisfaction	yemek, ekmek, yemeği, kaymakam, sonra, fasulye, peynir, beğenmedim, tatlı, aç	363 3.02
Gender-based student discourses/debates	kızlar, erkekler, erkek, kız, erkeklerin, adalet, burda, kadın, yapsa, hala	290 2.41
Complaints related to health services/health problems	hastane, doktor, sağlık, hasta, devlet, hastanesi, hastahane, dr	262 2.18

Note. Theme labels were translated into English for reporting purposes, whereas the representative words were retained in their original Turkish form in order to preserve the semantic specificity of the BERTopic outputs and to avoid distortions that might arise from literal translation.

The presence of post-event consolation and good-wish expressions among negative comments also suggests that negative sentiment did not always correspond to anger or complaint. In some cases, it reflected sorrow, loss, concern, or support-oriented discourse. Therefore, the negative comment structure should be interpreted as a broad expression space covering complaint, reaction, concern, and social sensitivity.

Thematic Structure in Positive Comments

The thematic distribution of positive comments is shown in Table 9. The most frequent theme was “Expressions of thanks, prayers, and good wishes” (41.19%). This indicates that a substantial part of positive sentiment was expressed through short, formulaic, and socially affirmative expressions.

The second and third most frequent themes were “Appreciation of natural beauty and scenery” (19.73%) and “Nature activities and open-air experience” (16.64%). These themes show that positive discourse related to Artvin was strongly associated with natural beauty, scenery, the Çoruh River, camping, outdoor activities, and place-based appreciation. This finding is consistent with Artvin’s identity as a city known for its natural and cultural landscape.

Other positive themes included service satisfaction and appreciation of effort, praise and success expressions, national/ethnic identity and homeland discourses, and evaluations related to city image and urban amenities. These findings suggest that positive comments were shaped mainly by gratitude, appreciation, nature-oriented satisfaction, and supportive expressions, while urban and service-related evaluations remained more limited.

Overall, the topic modeling results show that negative and positive comments differed not only in sentiment polarity but also in their thematic foundations. Negative discourse was more closely related to complaint, reaction, infrastructure, governance, environmental issues, and content trustworthiness. Positive discourse, on the other hand, was largely shaped by gratitude, approval, natural

beauty, scenery, and open-air experiences. This confirms the importance of combining sentiment analysis with topic modeling when interpreting local online discourse.

Table 9. Main themes and representative terms in positive comments

Theme	c-TF-IDF-based representative terms	Total Share (n)	(%)
Expressions of thanks, prayers, and good wishes	olsun, allah, razı, razı olsun, hayırlı, inşallah, allah razı, emeği, teşekkür, maşallah, helal, helal olsun, seni, bende, gerek, hep, iyi, canım, teşekkür ederim, ederim, öyle, hayırlısı, yakında, hasret, hesap, ilk, tam	2217	41.19
Appreciation of natural beauty and scenery	yer, doğa, harika, muhteşem, gereken, manzara, güzel yer, görülmesi, harikası, görülmesi gereken, atatürk, büyük, heykeli, atatürk heykeli, dünyanın, dünyanın büyük, heykel, heykelini, atepe, kahveci-oğlu	1062	19.73
Nature activities and open-air experience	çoruh, su, kamp, nehri, çoruh nehri, alanı, göl, kadar, yer, baraj, çay, tavsiye, temiz, kahvaltı, uygun, kahve, ücretsiz, yemek, tavsiye ederim, cam, teras, cam teras, teras, cam teras, mutlaka, manzarası, yükseklikte, km	896	16.64
Service satisfaction and appreciation of effort	sağlık, emeğinize, emeğinize sağlık, emeğine, elinize, emeğine sağlık, ellerine, ellerine sağlık, eline, elinize sağlık, belediye, teşekkür, teşekkürler, ederiz, ediyorum, teşekkür ederiz, teşekkür ediyorum, emeği, hopa, iyi	635	11.80
Expressions of appreciation, praise, and success	başkanım, gelsin, kolay gelsin, kolay, başkan, başarılar, sayın, diliyorum, olsun, başarılar diliyorum	330	6.13
National/ethnic identity and homeland discourses	türkiye, türk, nin, ye, türki, türki ye, türkiye nin, dir, vatan, arhavi	137	2.55
Evaluations related to city image and urban amenities	şehir, avm, küçük, tek, tek avm, si, şehri, şehre, nasip olur, gün	106	1.97

Note. Theme labels were translated into English for reporting purposes, whereas the representative words were retained in their original Turkish form in order to preserve the semantic specificity of the BERTopic outputs and to avoid distortions that might arise from literal translation.

Conclusion and Recommendations

This study examined multi-platform social media comments related to Artvin through the combined use of sentiment analysis, time series analysis, change point detection, and topic modeling. By analyzing comments collected from Facebook, Twitter/X, TikTok, Google Maps, Instagram, and YouTube, the study moved beyond a single-platform perspective and provided a broader view of local digital discourse. The findings show that negative sentiment was more dominant in online comments related to Artvin. However, this result

should not be interpreted as a direct or absolute measure of public attitude toward the city. Rather, it should be evaluated by considering that social media users often produce content with motivations such as complaint, criticism, reaction, evaluation, information sharing, and experience reporting.

One of the central findings of the study is the statistically significant structural break observed in monthly average sentiment scores as of February 2023. After this point, sentiment scores shifted to a more negative level, indicating that online discourse related to Artvin did not remain stable over time. This finding shows that social media data can function not only as a descriptive source of public perception, but also as a dynamic source for monitoring temporal changes and possible points of social sensitivity. The February 2023 break temporally overlaps with the national atmosphere following the February 6, 2023 Kahramanmaraş-centered earthquakes. Studies showing the visibility of individual and collective trauma in post-earthquake social media discourse (Kaya Kılıç et al., 2025), the emergence of psychological distress even among individuals not directly affected by the earthquake (Çınaroğlu et al., 2025), and the pressures created by the earthquake on health, economic life, and social welfare (Kahraman, 2023; Nasir et al., 2025) support this contextual interpretation. In the specific context of Artvin, the Yusufeli Dam and relocation process, infrastructure and transportation problems, environmental transformation, economic difficulties, and local governance debates also appear to have contributed to the local content basis of this negativity. The Yusufeli relocation process has also been discussed in relation to cultural change (Çelik Kan, 2024), official resettlement practices (Republic of Türkiye Ministry of Environment, Urbanization and Climate Change, 2022), and local reactions reported during the relocation period (Haberler.com, 2022). Nevertheless, since the study was not designed as an event-based causal analysis, the February 2023 break should be interpreted as a contextual indicator rather than as evidence of a direct causal relationship.

The topic modeling results show that negative and positive discourse differed not only in sentiment polarity, but also in their thematic foundations. Negative comments were concentrated around reactions to artificial or automatically generated content, transportation and infrastructure problems, local administration and political evaluations, dam/expropriation and environmental impacts, economic difficulties, health services, and everyday life problems. Positive comments, on the other hand, clustered around thanks, prayers, good wishes, natural beauty, scenery appreciation, nature activities, and open-air experiences. This differentiation indicates that Artvin's digital representation has a dual structure: a positive image shaped by nature, scenery, cultural values, and appreciation, and a critical discourse shaped by infrastructure, public services, environmental change, economic difficulties, and local governance.

When evaluated together with previous studies on Artvin, the contribution of this study becomes clearer. Aksu and Karaman (2022) showed that social media data can be used in relation to touristic places and artificial intelligence-supported tourism decision support, while Durmuş (2022) demonstrated that restaurant reviews can reveal service experience components through sentiment analysis and topic extraction. The present study extends this line of research by examining Artvin through a long-term, multi-platform, and sentiment-time-theme integrated framework. In this respect, it shows that local digital discourse is shaped not only by tourism or service satisfaction, but also by a broader agenda involving transportation, infrastructure, environmental transformation, economic conditions, health services, governance, and digital content credibility.

Methodologically, the study demonstrates that local public perception can be monitored by integrating BERT-based sentiment classification, monthly sentiment time series, PELT-based change point analysis, and BERTopic-based topic modeling. The manual validation results also support the reliability of the sentiment analysis outputs for subsequent temporal and thematic analyses. This integrated framework can be used as a

complementary approach in local governance, public communication, tourism planning, and regional decision-support processes.

The findings also offer practical implications for local decision-makers. Transportation and infrastructure problems, economic difficulties, health-service complaints, environmental debates, dam and relocation processes, and criticisms of local administration should be monitored regularly as priority areas. Social media and online review data can support the early detection of such problem areas and help institutions develop data-driven communication strategies. At the same time, Artvin's positive digital image is strongly associated with natural beauty, scenery, open-air experiences, cultural values, gratitude, and good wishes. Therefore, local institutions should not only respond to negative discourse, but also strengthen communication and promotion strategies that make these positive themes more visible.

The study has several limitations. First, the findings are based only on users who share content on social media and online review platforms; therefore, they should not be treated as a representative public opinion survey. Second, the platforms differ in terms of user profiles, interaction styles, algorithmic visibility, and expression cultures. Third, the dataset was limited to publicly accessible content collected through third-party API-based tools, and platform-specific access restrictions may have influenced the scope of the data. Fourth, sentiment analysis and topic modeling results are affected by model performance, preprocessing choices, embedding models, and clustering parameters. Finally, although the February 2023 break indicates a meaningful temporal shift, its causes were not tested causally.

Future studies may compare social media findings with survey, interview, or institutional complaint data to obtain a more comprehensive understanding of local public perception. In addition, platform-specific analyses, event-based designs, larger manually annotated validation datasets, and comparative studies involving different provinces may strengthen the generalizability and explanatory power of similar research. Overall, this study

shows that multi-platform social media data, when analyzed through an integrated sentiment-time-theme framework, can provide valuable evidence for monitoring local digital discourse and supporting data-driven regional decision-making.

Declarations

Funding: This research was supported by TÜBİTAK 2209-A Research Project Support Programme for Undergraduate Students under grant/project number 1919B012425078. The role of the funding body was as follows: The funding provided by TÜBİTAK supported the data collection process only. The funding body had no role in the study design, data analysis, interpretation of the findings, writing of the manuscript, or decision to submit the article for publication.

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

Declaration of Previous Presentation and Publication: The preliminary findings of this study were presented as an oral abstract at the Çukurova International Congress of Basic and Applied Sciences on 13 March 2026.

Ethical Approval: Ethical committee approval for this study was obtained from Scientific Research and Publication Ethics Committee of Artvin Çoruh University on June 13, 2025 (decision/document number: E-18457941-050.99-181928).

Informed Consent: Informed consent was not applicable because the study did not involve human participants.

Data Availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions: Conceptualization, M.Ç.A. (60%), S.S. (10%), E.N.İ. (10%), I.Ç. (10%), B.U. (10%); introduction and conceptual analysis, M.Ç.A.

(80%), S.S. (5%), E.N.İ. (5%), I.Ç. (5%), B.U. (5%); investigation and data analysis, M.Ç.A. (20%), S.S. (20%), E.N.İ. (20%), I.Ç. (20%), B.U. (20%); writing – original draft preparation, M.Ç.A. (60%), S.S. (10%), E.N.İ. (10%), I.Ç. (10%), B.U. (10%); writing – review and editing, M.Ç.A. (40%), S.S. (15%), E.N.İ. (15%), I.Ç. (15%), B.U. (15%).

AI Disclosure: ChatGPT (GPT-5.5) was used for language editing, translation support, and improving the clarity of the manuscript. The authors reviewed, edited, and verified all AI-assisted outputs and take full responsibility for the content of the manuscript.

References

- Aksu, M. Ç., & Karaman, E. (2022). Turistik Mekanlara Yönelik Sosyal Medya Paylaşımlarının Yapay Zekâ Yöntemleriyle Değerlendirilmesi: Artvin İli Örneği (Evaluation of Social Media Posts for Touristic Places with Artificial Intelligence Methods: The Case of Artvin Province). *Journal of Tourism and Gastronomy Studies*, 10(1), 505-524. <https://doi.org/10.21325/jo-tags.2022.1001>
- Aminikhanghahi, S., & Cook, D. J. (2017). A Survey of Methods for Time Series Change Point Detection. *Knowledge and Information Systems*, 51(2), 339-367. <https://doi.org/10.1007/s10115-016-0987-z>
- Apify. (n.d.). Apify: Full-stack web scraping and data extraction platform. Retrieved June 6, 2026, from <https://apify.com/>
- Arzu, M., & Aydoğan, M. (2023). Comparative Analysis of Transformers Based Architectures for Turkish Sentiment Classification. *Computer Science (IDAP-2023: International Artificial Intelligence and Data Processing Symposium Issue)*, 1-6. <https://doi.org/10.53070/bbd.1350405>
- Aydın, B. (2016). Sosyal Medyada Restoran İmajı: TripAdvisor Örneği. *Disiplinlerarası Akademik Turizm Dergisi*, 1(1), 13-30. <https://doi.org/10.31822/jomat.287998>
- Basseville, M., & Nikiforov, I. V. (1993). Detection of abrupt changes: theory and application. Prentice Hall.
- Belber, B. G., & Eker, S. (2022). Sosyal medyanın destinasyon tercihinin etkisi: Kapadokya örneği. *Yeni Medya* (12), 205-233. <https://doi.org/10.55609/yeni-medya.1090256>
- Bianchi, F., Terragni, S., & Hovy, D. (2021). Pre-training is a hot topic: Contextualized document embeddings improve topic coherence. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers) (pp. 759–766). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.acl-short.96>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022. <https://www.jmlr.org/papers/v3/blei03a.html>
- Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114(3), 494-509. <https://doi.org/10.1037/0033-2909.114.3.494>
- Çelik Kan, A. (2024). Eskiden Yeniye Geçiş Sürecinde Kent Kültürünün Değişimi (Yusufeli Örneği). *Asya Studies*, 8(30), 227-236. <https://doi.org/10.31455/asya.1533086>
- Çınaroğlu, M., Yılmaz, E., & Hızlı Sayar, G. (2025). Psychological impact of the 2023 Kahramanmaraş earthquakes on non-victims: a cross-sectional study. *BMC Public Health*, 25, 597. <https://doi.org/10.1186/s12889-025-21901-6>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293), 52-64. <https://doi.org/10.1080/01621459.1961.10482090>

- Durmuş, Y. (2022). Sentiment Analysis of Restaurant Reviews in Artvin Province by Rule-based Sentiment Analysis and Machine Learning. *GSI Journals Serie A: Advancements in Tourism Recreation and Sports Sciences*, 5(2), 134-144. <https://doi.org/10.53353/atrss.1090401>
- Eren, R., & Kuvvetli, M. (2017). Çevrimiçi Gastronomi İmajı: Türkiye Restoranlarının Tripadvisor Yorumlarının İçerik Analizi. *Tourism Academic Journal*, 4(2), 121-138. <https://dergipark.org.tr/en/pub/touraj/article/364267>
- Görmez, Y., Arslan, H., & Atak, B. (2024). Türkçe Metinlerde Duygu Analizi: Derin Öğrenme Yaklaşımlarının ve Ön İşlem Süreçlerinin Model Performansına Etkisi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 36(1), 509-520. <https://doi.org/10.35234/fumbd.1429040>
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv. <https://arxiv.org/abs/2203.05794>
- Haberler.com. (2022, October 18). Artvin haberleri: Baraj nedeniyle yeni yerleşim yerine taşınma süreci başlayan Yusufeli’nde esnaf ve vatandaşlar tepkili: “Manzara hiç iç açıcı değil”. <https://www.haberler.com/artvin-haberleri-ba-raj-nedeniyle-yeni-yerlesim-yerine-tasinma-15367804-haberi/>
- İncidelen, M., & Aydoğan, M. (2025). Sentiment analysis in Turkish using language models: A comparative study. *European Journal of Technique (EJT)*, 15(1), 68-74.
- Kahraman, S. (2023). A Comprehensive Health Economic Analysis of the 2023 Turkey Earthquake. *Disaster Medicine and Public Health Preparedness*, 17, e503. <https://doi.org/10.1017/dmp.2023.160>
- Karadaş, D. C., Aydın, T., & Özyer, G. (2025). 6 Şubat 2023 Depremi Üzerine X (Twitter) Verilerinin Konu Analizi: LDA ve BERTopic Yaklaşımlarının Karşılaştırması. *Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 18(1), 57-63. <https://doi.org/10.54525/bbmd.1633740>
- Kaya Kılıç, A., Bademli, K., & Kayakuş, M. (2025). Examination of Individual and Collective Trauma After the Kahramanmaraş Earthquake. *Clinical Social Work Journal*. <https://doi.org/10.1007/s10615-025-00993-y>
- Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal Detection of Change-points With a Linear Computational Cost. *Journal of the American Statistical Association*, 107(500), 1590-1598. <https://doi.org/10.1080/01621459.2012.737745>
- Koruyan, K. (2022). BERTopic Konu Modelleme Tekniği Kullanılarak Müşteri Şikayetlerinin Sınıflandırılması. *İzmir Sosyal Bilimler Dergisi*, 4(2), 66-79. <https://doi.org/10.47899/ijs.1167719>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), 50-60. <https://doi.org/10.1214/aoms/1177730491>
- Najafi, A., & Varol, O. (2024). TurkishBERTweet: Fast and reliable large language model for social media analysis. *Expert Systems with Applications*, 255, 124737. <https://doi.org/10.1016/j.eswa.2024.124737>
- Nasır, A., Alymany, G., Brimo, F., Tekin, S., & Drury, J. (2025). Understanding the socio-economic and wellbeing impacts of the 2023 earthquakes in Turkey on Syrian refugees. *International Journal of Disaster Risk Reduction*, 125, 105548. <https://doi.org/10.1016/j.ijdrr.2025.105548>
- Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135. <https://doi.org/10.1561/15000000011>
- Republic of Türkiye Ministry of Environment, Urbanization and Climate Change. (2022, October 10). Türkiye’nin en büyük iskân projesi Yusufeli’nde taşınmalar başladı. <https://csb.gov.tr/haberler/turkiye-nin-en-buyuk-isk-n-projesi-yusu-feli-nde-tasinmalar-basladi-300575>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) (pp. 3982-3992). Association for Computational Linguistics. <https://doi.org/10.18653/v1/D19-1410>
- Sigurdsson, V., Larsen, N. M., Gudmundsdottir, H. K., Alemu, M. H., Menon, R. G. V., & Fagerstrøm, A. (2021). Social media: Where customers air their

troubles—How to respond to them? *Journal of Innovation & Knowledge*, 6(4), 257–267.
<https://doi.org/10.1016/j.jik.2021.07.001>

Theocharopoulos, P. C., Tsoukala, A., Georgakopoulos, S. V., Tasoulis, S. K., & Plagianakos, V. P. (2023). Analysing sentiment change detection of Covid-19 tweets. *Neural Computing and Applications*, 35, 21433–21443.
<https://doi.org/10.1007/s00521-023-08662-2>

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining* (pp. 29–39). Practical Application Company.

Yürütücü, Ö. Y., & Demir, Ş. (2023). Ön Eğitimli Dil Modelleriyle Duygu Analizi. *İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 5(1), 46-53. <https://doi.org/10.47769/izufbed.1312032>