Turkish Tweet Classification with Transformer Encoder

Atıf Emre Yüksel* Boğaziçi University İstanbul, Turkey
Yaşar Alim Türkmen* Boğaziçi University İstanbul, Turkey
Arzucan Özgür Boğaziçi University İstanbul, Turkey
{ atif.yuksel, yasar.turkmen, arzucan.ozgur }@boun.edu.tr

Ayşe Berna Altınel Marmara University İstanbul, Turkey
berna.altinel@marmara.edu.tr

Abstract

Short-text classification is a challenging task, due to the sparsity and high dimensionality of the feature space. In this study, we aim to analyze and classify Turkish tweets based on their topics. Social media jargon and the agglutinative structure of the Turkish language makes this classification task even harder. As far as we know, this is the first study that uses a Transformer Encoder for short text classification in Turkish. The model is trained in a weakly supervised manner, where the training data set has been labeled automatically. Our results on the test set, which has been manually labeled, show that performing morphological analysis improves the classification performance of the traditional machine learning algorithms Random Forest, Naive Bayes, and Support Vector Machines. Still, the proposed approach achieves an F-score of 89.3% outperforming those algorithms by at least 5 points.

1 Introduction

Short-text usage is increasing day by day and we encounter short-text messages on many social media platforms in different forms such as tweets, Facebook status posts, or microblog entries. Twitter is one of the most widely used platforms, where a huge amount of short-texts are produced. More than 500 million tweets are posted on a typical day (Aslam, 2018). People use Twitter in order to produce and reach information faster about the topic they are interested in. Therefore, tweet classification becomes an important task to improve tweet filtering and tweet recommendation.

Traditional machine learning algorithms such as K-Nearest Neighbor, Support Vector Machines (SVM), and Naive Bayes have been widely used for text classification tasks and accuracy levels of over 80% have been reported in the literature for various data sets (Kadhim, 2019). However, obtaining a similar level of success for short-text classification is difficult, since short-texts contain smaller number of words compared to lengthy texts, which makes classifying them effectively a challenging task (Taksa et al., 2007). While a number of studies have been conducted for short-text classification, most of them have addressed the task of English tweet classification (Batool et al., 2013; Selvaperumal & Suruliandi, 2014).

In this paper, we tackle the task of Turkish tweet classification. The grammatical and syntactic features of the Turkish language pose additional challenges for short-text classification. The agglutinative nature of Turkish results in a high number of different word surface forms, since a root word can take many different derivational and inflectional affixes. This leads to the data sparseness problem. We propose a Transformer-Encoder based model for Turkish topic-based tweet classification and compare it with the traditional machine learning algorithms Naive Bayes, SVM, and Random Forest. The results show that morphological analysis enhances the performance of the traditional classification algorithms. However, the Transformer-Encoder model achieves the best F-score performance, even without any morphological analysis. Another contribution of this study is the constructed Turkish tweet data set on nine different topics. The tweet IDs and the corresponding topics are made available for future studies. [1]

The rest of this paper is organized as follows: in Section 2, we give a review of the related work

*Equal contribution

[1]The dataset can be obtained by e-mailing the authors.

Proceedings of Recent Advances in Natural Language Processing, pages 1380–1387,
Varna, Bulgaria, Sep 2–4, 2019.
https://doi.org/10.26615/978-954-452-056-4_158

on short-text classification, especially for Turkish. In Section 3, the data set creation and the steps of tweet preprocessing are explained in detail. Section 4 describes the proposed Transformer Encoder model and its usage. The experimental results and error analysis are presented in Section 5. Finally, Section 6 includes the conclusion and future works.

2 Related Work

Traditional text classification methods based on the BoW (Bag of words) model (Harris, 1954) suffer from high dimensionality and sparse feature sets, particularly in short-texts. Other limitations of BoW based models are that semantic features are not captured, the positions of the words in the text are not considered, and the words are assumed to be independent from each other.

In order to overcome the weaknesses of BoW, Latent Dirichlet Allocation (LDA) (Blei et al., 2003), where the terms are mapped to distributional representations based on latent topics within documents, has been used for building classifiers that deal with short and sparse text (Phan et al., 2008; Song et al., 2014).

One of the most commonly used algorithms for short text classification is Naive Bayes, which is based on word occurrence and class priors. For example, Kiritchenko & Matwin (2011) used this algorithm on an email dataset and Kim et al. (2006) offered powerful techniques for improving text classification by Naive Bayes. Sriram et al. (2010) extracted different features from short-texts and used these with Naive Bayes to classify them.

Support Vector Machine (SVM) is another commonly used algorithm in short text classification studies. Pointing to the weaknesses of the BoW approach, different kernels have been developed for SVM such as semantic kernels that use TF-IDF (Salton & Buckley, 1988) and its variants that apply different term weighting functions on the term incidence matrix. Wang & Manning (2012) showed that Naive Bayes obtains higher scores than SVM for short-text sentiment classification tasks and the combination of SVM and Naive Bayes outperforms SVM and Naive Bayes for some of the datasets that they used. More relevant to our study, Lee et al. (2011) used SVM for text-based classification of trending topics under 18 classes and reached 61.76% accuracy. They obtained 65.36% accuracy with Multinomial Naive Bayes.

The advances of deep learning in NLP led to the use of Artificial Neural Network (ANN) based models in recent studies. Convolutional Neural Networks (CNNs) and Dynamic CNNs have been applied to text classificatioon and promising results have been obtained (Kim, 2014; Kalchbrenner et al., 2014). In a recent study, Le & Mikolov (2014) successfully added sequential information by using Recurrent and Convolutional Neural Networks for sequential short-text classification.

In this study, we address the task of Turkish short text classification. Turkish is an agglutinative language, which may result in the same word to map to different features when it takes different inflectional affixes or suffixes. Possessive pronouns, tenses, and auxiliary verbs can be encoded as affixes of the words. For this reason, a word may have many different forms and this poses challenges for classical term weighting methods to construct strong relations between the text and its topic. Most prior work on Turkish short text classification is on sentiment analysis. Demirci (2014) used Naive Bayes, SVM and k Nearest Neighbor (kNN) for emotion analysis on Turkish tweets and obtained accuracy levels of up to 69.9%. Similarly, Yelmen et al. (2018) showed that SVM reaches 80% accuracy for sentiment classification of Turkish tweets for GSM operators, whereas an Artificial Neural Network (ANN) results in slightly lower performance. Türkmenoğlu & Tantuğ (2014) compared lexicon based sentiment analysis and machine learning based sentiment analysis on tweet and movie review datasets and concluded that the machine learning based algorithms SVM, Naive Bayes, and decision trees achieve better scores. Yıldırım et al. (2014) combined lexicon and machine learning based methods to improve sentiment analysis of Turkish tweets.

Unlike prior studies on Turkish short text classification that address the task of sentiment analysis, we tackle the task of topic-based tweet classification and create a Turkish tweet data set for nine different topics. We propose using Transformer Encoder architecture for Turkish short-text classification. Transformer Encoder is a recently proposed model that offers a better understanding of the language structure by protecting the semantic values and meanings of word sequences (Vaswani et al., 2017). Furthermore, the positions of the terms in text are taken into account and the terms are not assumed to be independent from each other by using a contextual word embedding representation.

3 Dataset

In this section, we briefly explain how we created the dataset and the main steps of data preprocessing, as well as the importance of lemmatization on Turkish short-text classification.

3.1 Dataset Creation

The dataset includes 164,549 Turkish tweets, written by 74 different users until February 2019. These tweets are retrieved from the users' profiles, who are known experts in their areas and mostly write tweets on the subjects of their expertise. The tweets of each user are automatically labelled with the topic of expertise of the user. The dataset contains nine topics, namely politics, economics & investment, health, technology & informatics, history, literature & film, sports, education & personal growth, and magazine. The selection of topics was made in a similar way the news sites categorise their content.

We observed that some of the tweets of a user could be mislabeled, since a user may also tweet about topics different than his/her area of expertise, which results in noise. In order to ensure that most of the tweets are related to their assigned topics, we selected a random sample of tweets and manually checked the percentage of the correctly labelled ones. The percentage of the correctly labelled tweets was around 80%. That is, the training dataset contains around 20% noise (i.e., incorrectly labeled tweets). We randomly selected 10% of the dataset as a test set. In order to report reliable results, we manually verified and corrected the labels of the test set.

3.2 Data Preprocessing

Preprocessing Turkish sentences is quite different from the English ones, since Turkish is an agglutinative language and we may encounter words in many different forms. In addition, tweets come with their own difficulties when they are used in natural language processing. They contain hashtags, mentions, emojis, and links which make tokenization difficult. To overcome those challenges we follow three main steps for preprocessing as described in the following subsections.

3.2.1 Data Cleaning

Users use some adhoc pattern such as hashtags, mentions, the link of website they want to refer to, and emojis in tweets.In our work, we are only interested in lexical terms in tweets. Hence, we drop the hashtags, mentions, links, emojis, numbers, and punctuations. On the other hand, hashtags can be quite useful when deciding the topic of a tweet. There are some studies to split hashtags (Çelebi & Özgür, 2018) into meaningful words in English. However, it is left as an open future work for Turkish hashtags.

Some tweets contain only response to a tweet like "greetings", "thank you", "okay", or "congratulations" for good news etc. When we examine the tweets which have less than 4 words, nearly all of these tweets are examples of such tweets. These tweets are removed from the dataset, since they are not about any specific topic.

In addition, when we examined the tweets with retweet and like statistics that are significantly different from the others, we discovered that those tweets are mostly retweets of campaigns or calling for help. For this reason, we filtered this type of tweets from the dataset.

3.2.2 Language Identification

In Twitter, users sometimes write and retweet tweets in languages different from their native language. When we analysed our dataset, we observed that the languages used by users are Turkish, English, Arabic, French, and German. For filtering non-Turkish tweets, we used the language identification model in (Lui & Baldwin, 2012).

After the dataset cleaning and language-based filtering steps, the training set contains 119,778 tweets and the test set contains 3050 tweets.

3.2.3 Lemmatization

The goal of lemmatization is to find the lemmas of the words by removing the prefixes, suffixes, and other types of affixes based on morphological analysis. This process is harder for the Turkish language, since it has more types of affixes than English.

In the Naive Bayes model, the frequency of a word is prominent for scoring the importance of the word for each class. Therefore, finding the correct lemmas of the words is a useful transformation to classify tweets correctly. Lemmatization is also important for the SVM classifier, in order to compute the similarity between the instances in the training set, as well as the support vectors and the test instances in the classification step more accurately. Yıldırım et al. (2014) showed the positive impact of morphological analysis for the sentiment analysis of Turkish tweets. Therefore, in order to increase the success of the Naive Bayes and SVM models, we use a lemmatization model (Sak et al., 2008) which is trained by nearly one million Turkish sentences. An example tweet and its lemmatized version are shown in Figure 1. The effects of lemmatization on the success of Turkish tweet classification are presented in Section 5.

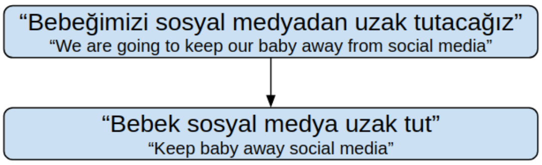

Figure 1: A sample tweet from the dataset and its lemmatized from.

4 Model Description

In this section, the main components of the proposed model are presented. First, the general structure of the input embeddings is explained. Then, the architecture of the Transformer Encoder model is described in detail.

4.1 Input Embeddings

One of the most widely used and robust word embedding models, word2vec, was proposed by Mikolov et al. (2013). Besides word embeddings, sentence embeddings (Logeswaran & Lee, 2018) and paragraph embeddings (Le & Mikolov, 2014) were studied in order to represent larger text snippets correctly and classify documents accordingly. Aiming to construct the word embeddings based on the context of each word, Peters et al. (2018) offered a new word embedding representation model named as ELMo. ELMo tries to keep the contextual features in word embedding representations so that the polysemy and homonymy problems are alleviated.

In our study, pretrained contextual word embedding representations of Turkish words (Devlin et al., 2019) are used. They are obtained by using the bidirectional approach with masked language model (Taylor, 1953) in training. Hence, the representation of each word contains the information of the left and right words in their own context. Moreover, in order to account for the word order in text, positional embeddings (Vaswani et al., 2017) are used and combined with the word embeddings. When using the input embeddings, they are fed into the Transformer Encoder by matching each word in the input tweet with the longest token in the vocabulary of the pretrained contextual word embeddings. An example from the dataset is shown in Figure 2.

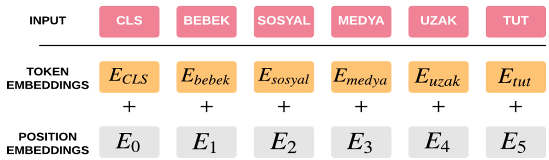

Figure 2: Input embeddings representation.

4.2 Transformer Encoder

This model is based on the Transformer architecture (Vaswani et al., 2017), which contains self-attention layers. What we aim to find in this architecture is reaching the importance of every word in tweets by training the query and key matrices of the attention layers. Hence, the attention score is calculated for every word in a tweet in order to determine its usefulness for each class. Lots of different attention patterns are obtained with multi-head self attention layers, which enable the model to decide which attention score is significant among the pairs of the terms in the tweet.

In general, the transformer architecture is constructed by combining the layers of multi self-attention heads, Dropout, Layer Normalization, and 2 fully-connected layers, respectively. A pictorial representation of this Transformer architecture is shown as a grey block in Figure 3.

In the general architecture of the Transformer Encoder, Layer Normalization and Dropout layers are applied to the contextual word embeddings, into which positional information is added before it enters into the Transformer. In the Transformer, multi-head self attention layers are used as we mentioned above. After the Transformer obtains the attention score matrices containing the score of every word with different attention patterns, these inter-layer features are connected into 2 fully-connected layers in order to reach the combination of different attention scores of the terms. Then, the class prediction of the tweet is found after passing the pooler, dropout, and the last fully-

Model	Accuracy	Precision	Recall	F1-Score
Random Forest (Baseline)	65.0	64.9	64.5	64.5
Random Forest + Lemmatization	71.1	70.9	70.9	70.7
Random Forest + Lemmatization + TF-IDF	69.9	69.8	69.6	69.6
Naive Bayes	82.7	82.4	82.4	82.6
Naive Bayes + Lemmatization	85.0	84.8	84.7	84.9
Naive Bayes + Lemmatization + TF-IDF	85.4	85.3	85.4	85.2
SVM	78.6	78.2	78.3	78.1
SVM + Lemmatization	80.3	80.3	80.2	80.1
SVM + Lemmatization + TF-IDF	84.4	84.2	84.3	84.2
Transformer Encoder + Input Embeddings	**89.4**	**89.3**	**89.4**	**89.3**

Table 1: Comparison of the models' scores over the manually labeled test set.

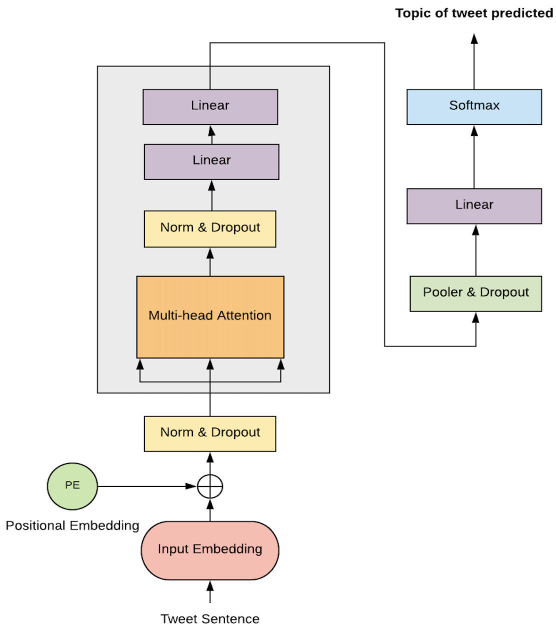

Figure 3: The architecture of the Transformer Encoder model.

connected layer. The class probabilities are calculated by using the Softmax classifier. The general architecture of the model is shown in Figure 3.

5 Experiments

We compared the performance of the Transformer Encoder model with three different baseline models, which are widely used in the area of short-text classification: Naive Bayes, SVM and Random Forest. The parameters of the models are tuned using cross-validation over the training set and the performance results are reported over the manually labeled test set (Table 1).

In the training phase of the Transformer Encoder, 10 epochs are run on the dataset. Batch size of training data is selected as 8 due to computational constraints. Maximum sequence length is 128, learning rate is equal to 2e-5, and the number of attention heads is equal to 12 in our configuration.

For Naive Bayes, SVM and Random Forest the most frequent 15000 words are selected as the features to represent the data. We also experimented with using the most frequent 10000, 20000, and 25000 words, however, the results in the cross-validation experiments were lower. So, we report the results with 15000 words over the test set. We also investigated the effect of lemmatization and TF-IDF weighting on the baseline models.

5.1 Comparison Results

The results of the compared models over the test set are shown in Table 1. Naive Bayes, Random Forest, and SVM obtain better scores when morphological analysis is performed.

However, in spite of these improvements, the proposed Transformer Encoder model achieves the highest scores in all metrics for tweet classification even though it does not use the morphological analysis. These results point out that using pretrained word embeddings, which are trained on large datasets, with a Transformer Encoder architecture is more effective than using morphological analysis and TF-IDF based term weighting with the traditional machine learning algorithms Random Forest, Naive Bayes, and SVM for Turkish short text classification.

5.2 Error Analysis

Figure 4 shows the confusion matrix for the Transformer Encoder model over the test set.

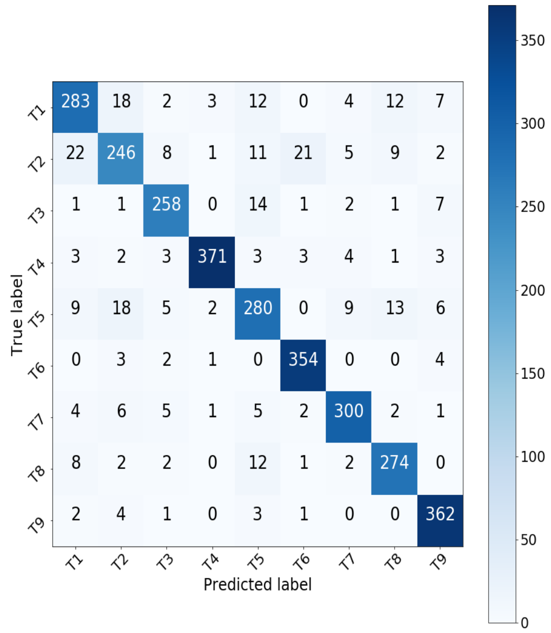

Figure 4: Confusion matrix for the Transformer Encoder. The names of the topics (from T1 to T9) are literature & film, education & personal growth, economy & investment, magazine, politics, health, sport, history, and technology & informatics, respectively.

It is observed that most of the errors are caused by the proximity of the topics like politics and history or literature & film and education & personal growth. The main reason is that there are many common words which are frequently used in close topics. In many cases, it is hard to differentiate a tweet about recent history from a tweet about politics. Similarly, an expression from a novel can be similar to an expression written by an educator. For example[1], the topic of the tweet "*Bachmann is so right, the real death of a man is not from diseases, but from what another man does to him.*" is predicted by the Transformer Encoder model as education & personal growth, whereas its correct label is literature & film. As another misclassified example the topic of the tweet "*The Lausanne Treaty debate should be discussed in public.*" is predicted as politics, whereas this tweet is in the scope of the recent history.

[1]The original tweets are in Turkish. Their English translations are provided here for easier comprehension.

6 Conclusion and Future Works

We proposed using a Transformer Encoder architecture with contextual word embeddings and positional embeddings for Turkish short text classification. In addition, we compiled a dataset of Turkish tweets for topic-based classification. The training set has been labeled automatically using a weakly supervised approach, whereas the test set has been manually labeled for reliable evaluation. Our results demonstrate that the Transformer Encoder model outperforms the widely used machine learning models for topic-based Turkish tweet classification. Also, this study shows the importance of morphological analysis for Turkish short-text classification with the widely used machine learning algorithms SVM, Naive Bayes, and Random Forest. It is interesting to note that, the Transformer Encoder model is able to outperform these algorithms, even though it does not involve a morphological analysis step.

One of the areas of usage of this study is guiding Twitter users and making easier for them to find correct accounts to follow. Twitter is one of the most used social media platforms in the world in order to get news and learn about what people think (Riquelme & Gonzlez-Cantergiani, 2016). Users may prefer Twitter for getting breaking news or reading about a social event. Therefore, it is necessary to find the correct people to follow and be aware of the popular and informative tweets. By using our model, Turkish tweets can be categorized and the most relevant ones can be suggested to users according to their interests, which would lead to saving time and manual effort.

References

Aslam, S. (2018). Twitter by the numbers: Stats, demographics fun facts. https://www.omnicoreagency.com/twitter-statistics. Accessed: 2019-01-06.

Batool, R., Khattak, A. M., Maqbool, J., & Lee, S. (2013). Precise tweet classification and sentiment analysis. In *2013 IEEE/ACIS 12th International*

Conference on Computer and Information Science (ICIS), (pp. 461–466).

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, *3*, 993–1022.

Çelebi, A. & Özgür, A. (2018). Segmenting hashtags and analyzing their grammatical structure. *Journal of the Association for Information Science and Technology*, *69*.

Demirci, S. (2014). Emotion analysis on Turkish tweets. Master's thesis, Middle East Technical University.

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)., Minneapolis, Minnesota. Association for Computational Linguistics.

Harris, Z. S. (1954). Distributional structure. *WORD*, *10*(2-3), 146–162.

Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, *52*(1), 273–292.

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A convolutional neural network for modelling sentences. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (pp. 655–665)., Baltimore, Maryland. Association for Computational Linguistics.

Kim, S. B., Han, K. S., Rim, H. C., & Myaeng, S. H. (2006). Some effective techniques for Naive Bayes text classification. *IEEE Transactions on Knowledge and Data Engineering*, *18*(11), 1457–1466.

Kim, Y. (2014). Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (pp. 1746–1751)., Doha, Qatar. Association for Computational Linguistics.

Kiritchenko, S. & Matwin, S. (2011). Email classification with co-training. In *Proceedings of the 2011 Conference of the Center for Advanced Studies on Collaborative Research*, CASCON '11, (pp. 301–312). IBM Corp.

Le, Q. & Mikolov, T. (2014). Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning - Volume 32*, ICML'14, (pp. II–1188–II–1196). JMLR.org.

Lee, K., Palsetia, D., Narayanan, R., Patwary, M., Agrawal, A., & Choudhary, A. (2011). Twitter trending topic classification. *2011 11th IEEE International Conference on Data Mining Workshops*.

Logeswaran, L. & Lee, H. (2018). An efficient framework for learning sentence representations. *CoRR*, *abs/1803.02893*.

Lui, M. & Baldwin, T. (2012). Langid.py: An off-the-shelf language identification tool. In *Proceedings of the ACL 2012 System Demonstrations*, ACL '12, (pp. 25–30)., Stroudsburg, PA, USA. Association for Computational Linguistics.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances In Neural Information Processing Systems*, (pp. 3111–3119).

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. In *Proc. of NAACL*.

Phan, X., Nguyen, L., & Horiguchi, S. (2008). Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, (pp. 91–100)., New York, NY, USA. ACM.

Riquelme, F. & Gonzlez-Cantergiani, P. (2016). Measuring user influence on Twitter: A survey. *Information Processing & Management*, *52*(5), 949–975.

Sak, H., Güngör, T., & Saraçlar, M. (2008). Turkish language resources: Morphological parser, morphological disambiguator and web corpus. In Nordström, B. & Ranta, A. (Eds.), *Advances in Natural Language Processing*, (pp. 417–427)., Berlin, Heidelberg. Springer Berlin Heidelberg.

Salton, G. & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513 – 523.

Selvaperumal, P. & Suruliandi, A. (2014). A short message classification algorithm for tweet classification. In *2014 International Conference on Recent Trends in Information Technology*, (pp. 1–3).

Song, G., Yunming, Y., Xiaolin, D., Huang, X., & Shifu, B. (2014). Short text classification: A survey. *Journal of Multimedia*, *9*(5), 635.

Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., & Demirbas, M. (2010). Short text classification in Twitter to improve information filtering. In *Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Taksa, I., Zelikovitz, S., & Spink, A. (2007). Using web search logs to identify query classification terms. *International Journal of Web Information Systems*, *3*, 315–327.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Bulletin*.

Türkmenoğlu, C. & Tantuğ, A. C. (2014). Sentiment analysis in Turkish media. In *Proceedings of Workshop on Issues of Sentiment Discovery and Opinion Mining, International Conference on Machine Learning (ICML)*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances In Neural Information Processing Systems*, (pp. 5998–6008).

Wang, S. & Manning, C. D. (2012). Baselines and bigrams: Simple, good sentiment and topic classification. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2*, ACL '12, (pp. 90–94)., Stroudsburg, PA, USA. Association for Computational Linguistics.

Yelmen, İ., Zontul, M., Kaynar, O., & Sönmez, F. (2018). A novel hybrid approach for sentiment classification of Turkish tweets for GSM operators. *International Journal of Circuits, Systems and Signal Processing*.

Yıldırım, E., Çetin, F. S., Eryiğit, G., & Temel, T. (2014). The impact of NLP on turkish sentiment analysis. *Turkiye Bilisim Vakfi Bilgisayar Bilimleri ve Muhendisligi Dergisi*, *7*, 43 – 51.