



POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: <http://dergipark.org.tr/politeknik>



LSTM network based sentiment analysis for customer reviews

Müşteri görüşleri için LSTM ağı tabanlı duygu analizi

Author(s) (Yazar(lar)): Burhan BİLEN¹, Fahrettin HORASAN²

ORCID¹: 0000-0002-3106-7369

ORCID²: 0000-0003-4554-9083

Bu makaleye şu şekilde atıfta bulunabilirsiniz (To cite to this article): Bilen B., Horasan F., "LSTM network based sentiment analysis for customer reviews", *Politeknik Dergisi*, 25(3): 959-966, (2022).

Erişim linki (To link to this article): <http://dergipark.org.tr/politeknik/archive>

DOI: 10.2339/politeknik.844019

LSTM Network Based Sentiment Analysis for Customer Reviews

Highlights

- ❖ Provision of a new dataset to this field to work on it.
- ❖ Showing the effects of noise normalization and preprocessing on the classification accuracy.
- ❖ Representation of texts in vector form in various ways to be able to work on them.
- ❖ Comparison of the LSTM Network with several machine learning algorithms.

Graphical Abstract

Sentiment analysis was performed on the Turkish Customer Reviews dataset and Stanford Large IMDB Movie Reviews dataset with the LSTM neural network after pre-processing on the data obtained.

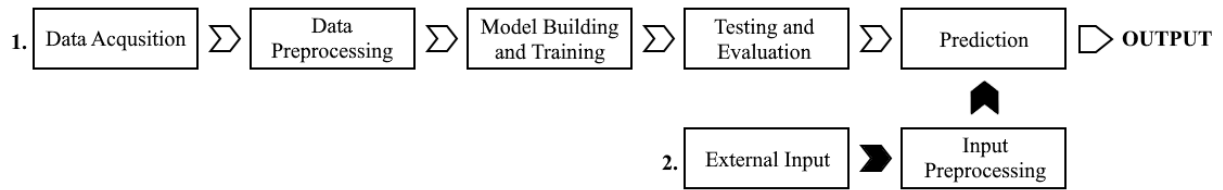


Figure. Flowchart of the Model

Aim

The main aim is to provide a different technique than mostly used machine learning algorithms. Providing a new dataset and new guide paper for Sentiment Analysis projects are also following aims.

Design & Methodology

An algorithm was used to gather raw data for the work, and also, a well-known dataset was used as extra. Preprocessing techniques were implemented on the data. Sentences were transformed into vector sequences. Hyperparameters were tuned. The Neural network was established using Embedding and LSTM Layers. The network was trained. Predictions were made on the test data. Model score was obtained. Model performance was checked with the use of some methods. New predictions/classifications were done on new external inputs.

Originality

This work is not the same as any other work in this field. Many similar approaches exist, but this study received a fairly high accuracy score, with the Long Short Term Memory Network, on the Turkish Customer Reviews dataset collected by the researchers.

Findings

The two most important findings are about how to get a higher accuracy score. The observation made in this study showed that as the amount of data increased, the accuracy score also increased. Also, another important increase in accuracy score was by text cleaning and noise normalization process.

Conclusion

The benefits of sentiment analysis were highlighted. A successful binary classification system was made. The LSTM Network provided a higher accuracy score compared to machine learning algorithms used in this study. Also, the highest accuracy score of the main paper of the IMDB dataset was passed in this study.

Declaration of Ethical Standards

The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

LSTM Network Based Sentiment Analysis for Customer Reviews

Araştırma Makalesi / Research Article

Burhan BİLEN, Fahrettin HORASAN*

Kirikkale University, Engineering and Architecture Faculty, Computer Engineering, 71450 Kirikkale, TURKEY

(Geliş/Received: 21.12.2020 ; Kabul/Accepted: 03.03.2021 ; Erken Görünüm/Early View : 17.03.2021)

ABSTRACT

Continuously increasing data bring new problems and problems usually reveal new research areas. One of the new areas is Sentiment Analysis. This field has some difficulties. The fact that people have complex sentiments is the main cause of the difficulty, but this has not prevented the progress of the studies in this field. Sentiment analysis is generally used to obtain information about persons by collecting their texts or expressions. Sentiment analysis can sometimes bring serious benefits. In this study, with singular tag-plural class approach, a binary classification was performed. An LSTM network and several machine learning models were tested. The dataset collected in Turkish, and Stanford Large Movie Reviews datasets were used in this study. Due to the noise in the dataset, the Zemberek NLP Library for Turkic Languages and Regular Expression techniques were used to normalize and clean texts, later, the data were transformed into vector sequences. The preprocessing process made 2% increase to the model performance on the Turkish Customer Reviews dataset. The model was established using an LSTM network. Our model showed better performance than Machine Learning techniques and achieved an accuracy of 90.59% on the Turkish dataset and an accuracy of 89.02% on the IMDB dataset.

Keywords: Deep learning, machine learning, sentiment analysis, LSTM, sequence embedding.

Müşteri Görüşleri için LSTM Ağı Tabanlı Duygu Analizi

ÖZET

Sürekli artan veriler yeni problemlerin ortaya çıkmasına neden olurlar. Problemler ise genellikle yeni bir araştırma alanı ortaya çıkarır. Ortaya çıkan yeni alanlardan biri ise Duygu Analizi konusudur. Bu alan bazı zorluklara sahiptir. İnsanların karmaşık duygulara sahip olması bu zorluğun ana sebebidir fakat bu durum çalışmaların ilerlemesine engel olmamıştır. Duygu analizi genel olarak kişilere ait herhangi bir yazı veya ifadenin toplanmasıyla yapılan çalışmalar sonucunda kişiye ait durum ve hisleri elde etmeye yarar. Bu tür analizler bazen ciddi faydalar sağlayabilir. Bu çalışmada, tekil etiket-çoğul sınıf yaklaşımıyla müşteri yorumları olumlu veya olumsuz olacak şekilde, ikili sınıflandırmada sınıflandırma işlemi yapılmıştır. Türkçe olarak topladığımız veri seti ve Stanford Large Movie Reviews adlı veri seti bu çalışmada kullanılmıştır. Veri setindeki gürültüden dolayı, sinir ağı eğitimi öncesinde Türk dilleri için yazılan Zemberek Doğal Dil İşleme kütüphanesi ve Regular Expression yöntemleriyle yazı normalizasyon ve temizleme işlemleri yapıp, veriler vektör dizileri haline getirilmiştir. Ön işleme sürecinin, Türkçe veri seti üzerinde, model performansına %2 düzeyinde katkısı olmuştur. LSTM ağı, kullanılan makine öğrenmesi modelleriyle kıyaslanmıştır. Kullanılan model, yapılan değerlendirmeler sonucunda genel Makine Öğrenmesi tekniklerinden daha iyi bir performans göstermiş, Türkçe Müşteri Görüşleri veri seti üzerinde %90.59 ve IMDB veri seti üzerinde ise %89.02 düzeyinde bir başarı skoru sağlamıştır.

Anahtar Kelimeler: Derin öğrenme, makine öğrenmesi, duygu analizi, LSTM, dizi vektörleştirme.

1. INTRODUCTION

The data, which have reached high dimensions and continue to increase in every field, reveal new problems and new research areas. Many research fields such as clustering, automatic text summarization, data classification, sentiment analysis can be given as examples of the areas created by this increase in data.

Today, the large amount of data available on the Internet have led researchers to work on automatic text categorization in order to organize texts well. Most of these studies have progressed by classifying documents according to their subjects. However, there is a huge

increase in the expression of sentiments in contents posted in many online communities [1].

Considering the classification process according to sentiments, this can be seen as a successful categorizing operation. Our study, which is in the same concept, is on the subject of analysis of sentiments and will be on opinion classification and will present some evaluation results at the end of the study.

With the increase in the use of the Internet, the concept of sentiment analysis has also gained importance. Opinions or expressions of individuals have become available to be obtained from any posts shared [2]. People who have left their marks in many fields today leave their sentiments there as well. Nowadays, by browsing many

*Sorumlu Yazar (Corresponding Author)
e-posta : fhorasan@kku.edu.tr

different web pages, we can easily see persons' reviews, what they like and dislike, what they want, their complaints, their health and economic status, and many other criteria. This facility has formed the subject of Sentiment Analysis, which is also known as Opinion Mining. Sentiment analysis can be divided into different types of problems, such as multiple class-single tags, multiple class-multiple tags, etc. We can understand how optimistic or pessimistic any person is based on their comment on a social media platform. Everyone can guess that there is no need to take any action for this operation. When we would read a comment, it would be easy for us to understand what kind of person he or she is. The main problem here is to obtain the positiveness or polarity of comments on a post with thousands of comments from a person because reading thousands of comments for just one post takes hours, which is too difficult to handle if the number of posts increases.

As the use of the Internet increases, it becomes more difficult to operate manually. Every day, new websites, electronic stores, or various platforms are emerging. In this way, the control of the big data caused by this increase has become uncontrollable even by special evaluation teams. In such a situation, it has become necessary to manage all kinds of analysis and similar processes with scientific techniques and ready-made models [3].

Deep Learning has become increasingly popular in these kinds of analysis processes. Deep Learning is a machine learning subtopic inspired by a real neural network structure [4]. It consists of mathematical artificial networks and imitates the structure of the human brain. This concept has now made many things possible. In this way, it has reached a level of performance that generally exceeds machine learning techniques.

In this study, Long-Short Term Memory Network, which is in the concept of recurrent neural networks, was used and compared with several machine learning models. LSTM was chosen because it is context-aware, that is, it can remember long sentences, and another reason is to avoid problems of general recurrent networks.

Recursive networks consist of four model types [5]:

- One-to-One
- One-to-Many
- Many-to-One
- Many-to-Many ($T_x = T_y$ or $T_x \neq T_y$)

The RNN model used in this study is the Many-to-One model due to the problem type, and it is shown in Figure 1. At this point, a state of sentiment is obtained for a sequence of sentences.

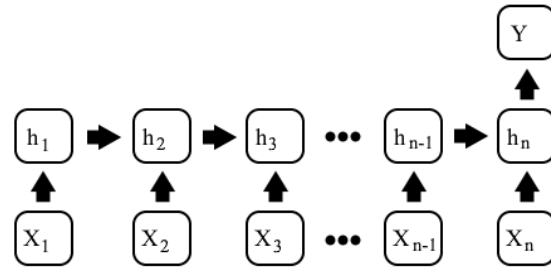


Figure 1. Many-to-One model type

This study aims to achieve more successful results by performing sequence processing and model training with recurrent neural networks, unlike frequently used machine learning techniques. The network structure used in this study was chosen considering the lower accuracy of machine learning techniques compared to the concept used in this study.

Some of the contributions of the proposed study are as follows:

- A new Turkish dataset has been added to the literature to be used in sentiment analysis researches [6].
- With the LSTM network structure, a higher classification accuracy has been achieved compared to machine learning techniques used.
- The effect of noise normalization and pre-processing on the classification accuracy on the data was emphasized.
- The accuracy score obtained in the main paper of the IMDB dataset was passed with LSTM.

In this study, two datasets were used; Turkish Customer Reviews [6] and Stanford Large Movie Reviews (IMDB) [25]. Turkish dataset contains approximately 8,500 customer reviews collected from various electronic stores, and the IMDB dataset contains 50,000 movie reviews (25000 positives and 25000 negatives). Reviews are divided into two as positive and negative. While collecting the customer reviews, sentiment tags were determined according to the number of stars that users gave to the products. In our own dataset, one and two stars represent negative reviews and four and five stars represent positive reviews.

At this point, the sentiment analysis topic encounters one of the biggest problems of its concept. For example, supposing a user saying that he/she likes the product but complains about the shipping company and therefore give two stars to the product. The question asked here is "Why the user is satisfied with the product and also gives it low stars?". Another example, let the person say that the product is very qualified, but it has shortcomings. In such cases, the machine learning or deep learning models will likely be making the wrong choice. This is still a problem for sentiment analysis studies.

More information about methods, techniques, and details can be obtained in chapter three.

2. RELATED WORKS

The first studies for sentiment analysis started in the 2000s and are still ongoing [7]. The studies conducted have created many experimental results and resources about this concept.

There are many Turkish studies in this field, but since Turkish is an agglutinative language and the use of letter points does not change the general meaning, some people do not write the words in accordance with their real spelling, and they even shorten many words sometimes. In this case, unless the data obtained goes through a good preprocessing stage, they cause poor performance scores from the model used.

Today, a lot of benefits have been gained from sentiment analysis. For example, in the United States elections, the presidential candidate's team reached millions and earned millions of donations as they successfully controlled and analyzed various social media platforms and people's data [2]. If we look at more concrete examples, the use of sentiment analysis concept under Social Media Marketing (SMM) can benefit many companies significantly. If a customer satisfaction rate is obtained from pre-sold products, it becomes easier than normal methods to understand whether to invest in the existing product. Another example is sentiment-based classification. For example, for survey studies, more efficient analysis results can be obtained by collecting sentences containing the same sentiment at certain points.

There is a lot of work done in this area. For example, in a polarity-based study, researchers took a new approach. The 28-features classifier, one of the techniques used in method comparisons to calculate contextual polarity, achieved the highest success score of 75.9% [8]. In a study conducted using the Stanford Sentiment Tree Bank (SSTb) dataset, an accuracy of 85.7% was achieved by using sentiment tags represented in binary format, both positive and negative. In addition, an 86.4% accuracy score was achieved on the Stanford Twitter Sentiment (STS) dataset [9]. Sebastian et al. conducted sentiment analysis using multiple techniques, in many languages. When they studied restaurant comments, they achieved an accuracy score of 73.6% when using the Turkish dataset and 81.4% when using the English dataset [10]. In another study, the researchers preferred the Naive Bayes and K-NN models. They achieved 82.43% of accuracy with Naive Bayes and 69.81% with K-NN. What is striking in that instance study is that the success rate increases with the increase in the size of the training dataset [11]. Palak et al. applied machine learning techniques on the IMDB dataset containing 2000 user comments in a similar study. When they used the Naive Bayes algorithm, they got the highest score, 81.4%. In addition, they achieved an accuracy score of 55.3% with the K-NN algorithm and 78.65% with the Random Forest algorithm [12]. In another analysis, researchers have made their studies by applying three different models. They achieved accuracy scores of 66.95% with the Naive Bayes algorithm, 81.77% with the SVM algorithm, and

74.75% with the Decision Tree algorithm. Later, the final score they obtained with the Cross-Validation technique was 81.75% with the SVM algorithm [13]. Dehkharghani et al. used different approaches such as Neural Networks, Logistic Regression, etc. They worked on polarity-based sentiment analysis with different combinations of feature subsets. The highest accuracy they have taken was 91.11% by a combination of classifiers they have chosen and it was taken by using all feature subsets [14]. Saglam et al. have worked on a project with a different approach to sentiment analysis. They have first translated the Turkish sentences into English, then done analysis using English sentiment lexicons. This work is an upgraded version of their previous work. In this way, they have increased the accuracy of 72.2% of previous work to 80.4%. The model they have built, SWNetTR++, achieved an accuracy of 75.0% in the polarity-based test and an accuracy of 80.4% in the tone-based testing [15]. Kamisli et al. have made an application of sentiment analysis for a university. They created a corpus by collecting tweets that are related to their university and performed an analysis on the corpus. They achieved an accuracy of 64.44% on negative tweets, 48.2% on neutral tweets, and 61.11% on the positive tweets about the university. Totally, they got an accuracy of 56.0% on the validation set.

If we check works based on Recurrent Neural Networks, in a study, researchers used the LSTM network model defined as Attention-based LSTM with Aspect Embedding (ATAE-LSTM) on the SemEval 2014 dataset, and when they classified the comments about restaurants as positive/negative/neutral, the highest score was obtained in the aspect-level polarity computation and it has reached the accuracy score of 84.0%. In the same work, they achieved 90.9% success with the aspect-term polarity computation in the positive/negative classification process [17]. In the work of Arras et al., when researchers used bi-LSTM model, they had a 46.3% of accuracy score when they have used five-class sentiment prediction (very negative, negative, neutral, positive, very positive) on 2210 sentences from the STSb dataset. Also, they obtained 82.9% of accuracy score when they used binary sentiment classification (positive, negative) [18]. As another instance, in the study of Pant et al., researchers used an RNN model on user tweets to predict Bitcoin prices, and they obtained 81.39% of accuracy, 82.90% of precision, 84.86% of recall and 83.86% of f-measure scores with a binary classification model. Also, they had a 77.62% of Bitcoin price prediction accuracy [19].

In our study, a similar approach to some of the examples mentioned was applied. Compared to the studies mentioned above, the proposed model has achieved a higher accuracy score in general, especially better than Machine Learning techniques. Also, higher or similar accuracy scores were obtained compared to similar approaches.

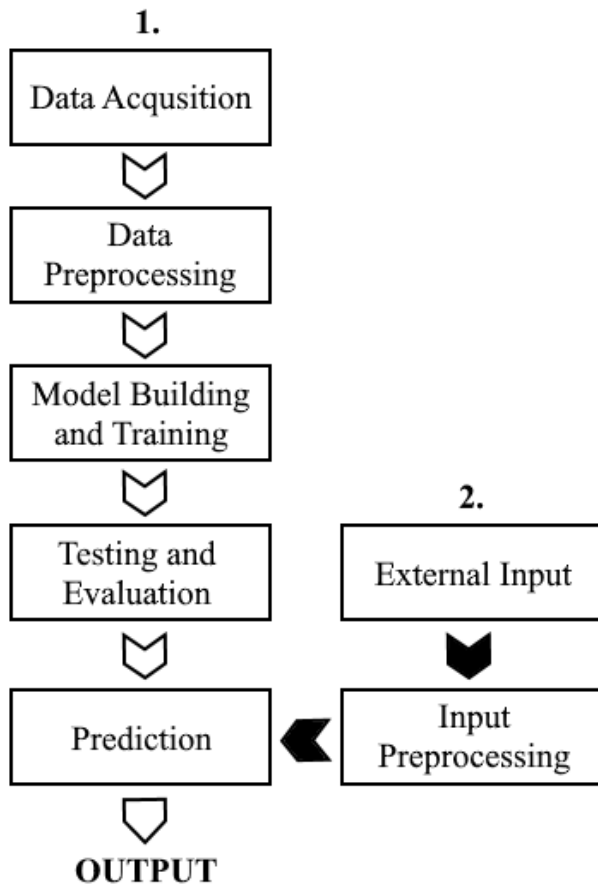


Figure 2. Steps of the model

3. METHODS AND APPLICATIONS

In order to understand the analysis process, a block diagram was used. In the scheme outlined in Figure 2, path number "1" represents pre-installation, and path number "2" represents post-installation. By saving the obtained model after completing the first path, the use of the second path begins.

3.1. Data Acquisition

The data acquisition process is one of the most important stages. It is important for the dataset used to have a balanced class distribution as they are collected because this process can cause major shortcomings in classification problems. The data used in this study were obtained from various electronic stores and IMDB reviews were obtained from the Stanford University website. The data distributions are shown in Figure 3 and Figure 4.

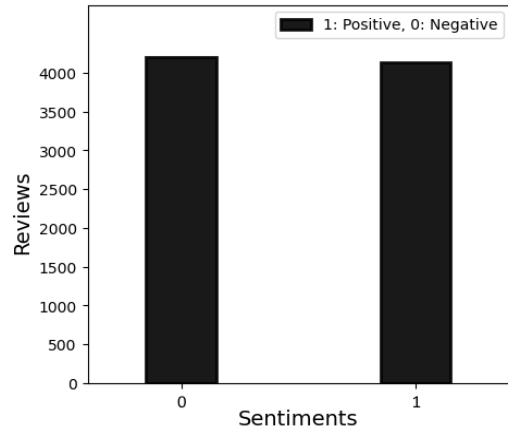


Figure 3. Turkish dataset class distribution

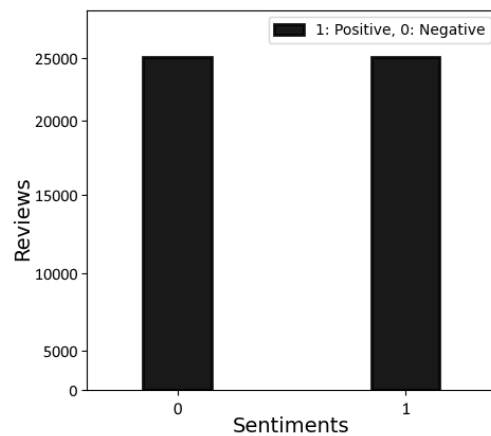


Figure 4. IMDB dataset class distribution

3.2. Data Preprocessing

The data preprocessing phase is one of the most crucial points of the work. In this way, working with preprocessed data can provide very high-performance gains compared to processing data in its raw form. Usually, the data in the Internet world is noisy. In this case, the desired performance may not be achieved if the data is not normalized. In some points, data elimination, deficiency correction, data reduction, and balancing may be required [20]. Depending on the data, preferred model structures may need noise normalization and necessary corrections to perform well. In this context, the Turkish dataset was first cleaned of unwanted characters with Regular Expression techniques. Besides, duplicate values were also deleted. Then, with the Zemberek NLP library, it was spell-checked and noise normalized. Zemberek is a natural language processing library written for Turkic languages, which is easy to use [21].

The implementation of these processes on the Turkish dataset resulted in a 2% increase in the performance of our model and also significantly reduced the amount of loss.

In the IMDB dataset, only Regular Expression techniques were used to clean unwanted characters in texts.

There are many types of problems to be encountered when dealing with texts. Since emojis are considered noisy during the model training phase, the role they play in classification is important. The emojis found should be removed from the data due to their negative effect on the performance. In the article by Alec et al, the effect of removing emojis on some machine learning methods is stated. For example, it has been explained that if SVM is used on data that is not purified from expressions, there will be a negative effect on the results, but the effect of emojis on Naive Bayes is smaller than SVM [22]. However, as another solution, phrases can also be represented by words determined by researchers.

Since the subject we are dealing with is text-based, the data in text form must be formed into vector sequences for models to be able to perform well.

As soon as the data which is supposed to be processed by any learning model must be transformed in a digitized/vectorized form. In this way, two techniques were used; CountVectorizer and Keras Tokenizer + Embedding Layer.

For machine learning models, CountVectorizer was preferred. It is a simple and useful technique for vectorizing text data. Simply, it returns a document-term matrix. Creating a matrix of $M \times N$ sizes where the M is the number of sentences in the dataset and the N is the number of unique words. Then, vectorizing data by assigning count values to the words due to their occurrences in sentences [26].

For the LSTM model, a different technique was preferred due to getting very low performance with CountVectorizer. Keras Tokenizer and Embedding Layer were used at this stage [27][28]. First, all data were represented by numbers (in a range of vocabulary size that is chosen by researchers). Then, the list of numbers of words was used in an Embedding Layer to make the data transform into vectors given by the layer network.

After all the data in the dataset have been digitized, their lengths must be equal to each other for the network model used. Therefore, the shortcomings of all sequences are complemented by adding zeros (padding) from the end of the sequence to the start, and as same as complementing with zeros, strings with longer lengths than fixed sequence length are truncated from end to start, too.

There are two types of padding and truncating features in sequencing, and they are called post/pre padding and post/pre truncating. Using post-padding for the LSTM network model, unlike truncating, often causes undesirable results because LSTM network model has a "forgetting" feature. Therefore, the use of pre-padding in LSTM networks produces more efficient results than the use of post-padding [23]. For example, if we continue by completing the sequences with zeros at the end, what the model will see continuously, will consist mostly of zeros. So, this will give undesired consequences.

3.3. Model Building

The reason of using the recurrent LSTM network is that it has a structure that is suitable for sequenced sentence vectors. The model, which has the structure shown in Figure 5, processes the elements of a sequence one by one and transfers them to the next stage, and gives the result at the end of the iteration. The elements of each array are processed and predicted one by one at the specified input size.

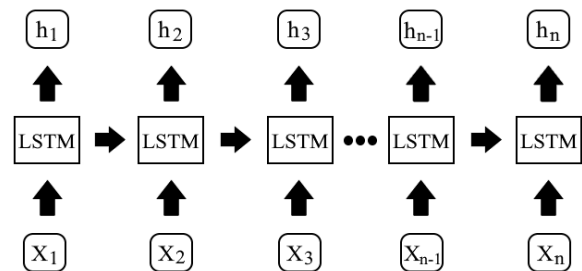


Figure 5. Long short-term memory network

For the training and test phases, data separations were made by random separation in the form of a 5-Fold Cross-Validation and the vocabulary size was chosen as 10000 for the Turkish dataset and 12000 for the IMDB dataset.

Later, the obtained vector sequences were used with the preferred learning models. LSTM network was then set up and sigmoid function was used as output activation. The reason for using Sigmoid is representing the input data as binary outputs. In this case, Binary Cross Entropy loss function was used. Also, the Rmsprop function was preferred as optimization function.

3.4. Testing and Evaluation

According to the tests performed after the compilation and training phase, our model has reached 90.59% of accuracy score on the test data of the Turkish dataset with five epochs and also it reached 89.02% of accuracy score on the test data of the IMDB dataset. At this stage, basic evaluation techniques were used. The classification performances of the models are shown in the following stage. In addition, the parameters of the LSTM models are shown in Table 1. Also, confusion matrices are given in Figure 6 to compare the models.

The case of classes separated as 0 and 1 in the reports shown in Table 2 and Table 3 are reported with precision, recall, f-score, and accuracy values. Precision; the ratio of true-positive predictions to all positive predicted values, Recall; the ratio of values predicted as true-positives to all positive values in the corpus, F-score; harmonic mean of precision and recall values and it gets a value between 1 and 0. In the accuracy, the measurement results are reported with the values of the ratio of correctly predicted values to the number of all values [24]. With the support values, the number of the values of the classes in the dataset is specified.

Table 1. LSTM model parameters for the datasets used

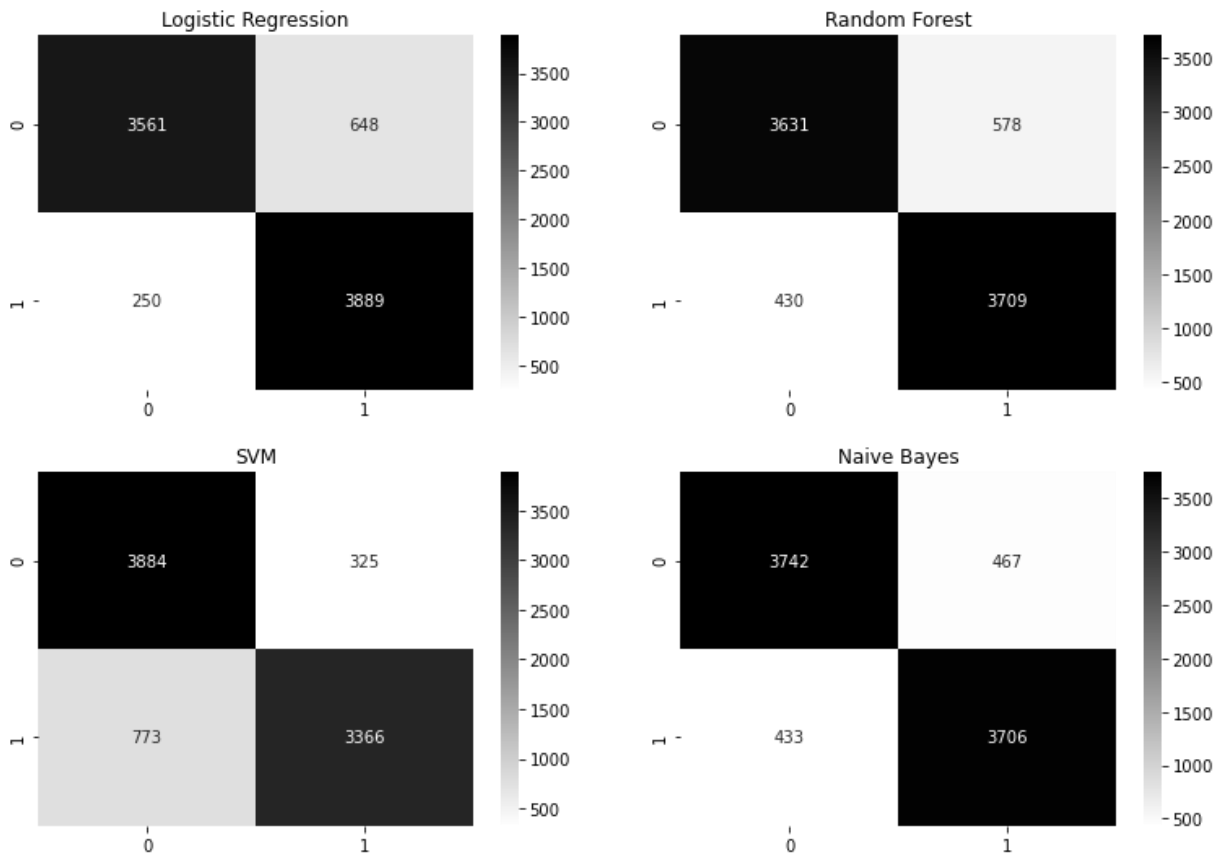
	Vocab. Size	Units	Activation	Optimizer	Loss Func.	Epochs	Batch Size	Dropout	Recc. Dropout
Turkish Customer Reviews	10000	196	Sigmoid	Rmsprop	Binary Crossentropy	5	32	0.3	0.3
Large Movie Reviews	12000	128	Sigmoid	Rmsprop	Binary Crossentropy	5	64	0.5	0.4

Table 2. Classification report for the IMDB dataset

Model	Class	Precision	Recall	F-score	Accuracy
LSTM Network	0	0.90	0.87	0.89	0.890200
	1	0.88	0.91	0.89	

Table 3. Comparison of learning models with 5-fold cross-validation

Model	Class	Precision	Recall	F-score	Accuracy
SVM	0	0.83	0.92	0.88	0.868470
	1	0.91	0.81	0.86	
Random Forest	0	0.90	0.86	0.88	0.878293
	1	0.87	0.90	0.88	
Logistic Regression	0	0.90	0.89	0.89	0.892189
	1	0.89	0.90	0.89	
Naive Bayes	0	0.93	0.85	0.89	0.892439
	1	0.86	0.94	0.90	
LSTM Network	0	0.90	0.92	0.90	0.905965
	1	0.91	0.90	0.90	



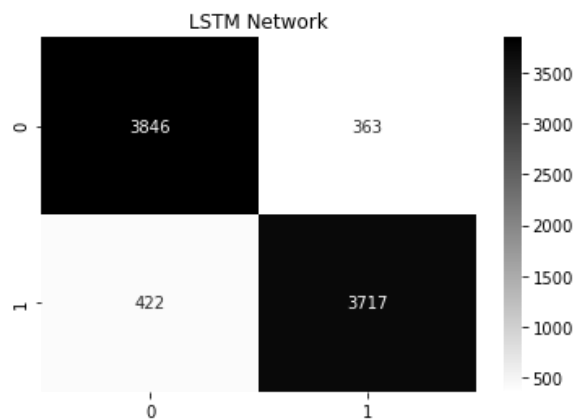


Figure 6. Classification matrices for comparison of the models

After the use of Turkish reviews, another dataset was used to compare the performance of the LSTM model. In this way, the article which has used the Large Movie Reviews dataset as origin was examined. Our model has obtained an accuracy score of 89.02%, without spell checking and noise normalization performing on the data. In the study of Maas et al., researchers have performed many approaches such as Bag of Words, LSA, etc. They have obtained the maximum accuracy of 88.89% with their hybrid featured model [25].

4. CONCLUSION

The subject of sentiment analysis, which is human-based, increasing significantly in the commercial sense, and is able to provide great benefits, has been studied in this paper and some test results were obtained and shared in the article. The sentiment analysis concept has a problem. As mentioned earlier, people can have complex sentiments or they may remain neutral, etc. many reasons can cause the problem. In our study, the classification was made in the binary system and an accuracy score of 90.59% was obtained on the Turkish dataset, and also an accuracy score of 89.02% was obtained on the IMDB dataset. These accuracies have been achieved with the preprocessing operation and the LSTM network used. This study achieved a higher accuracy score compared to other machine learning models used in the study. In addition, a new dataset has been added to this field in order to work on it.

DECLARATION OF ETHICAL STANDARDS

The authors of this article declare that the materials and methods they use in their studies do not require ethics committee permission and/or legal-specific permission.

AUTHORS' CONTRIBUTIONS

Burhan BİLEN: contributed to the design and implementation of the research, to the analysis of the results. He wrote the manuscript in consultation with Horasan.

Fahrettin HORASAN: supervised the project, contributed to the design of the research and analysis of the results.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

- [1] Pang B., Lee L. and Vaithyanathan S., "Thumbs up? Sentiment Classification Using Machine Learning Techniques", *Proceedings of EMNLP*, 10: 79-86, (2002).
- [2] Yıldırım S., Salman Y. B. and Ayvaz S., "Türkçe Duygu Kütüphanesi Geliştirme: Sosyal Medya Verileriyle Duygu Analizi Çalışması", *Avrupa Bilim ve Teknoloji Dergisi*, 16: 51-60, (2019).
- [3] Ayvaz S. and Shiha, M. O., "A Scalable Streaming Big Data Architecture for Real-Time Sentiment Analysis", *ICCBDC'18*, 47-51, (2018).
- [4] Brownlee J., "What is Deep Learning?", Retrieved From: <https://machinelearningmastery.com/what-is-deep-learning/>, (2019).
- [5] Amidi A. and Amidi S., "Recurrent Neural Networks cheatsheet", Retrieved From: <https://stanford.edu/~shervine/1/tr/teaching/cs-230/cheatsheet-recurrent-neural-networks>, (2019).
- [6] "Duygu Analizi için Türkçe Veri Seti", From: <https://www.kaggle.com/burhanbilenn/turkish-customer-reviews-for-binary-classification>, (2020).
- [7] Nałçakan Y., Bayramoğlu Ş. and Tuna S., "Sosyal Medya Verileri Üzerinde Yapay Öğrenme ile Duygu Analizi Çalışması", (2015).
- [8] Wilson T., Wiebe J. and Hoffmann P., "Recognizing contextual polarity in phrase-level sentiment analysis", *HLT/EMNLP*, 05: 347-354, (2005).
- [9] Dos Santos C. and Gatti de Bayser M., "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts", *Proceedings of COLING, 25th International Conference on Computational Linguistics: Technical*, 69-78, (2014).
- [10] Ruder S., Ghaffari P. and Breslin J., "A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis", *EMNLP*, 999-1005, (2016).
- [11] Dey L., Chakraborty S., Biswas A., Bose B. and Tiwari S., "Sentiment Analysis of Review Datasets Using Naïve Bayes and K-NN Classifier", *International Journal of Information Engineering and Electronic Business*, 8(4): 54-62, (2016).
- [12] Baid P., Gupta A. and Chaplot N., "Sentiment Analysis of Movie Reviews using Machine Learning Techniques", *International Journal of Computer Applications*, 179(7): 45-49, (2017).
- [13] Singla Z., Randhawa S. and Jain S., "Sentiment analysis of customer product reviews using machine learning", *International Conference on Intelligent Computing and Control (I2C2)*, 1-5, (2017).
- [14] Dehkharghani R., Saygin Y., Yanikoglu, B. and Oflazer, Kemal., "SentiTurkNet: a Turkish polarity lexicon for sentiment analysis", *Language Resources and Evaluation*, 50: 667-685, (2015).
- [15] Sağlam F., Genç B. and Sever H., "Extending a sentiment lexicon with synonym-antonym datasets: SWNetTR++", *Turkish Journal of Electrical Engineering and Computer Sciences*, 27: 1806-1820, (2019).
- [16] Kamisli Ozturk Z., Erzurum Cicek Z. and Ergul Aydin Z., "Sentiment Analysis: an Application to Anadolu University", *Acta Physica Polonica A*, 132: 753-755, (2017).
- [17] Wang Y., Huang M., Zhu X. and Zhao L., "Attention-based LSTM for Aspect-level Sentiment Classification",

- Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 606-615, (2016).
- [18] Arras L., Montavon G., Müller K., and Samek W., “Explaining Recurrent Neural Network Predictions in Sentiment Analysis”, *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 159-168, (2017).
- [19] Pant D., Neupane P., Poudel A., Pokhrel A. and Lama B., “Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis”, *IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 128-132, (2018).
- [20] Oğuzlar A., “Veri Ön İşleme”, *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21: 67-76, (2003).
- [21] Akin A. A. and Akin M. D., “Zemberek, an open source nlp framework for Turkic languages”, *Structure*, 10:1-5, (2007).
- [22] Go A., Bhayani R. and Huang L., “Twitter sentiment classification using distant supervision”, *Processing*, 1-6, (2009).
- [23] Dwarampudi M. and Reddy N. V., “Effects of padding on LSTMs and CNNs”, *arXiv*, (2019).
- [24] Tripathy A., Agrawal A. and Rath S., “Classification of Sentimental Reviews Using Machine Learning Techniques”, *Procedia Computer Science*, 57:821-829, (2015).
- [25] Maas A., Daly R., Pham P., Huang D., Ng Andrew and Potts C., “Learning Word Vectors for Sentiment Analysis”, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 1:142-150, (2011).
- [26] CountVectorizer, *Scikit-learn*, Retrieved From: https://www.scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.CountVectorizer.html
- [27] Text Tokenizer, *Keras*, Retrieved From: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/text/Tokenizer
- [28] Embedding Layer, *Keras*, Retrieved From: https://keras.io/api/layers/core_layers/embedding/