



Metaverse ile ilgili Türkçe Dilindeki Çeşitli Sosyal Medya Platformu Verileri ile Duygu Analizi

Ulaş Naki TURAN^{*a}, İlkim Ecem EMRE^b Selçuk KIRAN^b

^a* Globit Global Bilgi Teknolojileri A.Ş.

^b Marmara Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri Bölümü, İstanbul, 34722, TÜRKİYE

MAKALE BİLGİSİ

Alınma: 23.08.2022
Kabul: 03.11.2022

Anahtar Kelimeler:

Duygu Analizi, NLP,
Metin Madenciliği,
Makine Öğrenmesi,
Metaverse, Youtube,
Twitter

Sorumlu Yazar

e-posta:
ecem.emre@marmara.edu.tr

ÖZET

Kısa zamanda popüler bir kavram olmayı başaran metaverse kavramı hakkında günümüzde olumlu veya olumsuz birçok yorum ve düşünce bulunmaktadır. Böylesine yeni bir girişim birçok insanı heyecanladırsa da birçok insanı da çekimser kılmaktadır. Bu çalışmada da metaverse hakkında Youtube ve Twitter olmak üzere iki sosyal medya platformundan elde edilen Türkçe veriler ile bir metin madenciliği çalışması yapılmıştır. Elde edilen veriler ilk önce gönüllü insanlar vasıtasıyla olumlu ve olumsuz olarak etiketlenmiş sonrasında ise her kaynaktan eşit sayıda veri olacak şekilde ayrıştırılıp, birleştirilmiştir. Bozuk ve nötr duyguya sahip veriler ayrıştırılan veriler arasındadır. Platform bazında incelendiğinde Youtube yorumlarında daha fazla sayıda olumlu görüş belirten içerik olduğu görülmüştür. Twitter'da ise olumsuz görüşleri içeren içerik sayısı daha fazladır. Analiz bölümünde makine öğrenmesi algoritmaları olarak naive Bayes, lojistik regresyon, destek vektör makineleri ve rassal orman sınıflandırıcıları kullanılmıştır. 1350 adet Youtube, 1350 adet de Twitter olmak üzere toplamda 2700 adet veriyle yapılan analiz sonucunda uygulanan bütün sınıflandırma algoritmaları yüzde seksen üzeri bir başarı göstermiş ve %88 ile naive Bayes en başarılı algoritma olmuştur. Bu yüzde değeri literatürdeki benzer çalışmalar ile kıyaslandığında en düşük değer olarak ortaya çıkmaktadır.

Sentiment Analysis with Various Social Media Platform Data in Turkish Language Related to Metaverse

ARTICLE INFO

Received: 23.08.2022
Accepted: 03.11.2022

Keywords:

Sentiment Analysis,
NLP, Text Mining,
Machine Learning,
Metaverse, Youtube,
Twitter

Corresponding Author

e-mail:
ecem.emre@marmara.edu.tr

ABSTRACT

There are many comments and thoughts, positive or negative, about the concept of Metaverse, which has become a popular concept in a short space of time. Although such a new initiative excites many people, it also makes many people hesitate. In this study, a text mining study was conducted on Metaverse using the Turkish data obtained from two social media platforms, Youtube and Twitter. The obtained data were first labeled as positive and negative by volunteers, and then they were separated and combined to have an equal number of data from each source. Also, data with distorted and neutral emotion are among the eliminated data. When examined on a platform basis, it was seen that there were more positive opinions in Youtube comments. On Twitter, the number of content containing negative opinions is higher. In the analysis part, Naive Bayes, Logistic Regression, Support Vector Machines and Random Forest classifiers were used as machine learning algorithms. As a result of the analysis made with a total of 2700 data, including 1350 Youtube and 1350 Twitter, all classification algorithms applied with a success of over eighty percent and Naive Bayes was the most successful algorithm with 88%. This percent value emerges as the lowest value when compared to similar studies in the literature.

1. GİRİŞ (INTRODUCTION)

İnsanların olaylar, ürünler, siyasi partiler veya spor takımları hakkında düşünceleri her zaman toplumun yönetici tabakasında bulunan insanlar için ilgi çekici olmuştur. Ülke çapında yapılacak bir genel seçim için siyasi partiler en iyi nasıl hazırlanır ya da bir ürünün piyasaya sürüldüğünde insanlar tarafından yüksek ilgiyle karşılanması nasıl sağlanır? İşte bu ve buna benzer soruların cevaplarını bulmak için günümüzde sayısız yöntem bulunmaktadır. Bu yöntemlerden ilki, günümüzde uygulamasıyla çok sık karşılaştığımız anket yöntemidir. Anket yönteminde çeşitli bölge, kültür ve gruplardan oluşan topluluklara sorular yöneltilir ve sonuçları analiz edilir. Gayet basit bir yöntem olmasına rağmen uygulamada bazı problemleri beraberinde getirebilir. Bu problemlerden en önemlisi soruların yöneltileceği kişilerin sayısıdır. Bilgi çıkarımı yapılmak istenen anket belli bir sayıya ulaşmazsa analiz için tatmin edici kesinlikte çıkarımlar yapmak zorlaşır. İkinci en önemli problem ise doğru kitle seçimidir. Bir anketin güvenilirliğinin en önemli göstergesi, soruların yönetildiği kişilerin farklı grup veya topluluklardan oluşmasıdır. Eğer sorular aynı demografik özelliklere sahip veya tek bir lokasyonda oturan insanlara sorulmuşsa, anketten elde edilen bilgilerden anlamlı bir sonuç çıkarılamaz.

Günümüzde daha farklı, daha ucuz ve daha ergonomik tekniklerden söz etmek mümkündür. Bu tekniklerden biri olan metin madenciliği bugünlerde araştırmacılara ve veri analistlerine fazlasıyla yardımcı olmaktadır. İnternet kökenli olsun veya olmasın her tür dökümandan elde edilen metin verilerini analiz sürecine metin madenciliği denmektedir ve oldukça başarılı bir teknik olduğu söylenebilir. Metin analizinin bir türü olan duygu analizi sosyal medya analizlerinde kullanılan yaygın bir tekniktir ve genellikle bu analizlerde veri kaynağı olarak sosyal medya platformlarından elde edilen veriler ile çalışılmaktadır.

Yeni bir kavram olarak hayatımıza girmeye başlayan metaverse daha şimdiden bir çok insanda olumlu veya olumsuz bir düşünce oluşturmuştur ve hangi duygu türünün daha ağır bastığı bir merak konusudur. Bu doğrultuda bu çalışmanın amacı metaverse kavramı hakkında sosyal medya platformlarında oluşan genel algının tespit edilmesidir. İnsanların duygu ve düşüncelerine ulaşmada API bağlantısı olarak erişim kolaylığı sunan ve birçok duygu analizi çalışmasında da kaynak olarak kullanılan Twitter dünyada en çok ziyaret edilen sosyal medya platformlarından biridir [1]. Youtube ise hem dünyada en popüler 2. web sitesi olması [2] hem de Eyipınar vd. [3]'ün çalışmasında da belirttiği üzere çok dilli, çok alanlı olmasıyla önemli bir içerik paylaşım ve yorumlama merkezidir. Bu

bilgiler ışığında bu çalışmada da veri kaynağı olarak Twitter ve Youtube platformları seçilmiştir. Ayrıca platform seçiminde platformun içerik toplulukları veya blog/mikro blog türüne dahil olması gözetilen kriterlerden biri olmuştur [4]. Bu duygu zenginliği barındırması açısından önemli görülmektedir.

Çalışmada bahsedilen amaç doğrultusunda, hangi sınıflandırma algoritmasının daha iyi bir başarı göstereceği test edilmiştir. Metin madenciliğinin önemli bir yapı taşı olarak sınıflandırma algoritmaları yaklaşımsal olarak farklılıklar göstermektedir. Yapılan birçok çalışmada olasılık yaklaşımı sınıflandırma algoritmalarının metin madenciliğinde çok daha başarılı sonuçlar verdiği gözlenmiştir. Nitekim Pilavcılar [5] tarafından yapılan metin madenciliği çalışmasında naive Bayes en başarılı algoritma olmuştur ve yazar bunun sebebini algoritmanın olasılık hesabı yapmasına dayandırmıştır. Bu çalışmada da birden fazla olasılık yaklaşımı makine öğrenmesi algoritması kullanılarak durumun bu şekilde olup olmadığı gözlemlenmek istenmiştir. Son olarak literatürde metaverse kavramı üzerine Türkçe duygu analizi çalışmalarının sınırlı olması, çalışmanın gerçekleştirilme sebeplerinden biri olmuştur.

Çalışma kapsamında metaverse kavramı ve bu kavramın geleceği hakkında bazı bilgiler verilmiştir. Ek olarak metaverse kavramının etimolojik kökeni hakkında araştırma yapılmıştır. Çalışmanın üçüncü kısmında ise makine öğrenmesi ve metin madenciliği hakkında önemli bilgiler derlenmiş ve alt kavramları ile net bir şekilde ayrıştırılıp, açıklanmıştır. Pratikte sık rastlanan anlam karmaşası özellikle metin madenciliğinde daha fazla görülmektedir. Örneğin metin madenciliği, metin sınıflandırma, doğal dil işleme ve duygu analizi kavramları net bir şekilde birbirlerinden ayrıştırılmaya çalışılmıştır.

Literatür taraması bölümünde de bu alanda yapılmış çalışmalar incelendikten sonra materyal ve metod bölümünün ilk kısmında çalışmada kullanılan verilerin nasıl elde edildiği, hangi kriterlere göre seçildiği anlatılmıştır. Sonraki kısımlarda da sırasıyla veri temizleme, veri etiketleme, kullanılan algoritmalar, başarı göstergeleri ve çapraz doğrulama işlemleri açıklanmıştır.

Bulgular kısmında ilk önce tanımlayıcı istatistikler verilmiş sonrasında ise analiz sonuçlarının karşılaştırma tabloları sunulmuştur. Son olarak bu tablolar tartışma ve sonuç kısmında yorumlandıktan sonra metaverse ve metin madenciliği konularında gelecek öngörülere yapılmıştır.

2. METAVERSE: TANIMLAR, GELECEĞİ (METAVERSE: DEFINITIONS, FUTURE)

İnsanlık yüzyıllar boyunca sayısız keşif ve yeniliğe tanık oldu ve tanık olmaya devam etmektedir. Bu günlerde çokça duyduğumuz metaverse kavramı, internetin sunduğu sunduğu yeniliklerden biridir. Yaygın inanışın aksine metaverse kelimesinin kökeni 30 yıl önceye, yani 1992'ye dayanmaktadır ve ilk olarak Neal Stephenson'ın bilim kurgu romanı Snow Crash'te bu kavramdan bahsedilmiştir [6].

Stephenson'ın [6] daha 1990'lı yıllarda romanında bu kelimeyi kullanması, bugün tanık olunan gelişmelerin çok da şaşırtıcı olmadığını bize göstermektedir. Ayrıca kitabında bu terimden bahsetmesi kadar bu terime verdiği anlam da değerlidir. Bunun sebebi günümüzde metaverse teriminin açıklanmasında bu anlamın büyük rol oynamasıdır. Kitabında "*İnsanlar, avatar denilen yazılım parçalarıydı. Metaevrende insanların birbiriyle iletişim kurmak için kullandığı görsel-işitsel bedenlerdi.*" olarak tanımladığı metaverse terimi bugüne sağlam bir temel oluşturmuştur.

Metaverse kelimesinin günümüzdeki gelişmelerle birlikte kazandığı yeni anlamları açıklamadan önce bu kelimeyi etimolojik açıdan değerlendirmek daha uygun olacaktır. Batı dillerindeki birçok kelime gibi, metaverse kelimesinin de kökleri eski Yunanca'dan gelmektedir. "Evren" ve "Alem" anlamına gelen *universum* kelimesinin son ekinin ve "ötesi", "ileri" ve "sonra" anlamlarına gelen *meta* öneklerinin birleştirilmesiyle elde edilmiştir. Bu da öte evren gibi bir anlam yaratmıştır [7].

Etimolojik olarak öte evren anlamına gelen metaverse, insanların avatar adı verilen sanal bir araç kullanarak birbirleriyle iletişim kurabildikleri ve etkileşimlerin internetin 2 boyutlu düzeyinden farklı olarak 3 boyutlu düzeyde gerçekleştiği, insanlara hareket etme ve eğlenme fırsatı sunan sanal bir dünyadır [8]. İlk olarak 30 yıl önce ortaya çıkan bir kavram olarak metaverse kavramının neden son zamanlarda bu kadar popüler bir konu haline geldiği merak edilen soruların başında gelmektedir. Bu sorunun cevabı günümüzün önemli teknoloji şirketlerinden biri olan "Meta" (eski adıyla Facebook) örneği ile açıklanabilir. 28 Ekim 2021'de Facebook şirket adını değiştirme kararı aldı ve adını Meta olarak değiştirdi. Bunun ardından dünya çapında ilgi odağı haline gelen Facebook, bu hamlesi ile yeni bir çağın başladığını dünyaya ilan etmiştir. Başlangıçta pek çok kişi tarafından anlamı bilinmeyen metaverse, zamanla kendini dünyaya tanıtmış ve giderek daha fazla insan bu yeni girişime ilgi duymaya başlamıştır [9].

Bazı araştırmacılar zamanlama olarak Facebook'un böyle bir atılımı gerçekleştirmesini dünyanın karşı karşıya kaldığı spesifik bir vaka ile ilgisi olduğunu düşünmektedir. Lee [10] ve Kang [11] metaverse atılımının ortaya çıkış zamanının ve hızlı bir popülarite kazanmasının COVID-19 küresel salgınıyla bağlantılı olduğunu düşünen araştırmacılardan bazılarıdır. İki araştırmacı da Facebook'un tüm dünyada fiziksel sosyal yaşantının durma noktasına gelip, dijital ortamda etkileşimin üst seviyelere çıkmasıyla insanların bu girişime sıcak bakabileceği düşüncesinde olduğu kanısındadır.

Metaverse'ün işlevi konusunda Facebook, metaverse evreninin hedeflerini birkaç basın açıklamasında şu şekilde sıralamıştır: İnsanların sevdikleriyle daha fazla birlikte zaman geçirebilmelerini sağlamak; iş, eğitim ve özel topluluk toplantıları gibi çeşitli alanlarda insanların rahat bir şekilde bir araya gelmelerini sağlamak; eğlence sektörünü yanı sıra fiziksel dünyada çok maliyetli olan bilimsel deney ve etkinlikleri mümkün kılmak [8].

İnsanların metaverse hakkında pozitif duyguya sahip olmasındaki en önemli metaverse tasavvurlarından biri olarak da fiziksel dünyadaki engeller olmadan etkileşime geçilecek bir ortam olmasıdır [12]. İnsanların üzerinde olumlu bir duygu oluşmasındaki bir başka katkı ise pandemi ile birlikte profesyonel iş hayatının dijital ortamda gerçekleştirilebileceği fikridir [13]. Ayrıca bu girişimin yeni uzmanlık alanları ve kariyer olanaklarını yaratma potansiyeli Kuş [9]'un çalışmasında dikkat çektiği konulardan biridir. Son 15 yılda blok zinciri (blockchain) (dolaylı olarak kripto para birimleri), sanal gerçeklik gözlükleri ve çip teknolojisinde birçok önemli ilerleme kaydedilmiştir. Pek çok çevreye göre bunlar metaverse teknolojisi için altyapı oluşturmada önemli adımlardır [11]. Geline nokta metaverse popüler bir terim haline gelmiş olsa da, diğer bir yandan, henüz insanlığa tatmin edici düzeyde somut bir gelişme sunamamaktadır. Bu günlerde bazı beta sürümü girişimleri insanlar tarafından ilgiyle karşılanırsa da bunlar şu an için Meta şirketinin hedeflerinden oldukça uzaktır. Ayrıca metaverse girişiminin bireyleri ve toplumları karşı karşıya bırakabileceği potansiyel tehditler de birçok araştırmacı ve yazar tarafından incelenen önemli bir konudur. Bazı köşe yazarları günümüzdeki sosyal medya regülasyonlarının bile gençler üzerindeki negatif etkilerinin sınırlandırılmamış olmasını önemli bir sorun olarak görmektedir [14]. Bireysel tehditlerin yanında bir grup araştırmacı da metaverse girişiminin toplumların politik ve ahlaki yapısına zarar verebileceği öngörüsünde bulunmaktadır [15].

Sonuç olarak günümüzde teknoloji ne kadar gelişmiş olursa olsun, metaverse gibi bir kavramı tam anlamıyla gerçeğe dönüştürecek kadar gelişmiş değildir. Köse [7], metaverse girişiminin vaat edilen noktaya gelmesi için teknolojinin en az 15-20 yıl gelişmesi gerektiğini belirtmiştir. Zamanımızın sanal gerçeklik endüstrisi hala gelişmekte olan bir endüstridir ve diğer birçok teknoloji endüstrisinin yanında henüz önemli bir pay sahibi değildir. Sanal gerçeklik gözlükleri ve türetilmiş ürünlerin Meta'nın hem teknik hem de işlevsel olarak gelecek tasviri yaptığı aşamaya gelmesi biraz zaman alacaktır [16].

Metaverse teknolojisinin başarısındaki etkili olması beklenen en önemli konular biri de metaverse araçlarının estetik açıdan şık ve hoş bir seviyeye ulaşmış olmasıdır. Bugün piyasada bulunan birçok sanal gerçeklik gözlüğü, kullanıcı deneyimi açısından kullanıcılar tarafından ergonomik bulunmamaktadır [17]. Kallman [18] da çalışmasında bu konuya değinerek metaverse kafa-takımını kullanan insanların fiziksel dünyada kendine zarar verebileceği tehlikesinden söz etmiştir.

3. MAKİNE ÖĞRENMESİ VE METİN MADENCİLİĞİ (MACHINE LEARNING AND TEXT MINING)

Makine öğrenmesi, öğrenen varlıklar olarak biz insanların tahmin ve sınıflandırma yeteneklerini taklit etmek için, bir araç olarak, veri analizi yapılmasını sağlayan bir metodlar sunar. Makine öğrenmesi günümüzde oldukça popüler bir analiz tekniği olsa da temellerinin 20. yüzyılın ikinci çeyreğine kadar uzandığı bilinen bir gerçektir [19]. Nitekim ünlü Türk matematikçi Cahit Arf bu konuyu bir çalışmasında ele almış ve şu soruyu sormuştur; "*Makineler düşünebilir mi?*" [20]. Makine öğrenmesi teknikleri günümüzde birçok problemin çözümünde kullanılmakta ve tatmin edici sonuçlar vermektedir.

3.1. Doğal Dil İşleme, Metin Madenciliği ve Duygu Analizi (Natural Language Processing, Text Mining and Sentiment Analysis)

Bu çalışmanın da konusu olan metin madenciliği kavramı makine öğrenmesinin çözüm sunduğu alanlardan biridir. En basit tanımıyla metin madenciliği, metin formatındaki sosyal medya verilerini, kurumsal belgeleri, tüketici yorumlarını ve internet sitelerinin yapısını oluşturan XML dosyalarını toplayıp, katma değer üretimine katkıda bulunacak kullanışlı ve anlamlı bilgileri keşfetme sürecidir [21].

Veri madenciliği bir üst kavram olarak metin madenciliğini kapsasa da metin madenciliği

kullandığı materyal ve uyguladığı metod olarak farklılıklar göstermektedir [22]. Geleneksel veri madenciliğinde veriler satır sütun ilişkisi bulunan tablo halinde ele alınırken metin madenciliğinde, yapısal olmayan veriler doğal dil metinlerinden alınmaktadır. Bu da veri ön işleme tarafında farklı teknikler kullanılmasını gerektirir. Metin madenciliğindeki veri ön işleme süreci için pratikte doğal dil işleme (*Natural Language Processing, NLP*) denmektedir [23]. Arık [24] çalışmasında "*Bilgisayarlar konuşulan dilde anlayabilir, yorumlayabilir, işleyebilir ve hatta cümleler üretebilir.*" yorumunda bulunarak doğal dil işleme (NLP) kavramının fonksiyonlarını özetlemiştir.

Duygu analizi kavramı da bu iki kavramın bir alt başlığı olarak literatürde yer almaktadır. Duygu analizi, çevrimiçi sosyal platformlardaki insanların veya arz talep dengesinin önemli bir bileşeni olan tüketicilerin duygularının ve düşüncelerinin (her türden metnin negatif, pozitif veya nötr kategorilere sınıflandırılması) belirlenmelerine yardımcı olmak için sosyal medya verilerine sıklıkla uygulanmaktadır [25]. Ayrıca duygu analizi birçok kaynakta fikir madenciliği (*Opinion Mining*) olarak da geçmektedir. Bu iki kavram arasında hiçbir fark yoktur.

Makine öğrenmesi algoritmalarının dışında duygu analizlerinin uygulanmasında başka yaklaşımlar da mevcuttur. Ancak makine öğrenmesi yaklaşımı duygu analizleri pratikte daha iyi sonuçlar vermektedir [26].

3.2. Twitter ve Youtube Verileri (Twitter and Youtube Data)

Twitter ve Youtube verilerinin özelliklerini incelemeye önce sosyal medya kavramından söz etmek faydalı olacaktır. Sosyal medya geçmiş yüzyılda internetin ortaya çıkıp genişlemesiyle birlikte, içinde bulunduğumuz yüzyılda insan hayatı için daha etkin bir rol oynamaya başlamıştır. Bu yüzyılın başlarında sosyal medya dediğimiz bir yapının oluşmasında elbetteki diğer teknolojik gelişmelerin de katkısı bulunmaktadır. Özellikle akıllı telefonların ortaya çıkıp, çok kısa bir zaman diliminde hayatlarımıza girmesiyle birlikte sosyal medya uygulamaları Web 2.0 çağındaki kullanıma göre önemli bir artış göstermiştir [27]. Statista [28]'ya göre 2023 yılından önce dünya üzerindeki sosyal medya aktif kullanıcı sayısı 3.43 milyara ulaşacaktır. Burada aktif ifadesi oldukça önemlidir. Diğer bir deyişle neredeyse dünya nüfusunun üçte biri düzenli bir şekilde sosyal medya kullanıyor olacak demektir. Tam da bu noktada sosyal medya verilerinin öneminden bahsedilmelidir. Zaten hacimsel olarak oldukça büyük boyutta olan sosyal medya verileri,

internetteki sosyal medya uygulaması olmayan uygulamalardan gelen verilerle birlikte „Big Data“ yani büyük veri kavramını oluşturmuştur [29]. Yukarıdaki bilgilerden de anlaşılacağı üzere bizi bekleyen yakın gelecekte de sosyal medya verileri üstel bir şekilde artmaya devam edecektir.

Çalışmanın ilk kısımlarında da değinilen sosyal medya verilerinin yüksek boyutlu ve artan hızda çoğaldığı gerçeği, özellikle son yıllarda Youtube kullanımındaki artış ile çok daha açıklanabilir bir hale gelmiştir [30].

Çalışmada kullanılan bu iki platformda kısaca bahsetmek gerekirse, Twitter kullanıcılarına 280 karakterlik bir metin iletisi (ek olarak video veya ses iletisi de dahil olmak üzere) izin veren bir sosyal medya uygulamasıdır. Dünyanın her yerinden insanlar bu platform üzerinde istedikleri bir konu, olay, fikir, veya düşünce hakkında kişisel duygu ve görüşlerini dile getirebilirler. Nitekim Dalkılıç & Çam [31] çalışmalarında son yıllarda Twitter verileriyle gerçekleştirilen çok sayıda çalışmanın öneminden bahsetmiştir. Youtube ise kurum/tüzel (profesyonel) kişilere kıyasla daha çok oranla tekil (amatör) şahısların video türünde içerik üretip paylaştığı bir video paylaşım platformudur. Her türden video içeriğinin (politik, eğitim, eğlence, müzik vb.) olduğu platform zengin içeriği ile bugün internet ansiklopedisi olarak adlandırılan Wikipedia'nın video türünden bir versiyonu olarak düşünülebilir.

4. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Günümüze kadar olan süreçte makine öğrenmesi tekniklerine dayanan birçok duygu analizi çalışması yapılmıştır. Bunların büyük çoğunluğunu denetimli makine öğrenmesi kategorisi altındaki sınıflandırma algoritmaları oluşturmaktadır. Aynı şekilde yine büyük bir çoğunluğunda kaynak olarak Twitter sosyal medya platformundan gelen verilerin kullanıldığı literatür taraması sonucunda gözlemlenmiştir. İnsanların fikir ve düşüncelerini analiz etmede önemli bir kaynak olan Twitter dışında spesifik alan bazlı birçok duygu analizi çalışması da mevcuttur. Örneğin film ve diziler hakkında yapılan duygu analizi için yapılan birçok çalışmada IMDB (Internet Movie Database) web sitesi veri kaynağı olarak alınmıştır.

Literatür taraması için ilk önce metin madenciliği ve duygu analizi alanında yapılan çalışmalar incelenmiş olup bu alanda ulaşılan teknolojik gelişmeler aktarılmaya çalışılmıştır. Bu kısım için geçtiğimiz 5 yıla (2017-2022) ait yapılan çalışmalar taranmıştır. Tarama sonucunda 7,255 adet çalışma listelenmiş ve en çok atıf alan 30 adet çalışma incelenmiştir. Bu kısımda kaynak taraması için SCOPUS veritabanı

kullanılmıştır. Bu veritabanındaki tarama için sorgu skripti aşağıdaki gibidir;

```
( TITLE ( "text mining" ) AND KEY ( "text mining" ) ) OR (
TITLE ( "sentiment analysis" ) AND KEY ( "sentiment analysis" )
) AND PUBYEAR > 2017
```

Son olarak da özel bir başlık olarak metaverse ile ilgili yapılan duygu analizleri incelenmiştir. Kriter bağımsız olarak gerçekleştirilen bu tarama işlemi sonucunda metaverse ile ilgili 8 adet duygu analizi çalışması ile karşılaşmıştır. Bu tarama SCOPUS, Google Scholar veri tabanları üzerinde gerçekleştirilmiştir. Taramada kullanılan anahtar kelimeler aşağıdaki gibidir: "metaverse", "sentiment analysis", "machine learning", "text classification", "opinion mining", "text mining". Bu anahtar kelimeler ikili ve üçlü kombinasyonlar halinde aratılmıştır.

4.1. Metin Madenciliği ve Duygu Analizi Alanında Yapılan Çalışmalar (Sentiment Analysis Studies Related to the Metaverse)

Zhang vd. [32], çalışmasında derin öğrenme tekniklerini ve bu tekniklerin duygu analizlerinde uygulanmasını incelemiştir. Ma vd. [33], hayat bilgisi (*commonsense knowledge*) çıkarımını bir yinelemeli sinir ağı olan uzun kısa süreli bellek (LSTM) algoritmasına uygulayarak hedef tabanlı duygu analizi çalışması yapmıştır ve metinler içinde geçen kelimelerin vektöre kodlanmasında hayat bilgisini daha etkin bir şekilde dahil edebilmesi için LSTM hücresinin bir uzantısını uygulamıştır. Yang vd. [34], güneş ışınımı ve fotovoltaik (PV) güç tahmini konusunda 2 farklı kaynaktan (Google Scholar, Science Direct) elde ettiği 1249 araştırma üzerinde bir metin madenciliği çalışması yapmıştır ve bu alanda kullanılan önemli terimleri ve güneş enerjisi tahminlerinin ortalamasını çıkartmıştır. Cambria vd. [35], bağlam yerleştirme araçlarını kullanarak duygu analizi için ilkel kavramları keşfetme çalışması gerçekleştirmiştir ve çalışmada sembolik ve alt sembolik yapay zeka metodlarını birleştirmiştir. Xue vd. [36], hedef tabanlı duygu analizi için bu analiz tipindeki diğer çalışmalara farklı olarak tekrarlı sinir ağları yerine evrimsel sinir ağlarını kullanmıştır ve bu sayede daha doğru ve verimli eğitim aşaması gerçekleştirilerek süreden tasarruf edilmesini sağlamıştır. Jianqiang vd. [37], duygu analizi çalışması için sözcüksel ve sözdizimsel özellikler analizini gerçekleştiren çoğu çalışmanın aksine denetimsiz öğrenme ile elde edilen bir kelime yerleştirme yöntemini derin öğrenme teknikleriyle uygulamıştır ve klasik yöntemlere kıyasla daha iyi bir performans elde etmiştir. Alaei vd. [38], turizm alanında büyük veri teknolojileri ile elde ettikleri

veriler üzerinde farklı duygu analizi yaklaşımlarını, temel değerlendirme ölçütleri üzerindeki performansları açısından incelemiş ve yorumlamıştır. Çalışmalarında uyguladıkları tüm duygu analizi metodlarının olumlu cümleleri sınıflandırmada olumsuz veya tarafsız cümlelerden daha iyi performans vermesini önemli bir bulgu olarak not düşmüşlerdir. Hussein [39], çalışmasında duygu analizi uygulamalarında karşılan zorlukları araştırmıştır. Do vd. [40], çalışmasında hedef tabanlı duygu analizi için farklı türdeki derin öğrenme algoritmalarının karşılaştırmasını yapmıştır. Basiri vd. [41], çalışmasında uzun kısa süreli bellek (LSTM) modeli ve kapılı tekrarlayan birim (GRU) tekniklerinin getirdiği yüksek boyutlu özellik uzayı ve eşit derecede görülen farklı özellikler gibi dezavantajları gidermek adına dikkate dayalı çift yönlü CNN-RNN derin öğrenme modeli geliştirmişlerdir ve duygu analizi için uygulamışlardır. Xu vd. [42], kullanıcıların ürünler veya hizmetler hakkındaki değerlendirmelerini ve sorularını anlayan ve yanıt veren makineleri inceledikten sonra bu makinelerin inşasında kullanılan teknikleri kendi hedef tabanlı duygu analizlerinde kullanmıştır. Ragini vd. [43], çalışmasında sosyal ağlardaki acil durum ve afet verilerini toplayarak bunlardan etkilenen insanların ihtiyaçlarına göre kategorilere ayırmıştır ve sonrasında kategoriler üzerinde duygu analizi çalışması yapmıştır. Sun vd. [44], çalışmasında yardımcı cümle oluşturma tekniğiyle BERT tekniğini birleştirerek belirli bir hedefe yönelik fikir kutupluluğunu belirlemeyi amaçlayan hedef tabanlı duygu analizi gerçekleştirmiştir. Al-Smadi vd. [45], destek vektör makineleri ve derin tekrarlayan sinir ağları kullanarak Arap otellerinin değerlendirmeleri için duygu sınıflandırması gerçekleştirmiştir. Xu vd. [46], çalışmasında geliştirmiş oldukları çift yönlü uzun kısa süreli bellek tekniği ile kelime ağırlık vektörlerini kullanarak bir duygu analizi gerçekleştirmiştir ve literatürde sıklıkla kullanılan LSTM, NB, RNN ve CNN algoritmalarıyla karşılaştırmıştır. Sonucunda ise geliştirmiş oldukları teknik çift yönlü uzun kısa süreli bellek mevcut algoritmalara göre daha başarılı bir sonuç vermiştir. Ma vd. [47], hedef tabanlı duygu analizi için LSTM mimarisine ek hücreler eklemiştir ve Sentic LSTM olarak adlandırdığı hibrit sinir ağı mimarisi önermiştir. Majumder vd. [48], çalışmasında duygu analizinin eğitimi sırasında hiyerarşik füzyon kullanarak çalışmada elde edilen özellikleri modlarına göre boyut indirgeme yapmıştır ve sonrasında bu modları da ikili ve üçlü olacak şekilde birleştirmiştir. Rezaenia vd. [49], önceden eğitilmiş kelime yerleştirmelerinin doğruluğunu artıran, geliştirilmiş kelime vektörleri (IWW) adlı yeni bir yöntem ile duygu analizi çalışması gerçekleştirmiştir ve bu teknik

ile Google'ın önceden eğitilmiş Word2Vec tekniğinin doğruluk başarı ölçütünü %2'nin üzerine çıkarmıştır. Sohangir vd. [50], çalışmasında evrişimli sinir ağları, uzun kısa süreli bellek (LSTM) ve doc2vec kullanarak finans alanındaki önemli sosyal ağlardan biri olan StockTwits'teki görüş ve değerlendirmeler üzerinde duygu analizi gerçekleştirmiştir. Duygu analizinde metin sınıflandırıcıları metni olumsuz ya da olumlu olarak sınıflamak için optimize edildiğinden ve bu nedenle, duygu içermeyen cümleleri bu iki kategoriden birine zorla sığdırdığı için bu metinlerin ayrıştırılması önemli bir problem olarak ortaya çıkmaktadır. Chaturvedi vd. [51] de çalışmasında, bu problem için öznellik tespiti için el yapımı ve otomatik modelleri incelemiştir ve el yapımı modellerin, güçlü duygular üzerinde iyi çalıştığını, ancak zayıf özne cümlelerde tatmin edici sonuçlar alınmadığı bulgusunu kaydetmiştir. Yadav vd. [52], çalışmasında duygu analizlerinde karşılaşılan sorunları çözmek için derin öğrenme mimarilerini kullanarak bu sistemlerin önemini vurgulamıştır. Amado vd. [53], pazarlama alanında büyük veriyle alakalı araştırmaları derleme ve özetleme yapabilmek adına bir metin madenciliği çalışması yapmıştır. Zhang vd. [54], çalışmasında genişletilmiş duygu sözlüğüne dayalı Çince dilindeki mikro blog metninin duygu analizini gerçekleştirmişlerdir. Öztürk vd. [55], 2011 yılında başlayan Suriye savaşının patlak vermesiyle başlayan mülteci krizi hakkında Türkçe ve İngilizce tweet verileriyle bir duygu analizi çalışması yapmıştır. Bu çalışma aynı zamanda mülteci krizine yönelik duyguları araştıran ilk çalışmadır ve bu çalışma göstermiştir ki mültecilere olan olumlu yaklaşım İngilizce tweet'lerden farklı olarak Türkçe tweet'lerde daha baskın taraftadır. Zhang vd. [56], doğal dil işleme ve metin madenciliği yöntemlerini kullanarak şantiyelerde meydana gelen kazaları incelemiştir ve optimize edilmiş topluluk modelleri, tekli modellerden F1-ölçütü açısından daha iyi performans gösterdiği görülmüştür. Wang vd. [57], çalışmasında duygu analizi için tekrarlayan sinir ağı'na (RNN) dayalı bir kapsül modeli olan RNN-Capsule geliştirmiştir ve özetle bu yöntem optimize edilmiş dinamik yönlendirme algoritmaları ve küresel parametre paylaşımı yoluyla kapsül ağ parametrelerini yinelemeli olarak güncellenmesine dayanmaktadır. Sailunaz vd. [58], Twitter platformu aracılığıyla elde ettiği verilerle duygu analizi gerçekleştirmiş ve benzer duyguları ifade eden kullanıcıların listesini içeren genel ve kişiselleştirilmiş bir liste oluşturmuştur. Jeong vd. [59], ürün planlaması için sosyal medya kullanıcılarından elde ettikleri veriler üzerinde duygu analizi gerçekleştirmiştir. Shen vd. [60], çalışmasında sosyal medyadaki verilerle iki dilli metin madenciliği aracılığıyla O2O (Online to Offline) ticaretinin trendini analiz etmiştir. Hemmatian vd. [61],

çalışmasında duygu analizi ve fikir madenciliğinin sınıflandırma teknikleri üzerinde sistematik, çok taraflı ve eksiksiz bir incelemesini yapmıştır.

4.2. Metaverse İle İlgili Duygu Analizi Çalışmaları (Sentiment Analysis Studies Related to the Metaverse).

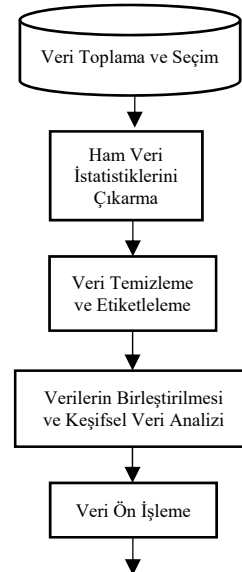
Literatürde metaverse ile ilgili yapılan duygu analizi çalışmaları incelendiğinde bu çalışmaların yaklaşım olarak ikiye ayrıldığı görülmüştür. Bazı çalışmalarda elde edilen veriler üzerinde keşifsel (*explorative*) veri analizi yapılırken bazı çalışmalarda ise bu çalışmada da olduğu gibi tahmine dayalı (*predictive*) veri analizi yapılmıştır. Bu çalışmalar uyguladığı yöntem ve sonuçlar ile birlikte aşağıda paylaşılmıştır.

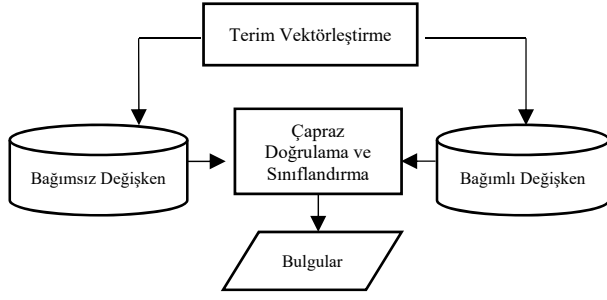
Ağralı & Aydın [62], çalışmasında metaverse ile ilişkili İngilizce tweet'leri analiz etmiştir ve tarih olarak Facebook şirketinin isim değişikliği işleminden bir hafta önceki ve bir hafta sonraki tweet'leri almıştır. Uygulama aşamasında Textblob, Vader ve Afinn duygu analizi araçlarını kullanmışlardır. Bu araçlardan Vader ve Textblob denetimsiz duygu analizi yöntemine girerken Afinn kural ve sözlük tabanlı yöntemine girmektedir. Bu üç araç da İngilizce metinler için kullanılan duygu analizi araçlarıdır. Facebook'un isim değişikliği açıklamasının 1 hafta öncesi ve sonrasında atılan tweet'ler için iki ayrı analiz çalışması yapmışlardır ve açıklama sonrası pozitif düşünce üç duygu analizi aracında da açıklama öncesine göre düşüş göstermiştir Lee vd. [63], metaverse uygulamalarının kullanıcı memnuniyeti tahmininin optimizasyonu için makine öğrenmesi algoritmalarını uygulamıştır. Bu algoritmalar K-En Yakın Komşu (KNN), Naive Bayes, Lojistik Regresyon, Hafif Gradyan Artırma Makinesi (LightGBM) ve Kategorik Artırma Algoritması (Catboost) olmak üzere 5 algoritmadan oluşmaktadır. Vektörleştirme için ise kullanılan araçlar Bag-of-Words (BoW), Terim Frekansı-Ters Belge Frekansı (TF-IDF) ve Word2Vec olarak sıralanabilir. Çalışmada %88,68 ile en yüksek doğruluğa sahip algoritma TF-IDF'li LightGBM olmuştur. Tunca vd. [64], ise çalışmasında metaverse ile ilgili Twitter platformundan elde ettiği 5447 İngilizce tweet'i metin madenciliği çalışmalarında sık kullanılan "socialbearing.com" üzerinde analiz etmiştir ve yüzde 58 ile pozitif tweet'ler birinci sınıf olmuştur. Hayawi vd. [65], sosyal medyadaki metaverse ile ilgili kamu görüşünü araştırmak için Twitter'dan elde ettiği 86.565 tweet ile bir duygu analizi çalışması gerçekleştirmiştir. Naive Bayes, K-En Yakın Komşu, Rassal Orman, Destek Vektör Makineleri, LSTM ve BERT algoritmalarıyla analizini gerçekleştirdiği çalışmada BERT %92,6 ile en başarılı doğruluk değerine sahip algoritma olmuştur. Kaya [66], çalışmasında metaverse kavramının çeşitli alternatif

kelimeleri hakkında Türkçe tweet'lerde duygu analizi ve insanların metaverse kavramı üzerinde ne kadar bilgisi olduğunu ve onu deneyimlemeye istekli olup olmadıklarını anlamayı amaçlayan bir anketin analizi gerçekleştirmiştir. Kullanılan tweet sayısının yer almadığı çalışmada araç olarak denetimsiz duygu analizi aracı olan Textblob kullanılmıştır ve yazar duygu analizinin sonucu olarak, toplanan tweet'lerin çoğunun olumlu olduğunu (Oran çalışmada yer almamaktadır.) ve insanların bu yeni kavramı deneyimlemeye istekli olduğunu belirtmiştir. Güler [67], Twitter'da metaverse hakkında atılan 1.572 İngilizce içerikli tweet üzerinde herhangi bir araç kullanmadan manuel bir duygu analizi çalışması gerçekleştirmiştir ve pozitif yorumların %33,1 çok iyi, %33,5 iyi ile baskın duygu olduğunu belirtmektedir. Kuş [9], çalışmasında Youtube'da "metaverse" kelimesi ile aratıldığında çok sayıda izlenme almış olan Türkçe bir videoyu veri kaynağı olarak seçmiş ve 3803 veri ile bir duygu analizi çalışması gerçekleştirmiştir. Rapidminer programıyla k-means kümeleme algoritmasını uyguladığı verileri 10 farklı kategoriye ayırabilmiştir. Beğeni sayılarının da hesaba katıldığı analiz sonucunda "Ekonomik Fırsatlara Yönelik Vurgu" ve "İçerik Üreticiye Geri Bildirim" en çok etkileşim alan temalardan iki tanesidir. Tunca vd. [68], 2021-2022 yılları arasında İngiliz The Guardian gazetesinde yazılmış metaverse konulu 201 makale üzerinde Leximancer yazılımını ve doğal dil işleme araçlarını kullanarak bir duygu analizi çalışması gerçekleştirmiştir ve %61 ile pozitif içerikli makaleler baskın tarafta yer almıştır.

5. MATERYAL VE METOT (MATERIAL AND METHOD)

Bu çalışmadaki tüm uygulamalar Google Colaboratory üzerinde Python programlama dili kullanılarak yapılmıştır. Uygulanan metot için adımlar Şekil 1'de verilmiştir.





Şekil 1. Metodoloji akış diyagramı
(Methodology flowchart)

5.1. Veri Seti (Dataset)

Bu çalışma için Twitter ve Youtube platformlarından veri çekilmiş ve veri seti oluşturulmuştur. Veriler iki ayrı sosyal medya platformundan web kazıma işlemi gerçekleştirilerek elde edilip, birleştirilmiştir. Verilerin elde edilmesi ve analizi Google Colab üzerinde Python dilinde yapılmıştır. Ön işleme aşamasında Microsoft Excel'den de yararlanılmıştır. Veri seti ve analizler ile ilgili yapılan işlemler aşağıda sırasıyla aktarılmıştır.

5.1.1. Veri Toplama ve Veri Seçim Kriterleri (Data Collection and Data Selection Criteria)

Twitter verilerinin elde edilmesinde Python programlama dilinin Selenium kütüphanesi kullanılmıştır. Python-Selenium, web tarayıcısı üzerinde robotik işlemler yapılmasına olanak sağlayan bir otomasyon aracı olarak tanımlanabilir [69]. Bu kütüphane sayesinde sıralı işlem adımlarının bilgisayara yaptırılması sağlanır. Bu işleme aynı zamanda literatürde „Web Kazıma“ (**Web Scraping**) da denilmektedir. Bir Python skripti oluşturulup, çalıştırıldıktan sonra sırasıyla aşağıdaki işlemleri gerçekleştirir;

1. Tarayıcıyı çalıştır. (Varsayılan tarayıcı) ve Twitter giriş sayfasına git.
2. Skriptte yazılan kullanıcı adı ve şifreyi gir.
3. Twitter arama kutusuna „#Metaverse lang:tr“ yaz ve arat.
4. Çıkan tweet'leri çek ve kaydırma çubuğunu aşağıya indir.
5. Bu işlemi aşağıya indirme (Eng: scroll down) denemesi 3 ü geçene kadar tekrarla. (Bu, sayfa daha fazla tweet yükleyemiyor demektir.)
6. Eğer deneme sayısı 3'ü geçmişse programı sonlandır.

2 gün 17 saat süren web kazıma işleminden sonra toplamda 8197 adet tweet verisi elde edilmiştir. Twitter verilerinin elde etme süreci 24.03.2022 tarihinde başlayıp 27.03.2022 tarihine kadar

sürmüştür. Twitter'dan elde edilen verilerden en eski tarihli atılan tweet 16.01.2022, en yeni atılan tweet ise 24.03.2022 tarihine aittir. Kazıma sürecinde internetin kesilip, süreci başa sardırma ihtimaline karşı kazıma işlemi gerçekleştirilen bilgisayarın interneti bir mobil hat üzerinden (akıllı telefon aracılığıyla) verilmiştir.

Youtube verilerinin elde edilmesinde Twitter kazıma işlemindeki gibi manuel olarak Selenium skripti oluşturulmamış, hali hazırda Youtube web kazıma işlemi yapılmasına olanak sağlayan Python'daki "youtube_comment_scraper_python" kütüphanesi kullanılmıştır. Video seçim kriteri için verilerin çekilmeye başlandığı tarih itibarıyla Youtube'da "metaverse" kelimesi ile yapılan arama sonucunda Türkçe içerikli en çok izlenen 3 video baz alınmıştır. 16 saat süren Youtube kazıma işleminin ardından toplamda 3043 (birinci video 1180, ikinci video 975, üçüncü video 888) adet Youtube video yorumu çekilmiştir. Youtube kazıma işlemi Twitter kazıma sürecine göre çok daha fazla tekrarlı veri oluşmuştur. Yani Youtube'dan çekilen bir yorum Excel dosyasına 3 veya 4 kez gelirken Twitter verilerinde bu oran yaklaşık 2'dir. Youtube verilerini elde etme süreci 19.03.2022 tarihinde gün ortası başlamış olup 20.03.2022 tarihine kadar sürmüştür. Kazıma işlemi sonrası tüm videolardan elde edilen yorumları tek bir veri setinde altında birleştirilmiştir. Burada en eski yorum 03.01.2022 tarihine, en yeni yorum ise 18.03.2022 tarihine aittir. Veri elde ediminde kullanılan videolar sırasıyla aşağıda verilmiştir.

1. "İnternette sonra bu gelecek! METAVERSE" – Youtube Kanalı: Barış Özcan
2. "Kimse bunları anlatmadı! METAVERSE hakkında her şey" – Youtube Kanalı: Mesut Çevik
3. "Metaverse: Evrensel Simülasyon'da Sizin Rolünüz Ne Olacak?" Youtube Kanalı: Evrim Ağacı

5.1.2. Veri Temizleme ve Etiketleme (Data Cleaning and Labeling)

Her iki kaynaktan da verileri elde ettikten sonra veri setleri temizlenmiştir. Twitter verilerinde; tekrarlayan, reklam içeren, bozuk, nötr ifade içeren örnekler çıkarılmıştır. Youtube yorumlarından ise bozuk ve nötr ifade içeren örnekler çıkarılmıştır. Burada bozuk veriyle kastedilen web kazıma işlemi sonucunda çekilen bazı tweet ve Youtube yorumlarının Html ayrıştırması (**Parsing**) sırasında meydana getirdiği hata olarak açıklanabilir. Bu aşamada elde edilen hatalı verilerin düzeltme işlemi uygulanabilir durumda olmasına rağmen tercih

edilmemiştir. Bunun sebebi şu şekilde açıklanabilir; uzun süren web kazıma işlemi sonrasında düzgün veya hatalı gelen yorum/tweet'ler karışık bir şekilde Excel dosyasına kaydedilmiştir ve bu da düzgün ve hatalı metinlerin elle ayıklanmasını gerektirmektedir. Sonuç olarak bu işlem efor ve zaman tasarrufu açısından göz ardı edilmiştir. Bozuk veriye örnek olarak Twitter ve Youtube verilerinden birer örnek verilebilir (Tablo 1).

Tablo 1. İki kaynaktan bozuk veri örneği
(Example of corrupted data from two sources)

Bozuk Veriler	
Twitter	„Bl ð ðmberg bÿle k ð nusuy ð r gelÿyor T¼arkleÿrn seş getÿrecek b ¼ay ¼k netaverse pr ð jesi Y¼lk ÿstelenmelerÿ 8. B ¼ay ¼k b ð rsÃ gÃte 🗨️🗨️🗨️🗨️🗨️“
Youtube	vr setler g ¼an ¼am ¼azde ç ð k etkÿsÿz. Bır Ã kın neta þ eysÿnÿ bÿyle þeylere ð d Ã kl Ã nn

İlk temizlik aşamasında Twitter'dan 5740, Youtube'dan 1623 veri kaydı atılmıştır. Geri kalan veri sayısı Twitter için 2457, Youtube için 1420'dir.

Temizlik aşamasından sonra her iki veri setindeki tweet ve yorumlar etiketlenmiştir. Verilerin neye göre ve nasıl etiketleneceği literatürdeki bazı çalışmalardaki adımlar baz alınmıştır. Meral & Diri [70] tarafından yapılan çalışmada veri etiketleme için birden fazla kişi etiketleme safhasında yer almıştır. Bu sayede duygu ve düşüncelerin insandan insana farklılık gösterebileceği ihtimali giderilmeye çalışılmıştır. Keza bu çalışmada da aynı yöntem uygulanmıştır. Veriler ana etiketleyici kişi tarafından kontrol etmesi için gönüllü ikinci kişiye verilmiştir. Veriler “Olumlu” ve “Olumsuz” olarak iki sınıf için etiketlenmiştir. Hedef değişken bu iki değerden oluşmuştur. “Nötr” tweet veya yorumlar temizlik aşamasında veri setinden çıkarıldığı için hesaba katılmamıştır.

Her iki kaynaktan elde edilen yorum ve tweet'lerin birleştirilmesi ve bu veri seti üzerinde analizlerin yapılması hedeflenmiştir. Ancak her iki kaynaktan elde edilen verilerin sayısı birbirinden farklıdır. Bu sebeple iki taraftan çekilen veriler arasından eşit sayıda veri alınması uygun görülmüştür. Çekilecek tweet ve yorum sayısı, en çok izlenen Youtube videolardan 1. videoya yapılan yorum sayısına göre belirlenmiştir. Bu videoya yapılan yorumlardan, veri temizleme sonrasında 459 tanesi etiketlenebilmiştir. Hesaplama kolaylığı olması açısından üç videodan da

450'şer adet ve eşitlik ilkesine bağlı olarak da Twitter'dan 1350 (450×3) adet tweet alınmıştır. Bu şekilde her iki kaynaktan gelen veri sayısının eşit olması sağlanmıştır. Görüldüğü üzere bunun sağlanabilmesi için her iki kaynaktan gelen bazı düzgün veriler elenmek zorunda kalmıştır. Bu eleme sürecinde rassallık ilkesi uygulanmış, her bir videoya yapılan yorum verileri ve tweet'ler kendi aralarında karıştırılarak ihtiyaç duyulan sayıda veri seçilmiştir. Rassal veri seçme işlemi Microsoft Excel üzerinde her bir metine 0 ile 1 aralığında sayılarla verilerle yapılmıştır. Youtube veri setinde her satır küçükten büyüğe doğru sıralanıp, en baştan 450 tanesi alınırken, Twitter veri setinde en baştan 1350 tanesi alınmıştır.

Veri birleştirme aşamasında Twitter verilerinden 1350 tweet, Youtube verilerinden de 1350 yorum olmak üzere toplamda 2700 veri kaydı etiketlenmiş, depolanmış ve birleştirilmiştir. Bu aynı zamanda sınıflandırma algoritmalarında kullanılacak verinin sayıdır.

Tablo 2. Veri Seti Detayları
(Dataset details)

Video	Çekildi	Etiketlendi	Kullanıldı
1. Video	1180	483	450
2. Video	975	478	450
3. Video	888	459	450
Platform			
Youtube	3043	1420	1350
Twitter	8197	2457	1350
Toplam	11240	3877	2700

5.1.3. Veri Ön İşleme (Data Preprocessing)

Bu aşamada duygu sınıflandırması için kullanılacak algoritmaların verileri anlayabilmesi için ön işleme çalışması yapılmıştır. İlk olarak bağımlı değişken olan “duygu” sütunundaki değerlere 0 ve 1 değerleri atanmıştır. Burada olumlu etikete sahip olanlar “0”, olumsuz etikete sahip olanlar da “1” olarak dönüştürülmüştür.

Metin madenciliği analizlerinin yapılabilmesi için elde edilen metinlerin (tweet ve yorum) ön işleminde literatürdeki diğer birçok çalışmada da sıkça kullanılan yöntemler kullanılmıştır. Bunlar Tablo 3'te belirtilmiştir.

Tablo 3. Veri ön işleme için kullanılan teknikler
(Techniques used for data preprocessing)

Tokenization:	Cümledeki kelimeleri ayırmak için kullanılır.
Stemming:	Kelimede var olan ekleri kaldırmak ve kökü bulmak için kullanılır.
Sayı ifadelerinin ve noktalama işaretlerinin kaldırılması:	Metinden noktalama işaretlerini ve sayıları çıkarmak için kullanılır.
Emoji, url, hashtag ve @ ile başlayan kelimelerin kaldırılması:	Youtube, Twitter gibi platformlarda yaygın olarak kullanılan ve metin sınıflandırmasında anlamsız olan veriler kaldırılır.
Stopword temizliği:	Metindeki anlamı değiştirmeyen ve hemen hemen tüm dillerde görünen kelimelerin çıkarılması (ve, veya, ancak, bazı, veya, gibi kelimeler)

Bu çalışmada “stopword” kelimeleri için hazır bir liste kullanılmamış, GitHub platformundan Selamettin Dirik adlı kullanıcının hazırlamış olduğu stopwords dokümanı alınarak terim eklemesi yapılmıştır [71]. Bunun sebebi bütün metinlerin aynı konu hakkında duygu ve düşünce belirtmesi sebebiyle eğitim sırasında oluşacak hatalardan kaçınmak için var olan stopwords listesine „Metaverse“ kelimesi gibi konuyla alakalı kelimeler de eklenmiştir.

5.2. Metin Vektörleştirme (Text Vectorization)

Metin vektörleştirme metin sınıflandırmada dolaylı olarak da duygu analizlerinde sıklıkla uygulanan bir işlemdir. Metin sınıflandırmada kullanılan algoritmaların başarısı çok büyük oranda terim vektörleştirme işlemine dayanmaktadır [72]. Bunun sebebi önceki bölümlerde de belirtildiği gibi makinelerin sadece sayısal değerlerden anlamlar çıkartabilecek yetenekleri olmasıdır.

Günümüzde kelime vektörleştirme için bir çok araç ve yöntem bulunmaktadır. Bu çalışmada TF-IDF Terim frekansı – Ters Metin Frekansı (**Term Frequency–Inverse Document Frequency**) istatistiksel metodu kullanılmıştır. TF-IDF metodunun pratikte tek başına Bag of Words aracından bir farkı bulunmamaktadır. Bu çalışmada da kullanılan 4 sınıflandırma algoritması için öncesinde bu işlem yapılmıştır.

5.3. Sınıflandırma Algoritmaları (Classification Algorithms)

Elde edilen veri setine sınıflandırma algoritmalarının uygulanabilmesi adına eğitim ve test veri seti olarak ayrılması noktasında çapraz geçirme (**Cross Validation**) yöntemi kullanılmıştır. 10 kat çapraz geçirme uygulanarak veri seti 10 parçaya ayrılmış, her seferinde bir parça test geriye kalan 9 parça ise eğitim veri olacak şekilde ayrıştırılmıştır. Aynı adımlar 10 kere tekrar edilmiştir. Sonrasında 4 farklı sınıflandırma algoritması modellenmiştir. Modellemede kullanılan sınıflandırma algoritmaları; naive Bayes (NB), lojistik regresyon (logistic regression-LR), destek vektör makineleri (support vector machine-SVM), rastgele orman (random forest-RF) şeklindedir. Algoritmaların seçiminde metin sınıflandırma çalışmalarında başarılı görülen ve incelemesi yapılmış algoritmalar önemli bir kriter olarak rol oynamıştır [73].

5.4. Model Performans Değerlendirme Ölçütleri (Model Performance Evaluation Metrics)

Model performanslarını değerlendirmesi için doğruluk ve F1-ölçütü metrikleri kullanılmıştır.

6. BULGULAR (FINDINGS)

İki ayrı veri kaynağından elde edilen verilere dair tanımlayıcı istatistik bulguları ve makine öğrenmesi modelleri aşağıda sunulmuştur:

6.1. Tanımlayıcı İstatistikler (Descriptive Statistics)

Veri seti için tanımlayıcı istatistikler Tablo 4’te verilmiştir. Twitter verilerinin %46’sı olumlu, %54’ü olumsuz duygu ifade ederken, Youtube yorumlarının %55’i olumlu, %45’i olumsuz duygu ifade etmektedir. Birleştirilmiş ve makine öğrenmesi analizlerinde kullanılmış veri setinde olumlu (%50,4) ve olumsuz (%49,6) yorumların birbirine yakın sayıda olduğu görülmektedir.

Tablo 4. Veri seti istatistikleri
(Twitter dataset statistics)

	Twitter	Youtube	Birleştirilmiş (Youtube+Twitter)
Toplam Veri	1350	1350	2700
Olumlu örnek	621 (%46)	740 (%55)	1361 (%50,4)

Olumsuz örnek	729 (%54)	610 (%45)	1339 (%49,6)
Ortalama kelime	13	14	14
Toplam kelime	18110	19316	37426

Twitter veri setinde geçen olumlu tweet'ler için kelime bulutu aşağıdaki (Şekil 2) gibidir:



Şekil 2. Twitter olumlu tweet'ler için kelime bulutu
(Word cloud for Twitter positive tweets)

Twitter veri setinde geçen olumsuz tweet'ler için kelime bulutu aşağıdaki (Şekil 3) gibidir:



Şekil 3. Twitter olumsuz tweet'ler için kelime bulutu
(Word cloud for Twitter negative tweets)

Youtube veri setinde geçen olumlu yorumlar için kelime bulutu aşağıdaki (Şekil 4) gibidir:



Şekil 4. Youtube olumlu yorumlar için kelime bulutu
(Word cloud for Youtube positive comments)

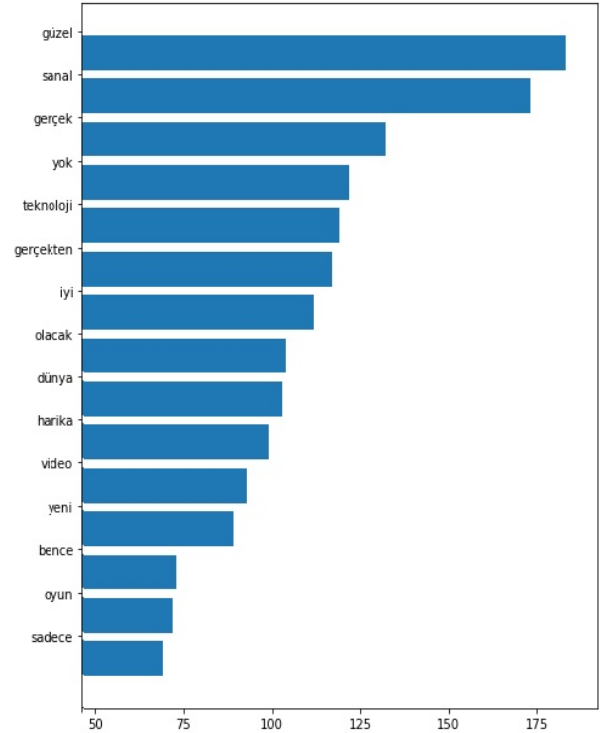
Birleştirilmiş Youtube veri setinde geçen olumsuz yorumlar için kelime bulutu aşağıdaki (Şekil 5) gibidir:



Şekil 5. Youtube olumsuz yorumlar için kelime bulutu

(Word cloud for Youtube negative comments)

Birleştirilmiş veri setinde ki popüler kelimelerin sıklık grafiği aşağıda (Şekil 6) verilmiştir. En sık geçen ilk 3 kelime görseldeki gibidir: güzel, sanat, gerçek.



Şekil 6. Birleştirilmiş veri setindeki popüler kelimeler

(Popular words in the combined data set)

6.2. Makine Öğrenmesi Bulguları (Analysis)

Sınıflandırmada kullanılan algoritmaların performans değerlendirme metrikleri Tablo 5'te verilmiştir. Kullanılan bütün algoritmalar %80 üzeri bir başarı değeri vermiştir ve en yüksek değeri Naive Bayes modeli vermiştir.

Tablo 5. Sınıflandırıcıların performansının karşılaştırılması

(Comparison of performance of classifiers)

Algoritma	Doğruluk	F1-ölçütü
NB	0,884	0,88
LR	0,866	0,87
SVM	0,841	0,84
RF	0,844	0,84

Tablo 6. Diğer performans metriklerinin karşılaştırılması

(Comparison of other performance metrics)

	Hedef Sınıf	Keskinlik	Duyarlılık	F1-Ölçütü
NB	Olumlu	0,87	0,9	0,89
	Olumsuz	0,9	0,87	0,88
LR	Olumlu	0,86	0,88	0,87
	Olumsuz	0,87	0,86	0,86
SVM	Olumlu	0,86	0,81	0,84
	Olumsuz	0,82	0,87	0,85
RF	Olumlu	0,86	0,82	0,84
	Olumsuz	0,83	0,87	0,85

Tablo 7. Çapraz geçiş sonrası standart sapmalar

(Standard deviations after cross validation)

Algoritma	Standart Sapma
NB	0,016
LR	0,017
SVM	0,022
RF	0,015

Literatür taramasında incelenen ve yöntem olarak bu çalışma ile benzerlik taşıyan çalışmaların doğruluk kriteri bakımından karşılaştırması Tablo 8'de sunulmuştur. Önemli bir detay olarak benzer çalışmalardaki algoritmaların en yüksek doğruluk oranları alınmıştır.

Tablo 8. Benzer çalışmalar ile kıyaslama

(Comparison with similar studies)

	Doğruluk			F1-ölçütü		
	Bu çalışma	[63]	[65]	Bu çalışma	[63]	[65]
NB	0,884	0,829	0,79	0,88	0,86	0,75
LR	0,866	0,883	-	0,87	0,90	-
SVM	0,841	-	0,909	0,84	-	0,909
RF	0,844	-	0,868	0,84	-	0,84
KNN	-	0,859	0,559	-	0,88	0,48
LIGHT-GBM	-	0,886	-	-	0,90	-
CATBOOST	-	0,875	-	-	0,89	-
LSTM	-	-	0,907	-	-	0,89
BERT	-	-	0,926	-	-	0,91

7. TARTIŞMA VE SONUÇ (DISCUSSION AND CONCLUSION)

Bu çalışmada sosyal medya platformlarında var olan Türkçe içerikli metinlerin makine öğrenmesi teknikleriyle duygu sınıflandırması gerçekleştirilmiştir. Çalışmada analiz safhasında 4 farklı sınıflandırma algoritması kullanılıp, başarı performanslarına göre karşılaştırılmıştır.

Platform bazında bakıldığında Youtube yorumlarında daha fazla sayıda olumlu görüş belirten içerik olduğu görülmüştür. Twitter'da ise olumsuz görüşleri içeren içerik sayısı daha fazladır. Buradan Youtube kullanıcılarının metaverse kavramına Twitter kullanıcılarından daha olumlu baktıkları yorumu yapılabilmektedir.

Kullanılan veri setindeki toplam kelime sayılarına bakıldığında Twitter'dan elde edilen veri setindeki kelime sayısının Youtube'dan elde edilen veri setindeki kelime sayısından daha az olduğu görülmüştür. Veri elde ediniminde iki platform için farklı metodların kullanılması, iki platformun içerik tipi bakımından farklılıklar içermesi (görsel ve yazılı) veya metaverse kavramına ilgi/yönelim bakımından iki platformun kullanıcı grupları arasında farklılık içerebilme ihtimali iki platform arasındaki kelime farkını açıklayabilecek faktörlerden bazılarıdır. Metin başına ortalama kelime sayıları da buna paralel olarak

Twitter’da daha az Youtube’da ise bir kelime daha fazladır. Burada çok büyük bir farklılık gözlemlenmemiştir.

Duygu sınıflandırması için bu çalışma bazında yapılan analizlerin sonucunda NB, %88,4 ile en yüksek doğruluk değeri veren algoritma olmuştur. Ancak bu değer literatürdeki benzer çalışmaların (Lee vd. [63], Hayawi vd. [65]) en yüksek değerleri ile karşılaştırıldığında en düşük değer olarak kendini göstermektedir. Benzer çalışmalarda analizin sayıca daha fazla veri üzerinde gerçekleştirilmiş olması ve derin öğrenme tekniğinin yüksek hesaplama gücünü kullanan algoritmalarından biri olan, aynı zamanda teknoloji devi Google tarafından da geliştirilmeye devam eden BERT’in analizde kullanılmış olması bu sonucun ortaya çıkmasında etkisi olabileceği düşünülen nedenler olarak sıralanabilir. Yine çalışma bazında diğer algoritmaların verdiği sınıflandırma sonuçları %80’nin üzerindedir.

NB birçok metin madenciliği çalışmasında kullanılan ve çoğunlukla yüksek başarı performansı veren algoritmalarından biri olarak, bu çalışmada en başarılı algoritma olmuştur. Diğer bir yandan en başarılı 3 sınıflandırıcı, olasılık yaklaşımlı makine öğrenmesi algoritması olurken diğerlerine kıyasla en düşük başarı performansı gösteren SVM, olasılık tabanlı bir algoritması olmamasıyla dikkat çekmektedir. Buradan hareketle de metaverse verileri üzerinde yapılan bu çalışmada olasılık tabanlı makine öğrenmesi algoritmalarının daha başarılı olduğu söylenebilir. 10 katlı çapraz geçirme işleminin ardından 4 algoritma için elde edilen değerlerin standart sapması alındığında en düşük değere (0,015) sahip RF algoritması en güvenilir çıktıyı veren analiz olmuştur.

Son olarak yeni bir kavram olarak hızlı bir popülerlik kazanan metaverse hakkında daha genellenebilir bir sonuç almak adına zamanla daha fazla Türkçe veriyle yeni duygu analizi çalışmalarının yapılabileceği düşünülmektedir. Literatürde metaverse ile ilgili çalışmaların artışıyla beraber bu çalışmadakine benzer araştırmaların yapılabileceği söylenebilir. Bu çalışma mevcut durumun tespitinin yapılabilmesi için verilerin yaygın olarak yöntemlerle analiz edilmesini sağlamıştır. Son olarak da metaverse teknolojisinin zamanla gelişerek bu girişimin öncüleri tarafından çok daha fazla desteklenmesiyle bu konuya olan duygu ve düşüncelerin sayısında artış gözlemlenebileceği düşünülmektedir.

KAYNAKLAR (REFERENCES)

[1] G. Resce ve D. Maynard, "What matters most to people around the world? Retrieving Better Life

Index“ *Technological Forecasting and Social Change*, cilt 37, pp. 61-75, 2018.

[2] SimilarWeb, “En Popüler Web Siteleri Sıralaması,” 09 2022. [Çevrimiçi]. Available: <https://www.similarweb.com/tr/top-websites/>. [Erişildi: 10 2022].

[3] C. D. Eyipinar, F. Büyükkalkan ve K. Semiz, “Sentiment Analysis Of Youtube Videos Comments On Sports Nutrition“ *Journal Of Physical Education Sport And Technologies*, cilt 2, no. 2, pp. 28-39, 2021.

[4] H. Cebeci, “Sosyal Medya Verileri ile Duygu Analizi“ *Mühendislikte Yapay Zeka Uygulamaları*, Sakarya, Sakarya Üniversitesi Yayınları, 2020, p. 192.

[5] İ. F. Pilavcılar, “Metin Madenciliği ve Metin Sınıflandırma,” Yüksek Lisans Tezi, 2007.

[6] N. Stephenson, *Snow Crash*. New York: A Novel; Random House Publishing Group, 2003.

[7] M. Köse, “Metaverse Nedir ve Neden Çok Önemlidir? Yaşamlarımızı Dijital Bir Evrene Taşıyabilir miyiz?,” 20 01 2022. [Çevrimiçi]. Available: <https://evrimagaci.org/meta-verse-nedir-veneden-cok-onemlidir-yasamlarimizi-dijital-bir-evrenetasiyabilir-miyiz-11135>.

[8] H. Duan, J. Li, S. Fan, Z. Lin, X. Wu und W. Cai, “Metaverse for social good: A university campus prototype,” *In Proceedings of the 29th ACM International Conference on Multimedia*, Chengdu, 2021.

[9] O. Kuş, “Metaverse: Perceptions Regarding Opportunities and Concerns in the 'Digital Big Bang’”, *Intermedia International e-Journal*, cilt 8, no. 15, pp. 245-266, 2021.

[10] K. B. Lee, “The Metaverse World and Our Future“, *Review of Korea Contents Association*, cilt 19, no. 1, pp. 13-17, 2021.

[11] M. Y. Kang, “Metaverse Framework and Building Block“, *Journal of the Korea Institute of Information*, cilt 25, no. 9, pp. 1263-1266, 2021.

[12] O. Güler ve S. Savaş, “Tüm Yönleriyle Metaverse Çalışmaları“, *Teknolojileri ve Geleceği*, cilt 8, no. 2, pp. 292-319, 2022.

[13] K. G. Nalbant ve Ş. Uyanık, “Computer Vision in the Metaverse“, *Journal of Metaverse*, cilt 1, no. 1, p. 9–12, 2021.

[14] Ü. Alan, “Metaverse Bir Kuş Değil, Biz De Havaya Bakmıyoruz“, 31 10 2021. [Çevrimiçi].

Available: <https://www.birgun.net/haber/metaverse-bir-kus-degil-biz-de-havaya-bakmiyoruz-363924>. [Erişildi: 10 2022].

[15] X. Shen, “Chinese state-owned think tank flags national security risks of metaverse, citing potential“, 01 11 2021. [Çevrimiçi]. Available: <https://www.scmp.com/tech/tech-trends/article/3154447/chinese-state-owned-think-tank-flags-national-security-risks..> [Erişildi: 10 2022].

[16] A. Aydın, “Sanal Gerçeklik Ve Artırılmış Gerçeklik içinde Eğitimde Dijitalleşme Ve Yeni Yaklaşımlar“, *İstanbul, Efe Akademi Yayınevi*, 2021, pp. 7-25.

[17] Y. Mengli, Z. Ronggang, W. Huiwen ve Z. Weihua, “An evaluation for VR glasses system user experience: The influence factors of interactive operation and motion sickness“, *Appl Ergon*, cilt 74, pp. 206-213, 2019.

[18] A. Kallman, “Cyber Security & The Metaverse“, *IBC*, 22 06 2018. [Çevrimiçi]. Available: <https://www.ibt.org/blog-cyber-security-and-themetaverse/2904.article>. [Erişildi: 10 2022].

[19] R. E. Neapolitan ve X. Jiang, “Artificial intelligence: With an introduction to machine learning“, Boca Raton: USA: Chapman & Hall/CRC, 2018.

[20] C. Arf, “Makine Düşünebilir Mi ve Nasıl Düşünebilir?“, 1958-1959 Öğretim Yılı Halk Konferansları, Erzurum, 1959.

[21] M. Hearst, “Untangling Text Data Mining“, *The 37th Annual Meeting of the Association for Computational Linguistics*, Maryland, 1999.

[22] D. Delen ve M. & Crossland, “Seeding the survey and analysis of research literature with text“, *Expert Systems with Applications*, cilt 34, no. 3, pp. 1707-1720, 2008.

[23] H. Yılmaz ve S. Yumuşak, “Open Source Natural Language Processing Libraries“, *İstanbul Sabahattin Zaim University Journal of the Institute of Science and Technology*, cilt 3, no. 1, pp. 81-85, 2021.

[24] K. Arık, “Social Media Content Review of MMORPG Games: Reddit Comment Scraping and Sentiment Analysis“, *Journal of Emerging Computer Technologies*, cilt 2, no. 1, pp. 13-21, 2022.

[25] M. K. Çelenli, “Sentiment Analizi (Duygu Analizi) Nedir?“, *TAVTechNews*, 2018.

[26] A. Bakirov, K. N. Çoğalmış ve A. Bulut, “Scalable sentiment analytics“, *Turkish Journal of Electrical Engineering and Computer Science*, cilt 24, no. 3, pp. 1560 - 1570, 2016.

[27] A. Yanık ve S. Özçiçek, “Akıllı Telefon Bağımlılığında Sosyal Medya ve Oyunların Etkilerini Anlamak“, *Uluslararası Halkbilimi Araştırmaları Dergisi*, cilt 4, no. 7, pp. 177-192, 2021.

[28] Statista, “Soziale Medien - Statistiken und Fakten“, *Statista Research Department*, 2022.

[29] A. Katal, M. Wazid ve R. H. Goudar, “Big data: Issues, challenges, tools and Good practices“, *2013 Sixth International Conference on Contemporary Computing (IC3)*, Noida, 2013.

[30] Statista, “Most popular social networks worldwide as of January 2022, ranked by number of monthly active users“, *Statista Research Dep.*, 2022.

[31] F. Dalkılıç ve A. Çam, “Automatic Movie Rating by Using Twitter Sentiment Analysis and Monitoring Tool“, *Journal of Emerging Computer Technologies*, cilt 1, no. 2, pp. 55-60, 2021.

[32] L. Zhang, S. Wang ve B. Liu, “Deep learning for sentiment analysis: A survey“, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, cilt 8, no. 8, 2018.

[33] Y. Ma, H. Peng ve E. & Cambria, “Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM“, *32nd AAAI Conference on Artificial Intelligence*, New Orleans, 2018.

[34] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro ve C. F. M. Coimbra, “History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining“, *Solar Energy*, 168 (68), pp. 60-101, 2018.

[35] E. Cambria, S. Poria, D. Hazarika ve K. & Kwok, “SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings“, *32nd AAAI Conference on Artificial Intelligence*, New Orleans, 2018.

[36] W. Xue ve T. Li, “Aspect based sentiment analysis with gated convolutional networks“, *56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, 2018.

[37] Z. Jianqiang, G. Xiaolin ve Z. Xuejun, “Deep Convolution Neural Networks for Twitter Sentiment Analysis“, *IEEE Access*, cilt 6, 2018.

- [38] A. R. Alaei, S. Becken ve B. Stantic, "Sentiment Analysis in Tourism: Capitalizing on Big Data", *Journal of Travel Research*, cilt 58, no. 2, 2019.
- [39] D. M. E. M. Hussein, "A survey on sentiment analysis challenges", *Journal of King Saud University - Engineering Sciences*, 30 (4), pp. 330-338, 2018.
- [40] H. H. Do, P. W. C. Prasad, A. Maag ve A. Alsadoon, "Deep learning for aspect-based sentiment analysis: A comparative review", *Expert Systems with Applications*, cilt 118, pp. 272-299, 2019.
- [41] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria ve U. R. Acharya, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis", *Future Generation Computer Systems*, cilt 115, pp. 279-294, 2021.
- [42] H. Xu, B. Liu, L. Shu ve P. S. Yu, "BERT post-training for review reading comprehension and aspect-based sentiment analysis", *NAACL HLT 2019*, Minneapolis, 2019.
- [43] J. R. Ragini, P. M. R. Anand ve V. Bhaskar, "Big data analytics for disaster response and recovery through sentiment analysis", *International Journal of Information Management*, cilt 32, pp. 13-24, 2018.
- [44] C. Sun, L. Huang ve X. & Qiu, "Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence", *NAACL HLT*, Minneapolis, 2019.
- [45] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh ve B. Gupta, "Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of arabic hotels' reviews", *Journal of Computational Science*, cilt 27, pp. 386-393, 2018.
- [46] G. Xu, Y. Meng, X. Qiu, Z. Yu ve X. Wu, "Sentiment analysis of comment texts based on BiLSTM", *IEEE Access*, 7, pp. 51522 - 51532, 2019.
- [47] Y. Ma, H. Peng, T. Khan, E. Cambria ve A. Hussain, "Sentic LSTM: a Hybrid Network for Targeted Aspect-Based Sentiment Analysis", *Cognitive Computation*, cilt 10, no. 4, pp. 1-12, 2018.
- [48] N. Majumder, D. Hazarika, A. Gelbukh, E. Cambria ve S. Poria, "Multimodal sentiment analysis using hierarchical fusion with context modeling", *Knowledge-Based Systems*, c.161, pp. 124-133, 2018.
- [49] S. M. Rezaeinia, R. Rahmani, A. Ghodsi ve H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings", *Expert Systems with Applications*, cilt 117, pp. 139-147, 2019.
- [50] S. Sohangir, D. Wang, A. Pomeranets ve T. M. Khoshgoftaar, "Big data: Deep learning for financial sentiment analysis", *Journal of Big Data*, 5(1), 2018.
- [51] I. Chaturvedi, E. Cambria, R. E. Welsch ve F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges", *Information Fusion*, cilt 44, pp. 65-77, 2018.
- [52] A. Yadav ve D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: A review", *Artificial Intelligence Review*, cilt 53, no. 6, pp. 4335-4385, 2020.
- [53] A. Amado, P. Cortez, P. Rita ve S. Moro, "Research trends on big data in marketing: A text mining and topic modeling based literature analysis", *European Research on Management and Business Economics*, cilt 24, no. 1, pp. 1-7, 2018.
- [54] S. Zhang, Z. Wei, Y. Wang ve T. Liao, "Sentiment analysis of chinese micro-blog text based on extended sentiment dictionary", *Future Generation Computer Systems*, cilt 81, pp. 395-403, 2018.
- [55] N. Öztürk ve S. Ayvaz, "Sentiment analysis on twitter: A text mining approach to the syrian refugee crisis", *Telematics and Informatics*, cilt 35, no. 1, pp. 136-147, 2018.
- [56] F. Zhang, H. Fleyeh, W. X. ve M. Lu, "Construction site accident analysis using text mining and natural language processing techniques", *Automation in Construction*, c.99, pp. 238-248, 2019.
- [57] Y. Wang, A. Sun, J. Han, Y. Liu ve X. Zhu, "Sentiment analysis by capsules", *WWW '18: Proceedings of the 2018 World Wide Web Conference*, Lyon, 2018.
- [58] K. Sailunaz ve R. Alhajj, "Emotion and sentiment analysis from Twitter text", *Journal of Computational Science*, cilt 36, no. 2, 2019.
- [59] B. Jeong, J. Yoon ve J. Lee, "Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis", *International Journal of Information Management*, Cilt 48, 280-290, 2019.
- [60] C. Shen, M. Chen ve C. Wang, "Analyzing the trend of O2O commerce by bilingual text mining on social media", *Computers in Human Behavior*, cilt 101, pp. 474-483, 2019.
- [61] F. Hemmatian ve M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis", *Artificial Intelligence Review*, cilt 52, no. 1, p. 1495-1545, 2019.

- [62] Ö. Ağralı ve Ö. Aydın, "Tweet Classification and Sentiment Analysis on Metaverse Related Messages", *Journal of Metaverse*, cilt 1, no. 1, pp. 25-30, 2021.
- [63] S. H. Lee, H. Lee ve J. H. Kim, "Enhancing the Prediction of User Satisfaction with Metaverse Service Through Machine Learning", *Computers, Materials and Continua*, 72(3), pp. 4983-4997, 2022.
- [64] S. Tunca, B. Sezen ve Y. S. Balcıoğlu, "Twitter Analysis For Metaverse Literacy", *International New York Academic Research Congress*, 2022.
- [65] K. Hayawi, S. Shahriar, M. A. Serhani ve E. Alothali, "Inevitable-Metaverse: A Novel Twitter Dataset for Public Sentiments on Metaverse", *ResearchSquare*, 2022.
- [66] G. A. Kaya, "Sentiment Analysis on the Metaverse: Twitter Data", *Sakarya University Journal Of Computer And Information Sciences*, cilt 5, no. 2, pp. 147-156, 2022.
- [67] H. N. Güler, "Metaverse (Sanal Evren) Kavramı İçin Twitter'da Yer Alan İçeriklerin Analizi", *İletişim Çalışmaları Dergisi*, cilt 3, no. 5, pp. 97-107, 2022.
- [68] S. Tunca, B. Sezen ve V. Wilk, "An Exploratory Content and Sentiment Analysis of The Guardian Metaverse Articles Using Leximancer and Natural Language Processing", *ResearchSquare*, 1-34, 2022.
- [69] E. Nalbant, "Yazılım Yaşam Döngüsünde Testin Önemi Ve Bir Test Otomasyonunun Gerçekleştirilmesi", Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2020.
- [70] M. Meral ve B. Diri, "Sentiment analysis on Twitter", Published in: 2014 22nd Signal Processing and Communications Applications Conference (SIU), Trabzon, 2014.
- [71] S. Dirik, "Turkish_Sentiment_Analysis_With_Multinomial_Naive_Bayes", 23 Ağustos 2019. [Çevrimiçi]. Available: https://github.com/slmtndrk/Turkish_Sentiment_Analysis_With_Multinomial_Naive_Bayes.
- [72] R. Patil ve S. Kolhe, "Supervised classifiers with TF-IDF features for sentiment analysis of Marathi tweets", *Social Network Analysis and Mining*, cilt 12, no. 1, 2022.
- [73] A. C. Tantuğ, "Metin Sınıflandırma", *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, cilt 5, no. 2, 2012.