

AFET YÖNETİMİNDE DOĞAL DİL İŞLEME: ACİL DURUM MÜDAHALELERİNİ HIZLANDIRMAK VE İYİLEŞTİRMEK

Tarkan Karaçay¹ Hakan Özbaşaran²

¹Doktora Öğrencisi, Eskişehir Osmangazi Üniversitesi, mail@tarkankaracay.com

²Doç. Dr., Eskişehir Osmangazi Üniversitesi, ozbasaran@ogu.edu.tr

Özet

Bu çalışma, afet yönetimine yönelik önemli bilgilerin derin öğrenme yöntemleri ile metin verisinden çıkarılabilmesini araştırmaktadır. Araştırma, ilgili metinlerdeki konum, talep, neden sonuç ilişkileri, tedarik malzemeleri, kişiler, durumlar ve zaman gibi bilgilerin dönüştürücü ağlar (transformer networks) ile çıkarılmasına odaklanmaktadır. Bunun için adlandırılmış varlık tanıma ve ilişki çıkarımı yöntemleri kullanılmıştır. x.com'dan (eski adı ile Twitter) deprem ile ilgili elde edilen 310 farklı metin verisi etiketlenmiştir. Verinin %70'i eğitim, %15'i doğrulama ve %15'i test için kullanılmıştır. Sonuçlar oldukça umut vericidir. Bu çalışma, daha fazla veri ile eğitildiğinde ve test edildiğinde metriklerin iyileşebileceği bir deneme çalışması olarak değerlendirilebilir.

Anahtar Sözcükler: Afet Yönetimi, Acil Durum Yönetimi, Deprem, Derin Öğrenme, Doğal Dil İşleme

Abstract

This study investigates the extraction of crucial information related to disaster management from textual data using deep learning methods. The research focuses on extracting information such as location, request, cause-and-effect relationships, supply materials, individuals, situations, and time from relevant texts using transformer networks. Named entity recognition and relationship extraction methods were used for this purpose. A total of 310 different textual data related to earthquakes, obtained from x.com (formerly Twitter), were labeled. 70% of the data was used for training, 15% for validation, and 15% for testing. The results are quite promising. This study can be considered as an experimental attempt, with the potential for metrics improvement when trained and tested with more data.

Keywords: Disaster Management, Emergency Management, Earthquake, Deep Learning, Natural Language Processing

Amaç

Afet yönetimi, afetlerin etkilerini azaltmak ve afetten etkilenen bireylerin kurtarılması için izlenen faaliyetlerin ve önlemlerin bir bütünüdür. Afet sonrası insanların en az zarar ile kurtarılması için hızlı tepkilerin alınması ve etkili yardımların yapılabilmesi oldukça önemlidir [1].

Eskiden uzun bir sürede ve kısıtlı bir çevreye yayılan afet haberleri, sosyal medya ile çok kısa bir sürede hemen herkese yayılmaktadır [2]. Günümüzde bir afet meydana geldiğinde bireyler sosyal medya platformlarını kullanarak çok sayıda metin verisi paylaşmaktadırlar [3]. Bu

nedenle sosyal medya afet zamanlarında iletişimde kalmanın ve bilgi aktarmanın etkili ve kolay bir yoldur [4].

Afet yönetimi için önemli ve kritik bilgiler sosyal medya kaynaklarında bulunmaktadır. Bu kaynaklar afetzedelerin isimleri, konumları, durumları, ihtiyaçları gibi bilgileri barındırmaktadır. Ancak bu metinlerin sayıları çok fazladır. İnsan eliyle bu verinin işlenmesi çok zor ve zaman alıcıdır [5].

Derin öğrenme yöntemleri, metin verisinden önemli bilgileri çıkarmak için kullanılabilecek güçlü bir araçtır. Derin öğrenme yöntemleri, metin verisinin karmaşık yapısını ve ilişkilerini öğrenerek bu veriden doğru ve verimli bilgiler çıkarılabilir [6]. Ayrıca, bu yöntemin insan hızıyla karşılaştırıldığında belirgin bir hız avantajı sunabildiği bilinmektedir.

Bu çalışma, doğal dil işleme yöntemleri ile afet yönetimi için önemli bilgilerin metin verisinden çıkarılmasını araştırmaktadır. Çalışmanın amacı dönüştürücü ağlar (transformer networks) kullanılarak metinlerden konum, yardım talebi, tedarik malzemelerinin talebi, kişiler, kişilerin durumları ve zaman gibi bilgilerin metin verisinden çıkarılmasını ve aralarındaki ilişkilerin elde edilmesini sağlamaktır.

Yöntem

Metin verisi elde etmek için x.com'dan (eski adı Twitter) yararlanılmıştır. 6 Şubat 2023'te yaşanan merkez üssü Pazarcık, Kahramanmaraş olan ve 20 Şubat 2023'te yaşanan merkez üssü Yayladağı, Hatay olan depremlerle ilgili çeşitli anahtar kelimeler belirlenmiştir. Bu anahtar kelimeler kullanılarak snsrape [7] isimli Python kütüphanesi ile depremler gerçekleşikten 7 gün sonrasına kadarki sürede paylaşılan çok sayıda gönderi çekilmiş ve depolanmıştır. Depolanan verideki tekrar eden veri silinmiş ve her gönderi bir metin dosyası haline getirilmiştir. İlgili metin verisinden 1400 gönderi tamamen rastgele olarak seçilmiş ve metin dosyaları Label Studio [8] isimli veri etiketleme platformuna aktarılmıştır. Daha sonra bu gönderilerden 310 tanesi etiketlenmiştir.

Etiketlemeler iki aşamadan oluşmaktadır. Bu aşamalardan birincisi adlandırılmış varlık tanıma [9] için yapılan etiketlemedir. İkincisi ilişki çıkarımı [10] için yapılan etiketlemedir.

Bu çalışma kapsamında belirlenen adlandırılmış varlık tanıma etiketleri şunlardır:

- Location (Konum)
- Request (Talep)
- Not (Değil)
- Reason (Neden)
- Activity (Eylem)
- Supply (Tedarik)
- Person (Kişi)
- Status (Durum)
- Time (Zaman)

Çalışma kapsamında oluşturulan ilişki çıkarımı etiketleri ise şunlardır:

- Location – Request
- Request – Reason
- Request – Not
- Request – Time
- Location – Activity
- Activity – Time
- Location – Supply
- Supply – Not
- Supply – Time
- Location – Person
- Person – Status
- Status – Time

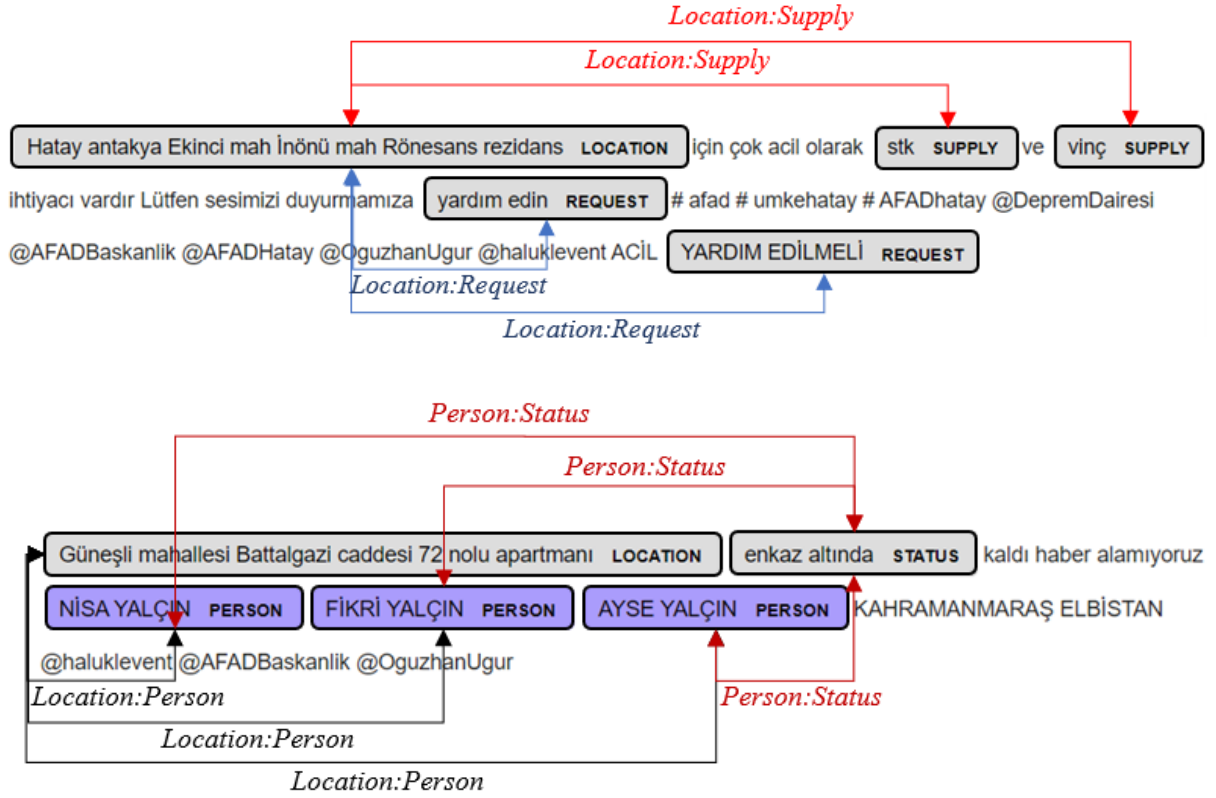
Belirlenen etiketler ile veri etiketlendikten sonra, verinin tamamen rastgele seçilen %70'i eğitim, %15'i doğrulama ve %15'i test için ayrılmıştır. BERTurk [11] isimli dönüştürücü ağ kullanılarak adlandırılmış varlık tanıma eğitimi ve testi yapılmıştır. Sonraki aşamada ilişki çıkarımı algoritmasının eğitimi ve testi yapılmıştır.

Çalışmada BERTurk kullanılmasının nedeni Türkçe veriler ile eğitilmiş bir model olmasıdır. Ayrıca Türkçe metinlerde oldukça iyi performans göstermesidir [12]–[15]. BERTurk modeli HuggingFace [16] platformundan indirilmiş ve SpaCy [17] kullanılarak eğitim ve testler gerçekleştirilmiştir.

Bulgular

BERTurk modeli ile adlandırılmış varlık tanıma başarı oranları test verisinde %66 kesinlik, %76 duyarlılık ve %70.5 F1-skoru olarak hesaplanmıştır.

Şekil 1'de test verisine ait iki farklı metin üzerinde, eğitilmiş BERTurk modelinin çıkardığı adlandırılmış varlıklar ve ilişkiler görülmektedir.



Şekil 1: Eğitilen BERTurk modelinin rastgele seçilen iki test verisindeki adlandırılmış varlık tanıma ve ilişki çıkarımı performansları

Sonuçlar ve Tartışma

Bu çalışma, afet yönetimi alanında metin verisinden önemli bilgilerin çıkarılması için bir derin öğrenme yöntemi olan dönüştürücü ağların kullanılabilirliğini incelemiştir. Bu çalışmanın ana sonuçları şunlardır:

- Çalışma; yeniden eğitilmiş BERTurk modeli aracılığıyla afet yönetimi metin verisinden başarılı bir şekilde konum, talep, neden-sonuç ilişkileri, tedarik malzemeleri, kişiler, durumlar ve zaman gibi önemli bilgilerin çıkarılabileceğini göstermektedir.
- BERTurk, 310 tane kısa metin verisi ile eğitilmesine rağmen adlandırılmış varlık tanıma için test verisinde yüksek bir doğruluk oranı elde etmiş ve %66 kesinlik, %76 duyarlılık ve %70.5 F1-skoru ile önemli bir başarı göstermiştir.
- Metin verisindeki ilişkileri doğru bir şekilde belirlemenin önemi vurgulanmıştır. İlişki çıkarımı, afet yönetimi süreçlerini daha iyi anlamayı sağlayacak ve hızlı tepkilerin planlanmasına yardımcı olacaktır.
- Daha fazla veri ile eğitilmiş modellerin daha yüksek performans sağlayabileceği öngörülmektedir.

Olası bir afet yönetimi senaryosunda, büyük metin havuzlarının doğal dil işleme yöntemleriyle incelenmesi, afet yönetimi ekiplerine hızlı bir şekilde bilgi sağlama potansiyeli taşımaktadır. Bu yaklaşım, afet durumlarında karar alma süreçlerini hızlandırabilir ve afet yönetimi ekiplerinin daha kesin ve etkili müdahalelerde bulunmalarına yardımcı olabilir.

Kaynaklar

- [1]-Erkal, T., Değerliyurt, M. Türkiye’de Afet Yönetimi, Doğu Coğrafya Dergisi, 14, 22, 147–164, 2011.
- [2]-Şahinsoy, K. Ş., Kriz Yönetimi Açısından Geleneksel ve Sosyal Medya, İstanbul Aydın Üniversitesi Dergisi, 9, 4, 1–19, 2017.
- [3]-Kankanamge, N., Yigitcanlar, T., Goonetilleke, A., How engaging are disaster management related social media channels? The case of Australian state emergency organisations”, International Journal of Disaster Risk Reduction, 2020.
- [4]-Ahmed, A., Use of social media in disaster management, International Conference on Information Systems, 2011, ICIS 2011, 5, 4149–4159, Şangay, 2011.
- [5]-Phengsuwan, J., Tejal, S., Thekkumma, N. B., Wen, Z., Sun, R., Pullarkatt, D., Thirugnanam, H., Ramesh, M. V., Morgan, G, James, P., Ranjan, R., Use of social media data in disaster management: A survey, Future Internet, 13, 2, 1–24, 2021.
- [6]-Vishwanath, T., Shirwaikar, R. D., Jaiswal, W. M., Yashaswini, M., Social media data extraction for disaster management aid using deep learning techniques, Remote Sensing Applications: Society and Environment, 30, 2023.
- [7]-URL: <https://github.com/JustAnotherArchivist/snsrape/>, Erişim Tarihi: 06.09.2023.
- [8]-URL: <https://labelstud.io/>, Erişim Tarihi: 06.09.2023.
- [9]-Li, J., Sun, A., Han, J., Li, C., A Survey on Deep Learning for Named Entity Recognition, IEEE Transactions on Knowledge and Data Engineering, 34, 1, 50–70, 2022.
- [10]-Luo, X., Zhou, W., Wang, W., Zhu, Y., Deng, J., Attention-Based Relation Extraction With Bidirectional Gated Recurrent Unit and Highway Network in the Analysis of Geological Data, Special Section on Human-Centered Smart Systems and Technologies, 6, 5705–5715, 2018.
- [11]-Budur, E., Özçelik, R., Gungor, T., Potts, C., Data and Representation for Turkish Natural Language Inference, Conference on Empirical Methods in Natural Language Processing (EMNLP), 8253–8267, Çevrimiçi, 2020.
- [12]-Özberk, A., Çiçekli, I., Offensive Language Detection in Turkish Tweets with Bert Models, 6th International Conference on Computer Science and Engineering, UBMK 2021, 517–521, Nanjing, 2021.
- [13]-Sahinuc, F., Toraman, C., Koc, A., Topic detection based on deep learning language model in Turkish microblogs, 29th IEEE Conference on Signal Processing and Communications Applications, 2021–2024, İstanbul, 2021.
- [14]-Celikten, A., Bulut, H., Turkish medical text classification using BERT, 29th IEEE Conference on Signal Processing and Communications Applications, 2021–2024, İstanbul, 2021.
- [15]-Koksal, A., Ozgur, A., Twitter dataset and evaluation of transformers for Turkish sentiment analysis, 29th IEEE Conference on Signal Processing and Communications Applications, 1–4, İstanbul, 2021.

[16]-URL: <https://huggingface.co/dbmdz/bert-base-turkish-128k-cased>, Erişim Tarihi: 12.11.2023.

[17]-URL: <https://spacy.io/>, Erişim Tarihi: 12.11.2023.