Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı 36, S. 305-315, Mayıs 2022
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue 36, pp. 305-315, May 2022
Copyright © 2022 EJOSAT
Research Article

Sentiment Analysis of Social Media Posts about Tourist Attractions: Black Sea Region Sample

Mertkan Sinoplu[1*], Eyüp Burak Ceyhan[2]

[1*] Bartın University, Graduate Education Institute, Department of Computer Technology and Information Systems, Bartın, Turkey, (ORCID: 0000-0003-4642-5090), mertkansinoplu@gmail.com

[2] Bartın University, Faculty of Engineering, Architecture and Design, Departmant of Computer Engineering, Bartın, Turkey, (ORCID: 0000-0002-5005-968X), eyupburak@gmail.com

(1st International Conference on Engineering and Applied Natural Sciences ICEANS 2022, May 10-13, 2022)

(**DOI:** 10.31590/ejosat.1107640)

ATIF/REFERENCE: Sinoplu, M. & Ceyhan, E. B. (2022). Sentiment Analysis of Social Media Posts about Tourist Attractions: Black Sea Region Sample. *European Journal of Science and Technology*, (36), 305-315.

Abstract

Social media has become one of the essential advertising channels with the increase of users and potential customers. With the data obtained from social media, analyses on advertisements and customers are conducted in many areas. The tourism industry is one of the areas where social media has a significant impact. In this study, we used Twitter and sentiment analysis for ranking tourist attractions in terms of positive tweet rates. We collected tweets shared on Twitter regarding certain tourist attractions in the Black Sea Region of Turkey using social media analytics. We conducted sentiment analysis using RapidMiner data processing software and Operator Toolbox extension. They were used for collecting and processing the data and sentiment analysis of Twitter tweets. The touristic places selected within the scope of the research were determined as Abant, Amasra, Ayder Plateau, Hattusa, Kartalkaya, Persembe Plateau, Sumela Monastery, Uzungol and Yedigoller. Total of 1985 tweets were used for this study. After the sentiment analysis, we identified that the tourist attraction with the highest percentage of positive tweets was Abant (%76.70). After Abant, Ayder Plateau (%76.47) and Uzungöl (%75.68) were identified that the tourist attraction with the highest percentage of positive tweets, respectively.

Keywords: Social Media, Sentiment Analysis, Classification, Black Sea Region, RapidMiner.

Turistik Yerler ile İlgili Sosyal Medya Paylaşımlarının Duygu Analizi: Karadeniz Bölgesi Örneği

Öz

Sosyal medya, kullanıcıların ve potansiyel müşterilerin artmasıyla birlikte vazgeçilmez reklam kanallarından biri haline geldi. Sosyal medyadan elde edilen veriler ile birçok alanda reklam ve müşteri analizleri yapılmaktadır. Turizm sektörü, sosyal medyanın önemli bir etkiye sahip olduğu alanlardan biridir. Bu çalışmada, turistik yerleri pozitif tweet oranlarına göre sıralamak için Twitter ve duygu analizini kullanılmıştır. Sosyal medya analizlerini kullanarak Türkiye'nin Karadeniz Bölgesi'ndeki bazı turistik yerler hakkında Twitter'da paylaşılan tweet'leri toplanmıştır. RapidMiner veri işleme yazılımı ve Operator Toolbox uzantısını kullanarak duyarlılık analizi yapılmıştır. Bu araçlar, Twitter tweetlerinin verilerinin toplanması, duygu analizini ve verilerin işlenmesi için kullanılmıştır. Araştırma kapsamında seçilen turistik yerler Abant, Amasra, Ayder Yaylası, Hattuşa, Kartalkaya, Perşembe Yaylası, Sümela Manastırı, Uzungöl ve Yedigöller olarak belirlenmiştir. Çalışma kapsamında toplam 1985 tweet kullanılmıştır. Duygu analizinden sonra, en yüksek olumlu tweet yüzdesine sahip turistik yerin Abant (%76.70) olduğunu belirlenmiştir. Abant'tan sonra en yüksek olumlu tweet yüzdesinin sırasıyla Ayder Yaylası (%76,47) ve Uzungöl (%75,68) olduğu tespit edilmiştir.

Anahtar Kelimeler: Sosyal Medya, Duygu Analizi, Sınıflandırma, Karadeniz Bölgesi, RapidMiner.

[*] Corresponding Author: mertkansinoplu@gmail.com

1. Introduction

The developments in technology have increased the use of social media (Akın and Gürsoy Şimşek, 2018), and it has become an essential platform for advertisements lately, creating new professional fields and teams. It is inevitable for companies to make use of social media since they wish to be liked and mentioned, and social media provides them with the opportunity to make their wish come true (Özdemir et al., 2014. Therefore, many institutions and enterprises develop advertisement and marketing strategies based on social media analytics.

People can use social media analytics in many areas. For example, in their study, Öztürkcan et al. (2017) focused on how social media analytics could be used for Twitter data. The study examined Twitter data to reveal the public reactions to the impactful events in Turkey and revealed that events could quickly spread when shared on social media.

In another study, Sabuncu and Atmis (2020) used Twitter data to evaluate the social media posts regarding Turkish Airlines as positive or negative. In this way, they facilitated the collection of feedback to the company. The study results revealed that the number of tweets increased upon an incident related to Turkish Airlines, and there was also an increase in the percentage of negative tweets.

Moreover, Akgül et al. (2016) conducted sentiment analysis on the data that they collected from Twitter. The study mentioned two different sentiment analysis models and developed a method for each model.

The purpose of this study was to evaluate certain tourist attractions in the Black Sea Region of Turkey using social media analytics and accordingly provide insight for companies in the tourism sector and potential visitors to these attractions. In addition, it was aimed to examine different classification algorithms to find the optimal algorithm for the 7-day dataset obtained from Twitter.

2. Material and Method

In the study, data collection and analysis were carried out using RapidMiner, a data mining software free for students. Both the company and users can develop the data mining tools in the software to be used as extensions (Verma et al., 2014).

The tourist attractions selected for the study included Abant, Ayder Plateau, Hattusa, Amasra, Kartalkaya, Perşembe Plateau, Sümela Monastery, Uzungöl and Yedigöller.

2.1. Data Collection

The example of the collected data can be seen in Figure 1. They contain tweets and user id numbers for each touristic location.

RapidMiner software was used to collect data in the study. The "Search Twitter" tool in the software enabled the collection of tweets containing the target words; however, one consideration in using this tool is that only the data can be retrieved for the last seven days. Therefore, the data in this study consisted of tweets shared between 15 January 2021 and 22 January 2021. In the sentiment analysis, a total of 1985 tweets were used including those about Abant (176), Amasra (112), Ayder Plateau (68), Hattusa (271), Kartalkaya (551), Perşembe Plateau (79), Sümela Monastery (194), Uzungöl (111) and Yedigöller (423).

1	Text	Id
2	Arkadaşlarımın toplanıp	1347602782002556933
3	Kartalkaya'daki otelciler	1347178369570500609
4	An itibari ile güncel tahı	1347130376339263488
5	Kartalkaya Kayak Merk	1347052140397813763
6	Kartalkaya'da kayağın t	1346734444716154881
7	Uludağ, Kartalkaya, Ka	1346733336027426816
8	Türkiye'nin önemli kış t	1346552034875617282
9	Bolu Kartalkaya Kayak	1346476721826979840
10	Kartalkaya'da kar yağışı	1346425450759933953

Figure 1. Example of the collected data.

RapidMiner operators used in the data collection process (Figure 2) are as follows:

- **Search Twitter**; allows collecting data from Twitter.
- **Select Attributes**; allows selecting the intended attribute from the table.
- **Nominal to Text**; changes the selected attributes to "text".
- **Write Excel**; allows writing the obtained results in an Excel spreadsheet file. This operator was later used to translate text.

Collected text are generally in Turkish. Operator Toolbox can only work on English text. Therefore, all text in the dataset was translated to English.

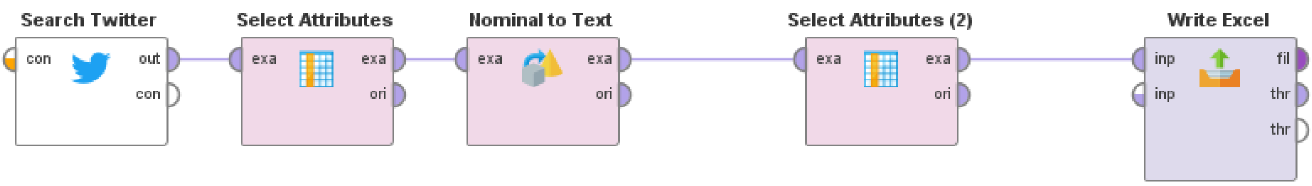

Figure 2. Data collection process.

RapidMiner operators used in the data obtaining process for algorithm tests (Figure 3) are as follows:

- **Read Excel**; allows reading data from a specified Excel file.
- **Extract Sentiment**; allows conducting sentiment analysis.

In this study, the tweets were analyzed using “sentiwordnet", one of the databases for sentiment analysis.

Upon the data acquisition (Figure 3), in order to test the algorithms, the sentiment analysis was performed with the “Extract Sentiment” operator in OperatorToolbox, which is a free extension included in RapidMiner software.

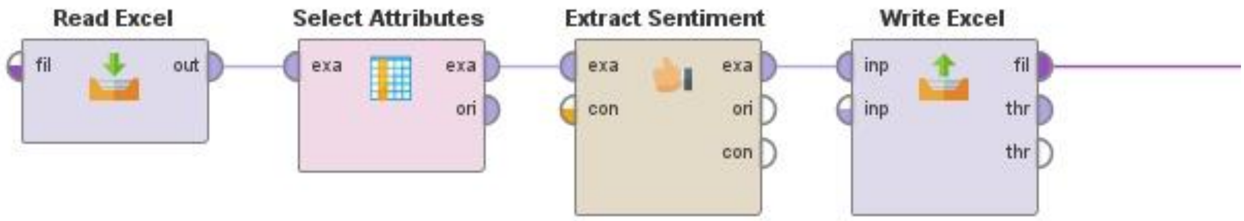

Figure 3. The data acquisition process for algorithm tests.

2.2. Classification Analysis

The algorithms to be used to classify sentiment analyses of tweets related to the tourist attractions in the study were tested using the RapidMiner software. Gradient Boosting, Deep Learning, Decision Tree, K-Nearest Neighbor, and Random Forest algorithms were used in the classification.

RapidMiner operators used in the classification process are presented in Figure 4. These operators are:

- **Generate Attribute**; constructs new attributes using the dataset. Here, sentiments were assigned as attribute names based on the attribute values.
- **Select Attributes**; allows selecting attributes from the dataset.
- **Set Role**; allows assigning a role to the specified attribute.
- **Nominal to Text**; changes the selected attributes to “text”.
- **Process Documents from Data**; allows creating a document from the data (Figure 5).

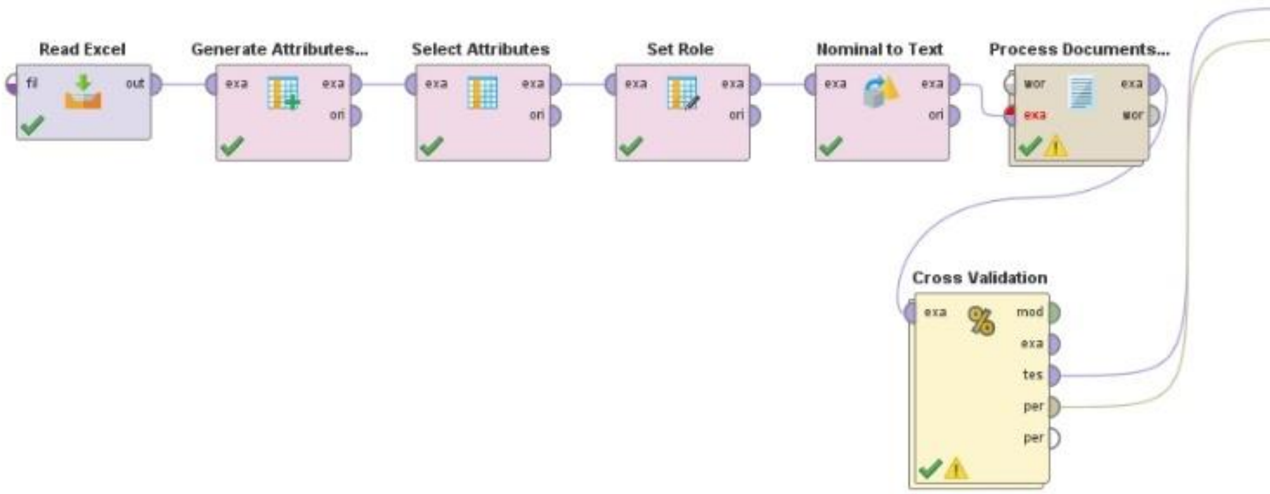

Figure 4. Classification process.

RapidMiner operators used in the documents from the data process (Figure 5) are as follows:

- **Tokenize**; splits a text into a sequence of words.
- **Transform Cases**; changes the characters in a text to uppercase or lowercase.
- **Filter Stopwords**; allows filtering the stop words in the text.

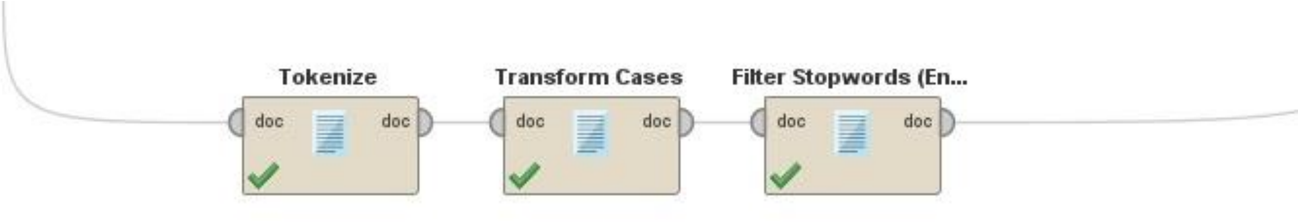

Figure 5. Process documents from the data process.

RapidMiner operators used in cross-validation process (Figure 7) are as follows:

- **Cross-Validation**; performs cross-validation as it is observed in Figure 6. The input data are partitioned into a specific number of equal-size subsets used as test and training data sets in this process. The algorithm specified in the process is used to obtain the optimal result. The partition number in the study was determined as 10.
- **Performance**: It shows the cross-validation results such as estimated negative/positive, real negative/positive, recall, precision, and accuracy. The recall is the ratio of correct guess size to the total class size. It is also known as the sensitivity of the model. On the other hand, precision is a positive predicted value. Recall and precision can be calculated with these formulas in Figure 6:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Figure 6. Calculation of precision and recall.

The algorithm to be used in the sentiment analysis of the tourist attractions was determined by considering the accurate classification rates obtained from the classification process. The performance rates of the classification algorithms are presented in Section 2.3.

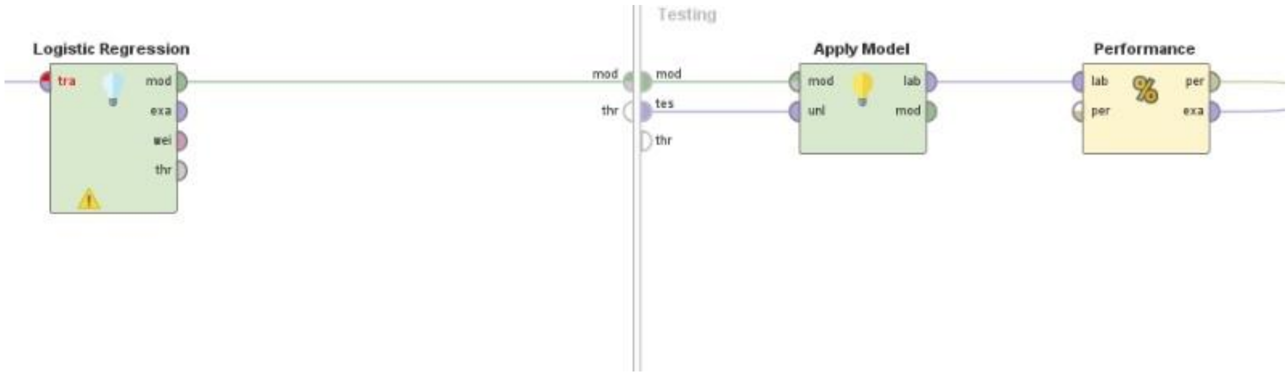

Figure 7. Cross-validation process.

2.3. Sentiment Analysis

The sentiment analysis of the data retrieved from Twitter and optimized was conducted using the OperatorToolbox, an extension of RapidMiner. Sentiment analysis can be performed with the "Extract Sentiment" tool included in this extension. It is also possible to use different databases of sentiment analysis in this operator.

RapidMiner operators used for the sentiment analysis process (Figure 8) are as follows:

- **Filter Examples**; filters the data based on a given condition and selects what to keep and remove.
- **Aggregate**; perform aggregate functions including group, sum, and count.
- **Merge**; merges datasets with different attributes.
- **Append**; merges datasets with the same attributes.
- **Rename**; allows renaming a specified attribute.
- **Normalize**; normalize the values of selected attributes using an intended method. The "proportion transformation" was used for normalization in the study. With this method, the data were divided into proportions such that their sum would be 1.
- **Generate Attributes**; constructs new attributes using the dataset. Here the data were calculated in percentage.

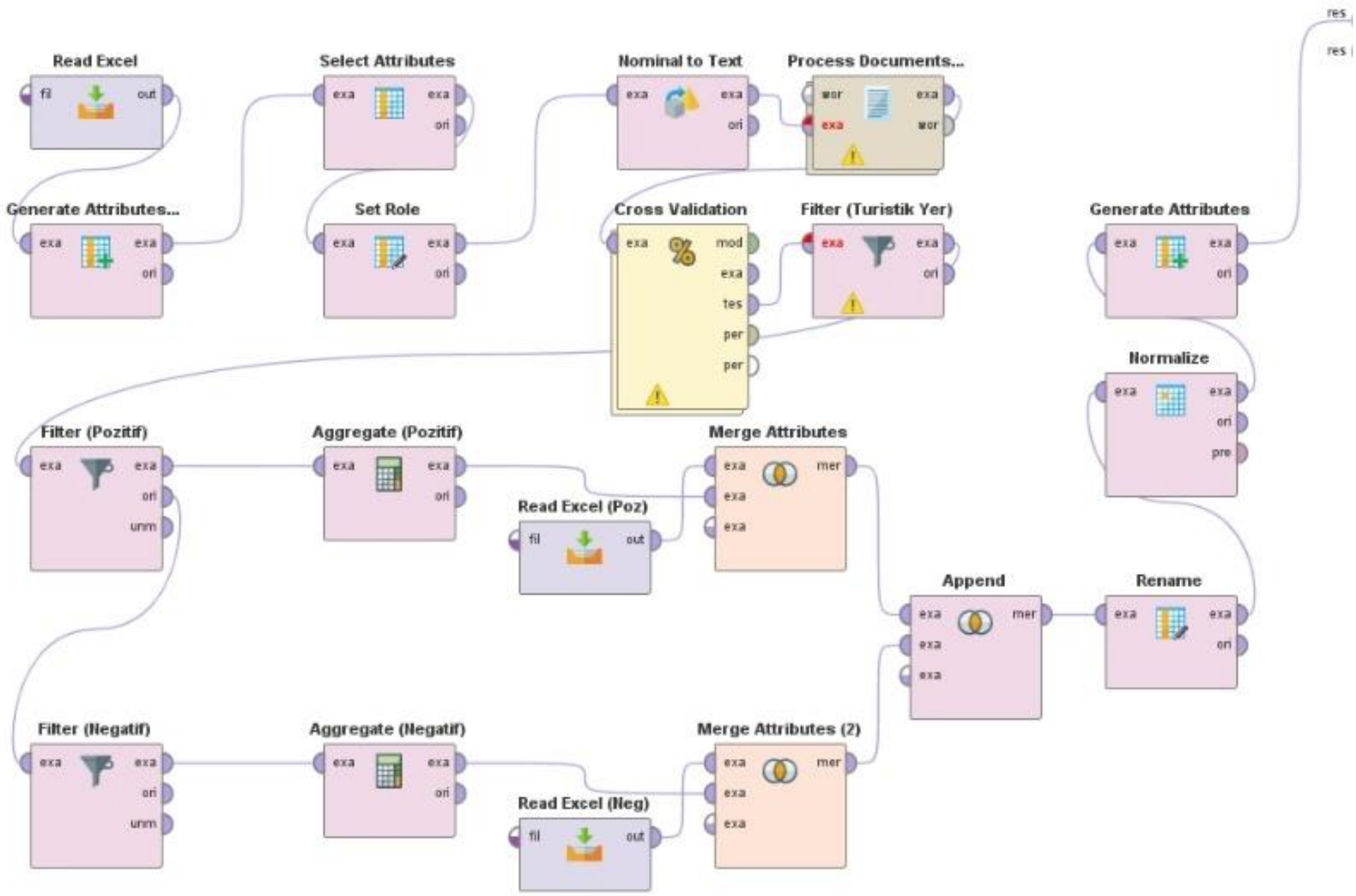

Figure 8. Sentiment analysis process.

In sentiment analysis, text data must first be preprocessed. For this reason, the data is filtered for sentiment analysis. Filtering means preparing features in text data because the features are not explicitly available in text data. After the filtering, tokenizing, stop words filtering, stemming, and lowercasing is done for better results. For tokenizing, the text is divided into words. After that, stop words such as "is, the, end," etc., are filtered out. After this process, all words can be stem for accuracy of the sentiment analysis. Stemming means reducing suffixes in words and making them close to the root word. This process can improve results, but stemming should be done carefully as the word's meaning can change.

The second part of sentiment analysis is attribute selection, which means giving features to words after the first step. There are many feature selection models, but the Bag of Words (BoW) model can be explained as the most preferred feature selection model for attribute selection.

The third step of sentiment analysis is training data. One part of the data is selected for training and obtaining trained classifiers in this process. Trained classifiers are needed for final tests. Classification can be done with different classification algorithms such as Naïve Bayes, Random Forest, Decision Trees, etc.

The last step of sentiment analysis is testing the dataset. After testing data with classification algorithms, the algorithm with the highest accuracy rate is selected. With this algorithm, all test data is classified and obtained.

The sentiment extraction process of this research was done with the help of the "Extract Sentiment" tool of the Operator Toolbox in RapidMiner. Then, classification algorithms are used to improve the accuracy and retrieve test results.

3. Results

The classification algorithms' performances were first compared to select the optimal algorithm for the sentiment analysis of the posts about the specified tourist attractions in the Black Sea Region of Turkey. Then the posts were analyzed using the sentiment analysis. The relevant results are presented in this section.

3.1. Performance of Classification Algorithms

The classification algorithms selected for the performance comparison in the study included Deep Learning, Gradient Boosting, Decision Tree, k-Nearest Neighbor, Logistic Regression, Naive Bayes, Random Forest, and Support Vector Machine. The performances of the relevant algorithms are presented in between Table 1 and Table 8, respectively. The values for accuracy (written in bold in the figures) were obtained using the following formula in Figure 9:

$$\text{Accuracy Rate} = \frac{\text{Accurately Estimated}}{\text{Accurate and Inaccurate Estimations}} \times 100$$

Figure 9. Calculation of accuracy rate.

The test results of the Deep Learning algorithm are given in Table 1. When we consider the precisions of the Deep Learning algorithm, it was seen that the negative precision was 75.14%, and the positive precision was 90.07%. When we consider the recalls of the algorithm, it was seen that the negative recall was 86.19%, and the positive recall was 81.46%. As a result of the test, it was seen that the accuracy rate of the Deep Learning algorithm was 83.32%.

Table 1. Performance of Deep Learning algorithm.

Deep Learning **Accuracy: 83.32%**	Real Negative	Real Positive	Precision
Estimated Negative	674	223	75.14%
Estimated Positive	108	980	90.07%
Recall	86.19%	81.46%	

As a result of classification algorithm tests, the Gradient Boosting algorithm values are given in Table 2. When considering precisions, it was seen that negative precision was 90.64%, and positive precision was 81.61%. When we consider the recalls of the algorithm, it was seen that the negative recall was 66.88%, and the positive recall was 95.51%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 84.23%.

Table 2. Performance of Gradient Boosting algorithm.

Gradient Boosting **Accuracy: 84.23%**	Real Negative	Real Positive	Precision
Estimated Negative	523	54	90.64%
Estimated Positive	259	1149	81.61%
Recall	66.88%	95.51%	

The information obtained from the classification algorithm test about the Decision Tree algorithm is given in Table 3. As a result of the classification algorithm test, the negative precision of the Decision Tree algorithm was 97.46%, and the positive precision was 76.80%. When recalls of the algorithm were considered, it was seen that the negative recall was 53.96%, and the positive recall was 99.09%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 81.31%.

Table 3. Performance of Decision Tree algorithm.

Decision Tree **Accuracy: 81.31%**	Real Negative	Real Positive	Precision
Estimated Negative	422	11	97.46%
Estimated Positive	360	1192	76.80%
Recall	53.96%	99.09%	

The classification algorithm test results about the k-Nearest Neighbor Algorithm are shown in Table 4. When we look at the test results, it was seen that the negative precision of the k-Nearest Neighbor algorithm was 84.92%, and the positive precision was 85.57%. When we consider the recalls of the algorithm, it was seen that the negative recall was 76.34%, and the positive recall was 91.19%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 85.34%.

Table 4. Performance of k-Nearest Neighbor algorithm.

k-Nearest Neighbor **Accuracy: 85.34%**	Real Negative	Real Positive	Precision
Estimated Negative	597	106	84.92%
Estimated Positive	185	1097	85.57%
Recall	76.34%	91.19%	

The test result of the Logistic Regression algorithm is shown in Table 5. As a result, the negative precision of the Logistic Regression was found 71.38%, and the positive precision was 85.20%. In terms of recall values, the negative recall of the algorithm was 78.77%, and the positive recall was 79.47%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 79.19%.

Table 5. Performance of Logistic Regression algorithm.

Logistic Regression **Accuracy: 79.19%**	Real Negative	Real Positive	Precision
Estimated Negative	616	247	71.38%
Estimated Positive	166	956	85.20%
Recall	78.77%	79.47%	

The test results of the Naïve Bayes algorithm are given in Table 6. Looking at the precision of the Naïve Bayes algorithm as a result of the test, it was seen that the negative precision was 78.75%, and the positive precision was 88.25%. When we consider the recalls of the algorithm, it was seen that the negative recall was 82.48%, and the positive recall was 85.54%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 84.33%.

The test results of the Naïve Bayes algorithm are given in Table 6. Looking at the precision of the Naïve Bayes algorithm as a result of the test, it was seen that the negative precision was 78.75%, and the positive precision was 88.25%. When we consider the recalls of the algorithm, it was seen that the negative recall was 82.48%, and the positive recall was 85.54%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 84.33%.

Table 6. Performance of Naïve Bayes algorithm.

Naive Bayes **Accuracy: 84.33%**	Real Negative	Real Positive	Precision
Estimated Negative	645	174	78.75%
Estimated Positive	137	1029	88.25%
Recall	82.48%	85.54%	

The classification algorithm test results of the Random Forest algorithm are given in Table 7. As a result of the test, when the precision of the Random Forest algorithm is examined, it is seen that the negative precision is 100%, and the positive precision is 72.08%. When we consider the recalls of the algorithm, it was seen that the negative recall was 40.41%, and the positive recall was 100%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 76.53%.

Table 7. Performance of Random Forest algorithm.

Random Forest **Accuracy: 76.53%**	Real Negative	Real Positive	Precision
Estimated Negative	316	0	100%
Estimated Positive	466	1203	72.08%
Recall	40.41%	100%	

After the classification algorithm tests, test values of the Support Vector Machine algorithm are given in Table 8. When we look at the precision of the Support Vector Machine algorithm as a result of the test, it was seen that the negative precision was 94.36%, and the positive precision was 85.63%. When we consider the recalls of the algorithm, it was seen that the negative recall was 74.94%, and the positive recall was 97.09%. In addition, as a result of the test, it was seen that the accuracy rate of the algorithm was 88.36%.

Table 8. Performance of Support Vector Machine algorithm.

Support Vector Machine Accuracy: 88.36%	Real Negative	Real Positive	Precision
Estimated Negative	586	35	94.36%
Estimated Positive	196	1168	85.63%
Recall	74.94%	97.09%	

The comparison of the accuracy of the classification algorithms examined in the study is presented in Figure 10. The relevant comparison revealed that Support Vector Machine had the highest accuracy (88.36%). It was followed by k-Nearest Neighbor (85.34%), Naive Bayes (84.33%), Gradient Boosting (84.23%), Deep Learning (83.32%), Decision Tree (81.31%), Logistic Regression (79.19%), and Random Forest (76.53%) respectively.

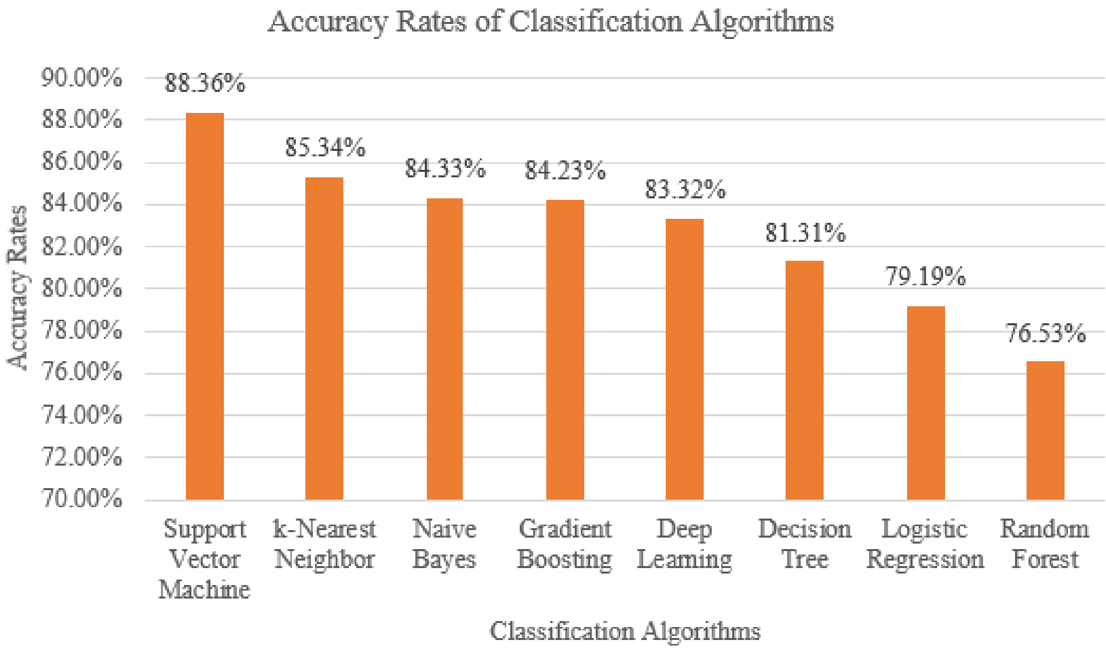

Figure 10. Accuracy of classification algorithms.

Consequently, in line with the results obtained, the Support Vector Machine was determined to be the optimal algorithm for the sentiment analysis of the relevant tourist attractions.

3.1. Performance of Classification Algorithms

This section conducted sentiment analysis on the Twitter posts about the specified tourist attractions. Support Vector Machine, the algorithm with the highest accuracy, was used in the analysis. The numbers and rates of the tweets with positive or negative content obtained as a result of the sentiment analysis are presented in Figure 11 and Figure 12, respectively.

The numbers of tweets with positive or negative content obtained from the sentiment analysis of the posts regarding the tourist attractions determined are presented in Figure 11.

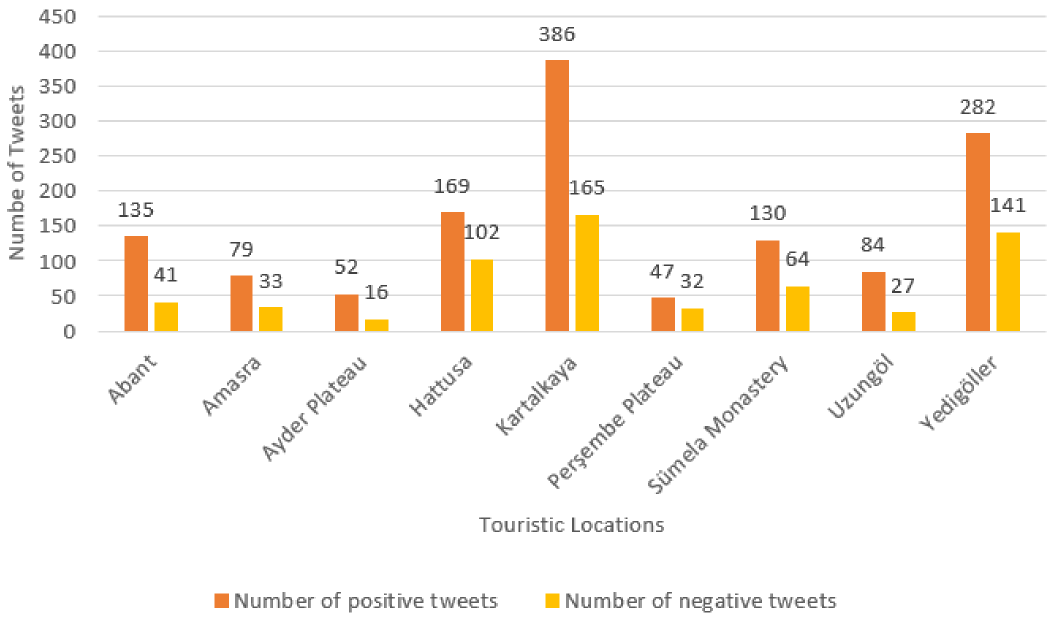

Figure 11. The numbers of positive and negative tweets about the tourist attractions.

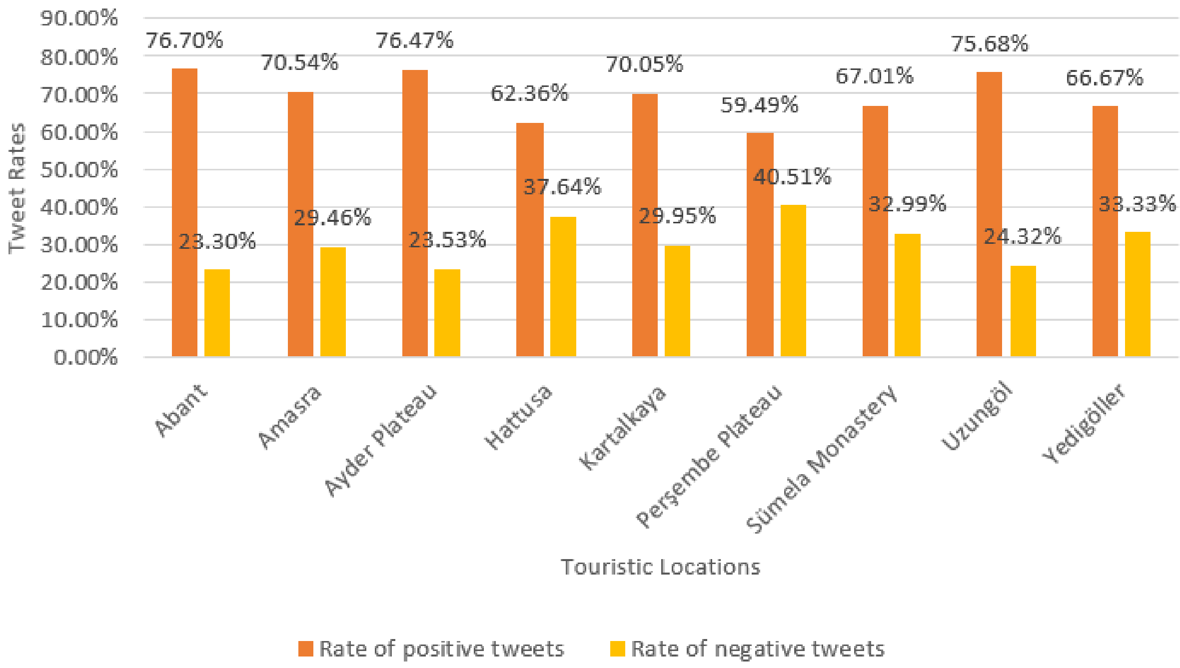

Figure 12. The rates of positive and negative tweets regarding the tourist attractions.

The rates of the tweets with positive or negative content acquired at the end of the sentiment analysis of the posts about the tourist attractions specified are presented in Figure 12.

The ranking of the tourist attractions based on the rate of the positive tweets is presented in Figure 13. When the figure is examined, it can be observed that Abant ranks first with the highest rate of tweets with positive content (76.70%). It is followed by Ayder Plateau (76.47%), Uzungöl (75.68%), Amasra (70.54%), Kartalkaya (70.05%), Sümela Monastery (67.01%), Yedigöller (66.67%), Hattusa (62.36%) and Perşembe Plateau (59.49%), respectively.

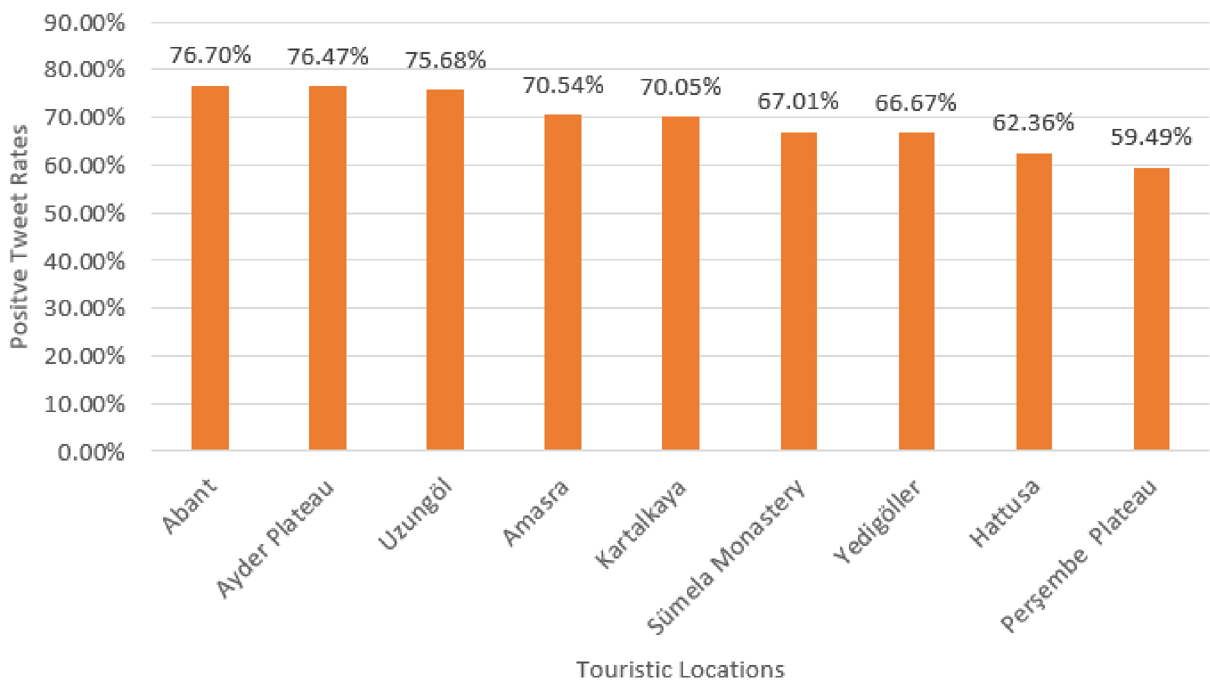

Figure 13. Ranking of the tourist attractions based on the rate of positive tweets.

4. Conclusion and Recommendations

With the development of technology, social media has become one of the biggest marketing channels for the tourism sector (Aktan, 2018). It is a fact that social media influences tourists' preferences to a great extent (Doğan et al.,2018). The present study collected tourism-related data from social media platforms, and sentiment analysis was conducted on the tweets about the specified tourist attractions. The analysis has revealed that, among all the tourist attractions, Abant has the highest rate of posts with positive content (76.70%), whereas Perşembe Plateau has the lowest rate (59.49%). In addition, the analysis has paved the way for an evaluation of the relevant destinations that need more care and improvement.

In classifying the dataset used in the study, Support Vector Machine has been determined as the highest accuracy algorithm (88.36%). Different classification algorithms may also be tested for future research to identify a better algorithm with higher accuracy.

Moreover, regarding the ranking of the tourist attractions, it can be considered unlikely to provide a precise ranking since the tourist attractions in the study do not belong in the same category, and the peak seasons are different for these destinations.

Furthermore, a website or application can be developed to present the data from the processes carried out in this study so that people who intend to visit the relevant tourist attractions can evaluate prior to their travel.

During the review of posts about the study's tourist attractions, different tweets were collected for each attraction. For future research, more data can be collected to achieve the same number of tweets for each tourist attraction to conduct a more effective study.

For future research, it is also recommended that longer-term data should be obtained, and tourist attractions should be ranked into categories. Besides, considering the peak seasons of the attractions, only the tweets posted during these seasons can be analyzed to obtain more accurate results.

In addition, since the data in the study were collected using the RapidMiner software, it should be considered that the limitation of the software allows retrieving data from Twitter for the last seven days from the date of retrieval. Therefore, it should be taken into account that the news and events during the week of the data collection could affect the social media posts positively or negatively, and it could consequently impact the analysis.

Finally, due to the 7-day data obtain the limitation of the RapidMiner software, it can be considered that the study results are based on a small-scale analysis. Therefore, for future research, it is suggested that the same subject be studied by retrieving a higher number of posts; in this way, potential tourists can benefit from this new research while making their travel decisions.

References

Akgül, E. S., Ertano, C., Diri, B. (2016). Sentiment Analysis with Twitter. *Pamukkale University Journal of Engineering Sciences*, 22(2), 106–110. https://doi.org/10.5505/pajes.2015.37268

Akın, B., Gürsoy Şimşek, U. T. (2018). Sosyal Medya Analitiği ile Değer Yaratma: Duygu Analizi ile Geleceğe Yönelim. *Mehmet Akif Ersoy Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi*, 797–811. https://doi.org/10.30798/makuiibf.435804

Aktan, E. (2018). Sosyal Medyanın turizm Pazarlamasındaki Rolünün Değerlendirilmesi (Assessing the Role of Social Media in Tourism Marketing*). Journal of Tourism and Gastronomy Studies*, 6(3), 228–248. https://doi.org/10.21325/jotags.2018.280

Doğan, M., Pekiner, A. B., Karaca, E. (2018). Sosyal Medyanın Turizm ve Turist tercihlerine Etkisi: Kars-Doğu Ekspresi Örneği. *Seyahat ve Otel İşletmeciliği Dergisi*, 669–683. https://doi.org/10.24010/soid.443504

Özdemir, S. S., Özdemir, M., Polat, E., Aksoy, R. (2015). Sosyal Medya Kavramı ve Sosyal Ağ Sitelerinde Yer Alan Online Reklam Uygulamalarının İncelenmesi. *EJOVOC : Electronic Journal of Vocational Colleges*, 4(4). https://doi.org/10.17339/ejovoc.96993

Özturkcan, S., Kasap, N., Çevik, M., Zaman, T. (2017). An analysis of the Gezi Park Social Movement Tweets. *Aslib Journal of Information Management*, 69(4), 426–440. https://doi.org/10.1108/ajim-03-2017-0064

Sabuncu, İ. & Atmis, M. (2020). Social Media Analytics for Brand Image Tracking: A Case Study Application for Turkish Airlınes. *Yönetim Bilişim Sistemleri Dergisi*, 6 (1), 26-41.

Verma, T., Renu, R., Gaur, D. (2014). Tokenization and Filtering Process in Rapidminer. *International Journal of Applied Information Systems*, 7(2), 16–18. https://doi.org/10.5120/ijais14-451139