

---

# DUYGU ANALİZİ VE DİJİTAL DÖNÜŞÜM ÜZERİNE ETKİLERİ

---

Muhammed KOTAN<sup>1</sup>

## 1. GİRİŞ

Dijital Dönüşüm (DD), organizasyonların hızla gelişen teknolojiler ve değişen ihtiyaçlar karşısında işleyişlerini düzenlemek üzere yeni teknolojileri uygulayarak ve/veya mevcut sistemlerini iyileştirerek bilgi ve iletişim teknolojisi (BİT) çözümlerinin kullanılmasıyla gerçekleştirdikleri dönüşüm sürecidir. Dördüncü sanayi devrimi döneminden itibaren başta iş dünyası, eğitim, hükümet, üretim vb. olmak üzere birçok farklı alanda DD uygulanmaktadır (Verina ve Titko, 2019). Veri etrafında işleyen organizasyonlar dünyasında dijital bir dönüşümden bahsedebilmek için verinin etkisinin anlaşılması büyük önem taşımaktadır.

Veriler üç başlık altında sınıflandırılabilir. Yapılandırılmış veriler, belirli bir formata sahip olan saklanması, erişilmesi, işlenmesi ve analiz edilmesi kolay verileri ifade eder. Verilerin, kurumsal kullanım için bilgi elde etmek amacıyla sorgulara yanıt verecek şekilde tanımlanabilir olduğu anlamına gelir (Eberendu, 2016; Doan vd, 2009). Belirli standartlar doğrultusunda tutuldukları için erişmesi, değiştirmesi ve kategorize etmesi kolaydır (Doğan ve Arslantekin, 2016). Kendini açıklayan veri olarak da adlandırılan yarı yapılandırılmış veriler, yapısal bir veri türünü yansıtmakla birlikte değişken bir yapıya sahip ancak sabit veya açık bir şemaya uymayan düzensiz verilerdir (Eberendu, 2016; Buneman, 1997). XML, JSON ve HTML gibi ilgili ancak farklı özelliklere sahip çeşitli kaynaklardan gelen bilgilerin bir bütün halinde sunulmasına olanak sağlar (Aktan, 2018; Hänig vd., 2010). Bu türlerin aksine yapılandırılmamış veriler ise e-posta, video, resim, metin dosyaları, sosyal medya iletileri gibi farklı kaynaklardan elde

---

<sup>1</sup> Dr. Öğr. Üyesi, Sakarya Üniversitesi, Bilişim Sistemleri Mühendisliği Bölümü, mkotan@sakarya.edu.tr, Orcid ID: 0000-0002-5218-8848

edilen ve standart olmayıp tekrar eden bir kalıba uymayan veri türünü ifade etmektedir.

DD yukarıda bahsi geçen yapılandırılmış, yarı yapılandırılmış veya yapılandırılmamış büyük hacimlerde olabilen verilerle çalışır. Üretilen ve depolanan veri miktarı, verilerin dijitalleşmesi nedeniyle son zamanlarda üstel bir şekilde katlanarak artmaktadır ve günümüzde organizasyonlar bu tür çeşitli verilerle dolup taşmaktadır. Tahminlere göre verilerin %80-%90 oranındaki bir çoğunluğunu metin, görüntü, video, ses, web kayıtları, sosyal medya ve daha fazlası gibi yapılandırılmamış veriler oluşturmaktadır (Harbert, 2021; Abdullah ve Ahmad, 2015). Yapılandırılmamış verilerin toplanması ve analizi, kuruluşlara işleyişlerine ilişkin içgörü sağlar (Eberendu, 2016). Bu nedenle, büyük yapılandırılmamış veri kümelerinden doğru bilgiyi çıkarmak amacıyla nasıl kullanacaklarını bulan ve yöntemler geliştiren organizasyonlar için bu veriler; rekabet avantajı sağlama, üretkenliği artırma ve yenilik oluşturma potansiyeline sahip DD süreçleri için değerli devasa bir kaynaktır.

Organizasyonlar, DD dönüşüm süreçlerinde ne kadar başarılı olduklarını anlamak üzere büyük ölçüde metrik raporlarını ve süreç takiplerini kullanmaktadırlar. Fakat DD; sosyal, mobil, bulut ve nesnelerin interneti gibi birçok yeni dijital teknolojilerin dönüşümsel etkisini kapsamaktadır ve farklı kaynaklardan, cihazlardan, telefonlardan ve sensörlerden gelen büyük verilerin harmanlanması gerekmektedir (Sebastian vd., 2020). Burada dikkat edilmesi gereken önemli noktalardan biri verilerin elde edilmesinin yanında, bu verilerin yorumlanması ve kullanılmasıdır. Yapılandırılmamış veriler, farklı bir şekilde sıralama, aranma, görselleştirme veya analiz gerektirdiğinden dolayı bilgi çıkarmak, paylaşmak ve değer sağlamak için yeni araç, süreç ve beceriler gerekecektir.

Günümüz teknolojisiyle birlikte kullanıcılar, aldıkları ürün veya hizmetlerle ilgili ya da organizasyonların işleyişine yönelik düşüncelerini, geri bildirimlerini ve beklentilerini çeşitli web sitelerinde ve sosyal medya platformlarında kolaylıkla paylaşabilmektedir. Kullanıcılar genellikle paylaşımlarında duygularının da izlerini bırakırlar. Tüm bu içerik türleri, şirketlere marka itibarlarını, hizmetlerini ve ürünlerini analiz etmek için önemli bilgiler sunar. Organizasyonlar, dağılmış durumda olan büyük miktardaki verileri tüm kaynaklarda ve platformlarda gerçek zamanlı olarak nasıl ele alınacağını anlamak, müşterilerini bilgilendirmek ve işi dönüştürmek için bu verilerden yararlanmaya odaklanırlar (Somekh, 2021).

DD, temel ticari işlemlerin, ürünlerin, süreçlerin, organizasyon yapılarının ve yönetim kavramlarının değişmesi amacıyla dijital teknolojilerin uygulanmasını içerir ve DD sürecinde dönüşümü gerçekleştirmeye yönelik nasıl bir yol izleneceği, ne tür stratejiler kullanılacağı ve organizasyonun durumunun nasıl değerlendirileceği de önem arz etmektedir (Sağlam, 2021). DD'nin temel faydaları arasında satışların ve üretkenliğin artması ile değer yaratılması ve müşteri etkileşiminde yenilikler sunulması gösterilebilir (Sağlam, 2021; Matt vd., 2015). Bu nedenle de kullanıcı duygularını, tepkilerini, tüketicilerin ürün tercihlerini, ürün kişiselleştirmesini ve diğer organizasyonel gereksinimleri belirlemek için verileri analiz etmeye ihtiyaç vardır (Eberendu, 2016). Doğal dil işleme (DDİ) yöntemleri sayesinde yapılandırılmamış metinler analiz edilebilecek verilere dönüştürülebilmektedir. En popüler DDİ tekniklerinden biri de bir durum hakkında nasıl hissedildiğini bulmak için kullanılan duygu analizidir (DA). Kullanıcı geri bildirimlerini, anket yanıtlarını ve ürün incelemelerini analiz etmek için yaygın olarak kullanılır. DA, DD süreçlerinde işletmelerin kullanıcıları hakkında daha iyi içgörüler elde etmelerine ve içgörü ile eylem arasındaki boşluğu doldurmalarına yardımcı olmakla ilgilidir. Gerekli verilere erişildiğinde doğru strateji ve platform entegrasyonu ile birlikte verilerle ilgili DD başlatılabilir.

Bu bölümde DA konusuna bir giriş yapılarak kullanılan yöntem ve teknikler tanıtılacaktır. DA'nın faydaları, kurumsal avantajlarının neler olabileceği ve DD süreçlerindeki etkileri hakkında değerlendirmeler sunulacaktır.

## 2. DOĞAL DİL İŞLEME VE DUYGU ANALİZİ

DA konusuna giriş yapmadan önce DDİ konusuna kısa bir değinmek faydalı olacaktır. DDİ; makinelere dil yorumlama, sınıflandırma, düzenleme ve üretme gibi özelliklerin öğretilmesini kapsar. Yapay zekâ (YZ) modelleri için temel teknolojilerden biri olan DDİ, işletmeler için en önemli ve yaygın araçlardan biri haline gelmiştir. DDİ'deki hızlı gelişmeler sayesinde dünya üzerinde anlamlı bir ekonomik etkiye sahip olmaya başlayan ve önemli dil yeteneklerine sahip YZ sistemleri geliştirilmeye başlanmıştır (Zhang vd., 2021).

DDİ, yakın zamanda derin öğrenmenin benimsenmesi ile birlikte, metin oluşturma, makine çevirisi, soru cevaplama, metin sınıflandırma, konu modelleme, duygu analizi ve diğer görevlerdeki başarılı uygulamalarıyla organizasyonlar için önemli bir teknoloji haline gelmiştir. Küçük

girişimlerden büyük işletmelere kadar çok sayıda kuruluş, müşterileri için daha verimli, kişiselleştirilmiş deneyimler oluşturmak için bu teknolojileri benimsemeye motive olmakta ya da gelecekte kuruluşlarına getirebileceği değeri kabul etmektedirler (Global AI Adoption Index, 2021).

Dünya çapında 7,502 işletme genelinde gerçekleştirilen IBM Global AI Adoption Index 2022 araştırmasına göre şirketlerin başlıca DDİ kullanma tercihleri arasında “Müşteri Hizmetleri”, “Güvenlik”, “İş Geliştirme”, “Satış”, “Pazarlama”, “İnsan Kanyakları veya Çalışan Hizmetleri” vb. listelenmektedir. DA kurumlar arasında en popüler DDİ yöntemleri arasında listelenmekte ve daha çok müşteri hizmetleri ve çalışan hizmetleri alanında tercih edilmektedir.

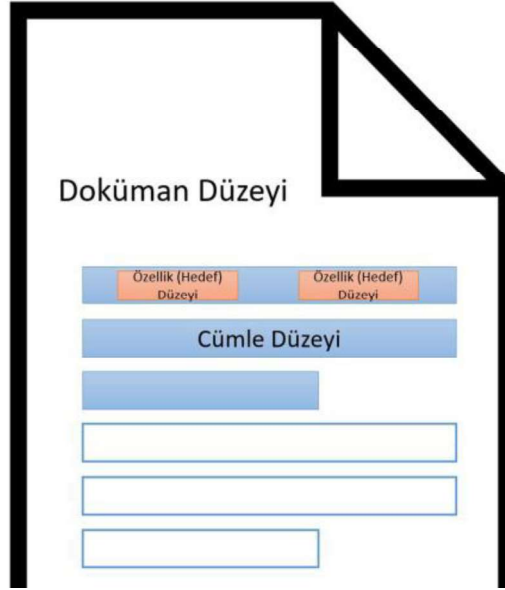
## 2.1. Duygu Analizi

DA, belirli metin, görüntü, ses verilerinden bilişim destekli olarak insan davranışını, duygu ve inanışlarını keşfetmek için kullanılan, genellikle DDİ ve makine öğrenmesi gibi alanlarla birlikte çalışan disiplinler arası bir analiz yöntemidir (Cebeci, 2020; Poria vd., 2018). "Fikir madenciliği" olarak da bilinen DA, bireyler, kuruluşlar, sorunlar, olaylar ve konular gibi varlıklar için fikir ve duyguların anlaşılması amacıyla gerçekleştirilen hesaplamalı çalışmadır. Bu bağlamda duygu (sentiment) kavramının ne anlama geldiğini anlamak faydalı olacaktır. Duygu, yazarın bir yargısı, ruh hali veya değerlendirmesi olarak düşünülebilir (Osimo ve Mureddu, 2012). Duygular genellikle çeşitli metinlerden çıkarılır ve buna göre olumlu, olumsuz veya tarafsız sınıflara ayrılır. Metnin yanı sıra duygular, konuşma ve yüz ifadeleriyle de belirtilebilir. Çok modlu (multimodal) duygu analizi, sesli ve görsel veriler gibi ek modaliteleri de dahil ederek standart metin tabanlı duygu analizinin farklı bir boyutudur (Wankhade vd., 2022).

DA çalışmaları; (i) doküman düzeyi, (ii) cümle düzeyi ve (iii) özellik düzeyi olmak üzere üç çıkarım düzeyinde gerçekleştirilebilir (Cebeci, 2020; Dang vd., 2020):

- i. **Doküman düzeyi analiz:** DA, tüm belge üzerinde gerçekleştirilerek tek kutupluluk verilir. Buradaki görev, belgenin genel görüşünü belirlemektir (Wankhade vd., 2022).
- ii. **Cümle düzeyi analiz:** Her bir cümlenin bir görüş ifade edip etmediğini belirler ve cümlenin tek bir varlık hakkındaki fikirleri ifade ettiğini varsayar. Her cümle analiz edilir ve karşılık gelen bir polarite bulunur (Wankhade vd., 2022).

- iii. **Özellik (Hedef) düzeyi analiz:** Duygu hedefi ve ilgili varlık tespit edilmek üzere daha ayrıntılı bir analiz gerçekleştirilerek duyguyu varlıkların belirli yönlerine göre sınıflandırmayı amaçlar (Cebeci, 2020). Öncelikle varlıklar ve özellikleri belirlenir. Cümlede kullanılan tüm yönlere dikkat edilir ve kutupluluk atanır. Ardından tüm cümle için bir toplu duygu hesaplanır (Schouten ve Frasincar, 2015).



Şekil 1. Duygu Analizi Seviyeleri

## 2.2. Duygu Analizi Türleri

**Standart DA:** Genelde ilk akla gelen DA şekli olarak bir görüşteki temel duyguyu tanımlar ve ardından genellikle olumlu, olumsuz veya tarafsız olarak etiketler. Bir otel hakkındaki aşağıdaki üç yorum örnek olarak gösterilebilir.

“Uygun fiyatlarda çok harika hizmetler sunuluyor.

Kesinlikle tekrar kalmak isterim.” 😊 (OLUMLU)

“Sadece birkaç gün konaklayabildim.” 😐 (TARAFSIZ)

“Otel çalışanları çok kaba ve odalar hiç temiz değil.” 😞 (OLUMSUZ)

**İnce Taneli (Fine-Grained) DA:** DA gerçekleştirilirken önceki bahsedilen “olumlu”, “olumsuz” ve “tarafsız” kategorilerin ötesinde farklı sınıflandırmalar da gerçekleştirilebilir. İnce taneli DA, daha verimli sonuçlar elde etmek için biraz daha kategori ekleyecektir (Saju vd., 2020). Kutupluluğa daha fazla odaklanır. Kutupluluk, duygu puanı olarak bilinen sayısal bir derecelendirme olarak ifade edilebilir. Örneğin, yıldız incelemelerindeki çok olumsuz ve çok olumlu arası derecelendirmeler

düşünülebilir. Ayrıca, belirli duygulara göre daha hassas durumlar da ele alınabilir. Örneğin olumlu bir duygu; mutlu, heyecanlı, etkilenmiş vb. olarak daha da irdelenebilir (Thematic).

“Önceki tasarım daha kullanışlıydı.” (OLUMSUZ)

“Aşırı ısıtıyor. Kesinlikle kimseye önermem.” (ÇOK OLUMSUZ)

**Duygu Tespiti (Emotion detection):** Birbirinden farklı duyguları tespit etmek için kullanılan diğer bir DA yöntemidir. Duygu analizinde (sentiment analysis) kutupluluk birincil kaygıdır, duygu tespitinde (emotion detection) ise duygusal veya psikolojik durum veya ruh hali algılanır (Nandwani ve Verma, 2021). İlgili kaynakta aranan heyecan, mutluluk, öfke, üzüntü vb. duyguları tespit etmek için tercih edilir. İnsan sesinden duygu çıkarımına yönelik çalışmalarda sıklıkla tercih edilmektedir.

“Bu uygulamayı kullanırken aşırı keyif aldım.” (Mutlu)

“Müşteri ile iletişim berbat!” (Öfkeli)

**Hedef tabanlı (Aspect-based) DA:** Verileri hedefe göre kategorize eden ve her birine atfedilen duyguyu tanımlayan bir tekniktir. Metin içinde yer alan farklı duyguların ilgili oldukları hedef varlıklar ile birlikte tespit edilmesi olarak tanımlanabilir (Çetin ve Eryiğit, 2018). Belirli duyguları bir ürün veya hizmetin farklı yönleriyle ilişkilendirerek bildirimlerin analizi için kullanılabilir. Belirli bir yön hakkında olumlu veya olumsuz görüşler ele alınırken hedef tabanlı işlendiğinde hakkında konuşulan kategoriler, özellikler veya konular daha kullanışlı bir şekilde ele alınır. Örneğin; bir ürüne ait fiyat, verimlilik, uyumluluk, kullanışlılık, arayüz vb. gibi farklı özellikleri hakkında gözden geçirenin görüşlerini içeren bir ürün incelemesi düşünülebilir.

“Hız yönünden harika fakat tasarımı pek kullanışlı değil.”

(Hız: OLUMLU, Tasarım: OLUMSUZ)

**Niyet Tespiti (Intent detection):** Belirli bir görüşün ardındaki eylemi işaret edebilen bir analiz türüdür (Saju vd., 2020). Görüşlerden niyetleri belirlemeyi amaçlayan niyet tespiti, hedefe yönelik diyalog sistemlerinde çok önemli görevlerden biridir (Tian vd., 2022) Bir müşteri sorununu ele almak, ürünü iyileştirmek veya şikayetleri ilgili departmanlara yönlendirmek gibi ihtiyaçların anlaşılmasına yardımcı olur (Saju vd.,2020). Sohbet botları ve müşteri desteği gibi alanlarda kullanılmaktadır.

“Bir yetkili ile görüşmek istiyorum.”

“Şifremini doğrulayamıyorum. Lütfen yardımcı olur musunuz?”

## 2.3. Duygu Analizi Süreci

Yapılandırılmamış metinsel verileri analiz etmek zorlu bir görevdir ve bu görevi başarıyla gerçekleştirmek için DDİ ve makine öğrenmesi alanlarıyla birlikte çalışmanız gerekebilir. Genel bir DA süreci; (i)veri toplama, (ii)ön işleme, (iii)öznitelik çıkarma, (iv)sınıflandırma ve (v)sonuçların yorumlanması gibi adımlar içerir.

Veri toplama, DA'nın ilk aşamasıdır. Duygu analizine bağlı olarak, metin verileri, video, ses, konum vb. gibi diğer veri türleri birleştirilebilir (Wankhade vd.,2022). Bir API yardımıyla veya Web kazıma (scraping) yöntemleriyle sosyal medya, haber kanalları, E-ticaret siteleri vb. aracılığıyla internetten veri elde edilebilir.

Metinlerin analiz edilebilmesi için öncelikle ön işleme aşamalarından geçirilmeleri ve makinelerin anlayabileceği şekilde biçimlendirilmeleri gerekir. Verilerin ön işleme tabi tutulmasındaki temel amaç metindeki gürültüyü azaltmak, sınıflandırıcının performansını iyileştirmek ve sınıflandırma sürecini hızlandırmaya yardımcı olmaktır (Haddi vd., 2013). Veri ön işlemedeki süreçler; metin temizleme, gereksiz boşluk türlerini kaldırma, kısaltmaları genişletme, köke indirgeme, durdurma kelimelerini kaldırma (Metinlerde sık kullanılan fakat çıkarıldıklarında ayrıldıkları cümlenin anlamında önemli değişiklikler oluşturmayan sözcüklerin elenmesi), olumsuzlama durumları, noktalama işaretlerinin kaldırılması, küçük harflere dönüştürme, dil bilgisi hatalarının düzeltilmesi ve öznitelik seçimi gibi aşamalar içerir.

Çevrimiçi metinlerde genellikle çok sayıda gürültü ve HTML etiketleri, komutlar gibi bilgi vermeyen kısımlar bulunur. Ayrıca kelime düzeyinde, metindeki birçok kelimenin metnin genel yönelimi üzerinde bir etkisi olmayabilir (Haddi vd., 2013). Özellikle sosyal medya verilerinde ek olarak özel anlamı bulunan karakterlerin, kısaltmaların ve emojilerin de uygun şekilde ele alınması gerekmektedir.

### 2.3.1. Öznitelik Çıkarımı ve Seçimi

Öznitelik çıkarma, verilerden değerli bilgilerin çıkarılmasını içerdiğinden ve modelin performansını doğrudan etkileyeceğinden dolayı duygu sınıflandırmasında önemli bir görevdir. Yaklaşım, metnin en temel özelliklerini kapsayan değerli bilgileri çıkarmaya çalışır.

Kelime Çantası (Bag of Words), metin özniteliklerini çıkarmak için en basit yaklaşımlardan biridir. Bir dokümandaki kelimelerin oluşumunu

tanımlar. Çanta, her cümle için bir vektörün oluşturulduğu kelimelerin kelime dağarcığını temsil eder. Bu modeldeki temel sorun, metnin sözdizimsel anlamını dikkate almamasıdır. Kelime çantası yaklaşımı, çoğu durumda daha iyi performans gösteren Terim Frekansı–Ters Doküman Frekansı (TF–IDF) kullanılarak değerlendirilir. Terim frekansı çeşitli DDİ uygulamalarında daha sık kullanılan öznitelikleri ifade etmenin en basit yollarından biridir. TF-IDF, verilen belgedeki herhangi bir dizgenin önemini vurgulayacak daha iyi sonuçlar için ağırlıklı bir şema olarak kullanılabilir. Bir kelimenin bir dokümanda ve bir derlemde ne sıklıkta geçtiğini dikkate alan ölçüler verir.

Kelime gömme modelleri, etiketli verileri kullanmadan sözdizimsel ve anlamsal bilgileri yakalayabilir ve bu nedenle DA, metin sınıflandırma, bilgi çıkarma, ve makine çevirisi gibi birçok DDİ görevinde uygulanmaktadır (Pham ve Le, 2018). Kelime gömme modellerinin arkasındaki fikir, benzer kelimelerin yakın değerler içeren vektörlere sahip olmasıdır (Pintas vd., 2021). Bu sayede kelimeler arasındaki sözdizimsel veya anlamsal benzerlik düzeyi, vektörlerinin uzaklığına göre ölçülebilir. Kelime vektörlerini tahmin etmek için Word2Vec, Glove ve FastText gibi teknikler kullanılmaktadır.

Öznitelik seçimi, öznitelik listesinde gereksiz ve alakasız olanları tanımlayıp eleyen ve böylece duygu sınıflandırma doğruluğunu artıran bir prosedürdür (Wankhade vd., 2022). Alakasız ve gereksiz öznitelikleri elemek için çeşitli öznitelik seçim yaklaşımları kullanılmaktadır (Ahmad vd., 2019).

Özellik seçimi için istatistiksel teknikler (i)filtreleme, (ii)sarmalayıcı, (iii)gömülü ve (iv)hibrit olmak üzere dört kategoriye ayrılır (Wankhade vd., 2022; Pintas vd., 2021):

- i. **Filtreleme yöntemi:** Öznitelik, çeşitli istatistiksel metrikler kullanılarak sıralanır ve ardından en yüksek sıralamaya sahip olanlar seçilir (Adomavicius, Kwon, 2011). Öznitelik ile ilgilenilen değişken arasındaki ilişki hesaplanarak bir filtreleme gerçekleştirilir ve belirli eşik değerler ile seçim yapılarak bir alt küme oluşturur (Cebeci, 2020). Bilgi Kazanımı, Ki Kare Skoru, Karşılıklı Bilgi ve Doküman Frekansı temel filtreleme tekniklerinden bazılarıdır.
- ii. **Sarmalayıcı yöntem:** Makine öğrenmesi algoritmalarına dayanmaktadır. Genellikle yinelemeli hesaplama gerektirir ve söz konusu modelleme algoritması için en uygun öznitelik kümesini belirlemeye çalışır (Wankhade vd., 2022). Bu yöntemlerde temelde oluşturulan alt modellerin her birinin kendine has öznitelik

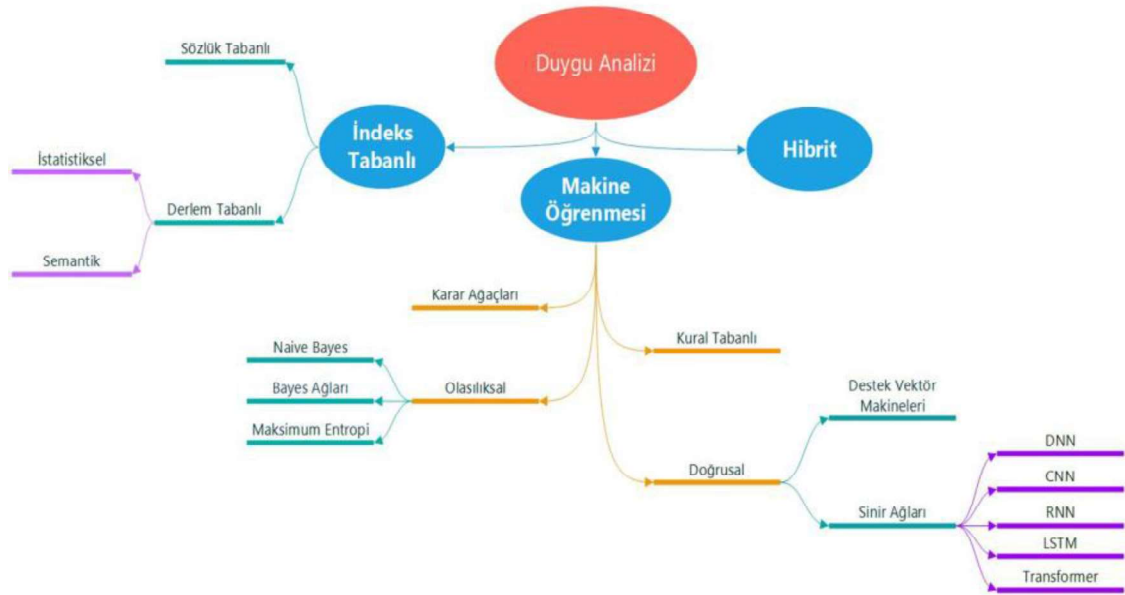


varyasyonları bulunur ve öznitelik seçimi bu alt model performanslarının karşılaştırmalı değerlendirilmesi ile gerçekleştirilir (Cebeci, 2020).

- iii. **Gömülü yöntem:** Yerleşik bir öznitelik seçme yeteneğine sahip sınıflandırma yöntemlerini kullanarak öznitelik seçim prosedürünü modelleme algoritmasının yürütülmesiyle birleştirir (Wankhade vd., 2022; Imani vd., 2013).
- iv. **Hibrit yöntem:** Optimum öznitelik alt kümesini üretmek için genellikle birden fazla yaklaşımın kullanılmasıyla performans ve doğruluk elde eder. DA için çok sayıda hibrit öznitelik seçim algoritması geliştirilmiştir (Wankhade vd., 2022).

### 2.3.2. Duygu Sınıflandırma Yaklaşımları

DA gerçekleştirmek için mevcut üç temel yaklaşım bulunmaktadır: (1) İndeks (Lexicon) tabanlı teknikler, (2) makine öğrenimi tabanlı teknikler ve (3) hibrit teknikler. Şekil 2’de bu yaklaşımlara yönelik örnek bir kategorizasyon sunulmuştur.



Şekil 2. Duygu Analizi Yaklaşımları

#### 2.3.2.1. İndeks Tabanlı Yaklaşım

Bu yaklaşım, anlamsal yönelimlerden tüm dokümanın veya cümlelerin duygu yönelimlerini hesaplar. Bir duygu indeksi (lexicon), genellikle anlamsal yönelimlerine göre ya olumlu ya da olumsuz olarak etiketlenen sözcüksel özniteliklerin bir listesidir. Genellikle önce duygu kelime listeleri

derlenerek bir duygu sözlüğü oluşturulur ve sözlükte tanımlanan olumlu ve olumsuz göstergelere dayalı olarak verilen incelemenin kutupluluk puanı belirlenir (Bonta ve Janardhan, 2019). Örneğin; olumlu sözlükler "hızlı", "iyi", "harika", "ekonomik" ve "kullanışlı" ifadelerini içerebilir. Olumsuz sözlükler ise "yavaş", "pahalı", "berbat" ve "karmaşık" gibi ifadeler içerebilir. Sözlükler manuel olarak oluşturulabileceği gibi otomatik olarak da oluşturulabilir (Gupta ve Agrawal, 2020).

Metin içerisindeki sözcükler çeşitli hesaplamalar yardımıyla özellik, cümle veya doküman bazlı olarak değerlendirilir ve tüm değer için tek bir puana dönüştürülür. Olumlu, tarafsız ve olumsuz için +1,0,-1 gibi kutuplara dayalı olarak bir puan atanır veya puan kutupsallaşma yoğunluğuna göre değerleri [+1,-1] aralığında değişecek şekilde atanabilir (Wankhade vd., 2022).

Bu yaklaşımın avantajlarından biri herhangi bir eğitim verisi gerektirmediğinden dolayı denetimsiz bir yaklaşım olarak kabul edilmesi ve kolay uygulanabilirliğidir. Başlıca dezavantajı ise alan bağımlılığıdır. Bir alanda olumlu sayılan bir sözcük diğer bir alanda olumsuz olabilir. Bu nedenle kutupluluk, alan dikkate alınarak kelimelere dikkatlice atanmalıdır (Wankhade vd., 2022). Örneğin; "Ürüne büyük bir ilgi var." cümlesinde "büyük" kelimesi olumluluk belirtirken "İletişim ağında büyük gecikmeler var." cümlesinde ise olumsuz olarak değerlendirilebilir. Alan özgü duygu sözlükleri geliştirilerek veya kelime hazneleri uyarlanarak bu sorun bir nebze aşılabılır.

İndeks tabanlı yaklaşım, başlangıçta hazırlanan listenin oluşturulma şekline bağlı olarak (i)sözlük tabanlı ve (ii)derlem tabanlı olmak üzere ikiye ayrılır:

- i. **Sözlük (Dictionary) Tabanlı** yaklaşımda duygu sınıflandırma bir terimler sözlüğü kullanılarak gerçekleştirilir. Önceden hazırlanmış, içerisinde sözcük ve tamlamalar ile duygu değerleri bulunan sözlükler yardımıyla duygu sınıflandırması yapılabilmektedir (Cebeci, 2020). Çevrimiçi sözlükler veya eş anlamlılar sözlüğü kullanılarak bu sözcükler genişletilebilir ve manuel incelemelerle geliştirilebilir. Örneğin SentiWordNet ve Senticnet sözlükleri gösterilebilir.
- ii. **Derlem (Corpus) Tabanlı** yaklaşımda ise içeriğe özgü sözcüklerin duygu yönelimleri bulunur. Derlem tabanlı olarak türetilmiş alan bilgisinin yoğunluğu yüksek olan analize özel listeler kullanılır (Cebeci, 2020). Yaklaşım, cümlelerin duygusunu belirlemek için

anlamsal ve sözdizimsel kalıpları kullanır (Wankhade vd., 2022). Bu yaklaşım, önceden tanımlanmış bir dizi duygu terimi ve bunların yönelimi ile başlar ve daha sonra, duygu dizgelerini ve bunların derlemdeki yönelimlerini keşfetmek için sözdizimsel veya benzer kalıpları araştırır (Wankhade vd., 2022). Bu yaklaşımda kendi içerisinde İstatistiksel ve Anlamsal olarak ikiye ayrılır (Gupta ve Agrawal, 2020).

### 2.3.2.2. Makine Öğrenimi Yaklaşımı

Duygu analizindeki diğer en sık kullanılan yaklaşım ise makine öğrenmesi yöntemleridir. Otomatik DA, makine öğrenimi tekniklerine dayanır. Yaklaşımın başarısı, kullanılan eğitim veri setine ve algoritmalara bağlıdır (Thematic).

Makine Öğrenimi yaklaşımlarında (i)Denetimsiz ve (ii)Denetimli öğrenme olmak üzere iki temel yaklaşım vardır.

Denetimsiz öğrenme yaklaşımlarında herhangi bir öğrenme setine bağlı kalınmaz ve derlem odaklı olarak bölümlene gerçekleştirilir (Cebeci, 2020; Sharda vd., 2014).

Denetimli öğrenme yöntemlerinin ise gerçek verilere uygulanmadan önce bir eğitim seti üzerinde eğitilmesi gerekir. Sınıflandırma modelinde etiketli bir eğitim seti kullanılarak modelin daha sonraki girdi verilerini en uygun karşılık gelen etiketle ilişkilendirmeyi öğrenmesi sağlanır. Model, eğitim verilerinden öğrenerek daha önce görmediği veriler için etiketleri veya sınıfları tahmin eder. Böylece, önceden tanımlanmış bir sözlük ihtiyacı kalmaz. Doğru sonuçlar elde etmeleri nedeniyle denetimli öğrenme yöntemleri daha yaygın olarak kullanılmaktadır (Wankhade vd., 2022). Sınıflandırma modelleri için sık kullanılan yöntemlerden bazıları aşağıda sunulmuştur.

**Naive Bayes:** Bayes Teoremine dayanmaktadır ve belirli bir metin için bir etiketin olasılığının hesaplandığı olasılıksal algoritmalarındandır. Kelimelerin eğitim setine benzerlikleri, şartlı olasılıkları dikkate alınarak hesaplanır (Cebeci, 2020).

**Lojistik Regresyon:** Bağımsız değişkenlere dayalı ikili bir sonucu tahmin eden bir sınıflandırma algoritmasıdır. Bir girdi değerini bir ağırlık değeriyle çarparak çalışır. Sınıflandırma görevleri için kullanılan olasılıksal bir regresyon analizidir (Wankhade vd., 2022).

**Destek Vektör Makineleri:** Verileri analiz etmek ve karar limitlerini belirlemek için hiper düzlemleri kullanılır. Sınıflandırma görevleri için başarıyı yüksek ve sıklıkla kullanılan bir olasılık dışı denetimli öğrenme tekniği türüdür (Cebeci, 2020). Birincil amaç, verileri farklı sınıflara en iyi şekilde ayıran hiper düzlemi belirlemektir ve doğrusal hesaplamalar ile sınıflara atama işlemi gerçekleştirilir (Wankhade vd., 2022).

**Karar Ağaçları:** Metnin kutupsallığını sınıflandırmak için eğitim örneği kullanılarak bir ağacın oluşturulduğu denetimli bir öğrenme tekniğidir. Köke yerleştirilmesi gereken en uygun özniteliği belirlemek için CART, ID3, C5.0, C4.5 gibi algoritmalar kullanılabilir. Rastgele Orman (Random Forest) yöntemi, doğruluğu artırmak için birden fazla karar ağacını birleştirir ve genellikle Karar ağaçlarından daha sık kullanılır (Wankhade vd., 2022).

**Derin Öğrenme:** Derin öğrenme, sinir ağının gizli katmanlarına çok katmanlı bir yaklaşım uyarlayan makine öğrenimin bir alt kümesidir. Bu sinir ağları, insan beyni işlevselliğinden ilham alır ve büyük miktarda veriden öğrenme gerçekleşmesine olanak tanır. Ham verileri alan bir dizi girdiye sahip olan sinir ağı katmanları aracılığıyla bilgileri sınıflandırır. Derin öğrenme algoritmaları, DA doğruluğunda ve verimliliğinde önemli artışlara yol açmıştır (Thematic). Geleneksel sınıflandırma algoritmalarına göre önemli avantajlara sahiptir (Kahraman vd., 2022). Özniteliklerin otomatik olarak öğrenilip çıkarılmasıyla daha iyi doğruluk ve performans elde edilir (Dang vd., 2020). Yapay sinir ağları ve derin öğrenme, günümüzde DDİ'nin yanı sıra görüntü ve konuşma tanıma alanlarındaki birçok soruna iyi çözümler sunmaktadır (Dang vd., 2020).

Derin öğrenme teknikleri arasında Derin Sinir Ağları (Deep Neural Networks - DNN), Evrişimli Sinir Ağları (Convolutional Neural Network - CNN) ve Tekrarlayan Sinir Ağları (Recurrent Neural Network - RNN) gibi teknikler bulunur. DNN modelinin bir girdi katmanı, çıktı katmanı ve birden çok gizli katmanı vardır. Çok katmanlı bir ileri beslemeli sinir ağı yapısı kullanır (Kahraman vd., 2021). CNN, Derin öğrenme alanında en ünlü ve yaygın olarak kullanılan algoritmalarındandır ve öncüllerine kıyasla temel faydası, herhangi bir insan denetimi olmadan ilgili öznitelikleri otomatik olarak tanımlamasıdır (Alzubaidi vd., 2021). Geleneksel sinir ağlarının aksine, RNN önceki bilgi hesaplamasını hatırlayabilir ve onu girdi dizisindeki bir sonraki öğeye uygulayarak yeniden kullanabilir (Dang vd., 2020). Özel bir RNN türü ise gizli katmandaki aktivasyon fonksiyonlarının girişi olarak uzun belleği kullanabilen Uzun-Kısa Süreli Bellek (LSTM)

ağıdır (Dang vd., 2020). LSTM'ler çok başarılı sonuçlar elde etse de özellikle uzun cümleler söz konusu olduğunda sınırlamaları vardır ve tekrarlayan bir temel yapıya sahip olduğundan dolayı öğrenme sürecini paralel hale getirmek zordur (Thematic, Gillioz vd., 2020).

2017'de dikkat katmanlarına dayalı Transformer mimarisi önerilmiştir (Vaswani vd., 2017). Cümlenin uzun menzilli bağımlılıklarını yakalama ve paralelleştirilebilir olma avantajına sahiptir (Gillioz vd., 2020). DDİ alanında hızla popüler hale gelerek LSTM ile Bi-LSTM kullanımının yerini almıştır (Wankhade vd., 2022). Transformer mimarisine dayanan GPT ve BERT gibi modeller, birçok önceki ağlardan daha iyi performans sergilemişlerdir (Gillioz vd., 2020).

### **2.3.2.3. Hibrit Yöntemler**

Hem makine öğrenmesi hem de indeks tabanlı yaklaşımları birleştiren hibrit DA algoritmaları da bulunmaktadır (Wankhade vd., 2022;Chang vd., 2020).

## **2.4. Duygu Analizi Uygulamaları, Önem ve Faydaları**

DA, yapılandırılmamış verileri anlamlı bilgilere dönüştürür. Sosyal medya analizi, sohbet robotları, çağrı merkezi verileri, anket yanıtları vb. ile toplanan yapılandırılmamış veriler, DA ile elde edilen görümler ile birleştirilerek, bir organizasyonun işlerini doğru bir şekilde değiştirmesine ve dönüştürmesine yardımcı olur. Daha rekabetçi olma, yeni müşteriler kazanma, mevcut müşterileri elde tutma, pazarlama stratejileri ve kampanyalar oluşturmaya yönelik stratejik önlemler alınmasıyla fayda sağlar.

DA'nın iş zekâsı alanındaki en tipik kullanımı, müşterilerin hizmet veya ürünlerle ilgili izlenimlerini analiz etmektir. Ancak bu çalışmalar üreticiler ile sınırlı olmayıp tüketiciler de bunları incelemek ve daha bilinçli kararlar vermek için kullanılabilir (Wankhade vd., 2022).

Aşağıda duygu analizinin özellikle iş analizi konusunda bazı uygulamalarına değinilmiştir.

### ***Ürün Tasarımı ve İyileştirme***

DA sayesinde kullanıcıların geri bildirimleri işlenerek ürün ve hizmetlere karşılık ne düşündükleri anlaşılabilir, şikayetler yakalanabilir ve bu sayede sorunların düzeltilmesini ve ürünlerin ihtiyaçlara yönelik

geliştirilmesini sağlayabilir. Örneğin; bir uygulamanın yeni arayüzü ile ilgili yapılan aşağıdaki iki yorumu ele alabiliriz.

“Sade ve kullanıcı dostu bir arayüz. Menü dizilimi harika.” 😊

“Yeni arayüz çok kötü. Hem göz alıyor hem de aradığımızı bulamıyorsunuz.” 😞

### **Memnuniyetin Ölçülmesi**

Organizasyonlar kullanıcıları tarafından birçok çevrimiçi platformlarda değerlendirilmekte ve bu değerlere bağlı olarak sayısal bir puan veya yıldızlı derecelendirmelere sahip olmaktadır. Bu derecelerin yanı sıra çevrimiçi incelemelerin ele alınması organizasyonların güçlü ve zayıf yönlerini tespit etmelerine olanak sağlar. Örneğin bir firma kendine ait çok sayıda müşteri/kullanıcı değerlendirmesini otomatik olarak analiz ederek en olumsuz duyguların taşıma ve ürünlerin gecikmesinden kaynaklandığını tespit edebilir ve buna bağlı olarak lojistik yönetimini iyileştirme kararı alabilir. Literatürde turizm, ulaşım, sağlık vb. birçok sektörde DA çalışmalarının örnekleri hızla artmaktadır. Mathayomchan ve Taecharungroj restoran özelliklerinin etkilerini ve genel müşteri deneyimini etkileyen altta yatan faktörleri incelemek amacıyla restoranların çevrimiçi incelemelerini analiz etmişlerdir (Mathayomchan ve Taecharungroj, 2020). 5010 restoranın 935.386 Google Haritalar incelemesini yemek, servis, atmosfer ve fiyat gibi dört temel unsur üzerinden ele almışlardır. Bu özellikler ve 5 yıldızlı derecelendirme arasındaki ilişkileri test edilmiş ve farklı yıldız derecelendirme seviyelerindeki oran değişiklikleri karşılaştırılmıştır. Kwon ve diğerleri Skytrax paylaşımlarındaki 27 havayolunun 14.000’den fazla çevrimiçi değerlendirmeleri üzerinde konu modelleme ve DA çalışması gerçekleştirmişlerdir (Kwon vd., 2021). Uçuşta koltuk, servis ve yiyecek temalarının önemli konulardan olduğunu ve müşteri memnuniyetsizliğini etkileyecek ana sorunun gecikmeler ve zamanlama ile ilgili olduğunu öne sürmüşlerdir. Çevrimiçi incelemelerden organizasyonların hangi alanlarda iyileştirmelere ihtiyaç duyduğuna ve hangi alanlara yatırım yapmaları gerektiğine dair, veri odaklı stratejiler oluşturulabilir.

### **Marka İtibarı**

Kullanıcılarının veya müşterilerinin arzu, beklenti, ilgilerini belirli bir sorumlulukla ele alabilen organizasyonlar DD süreçlerinde başarılı olup gelecekte de büyüyeceklerdir. Markalar, şirketlerin varlık sebeplerinin ve değer sistemlerinin başlıca sembollerindedir (Kaptanoğlu vd., 2019). Marka

aynı zamanda duygu içeren bir kavram olmakla beraber birçok marka tanımı bu ifadeye yer vermektedir:

“Markanız, siz odada olmadığınız zaman hakkınızda söylenenlerdir.” — Jeff Bezos

“Bir misyon, logo, slogan, sembol, ürün ya da servis, reklam veya pazarlama faaliyeti olmanın ötesinde, hedef kitle ve ürün arasında kurulan ilişki, her seferinde hedef kitlelerle kurulan soyut ve duygusal bağ, bir kalite güvencesi ve tutkudur.” (Borça, 2004)

Bu tanımlardan yola çıkıldığında marka, satın alma ve kullanma sonrasındaki duyguların birer ifadesidir ve itibarınız, insanların sizin hakkınızda nasıl hissettikleridir (Somaklar, 2006; Kaur vd., 2021). Bir markanın kullanıcıları veya tüketicilerinin zihinlerinde edindiği yer, düşünceler ve kazandığı bu değeri sürekli kılabilirdiği zamanın önemi anlaşılmaktadır.

Başta sosyal medya akışlarının patlayıcı yayılımı nedeniyle, kullanıcılar yeni dijital pazarlama kanalları tarafından sağlanan özelliklerden yararlanarak markaların itibarını ve işletmelerin kalitesini değerlendirebilmektedir. Kullanıcı nezdinde itibar arttırımı için büyük miktarda dijital verinin tespit ve analizi gerekecektir.

Özellikle sosyal medyada doğruluğu araştırılmadan hızla yayılan kötü propagandalar, markaların değer yitirmesinde çok etkili olmaya başlamıştır (Altındal, 2013) . DA, marka itibarı izlemelerine olanak tanıyarak gönderilerin arkasındaki duyguları ön plana çıkarır. Şirketler, ürünleri, hizmetleri ve markaları hakkında değerli ve ileriye yönelik stratejik bilgiler elde eder. Böylece, belirli kitlelerin markaya olan duyarlılığının tespit edilmesi, markayla ilgili konuların öneme göre derecelendirilmesi, marka itibarının zamana yönelik gelişiminin anlaşılması, rakiplerin gelişimini anlama, belirli zamanlara yönelik marka ile ilgili duyguların analiz edilmesi ve bir kampanyanın veya yeni bir ürün lansmanının genel marka duyarlılığını nasıl etkilediğini değerlendirmesi gibi birçok yönden yardımcı olur.

### ***Sosyal medya takibi***

Günümüzde yeni ve mevcut müşterilerle etkileşim kurmanın en etkili yollarından biri sosyal medyadır. Organizasyonlar hakkındaki iyi yorumlar ve sosyal medya gönderileri, diğer potansiyel kişileri teşvik eder (Thematic). Ayrıca olumsuz sosyal medya gönderileri veya incelemeleri aynı şekilde organizasyon için çok büyük kayıplara neden olabilir. Sosyal medyada öfkeli

bir şekilde bildirim yapan bir kullanıcı, memnun bir müşteriden gelen bir teşekkür bildiriminden daha önemli olabilir (Saju vd., 2020). Sosyal medya verileri üzerinde DA, kullanıcı duygularını 7/24 gerçek zamanlı olarak izlemeye ve durumların hızla ele alınmasına imkan sunar (Wankhade vd., 2022). DA ile bu gibi problemler henüz büyümeden önce gerçek zamanlı olarak belirlenebilir ve işletmeler marka itibarlarına yönelik önlemler alabilir.

### ***Pazar araştırması***

Pazar araştırması en yaygın DA uygulamalarından biridir (Wankhade vd., 2022). DA, yeni ortaya çıkan eğilimlerin belirlenmesi, rakiplerin analiz edilmesi ve yeni pazarların araştırılması konularında etkilidir. Yeni bir ürün piyasaya sürülmesi planlanıyorsa farklı platformlardan toplanan veriler incelenerek hedeflenen pazar segmentindeki beklenti ve gereksinimler anlaşılabilir (Saju vd., 2020). Ayrıca, kullanıcıların rakip organizasyonların faaliyetleri ve ürünleri hakkında olumlu/olumsuz buldukları yönler tespit edilebilir. Bu bilgiler, rekabet avantajı elde edilmesinde ve ileriye yönelik kararlarda avantaj sağlar. Örneğin; rakip organizasyon müşterilerinin bataryaların aşırı ısınmalarından dolayı şikayetçi olduğunu tespit eden bir şirket, sonraki pazarlama stratejilerinde kendi üstün soğutma sistemlerini ön plana çıkarabilir.

### ***Zaman Tasarrufu***

Duygu analizine yönelik yaklaşımlar sürekli olarak geliştirilmekte ve daha performanslı hale gelmektedir. DA yöntemleri ile büyük boyutlardaki metin verileri çok kısa süreler içerisinde analiz edilebilir. DA sayesinde elde edilen bildirimler, hızlı ve daha az bir çaba ile elde edildiğinde doğru stratejiler oluşturmak ve zamanında karar almak için vakit elde edilmesini sağlayacaktır.

### ***Hızlı Değerlendirme***

Gerçek zamanlı bildirimler, organizasyonlar için süratle ele alınması gereken sorunların belirlenmesinde oldukça faydalı olacaktır. Örneğin; sosyal medyada ön plana çıkan olumsuz bir bildirim, anında ele alınabilir ve hızlı bir şekilde incelenebilir. Olumsuz bildirim yapılan bir konu hakkında diğer kullanıcıların da aynı sorunu yaşama olasılığına yönelik organizasyonlar çözüm yollarını hızlı bir şekilde değerlendirebilir ve doğru yönlendirmeler ile ilgili birimler anında uyarılarak kötü deneyimlerin yaşanması önlenir.



## 2.5. Duygu Analizi için Mevcut Zorluklar

DA gerçekleştirilirken hesaplama maliyetleri, alan bağımlılığı, insanların doğal dil kullanımından kaynaklanan çeşitli zorluklar, dil farklılıkları gibi problemlerle karşılaşılır.

Duygu analizinde karşılaşılan birkaç önemli zorluk şunlardır:

**Belirsiz İfadeler:** Belirsiz ifadeler, doğal dil metinlerinde yaygın bir olgudur. Anlamsal belirsizlikler karışıklıklara yol açabilir. Belirsiz ifadelerin doğru ayrıştırılması DDİ için zorlu bir görevdir. Örneğin “El” kelimesi birçok farklı anlamlara gelebilmektedir (Kolun bir bölümü, sahiplik, kez, oyun sırası ve tur, nesne ve araçları tutmaya yarayan bölüm, yabancı, ülke-yurt-il-halk-ahali) ve kastedilen mananın bilinmesi gerekir. Ayrıca birçok sözdizimsel belirsizlik olgusu vardır. Örneğin; “Kitap okumayı benden çok seviyor” gibi bir ifadede karşılaştırılan iki kişinin kitap okuma sevgisi mi yoksa kişi ile kitap okuma mı olduğu belirsizdir.

**Öznellik:** DA gerçekleştirilirken öznel yorumların dikkatli bir şekilde ele alınması gerekir. Örneğin; “Tasarım rengi çok iyi” gibi bir yorum olumlu olarak değerlendirilebilirken “Tasarım rengi siyah” gibi bir ifade kişiye özel olacaktır.

**Resmi olmayan ifadeler:** Gayri resmi yazı stili DA için en büyük zorluklardan biridir (Wankhade vd., 2022). Metinlerinde anlaşılması zor olan kısaltmalar çok sık olarak kullanılmakta ve birçok bölgesel kısaltmalar da gittikçe artmaktadır.

**Dilbilgisi ve yazım hataları:** Dilbilgisi ve yazım hataları özellikle resmi olmayan sosyal medya iletileri gibi metinlerde çok yaygındır. DA görevlerinin doğruluğu için bu hataların ele alınıp düzeltilmesi gerekecektir.

**Mizah ve Alay etme:** Mizah ve alay ifadeleri ile kişiler genellikle olumsuz deneyimleri tanımlamak için olumlu kelimeler kullanırlar. Gerçekte övgü gibi görünse de asıl itibarı ile yergi içerirler. Alay ifadeleri kullanılan ton, mevcut arka plan bilgisi vb. gibi pek çok faktörden etkilendiğinden dolayı da makinelerin alayı algılaması zordur (Wankhade vd., 2022). Bununla birlikte, mizah da kültürel olarak çok özel olduğundan ve bir makinenin benzersiz kültürel imaları anlaması zor olduğundan çözümü zordur (Wankhade vd., 2022). Örneğin bir kargo şirketi hakkında “Paketleme çok güzeldi sallana sallana gelmiş. Teşekkürler.” gibi bir ifadeyi ve makinelerin buradaki duyguyu anlaması zor olabilir. ‘Güzel’ ve ‘Teşekkürler’ gibi kelimeler duygu açısından olumlu olsa da genel duygu olumsuzdur.

**Karşılaştırmalar:** Karşılaştırma, duygu sınıflandırmasını ele almak için başka bir potansiyel engeldir. Karşılaştırılan nesnenin özellikleri bilinmeden sınıflandırma görevini gerçekleştirmek zor olabilir. Örneğin “Alternatiflerine göre daha iyi” şeklinde bir yorum ele alındığında alternatif kelimesi ile vurgulanan nesnenin iyi veya kötü mü olduğuna ait duygu da ifade edilen yorumun sınıflandırmasını etkileyecektir.

**Emojiler:** Emojiler, özellikle sosyal medya platformları gibi veri kaynaklarını kullanırken kapsamlı bir ön işleme gerektirebilir. Benzer şekilde, emojilerin farklı versiyonları örneğin farklı gülme emojileri bir duygunun farklı yoğunluğunu ifade edebilir. 😊(ツ)☺

**Deyimler:** Deyimler kullanılan algoritmalarda karışıklıklara yol açabilir. Örneğin; yeni bölümü nasıl buldunuz gibi bir soruya verilen “Ağzım açık kaldı” gibi bir cevabın sınıflandırılması düşünülebilir.

**Tarafsızlık:** Tarafsız etiketini uygun şekilde tanımlamak önemlidir. Kaliteli ve güvenilir analiz oluşturmak için kriterlerin tutarlı olması gerekir (Thematic). Bazı durumlarda, tarafsız duygu da dahil edilir fakat belirsizlikleri ve bilgi eksikliği nedeniyle birçok DA görevinde tarafsız değerlendirmeler sıklıkla göz ardı edilir (Wankhade vd., 2022).

**Görsel-İşitsel İçerikler:** Video, görüntü ve ses verileri metinlerden farklı yapıda oldukları için uygun şekilde ele alınmaları gerekir.

**İnsan değerlendirmesi kaynaklı sınırlamalar:** Bazı ifadelerin insanlar tarafından bile tanımlanması zordur. Alan özgü uzman bilgisi ve çoğu durumda birden fazla kontrolcünün tutarlılığının ölçülmesi gerekebilir.

**Olumsuzlamalar:** Olumsuzlamalar görüşün kutupluluğunu tersine çevirebilen ve bir cümlenin anlamını değiştirebilen kelimelerdir. Olumsuzluk sözcükleri DA modelleri için zorluk çıkarabilir ve çok dikkatli bir şekilde ele alınmaları gerekir. Olumsuzlama sözcükleri belirli bir metnin polaritesini tersine çevirebildikleri için DA için kritik öneme sahiptir. “Ürün güzel” gibi bir ifade olumlu bir şekilde ele alınırken “Ürün güzel değil” olumsuz bir ifadedir. “Kullanışlı” ve “uygun” gibi kelimeler genel itibari ile olumlu duygu belirtirken “ürün ne uygun ne de kullanışlı” gibi bir ifade tüm duyguyu tersine çevirebilir. Bazı durumlarda durdurma kelime listelerine dahil edildikleri veya bir sözlükte tarafsız bir duygu değerine sahip oldukları ve mutlak kutupsallığı etkilemedikleri için olumsuzlama kelimeleri kaldırılır (Wankhade vd., 2022). Ayrıca, cümlede geçen her sözcük kutupluluğu tersine çevirmez; bu nedenle de durdurma sözcüklerinden tüm olumsuzlama sözcüklerinin çıkarılması, hesaplama maliyetini artırabilir ve modelin doğruluğunu da azaltabilir (Wankhade vd., 2022). LSTM veya ön eğitilmiş

Transformer modelleri bu gibi durumların etkili ve doğru bir şekilde ele alınmasında kullanılabilir modellerdendir (Thematic).

**Veri eldesi ve kullanılabilirliği:** Geleneksel istatistiksel makine öğrenimi yöntemi ve derin öğrenme yöntemleri de dahil olmak üzere, büyük ölçekli veri kümelerine dayanan makine öğrenimi yöntemleridir (Kaur vd., 2021). Ancak büyük ölçekli örnekleri toplamak ve açıklama eklemek genellikle zor bir iştir. Ayrıca aşırı öğrenme sorunlarının da dikkate alınması gerekir. Yayının açılmasını ve paylaşılmasını engelleyen telif hakkı veya gizlilik sorunları nedeniyle de çok fazla çevrimiçi içerik elde etmek zordur (Zong vd., 2021). Veriler elde edilse bile çoğunlukla tek tip bir formattan yoksundurlar. Veriler genellikle belirli bir alana aittir ve bir alanın eğitim verileri, diğer alanlar için geçerli ve değerli de olmayabilir. Alan değişirse, veri toplama, işleme ve açıklama çalışmaları yeniden başlamak zorunda kalabilir (Zong vd., 2021).

**Sahte ve yanlış bilgiler:** Sahte incelemeler veya botlar tarafından oluşturulan içerikler de büyüyen bir endişe kaynağıdır. Bu tür içerikleri ve analiz sonuçlarını saptırmalarını belirlemek zor olabilir.

## SONUÇ

Günümüzde birçok endüstri ve kuruluş giderek daha fazla miktarda veri üretilmesiyle sonuçlanan bir tür dijital dönüşümden geçmektedir. Karşılaşılan en büyük zorluklardan biri kullanıma uygun olmayan durumdaki yapılandırılmamış verileri, veriye dayalı kararlar alınmasında ve operasyonel verimlilik, değer iyileştirme ve genel rekabet avantajı oluşturma da yardımcı olabilecek faydalı kazanımlara dönüştürmektir. DA, DD süreçlerindeki veya başlama niyetindeki herhangi bir kuruluş için çok önemli tekniklerden biridir. Sunulan ürün veya hizmetlerle ilgili olumsuz veya olumlu duyguların hassas bir şekilde tespit edilmelerini ve gerekli adımların atılması için ihtiyaç duyulan görüşü sağlar. Organizasyon içinde değişiklik gereken alanların belirlenmesi, dönüşüm için hangi araçlara ihtiyaç duyulabileceğini anlamak, mevcut durumu göz önünde bulundurarak bir sonraki eylemleri daha iyi planlamak için DA yardımcı olacaktır. DD hedeflerine yardımcı olabilecek görüşleri ve tepkileri takip etmek amacıyla sosyal medya gönderilerini ve çevrimiçi incelemeleri analiz etmek için DA uygulanabilir. Pazar araştırması, marka ve itibar takibi, müşteri deneyimi analizi için kullanışlı bir araçtır. Piyasada hayatta kalmak, kullanıcılara ve müşterilere sunulan ürünlerden, hizmetlerden memnuniyetleri ölçülerek gerçekleştirilebilir. Farklı kaynaklardan verilerin elde edilmesi ve DD için

uygulanabilir hale getirilmelerinde de DA'dan faydalanılabilir. DA yöntemleri sürekli gelişen ve hızlı hareket eden bir alandır bu nedenle de son eğilimleri ve teknolojileri takip etmek son derece önemlidir.

## KAYNAKLAR

- Abdullah, M. F., & Ahmad, K. (2015, August). Business intelligence model for unstructured data management. In *2015 International Conference on Electrical Engineering and Informatics (ICEEI)* (pp. 473-477). IEEE.
- Adomavicius, G., & Kwon, Y. (2011). Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering*, 24(5), 896-911.
- Ahmad, S. R., Bakar, A. A., & Yaakub, M. R. (2019). A review of feature selection techniques in sentiment analysis. *Intelligent data analysis*, 23(1), 159-189.
- Aktan, E. (2018). Büyük veri: Uygulama alanları, analitiği ve güvenlik boyutu. *Bilgi Yönetimi*, 1(1), 1-22.
- Altındal, M. (2013). Dijital pazarlamada marka yönetimi ve sosyal medyanın etkileri. *Akademik Bilişim Konferansı, Akdeniz Üniversitesi*, 23, 25.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, 8(1), 1-74.
- Buneman, P. (1997, May). Semistructured data. In *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems* (pp. 117-121).
- Bonta, V., & Janardhan, N. K. N. (2019). A comprehensive study on lexicon based approaches for sentiment analysis. *Asian Journal of Computer Science and Technology*, 8(S2), 1-6.
- Borça, G. (2004). İnsanları Daha Çok Para Ödeyip Markalı Ürünleri Almaya İten Nedir? Erişim Adresi: <http://www.guvenborca.com/guven-borca-kulliyati/makaleler/insanlari-daha-cok-para-odeyip-markali-urunleri-alamaya-iten-nedir-01-01-2004/>
- Cebeci, H.İ. (2020). Sosyal medya verileri ile duygu analizi. In Torkul, O., Gülseçen, S., Çağıl, G., Uçar, M. K., İnal, M. M., Erol, Ç., ... & Selvi, İ. H. (Eds.). *Mühendislikte yapay zeka ve uygulamaları 3.* (pp.191-211). Sakarya Üniversitesi Yayınevi.

- Chang, J. R., Liang, H. Y., Chen, L. S., & Chang, C. W. (2020). Novel feature selection approaches for improving the performance of sentiment classification. *Journal of Ambient Intelligence and Humanized Computing*, 1-14.
- Çetin, F. S., & Eryiğit, G. (2018). Türkçe hedef tabanlı duygu analizi için alt görevlerin incelenmesi–hedef terim, hedef kategori ve duygu sınıfı belirleme. *Bilişim Teknolojileri Dergisi*, 11(1), 43-56.
- Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, 9(3), 483.
- Doan, A., Naughton, J., Baid, A., Chai, X., Chen, F., Chen, T., ... & Vuong, B. Q. (2009). The case for a structured approach to managing unstructured data. *arXiv preprint arXiv:0909.1783*.
- Doğan, K., & Arslantekin, S. (2016). Büyük veri: önemi, yapısı ve günümüzdeki durum. *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi*, 56(1), 15-36.
- Eberendu, A. C. (2016). Unstructured Data: an overview of the data of Big Data. *International Journal of Computer Trends and Technology*, 38(1), 46-50.
- Gillioz, A., Casas, J., Mugellini, E., & Abou Khaled, O. (2020, September). Overview of the Transformer-based Models for NLP Tasks. In *2020 15th Conference on Computer Science and Information Systems (FedCSIS)* (pp. 179-183). IEEE.
- Global AI Adoption Index 2021 (2021). Erişim Adresi: [https://filecache.mediaroom.com/mr5mr\\_ibmnews/190846/IBM%27s%20Global%20AI%20Adoption%20Index%202021\\_Executive-Summary.pdf](https://filecache.mediaroom.com/mr5mr_ibmnews/190846/IBM%27s%20Global%20AI%20Adoption%20Index%202021_Executive-Summary.pdf)
- Gupta, N., & Agrawal, R. (2020). Application and techniques of opinion mining. In *Hybrid Computational Intelligence* (pp. 1-23). Academic Press.
- Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, 17, 26-32. <https://doi.org/10.1016/j.procs.2013.05.005>
- Harbert, T. (2021). Tapping the power of unstructured data. *MIT Sloan. Feb, 1, 3*. Erişim Adresi: <https://mitsloan.mit.edu/ideas-made-to-matter/tapping-power-unstructured-data>

- Hänig, C., Schierle, M., & Trabold, D. (2010, October). Comparison of structured vs. unstructured data for industrial quality analysis. In *Proceedings of the World Congress on Engineering and Computer Science* (Vol. 1, pp. 20-22).
- IBM Global AI Adoption Index 2022 (2022). Erişim Adresi: <https://www.ibm.com/watson/resources/ai-adoption>
- Imani, M. B., Keyvanpour, M. R., & Azmi, R. (2013). A novel embedded feature selection method: a comparative study in the application of text categorization. *Applied Artificial Intelligence*, 27(5), 408-427.
- Kahraman, A., Kantardzic, M., Kahraman, M. M., & Kotan, M. (2021). A data-driven multi-regime approach for predicting energy consumption. *Energies*, 14(20), 6763.
- Kahraman, A., Kantardzic, M., & Kotan, M. (2022). Dynamic Modeling With Integrated Concept Drift Detection for Predicting Real-Time Energy Consumption of Industrial Machines. *IEEE Access*, 10, 104622-104635.
- Kaptanoğlu, R. Ö., KILIÇARSLAN, M., & Tosun, A. (2019). Marka ve marka farkındalığı. *The Journal of Social Science*, 3(5), 248-266.
- Kaur, N., Sahdev, S. L., Singh, G., & Tokas, N. (2021, January). Monetizing personal brand for business success, financial security and career longevity-sentiment analysis in COVID 19 Era. In 2021 2nd International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 321-325). IEEE.
- Kwon, H. J., Ban, H. J., Jun, J. K., & Kim, H. S. (2021). Topic modeling and sentiment analysis of online review for airlines. *Information*, 12(2), 78.
- Mathayomchan, B., & Taecharungroj, V. (2020). "How was your meal?" Examining customer experience using Google maps reviews. *International Journal of Hospitality Management*, 90, 102641.
- Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. *Business & information systems engineering*, 57(5), 339-343.
- Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 1-19.

- Osimo, D., & Mureddu, F. (2012). Research challenge on opinion mining and sentiment analysis. *Universite de Paris-Sud, Laboratoire LIMSI-CNRS, Bâtiment, 508*.
- Pham, D. H., & Le, A. C. (2018). Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis. *International Journal of Approximate Reasoning*, 103, 1-10.
- Pintas, J. T., Fernandes, L. A., & Garcia, A. C. B. (2021). Feature selection methods for text classification: a systematic literature review. *Artificial Intelligence Review*, 54(8), 6149-6200.
- Poria, S., Hussain, A., & Cambria, E. (2018). *Multimodal sentiment analysis (Vol. 8)*. Cham, Switzerland: Springer.
- Sağlam, M. (2021). İşletmelerde geleceğin vizyonu olarak dijital dönüşümün gerçekleştirilmesi ve dijital dönüşüm ölçeğinin türkçe uyarlaması. *İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi*, 20(40), 395-420.
- Saju, B., Jose, S., & Antony, A. (2020, July). Comprehensive Study on Sentiment Analysis: Types, Approaches, Recent Applications, Tools and APIs. In *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)* (pp. 186-193). IEEE.
- Schouten K, Frasinca F (2015) Survey on aspect-level sentiment analysis. *IEEE Trans Knowl Data Eng* 28(3):813–830
- Sebastian, I. M., Ross, J. W., Beath, C., Mocker, M., Moloney, K. G., & Fonstad, N. O. (2020). How big old companies navigate digital transformation. In *Strategic information management* (pp. 133-150). Routledge.
- Sharda, R., Delen, D., Turban, E., Aronson, J., & Liang, T. (2014). Business intelligence and analytics. *System for Decesion Support*.
- Somaklar, F. Ö. (2006). İşletmelerde marka yönetimi süreci ve bir uygulama (Doctoral dissertation, DEÜ Sosyal Bilimleri Enstitüsü).
- Somekh, S. (2021, July 23). The role of sentiment analysis in digital transformation. Retrieved November 15, 2022, Erişim Adresi: <https://www.revuze.it/blog/the-role-of-sentiment-analysis-in-digital-transformation/>
- Thematic. Sentiment Analysis: Comprehensive Beginners Guide. Erişim Adresi: <https://getthematic.com/sentiment-analysis/>

- Tian Xie, Xinyi Yang, Angela S Lin, Feihong Wu, Kazuma Hashimoto, Jin Qu, Young Mo Kang, Wenpeng Yin, Huan Wang, Semih Yavuz, et al. 2022. Converse—a tree-based modular task-oriented dialogue system. arXiv preprint arXiv:2203.12187
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Verina, N., & Titko, J. (2019, May). Digital transformation: conceptual framework. In *Proc. of the Int. Scientific Conference “Contemporary Issues in Business, Management and Economics Engineering’2019”*, Vilnius, Lithuania (pp. 9-10).
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 1-50.
- Zhang, D., Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., ... & Perrault, R. (2021). The AI index 2021 annual report. *arXiv preprint arXiv:2103.06312*.
- Zong, C., Xia, R., & Zhang, J. (2021). *Text Data Mining* (Vol. 711, p. 712). Singapore: Springer.