



BÜYÜK VERİ'NİN V'leri ve VERİ ANALİTİĞİ

Ayşe YILDIZ*

Öz

Bilgisayar teknolojisindeki gelişmelerle birlikte büyük veri günümüzün en dikkat çekici konularından biridir. 2000'li yıllarda Hacim (Volume), Hız (Velocity) ve Çeşitlilik (Variety) özellikleri ile tanımlanan bu kavram 20 yıllık bir dönem içinde her geçen zamanda kendisine yeni bir V ekleyerek 50'ye yakın V'li özelliğe sahip olarak karşımıza çıkmış durumdadır. Belli ki bu V'lerde artış devam edecektir. Çalışmada literatürde en çok karşılaşılan 32 adet V'li kavram tanıtılmış ve diğer kavramlar için kaynak önerilerinde bulunulmuştur. Ancak büyük veriyi bu kadar önemli kılan hiç şüphesiz onun bu ham hali değil, veri analitiği denilen tekniklerle işlenerek bilgiye dönüştürülmesi ve karar verme sürecinde kullanılan en önemli girdilerden birini oluşturmasıdır. Bu çalışmada veri analitiği kavramına ilişkin bilgiler sunulmuş, veri kaynakları ile kullanılma amaçları bakımından veri analitiği sınıflandırmaya tabi tutulmuştur. Son olarak da işletmeler başta olmak üzere kuruluşların bu büyük veriden ve veri analitiğinden sağlayabilecekleri faydalar sıralanmıştır. Çalışma bu kadar büyük bir konunun belirli boyutlarını ele almıştır, ancak bu konu gelişime açıktır. İleriki çalışmalarda veri analitiğinde kullanılan araçlar ve tekniklere, veri analitiğinde karşılaşılan sorunlara ve ileride gözlemlenebilecek trendlere ilişkin bilgiler sunulularak çalışma geliştirilebilir.

Anahtar Kelimeler: *Büyük veri, Veri analitiği, Veri analitik türleri.*

V's of BIG DATA and DATA ANALYTICS

Abstract

With the developments in computer technology, big data is one of the most striking issues of today. This concept, which was defined with the features of Volume, Velocity and Variety in the 2000s, has emerged with nearly 50 V features by adding a new V to itself over a 20-year period. Obviously, these Vs will continue to increase. In the study, 32 concept with V, which are the most common in the literature, were introduced and resource suggestions were made for other concepts. However, what makes big data so important is undoubtedly not its raw form, but the fact that it is transformed into information by processing with techniques called data analytics and constitutes one of the most important inputs used in the decision-making process. In this study, information on the concept of data analytics has been presented, and data analytics has been classified in terms of data sources and purposes of use. Finally, the benefits that organizations, especially businesses, can derive from this big data and data analytics are listed. The study has addressed certain aspects of such a large topic, but it is open to improvement. The study can be improved by presenting information about the tools and techniques used in data analytics, the problems encountered in data analytics, and the trends that can be observed in the future in future studies.

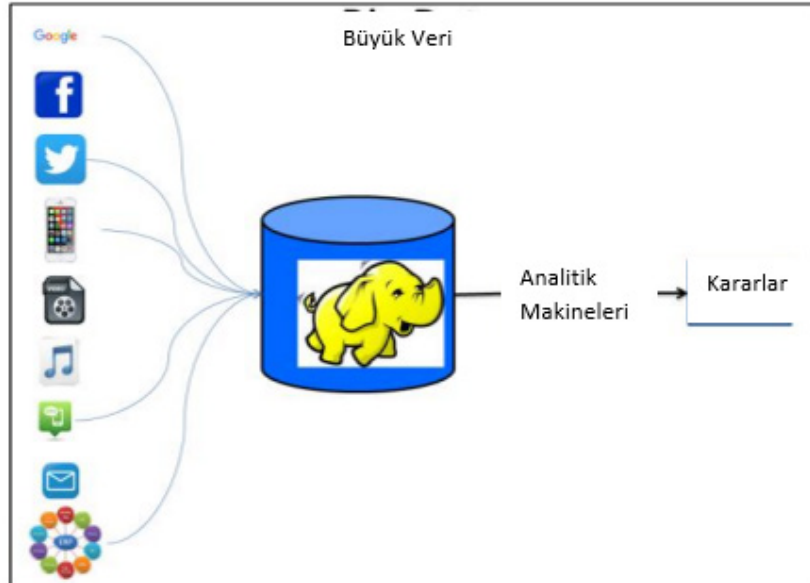
Key Words: *Big data, Data analytics, Types of data analytics.*

*Doç. Dr., Ankara Hacı Bayram Veli Üniversitesi, İİBF, İşletme Bölümü, Ankara
e-posta: ay.yildiz@hbv.edu.tr (<https://orcid.org/0000-0003-1165-3915>)

1. GİRİŞ

Bilgisayar teknolojisindeki gelişmeler birçok işletmenin büyük veri olarak adlandırılan bir veriye sahip olmasını sağlamıştır. Büyük veri, devrim niteliğindeki içgörülerini gerçek zamanlı olarak kullanıma sunan ve geleneksel veri işleme uygulamalarının yetersiz kaldığı çok büyük ve çok karmaşık olan veri kümelerini ifade eder. Bir iş için çözüm bulmak amacıyla yapılandırılmamış verilerden yapılandırılmış olanlara ilişkin kalıpları, trendleri ve ilişkilendirilmiş yapıları ortaya çıkarmak için analiz edilebilen büyük veri kümeleri günümüz iş dünyasının anahtar faktörlerinden biri olarak kabul edilmektedir. Artık akıllı saatler, gözlükler ve hatta akıllı kıyafetler ile dünya bir veri toplama mekanizması haline gelmiştir. Veriye bakıldığında aslında verinin her yerde olduğu görülmektedir. Google'da yapılan aramalar, Facebook'ta atılan mesajlar, insanların sosyal medyalarında yayınladıkları tüm durum güncellemeleri, resimler ve videolar, demografik özellikleri, beğenileri, beğenmedikleri, satın alma tercih bilgileri her yerdedir (Davenport, 2014: 45).

Son birkaç on yılda bu bilgiyi elde etmek için dünyanın dört bir yanındaki kuruluşlar tarafından toplanan ve depolanan veri miktarındaki büyük artış tartışılmazdır. Verilere erişim de üstel olarak artmıştır. Bu verilere erişme ve bunları analiz etme yeteneği de hızla giderek daha önemli hale gelmektedir. Basit istatistiksel ve matematiksel teknikler bu verilerin analizinde yetersiz kalmıştır. Büyük verilerin analizi için daha ileri düzey, kapsamlı ve tamamen bilgisayar teknolojisine dayalı tekniklerin kullanılması zorunlu hale gelmiştir. Bu teknik veri analitiği olarak bilinmektedir ve birçok kuruluşun büyük verilerden değerli bilgiler oluşturmak amacıyla benimsediği bu uygulama giderek hızla yükselen bir trend haline gelmiştir (Naganathan, 2018:1948). Büyük verinin farklı kaynaklardan nasıl oluşturulduğu ve karar verme için kullanılan veri analitiğine ilişkin süreç Şekil 1'de gösterilmiştir.



Şekil 1: Büyük Veri ve Analitiğe Genel Bakış

Kaynak: Jebble ve Patil, 2016: 4

Büyük veri ve veri analitiği konusu iş dünyasının gündemindeki ilk sıralarda yer almaya devam etmektedir. İş dünyası, hükümetler ve hatta Sivil Toplum Kuruluşları (STK), büyük veri analizinin sunduğu iç görülerden yararlanarak çok fazla sayıda fayda sağlayabilirler. Alışveriş merkezleri veya mağaza sahipleri, hangi ürünlerin en fazla talep edildiğini ve seyahat acenteleri de en çok hangi mekanların tercih edildiğini belirleyebilirler. Bu ilgi Google, Amazon ve benzeri şirketlerin verileri kullanarak geliştirdikleri yeni güçlü iş modelleri ve açık ara rakiplerinin önüne geçmelerini anlatan başarı hikayeleri ile daha da güçlenmektedir. Bunun yanında IBM ve Hewlett Packard gibi teknoloji devlerinin büyük veri konusundaki ciddi yatırımları da bu ilgiyi pekiştirmektedir (Altunışık, 2015: 66).

Bu çalışma, büyük veri kavramı ile birlikte büyük verinin birçok özelliğini ele almakta ve bu özellikler dikkate alınarak büyük verinin sunduğu olanakların neler olabileceğine dair bilgiler sunmaktadır. Bu kapsamda çalışma beş temel bölümden oluşmaktadır. Devam eden ikinci bölümde büyük veri kavramı tanımlamalarına ve belirleyici özelliklerine yer verilmiştir. Üçüncü bölümde veri analitiği ile türlerine ilişkin bilgiler sunulmuştur. Dördüncü bölümde veri analitiğinin özel bir durumunu ifade eden iş analitiği kısmı üzerinde yoğunlaşarak işletmelere ve kurumlara sundukları olanaklar belirtilmiştir. Beşinci bölüm sonuç kısmı olup bu kısımda genel değerlendirmeler yapılmış ve önerilerde bulunulmuştur.

2. BÜYÜK VERİ KAVRAMI ve ÖZELLİKLERİ

2.1. Büyük Veri Kavramı

Büyük veri tüm farklı veri yapılarını kullanan ve ilk kez Michael Cox ve David Ellsworth tarafından 1997 yılında düzenlenen 8. IEEE Görüntüleme Konferansı'nda (Proceedings of the 8th Conference on Visualization), "Application-Controlled Demand Paging for Out-of-core Visualization" adlı çalışmada kullanılan bir kavramdır. Büyük veri oldukça geniş, oldukça karmaşık ve bir o kadar da dinamik ve Microsoft gibi standart bir yazılım ile yönetilemeyecek kadar büyük olup terabaytlar ve petabaytlar gibi birimlerle ölçülmektedir (Vassakis ve diğerleri, 2018; Bilik ve Aydın, 2018).

Büyük veri kavramını daha iyi anlayabilmek için yapısal, yarı-yapısal ve yapısal olmayan veri türlerinden bahsetmek faydalı olacaktır (Jeble ve Patil, 2016: 4-5; Aktan, 2018: 3-4).

▪ *Yapısal (Yapılandırılmış) Veriler*

Yapısal verilerin kaynağını, satış noktası verileri, toplu işlemler, Kurumsal Kaynak Planlaması (ERP -Enterprise Resource Planning), Tedarik Zincir Yönetimi (SCM- Supply Chain Management) ve Müşteri İlişkileri Yönetimi (CRM-Customer Relationship Management) gibi genişletilmiş kurumsal bilgi sistemleri oluşturmaktadır. Bu veriler, ilişkisel bir veritabanında iyi tanımlanmış tablo yapılarında düzenlenir. Geleneksel ilişkisel veritabanı yönetim sistemleri (RDMS- Relational Data Management Systems) ile verileri çıkarmak (Extract), dönüştürmek (Transform) ve veri ambarına yükleme (Load) olarak belirtilen ve kısaca ETL olarak tanımlanan araçlar ile işlemler gerçekleştirilmektedir.

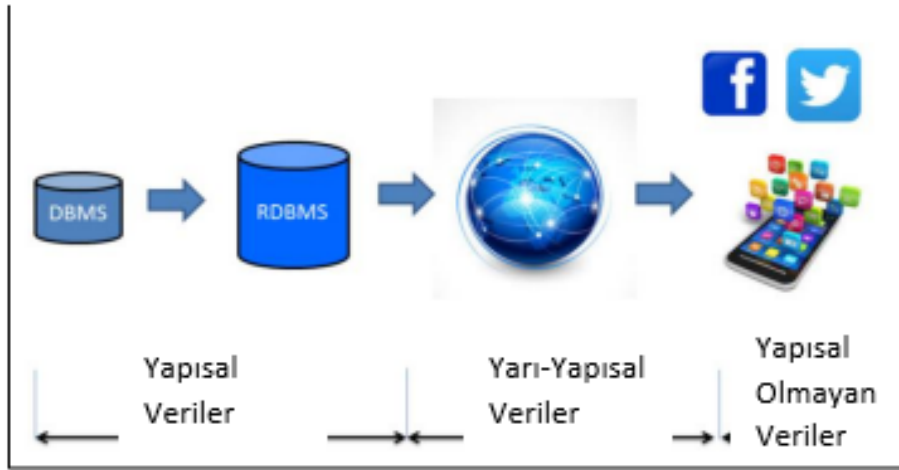
▪ *Yapısal Olmayan (Yapılandırılmamış) Veriler*

Yapısal olmayan veriler mesajlar, e-postalar, eklentiler gibi serbest formatlardaki metinler ile videolar, resimler, ses dosyaları gibi medya dosyalarından elde edilmektedir. Firmaların iletişim için kullandıkları e-postaları, lojistikte daha hızlı işleme için kullanılan Radyo Frekans Tanıma (Radio Frequency Identification – RFID) teknolojilerinin veya nesnelere takip etmek için geliştirilen sensör cihazlarının ürettiği veriler diğer yapısal olmayan veri kaynaklarıdır. Hacim ve hız açısından bunlar çok büyüktür ve büyük veri kümelerinin %80'inden fazlasını oluşturmaktadırlar. Bu verilerin analizi ve madenciliği yapılandırılmış verilerden çok daha zordur. NoSQL (Not only SQL) gibi yeni veritabanı yönetim sistemleri ile veri işleme süreçleri sağlanmaktadır. Bu verilerin analizinde ise Hadoop (Hadoop Dağıtılmış Dosya Sistemi) ve Eşle-İndirge (MapReduce) gibi ileri düzey bilişim teknolojileri kullanılmaktadır.

▪ *Yarı-Yapısal (Yarı-yapılandırılmış) Veriler*

Bu veriler yapısal ve yapısal olmayan kaynakların bileşiminden elde edilen verileri ifade etmektedir. Verideki farklı alanların hiyerarşisini göstermek amacıyla etiketler ve işaretler gibi çeşitli meta modellerini içermektedir. Genişletilmiş İşaretleme dili (Extendible Markup Language-XML) ve Javascript programlama dilleri bu veri yapısının en çok bilinen örnekleridir.

Şekil 2 bu veri türlerini göstermektedir.



Şekil 2: Büyük Verideki Veri Türleri

Kaynak: Jeble ve Patil, 2016: 4

Büyük veriyi karakterize eden yalnızca hacim değil, aynı zamanda elde edilebilecek bilgilerin heterojenliğinden kaynaklanan karmaşık özelliğidir. Ayrıca bu verilerin toplanması, işlenmesi ve kaydedilmesi hızlı ve değişken olarak ortaya çıkmaktadır. Büyük verinin %80-90'ı yapısal olmayan türdendir ve yıllık %40-60 oranında büyümekte olduğu öngörülmektedir (Altunışık, 2015: 49).

Günümüz dünyasında bu kavramın bu kadar çok öne çıkmasında birçok faktörün etkili olduğu görülmektedir: Bunlar şu şekilde belirtilebilir:

▪ *Artan Rekabet*

Artan pazar dinamizmi işletmeleri değişimlere karşı daha proaktif, daha esnek ve hızlı bir şekilde tepki vermeye ve çevresel değişimlere daha hızlı adapte olmaya zorlamaktadır. Bu durum işletmeleri tüketici davranışları, rakipler ve pazarla ilgili konularda daha fazla bilgi sahibi olmaya itmekte ve bu bilgiye sahip olmak için de daha fazla sayıda, türde ve sıklıkta veri toplamalarını gerektirmektedir (Altunışık, 2015: 56-57).

▪ *Multi-medya İçeriğindeki Artış*

Günümüzde multi-medya verisi İnternet yapısının %70 ten fazlasını teşkil etmekte olup, birçok sektörde de multi-medya içeriği giderek artmaya devam etmektedir.

▪ *Sosyal-Medya*

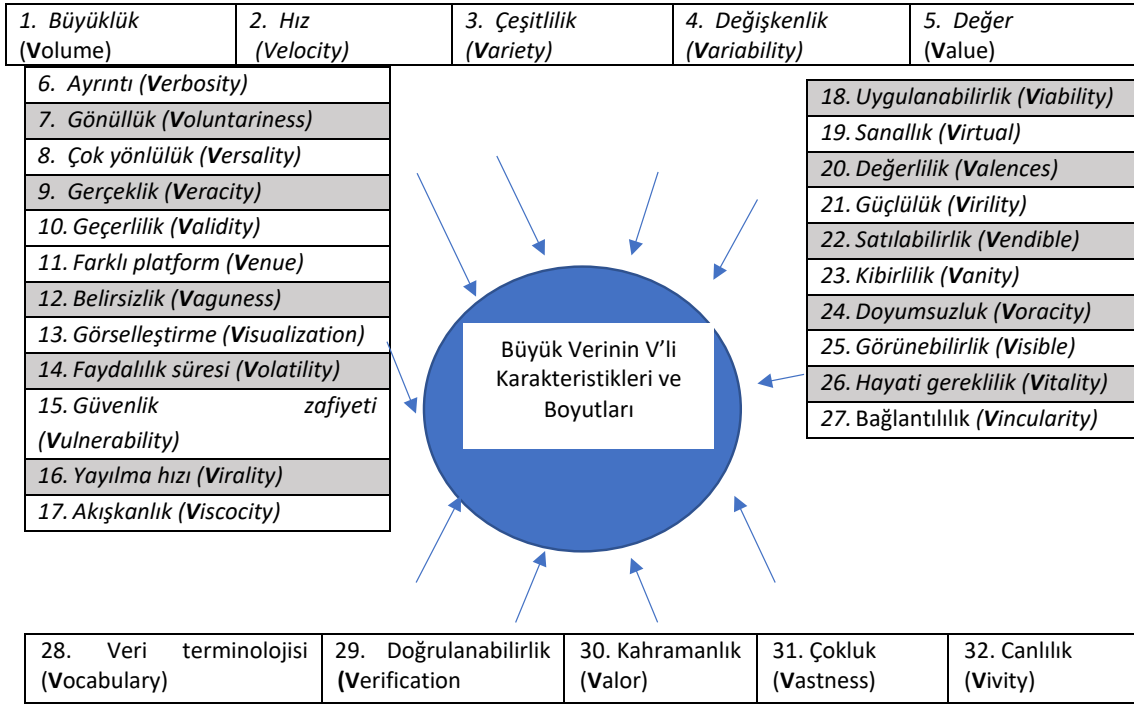
Sosyal medya, insanların iletişim kurma ve iş yapma şeklini değiştirmiştir. Sosyal medya dünyasında başarının şartı hızlı ve verimli olmaktır. Hedef kitle olan müşterilerle iletişim kurma şekli başarılı olmanın anahtarı haline gelmiştir. Büyük veri, birçok işletmenin hedef kitleleriyle iletişim kurmak için aradığı yeni yoldur. Bu yol şirketlerin hedef kitlelerine toplu e-postalar veya metin mesajları göndermelerine olanak tanımaktadır ve hedef kitlenin gerçek zamanlı olarak izlenmesini sağlamaktadır (Malik, 2022).

▪ *Nesnelerin İnterneti (IoT)*

IoT ve büyük veri, iyi bilinen teknolojik gelişmeleri sağlamak için birbirinden ayrılmaz iki bağımsız teknolojidir. IoT, farklı sensörler aracılığıyla fiziksel nesnelere büyük ölçüde veri toplarken, bu verilerin daha hızlı ve verimli bir şekilde depolanmasına ve işlenmesine de olanak sağlamaktadır (Altunışık, 2015:56-57).

2.2. Büyük Verinin Özellikleri

Büyük veri kavramı 2000'lerin başında endüstri analisti Doug Laney'in büyük verinin Hacim (Volume), Hız (Velocity) ve Çeşitlilik (Variety) olarak üç özelliğini ifade etmesiyle ortaya çıkmıştır. Her geçen zamanda bu özelliklere yenilerinin eklendiği görülmektedir. Bu özelliklerden öne çıkanlar Şekil 3'de gösterilmiştir.



Şekil 3: Büyük Verinin V'li Özellikleri

Kaynak: (Hussein, 2020)

Devam eden kısımda bu özelliklerden on tanesine ilişkin detaylı açıklamalar yapılmış, diğer yeni eklenen özelliklere ilişkin kısa tanımlamalara yer verilmiştir.

1) *Hacim (Volume)*

Hacim (büyükük) Facebook, Google, oyunlar ve Netflix gibi sınırsız kaynaklarla sosyal medya kullanıcılarının ürettiği çıktılardan elde edilen verileri ve bunlardaki sürekli büyümeyi ifade eden özelliktir. Veri hacmi ise tamamına sahip olunmasını gerektirmeyen ancak erişilebilir tüm mevcut olan veri miktarını ölçen bir değer olarak tanımlanmaktadır (Kaisler ve diğerleri, 2013: 996).

2) *Hız (Velocity)*

Büyük verinin hız özelliği verilerin daha önce görülmemiş bir hızda üretildiği, depolandığı veya yenilenmekte olduğu anlamına gelmektedir (Gandomi ve Haider, 2015: 138). Büyük verinin üretilme hızı çok yüksektir ve her geçen gün de hızla artmaktadır. Hız verilerin çevrimiçi işlem, sosyal ağ, video, ses veya harita görselleştirme tarafından oluşturulduğu gerçek zamanlı veri akışında çok önemli bir rol oynamaktadır. E-ticaret, farklı ticari işlemler (örneğin web sitesi tıklamaları) verilerin hızını ve zenginliğini hızla artırmıştır. Veri hızı yönetimi, bir bant genişliği sorunundan çok daha fazlasıdır; aynı zamanda bir işleme sorunudur (Kaisler ve diğerleri, 2013:996)

3) *Çeşitlilik (Variety)*

Çeşitlilik ise bir verideki yapısal heterojenliği belirtmektedir. Çeşitlilik, tipik olarak farklı bir kaynaktan, formattan ve türden toplanan disiplinler arası veri türü ile ilgilidir. Bu veriler reklamlardan, sensörlerden, mobil cihazlardan, kurumsal belgelerden, uydu görüntülerinden veya sosyal ağlardan gelebilir. Bu veriler genellikle yapısal, yarı-yapısal ve yapısal olmayan veriler şeklindedir (Amalina ve diğerleri, 2019: 3631). Analitik bir bakış açısından, büyük hacimli verilerin etkin bir şekilde kullanılmasının önündeki en büyük engeldir. Uyumsuz veri formatları ve veri yapıları ile tutarsız veri dili, analitik süreçte karşılaşılabilecek en büyük zorluklardır (Kaisler ve diğerleri, 2013:996)

Oguntimilehin ve Ademola (2014) tarafından ortaya konan özellikler ise diğer özelliklerinin yanısıra Değişkenlik ve Değer olarak ifade edilmiştir.

4) Değişkenlik (*Variability*)

Değişkenlik özelliği, aşırı belirsizlik ve öngörülemez koşullar olduğunda, verilerin yapısının düzenli ve güvenilir olmadığını, diğer bir ifadeyle verilerde ortaya çıkan tutarsızlığı belirtmektedir. Değişkenlik çeşitlilikten farklı olup, verideki heterojenliği değil anlamın değişmesini ifade etmektedir. Bu özellik verilerin sürekli olarak farklı kaynaklardan gelmesi ve gürültülü veriler ile önemli verileri ne kadar verimli bir şekilde ayırt ettiği ile ilgilidir (Panimalar ve diğerleri, 2017:331).

5) Değer (*Value*)

Çok büyük veriye sahip olmak, her zaman yüksek değerli verilere sahip olduğu anlamına gelmemektedir. Değer özelliği doğrudan nihai amaçla ilgilidir. Büyük veri değeri herhangi bir yeni veri kaynağının ve veri biçiminin işletmeye ne gibi ticari değer katabileceği ile ilgili olup karar vermede verilerin kullanılabilirliğini ölçer (Akay, 2018: 44). Büyük verilerin değerini tespit etmek zor olabilir. Ancak eğer bilginin eyleme dönüştürülmesi, bir yatırım getirisinin (ROI-Return on investment) hesaplanması, metin analizi ve duygu analizi ile müşteri ilişkilerinin düzeyi belirli ölçümlerle gerçekleştirilebiliyorsa verinin değeri belirlenebilir (Storey ve Song, 2017:52).

Kapil ve Khan (2016) tarafından ortaya konan Ayrıntı, Gönüllük ve Çok Yönlülük özelliklerinin kısa tanımlarına ise devam eden kısımda verilmiştir.

6) Ayrıntı (*Verbosity*)

Büyük verinin ayrıntılı doğasını göz önünde bulundurarak, “farklı kaynaklarda bulunan bilgilerin fazlalığı” olarak tanımlanabilir.

7) Gönüllük (*Voluntariness*)

Büyük veri, farklı kuruluşlar tarafından herhangi bir çıkar gözetilmeden gönüllü olarak kullanılan çok büyük miktarda veriyi ifade eder ve çok sayıda işletmeye yardımcı olmaktadır.

8) Çok Yönlülük (*Versatility*)

Büyük verinin becerikli/uyarlanabilir doğasını göz önünde bulundurarak, “büyük verinin farklı bağlamlar için farklı şekilde kullanılabilir kadar esnek olabilmesi” olarak tanımlanabilir.

Diğer ortaya konan özellikler de aşağıda sıralanmıştır (Shafer, 2017; Hussein, 2020; Wang, 2012; Dinh ve diğerleri, 2020):

9) Gerçeklik (*Veracity*)

Gerçeklik, toplanan büyük verilerin gerçeğe uygunluğunu ve tutarlığını ifade etmektedir. Diğer bir deyişle, gerçeklik bir veri setinin ne kadar güvenilir ve hatasız olabileceğini gösteren bir özelliktir.

10) Geçerlilik (*Validity*)

Doğruluk özelliğine benzeyen geçerlilik özelliği, verilerin amaçlanan kullanım için ne kadar doğru ve geçerli olduğunun ölçütüdür. Bir kurumun başarmak istediği hedefle daha fazla ilişkili olmasını ve amaçlanan kullanımı için daha yararlı bilgiler içermesini belirtmektedir.

11) Farklı Platform (*Venue*)

Farklı platform özelliği kişisel, özel ve genel bulut gibi farklı platformlar ile farklı geliştirilen sistemleri tanımlamaktadır

12) *Belirsizlik (Vaguness)*

Bir verideki varlığın net olmayışını belirtmektedir. Belirsizlik, her bir verinin ne ilettiği yani içeriği ile ilgili çok az bilgiye sahip olunmasını ifade etmektedir.

13) *Görselleştirme (Visualization)*

Görselleştirme birçok veri değişkenini içerebilen karmaşık grafikler anlamına gelmektedir. Büyük verinin çeşitliliği ve hızından kaynaklanan değişken sayısındaki muazzam artış ve bunlar arasındaki karmaşık ilişkilerin birleştirilerek görselleştirilmesi her geçen gün daha değerli bir özellik haline gelmektedir.

14) *Faydalılık Süresi (Volatility)*

Büyük verinin faydalılığı, saklanan verinin ne kadar süreyle kullanıcı için faydalı olduğu anlamına gelir. Bu durum, veri akışları değişen ve gürültü oranlarına sahip çeşitli kaynaklardan sağlandığında ortaya çıkar.

15) *Güvenlik Zaafiyeti (Vulnerability)*

Bu herhangi bir verinin ele geçirilebileceği veya manipüle edilebileceği anlamına gelmektedir. Büyük veri sistemleri, kişisel kullanıcılardan profesyonel kişilere kadar herkesin düşündüğünden daha savunmasızdır.

16) *Yayılma Hızı (Virality)*

Verinin bir kullanıcı tarafından yayınlanma/yayılma ve farklı kullanıcılar tarafından alınma hızı olarak tanımlanır.

17) *Akışkanlık (Viscosity)*

Bu veri hacmindeki akışa karşı direnci ölçer. Bu direnç, farklı veri kaynaklarından, entegrasyon akış hızlarından kaynaklanan gecikmeden ve verileri içgörüyü dönüştürmek için gereken işlem den kaynaklanabilir.

18) *Uygulanabilirlik (Viability)*

Uygulanabilirlik, kuruluşlar için tahmin etme olasılığı en yüksek olan verilerin dikkatli bir şekilde seçilmesi olarak görülebilir.

19) *Sanallık (Virtual)*

Büyük veri organizasyonları için hayali yapılar oluşturmaya dayanmaktadır. Kuruluşlar ve diğer gruplar, çeşitli amaç ve hedeflere ulaşmak için topladıkları tüm veri varlıklarını kullanma yetkisine sahip olduklarından büyük veri sanallaştırmasından yararlanabilirler.

20) *Değerlilik (Valences)*

Büyük veri çok büyük olabilirken, veri öğeleri arasındaki bağlantılar kurulmazsa, ilişkileri tam olarak anlayamayan veya kullanılmayan farklı veri alanlarına veya bölgelerine sahip olunabilir. Bir değerlik ölçüsü, gerçekte bağlı veri öğeleri ile veri toplama içinde kurulabilecek olası bağlantıların sayısı arasındaki oran olarak belirlenir.

21) *Güçlülük (Virility)*

Temel anlamıyla güçlü, çevik ve sağlıklı bir yetişkinin özelliklerine sahip olma özelliğini ifade eder. Büyük veri ile kendini yaratması anlamına gelir. Sahip olunan büyük veri ne kadar fazlaysa, verinin gücü de o kadar artmaktadır.

22) *Satılabilirlik (Vendible)*

Büyük verinin satılabilirlik varlığını kanıtlanmasını belirtir. Büyük veri için müşterinin varlığı, bunun kayda değer olduğunu gösterir. Örneğin abone verileriyle bilinen bazı ticaret araçlarının iletişiminden bu özellik açıkça görülebilir.

23) *Kibirlilik (Vanity)*

Boş veri, diğer bireyler üzerinde yarattığı etkiden memnun olduğu anlamına gelir

24) *Doyumsuzluk (Voracity)*

Doyumsuzluk özelliği yaygın ve ısrarlı bir şekilde kendini gerçekleştirme arzusuna sahip olmasını belirtir. Büyük veri, potansiyel olarak o kadar doyumsuzdur ki etkileme, yönetme ve kendini tüketme olanağını elde edebilir.

25) *Görünebilirlik (Visible)*

Görünebilirlik büyük veri depoları içinde saklanan verilerin sadece burada mevcut olmaması, aynı zamanda ihtiyaç duyulduğunda kullanıcılar tarafından erişilebilir olmasını belirtir. Veri analitiğinde karşılaşılan zorluklardan biri olarak görülür.

26) *Hayati Gereklilik (Vitality)*

Verilerin canlılığı, hayati önem taşıyan ve değer kavramı içinde yer alan önemli bir algıdır. Temel iş hedefi için daha anlamlı veya kritik olan bilgiler için önceliklerin belirlenmesi gerekir.

27) *Bağlantılılık (Vincularity)*

Bağlantı günümüzün internet üzerinden birbirine bağlı dünyasında çok uygun bir tanımlamadır. Çeşitli bilgi gruplarını birbirine bağlayarak önemli bir değer elde edilmesi mümkündür.

28) *Veri Terminolojisi (Vocabulary)*

Veri terminolojisi, veri modelini ve veri yapılarını ifade eder. Veri bilimi, çeşitli sorunları ele almak için bir kelime dağarcığı sağlar. Farklı modelleme yaklaşımları farklı problem alanlarını ele alır ve farklı doğrulama teknikleri bu yaklaşımları farklı uygulamalarda güçlendirir.

29) *Doğrulanabilirlik (Verification)*

Verilerin gerçekliğini, kesinliğini veya geçerliliğini ifade eder. Resmi belgelerin doğrulanması, böylece verilerin güvenilirliğini göstermek için birçok yönden standardizasyonu sağlar.

30) *Kahramanlık (Valor)*

Büyük veri karşısında, büyük sorunların kahramanca ele alınmasıdır. Değer üretme olasılığı olan ve bunun nasıl başarılacağına rehberlik eden spesifik verileri tanımlar.

31) *Çokluk (Vastness)*

Nesnelerin internetinin ortaya çıkmasıyla birlikte, büyük verilerin “büyüklüğü” hızlanmaktadır.

32) *Canlılık (Vivify)*

Veri bilimi, reklamcılıktan dolandırıcılık tespitine kadar her türlü karar verme ve iş sürecini canlandırma potansiyeline sahiptir.

2.3. Büyük Verinin İşletmelerde Kullanım Alanları

Veriler büyüdükçe kullanım alanlarının da genişlediği görülmektedir. Büyük veri teknolojilerine yapılan harcamalardaki en hızlı büyüme bankacılık, sağlık, sigorta, menkul kıymetler ve yatırım hizmetleri ile telekomünikasyon sektörlerinde yaşanmaktadır. İşletmeler artık birçok kararlarını sahip oldukları verilere dayanarak vermekte ve bu şekilde rekabet avantajı elde etmeye çalışmaktadırlar. Yılmaz (2021) tarafından gerçekleştirilen çalışmada büyük verinin işletmeler tarafından kullanım alanları sıralanmıştır. Bu alanlar Tablo 1’de gösterilmiştir. Tablodan görüldüğü gibi büyük veri işletmenin temel fonksiyonlarında ve kritik noktalarda yer almaktadır.

Tablo 1: Büyük Verinin İşletmede Kullanım Alanları

<i>Pazarlama alanında kullanımı</i>	<i>Üretim /hizmet alanında kullanımı</i>	<i>Finans/Bankacılık alanında kullanımı</i>
<ul style="list-style-type: none">- Kişiselleştirilmiş ürün ve hizmet sunma,- Müşteri memnuniyeti,- Müşteri kaybının önlenmesi,- Mobil pazarlama uygulamaları,- Pazarlama bilgi sistemleri,-Pazarlama karması ve pazarlama stratejileri geliştirme,- Pazar bölümlendirme,- Sosyal medya uygulamaları ve reklam,- Marka iş birlikleri.	<ul style="list-style-type: none">- E-ticaret uygulamaları,- Gerçek zamanlı teklifler,- Stok yönetimi,- Süreç optimizasyonu,- Ürün talep tahmini,- Ürün ve iş geliştirme,- Maliyetleri düşürme,- Verimliliği artırma.	<ul style="list-style-type: none">- Yatırım kararları alma,- Karmaşık borsa hareketlerinin analizi,- Küresel finansal kararlar alma,- Hisse senedi alış ve satışı,- Hile-suiistimali önleme,- Saldırı tespiti,- Risk analizleri,-Yeni kredi başvurularını değerlendirme,- Mevcut müşteriler için kredi puan analizi,- Yatırımcı analizi,- Hisse senedi tahmini.

3. BÜYÜK VERİ ANALİTİĞİ ve ANALİTİK TÜRLERİ

3.1. Büyük Veri Analitiği

Büyük veri analitiği, farklı türlerde içerik barındıran çok geniş ve farklı kayıtları işlemek adına geliştirilmiş analitik ve paralel tekniklerin kullanılmasıdır. Bu bağlamda büyük veri analitiği araçları, geleneksel veritabanları teknikleri kullanılarak işlenmesi zor olan, hızla değişen ve çok miktardaki yapısal, yarı yapısal ve yapısal olmayan verinin bir bütün olarak analizi ile veriden değerli bilgiler elde edilmesini amaçlayan teknikler olarak tanımlanabilir (Gahi ve diğerleri, 2016). Dysche (2014), büyük veri keşif çabasının, daha önce bilinmeyen bulguları ortaya çıkarabileceğini ve bunun da yönetsel karar vermede yardımcı olan içgörülerle sonuçlanabileceğini öne sürmüştür.

Mevcut rekabetçi iş ortamı, şirketleri yüksek hızlı verileri işlemeye ve değerli bilgileri üretim süreçlerine entegre etmeye zorlamaktadır. Arena ve Pau (2020) Forrester Research'ün alanında uzman kişilerle yapılan görüşme sonuçlarını gözönünde bulundurarak gerçekleştirdikleri çalışmalarında büyük veri analizinin kullanım durumlarını temel olarak üç gruba ayırmıştır:

- Potansiyel risklerin etkinliği ve yönetimi: Finansal süreçler, varlık yönetimi, insan kaynakları ve tedarik zincirinde risk minimizasyonu için uygulanan büyük veri çalışmaları,
- Tahmine dayalı güvenlik performans uygulamaları,
- Müşterilere sunulan bilgi ve hizmetler: Pazarlama ve satış projelerinde, ürün geliştirmede ve aynı zamanda dijital deneyimi optimize etmede kullanılan büyük veri analizine yönelik tüm çözümler ve uygulamaları.

Kullanılan istatistiksel, matematiksel ve optimizasyon teknikleri rekabet avantajı elde etmek ve sürdürmek için gerekli olan benzersiz müşteri gereksinimlerini ele almak için gittikçe daha önemli hale gelmektedir. Bu nedenle rekabetçi işletmeler için bu teknikleri kullanmak aşağıda örnekleri verilen amaçları gerçekleştirmek için kaçınılmazdır.

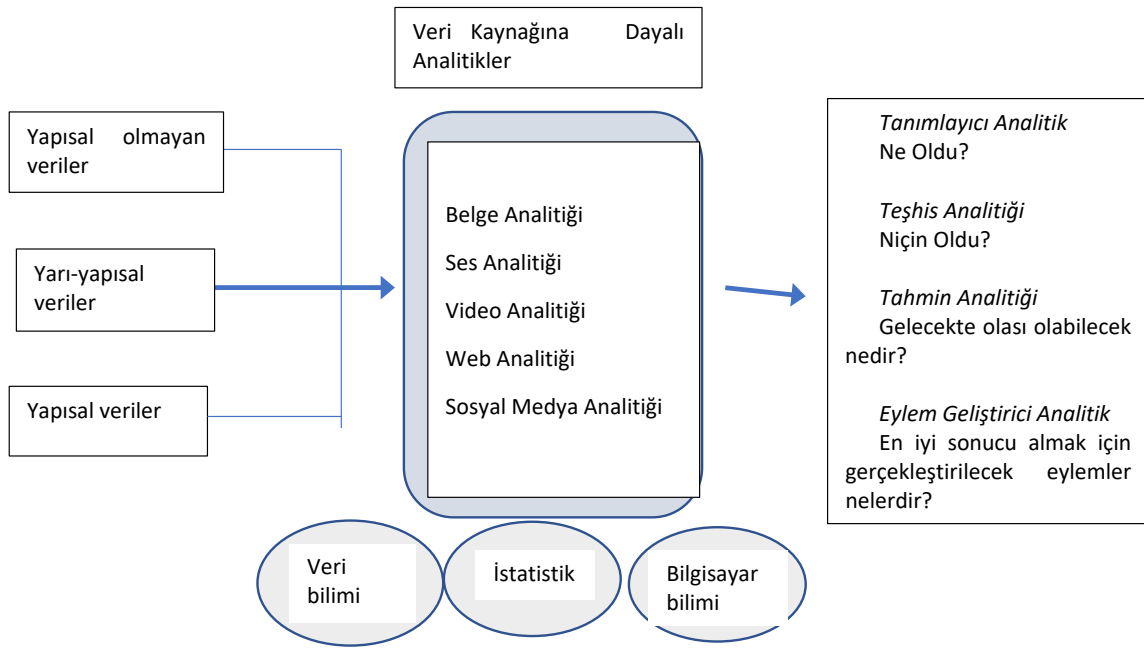
- Gelecekteki satın almalar için daha iyi ve daha fazla kişiselleştirilmiş önerilerde bulunmak veya indirimler sunmak,
- Arızaların, engellerin ve kusurların temel nedenlerini neredeyse gerçek zamanlı olarak belirlemek ve hatta olası arızaları gerçekleşmeden önce tahmin edip düzeltmek,
- Kalite iyileştirme ve ürün yeniliği için çevrimiçi tüketici incelemelerinin veya çağrı merkezi verilerinin analizi yoluyla tüketicilerin ürün veya hizmetlerle ilgili deneyimlerini anlamak,
- Kriz zamanında hızlı tepkiler geliştirmek ve anormallikleri tespit etmek,

- Dahili süreçlerde gerekli düzeltmeleri yapmak ve bir işletme içindeki operasyonel engelleri saptamak.

Yapısal ve yapısal olmayan veri akışlarının analitiğinden elde edilen içgörüler, işletmelerin daha önce düşünmediği soruları yanıtlayabilir. Muhtemelen, son on yılda hiçbir iş trendi, mevcut bilgi teknolojileri yatırımları üzerinde büyük veri analitiği kadar potansiyel bir etkiye sahip olmamıştır (Grover ve diğerleri, 2018: 390).

3.2. Büyük Veri Analitik Türleri

Büyük veri analitik türleri kullanılan veri kaynaklarına ve yapılan analitiğin amacına bağlı olarak farklı şekillerde ele alınabilir. Şekil 4 veri analitiğinin çerçevesini ortaya koymaktadır. Bunlara ilişkin detaylı açıklamalara devam eden kısımda yer verilmiştir.



Şekil 4: Veriden Analitiğe

Kaynak: Jebble ve Patil, 2016: 10

3.2.1. Veri Kaynağına Dayalı Analitikler

Veri kaynakları analiz için kullanılacağından elde edilecek çıktılar açısından önemli bir role sahiptir. Veri kaynağına dayalı olarak veriler Belge, Ses, Video, Web veya Sosyal medya analitiği olarak sınıflandırılabilir.

1) Belge Analitiği (Text Analytics)

Belge analitiğinde farklı analitik yaklaşımlar mevcuttur. Bunlardan öne çıkanlar aşağıda belirtilmiştir.

- Metin analizi: Metinsel verilerden bilgi çıkaran teknikleri ifade eder. Sosyal Ağ beslemeleri, e-postalar, bloglar, çevrim-içi forumlar, anket yanıtları, kurumsal belgeler, haberler ve çağrı merkezi günlükleri, kuruluşlar tarafından tutulan veriler metinsel verilerin örnekleridir. Metin analitiği, istatistiksel analiz, hesaplamalı dilbilim ve makine öğrenimini içerir. Bu analitik, işletmelerin, insan tarafından oluşturulan büyük hacimli metinleri, kanıta dayalı karar vermeyi destekleyen anlamlı özetlere dönüştürmelerini sağlar. Örneğin, metin analitiği, finans haberlerinden çıkarılan bilgiler üzerinde borsa tahmininde bulunmak için kullanılabilir.

- Metin madenciliği: Bu teknolojiyle daha önce fark edilmeyen içgörülerini ortaya çıkarmak için web'den, yorum alanlarından, kitaplardan ve diğer metin tabanlı kaynaklardan gelen metin verileri analiz edilebilir. Metin madenciliği, büyük miktarda bilginin analiz edilmesine ve yeni konular ve terim ilişkilerinin keşfedilmesine

yardımcı olmak için belgeleri (e-postalar, bloglar, Twitter beslemeleri, anketler, rekabet istihbaratı ve daha fazlası) taramak için makine öğrenimi veya doğal dil işleme teknolojilerini kullanmaktadır (SAS, 2022). Belge gösterimi, kurumsal arama sistemi, arama motorları, kullanıcı modelleri, geri bildirim alaka düzeyi, sorgu işleme, belirli bir ürün için Google'da milyarlarca müşteri araması, Amazon'un web sitesindeki aramalar ile müşterinin ürünü satın alma niyetinin göstergesini sağlamaktadır. Amazon, Jet Airways diğer birçok e-ticaret firması bu özelliği müşterilerin bir dahaki sefere web sitelerine göz atacağı zaman ürün veya uçuş önermek için kullanmakta ve böylece müşterinin satın alma kararının olasılığını artırmayı hedeflemektedir (Jebble ve diğerleri, 2018: 39).

- Metin özetleme: Metin özetleme teknikleri, tek veya birden fazla belgenin kısa fakat anlamlı bir özeti otomatik olarak üretilmesini sağlar. Ortaya çıkan özet, temel bilgileri orijinal metinden aktarmaktadır. Başvurular arasında bilimsel ve yeni makaleler, reklamlar, e-postalar ve bloglar yer almaktadır. Orijinal metni ayrıştırmak ve özeti oluşturmak için soyutlayıcı özetleme, gelişmiş Doğal Dil İşleme (Natural Language Processing-NLP) tekniklerini kullanmaktadır.

- Soru cevaplama: Bu teknikler doğal dilde sorulan sorulara cevaplar sağlamaktadır. Apple'ın Siri ve IBM'in Watson'ı ticari soru-cevaplama sistemlerine örnektir. Bu sistemler sağlık, finans ve pazarlama alanlarında kullanılmaktadır.

- Duygusal analiz: Bu teknikler insanların ürünler, organizasyonlar, bireyler ve olaylar gibi varlıklara yönelik görüşlerini içeren kanaate dayalı metinleri analiz eder. İşletmeler, müşterilerinin duyguları hakkında giderek daha fazla veri toplamaktadır ve bu da duygu analizinin aşırı yaygınlaşmasına yol açmaktadır (Gandomi ve Haider, 2015: 140; Chung, 2014; Liu, 2012).

2) Ses Analitiği (Audio Analytics)

Ses analitiği, yapılandırılmamış ses verilerini ortaya çıkarır ve çıkan verileri analiz eder. Ses analitiği, insan konuşma diline uygulandığında, konuşma analitiği olarak da anılır. Örneğin çağrı merkezleri, binlerce, hatta milyonlarca saatlik kayıtlı aramanın verimli analizi için ses analitiğini kullanmaktadırlar. Bu teknikler, müşteri deneyimini iyileştirmeye, temsilci performansını değerlendirmeye, satış devir oranlarını artırmaya yönelik olarak tasarlanabilir. Buna ek olarak, otomatik çağrı merkezleri şikayetleri belirlemek ve yönetmek için Etkileşimli Sesli Yanıt (IVR-Interactive Voice Response) platformlarını kullanabilmektedirler (Gandomi ve Haider;2015: 141; Chung, 2014).

3) Video Analitiği (Video Analytics)

Video içerik analizi (VCA) olarak da bilinen video analitiği, video akışlarından anlamlı bilgileri izlemek, analiz etmek ve çıkarmak için kullanılan çeşitli teknikleri içermektedir. Video analitiği diğer veri madenciliği türleriyle karşılaştırıldığında henüz emekleme aşamasında olsa da, gerçek zamanlı ve önceden kaydedilmiş videoları işlemek için çeşitli teknikler zaten geliştirilmiştir. Perakende satış noktalarındaki kameralar tarafından üretilen veriler iş zekası için kullanılabilir. Pazarlama ve operasyon yönetimi birincil uygulama alanlarıdır. Örneğin, akıllı algoritmalar müşteriler hakkında cinsiyet, yaş ve etnik köken gibi demografik bilgileri toplayabilmektedir. Benzer şekilde, perakendeciler müşteri sayısını sayabilir, mağazada kaldıkları süreyi ölçebilir, hareket modellerini tespit edebilir, farklı alanlarda kalma sürelerini ölçebilir ve kuyrukları gerçek zamanlı olarak izleyebilirler. Ürün yerleştirme, fiyat, ürün çeşitliliği optimizasyonu, promosyon tasarımı, çapraz satış, yerleşim düzeni optimizasyonu ve personel alma kararlarını yönlendirmek için bu bilgileri müşteri demografisi ile ilişkilendirerek değerli bilgiler elde edilebilir. Perakendede video analizinin bir başka potansiyel uygulaması, grupların satın alma isteği üzerine yapılan araştırmalarda yatmaktadır. Birlikte alışveriş yapan aile üyelerinden sadece biri kasadaki mağazayla etkileşime girerek geleneksel sistemlerin diğer üyelerin satın alma alışkanlıklarına ilişkin verileri kaçırmamasına neden oluyor. Video analizi, grubun demografisinin boyutu ve bireysel üyenin satın alma davranışı hakkında bilgi sağlayarak perakendecilerin bu kaçırılan fırsatı ele almasına yardımcı olabilir (Gandomo ve Haider, 2014: 141-142).

4) Web analitiği (Web Analytics)

Son on yılda, web analitiği, iş zekası (BI-Business Intelligence) ve yapay zeka (AI-Artificial Intelligence) içinde aktif bir araştırma alanı olarak ortaya çıkmıştır. Veri analitiğinin veri madenciliği ve istatistiksel analiz temelleri

ve metin analitiğindeki bilgi alma ve NLP modelleri üzerine inşa edilen web analitiği, benzersiz analitik zorluklarla birlikte inanılmaz fırsatlar sunmaktadır. http /html tabanlı hiper bağlantılı web siteleri ve web içeriğini bulmak için ilgili web arama motorları ve dizin sistemleri, web sitesi taraması/örümceği, web sayfası güncellemesi, web sitesi sıralaması ve arama günlüğü analizi için benzersiz İnternet tabanlı teknolojilerin geliştirilmesine yardımcı olmuştur (Chen ve diğerleri, 2012: 1176). Çevrimiçi perakendeci Amazon şirketi tıklama akışları, web aramaları, sipariş geçmişi vb. gibi büyük verileri çıkarmak için veri madenciliği tekniklerini kullanmaktadır. Bu tekniklerle geçmişteki benzer satın almalara dayalı olarak önceki satın alma geçmişi ile potansiyel yeni satın alma arasında bir ilişki elde edilebilmektedir. Bu ilişki, potansiyel müşterileri belirlemek ve bu müşterilere e-posta, Facebook gibi dijital medyayı kullanarak veya Amazon.com’da yanıp sönen mesajlarla farklı ürünleri tanıtmak için kullanılabilir (Jeble ve diğerleri, 2018: 39).

5) Sosyal Medya Analitiği (Social Media Analytics)

Sosyal medya, kullanıcıların içerik oluşturmaya ve değiş tokuş etmesine olanak tanıyan çeşitli çevrimiçi platformları kapsayan geniş bir terimdir ve aşağıdaki türlere ayrılabilir:

- Sosyal ağlar (ör. Facebook ve LinkedIn),
- Bloglar (ör. Blogger ve WordPress), mikrobloglar (ör. Twitter ve Tumblr),
- Medya paylaşımı (ör. Instagram ve youtube),
- Wikiler (ör. Wikipedia ve Wikihow),
- Soru-cevap siteleri (ör. Yahoo! Answers ve Ask.com),
- Yorum siteleri (ör. TripAdvisor).
- Ayrıca Find My Friend gibi birçok mobil uygulama, sosyal etkileşimler için bir platform sağlamakta ve dolayısıyla sosyal medya kanalları olarak hizmet etmektedir.

Sosyal medya analitiği de sosyal medyadan gelen bu verilerin analizine yönelik olarak geliştirilen teknikleri içermektedir. Bu analitik tekniklerini kullanan firmalar için elde edilen çıktılar değerli bilgiler içermektedir. Bu bilgiler tüketicilerin arama motorlarını ve sosyal medyayı kullanarak ürünleri nasıl araştırdığını göstermektedir.

3.2.2. Amaca Yönelik Analitikler

Verileri gerçek zamanlı olarak analiz etmek, kuruluşların geçmişi değerlendirmesine ve geleceği öngörmesine yardımcı olmaktadır. Bu durum akış analitiğinin özelliğidir. Bu analitikler sayesinde ne olduğunu bilmek (tanımlayıcı), neden olduğunu anlamak (teşhis), neler olabileceğine yönelik çıkarımlarda bulunmak (tahmin edici) ve nihayetinde gelecekteki eylemlerin nasıl oluşturulabileceğini belirlemek (eylem geliştirici) mümkün olabilmektedir (Ajah ve Nweke, 2019: 2). Chaudhari ve Patel (2017) bu durumu Şekil 5’de olduğu gibi göstermiştir. Bu analitiklere ilişkin açıklamalara devam eden kısımda yer verilmiştir.



Şekil 5: Veri Analitikleri

Kaynak: Chaudhari ve Patel, 2017: 596

1) Tanımlayıcı Analitik (Descriptive Analytics)

Tanımlayıcı analitik, ne olduğunu anlamaya yardımcı olan raporlar, gösterge tabloları aracılığıyla geçmiş verilerden bir olguyu açıklamada kullanılmaktadır (Jeble ve Patil, 2016:10). Bu analitik bir veri setinde veya veri tabanında nelerin bulunduğunu açıklayan merkezi eğilim ölçüleri (ortalama, medyan, mod), dağılım ölçüleri (standart sapma), çizelgeler, grafikler, sıralama yöntemleri, frekans dağılımları, olasılık dağılımları ve örnekleme yöntemleri dahil olmak üzere basit tanımlayıcı istatistiksel teknikleri içermektedir (Rajaman, 2016: 701). Bu analitik türünde temel amaç gerçek popülasyon değerinin bir tahmini yerine, belirli bir örnek için tanımlayıcı istatistiklerin bir özetini sağlamaya yöneliktir. Ortak tanımlayıcı sonuçlar genellikle grafikler, çizelgeler ve çizgiler aracılığıyla görüntülenmektedir. Örneğin, yapılandırılmış veritabanında tutulan şirketin ürünleri, müşterileri ve satış bilgileri periyodik olarak incelenerek verinin tüm özellikleri, veriler arası ilişkiler veya veride olası görülen eğilimler belirlenebilir. Böylece işletmeler popüler ürünleri cinsiyet ve yaşa göre ortaya çıkarabilirler. Bu tür veri analizi, rafların nasıl düzenleneceğine, ürünlerin nasıl tavsiye edileceğine ve indirimlerin nasıl teklif edileceğine karar vermeye yardımcı olmaktadır. Bu basit tanımlayıcı istatistik türleri, iş zekası ve OLAP ile ilgilidir (Duan ve Xiong, 2015:6).

Banerjee ve diğerleri (2013) tarafından işaret edilen başka bir tanımlayıcı analiz biçimi ise gösterge panosu türündeki uygulamalardır. Bu uygulamalarda işletmelerin bir süreci veya farklı zamanlardaki birden çok süreci izlemek için rutin olarak farklı metrikler oluşturulması söz konusudur. Oluşturulan bu metriklerle işletme belirli bir zaman noktasındaki mali gücünü anlama veya farklı zaman noktalarında başkaları veya kendileriyle karşılaştırma yapma olanağına sahip olabilecektir. Büyük veri analitiğinin birçoğu, doğası gereği genellikle tanımlayıcıdır (keşif amaçlıdır) ve tanımlayıcı istatistiksel yöntemlerin (veri madenciliği araçları) kullanımı, işletmelerin, ticari kararlar almak için kullanılacak faydalı kalıpları veya tanımlanamayan ilişkileri keşfetmesine olanak tanımaktadır (Naganathan, 2018: 1952).

2) Teşhis Analitiği (Diagnostic Analytics)

Teşhis analitiği, belirli olayların nedenini belirlemek için geçmiş verilerin analizi için kullanılmasıdır. Bu nedenle, teşhis analitiği, toplanan verilerdeki kalıpları kullanarak belirli olayların neden meydana geldiğini sorarak tanımlayıcı analitiği güçlendirir. Teşhis analizi, bir şeyin “neden” olduğunu değerlendirir. Bir sorunun temel nedenlerini keşfetmek için mevcut verilerin keşifsel veri analizine veya gerekirse görselleştirme teknikleri gibi araçlar kullanılarak ek verilerin toplanmasına ihtiyaç duyulabilir (Banerjee, 2013:2).

3) Tahmin Analitiği (Predictive Analytics)

Doğru ve etkin karar verme, gelecekte ne olacağına dair en iyi değerlendirmeyi sağlamakla ilgilidir. Tahmine dayalı analitik, neler olabileceğinin anlaşılmasına yardımcı olmaktadır. Bu analitik türünde amaç değişkenleri belirlemek ve tanımlayıcı bir analizde tahmine dayalı modeller oluşturmaktır. Bunun için geliştirilmiş istatistik,

yazılım, işlem ve yöneylem araştırması uygulamalarından faydalanılması gerekir. Buradan elde edilen sonuçlar, firmanın ürün ve hizmetlerini geliştirmek için yararlanabileceği fırsatları öngörmektedir (Rajaraman, 2016). Örneğin kullanım kolaylığı ile mobil para ödemesini kabul etme arasındaki ilişkinin ve etkinin belirlenmesi için regresyon analizi kullanılabilir. İlişkilerin ve olası etkilerin var olduğunu bilmek, bir dizi bağımsız değişkenin bağımlı değişkenleri nasıl etkilediğini açıklamaya yardımcı olmaktadır. Uygulamada, tahmine dayalı analitik, birkaç bin sensörden gelen veri akışına dayalı jet motorlarının arızasını tahmin etmekten, müşterilerin ne satın aldıklarına, ne zaman satın aldıklarına ve hatta sosyal medyada söylediklerine göre sonraki hareketlerini tahmin etmeye, dolandırıcılık tespitine, risk operasyonlarını belirlemeye kadar neredeyse tüm disiplinlere uygulanabilmektedir (SAS, 2022).

4) Eylem Geliştirici Analitik (Prescriptive Analytics)

Eylem geliştirici analitik, yönetici karar verme sürecini destekleyen başka bir güçlü araçtır. Farklı senaryolar altında farklı sonuçların anlaşılmasına yardımcı olmaktadır. Optimizasyon, simülasyonlar, eğer... ise analiz senaryosu gibi girdi parametrelerinin değiştirilmesi gibi çeşitli araçlardan oluşmaktadır. Yöneticiler, beklenen sonuçları doğru bir şekilde anlayarak karar verebilmekte ve beklenmedik durumları önceden planlayabilmektedirler. Bu analitik, tahsis edilen kaynakları en iyi şekilde kullanmak için karar bilimi, yönetim bilimi ve yöneylem araştırması metodolojilerinin (uygulamalı matematiksel teknikler) gücünü kullanmaktadır. Kaynaklar, öngörülen fırsatlardan yararlanmak için tahsis edilir. Örneğin, hedef müşterilere ulaşmayı hedefleyen ve sınırlı bir reklam bütçesine sahip olan bir mağaza, programlama modellerini ve karar teorisini kullanarak bütçeyi çeşitli reklam medyalarına en uygun şekilde tahsis etmek için sıralayabilir. Tedarik zincirlerinin tasarımında karı maksimize etmek için doğrusal programlama (kısıtlı bir optimizasyon metodolojisi) kullanılabilir. Genel olarak, ortaya konan çözümler eylemlerin belirlenmesi ve iş hedefleri, gereksinimler ve kısıtlamalarla ilgili etkilerin değerlendirilmesi ile karar vermede iş analistlerine yardımcı olmaktadır. Örneğin, simülatörler, bir işletmenin pazardaki mevcut konumunu korumak veya güçlendirmek için uygulanabilecek makul seçeneklere ilişkin içgörüler sağlanmasında kullanılabilirler (Naganathan, 2018: 1953-1954).

4. VERİ ANALİTİĞİNİN SUNDUKLARI

Veri analitiği gizli kalıpları, ilişkileri ve diğer içgörülerini ortaya çıkarmak için büyük miktarda veriyi incelemektedir. Günümüz teknolojisiyle, verileri analiz etmek ve ondan neredeyse anında yanıt almak mümkündür; oysa bu, daha geleneksel iş zekası çözümleriyle daha yavaş ve daha az verimli bir çabaydı. Veri analitiğinde veri miktarının fazla olmasından ziyade kuruluşların verilerle ne yaptığı daha önemlidir. (Jebble ve Patil, 2016). Büyük verilerin analitiği ile işletmelerin sağlayacağı avantajlar aşağıdaki şekilde sıralanabilir:

1) Maliyet ve Zaman Azaltma

Hadoop ve bulut tabanlı analitik gibi büyük veri teknolojileri, konu büyük miktarda veri depolamak olduğunda önemli maliyet avantajları, zaman tasarrufu ve ayrıca daha verimli iş yapma yöntemlerini belirleyebilirler (SAS, 2022). Hadoop kümeleri gibi büyük veri teknolojileri, geleneksel veritabanlarına kıyasla önemli ölçüde düşük maliyetli bir seçenek olarak ortaya çıkmış durumdadır. Bu teknolojiler müşterilere mevcut konumlarına göre teklif ve hizmetlerin tanıtımına ilişkin gerçek zamanlı kararların verilmesinde önemli rol oynayabilirler (Jebble ve diğerleri, 2018:41; Davenport ve Dyche, 2013).

2) Daha Hızlı, Daha İyi Karar Verme

Hadoop'un hızı ve bellek içi gücü ile yeni veri kaynaklarını analiz etme yeteneği bir araya geldiğinde, işletmeler bilgileri anında analiz edebilir ve öğrendiklerine göre kararlar alabilirler. Toplu olarak gerçekleştirilen veri süreçleri, yüksek performanslı analitiklerin ayrı ancak yüksek düzeyde entegre edilmiş işlevleridir. Büyük veri araçlarını ve yazılımını kullanmak, bir kuruluşun hangi verilerin alakalı olduğunu ve gelecekte daha iyi iş kararları almak için analiz edilebileceğini belirlemek için bir işletmenin topladığı son derece büyük hacimli verileri işlemesine olanak tanımaktadır (SAS, 2022; Nagenathan, 2018: 1949-195). Örneğin Netflix gibi firmalar, müşterilerin hangi ürünleri satın alabileceğini tahmin etmek için müşteri aramaları ve satın alma geçmişi arasındaki ilişkiyi bulmak için algoritmalar geliştirmişlerdir. Bu algoritmalar kullanılarak müşterilerin satın alma geçmişine göre

önerilen ürünler hakkında hatırlatma yapılmaktadır. Böylece müşterilerin önerilen ürünlerden bazılarını satın almasına ve böylece işletme için satışları artırma fırsatı yaratılmış olunur. Bu teknik, yeni satışların üçte biri kadarını sağlamaktadır (Jebble ve diğerleri, 2018: Artun ve Levin, 2015).

3) Yeni Ürünler ve Hizmetler

Müşteri ihtiyaçlarını ve memnuniyetini analitik yoluyla ölçme yeteneği, müşteriye istediklerini verme gücünü ortaya çıkarmaktadır. Daha fazla şirket müşteri istek ve ihtiyaçlarını karşılamak için yeni ürün yaratmada büyük veri analitiğinden yararlanmaktadır (SAS, 2022; Nagenathan, 2018: 1949-19).

4) Müşterileri Anlamak

Artık günümüzde işletmeler bireysel müşteri tercihlerini öğrenerek her müşteriye yönelik özel sunumlar ve teklifler yapabilmektedir. Hatta bazı işletmeler tüketici davranışlarına ilişkin verileri kullanıp tüketicilerin belirli davranışlarına ilişkin nedenlerini ortaya çıkarabilmektedirler. Bu şekilde ürün satışlarını ve tutundurma bağlamında yapılan her bir kuruluş pazarlama yatırımının geri dönüşümünü yükseltmenin yollarını aramaktadır (Altunışık, 2015: 67). Ayrıca çağrı merkezi günlüklerinden, web sitesi tıklamalarından, işlem kayıtlarından, ATM işlemlerinden, tıklama akışlarından vb. kaynaklardan elde edilen yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış verilerin karışımını kullanarak müşteri yolculuğunun eksiksiz profilini oluşturmaktadırlar. Bu profil, müşteri ayrılmalarının nedenlerini anlamalarına, yolculukları ilişkilendirmelerine yardımcı olmaktadır (Davenport ve Dyche, 2013). Örneğin, Wells Fargo, Bank of America ve Discover gibi önde gelen bankalar, aşağıda açıklandığı gibi çeşitli kaynaklardan toplanan büyük verileri kullanarak olası müşteri ilişkileri oluşturmaya çalışmaktadır (Jebble ve diğerleri, 2018: 41).

5) Rekabetçi İstihbarat

Sosyal ağ sitelerini kullanarak tüketici duygularını, tutumlarını ve görüşlerini anlamak için yürütülen çeşitli çalışmalar vardır. Ürünleriyle ilgili tüketici duyarlılığına ek olarak, şirket yöneticilerinin müşterilerin rakiplerinin ürünleri hakkında ne düşündüklerini de bilmeleri gerekir. Bu zeka, gelecekteki ürünlerdeki yenilikleri planlamaya veya ürünleri pazarlamak için bir strateji tasarlamaya yardımcı olacaktır. Bir sosyal medya verisinin madenciliği, bir işletmenin ve rakiplerinin tüketici görüşleri ve satış performansının karşılaştırmalı analizini gerçekleştirebilir. Benzer şekilde, Google tarafından sağlanan trend analizi, iki veya daha fazla rakip ürünün ürün aramalarını karşılaştırmak için iyi bir mekanizma sağlar. Bu analizler, farklı ürünlerin, hizmetlerin veya kişilerin web üzerinden farklı coğrafi konumlarda nasıl arandığına dair içgörüler sağlar. Bu, ürün farkındalığı ve gelecekteki pazarlama stratejilerinin veya yeni ürün lansmanlarının tasarlanmasıyla ilgili değerli bilgiler sağlayabilir (Jebble ve diğerleri, 2018:41).

6) Optimal Tedarik Zincir Yönetimi

Tedarik zincirleri, çok sayıda tedarikçi ve iş ortağı ile giderek daha karmaşık hale gelmektedir. Malların tedarik zincirleri arasında verimli hareketi için teknoloji önemli bir rol oynamaktadır. Tedarik zinciri analitiği, tedarik zinciri içindeki verilerin entegre bir görünümünü elde ederek karar vericilerin kapasitesini artırmaktadır. Bu sistem içindeki veri formundaki veri kaynakları ayıklanabilir, dönüştürebilir, analiz edebilir ve istihbarat elde etmek için analitik kullanılabilir. Tedarik zinciri analitiği, panolar, model ve eğilim analizi, ayrıntılı görünüm, tahminler, bilgi tabanı, senaryo ve durum analizi, simülasyonlar ve optimizasyon yetenekleri gibi çeşitli gelişmiş yetenekler sağlar. Bunlar, rekabetçi iş ortamlarında firmalar için çok önemli olan karar verme yeteneklerini ve durumların yorumlanmasını geliştirmektedir (Souza, 2014; Jebble ve diğerleri, 2018:41).

7) Gelecekteki Sonuçları Tahmin Etme

Bu kadar büyük miktarda veriyi analiz etmek için, büyük veri analitiği tipik olarak tahmine dayalı analitik, veri madenciliği, metin madenciliği, tahmin ve veri optimizasyonu için özel yazılım araçları ve uygulamaları kullanmaktadır. Gelecekteki sonuçları tahmin etmek için veri setlerini kullanmanın birkaç yaklaşımı vardır. Bu yaklaşımlar aşağıda belirlenen çerçevelerde geliştirilebilir;

a) Geçmiş işlem verilerine dayanarak, regresyon gibi tahmin modellerini kullanmak, bir firma için ürün veya hizmetler için gelecekteki satışları tahmin eder.

b) Geçmişteki satın almalarda bulunan ilişkilere dayanarak, müşteriler tarafından birlikte satın alınan ürünler belirlenebilir. Bu ilişkilere ve bir müşterinin satın alma geçmişine bakılarak, müşterinin hangi ürünleri satın alma olasılığının daha yüksek olduğu tahmin edilip önerilerde bulunabilir. Ayrıca geçmiş veriler analiz edilip, geçmiş verilerden ayrılan müşterilerin özellikleri ortaya çıkarılıp hangi değişkenlerin müşteri ayrılmasına neden olduğu belirlenebilir. Hangi müşterilerin ayrılma olasılığının daha yüksek olduğunu tahmin edildiğinde, onları elde tutmak için proaktif adımlar atılması mümkün olabilecektir (Jebe ve diğerleri, 2018:41; Artun ve Levin, 2015).

Büyük verilerden ve tahmine dayalı analitikten elde edilen iş değeri büyük, dönüştürücü ve oldukça görünürdür. Tahmine dayalı analitik teknikleri istatistiksel, veri madenciliği, makine öğrenimi ve derin öğrenme modellerini içerir. Veriler ayıklandıktan, organize edildikten ve hazırlandıktan sonra günümüz modelleri, milyonlarca veri kaydı üzerinden ölçülen binlerce tahmine dayalı değişkene dayalı olarak hızlı tahminler yapabilmektedir. Örneğin, makine öğrenimi teknikleri, örüntü tanıma konusunda uzmanlardır ve yapay zeka bilişsel modelleri (derin öğrenme veya sinir ağı tekniklerini kullanır) insanlardan daha hızlı izleyebilen, işleyebilen ve tepki verebilen karar verme araçları olarak kullanılabilirler (Naganathan, 2018: 1949).

8) Gerçek Zamanlı Karar Verme

Birkaç vizyon sahibi şirket, arz ve talep tarafı bilgilerini kullanarak geleneksel iş modellerinden çok daha etkin karar verme araçları geliştirmişlerdir. Bu araçları kullanan işletmeler arasında gerçek zamanlı karar verme yeteneklerine sahip olabilmektedirler. Örneğin, Uber gibi ulaşım hizmeti sağlayıcıları, araç alma sürelerini en aza indirmek ve yolculukla ilgili müşteri deneyimini optimize etmek için gerçek zamanlı araba rotası için büyük veri analitiğini kullanmaktadırlar (Jebe ve diğerleri, 2018:42).

5) SONUÇ

Bu çalışmada, büyük verinin yalnızca üç V (hacim, çeşitlilik ve hız) ile tanımlanan içeriği genişleterek büyük veri ile ilişkilendirilen yeni özellikleri tanımlanmıştır. Bu özelliklerin bilinmesi ve anlaşılması ile başta işletmeler olmak üzere büyük veri kullanıcıları büyük verinin sunduğu olanakları belirlemede daha başarılı olabilecek ve bu verileri kullanarak daha işe yarar veri analitiğini gerçekleştirebileceklerdir. Veri analitiği bugün birçok işletme için en güçlü rekabet silahı ve değerli varlık olarak görülmektedir. Özellikle değişime ve gelişime açık finans, bankacılık, pazarlama, sağlık alanında faaliyette bulunan birimler için müşteri ve hastalarına en iyi hizmeti sunmak için büyük veri analitiğinin potansiyel gücünden yararlanmak önemli bir rekabet unsuru haline gelmiş durumdadır. Bunun için işletmeler metin, ses, video, web aracılığıyla elde ettikleri verileri en gelişmiş teknolojilerle işleyerek bilgiye dönüştürmeyi ve bu bilgiyi karar verme sürecinde kullanmayı hedef edinmişlerdir. Hiç şüphesiz büyük veri insanların yaşamlarının her alanındaki verilere ulaşarak büyümeye devam edecek ve hergün kendine yeni özellikler ekleyecektir. Dolayısıyla bu dinamik ve büyük yapısıyla büyük veri ve büyük veri analitiğini her boyutuyla tek bir çalışmada ele almak mümkün olamamaktadır. İncelenen çalışmalar bu konunun bir veya birkaç boyutunu ele almakla sınırlı kalmışlardır. Nitekim bu çalışma da büyük verinin dinamik yapısını ortaya koymak amacıyla belirgin özelliklerin bir kısmını inceleyebilmiş; büyük veri analitiğinin sunduğu olanaklardan ve bu olanaklardan faydalanabilmek için gerçekleştirebilecek tanımlayıcı, tahmin edici, teşhis koyucu, eyleme geçirci analitik türlerinden kısmen bahsedebilmiştir.

Ancak birçok konuda olduğu gibi büyük veriyi elde etmede, işlemede, analiz edilerek bilgiye dönüştürülmesinde ve yönetsel bağlamda karar aracı olarak kullanılmasında birçok sorun mevcuttur. Bu sorunların güvenlik, hukuki, insan kaynakları ve yönetim gibi boyutlarda incelenmesi gerekmektedir. Ayrıca işin teknik boyutu ele alınarak kullanılacak matematiksel, istatistiksel tekniklere ve kullanılacak araçlara yönelik bilgiler verilebilir. Öte yandan bu alandaki trendlerin ve olası gelişmelere ilişkin öngörülerin ortaya konması da önemlidir. Gelecek çalışmalar konuyu bu boyutlarda inceleyerek, her geçen gün daha fazla dikkat çeken bu alana katkıda bulunabilir.

KAYNAKÇA

- Akay, E. Ç. (2018). "Ekonometride yeni bir ufuk: Büyük veri ve makine öğrenmesi". *Sosyal Bilimler Araştırma Dergisi*, 7(2), 41-53.
- Aktan, E. (2018). "Büyük veri: Uygulama alanları, analitiği ve güvenlik boyutu". *Bilgi Yönetimi*, 1(1), 1-22.
- Altunışık, R. (2015). "Büyük Veri: Fırsatlar Kaynağı mı Yoksa Yeni Sorunlar Yumağı mı?" *Yıldız Social Science Review*, 1(1), 45-76.
- Amalina, F., Hashem, I. A. T., Azizul, Z. H., Fong, A. T., Firdaus, A., Imran, M., ve Anuar, N. B. (2019). "Blending big data analytics: Review on challenges and a recent study". *IEEE Access*, 8, 3629-3645.
- Arena, F., ve Pau, G. (2020). "An overview of big data analysis". *Bulletin of Electrical Engineering and Informatics*, 9(4), 1646-1653.
- Artun, O., ve Levin, D. (2015). Predictive marketing: easy ways every marketer can use customer analytics and big data. John Wiley ve Sons.
- Banerjee, A., Bandyopadhyay, T., ve Acharya, P. (2013). "Data analytics: Hyped up aspirations or true potential?". *Vikalpa*, 38(4), 1-12.
- Big Data Analytics (Mayıs, 2018). "What it is and why it matters, SAS", (Erişim Tarihi: 11 May 2022, https://www.sas.com/en_us/insights/analytics/bigdata-analytics.html)
- Bilik, M., ve Aydın, Ü. (2018). Finansal Hizmetlerde Dijital Dönüşüm ve Etkileri. In Book of Proceedings 3rd International Congress on Economics, Finance and Energy. ISBN: 978-601-7805-32-6:22.
- Chaudhari, P., ve Patel, B. (2017). "Future of big data. International research journal of engineering and technology, 4(1), 595-597.
- Chen, H., Chiang, R. H., ve Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact". *MIS quarterly*, 1165-1188.
- Cox, M., ve Ellsworth, D. (1997). "Application-controlled demand paging for out-of-core visualization". In *Proceedings. Visualization'97, IEEE*, 235-244.
- Davenport, T. H. (2014). How strategists use "big data" to support internal business decisions, discovery and production. *Strategy ve Leadership*.
- Davenport, T. H., ve Dyché, J. (2013). "Big data in big companies". *International Institute for Analytics*, 3(1-31).
- Dinh, L. T. N., Karmakar, G., ve Kamruzzaman, J. (2020). "A survey on context awareness in big data analytics for business applications". *Knowledge and Information Systems*, 62(9), 3387-3415.
- Duan, L., ve Xiong, Y. (2015). "Big data analytics and business analytics". *Journal of Management Analytics*, 2(1), 1-21.
- Gandomi, A., ve Haider, M. (2015). "Beyond the hype: Big data concepts, methods, and analytics". *International journal of information management*, 35(2), 137-144.
- Grover, V., Chiang, R. H., Liang, T. P., ve Zhang, D. (2018). "Creating strategic business value from big data analytics: A research framework". *Journal of Management Information Systems*, 35(2), 388-423.
- Hussien, A. A. (2020). How many old and new big data v's characteristics, processing technology, and applications (bd1). *International Journal of Application or Innovation in Engineering ve Management*, 9(9), 15-27.
- Jebble, S., ve Patil, Y. (2016). "Role of big data and predictive analytics". *International Journal of Automation and Logistics*, 2(4), 307-331.
- Jebble, S., Kumari, S., ve Patil, Y. (2018). Role of big data in decision making. *Operations and Supply Chain Management: An International Journal*, 11(1), 36-44.
- K. Borne, "Top 10 Big Data Challenges – A Serious Look at 10 Big Data V's," 11 Mayıs 2022 tarihinde alındı. <https://www.mapr.com/blog/top-10-big-datachallenges-%E2%80%93-serious-look-10-big-data-v%E2%80%99s>.
- Kaisler, S., Armour, F., Espinosa, J. A., ve Money, W. (2013). "Big data: Issues and challenges moving forward. In 2013 46th Hawaii international conference on system sciences (pp. 995-1004). IEEE.
- Kapil, G., Agrawal, A., ve Khan, R. A. (2016). "A study of big data characteristics". In 2016 International Conference on Communication and Electronics Systems (ICCES) (pp. 1-4). IEEE.

- Laney, D. (2001). "3D Data management: Controlling data volume, velocity and variety". *META group research note*, 6(70), 1.
- Malik, R. S.2022, Big data in Social Media: How Big Data Can Be Used to Better Analyze Social Media Participation <https://medium.datadriveninvestor.com/big-data-in-social-media-how-big-data-can-be-used-to-better-analyze-social-media-participation-9ff2702de43d>
- Naganathan, V. (2018). "Comparative analysis of Big data, Big data analytics: Challenges and trends". *International Research Journal of Engineering and Technology (IRJET)*, 5(05), 1948-1964.
- Oguntimilehin A., ve Ademola E.O. (2014). "A Review of Big Data Management, Benefits and Challenges," *Journal of Emerging Trends in Computing and Information Sciences*, Vol-5, 33-437.
- Panimalar, A., Shree, V., ve Kathrine, V. (2017). "The 17 V's of big data". *International Research Journal of Engineering and Technology (IRJET)*, 4(9), 3-6.
- Rajaraman, V. (2016). "Big data analytic"s. *Resonance*, 21(8), 695-716.
- Shafer, T. (April, 2017). The 42 V's of Big Data and Data Science. 9 Mayıs 2020 tarihinde alındı. <https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html>
- Souza, G. C. (2014). "Supply chain analytics". *Business Horizons*, 57(5), 595-605.
- Storey, V. C., ve Song, I. Y. (2017). "Big data technologies and management: What conceptual modeling can do". *Data ve Knowledge Engineering*, 108, 50-67.
- Sun, Z., Strang, K., ve Firmin, S. (2017). "Business analytics-based enterprise information systems". *Journal of Computer Information Systems*, 57(2), 169-178.
- Vassakis, K., Petrakis, E., ve Kopanakis, I. (2018). "Big data analytics: applications, prospects and challenges". In *Mobile big data* (pp. 3-20). Springer, Cham.
- Waller, M. A., ve Fawcett, S. E. (2013). "Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management". *Journal of Business Logistics*, 34(2), 77-84.
- Wang, R. (2012). 10 Mayıs 2022 tarihinde alındı. <https://enterpriseirregulars.com/46120/beyond-the-three-vs-of-big-data-viscosity-and-virality/>
- Yılmaz, Ş. K. (2021). Türkiye'de Büyük Veri İle Değer Yaratan İşletmeler: Özel Sektör Uygulamaları Üzerine Bir İnceleme. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 5(1), 44-62.

Beyan ve Açıklamalar (Disclosure Statements)

1. Bu çalışmanın yazarları, araştırma ve yayın etiği ilkelerine uyduklarını kabul etmektedirler (The authors of this article confirm that their work complies with the principles of research and publication ethics).
2. Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir (No potential conflict of interest was reported by the authors).
3. Bu çalışma, intihal tarama programı kullanılarak intihal taramasından geçirilmiştir (This article was screened for potential plagiarism using a plagiarism screening program).