

Kontrole Değer İddiaların Tespiti

1. Özet

Günümüzde sosyal medya kullanımının yaygınlaşmasından ve yalan haberlerin artmasından dolayı ortaya atılan iddiaların teyiti gerekmekte ve teyit edilecek metinlerin doğru bir şekilde bulunması gerekmekte olduğundan dolayı, kontrole değer iddiaların tespiti araştırması yapılmıştır. Çalışmada Türkçe için optimize edilmiş doğal dil işleme yöntemleri, öznitelik olarak POS, NER, BOW ve cümle uzunluğu kullanılmış, bu özniteliklerin farklı kombinasyonları 2 farklı denetimli makine öğrenmesi algoritması ile test edilmiş ve sonuçları karşılaştırılmıştır. Böylece Türkçe dilindeki ilk kontrole değer iddiaların tespiti çalışması yapılmıştır. Sonucunda taban çizgimiz olan ClaimBuster^[20] araştırmasına yakın bir başarı Türkçe için elde edilmiştir.

2. Giriş

Teknolojinin ve internet altyapısının gelişmesiyle dünyanın her tarafında internet kullanımı sürekli olarak artmaktadır. Çoğalan kullanıcı sayısı ile birlikte sosyal medyada her gün milyonlarca paylaşım yapılmaktadır. Yapılan araştırmalara göre bu paylaşımlar çoğu zaman kişisel olmaktadır ve nadiren de bir iddia ortaya atmaktadır. Yüzdelik oranı az olsa da milyonlarca paylaşım içinde bu paylaşımlar göz ardı edilemez. Sosyal medyayı 7'den 70'e herkes kullanabildiği için bilgi kirliliği olması kaçınılmazdır. İnsanlar bir iddia ortaya attığı zaman bu iddia hızla yayılıp toplum üzerinde büyük etkilere sahip olabilir. Bu sebeple iddiaların öncelikle tespit edilmesi ve daha sonra doğru olup olmadığının araştırılması toplum sağlığı açısından oldukça önemlidir. Sosyal medyada bir bilgi dakikalar içinde tüm dünyaya yayılabilir. Bu yüzden tespit ve teyit süresi oldukça kısa olmalıdır. Klasik yöntemlerle yapılan arama, tarama ve araştırma faaliyetleri milyonlarca paylaşımın içerisinde yeterince uygun olmayacağı için bu işlemlerin bir mekanizma yardımıyla otomatikleştirilmesi gerekmektedir.

Kontrole değer iddiaların tespiti araştırması, girdi olarak verilen büyük verilerin içinden iddia ortaya atan cümleleri tespit etmeyi, daha sonra bu verilerin içindeki kontrole değer iddiaların yüzdelik diliminin hesaplamayı kapsamaktadır. Bu araştırma, şu ana kadar yapılan araştırmaların başarı oranları daha iyiye gidebilir mi, şu ana kadar yapılan araştırmaların başarı oranı Türkçe dilinde de elde edilebilir mi, elde edilen başarı farklı veri setlerinde yaklaşık olarak aynı sonuca ulaşılabilir mi gibi sorulara cevap aramıştır.

Bu projede daha önce geliştirilen algoritma ve yöntemler araştırılıp, en yüksek başarı oranına sahip olan metotlar belirlenip, uygulanabilecek ve başarıyı arttırabilecek çözümler kullanılmıştır. Aynı zamanda, daha önce İngilizce veri setleri için geliştirilen algoritmalar Türkçe diline uyarlanabilir mi gibi sorulara cevap aramıştır. Araştırma problemi temel olarak bu konuda yapılmış çalışmaları Türkçe dili için uyarlamak ve üzerine Türkçe için yapılabilecek optimizasyonları eklemek olmuştur. Bu araştırmayı Türkçe dili için yapmak ve algoritma geliştirmek çok önemli olmuştur. Çünkü daha önce Türkçe için bu tarzda yapılmış çalışmalara rastlanılmamıştır.

Yapılan çalışma için 10.000 cümlelik bir veri seti, Türkçe için hazırlanmış bir “dolgu kelimeleri” havuzu, ve yine Türkçe dili için hazırlanmış ve optimize edilmiş ön işleme algoritmaları kullanılmıştır. Yapılan preprocessing işlemleri sırasıyla tokenlere ayırma, normalleştirme, kök çıkarma ve dolgu kelimelerini atma olmuştur. Hazırlanan stop words havuzu ve normalizasyon algoritması açık kaynaklı olarak paylaşılmıştır^[23]. Bu çalışmalar sayesinde Türkçe dili için yapılmış çalışmalar artacak ve daha sonraki çalışmalara yol açacaktır.

Diğer taraftan “Trump bu akşam hamburger yedi.” ve “Trump, İran’a savaş açtı.” cümlelerini değerlendirdiğimizde Trump’ın hamburger yemesi gibi kontrol edilmeye değmeyen haberlerin sayısı çok fazla olduğundan dolayı ve teyit.org^[14] gibi web siteleri bu işlemi manuel olarak uyguladığı için, bu durum iddia kontrolü mekanizmasının gereksiz çaba sarf etmesine ve zaman israfına yol açmaktadır. Yapılan çalışmalar ile bu işlem akış halinde yürütülmüş ve kontrol edilmeye değer cümleler gerektiğinde bu tarz web sitelerine veya başka doğruluk kontrolü yapan hizmetlere sunulabilir ve bu sayede süreç otomatikleştirilebilir.

Kontrole değer iddiaların tespiti araştırması, veri seti olarak eğitim için teyit.org ve buna benzer iddia havuzları içeren siteleri ve test için Twitter sosyal medya sitesini kullanarak, elde edilen veriler üzerinde çeşitli doğal dil işleme yöntemleri ve ön işleme yöntemleri uygulayarak makine öğrenmesi için yukarıda belirtilen gerekli öznitelikleri çıkarmak ve daha sonra eğitilmiş verileri test için oluşturulan veri setinde test etmeyi ve milyonlarca veri içerisinde iddia içerenlerin tespit edilme işlemini kapsamaktadır. Kontrole değer iddiaların tespiti araştırmasının temel problemi, klasik yöntemlerle yapılan arama, tarama ve araştırma faaliyetleri milyonlarca paylaşımın içerisinde yeterince uygun olmayacağı için bu işlemlerin bir mekanizma yardımıyla otomatik hale getirilmesi gerektiği için ortaya çıkmıştır.

Formal bir ifade ile bu çalışmanın amacı, girdi olarak verilen keyfi bir metnin kontrole değer olup olmadığına karar veren bir ikili sınıflandırıcı geliştirmektir.

Bu araştırma daha önce farklı diller için çalışılmış ve belirli sonuçlara ulaşmıştır. Fakat Türkçe dili için yapılan ilk araştırma bu çalışma olmuştur. Bu çalışmanın bilime katkısı Türkçe dili için hazırlanan ilk tespiti değer iddia bulma sistemi olması olmuştur.

3. İlgili Çalışmalar

Check-worthy iddia tespiti alanında yapılan ilk çalışma ClaimBuster’dır^[20]. ClaimBuster projesinde öğrenciler, profesörler ve gazeteciler tarafından işaretlenmiş verileri eğitim verisi olarak kullanılmıştır. Öznitelik olarak TF-IDF ağırlıklarını, POS etiketlerini, varlık türlerini, kelime sayılarını öznitelik olarak kullanmışlardır. Denetimli makine öğrenmesi kullanılarak 4-fold çapraz doğrulama uygulanmıştır. Multinomial Naive Bayes Classifier (NBC), Support Vector Classifier (SVM) ve Random Forest Classifier (RFC) algoritmaları hepsi ile denenerek toplam 12 deneme yapılmış ve sonuçlar karşılaştırılmıştır.

ClaimBuster’ın üzerine ClaimRank^[1] daha fazla öznitelik kullanarak, İngilizce ve Arapça olmak üzere iki dil desteği sunarak, farklı fact-checking organizasyonlarından veri toplayarak her sonucu bu

organizasyonların ayrı ayrı neler düşüneceğini tahmin edecek şekilde hazırlanmıştır. ClaimRank araştırmasında ise iki adet gizli katmanı olan bir sinir ağı kullanılmıştır. Stochastic Gradient Descent (SGD) kullanılarak check-worthy buldukları iddialar arasından sıralama yapılmaktadır ve en olası olanlar en başa gelmektedir. ClaimBuster ve ClaimRank çalışmaları bu alanda yapılan state of the art çalışmalardır.

Tathya çalışması biraz daha kompleks bir çözüm önerisi sunmuştur. Örneğin “They found this.” cümlesindeki they’in kim olduğunu bulmak için önceki cümlelere bakılmaktadır. Öznitelikler için ek olarak varlık geçmişine bakarak bunun bir tartışma mı, tekrarlama mı olduğunu incelenmektedir. Scikit-learn^[19] kullanılarak eğitilmiş doğrusal SVM sınıflandırıcılarını ve çoklu sınıflandırma sistemlerinde verileri kümelemek için k-means kullanılmaktadır.

IPIPAN ekibinin^[24] CheckThat^[31]’e gönderdiği çalışması iddiaların otomatik olarak belirlenmesi ve doğrulanmasına odaklanmıştır. Politika tartışmalarındaki iddiaların kontrol edilebilirliğinin tespit edilmesi ve derecelendirilmesiyle ilgilendiler. Cümlelerin vektör gösterimi, cümlenin öğeleri (POS), NER ve duyarlılık puanları gibi özellikleri kullanarak lojistik regresyona dayalı bir sınıflandırıcı önermişlerdir.

Zuo ve ark.^[25] yaptığı çalışmada politik görüşmelerde ve konuşmalarda yapılan iddiaları tanımlamak için denetimli makine öğrenimi ile basit sezgisel buluşları birleştiren ve bunları “çekiciliği” açısından sıralamak için bir mekanizma sağlayan karma bir yaklaşım geliştirilmiştir.

Gencheva ve ark.^[26] verilen bir belgede hangi iddiaların en değerli olduğunu ve gerçeklerin kontrolü için öncelikli olması gerektiğini otomatik olarak belirleme sorununu ele almışlardır.

Hansen ve ark.^[27] sahte haberler gibi yanlış bilgileri tespit etmektedirler. Rule-based özellikler kullanarak ve sözdizimsel bağımlılıklarını temsil eden bir sinirsel denetim değeri cümlesi sıralama modelini kullanmışlardır.

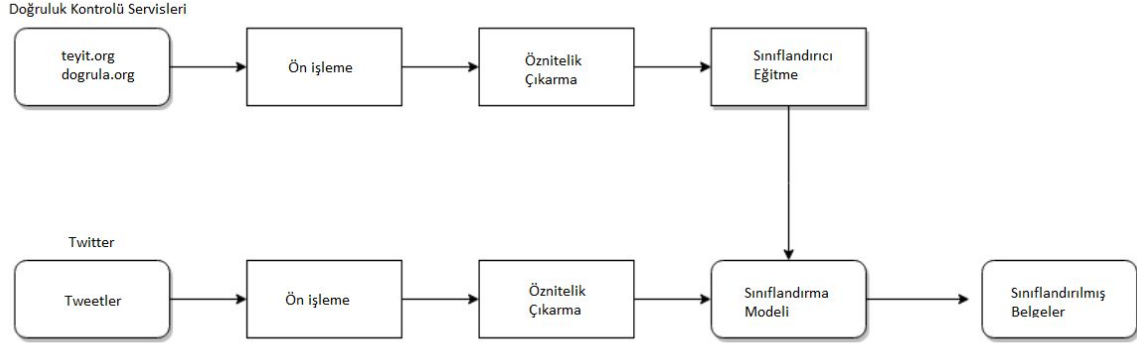
M. Kutlu ve ark.^[28], kontrol değerlerine göre iddiaları sıralamak için, NER ve duyarlılık analizi gibi doğal dil işleme yöntemleriyle elde edilen öznitelikleri içeren bir learning-to-rank yaklaşımı sunmuşlardır. Doğruluk tahmini için, potansiyel olarak alakalı web sayfalarını almak ve doğruluğu tahmin etmek için bu sayfaların ilgili bölümlerinden özellikleri çıkarmak için harici bir web arama motoru kullanmışlardır.

Rashkin ve ark.^[29] siyasi gerçeklerin kontrolü ve sahte haber tespiti bağlamında haber medyasının dili üzerine analitik bir çalışma yapmışlardır. Araştırma deneyleri stilistik işaretlerin metnin doğruluğunu belirlemeye yardımcı olabileceğini göstermektedir.

Thorne ve ark.^[30] doğal dil işleme, makine öğrenmesi, bilgi sunumu, veritabanları ve gazetecilik dahil olmak üzere çeşitli alanlarda bu çalışmayı yürütmüştür. Ayrıca doğal dil işleme ve ilgili disiplinlerden kaynaklanan araştırmaları otomatik olarak inceleyerek görev formülasyonlarını ve metodolojileri makale ve yazarlar arasında birleştirerek araştırmışlardır.

Bizim önerilen yöntemimiz burada bahsedilen çalışmalar gibi, NLP sürecinden geçmiş metinlere ait NER, POS, Length ve BOW özniteliklerinin makine öğrenmesi yöntemleri ile sınıflandırılmasından oluşmaktadır.

4. Önerilen Çözüm



Şekil 1. Sistemin genel yapısı

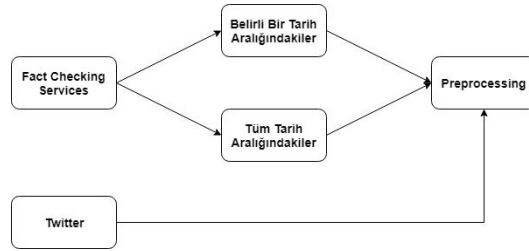
Araştırma problemi temel olarak veri toplama, ön işleme, öznitelik çıkarma, sınıflandırma ve değerlendirme alt problemlerine ayrıldı. Bu alt alanların her biri proje sonucunun başarısını etkileyecek temel çözümlerden oluşmaktadır.

Ön işleme problemi sırasıyla normalleştirme, tokenlara ayırma, kök bulma, dolgu kelimelerini atma adımlarından geçiriliyor. Ön işlemlerden geçmiş veriler kullanılarak gerekli öznitelikler çıkarılıyor. Bu öznitelikler cümle uzunluğu, kelime torbası, varlık tanıma ve cümlelerin öğeleridir.

İddia kontrolü servislerinden elde edilen veri kümesi sınıflandırıcıyı eğitme amacıyla kullanılıyor. Oluşan bu sınıflandırıcı model de girdi olarak verilen tweetleri, ‘kontrole değer bir iddia’ ve ‘kontrole değer bir iddia değil’ olmak üzere sınıflandırılıyor. Bu sınıflandırma için belirtilen denetimli makine öğrenmesi metotları kullanılıyor.

Araştırma, Twitter’daki belirli bir tarih aralığındaki tweetleri ve teyit.org gibi iddia kontrolü yapan servislere ait internet sitelerindeki verileri çekerek Türkçe veri kümesi oluşturma; ön işleme olarak tüm verileri tokenlara ayırma, normalleştirme, kök çıkarma, dolgu kelimelerini atma; ve nihai olarak da makine öğrenmesi yöntemlerinden Multinomial Naive Bayes Classifier (NBC), Support Vector Classifier (SVM) ve Random Forest Classifier gibi metodları kullanarak ikili sınıflandırıcı geliştirme şeklinde adımlara ayrılmaktadır.

4.1 Veri Toplama

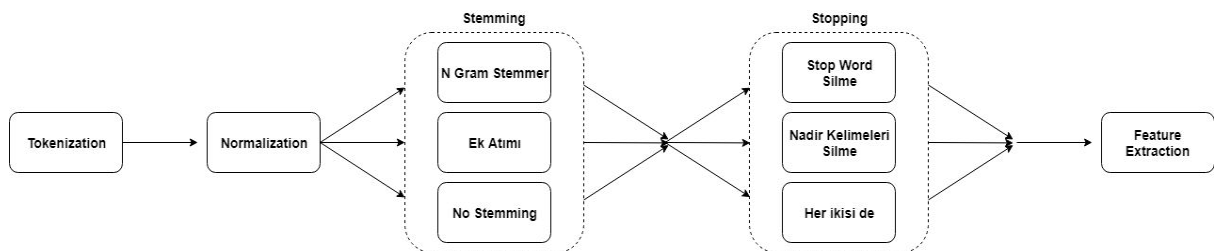


Şekil 2. Veri toplama akışı

Bu çalışmada iddia verileri ve tweetler olmak üzere iki farklı veri seti mevcuttur. Bu veriler tamamen Türkçe metinlerden oluşmaktadır. İddia verileri makine öğrenmesi modelinin eğitilmesinde kullanılmıştır ve Twitter verileri ise bu model tarafından sınıflandırılan test verileridir. Bu çalışmada teyit.org gibi servislerin doğruluk kontrolü yapmış olduğu iddiaların check-worthy olduğu varsayılmıştır.

İddia verilerinin toplanması için 6 farklı fact-checking sitesi Selenium^[9] kullanılarak crawl edildi. Toplanan tüm bu verilerin formatı düzenlenerek “tarih, iddia, doğruluk, kaynak” şeklinde aynı hale getirildi. Yani iddia tarihi, iddianın metni, iddianın doğru olup olmadığı ve iddia kontrolünün yapıldığı kaynak bilgileri tutuldu. ClaimRank araştırmasında bu çalışma İngilizce ve Arapça olarak yapılmıştır ve PolitiFact^[32], FactCheck^[33], ABC^[34], CNN^[35], NPR^[36], NYT^[37], Chicago Tribune^[38], The Guardian^[39], ve Washington Post^[40] olarak 9 saygın iddia doğrulama servisindeki gerçek belirtiler kullanılmıştır^[1]. Biz ise iddia verilerini teyit.org, dogrula.org^[10], dogrulukpayi.com^[11] ve gununyalanlari.com^[12] web sitelerinden elde ettik. Bu veriler düzenli hale getirildi ve öznitelik çıkarma yapılabilecek hale getirildi. Bu sitelerden hem belirlenen tarih aralığındaki verileri hem de sitedeki tüm verileri kullanarak iki farklı veri kümesi oluşturuldu. Bu iki veri kümesinden hangisinin daha başarılı olduğu parametrik olarak araştırma sonucunda incelendi. Tarih aralığı kısa olan veri kümesinin konu özniteliğinin önemini arttırdığı düşünülmektedir.

4.2 Ön İşleme



Şekil 3. Ön işleme yapısı

Elde edilen tüm veriler ön işleme aşamasından geçirildi. Bunun için öncelikle veri kümelerindeki tüm metinler tokenize edildi. Daha sonra sosyal medyada yazım hataları çok fazla olduğu için tüm kelimeler, Zemberek^[5] kütüphanesi tarafından sunulan algoritmalarla normalizasyon işleminden

geçirildi. Her bir veri boru hattı şeklinde tek tek bütün işlemlerden geçirildi. Bu boru hattı, özellikle ön işleme konusunda farklı yaklaşımların birden test edilebilmesi için parametrik olarak tasarlandı.

ClaimBuster araştırmasında stemming veya dolgu kelimelerin kaldırılması işlemleri uygulanmamış, yalnızca yeterince az geçen kelimeler kaldırılmıştır^[2]. Buna karşılık diğer birçok çalışmada ise ön işlem olarak stemming ve stop-words kaldırma işlemlerinin yapıldığı görülmektedir^{[1][3]}. Kelimenin başından ilk n karakteri alıp geri kalan kısmını atma yöntemi ise Türkçe için uygun olabilecek yaklaşık kök bulma yöntemi olarak önerilmiştir^[4].

4.3 Öznitelik Çıkarma

Bu çalışmada yeni öznitelikler ortaya konulmamış ancak geçmişte farklı araştırmacılar^{[1][2][3]} tarafından ortaya konulan öznitelikler kullanılmıştır.

Metin Uzunluğu: Metindeki toplam token sayısıdır. Metin üzerinde tokenization işlemi uygulandıktan sonra tokenlar sayılarak bu değer elde edildi.

Kelime Torbası (BOW): Training veri kümesindeki metinlerde geçen kelimeler öznitelik olarak kullanıldı. Bu kelimeler ön işleme kısmında anlatılan işlemlere tabidir.

Cümlelerin Öğeleri (POS): Cümleleri öğelerine ayırmak için açık kaynak Turkish POS Tagger^[6] isimli kütüphane kullanılacak ve işlem sonucunda hesaplanan etiket vektörü bir öznitelik olarak kullanıldı.

Varlık Türü: Zemberek tarafından sunulan varlık türü tanımlayıcıyı kullanarak metinlerdeki terimlerin tipleri tespit edilip öznitelik olarak kullanılıyor. Bu özellik n-boyutlu ikili bir vektör olarak temsil edildi.

5. Deneysel Değerlendirme

Sınıflandırma yöntemi olarak Naeemul Hassan tarafından önerildiği gibi^[2], denetimli öğrenme yöntemlerinden Multinomial Naive Bayes Classifer (NBC) ve Support Vector Classifer (SVM) metodları kullanıldı. Bu sınıflandırıcıların performansı precision, recall ve f-measure yönünden değerlendirildi.

Değerlendirme sürecinde bütün öznitelikler tek bir şekilde kullanılmayıp ve özniteliklerin farklı kombinasyonları test edilip başarı oranı araştırıldı.

Bu araştırmayı değerlendirirken, “*Hangi öznitelik kombinasyonları başarıyı maksimum seviyeye çıkarır?*” sorusuna, “*Ön işleme aşamasındaki farklı yöntem ve tercihlerin başarıya etkisi nedir?*” sorusuna ve “*Denetimli makine öğrenmesi sınıflandırma yöntemlerinden hangileri daha başarılı sonuç verir?*” sorusuna cevap aranmaktadır.

Projede denetimli makine öğrenmesi yöntemi kullanıldı. Bunun için 4 adet öznitelik çıkarıldı. Bu öznitelikler metin uzunluğu (length), cümlelerin öğeleri (POS), varlık tanıma (NER) ve kelime torbasıdır (BOW). Her biri bu makalede açıklanan çeşitli yöntemler kullanılarak elde edildi. Bu özniteliklerin projeye katkı sağlayıp sağlamadığı ve katkısının ne ölçüde olduğu bilinmediğinden

dolayı araştırılması gerekmektedir. Proje başarısını maksimum seviyeye çıkaran öznitelik kümesini bulmak için çeşitli kombinasyonlar Naive Bayes Classifier (NBC) ve Linear Support Vector Machine (SVM) ile denenmiştir.

5.1. Deneysel Kurulum

Sistemin temel yapısı Python ile yazılmış bir boru hattından oluşmaktadır. Boru hattı sırasıyla veri toplama, ön işleme, öznitelik çıkarma ve sınıflandırma aşamalarından oluşmaktadır.

Eğitim veri seti için iddia doğrulama web siteleri kullanıldı. Toplam 4 adet web sitesi Selenium^[9] kullanılarak crawl edildi. Bu sitedeki veriler iddia olarak etiketlenmiştir. İddia olmayan metinler Twitter’den elle etiketlenerek elde edilmiştir. Elde edilen eğitim veri setinde 2400 adet iddia olan, 7000 adet iddia olmayan metin bulunmaktadır. Eğitim setinde tarih aralığı sınırı bulunmamaktadır, ancak tarihleriyle birlikte tutulduğu için tarih filtrelemesi yapılabilir. Sitelerden elde edilen verilerde iddia tarihi, iddianın metni, iddianın doğru olup olmadığı ve iddianın yer aldığı teyit sitesi bilgisi tutulmaktadır. Bu verilerin hepsinin iddia olduğu varsayılmaktadır. Bu aşamada test için yalnızca rastgele seçilmiş yaklaşık 500 tweet kullanılmıştır.

teyit.org	998
dogrula.org	583
dogrulukpayi.com	744
gununyanlari.com	80

Şekil 4. Kaynaklara göre iddia verilerinin sayısı

	Check-worthy olan	Check-worthy olmayan
Eğitim	2400	7000
Test	100	400

Şekil 5. Eğitim ve test verilerinin dağılımı

Ön işleme kısmı normalleştirme, tokenlara ayırma, dolgu kelimelerini atma ve kök çıkarma olarak 4 bölümden oluşmaktadır. Normalleştirme ve kök çıkarma için Zemberek^[15] kütüphanesi, tokenlara ayırma ve dolgu kelimelerini atma için NLTK^[16] kütüphanesi kullanılmıştır. Ayrıca dolgu kelimeleri veri setinin optimize edilmesi amacıyla elle genişletilmiştir.

Metin uzunluğu, cümlelerin öğeleri, varlık tanımı, duygu analizi, konu çıkarma ve kelime torbası olarak 4 adet öznitelik çıkarılmıştır. Metin uzunluğu, kelime torbası ve cümlelerin öğelerini bulmak için bir kütüphane kullanılmadı. Cümlelerin öğelerini bulmak için İTÜ POS Tagging^[17] veri seti kullanılmıştır. Varlık tanıma (NER) için Zemberek kütüphanesi kullanılmıştır.

Makine öğrenmesi kütüphanesi olarak içinde NBC, SVM, RFC gibi denetimli makine öğrenmesi yöntemlerinin hazır bulunduğu Sklearn^[19] kütüphanesi kullanıldı. Bu yöntemler kullanılarak 4-fold çapraz doğrulama uygulandı ve precision, recall ve f measure hesaplandı. Özniteliklerin 4 farklı kombinasyonu ayrı ayrı test edildi. Bu kombinasyonlar “BOW”, “BOW + POS”, “BOW + POS + NER”, “BOW + POS + NER + Length” şeklindedir.

Bu konuda Türkçe olarak yapılan çalışmalara rastlanmadı. Bu sebeple taban çizgisi olarak İngilizce dili için yapılan ClaimBuster araştırması seçildi. ClaimBuster, İngilizce dili için hazırlandığından dolayı, bu araştırmayla karşılaştırmak için araştırmaları aynı veri seti üzerinde çalıştırmak gerektiğinden dolayı tam olarak bir kıyaslama yapılamayacağından, bu araştırmanın başarısını ölçmek için elde edilen test veri seti Google Translate^[22] kullanarak İngilizce'ye tercüme edildi. 6000 adet iddia olan ve 2000 adet iddia olmayan metin tercüme edilerek veri seti hazırlandı. Bu veri seti ClaimBuster API^[21] üzerinde denenerek sonuçlar elde edildi. Test için hazırlanan eğitim seti aynı zamanda bu araştırmanın sonucunda ortaya çıkan boru hattına verilerek sonuçlar alındı ve taban çizgisi karşılaştırması yapıldı.

6. Sonuçlar ve Gelecek Çalışmalar

Aşağıda Naive Bayes Classifier (NBC) ve Support Vector Machine (SVM) metodlarına ait Metin Uzunluğu (Length), Varlık Tanıma (NER), Cümlelerin Öğeleri (POS), Kelime Torbası (BOW) özniteliklerinin farklı kombinasyonları ile yapılan test sonuçları precision, recall ve f1 puanı olarak görülmektedir.

	precision (0)	recall (0)	f-measure (0)	precision (1)	recall (1)	f-measure (1)	accuracy
NBC (Length + NER + POS)	0,90	0,96	0,93	0,31	0,15	0,20	0,86
NBC (BOW)	0,89	0,92	0,91	0,16	0,11	0,13	0,83
NBC (BOW + Length + NER + POS)	0,89	0,92	0,91	0,16	0,11	0,13	0,83
SVM (Length + POW + NER)	0,89	1,00	0,94	1,00	0,04	0,07	0,89
SVM (BOW + Length)	0,89	0,98	0,93	0,38	0,11	0,17	0,88
SVM (BOW)	0,89	0,98	0,93	0,38	0,11	0,17	0,88

Şekil 6. Deney sonuçları

Yapılan deneylerde Türkçe kaynak yetersizliği ve zaman kısıtından dolayı duygu analizi (sentiment analysis) ve konu çıkarma (topic extraction) kullanılmayıp, kelime torbası (BOW), metin uzunluğu (Length), cümlelerin öğeleri (POS) ve varlık tanıma (NER) öznitelikleri; Naive Bayes Classifier (NBC) ve Linear Support Vector Machine (SVM) metodları üzerinde denenmiştir. Bu aşamada, eğitim (training) hızının yavaş olması sebebiyle Random Forest Classifier (RFC) üzerinde deney yapılmamıştır.

Yapılan deneyler, toplamda 8000 tweet'ten oluşan bir eğitim veri seti ve 300 tweet'ten oluşan bir test veri seti üzerinde yürütüldü. Bu veri kümeleri üzerinde daha önce bahsedildiği şekliyle standart ön işleme uygulandı. Farklı ön işleme yöntemlerinin ve parametrelerinin başarıya olan etkisi bu aşamada test edilmedi. Araştırmanın ilerleyen safhalarında bu soruya yanıt bulmak amacıyla gerekli deneyler yapılacaktır.

Yapılan deneylerde kelime torbasının diğer özniteliklere kıyasla büyük bir ağırlığı olduğu görülüyor. Bu sebeple hem NBC hem de SVM yöntemlerinde, öznitelik kombinasyonu olarak kelime torbasının tek başına kullanılmasıyla diğer öznitelikler ile beraber kullanılması arasında bir fark görünmüyor. Ancak deney sonuçlarından elde edilen ilginç bir sonuç olarak, kelime torbası özniteliği ile diğer üç özniteliğin kombinasyonu karşılaştırıldığında, kelime torbası özniteliğinin kullanılmasının, sistemin performansını precision değerinde düşürdüğü ve recall değerinde ise belirgin bir etki oluşturmadığı

görülüyor. Bir başka sonuç olarak da SVM yönteminin NBC yönteminden daha yüksek bir başarı sunduğu precision değerindeki farktan anlaşılmaktadır.

Test veri kümesi İngilizceye tercüme edilip taban çizgimiz olan ClaimBuster tarafından sınıflandırıldığında aşağıdaki sonuçlar elde edilmiştir:

	precision	recall	f1-score
0	0,91	0,78	0,84
1	0,18	0,37	0,24
accuracy			0,73
macro avg	0,54	0,57	0,54
weighted avg	0,82	0,73	0,77

Şekil 7. ClaimBuster sonuçları

Sonuçlar taban çizgimiz ile karşılaştırıldığında kontrole değer olmayan iddiaların tespitinde precision bazında yaklaşık aynı başarının sağlandığı, recall bazında ise taban çizgimizden daha yüksek bir başarı elde edildiği; kontrole değer iddiaların tespitinde ise precision bazında kendi sistemimizin, recall bazında ise ClaimBuster'ın daha başarılı olduğu görülüyor. Orijinal veri kümesi İngilizceye tercüme edilirken meydana gelmesi muhtemel olan hatalardan ve bu sınıflandırıcının ABD politikalarına yönelik veri kümesi ile eğitilmiş olmasından kaynaklı olarak tam anlamıyla bir kıyaslama yapılması mümkün olmasa da, genel olarak bu projede taban çizgimiz olan ClaimBuster ile benzer bir başarı elde edildiği sonucuna varılabilir.

Bundan sonra yapılabilecek çalışmalar arasında en öncelikli olan konulardan birisi önceden eğitilmiş model üzerinde Almanca, Fransızca vb. farklı dillerdeki veri kümelerinin test edilmesi olabilir. Bu şekilde ön işleme, öznitelik çıkarımı ve kültürel farklılıkların iddia tespiti üzerindeki sonucu test edilebilir.

Şu anki çalışmamızı geliştirmek adına, LDA ile topic detection yapılarak word2vec kullanılabilir. Retweet miktarının da sosyal medyada önemli olduğunu düşünüyoruz. Bu yüzden veri kümesini retweet gibi ekstra bilgiler barındıracak şekilde tekrar çekmeyi planlıyoruz. Ek öznitelik olarak duygu analizini, ön işlemenin farklı varyasyonlarını kullanmayı ve NER tagging, POS tagging setimizi geliştirmeyi düşünüyoruz. BOW özniteliklerini tf-idf olarak ağırlıklandırmayı planlıyoruz.

7. Referanslar

- 1) Israa Jaradat et al. 2018. ClaimRank: Detecting Check-Worthy Claims in Arabic and English. Qatar Computing Research Institute, HBKU, Qatar
- 2) Naeemul Hassan et al. 2017. Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster.
- 3) Ayush Patwari et al. 2017. Tathya: A Multi-Classifer System for Detecting Check-Worthy Statements in Political Debates
- 4) H. Sever, Y. Tonta, Y. Bitirim. 2002. Information retrieval effectiveness of Turkish search engines. Lecture Notes in Computer Science, 2457, 93-103.
- 5) Zemberek. <https://github.com/ahmetaa/zemberek-nlp>
- 6) Turkish POS Tagger. <https://github.com/onuryilmaz/turkish-pos-tagger>
- 7) Gensim. <https://github.com/RaRe-Technologies/gensim>
- 8) Snowball. <https://github.com/snowballstem/snowball>
- 9) <https://www.seleniumhq.org/>
- 10) <https://dogrula.org/>
- 11) <https://www.dogrulukpayi.com/>
- 12) <https://gununyalanlari.com/>
- 13) <https://www.malumatfurus.org>
- 14) <https://teyit.org/>
- 15) <https://github.com/ahmetaa/zemberek-nlp>
- 16) <https://www.nltk.org/>
- 17) <http://tools.nlp.itu.edu.tr/Datasets>
- 18) <https://radimrehurek.com/gensim/>
- 19) <https://scikit-learn.org/stable/>
- 20) Naeemul Hassan et al. 2017. Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster.
- 21) <https://idir.uta.edu/claimbuster/>
- 22) <https://translate.google.com/>
- 23) <https://github.com/doruksahin/KontrolDegerMi>
- 24) Jakub Gasiot, Piotr Przybyla. The IPIAN Team Participation in the Check-Worthiness Task of the CLEF2019 CheckThat! Lab. CLEF (Working Notes) 2019
- 25) Banerjee, Ritwik & Zuo, Chaoyuan & Karakas, Ayla. (2018). A Hybrid Recognition System for Check-worthy Claims Using Heuristics and Supervised Learning.
- 26) Gencheva, Pepa & Nakov, Preslav & Marquez Llu'is & Barron-Cede'no, Alberto & Koychev, Ivan. Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017
- 27) Casper Hansen, Christian Hansen, Stephen Alstrup, Jakob Grue Simonsen, and Christina Lioma. 2019. Neural Check-Worthiness Ranking with Weak Supervision: Finding Sentences for Fact-Checking. In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19), Ling Liu and Ryen White (Eds.). ACM, New York, NY, USA, 994-1000. DOI: <https://doi.org/10.1145/3308560.3316736>
- 28) Khaled Yasser, Mucahid Kutlu, Tamer Elsayed. bigIR at CLEF 2018: Detection and Verification of Check-Worthy Political Claims. CLEF (Working Notes) 2018

- 29) Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. Truth of varying shades: Analyzing language in fake news and political factchecking. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2931–2937, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1317
- 30) James Thorne and Andreas Vlachos. Automated Fact Checking: Task formulations, methods and future directions. CoRR, abs/1806.07687, 2018
- 31) http://clef2019.clef-initiative.eu/index.php?page=Pages/labs_info.html
- 32) <https://www.politifact.com>
- 33) <https://www.factcheck.org>
- 34) <https://abcnews.go.com>
- 35) <https://edition.cnn.com>
- 36) <https://www.npr.org>
- 37) <https://www.nytimes.com>
- 38) <https://www.chicagotribune.com>
- 39) <https://www.theguardian.com/international>
- 40) <https://www.washingtonpost.com/>