

An Artificial Intelligence Application in Health Developed on Covid-19 Documents

Murat Aydogan (Corresponding author)
Genc Vocational School, Bingol University
12000, Bingol, Turkey
E-mail: maydogan@bingol.edu.tr

Abdullah Sener
Genc Vocational School, Bingol University
12000, Bingol, Turkey
E-mail: asener@bingol.edu.tr

Abstract

With the developments in computer science, the concept of artificial intelligence has appeared more frequently in recent years. The concept of artificial intelligence, which is basically defined as computers (machines) thinking like people and making decisions, has become very popular today. Artificial intelligence is used in many fields, especially computer science, education, law, trade, tourism and economy. The health sector is one of these areas. The importance of the applications developed in health sciences has emerged once again, especially during the pandemic process. The development of systems that help reduce the workload of healthcare professionals and make decisions by processing medical data is also an important and real problem that can be solved with artificial intelligence. In this study, natural language processing which is one of the main study subjects of artificial intelligence, has been developed a system that automatically determines the concepts such as disease, medication and treatment on medical data with artificial intelligence by the system. During the experimental studies, it was observed that 91% accurate estimation was made with the model developed. For this study, a Turkish dataset was created by scanning medical articles and studies related to Covid-19 disease.

Keywords: Artificial Intelligence, Turkish natural language processing, Name entity recognition, Covid-19

DOI: 10.7176/JHMN/75-08

Covid-19 Dokümanları Üzerinde Geliştirilen Sağlıkta Bir Yapay Zekâ Uygulaması

Özet

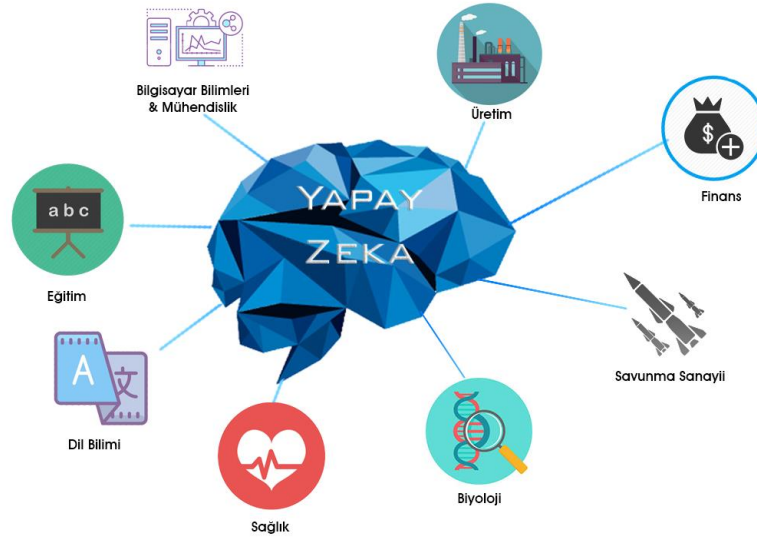
Bilgisayar bilimlerinde yaşanan gelişmeler ile birlikte son yıllarda yapay zekâ kavramı daha sık karşımıza çıkmaktadır. Temelde, bilgisayarların (makinelere) insanlar gibi düşünüp karar vermeleri olarak tanımlanan yapay zekâ kavramı günümüzde oldukça popüler hale gelmiştir. Başta bilgisayar bilimleri olmak üzere, eğitim, hukuk, ticaret, turizm, ekonomi olmak üzere birçok alanda yapay zekâdan faydalanılmaktadır. Sağlık sektörü de bu alanlardan biridir. Özellikle içinde bulunduğumuz pandemi sürecinde sağlık bilimlerinde geliştirilen uygulamaların önemi bir kez daha ortaya çıkmıştır. Tıbbi verilerin işlenerek sağlık çalışanlarının iş yükünün azaltılması ve karar verilmesine yardımcı sistemlerin geliştirilmesi de yapay zekâ ile çözülebilecek önemli ve gerçek bir problemdir. Bu çalışmada, yapay zekânın temel çalışma konularından biri olan doğal dil işleme ile tıbbi dokümanlar üzerinde hastalık, ilaç ve tedavi gibi kavramların sistem tarafından yapay zekâ ile otomatik olarak tespit edilen bir sistem geliştirilmiştir. Deneysel çalışmalar sırasında, geliştirilen model ile % 91 oranında doğru tahmin yapıldığı görülmüştür. Bu çalışma için Covid-19 hastalığı ile alakalı makale ve çalışmalar taranarak Türkçe bir veri seti üretilmiştir.

Anahtar Kelimeler: Yapay zeka, Türkçe doğal dil işleme, Varlık ismi tanıma, Covid-19

1. Giriş

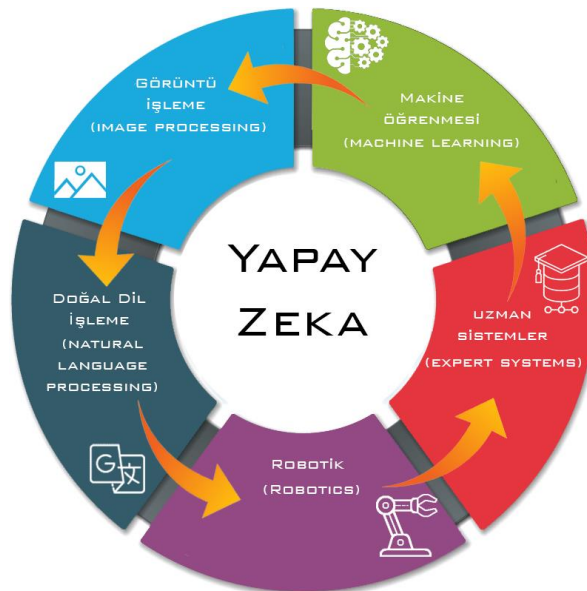
Yapay zekâ; düşünme, akıl yürütme, algılama, sonuç çıkarma gibi insana özgü bu özelliklerin analiz edilerek makinelere kazandırılması ve insan gibi düşünebilen, yorumlayabilen, karar verebilen bilgisayar yazılımlarının geliştirilmesidir (Görz, 2002).

Yapay zekânın kullanım alanlarına bakıldığında; Sohbet robotları, yapay zekâ yardımıyla müşteri sorunlarının çözümünde hem daha hızlı hem de daha anlamlı cevaplar verilmesinde kullanılmaktadır. Akıllı asistanlar, veri kümelerindeki önemli bilgilerin çekilmesinde yapay zekâyı kullanarak zamanlamayı iyileştirmektedir. Öneri motorları, kullanıcıların daha önceden deneyimledikleri izleme alışkanlıklarına uygun TV programları önerebilir. Şekil 1’de yapay zekânın kullanım alanlarından bazıları verilmiştir. (Bozüyük ve ark.,2005). Şekil 1’de kullanım alanları verilmiştir.



Şekil 1. Yapay zekânın kullanıldığı alanlar

Yapay zekânın kullanıldığı önemli alanlardan biride sağlık alanıdır. Yapay zekâ teknolojileri sağlık hizmetlerinde başta erken teşhis, tanı koyma, karar verme ve tedavi süreci olmak üzere birçok alanda kullanılmaktadır. (Büyükgöze ve Dereli, 2019). Yapay zekâ teknolojileri geliştirilirken kullanılan farklı çalışma alanları vardır.



Şekil 2. Yapay zekâ alt dalları

Yapay zekâ teknolojileri ile Şekil 2’de yer alan alt dalları kullanılarak birçok alanda olduğu gibi sağlık alanında önemli uygulamalar geliştirilmiştir. Bu çalışmalardan bazıları ve kullandıkları yapay zekâ teknolojisi Tablo 1’de görülmektedir (Macdonald,2016), (Johnson, 2018), (Markoff, 2016), (Spencer, 2018), (Webster, 2018), (Lock, 2018).

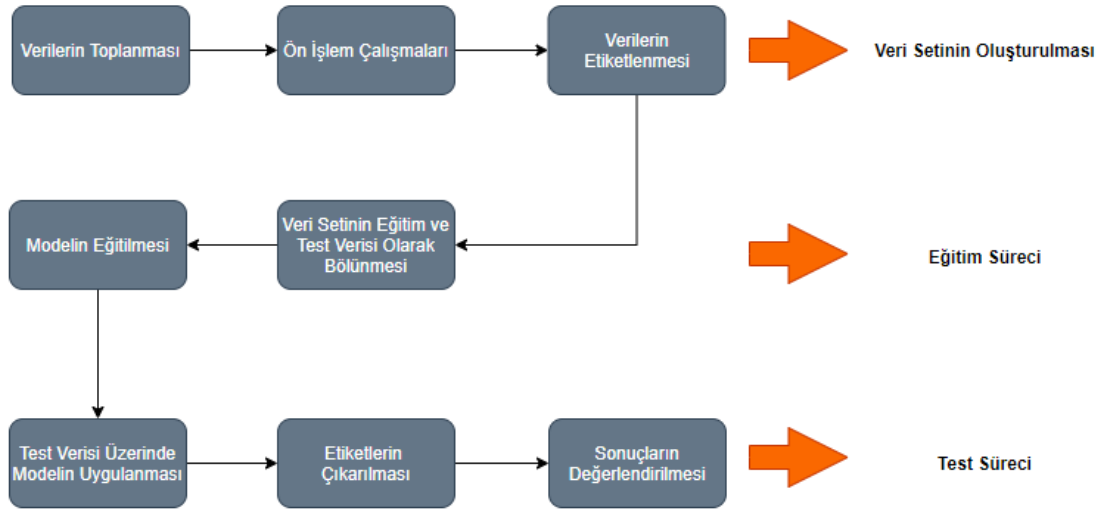
Tablo1: Yapay zekâ teknolojilerinin sağlık alanında kullanım örnekleri

Kullanım örneği veya Kullanıcı Grubu	Kategori	Uygulama Örnekleri	Kullanılan Yapay Zekâ Teknolojisi
Hastalar ve aileleri	*Sağlık İzleme *Fayda/Risk Değerlendirme	*Cihazlar ve giyilebilir cihazlar *Akıllı telefonlar ve tablet uygulamaları, web siteleri	Makine öğrenmesi Doğal Dil İşleme
	Hastalığın Önlenmesi	*Obezite azalması *Diyabetin önlenmesi ve yönetimi *Duygusal ve zihinsel sağlık desteği	Doğal Dil İşleme Makine öğrenmesi
	İlaç Tedavisi Yönetimi	İlaç uyumu	Robotik
	Rehabilitasyon	*Rehabilitasyon uygulamaları ve robotları kullanma	Robotik
Klinik bakım ekipleri	Erken teşhis, tahmin ve teşhis araçları	*Erken kanser tespiti	Makine Öğrenmesi
	Cerrahi işlemler	*Uzaktan kontrollü robotik cerrahi *Yapay zekâ destekli cerrahi yol haritaları	Robotik Makine Öğrenmesi
	Hassas tıp	*Kişiselleştirilmiş kemoterapi tedavisi	Makine Öğrenmesi Pekiştirmeli Öğrenme
Halk sağlığı programı yöneticileri	Risk altındaki bireylerin belirlenmesi	*Sosyal medya kullanarak intihar riski tanımlama	Derin Öğrenme
	Nüfus sağlığı	*Yaşlı bakımı izleme	Yapay zekâ sensörleri
	Nüfus sağlığı	*Hava kirliliği epidemiyolojisi * Su mikrop tespiti	Derin Öğrenme Makine Öğrenmesi
İşletme yöneticileri	Uluslararası Hastalık Sınıflaması,	* Tıbbi kayıtların geri ödeme için otomatik olarak kodlanması	Makine Öğrenmesi Doğal Dil İşleme
	Sahtecilik tespiti	*Sağlık faturalandırma dolandırıcılığı *Lisanssız sağlayıcıların tespiti	Denetimli, denetimsiz ve hibrit makine öğrenimi
Araştırmacılar	Siber güvenlik	*Kişisel sağlık bilgilerinin korunması	Makine Öğrenmesi Doğal Dil İşleme
	Doktor yönetimi	*Hekim yeterliliğinin değerlendirilmesi	Makine Öğrenmesi Doğal Dil İşleme
	Hastalık tahmini	*Yumurtalık kanserinin tahmini	Sinir Ağları
	Keşif	* İlaç keşfi ve tasarımı	Makine Öğrenmesi

2. Materyal ve Metot

Bu çalışmada metodoloji olarak Şekil 3’te görünen yapı uygulanmıştır. İlk olarak Google Scholar platformunda yer alan sağlık alanındaki akademik çalışmalar taranarak veriseti olarak kullanılmak üzere bir doküman (derlem) oluşturulmuştur. Daha sonra üretilen bu veriseti üzerinde ön işleme çalışmaları yapılarak veriler işlenebilir hale getirilmiştir. Bir sonraki adımda ise verilerin etiketlenme işlemi yapılmıştır.

ve veriler eğitim verisi - test verisi olarak ayrılmıştır. Eğitim modeli oluşturularak eğitim süreci işletilmiş daha sonra test süreci uygulanmıştır.



Şekil 3. Geliştirilen modelin genel yapısı

2.1 Materyal

Sağlık alanındaki terimleri içeren Türkçe bir veri seti literatür de yer almamaktadır. Bu nedenle çalışmanın amacı doğrultusunda Türkçe bir veri seti üretilmiş ve bundan sonraki çalışmalara katkı sağlanması hedeflenmiştir.

Verilerin Toplanması- Bu çalışmada kullanılmak üzere Google Scholar platformunda yer alan Türkçe 100 adet sağlık alanındaki akademik çalışma ve bazı web sayfalarının içerikleri çekilerek bir derlem oluşturulmuştur.

Ön İşlem Aşaması- Oluşturulan veri seti üzerinde ön işlem çalışmaları yapılarak veriler işlenebilir hale getirilmiştir. Bu çalışmalardan ilki metinlerin tamamının küçük harflere çevrilmesi, cümle sonundaki noktalar referans alınarak metinlerin cümlelere daha sonra cümle içindeki boşluklara göre de cümlelerin kelimelere dönüştürülmesidir. Son olarak noktalama işaretleri temizlenmiş, normalizasyon işlemleri uygulanmıştır.

Verilerin Etiketlenmesi- Ön işlem aşamasından sonra kelimelere dönüştürülmüş verilerin etiketlenme işlemi yapılmıştır. Etiketleme için kullanılan varlık isimleri olarak şunlar tercih edilmiştir. Kişi, Kurum, Yer, Hastalık, Belirti, Tedavi, Anatomik Yapı Adı (AYA). Varlık isimlerinin tespit edilmesinde ve verilerin etiketlenmesinde enfeksiyon hastalıkları doktoru bir uzmandan destek alınmıştır.

2.2. Metot

2.2.1 Doğal Dil İşleme

Doğal dil işleme yapay zekânın en fazla çalışılan dolayısıyla da en önemli alt dallarından biridir. Doğal dil işleme en basit tanımla Türkçe, İngilizce, Almanca gibi doğal dillerin işlenmesi ve üzerinde çalışmalar yapılması ile ilgili bir bilim dalıdır (Kocatekin ve Ünay, 2013). Doğal dil işlemenin temel görevleri arasında bir metnin bir dilden başka bir dile çevrilmesi, bir metnin ait olduğu kategoriye göre sınıflandırılması, soru-cevap sistemleri ve akıllı asistanların geliştirilmesi gibi belli başlı konular gösterilebilir (Küçük ve Arıcı, 2018). Kısaca bir metin üzerinde yapılabilecek her türlü araştırma ve geliştirme uygulamaları doğal dil işlemenin çalışma alanına girmektedir. Bu nedenle günümüzde oldukça popüler ve önemli bir çalışma alanıdır bu durumun üretilen veri miktarlarının artmasıyla birlikte gelecekte daha da önem kazanacağı tahmin edilmektedir.

2.2.2 Varlık İsmi Tanıma

Varlık ismi tanıma, yapay zekânın popüler alt başlıklarından biri olan doğal dil işlemenin önemli bir çalışma konusudur. Bu problem ile ilgili çeşitli alanlara yönelik çalışmalar yapılmıştır. Ancak sağlık alanı ile alakalı çalışmalar olsa da bunların sayısı yeterli düzeyde değildir. Ayrıca yapılan çalışmalar incelendiğinde başta İngilizce dili olmak üzere farklı dillerde uygulamalar bulunmakta iken Türkçe diline ait çalışmaların sayısı çok sınırlıdır. Bu nedenle bu çalışmanın dili de Türkçe olarak seçilmiş ve Türkçe bir veri seti üretilmiştir (Eken,2015).

Bu çalışmada, varlık isimlerinin etiketlenmesi işlemi tamamlandıktan sonra geliştirilen model ile bu varlık isimlerinin model tarafından yapay zekâ ile tespit edilmesi hedeflenmiştir. Etiketleme işlemi

aşamasında şöyle bir format kullanılmıştır. Tablo 2’de etiketleme örneği görülmektedir.
B=> Bir varlık ismi başladı, I=> Varlık ismi devam ediyor (içinde), O=> Varlık isminin dışında

Tablo 2. Varlık isimlerinin etiketlenme yönteminin örneği

Örnek-1	Kelime	kalp	hastası	olabilir	bronşit	iltihabı	var	antibiyotik	tedavisi
	Etiket	B-AYA	O	O	B-Hastalık	O	O	B-Tedavi	O
Örnek-2	Kelime	hemodiyaliz	kronik	böbrek	yetmezliği	hastalarında	uygulanır		
	Etiket	B-Tedavi	B-Hastalık	I- Hastalık	I- Hastalık	O	O		

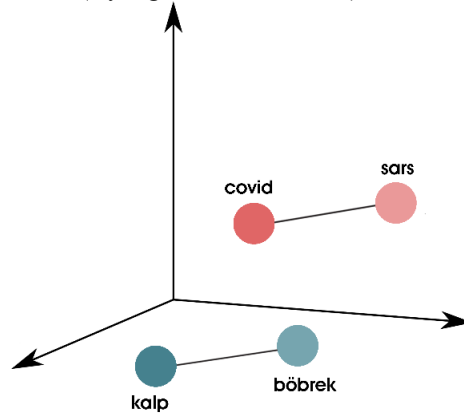
Etiketleme işlemi sonrasında oluşturulan veri setinde yer alan hastalık, tedavi ve diğer etiketlerine sahip varlık isimlerinden bazıları tabloda verilmiştir. Daha öncede değinildiği üzere *B* bir varlık isminin başladığını, *I* geçerli varlık isminin devam ettiğini son olarak *O* ise varlık isminin dışında olduğunu yani herhangi bir varlık ismine ait olmadığını temsil etmektedir. Tablo 3’te örnek etiketler görülmektedir.

Tablo 3. Etiketlenmiş veri örnekleri

B-Hastalık	kanser boğmaca frengi bronşit ülser dizanteri egzama diyabet guatr hepatit.....
I-Hastalık	iltihap ağrı kanama yara bulantı felci şişliği çarpması yanığı zorluğu darlığı
B-Tedavi	kemoterapi antibiyotik mezoterapi osteopati akupunktur
I-Tedavi	uygulamalar terapi sonda nakli küçültmek baypas protez hortum(tüp).....
O	modern hizmet ülkemiz çok ve yöntem sistem ayrıca insan durum

2.2.3. Kelime Vektörleri

Doğal dil işleme problemlerinde bir kelimenin gösterim biçimi bir başka deyişle bilgisayarın o kelimeyi nasıl anlayacağı bu problemin en kritik noktasıdır. Bu gösterim şeklinin geleneksel ve modern olarak sınıflandırılabilir çeşitleri vardır (Goldberg ve Levy, 2014). Günümüzde kullanılan en popüler yöntem Google yapay zekâ araştırmaları ekibi tarafından geliştirilen, kelimelerin vektör biçiminde sayısal ifadelerle temsil edildiği Word2Vec yöntemidir (Rong, 2014). Bu yöntemde göre her bir kelime uzayda vektörlerle doğru bir şekilde sembolize edilirse, bir vektöre dolayısıyla vektörün temsil ettiği kelimeye en yakın vektör en yakın anlamlı kelime olarak düşünülebilir. İki vektör arasında ki yakınlık kosinüs benzerliği ile tespit edilebilmektedir (Aydoğan ve Karıcı, 2019).

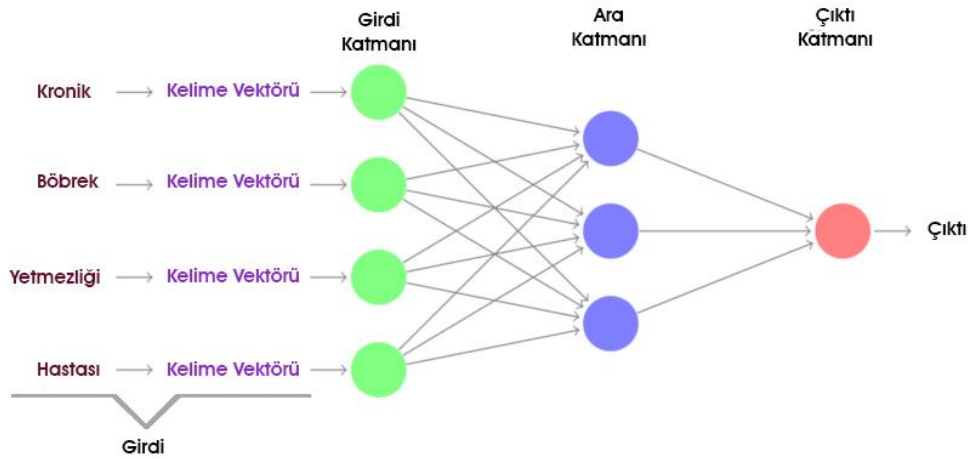


Şekil 4. İlişkili kelimelerin düzlem üzerindeki gösterimi

Şekil 4’te vektörel forma dönüştürülmüş yakın anlamlı kelimelerin tespit edildiği görülmektedir. Bu çalışmada da yakın anlamlı kelimelerin belirlenebilmesi için oluşturulmuş Türkçe veri seti üzerinde Word2Vec metodu kullanılarak çalışmalar yapılmıştır.

2.2.4. Yapay Sinir Ağları

Yapay sinir ağları, günümüzde kullanılan birçok yapay zekâ yaklaşımının temelini oluşturmaktadır. Esasen yapay sinir ağlarında, kendi kendine öğrenilme ve değerlendirme yapabilme özellikleri ile insan beyninin modellenmesi hedeflenmiştir. Daha öncede değinildiği üzere bu durumda yapay zekâ denilen yaklaşımın temelidir. Çünkü makinelerin de insanlar gibi eğitilebilmesi, öğrenbilmesi ve çıkarımlar yapabilmesi amaçlanmıştır (Öztürk ve Şahin, 2018). Şekil 5’te bir yapay sinir ağının yapısı görülmektedir.



Şekil 5. Yapay sinir ağı yapısı

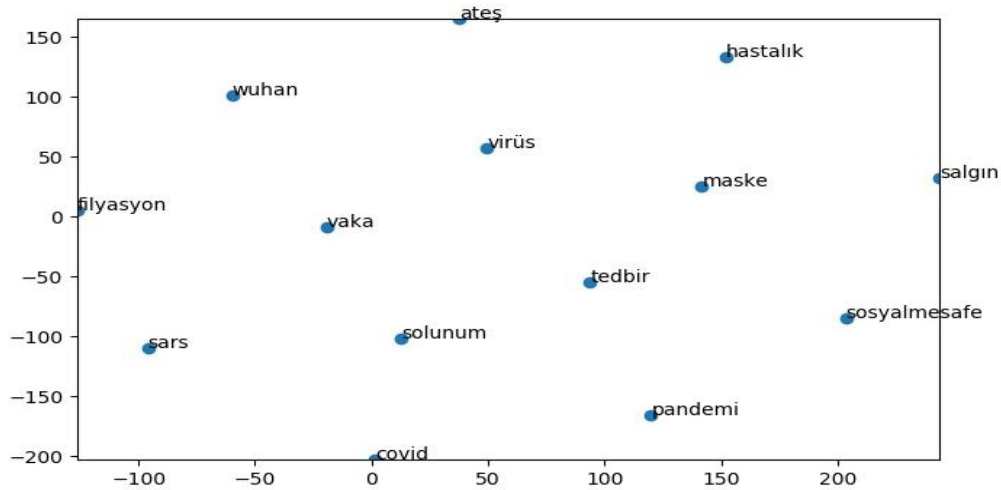
3. Deneysel Çalışmalar

Bu çalışma için üretilen Türkçe veri seti eğitim ve test verisi olarak iki kısma bölünmüş eğitim verisi üzerinde geliştirilen modelin eğitim işlemi yapılmış, test verisi üzerinde ise modelin eğitimi tamamlandıktan sonra hiç görmediği veriler üzerinde tahminler yapması sağlanmıştır. Üretilen veri setinin %80'lik kısmı eğitim verisi %20'lik kısmı ise test verisi olarak bölünmüştür. Tablo 4'te eğitim ve test verisinde yer alan toplam kelime sayısı ayrıca hastalık, belirti, tedavi, anatomik yapı adı (AYA) gibi kategorilerden yer alan kelime sayıları da yer almaktadır.

Tablo 4. Kategorilere göre veri setinde bulunan kelime sayıları

		Kelime Sayısı	Kişi İsmi	Kurum	Yer	Hastalık	Belirti	Tedavi	AYA
Veri Seti	Eğitim	742400	10811	951	3830	16718	18265	9534	30269
	Test	185600	3099	118	500	11921	6331	1056	5067
		928000							

Üretilen Türkçe veri seti üzerinde Word2Vec yöntemi kullanılarak kelime vektörleri oluşturulmuştur. Deneysel çalışmalar sırasında bu vektörlerin dolayısıyla temsil ettikleri kelimelerin birbirleri ile olan anlamsal ilişkileri incelenmiştir. Örneğin "covid" kelimesi ile yakın anlamlı kelimeler listelenmek istendiğinde *virüs*, *maske*, *sars*, *sosyal mesafe* gibi kelimelerin tespit edildiği görülmüştür. Dilimizi bilmeyen bir sistem belli bir kelime sorgulatıldığında başarılı bir şekilde o kelime ile benzer kelimeleri yakalamıştır. Bu durum Şekil 6'da ki gibi görselleştirilmiştir.



Şekil 6. Covid kelimesi ile ilişkili kelimelerin 2 boyutlu düzlemde gösterimi

Yapay zekâ teknolojilerinin sağlık alanında kullanımı ile birlikte geliştirilebilecek uygulamalar ile sağlık çalışanları üzerinde ki iş yükünün azaltılması daha kısa sürede daha çok bilgiye erişimin sağlanarak günümüz teknolojilerinin bu tür kritik alanlarda daha fazla yer alması hedeflenmektedir. Bu kapsamda covid-19 hastalığı ile alakalı akademik çalışmalardan üretilmiş kelime vektörleri ile semptom ve tedbir kelimeleri sorgulandığında elde edilen sonuçlar heyecan vericidir. Örneğin *semptom* kelimesi ile elde edilen sonuçlar incelendiğinde en yüksek skor ile elde edilen ilk kelimeler ateş, öksürük, ishal gibi sağlık bakanlığı tarafından covid-19 hastalığının en belirgin belirtileri olarak kabul edilen semptomlar yakalanmıştır. Şekil 7’de yer alan bir diğer örnekte ise *tedbir* kelimesi sorgulanmış yine en yakın ilişkili 10 kelime skor değerlerine göre listelendiğinde maske, sosyal mesafe gibi kelimeler elde edilmiştir. Pandemi sürecinin en yaygın sloganlarından olan *evdekal* ifadesinin de burada yakalanmış olması dikkat çekicidir. Bu durumda yapay zekânın gücünü bir kez daha ortaya koymaktadır.

vector(“semptom”)	vector(“tedbir”)
[('ateş', 0.7206), ('öksürük', 0.6463), ('ishal', 0.6445), ('ağrı', 0.6126), ('boğaz', 0.5989), ('göğüs', 0.5809), ('solunum', 0.5762), ('kuru', 0.5708), ('belirti', 0.5695), ('covid', 0.5689)]	[('önlem', 0.7751), ('maske', 0.7654), ('sosyalmesafe', 0.7424), ('temas', 0.7126), ('vaka', 0.6989), ('mesafe', 0.6816), ('evdekal', 0.6629), ('salgın', 0.6073), ('pandemi', 0.5561), ('yasak', 0.5430)]

Şekil 7. İlişkili kelimelerin listesi

Bu çalışmanın temel amacı ise üretilen Türkçe veri seti üzerinde etiketlenen varlık isimlerinin geliştirilen model ile eğitilmesinin ardından sistemin görmediği verileri tahmin edip varlık isimlerini otomatik olarak çıkarabilmesidir. Örneğin Şekil 8’de görülen program çıktısında eğitim işlemi sonrasında test verisi olarak sisteme verilen verilerden varlık ismi olarak tahmin edilen kelimeler model tarafından çıkarılmıştır. Öksürük, ateş gibi kelimeler belirti olarak tanımlanırken, kanser, bronşit ve diyabet gibi kelimelerin hastalık olduğu doğru olarak tahmin edilmiştir. Bu uygulamanın özellikle covid-19 gibi yeni hastalıklar ile alakalı yapılabilecek akademik taramalarda büyük kolaylık sağlayacağı düşünülmektedir.

Kanser olmuş, kalp hastası da olabilir. Karaciğer tedavisi için hastaneye yattım. Soğuk algınlığı, diş çürüğü, böbrek ağrısı, bronşit ve diyabet var. Antibiyotik tedavisi uygulanmaktadır. Öksürük ve yüksek ateş nedeniyle derhal müdahale

Renk Kodları: hastalık, anatomik yapı adı, tedavi, belirti

Şekil 8. Uygulamanın ekran çıktısı

Modelin performansının değerlendirilmesinde ise yapay zekâ çalışmalarında en sık kullanılan precision, recall ve f-score ölçütleri kullanılmış ve bu değerler aşağıdaki 1, 2 ve 3 numaralı formüller ile hesaplanmıştır.

$$\text{Precision} = \frac{\text{Modelin Doğru Tahminleri}}{\text{Modelin Tüm Tahminleri}} \quad (1)$$

$$\text{Recall} = \frac{\text{Modelin Doğru Tahminleri}}{\text{Modelin Bilmesi Gereken Tüm Varlık İsimleri}} \quad (2)$$

$$\text{F – Score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Tablo 5. Geliştirilen modelin başarıml oranları

	Precision (%)	Recall (%)	F-Score (%)
B-Kişi	82.32	72.92	77.33
I-Kişi	78.40	74.80	76.66
B-Kurum	82.61	71.03	76.38
I-Kurum	85.77	74.65	79.82
B-Yer	78.90	75.20	77.00
I-Yer	71.11	69.88	70.49
B-Hastalık	72.09	62.63	67.03
I-Hastalık	87.31	69.20	77.21
B-Belirti	83.25	76.75	79.87
I-Belirti	86.01	84.25	85.12
B-Tedavi	86.55	73.18	79.31
I-Tedavi	86.88	83.64	85.23
B-AYA	85.33	86.56	85.94
I-AYA	77.50	75.80	76.67
O	95.86	94.19	95.01

Tablo 5’te veri setinde etiketlenen varlık isimlerinin başladığı (B) ve devam ettiği (I) durumları farklı etiketler halinde precision, Recall ve f-score ölçütlerine göre elde ettiği başarıml oranları görülmektedir. Bir başka deyişle modelin eğitim süreci sırasında hiç görmediği test verisi üzerinde yaptığı tahminlerin sonuçlarıdır denilebilir. Örneğin model herhangi bir varlık ismi olmayan ve O etiketi ile temsil edilen verilerin yaklaşık %95’ini doğru olarak tahmin etmiştir. Hastalık isimlerinin başladığı B-Hastalık etiketlerde precision değeri %72 iken, devamını temsil eden I-Hastalık etiketinde ise bu değer %87 olmuştur. Yani model hastalık isimlerinin devam ettiği etiketi tahmin etmede, başladığı etiketi tahmin etmeye göre daha başarılı olmuştur.

Deneyisel çalışmaların tümü Python 3 programlama dili ve Anaconda ortamı ayrıca Microsoft Windows Server 2012 R2 işletim sistemi ile Intel Xeon E5-2630 2.20 CPU ve 64 GB bellek kullanılarak geliştirilmiştir.

4. Sonuçlar

Yapay zekâ teknolojileri başta bilgisayar bilimleri ve mühendislik alanında olmak üzere birçok alanda kullanılmakta ve oldukça başarılı performanslar sağlamaktadır. Sağlık alanında özellikle son yıllarda yapay zekânın kullanımı ile birlikte önemli gelişmelerin yaşandığı bir alan haline gelmiştir. Toplum için hayati bir öneme sahip olan sağlık alanında yapılan tüm araştırmalar başta sağlık çalışanlarının iş yükünü azaltmayı, karar verebilmelerini kolaylaştırmayı hedeflemekte olduğu için oldukça kıymetli olarak kabul edilmektedir. Bu nedenle yapay zekânın önemli alt alanlarından biri olan doğal dil işleme teknikleri ile yapay zekâ temelli bir uygulama geliştirilmiştir.

Çalışmada kullanılmak üzere Google Scholar platformundan sağlık alanına yönelik Türkçe akademik makaleler ve çeşitli web sayfalarından toplanan Türkçe içerikler ile literatürde bulunmayan özgün ve yeni Türkçe bir veri seti üretilmiştir. Bu veri seti üzerinde yapılan çalışmalar neticesinde geliştirilen model ile sisteme yüklenen metinlerde hastalık, tedavi, belirti gibi etiketlerin otomatik olarak tanımlanması hedeflenmiştir. Bu durumun özellikle covid-19 gibi yeni hastalıklar hakkında bilgi toplanması, literatür taranması, araştırma yapılması safhasında büyük kolaylık sağlayacağı düşünülmektedir.

Yapılan deneysel çalışmalar sonrasında, metinler içerisinden kişi, kurum, yer, hastalık, belirti ve tedavi gibi varlık isimlerinin %91 başarıml oranında doğru olarak çıkarıldığı görülmüş ve bu performans precision, recall ve f-score gibi ölçütler ile desteklenmiştir. Ayrıca bu çalışma kapsamında üretilen Türkçe veri setinin, doğal dil işleme ve yapay zekâ çalışmalarına katkı sağlaması amacıyla araştırmacılar ile paylaşılması planlanmaktadır.

Referanslar

- Aydoğan, M, Karci, A. (2019). Kelime Temsil Yöntemleri ile Kelime Benzerliklerinin İncelenmesi. Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 34 (2), 181-196. DOI: 10.21605/cukurovaummfd.609119

- Bozüyük, T., Yağci, C., Gökçe, İ., & Görkem, A. K. A. R. (2005). Yapay Zeka Teknolojilerinin Endüstrideki Uygulamaları.
- Büyükgöze, S., & Dereli, E. (2019). Dijital Sağlık Uygulamalarında Yapay Zeka. *VI. Uluslararası Bilimsel ve Mesleki Çalışmalar Kongresi-Fen ve Sağlık*, 07-10.
- Eken, B. (2015). *Kısa Metinlerde Varlık İsmi Tanıma* (Doctoral dissertation, Fen Bilimleri Enstitüsü).
- Goldberg, Y., & Levy, O. (2014). word2vec Explained: deriving Mikolov et al.'s negative-sampling
- Görz, G., & Nebel, B. (2002). Yapay zeka. İstanbul: İnkılap Kitapevi.
- Johnson, Khari; (2018), “Baidu Researcher’s breast cancer dedection algorithm outperforms human pathologists”, Venture Beat, (18 Haziran 2018), <https://venturebeat.com/2018/06/18/baidu-researchers-breast-cancer-detection-algorithm-outperforms-human-pathologists/>. (Erişim Tarihi: 05.06.2020)
- Kocatekin, T., & Ünay, D. (2013, April). Text mining in radiology reports. In *2013 21st Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- Küçük, D., & Arici, N. (2018). Doğal Dil İşlemede Derin Öğrenme Uygulamaları Üzerine Bir Literatür Çalışması. *Uluslararası Yönetim Bilişim Sistemleri Ve Bilgisayar Bilimleri Dergisi*, 2(2), 76-86.
- Linn, Allison; “How Microsoft computer scientists and researchers are working to ‘solve’ cancer”, Microsoft, <https://news.microsoft.com/stories/computingcancer/>. (Erişim Tarihi: 05.06.2020)
- Lock, Stan; (2018), “Danish startup Corti uses AI to save people from cardiac arrest “, Business Insider Nordic, (30 Nisan 2018), <https://nordic.businessinsider.com/danish-speech-recognition-ai-corti-saves-lives-by-detecting-cardiac-arrest-in-real-time--/>. (Erişim Tarihi: 05.06.2020)
- Macdonald, Fiona; (2016), “AI Can Analyse Mammogram Results 30 Times Faster Than Doctors, And With 99% Accuracy”, Science Alert, (19 Eylül 2016), <https://www.sciencealert.com/ai-analyses-mammograms-30-times-faster-and-20-more-accurately-than-doctors>. (Erişim Tarihi: 05.06.2020)
- Markoff, John; (2016), “Microsoft Finds Cancer Clues in Search Queries”, New York Times, (7 Temmuz 2016), <https://www.nytimes.com/2016/06/08/technology/online-searches-can-identify-cancer-victims-study-finds.html>. (Erişim Tarihi: 05.06.2020)
- Öztürk, K., & Şahin, M. E. Yapay Sinir Ağları ve Yapay Zekâ’ya Genel Bir Bakış. *Takvim-i Vekayi*, 6(2), 25-36, (2018).
- Rong, X. (2014). word2vec parameter learning explained. *arXiv preprint arXiv:1411.2738*.ord-embedding method. *arXiv preprint arXiv:1402.3722*.
- Spencer, Geoff; (2018), “AI and preventative healthcare: Diagnosis in the blink of an eye”, Microsoft Asia News Center, (17 Eylül 2018), <https://news.microsoft.com/apac/features/ai-and-preventative-healthcare-diagnosis-in-the-blink-of-an-eye/>. (Erişim Tarihi: 05.06.2020)
- Webster, William; (2018), “Smartphones and Selfies sourcing AI Healthcare Diagnostics”, Avoncourt, (30 Kasım 2018), <http://avoncourt.de/smartphones-and-selfies-sourcing-ai-healthcare-diagnostics/>. (Erişim Tarihi: 05.06.2020)