ESTUDAM Bilişim Dergisi
Cilt 1, Sayı 1, 8-11, 2019

Journal of ESTUDAM Information
Volume 1, Issue 1, 8-11, 2019

DOI: *****

(Geliş Tarihi / Received Date: 03.04.2019, Kabul Tarihi/ Accepted Date: 11.04.2019)

Orange 3 ile Türkçe ve İngilizce SMS Mesajlarında Spam Tespiti

Özlem ÖRNEK[*1]

[1]Eskişehir Osmangazi Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, Eskişehir

Anahtar Kelimeler:
SMS Spam Tespiti,
Metin Madenciliği,
Sınıflandırma

Özet: Kısa mesajlaşma servisi (SMS), haberleşme ve bilgilendirme için hızlı ve etkili bir yoldur. Kısa süre içerisinde büyük kitlelere ulaşabilme imkanı sağlayabilir. Ancak bu özelliklerin kötü amaçlara yönelik kullanılması kullanıcılar için problem oluşturmaktadır. İstenmeyen, kandırma amaçlı, kötü içerikli ve yanlış bilgi içeren vb. mesajlar gönderilebilmektedir. Bu problemlerin giderilmesi ve daha güvenli bir ortam sağlanması amacıyla bu çalışmada TurkishSMS mesaj ve UCI SMS Spam koleksiyonları kullanılarak Türkçe ve İngilizce içeriklere sahip SMS'ler için spam tespiti yapılmıştır. Çalışmada doğruluk ve hata oranları baz alındığında TurkishSMS veri kümesi için Sinir Ağları, UCI SMS Spam veri kümesi için ise Naive Bayes algoritmasının büyük bir doğruluk ve daha az kaçırma oranına sahip olduğu tespit edilmiştir.

Spam Detection in Turkish and English SMS Messages with Orange 3

Keywords:
SMS Spam Detection,
Text Mining,
Classification

Abstract: Short messaging service (SMS) is a fast and effective way to communicate and inform. Also, it can provide access to large masses in a short time. However, the use of these features for malicious purposes is a problem for users. Unsolicited messages can be sent like cheating, bad content and misinformation etc. In order to solve these problems and to provide a safer environment, spam detection was done for Turkish and English SMS messages by using TurkishSMS message and UCI SMS Spam collections. Based on the accuracy and error rates in the study, it was determined that the Neural Network for the TurkishSMS dataset and the Naive Bayes algorithm for the UCI SMS Spam data set had a great accuracy and less missing rate.

1. GİRİŞ

Kısa mesajlaşma servisi (SMS) spam, mobil ağ üzerinden gönderilen istenmeyen mesajlar olarak tanımlanabilir. İstenmeyen SMS mesajları günlük yaşamımızda sık rastlanan bir durumdur ve iletişim süresini, bant genişliğini ve kaynakları tüketir [1]. Ayrıca spam SMS'ler nedeniyle kötü amaçlı yazılım veya virüs sorunları ile karşı karşıya kalınabilir. Bu nedenlerle mobil SMS spam tespiti önemli bir konudur.

Sıkıntıları azaltmak amacıyla SMS spamlarının tespit edilmesi için çeşitli yöntemler ve teknikler önerilmiştir [2-6]. Çalışmada SMS spamlarının tespit edilmesi ile ilgili bilgiler verilmektedir. Spam SMS tanımlamak ve sınıflandırmak için Orange 3 uygulaması iki farklı veri kümesinde sınıflandırma algoritmaları uygulanmıştır. Amaç doğruluk ve hata oranı temel alınarak uygun bir algoritma bulmaktadır.

Bölüm 2'de metin madenciliği ön işlemleri ve SMS spam sınıflandırılması, bölüm 3'de önerilen yöntem ve son olarak bölüm 4'de sonuçlar verilmektedir.

2. SMS SPAM TESPİTİ

SMS spam tespiti için filtreleme veya sınıflandırma gibi çeşitli teknikler kullanılmaktadır. Teknikler, Erişim Katmanı (AL) veya Hizmet Sağlayıcı Katmanı (SPL) içinde çalışmak üzere tasarlanmaktadır [1]. AL ve SPL'de Destek Vektör Makineleri (SVM), Bayesian Ağları ve Makine Öğrenmesi tabanlı vb. algoritmalar kullanılmaktadır [7-9].

Modupe vd., Naive Bayes ve J48 algoritmasını kullanarak SMS spam tespitini gerçekleştirmiş ve elde edilen algoritma sonuçlarını karşılaştırmıştır [10]. Sonuçların birbirine yakın olduğu görülmüştür. Alzahrani vd., SMS tespiti için Bayes, SVM, C4.5 VE PART algoritmalarının sonuçlarını karşılaştırmış, SVM algoritmasının daha iyi performans gösterdiğini tespit etmiştir [11]. Yadav vd.

*İlgili yazar: ozlemmornek@gmail.com

çalışmalarında Bayes algoritmasını kullanarak SMS spam tespiti yapmıştır [12].

2.1. Metin madenciliği ile ön işlem

Metin (Text) madenciliği metinleri analiz etmek ve bilgi çıkarmak için kullanılan veri madenciliği dallarından biridir. Metin madenciliği, SMS'lerin spam ve spam olmayan olarak sınıflandırılması için kullanılmaktadır. Sınıflandırma işleminin gerçekleştirilmesi için önce metin madenciliği ile metnin hazırlanması için bazı ön işlemler uygulanmaktadır. Bunlar:

- Tokenizasyon,
- Lemmatizasyon,
- Terim ağırlıklandırma ve
- Özellik seçimi olarak verilebilir [13-14].

Tokenizasyon, mesajdaki kelimeleri çıkarmak için kullanılan bir süreçtir [13]. Lemmatizasyon, aynı kelimeleri gruplamak ve hesaplamak için bir süreçtir. Çalışmalara göre, lemmatizasyon süreci benzer ve gereksiz kelimeleri belirlemeyi amaçlamaktadır [13]. Bu tekrarları belirleyerek, ihtiyaç fazlalığının kaldırılması ve SMS için algılama doğruluğunun geliştirilmesi sağlanabilir [13]. Ancak kullanıcı diğer karakterleri kaldırırken kök kelimeyi veya temel kelimeleri seçerse, bu işlem algılama doğruluğunu azaltabilir [13]. Terim ağırlıklandırma, bir özniteliğin bir dokümanda gözlemlenme sıklığıdır [14]. Bir terimin bir dokümandaki ağırlığı ne kadar yüksekse bu terim o doküman için o kadar ayırt edici özelliğe sahip demektir [14]. Bu sebeple terimlerin ağırlıklandırılması sınıflandırma başarısını etkileyen önemli bir faktördür. Özellik seçim tekniğinin yararı, SMS saldırılarını daha ayrıntılı ve doğru bir şekilde sınıflandırmak ve saptamak için gereksiz özelliklerin ortadan kaldırılmasına yardımcı olmasıdır [13].

2.2. Sınıflandırma

Sınıflandırma, modele göre sınıf etiketini belirli bir örnek için tahmin etmektir. Sınıf etiketi kullanıldığı için denetimli öğrenme (supervised learning) tekniğidir. Sınıflandırma iki adımda çalışmaktadır. Bunlar:

- Model yapımı ve
- Bir modelin kullanılmasıdır.

İlk adımda model, sınıf etiketinin bulunduğu örnek veri kümesinden oluşturulur. Bu adımın çıktısı algoritmaya göre farklı olabilen bir çeşit kural, matematiksel veya istatistiksel formüldür [15]. Bir sonraki adımda oluşturulan model, sınıf etiketlerinin atanmadığı, görünmeyen veriler veya test verilerine uygulanır [15]. Bu adımda, algoritmanın doğruluğu bulunur ve doğruluk kabul edilebilir bir değer ise, sınıf etiketlerini tahmin etmek için test verilerine model uygulanır.

3. ÖNERİLEN YÖNTEM

3.1. Veri kümesi

Çalışmada UCI SMS Spam koleksiyonu [16] ve TurkishSMS mesaj koleksiyonu [17] kullanılmıştır. UCI SMS Spam Koleksiyonu, mobil SMS spam çalışmaları için oluşturulan ve toplanan ücretsiz spam metin mesajı veri kümesidir. Koleksiyon, mesaj ve mesajın hangi sınıfa ait olduğu bilgisini içeren satırlardan oluşan metin dosyası formatındadır ve dili İngilizcedir. Toplamda 5574 adet örnek bulunmaktadır. Veri kümesi içerisinde 4827 adet spam olmayan 747 adet spam olan SMS mesajı bulunmaktadır.

TurkishSMS mesaj koleksiyonu, akademik literatür için oluşturulan ilk Türkçe mesaj koleksiyonudur. Koleksiyon, 1103F054 nolu hibe kapsamında Anadolu Üniversitesi tarafından finanse edilen Bilimsel Araştırma Projesi kapsamında Alper Kürşat Uysal, Serkan Günal, Semih Ergin ve Efnan Sora Günal tarafından hazırlanmıştır. Toplamda 850 adet örnek bulunmaktadır. Veri kümesi içerisinde 430 adet spam olmayan 420 adet spam olan SMS mesajı bulunmaktadır.

3.2. Yöntem

İlk olarak SMS spam tespiti için SMS veri kümeleri transformasyon, tokenizasyon ve filtreleme ön işlemlerinden geçirilir. Transformasyon ile tüm SMS mesajları küçük harfe dönüştürülür [18]. Tokenizasyon aşaması, yalnızca varsayılan olarak noktalama işaretlerini çıkarır ve sözcüklere ayırır [18]. Filtreleme, durma sözcüklerini metinden kaldırır. Filtre uygulanacak dil, UCI SMS Spam koleksiyonu için İngilizce, TurkishSMS için Türkçe olarak ayarlanmıştır. Ön işlemden geçirilen veri kümelerine "Bag of Words" işlemi uygulanır. Bu işlem ile veri kümesindeki SMS mesajların içerdiği kelimeler ve kelimelerin frekansları (sayıları) elde edilir. İşlemlerden sonra elde edilen veri kümesi UCI SMS Spam koleksiyonunun metin özniteliği 5574 örnek ve 8609 adet kelime, TurkishSMS koleksiyonunun metin özniteliği 850 örnek ve 3750 adet kelime içermektedir.

Son aşama sınıflandırmadır. Spam algılama sınıflandırma süreci eğitim ve test aşamasından oluşmaktadır. Eğitim aşamasında spam ve spam olmayan SMS içeren veri kümeleri kullanılmaktadır. Eğitim sonucunda model elde edilir. Daha sonra test aşamasına geçilir. Test aşamasında oluşan model ile mesajlar kullanılarak sınıflandırma işlemi gerçekleştirilir. Tüm algoritmalar için "10 kat çapraz doğrulama (10 fold cross validation)" yöntemi kullanılmıştır. 10 kat çapraz doğrulama ile sınıflandırma yönteminde, veri kümesi 10'a bölünür. Biri, test veri seti olarak kullanılırken, diğer dokuzu, eğitim veri seti olarak kullanılır. Amaç, bu veri kümesi için en iyi sınıflandırıcıyı tanımlamaktır. Son olarak, sınıflandırıcıların performansı analiz edilir ve doğruluk, zaman ve hata oranı temelinde en iyi performansa sahip sınıflandırıcı seçilir [15].

Çalışmada literatürde de SMS spam tespiti için kullanılmış olan Sinir Ağları, Naive Bayes, Lojistik Regresyon, Random Forest, AdaBoost, Tree, SVM ve kNN sınıflandırma algoritmaları Orange 3 uygulaması ile iki veri kümesi içinde uygulanmaktadır. Orange 3, acemi ve uzmanlar için açık kaynaklı makine öğrenimi ve veri görselleştirmesini etkileşimli olarak gerçekleştiren bir uygulamadır [19].

4. DENEY SONUÇLARI

TurkishSMS ve UCI SMS Spam veri kümeleri sınıflandırma işlemi için gerekli ön işlemlerden geçirildikten sonra SMS spam tespitinin yapılacağı sınıflandırma aşamasına geçilmektedir. Sınıflandırma için Sinir Ağları, Naive Bayes, Lojistik Regresyon, Random Forest, AdaBoost, Tree, SVM ve kNN sınıflandırma algoritmaları kullanılmış ve algoritmalardan elde edilen sonuçlar incelenmiştir. Tablo 1'de TurkishSMS, Tablo 2'de ise UCI SMS Spam veri kümesinin sınıflandırma algoritmaları ile eğitim ve test işlemleri sonucu elde edilen doğruluk ve spam kaçırma oranları (FN Oranı) verilmektedir.

Tablo 1'de görüldüğü üzere TurkishSMS veri kümesi için en yüksek doğruluk oranı sinir ağları, en küçük kaçırma oranı ise SVM algoritması ile elde edilmiştir. En düşük doğruluk ve en yüksek kaçırma oranını kNN algoritması vermiştir. Tablo 2'de ise UCI SMS Spam veri kümesi için en yüksek doğruluk oranı Naive Bayes ve en küçük kaçırma oranı SVM algoritması ile elde edilmiştir. En düşük doğruluk oranı SVM, en yüksek kaçırma oranını ise kNN algoritması vermiştir.

Her iki veri kümesi için uygulanan algoritmaların doğruluk ile kaçırma oranları arasında belli bir ilişki görülmemiştir.

Tablo 1. TurkishSMS veri kümesi sınıflandırma algoritmaları değerlendirme metrikleri

Algoritma	Doğruluk Oranı	FN Oranı
Sinir Ağları	**98.4**	**0.017**
Naive Bayes	97.8	0.031
Random Forest	96.9	0.033
Lojistik Regresyon	96.8	0.048
AdaBoost	95.2	0.055
Tree	92.5	0.105
SVM	80.5	0.002
kNN	61.2	0.786

Tablo 2. UCI SMS Spam veri kümesi sınıflandırma algoritmaları değerlendirme metrikleri

Algoritma	Doğruluk Oranı	FN Oranı
Naive Bayes	**98.4**	**0.064**
Sinir Ağları	98.4	0.118
Lojistik Regresyon	98.3	0.116
Random Forest	97.8	0.158
AdaBoost	97.3	0.131
Tree	95.0	0.336
kNN	91.6	0.627
SVM	50.6	0.033

5. TARTIŞMA VE SONUÇ

Çalışmada SMS spam mesajı sınıflandırma, tanımlama için Orange 3 uygulaması kullanılarak en iyi algoritmayı bulmak amaçlanmıştır. Değerlendirme sonuçları, farklı algoritmalar için doğruluk ve hata oranlarının farklı olduğunu göstermektedir. Bu çalışma doğruluk ve hata oranları baz alındığında TurkishSMS veri kümesi için Sinir Ağları, UCI SMS Spam veri kümesi için ise Naive Bayes algoritmasının büyük bir doğruluk ve daha az kaçırma oranına sahip olduğunu göstermektedir. Veri kümelerin en temel farkı olan metin dilinin, algoritma sonuçları üzerinde etkili olduğu görülmüştür. TurkishSMS ve UCI SMS Spam veri kümelerinde uygulanan SVM algoritmasında doğruluk açısından farklı sonuçlar elde etmiştir. Ancak her iki veri kümesi için de en az kaçırma oranını veren algoritma olduğu görülmüştür. Her ne kadar kaçırma oranın az olması önemli bir faktör olsa da spam olmayan SMS'leri spam olarak tespit etmekte uygun değildir. Bu nedenle veri kümeleri için uygun algoritmaların seçiminde doğruluk ve kaçırma oranı arasındaki denge göz önüne alınmıştır. Bu çalışmada optimum sonuçların elde edilmesi amaçlanmış ve SMS spam tespiti gerçekleştirilmiştir.

KAYNAKÇA

[1] Shafi'I, M. A., Latiff, M. S. A., Chiroma, H., Osho, O., Abdul-Salaam, G., Abubakar, A. I., Herawan, T. 2017. A review on mobile SMS spam filtering techniques. IEEE Access, 5, 15650-15666.

[2] Joe, I., Shim, H. 2010. An SMS spam filtering system using support vector machine. In International Conference on Future Generation Information Technology, Springer, Berlin, Heidelberg, 577-584.

[3] Liu, J. Y., Zhao, Y. H., Zhang, Z. X., Wang, Y. H., Yuan, X. M., Hu, L., Dong, Z. J. 2012. Spam short messages detection via mining social networks. Journal of Computer Science and Technology, 27(3), 506-514.

[4] Nuruzzaman, M. T., Lee, C., Choi, D. 2011. Independent and personal SMS spam filtering. In 2011 IEEE 11th International Conference on Computer and Information Technology, IEEE, 429-435.

[5] Almeida, T. A., Hidalgo, J. M. G., Yamakami, A. 2011. Contributions to the study of SMS spam filtering: new collection and results. In Proceedings of the 11th ACM symposium on Document engineering, ACM, 259-262.

[6] Mathew, K., Issac, B. 2011. Intelligent spam classification for mobile text message. In Proceedings of 2011 International Conference on Computer Science and Network Technology, IEEE, Vol. 1, 101-105.

[7] Uysal, A. K., Gunal, S., Ergin, S., Gunal, E. S. 2012. A novel framework for SMS spam filtering. In 2012 International Symposium on Innovations in Intelligent Systems and Applications, IEEE, 1-4.

[8] Uysal, A. K., Gunal, S., Ergin, S., Gunal, E. S. 2013. The impact of feature extraction and selection on SMS spam filtering. Elektronika ir Elektrotechnika, 19(5), 67-73.

[9] Kim, S. E., Jo, J. T., Choi, S. H. 2015. SMS spam filterinig using keyword frequency ratio. International Journal of Security and Its Applications, 9(1), 329-336.

[10] Modupe, A., Olugbara, O. O. Ojo, S. O. 2014. Filtering of mobile short messaging service communication using latent Dirichlet allocation with social network analysis. In Transactions on Engineering Technologies, Springer, Dordrecht, 671-686.

[11] Alzahrani, A. J., Ghorbani, A. A. 2014. SMS mobile botnet detection using a multi-agent system: research in progress. In Proceedings of the 1st International Workshop on Agents and CyberSecurity, ACM, 2.

[12] Yadav, K., Saha, S. K., Kumaraguru, P., Kumra, R. 2012. Take control of your SMSes: Designing an usable spam SMS filtering system. In 2012 IEEE 13th International Conference on Mobile Data Management, IEEE, 352-355.

[13] Foozy, C. F. M., Ahmad, R., Abdollah, M. F., Wen, C. C. 2017. A Comparative Study with RapidMiner and WEKA Tools over some Classification Techniques for SMS Spam. In IOP Conference Series: Materials Science and Engineering, IOP Publishing, Vol. 226, No. 1, 012100.

[14] Bozan, Y. S., Çoban, Ö., Özyer, G. T., Özyer, B. 2015. SMS spam filtering based on text classification and expert system. In 2015 23nd Signal Processing and Communications Applications Conference (SIU), IEEE, 2345-2348.

[15] Kawade, D. R., Oza, K. S. 2015. SMS spam classification using WEKA. International Journal of Electronics Communication and Computer Technology, 5, 43-7.

[16] UCI Machine Learning Repository. SMS Spam Collection Data Set. https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection (Erişim Tarihi: 15.03.2019).

[17] Pattern Analysis and Recognition Group. TurkishSMS Dataset. http://ceng.eskisehir.edu.tr/par/ (Erişim Tarihi: 15.03.2019).

[18] Orange. https://orange.biolab.si/ (Erişim Tarihi: 15.03.2019).

[19] Orange 3 Text Mining. Preprocess Text. https://orange3-text.readthedocs.io/en/latest/widgets/preprocesstext.html (Erişim Tarihi: 15.03.2019).