



Türk Turizm Araştırmaları Dergisi

2020, 4(4): 3465-3485.

DOI: [10.26677/TR1010.2020.573](https://doi.org/10.26677/TR1010.2020.573)

ISSN: 2587-0890 Dergi web sayfası: <https://www.tutad.org>



ARAŞTIRMA MAKALESİ

Turizm Sektöründe Veri Madenciliği

Dr. Öğr. Üyesi Şeyma BOZKURT UZAN, İstanbul Gelişim Üniversitesi, İktisadi İdari ve Sosyal Bilimler Fakültesi, İstanbul, e-posta: suzan@gelisim.edu.tr

ORCID: <https://orcid.org/0000-0003-3527-3730>

Mehmet Ömer ÖZÜÇAĞLIYAN, Doktora Öğrencisi, İstanbul Gelişim Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul, e-posta: mehomerozcag05@hotmail.com

ORCID: <https://orcid.org/0000-0002-0196-8908>

Öz

Teknolojik gelişmelere bağlı olarak veri üretimi katlanarak artmaktadır. İnternet kullanımı, insanların etkileşimlerine bağlı olarak sosyal medya ve internete bağlı cihazlar, veri üretimini artıran etkenlerdir. Yığınla verinin oluştuğu dijital dünyada bu veri yığınları depolanmakta veya anlık olarak akmaktadır. Bu veri yığınlarından potansiyel bilgiyi çıkarmak amacıyla yapılan işlemler silsilesi, veri madenciliği olarak adlandırılmaktadır. Veri madenciliği, günümüzde her alanda etkin bir biçimde kullanılmaktadır. Veri Madenciliğinin kullanıldığı alanlardan biri de turizmdir. Turizmde müşterilere ilişkin kişisel veri, tüketim alışkanlıkları, aylara göre turist hareketliliği, konaklama süresi ve müdavimliğe ilişkin birçok veri bulunmaktadır. Veri yığınları, "Veri Madenciliği" süreci ile işlenerek müşteri profili belirleme, pazarlama stratejileri geliştirme, e-ticaret seçenekleri belirleme, süreç yönetimi, lojistik yönetimi, müşteri memnuniyetini artırma, seyahat güzergahları belirleme, hizmet kalitesini artırma gibi bir dizi temel işlev için kullanılabilir. Bu çalışmada 2015-2019 yılları arasında turizm sektöründe yapılan veri madenciliği çalışmaları incelenmiştir. Bu çalışmaların yıllara, ülkelere, kullanılan modele, yayımlandıkları dergilere, veri kaynaklarına ve çalışma konularına göre incelenmiştir. Araştırmanın sonuçlarına göre en çok makale, 2019 yılında yayımlanmıştır. Son iki yılda bu çalışmaların arttığı söylenebilir. En çok araştırma yapılan ülke Çin olduğu görülmektedir. Uygulama çalışmalarında tahmin edici ve tanımlayıcı modellerin oranı birbirine yakındır. En çok çalışılan konular, turist memnuniyeti ve müdavimlik, turist seyahat davranışlarını, kararlarını belirlemek ve turistler için gezinti, tur rotası belirleme ve önerilerdir.

Anahtar Kelimeler: Veri Madenciliği, Turizm Sektörü, Literatür Taraması.

Makale Gönderme Tarihi: 30.05.2020

Makale Kabul Tarihi: 08.10.2020

Önerilen Atıf:

Bozkurt Uzan, Ş. ve Özüçaglıyan, M. Ö. (2020). Turizm Sektöründe Veri Madenciliği, *Türk Turizm Araştırmaları Dergisi*, 4(4): 3465-3485.

© 2020 Türk Turizm Araştırmaları Dergisi.



Journal of Turkish Tourism Research

2020, 4(4): 3465-3485.

DOI: [10.26677/TR1010.2020.573](https://doi.org/10.26677/TR1010.2020.573)

ISSN: 2587-0890 Journal Homepage: <https://www.tutad.org>



RESEARCH PAPER

Data Mining in Tourism Sector

Assistant Prof. Dr. Şeyma BOZKURT UZAN, İstanbul Gelişim University, Faculty of Economics and Administrative Sciences, İstanbul, e-mail: suzan@gelisim.edu.tr
ORCID: <https://orcid.org/0000-0003-3527-3730>

Mehmet Ömer ÖZÜÇAĞLIYAN, Ph.D. Student, İstanbul Gelişim University, Social Sciences Institute, İstanbul, e-mail: mehomerozcag05@hotmail.com
ORCID: <https://orcid.org/0000-0002-0196-8908>

Abstract

Technological developments are increasing by multiplying data production. Internet use, social media interactions of people and internet-connected devices are factors that increase data production. In the digital world where the data is stored with the stack, these data stacks are stored or flowed instantaneously. The series of transactions made to extract potential information from these data stacks is called Data Mining. Data mining is used effectively in every field today. One of the areas where Data Mining is used is tourism. There are many data regarding personal data, consumer habits, tourist mobility by months, duration of stay and regularity in tourism. Data stacks can be used for a number of basic functions such as customer profile determination, developing marketing strategies, choosing e-commerce options, process management, logistics management, increasing customer satisfaction, determining travel routes, improving service quality by processing with the Data Mining process. In this research, data mining studies in the tourism sector between 2015-2019 have been examined. These studies were examined according to years, countries, models used, journals, data sources, and study topics. According to the results of this study, most articles were published in 2019. It can be said that studies of DM on Tourism have increased in the last two years. It is seen that the most researched country is China. The proportion of predictive and descriptive models in application studies is close to each other. The most studied topics are tourist satisfaction and regularity, determining tourist travel behaviors, decisions, and touring, touring and suggestions for tourists.

Keywords: Data Mining, Tourism Sector, Literature Review.

Received: 30.05.2020

Accepted: 08.10.2020

Suggested Citation:

Bozkurt Uzan, Ş. and Özüçaglıyan, M. Ö. (2020). Data Mining in Tourism Sector, *Journal of Turkish Tourism Research*, 4(4): 3465-3485.

© 2020 Türk Turizm Araştırmaları Dergisi.

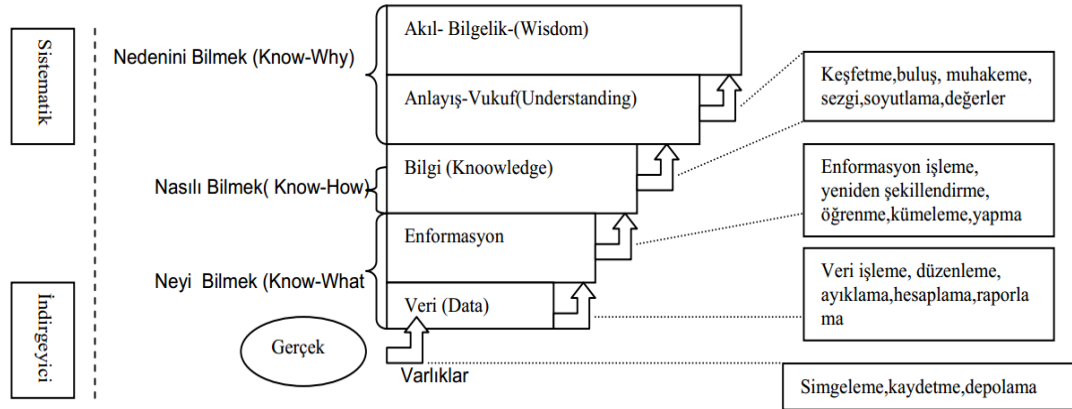
GİRİŞ

Gelişen teknoloji, dünyada gerçekleşen iş ve işlemlerin dijital olarak yapılmasını ve dijital ortamda kayıt altına alınmasını yaygınlaştırmıştır. Bulut teknolojileri ve dijital depolama aygıtlarına kolay ulaşım gibi olanaklar, hem dijital verinin kolay saklanabilmesini sağlamakta, hem de veriye erişimi kolaylaştırmaktadır. Dünya genelinde IDC tarafından yapılan bir çalışmaya göre 2013-2020 yıllarında üretilen dijital veri, 44 ZB'ye (Zetta Bayt) ulaşacaktır. Üretilecek bu veri, dijital olarak 2012 yılına dek üretilen tüm verinin yaklaşık on katı kadar olacaktır (IDC, 2020). Çevrimiçi Science Express isimli dergideki bir makalede dünya üzerinde dijital veri depolama sığasının her üç yılda bir katlandığı ortaya konmuştur (SDN, 2016). Dünya üzerinde bilgisayarların ve akıllı cihazların artmasıyla dünya dijitalleşmekte, tüm iş, işlem ve etkileşimler kayıt altına alınmaktadır. İnternet üzerinden sürekli ve çok yoğun bir veri akışı olmaktadır. Teknolojik gelişmelerle birlikte sosyal medya verileri, çoğalarak büyük hacimli veri yığınlarına (big data) dönüşmektedir. (Karaöz Akın ve Gürsoy, 2018) Büyük ölçekteki veriyi işlemek çok önemli bir hale gelmiştir. Veriden elde edilen bilgiler, günlük hayatımızda alışveriş önerileri, hava durumu tahminleri ve COVID-19'da olduğu gibi virüs salgınına ilişkin tahmin modelleri geliştirmeye kadar birçok alanda kullanılmaktadır.

Veri Madenciliğinin kullanıldığı alanlardan biri de turizmdir. Veri yığınları, "Veri Madenciliği" süreci ile işlenerek müşteri profili belirleme, pazarlama stratejileri geliştirme, e-ticaret seçenekleri belirleme, süreç yönetimi, lojistik yönetimi, müşteri memnuniyetini artırma, seyahat güzergahları belirleme, hizmet kalitesini artırma gibi bir dizi temel işlev için kullanılabilir. Bu çalışmada 2015-2019 yılları arasında turizm sektöründe yapılan veri madenciliği çalışmaları incelenmiştir. Bu çalışmaların yıllara, ülkelere, kullanılan modele, yayımlandıkları dergilere, veri kaynaklarına ve çalışma konularına göre incelenmiştir. Araştırmanın sonuçlarına göre en çok makale, 2019 yılında yayımlanmıştır. Son iki yılda bu çalışmaların arttığı söylenebilir. En çok araştırma yapılan ülke Çin olduğu görülmektedir. Uygulama çalışmalarında tahmin edici ve tanımlayıcı modellerin oranı birbirine yakındır. En çok çalışılan konular, turist memnuniyeti ve müdavimlik, turist seyahat davranışlarını, kararlarını belirlemek ve turistler için gezinti, tur rotası belirleme ve önerilerdir.

VERİ MADENCİLİĞİ

Verinin değerini anlayabilmek, veri ve bununla ilgili kavramlar üzerinde durmak yararlı olacaktır. Veri, enformasyon, bilgi, bilgelik kavramları ve aralarındaki ilişki Şekil 1'de gösterilmiştir. Bu ilişki, veriden bilgiye uzanan sıradüzen içinde bir zincirdir. Veriden bilgiye yükseldikçe elimizdeki yapı sistematikleşmektedir. Bu sıra dizinde, somut gerçeklerden zihinsel soyutlamaya ve sezgisel bir anlayışa doğru evrilme görülmektedir. Veri, olgu, kavram veya komutların, iletişim, yorum ve işlem için elverişli biçimli gösterimidir (TDK, 2020). Veri, sembol, sayı veya sözcüklerle gösterilen işlenmemiş, ham gerçekliklerdir. Ölçüm yapılarak veya sayısal olarak gösterilen bir veriye "nicel veri", sayısal bir gösterimi olmayan, yazı veya sözcüklerle ifade edilen veriye ise "nitel veri" denir (Özen, 2014). Veriden başlayarak enformasyona doğru insan anlamlandırması artar. Veri, bilginin ham şekli, işlenmemiş durumu olarak ifade edilir.



Şekil 1. Anlam Şeması

Kaynak: (Aktan ve Vural, 2005:31).

Enformasyon, insan zihninin karşılaştığı durumlara ilişkin bilinmezlikleri azaltan bir faktördür. Belirsizlik seviyesini azaltan, verinin biçimlendirilmiş ve yapılandırılmış bütünsel bir formu olarak ifade edilebilir (Yılmaz, 2011). Haber alma, haber verme ve haberleşmedir. Bilgi, veriye yöneltilen anlam, insan zihninin algılayabileceği olgu, realite ve bu anlam ilişkisine yön veren ilkeler olarak ifade edilmektedir. Anlayış, “bakış açısı ve değer etmenlerinin şekillendirdiği düşünme şekli, düşünüş biçimi, zihniyet, mantalite” olarak tanımlanmaktadır. Bilgelik ise herkesin ulaşamadığı derin, kapsamlı, bütünsel bilgi olarak tanımlanmakta (TDK, 2020), farklı durumlarda geleceği kestirebilme yeteneği olarak ifade edilmektedir. Bilgelik, özel bir alandaki deneyimleri sonucu kişinin farklı durumlarda değerlendirme ve karar alma konusunda bilgisini kullanma şeklini tanımlanmaktadır (Özen, 2014). Günlük hayatta bu kavramların kullanımında sıkça hatalar yapılmaktadır. Özellikle bilgi ve enformasyon kavramları, sıkça birbiri yerine kullanılmaktadır. Enformasyon, bir durum veya bir konuda var olan belirsizliği gideren betimleyicilerdir, verinin işlenmiş halidir. Bilgi, olayları ve olayları tanımlamanın, olayların nedenlerini veya yol açtığı sonuçları anlamamanın, açıklamanın bir yoludur. Bilgi edinmenin yolları, eğitim, gözlem, araştırma veya deneyim olarak sayılabilir.

Veri Madenciliği Kavramı

Veri madenciliği terimi, literatürde farklı biçimlerde kullanılmaktadır. Kullanılan terimlerden biri, veri tabanlarından bilgi madenciliği (knowledge mining from databases) olarak ifade edilmektedir. Alanyazında veri madenciliği, bilgi çıkarımı-knowledge extraction olarak ifade edilmektedir. Veri madenciliği için veri arkeolojisi ve örüntü (desen, pattern) analizi terimleri de kullanılmaktadır. Bunların dışında veri madenciliği, veri tabanlarından bilgi keşfi (KDD-Knowledge Discovery in Databases) olarak yaygın bir biçimde kullanılmaktadır (Akgöbek ve Kaya, 2011).

Veri madenciliğinin başlangıcı, verinin elektronik olarak depolanmasına dayanmaktadır. Verinin dijital olarak depolanması, ilk defa 1960’lı yıllarda gerçekleşmiştir. Bilgisayarda veri toplanması sayesinde veri üzerinde bilgisayarları kullanarak analiz yapılmaya başlanmıştır. İlişkisel veri tabanları ve onlar üzerinde sorgulama yapmayı sağlayan programlama dillerinin gelişimi, 1980’lerde veri üzerinde dinamik ve senkronize olarak analiz yapılmasını sağlamıştır. Sürekli gelişen teknoloji, 1990’larda çok büyük miktarda veri toplanmasına olanak vermiştir. Veri yığınları, veri ambarları içerisinde depolanmıştır. (Argüden ve Erşahin, 2008). 2000’li yıllarda ise

tüm sektörlerde veri üzerinde yoğun çalışmalar yapılmakta ve veri madenciliği yaygın bir şekilde kullanılmaktadır.

Veri madenciliği, büyük veri yığınlarından bilgi çıkarımı yapılması amacıyla matematik, istatistik ve yapay zekâ yöntemlerinin kullanılması ile gelişmiştir (Argüden ve Erşahin, 2008). Veri madenciliği, genel olarak değişik açılardan veriyi analiz eden ve veriyi gelirleri artırmak, giderleri düşürmek veya her ikisini birden yapmak için kullanışlı bilgiler şeklinde özetleme sürecidir (Frاند, 2016). Veri madenciliği için başka bir tanımda büyük veri yığınlarından geleceğe yönelik kestirimlerin yapılmasına olanak sağlayan ilişki, desen ve kuralların bilgisayarlarda özel programlar yardımıyla bulunmasıdır. (Alpaydın, 2000) Veri madenciliğinde büyük veri depolarından, verilerdeki örüntüleri ve eğilimleri keşfetmek için basit analizlerin ötesinde analizler uygulanmaktadır. Veriden bilgi keşfi olarak da bilinir. Verileri bölümlenmek ve geleceğe yönelik tahminlerde bulunmak için komplike matematiksel algoritmalar kullanır ve dört anahtar kavrama sahiptir (Oracle, 2020). (Desenlerin keşfi, olası çıktılar tahmin etme, işe yarar bilgi üretilmesi, büyük veri setlerine/veri tabanları üzerinde çalışma.)

Veri madenciliğinden karar verme sürecinde de önemli ölçüde yararlanır. Büyük veri içinden potansiyel olarak anlamlı, yararlı ve uygulanabilecek bilgilerin elde edilmesi amacıyla veri madenciliğinden yararlanılmaktadır. Veri madenciliği süreci ve sürecin sonunda hedeflenen bilgiler, kritik özelliktedir. Bu bilgi, ilk başta örtük, veri madenciliği sürecinden önce bilinmeyen, önceden keşfedilmemiş bir bilgidir. Elde edilen bilgi, kullanım potansiyeli olan anlamlı bir bilgidir (Argüden ve Erşahin, 2008).

Veri Tabanlarından Bilgi Keşfi Olarak Veri Madenciliği

Veri madenciliği, ilk kez veri tabanları üzerinden bilgi keşfetme amacıyla kullanılmıştır. Veri tabanlarından bilgi keşfetme süreci beş aşamada özetlenebilir.

Veri Seçimi (Selection): Veri seçimi aşamasında süreçte kullanılmayacak veri ayıklanarak analiz dışı bırakılır.

Veri Ön işleme (Data Preprocessing): Veri seçimi aşamasında seçilen verinin incelenerek hatalı kayıtların, değerlerin, eksik bölümlerin ayıklandığı, giderildiği aşamadır. Veri seti üzerinde ilgisiz özellikler çıkarılır ve kendini tekrarlayan veri ayıklanır. Bu aşamada analiz için uygun, stabil bir veri seti elde edilmesi, süreç sonunda elde edilecek bilginin kalitesini artırır.

Dönüştürme (Transformation): Bu adımda veriyi daha kullanışlı hale getirmek için gerekli dönüşümler yapılır. Bu aşamada veri setinde sadece ilgili özelliklerin bırakılması, veri madenciliği işlemlerinde sürecin daha hızlı işlenmesini sağlar, analiz dışı özelliklerin sistemdeki yükünü kaldırır.

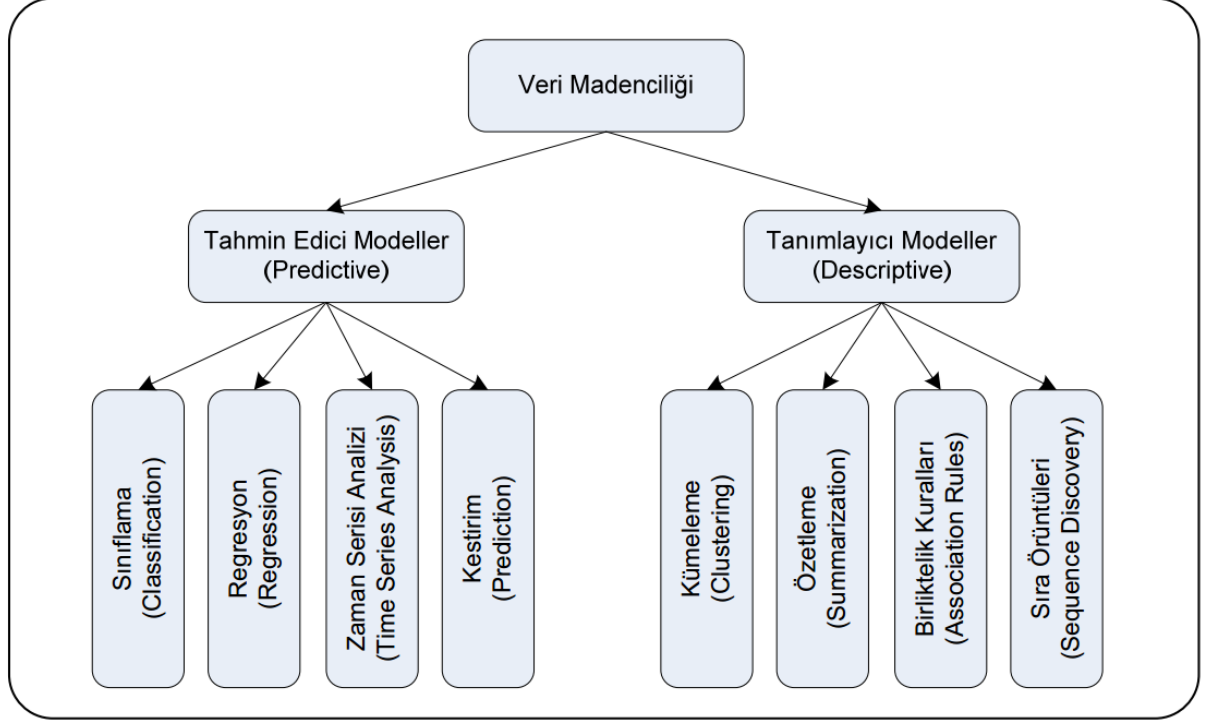
Veri Madenciliği (Data Mining): Modelleme olarak da adlandırılan bu aşamada veri üzerinde uygun teknikler kullanılarak potansiyel bilgiye ulaşmayı sağlayacak modeller oluşturulur. Bu, veri madenciliğinin amacına yönelik olarak bir tahmin, betimleme veya birliktelik modeli olabilir.

Değerlendirme (Evaluation): Bu aşama, modelleme sonucu elde edilen bilginin değerlendirilme aşamasıdır. Bu bilginin kritikliği, yeniliğin sağlayacağı katkı ve değerlendirmede önemlidir (Sever ve Oğuz, 2002). Veri madenciliği, bilgi yönetimi sürecinin bir parçası olarak da ele alınmaktadır. Literatürde bilgi yönetim süreci: Bilginin düzenlenmesi, damıtılması ve yayılması olarak el alınmaktadır. Bilginin damıtılması sürecinde en temel işlem, veri madenciliği uygulamalarıdır (Sağsan, 2014).

Veri madenciliği süreçlerinde bu adımlar izlenerek birçok avantaj elde edilmektedir. Bu adımlar sayesinde sürecin hızlı ve verimli yönetimi sağlandığı gibi düşük maliyetleri de meydana getirmektedir.

Veri Madenciliğinin İşlevleri ve Kullanılan Modeller

Dunham (2003)'e göre veri madenciliğinin kestirim ve betimleme olmak üzere iki ana işlevi vardır. Bu iki ana işlevinin altında yer alan farklı modeller Şekil 2'de gösterilmiştir.



Şekil 2. Veri Madenciliği Modelleri Dunham 2003'den uyarlanmıştır.

Tahmin edici modellerde sonuçlar, bilinen olay, durum ve özellikler kullanılır. Veri üzerinden özelliklerle sonuçları eşleştiren modeller, kuralları, desenleri veya örtük yapıları bulmayı sağlar. Modeller, yeni durumların, olayların sonuçlarını tahmin etmek için kullanılır (Aydın, 2007). Tanımlayıcı modellerin kullanımında amaç, veri setinin durumunu en iyi şekilde betimlemektir. Veri setindeki nitelikler arasındaki örüntü, desen ve bağlantıları bulmaktır. Tanımlayıcı modeller, veri setindeki kayıtların nasıl özellikler gösterdiğini ve hangi kümelere dahil olduğunu gösterir. Bu şekilde eldeki verinin ne ifade ettiği ortaya çıkar (Argüden ve Erşahin, 2008). Bu modellere ilişkin ayrıntılı bilgi aşağıda verilmiştir.

Tahmin (Predictive) Modelleri

Tahmin ya da diğer adıyla kestirim modelleri, geleceğe yönelik bir bakıştır. Bu modeller altında yer alan alt modeller aşağıda sıralanmıştır.

Sınıflama (Classification)

Veri üzerinde yapılan işlem sonucu bir sınıflandırma yapılırsa bu, olay, kayıt, veri veya kişinin sahip olduğu sınıfın özelliklerine uygun hareket edeceğinden hareket ile önemli bilgiler elde edilir. "Bir bankanın 3 yıldan uzun süredir evli, en az 5 yıldır bir işte çalışan, orta gelirli ve faturalarını düzenli ödeyen erkekleri kredi notu yüksek müşteriler sınıfına atması" buna bir örnektir

Regresyon (Eğri Uydurma)

Bir öğrencinin ara sınav notlarıyla mezuniyet notu arasında bir ilişkiyi ifade edebilecek bir fonksiyon, bir matematiksel bağlantı bulmak, regresyona örnek olarak verilebilir. Regresyon, bağımsız değişkenin bağımlı değişkenlerle ilişkisini bir fonksiyon yardımıyla ifade etmek için kullanılır. Bu model kullanılarak bir öğrencinin sadece ara sınav puanlarına bakılarak mezuniyet notu kestirilebilir (Altun, 2019)

Zaman Serisi Analizleri (Time Series Analysis)

Veri üzerinde yıllara göre değişimleri ve gelecek yıllardaki olası durumları inceleyen modellerdir. Bir şirketin gelecek yıldaki kâr miktarının tahmin edilmesi örnek olarak verilebilir (Özcan, 2014). Borsadaki değişimler ve yabancı para birimlerindeki kur değişiklikleri, zaman serisi analizlerinin kapsadığı konulardır.

Kestirim (Prediction)

Geleceğe yönelik diğer kestirimleri içeren modellerdir. Bir yerde meydana gelebilecek sel baskını kestirebilmek buna örnektir. Bunun için yağış miktarı, nem, dere yataklarının durumu, ağaç miktarı ve eğim gibi özelliklere ve sel durumuna ilişkin veri kullanılarak bir kestirim modeli geliştirilebilir (Aydın, 2007).

Tanımlayıcı (Descriptive) Modeller

Bu kategoride yer alan veri setinin özelliklerini ortaya koymak, onu tanımlamak için kullanılan modellere ilişkin bilgiler aşağıda sıralanmıştır.

Kümeleme (Clustering)

Benzer özellikler gösteren kayıt, kişi, durum ve verinin aynı grupta yer alması, benzer özellikler göstermeyenlerin ise farklı gruplarda yer almasını sağlayan, kendi içinde homojen gruplar oluşturan bir modeldir (Özcan, 2014). Bu model ile her gruba ait belirgin özellikler ortaya çıkarılmakta ve her grup bu özellikler kullanılarak tanımlanabilmektedir.

Özetleme (Summarization)

Bir bakışta veri setine ilişkin enformasyon sağlayan modeldir. Bir veri setinin ortalaması, taban ve tavan değerleri buna örnek olarak verilebilir. Tüm veri seti incelenmeden verinin ne ifade ettiği anlaşılabilir. (Çüllüoğlu Gülen, 2014)

Birliktelik Analizi (Association Rules)

Özellikle marketing çalışmalarında kullanılan, A ürününü alan kişilerin yüzde x kadarı aynı zamanda B ürününü de alır şeklinde modellemelerdir (Argüden ve Erşahin, 2008). Birliktelik analizi sahtekarlık tespitinde de kullanılmaktadır (Özcan, 2014).

Sıra örüntüleri (Sequence Discovery)

Birliktelik kurallarına benzemekle birlikte farkı, iki olayın veya durumun aynı anda değil birbiri ardına gerçekleşme bağlantısının modellenmesidir. Durumlar arasındaki ilişki, zamana bağlıdır (Aydın, 2007). Beyaz eşya satan bir işletmede buzdolabı alanların belirli bir oranının daha sonra bulaşık makinesi almasını tespit edilmesidir (Argüden ve Erşahin, 2008).

Literatürde tahmin modelleri ve tanımlayıcı modeller ile ilgili birçok çalışma yapılmıştır. Kullanım alanlarına göre model seçimleri değişiklik gösterebilmektedir.

Veri Madenciliğinin Kullanım Alanları

Veri madenciliği, birçok alanda ve sektörde yaygın bir biçimde kullanılmaktadır. Kullanıldığı başlıca alanlar ve kullanım amaçları aşağıdaki gibi özetlenebilir.

- **Bilimsel Çalışmalarda:** Yapılan bilimsel çalışmalarda çok miktarda veri üretilmektedir. Bu veri yığınının içinde yapılan çalışmalarda sonuçlara ulaşmak için yeni yöntemler denenmektedir. Bilimsel çalışmalarda elde edilen büyük verinin işlenebilmesi için veri madenciliği yöntemlerinden sıkça yararlanılmaktadır. (Akgöbek ve Çakır, 2009)
- **Bankacılık ve Finans:** Bu alanda müşterilerin profillerinin belirlenmesi, kredi skorlarının hesaplanması, risk durumlarının belirlenmesi, ürün sınıflandırmaları ve satışa sunma gibi işlemleri yapmak için veri madenciliğinden yararlanılmaktadır. Özellikle bankalarda veri madenciliği tekniklerini kullanarak müşterilerinin ödeme durumlarını kestirmek, kredi ihtiyaçlarını belirlemek ve buna uygun paketler sunmak, kredi onayı vermek gibi durumlarda veri madenciliğinden sıkça yararlanılmaktadır (Özcan, 2014). Bu sektörde veri madenciliği, müşterileri profillerine göre sınıflandırmak ve bu sınıflara özel kampanyalar düzenlemek gibi çalışmalar için kullanılmaktadır. Finansal göstergelerin takip edilerek finansal durumun kestirilmesi önem arz etmektedir. Hisse senetlerinin değerlerinin değişiminin kestirilmesi, kredi kartı dolandırıcılıklarının tespiti ve döviz kurlarının değişiminin kestirilmesi gibi durumlarda veri madenciliğinden yararlanılmaktadır. (Aydın, 2007)
- **Müşteri İlişkileri Yönetimi:** Veri madenciliği, müşteri profillerinin belirlenmesi, sistemi terk edecek müşterilerin tespiti, müşteri gruplarına yönelik özel kampanyaların belirlenmesi gibi amaçlar için kullanılmaktadır (Özcan, 2014:43-44). Müşterilerin iyi tanınmasını sağlayacak modeller geliştirilerek işletmelerde müşteri memnuniyetini artıracak çalışmalar yapılabilir.
- **Sahtekârlık Tespiti:** Tüm sektörlerde oluşabilecek anormalliklerin bulunması ve sahtekârlıkların tespiti amacıyla veri madenciliği etkin bir şekilde kullanılmaktadır. Sigortacılık sektöründe yapılan şüpheli işlemlerin analizi, kredi kartı ve nakit para çekme makinelerinde yapılan sahtekarlıkların tespiti veri madenciliğinin kullanıldığı alanlara örnektir (Özcan, 2014:43-44).
- **Güvenlik/İstihbarat:** Terör olaylarında, istihbarat elde etmek amacıyla, sosyal ağlardaki akışı, kilit kişileri bulmak için veri madenciliği kullanılmaktadır. Suç örgütleri ve olayları arasındaki ilişkinin ortaya çıkması için toplanan verinin işlenmesi, sonuca ulaşılması için veri madenciliği etkili bir yöntem olarak kullanılmaktadır. (Özcan, 2014; Aydın, 2007)
- **Sağlık Alanında Kullanımı:** Sağlık, veri madenciliğinin en yaygın kullanıldığı alanlardan biridir. Hastalara ilişkin tutulan verinin kullanılarak, hastalığın seyrinin kestirilmesi veya tanı konulması gibi işlemlerde veri madenciliği önemli bilgiler sağlamaktadır (Dalkılıç ve Türkmen, 2002). Yine hastalıklar ile bireyin fiziksel durumu, beslenme alışkanlıkları ve diğer özellikleriyle bağlantıların tespitinde veri madenciliği kullanılmaktadır.
- **Eğitim Alanında Kullanım:** Eğitim alanında veri madenciliği kullanımı, öğrencilerin gelecekteki akademik başarı durumlarının kestirilmesi, öğrenci profillerinin belirlenmesi ve öğrenci risk gruplarının belirlenmesi gibi amaçlarla kullanılmaktadır (Altun, 2019).

Kullanım alanları yukarıda sayılanlarla sınırlı olmamakla birlikte, en çok kullanılan alanlar sıralanmıştır. Bu çalışmada ise turizm sektöründe yapılan veri madenciliği uygulamaları ele alınmıştır ve incelenmiştir.

Turizmde Veri Madenciliği Uygulamaları

Veri Madenciliği çalışmaları, genel olarak tüm sektörlerde etkin bir biçimde kullanılmaktadır. Bu araştırmada özel olarak veri madenciliğinin turizm sektöründe kullanımına ilişkin çalışmalar incelenmiştir. Özet tablolar oluşturularak, araştırmacılara sunulmuştur.

YÖNTEM

Bu araştırmada turizm sektöründe yapılan veri madenciliği çalışmalar incelenmiştir. Yapılan çalışmalar, Scopus, Ebsco, Emerald ve Taylor Francis veri tabanlarında, hakemli dergilerde, tam metin erişime açık makaleler içinde, 2015-2019 yılları arasında yayımlanan makalelerde taranmıştır. Tarama yapılırken özet alanında “data mining” ve “tourism” anahtar kelimeleri birlikte aratılmıştır. Tarama sonucu çıkan ve listelenen makalelerin özetleri okunarak konuyla ilgili olanlar ayrılmıştır.

BULGULAR

Yapılan literatür taraması sonucu Scopus, Ebsco, Emerald ve Taylor Francis veri tabanlarında erişilen makaleler arasından “turizmde veri madenciliği” konusuyla ilgili olanlar Tablo 1’de listelenmiştir.

Tablo 1. Turizm Sektöründe VM Makaleleri

| SN | Yazar / Yazarlar | Konu | Kullanılan Teknik | Veri Kaynağı | Dergi | Yıl | Ülke |
|----|-------------------------|--|--|--|--|------|-----------|
| 1 | Zhang | Turistler için gezinti, tur rotası belirleme/öneri | Tanımlayıcı/Kümeleme (karınca kolonisi optimizasyonu) | Coğrafi veri seti, haritalar ve rotalar | Journal of Coastal Research | 2019 | Çin |
| 2 | Amjad ve Croock | Turist çeken bölgelerin belirlenmesi | Tahmin Edici/Sınıflama (Lojistik Regresyon) | Sosyal Ağ | Telkomnika | 2019 | Irak |
| 3 | Bin, Gu, Sun, Chang | Turistler için gezinti, tur rotası belirleme/öneri | Tanımlayıcı/Sıra Örüntüleri | Web sayfaları | Multimedia Tools and Applications | 2019 | Çin |
| 4 | Zhang, Chen ve Li | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tahmin Edici/Sınıflama (derin öğrenme) | Coğrafi etiketli Sosyal Ağ Verisi (Flickr) | Tourism Management | 2019 | Çin |
| 5 | Padmaja, N., Sudha, T | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tahmin edici/Birliktelik Kuralları | Web sayfaları (yorumlar) | International Journal of Recent Technology and Engineering | 2019 | Hindistan |
| 6 | Tao, Zhang, Shi ve Chen | Turist memnuniyeti-müdavimlik | Tahmin edici/Sınıflama (Doğal dil işleme, duygu analizi, yapay sinir ağları) | Web sayfaları (yorumlar) | Sustainability (Switzerland) | 2019 | Çin |

| | | | | | | | |
|----|--|--|--|---|--|------|------------|
| 7 | Song, Kawamura, Uchida ve Saito | Turist memnuniyeti-müdavimlik | Tahmin edici/Sınıflama (duygu analizi, metin madenciliği) | Web sayfaları (yorumlar) ve anket | Information Technology and Tourism | 2019 | Japonya |
| 8 | Li, Li, Zhang, Hu ve Hu | Turist memnuniyeti-müdavimlik | Tahmin edici/Sınıflama (duygu analizi, metin madenciliği) | Literatür taraması | Applied Sciences (Switzerland) | 2019 | Çin |
| 9 | Su, Stepchenkova, ve Kirilenko | Turist memnuniyeti-müdavimlik | Tahmin edici/Sınıflama (duygu analizi, metin madenciliği) | Web sayfaları (yorumlar) | Tourism Management | 2019 | Çin |
| 10 | Tkach, Pavlenchuk, Sadchenko, Ni kola, Drozdova ve Davydenko | Turizm talebinin tahmini | Tahmin edici/Kestirim | Veri tabanı | International Journal of Recent Technology and Engineering | 2019 | Ukrayna |
| 11 | Zhou, Zhan, Feng, Zhang ve Li | Turistler için gezinti, tur rotası belirleme/öneri | Tahmin Edici/Sınıflama (karar ağaçları) | Web sayfaları (yorumlar) | ISPRS International Journal of Geo-Information | 2019 | Çin |
| 12 | Zeng, Nakatoh, Hirokawa ve Eguchi | Turist memnuniyeti-müdavimlik | Tanımlayıcı/Kümeleme (metin madenciliği) | Web sayfaları (yorumlar) | IEEE Transactions on Electrical and Electronic Engineering | 2019 | Japan |
| 13 | Bigné, Oltra ve Andreu | Turizm talebinin tahmini | Tahmin edici/Sınıflama (metin madenciliği, yapay sinir ağları) | Coğrafi etiketli Sosyal Ağ Verisi (Twitter) | Tourism Management | 2019 | İspanya |
| 14 | Zhu, Sun, Yuan, Hu ve Miao | Turistler için gezinti, tur rotası belirleme/öneri | Tahmin edici/Sınıflama | Telefon arama kayıtları | ISPRS International Journal of Geo-Information | 2019 | Çin |
| 15 | Karapınar, Papapanos ve Kizos | Turist memnuniyeti-müdavimlik | Tanımlayıcı/Kümeleme (metin madenciliği) | Görüşme formu | Sustainability (Switzerland) | 2019 | Yunanistan |
| 16 | Fei, Zhenlong, Chaowei ve Yongyao | Turistler için gezinti, tur rotası belirleme/öneri | Tanımlayıcı/Kümeleme | Coğrafi etiketli Sosyal Ağ Verisi (Twitter) | Cartography and Geographic Information Science | 2019 | ABD |
| 17 | He ve Zhaojun | Turizm talebinin tahmini | Tahmin Edici/Kestirim (Naive Bayesian) | İstatistikler | Enterprise Information Systems | 2019 | Çin ABD |
| 18 | Ghalekhonda bi, Ardjmand, Young ve Weckman | Turizm talebinin tahmini | Literatür Taraması | Literatür | Journal of Tourism Futures | 2019 | ABD |

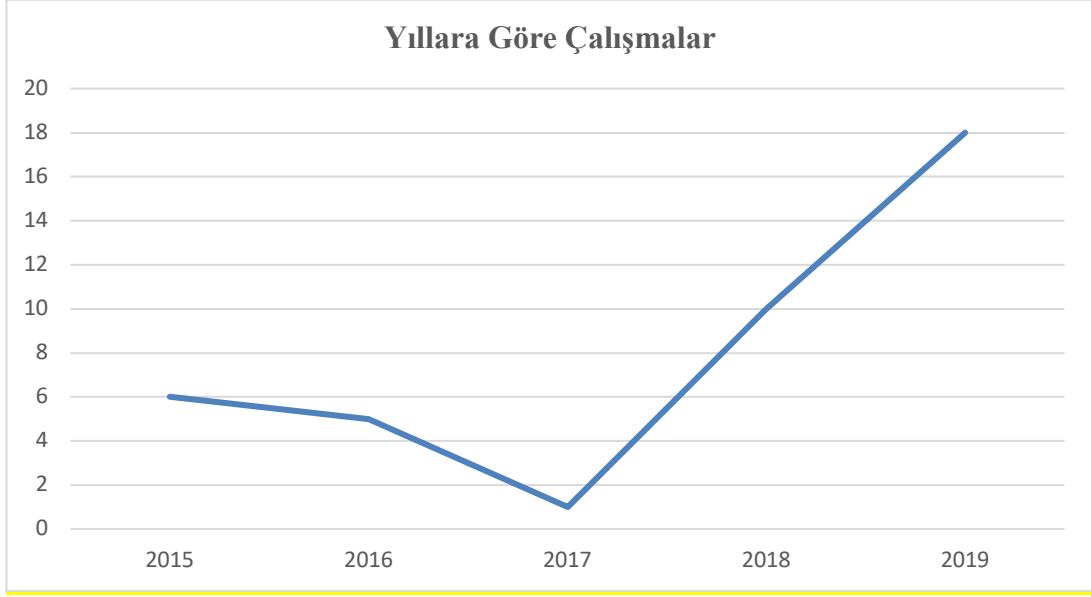
| | | | | | | | |
|----|---|--|---|--|---|------|------------|
| 19 | Kozak, Rita ve Bigné | Turist memnuniyeti-müdavimlik | Tanımlayıcı/Kümeleme (metin madenciliği) | Web sayfaları (yorumlar), mobil cihaz verisi | European Journal of Management and Business Economics | 2018 | İspanya |
| 20 | Wan, Hong, Huang, Peng ve Li | Turistler için gezinti rotası belirleme/öneri | Tahmin edici/Sınıflama Algoritmaları (Naive Bayesian) | Coğrafi etiketli Sosyal Ağ Verisi (Flickr) | International Journal of Geographical Information Science | 2018 | Çin |
| 21 | Lenormand vd. | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Kümeleme (Hiyerarşik kümeleme) | Coğrafi etiketli Sosyal Ağ Verisi (Flickr) | PLoS ONE | 2018 | Avrupa |
| 22 | Pineda-Bravo, Pérez-García ve Cristóbal Martínez-Martínez | Turizm yönetiminde VM uygulaması | Tahmin Edici/Sınıflama | Veri tabanı (Web olay günlükleri) | Retos Turísticos | 2018 | Küba |
| 23 | Huang, Wang, Yang ve Xu | Turist çeken bölgelerin belirlenmesi | Tahmin Edici/Sınıflama (Naive Bayesian, metin madenciliği) | Web sayfaları (yorumlar) | Intelligent Data Analysis | 2018 | Çin |
| 24 | Kirilenko ve Stepchenkova | Turizm alanındaki çalışmaların VM ile analiz edilmesi | Tahmin Edici/Sınıflama (Naive Bayesian, metin madenciliği) | Turizm alanında yapılan araştırmalar | PLoS ONE | 2018 | ABD |
| 25 | Qi, Wong, Chen, Rong ve Du | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Kümeleme (metin madenciliği) | Web sayfaları (yorumlar) | Information Technology and Tourism | 2018 | Avustralya |
| 26 | Fazzolari, Petrocchi (2018). | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Kümeleme ve Birliktelik (doğal dil işleme, metin madenciliği) | Web sayfaları (yorumlar) | Information Technology and Tourism | 2018 | İtalya |
| 27 | Lin, Yuming, Zhou, Xia ve Ran | Turistler için gezinti, tur rotası belirleme/öneri | Tanımlayıcı/Kümeleme Tahmin edici/Sınıflama | Coğrafi etiketli Sosyal Ağ Verisi (Flickr) | International Journal of Geographical Information Science | 2018 | Çin |
| 28 | Cavagnaro, Staffieri ve Postma | Turist memnuniyeti-müdavimlik | Tanımlayıcı/Kümeleme | Anket | Journal of Tourism Futures | 2018 | Hollanda |
| 29 | Richard | Turizm yönetiminde VM uygulaması | Literatür Taraması | Literatür | Journal of Tourism Futures | 2017 | ABD |
| 30 | Cankurt ve Subaşı | Turizm talebinin tahmini | Tahmin Edici/Regresyon (Regresyon, Yapay sinir Ağları, Destek Vektörleri) | İstatistikler | Turkish Journal of Electrical Engineering & Computer Sciences | 2016 | Türkiye |

| | | | | | | | |
|----|---|--|--|---|--|------|------------|
| 31 | Chua, Servillo, Marcheggiani ve Moere | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Kümeleme ve Birliktelik Kuralları (doğal dil işleme, madencilik) | Coğrafi etiketli Sosyal Ağ Verisi (Twitter) | Tourism Management | 2016 | İtalya |
| 32 | Li, Xiao, Ye, Xu ve Law | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Kümeleme (metin madencilik) | Coğrafi etiketli Sosyal Ağ Verisi | Tourism Management | 2016 | Çin |
| 33 | Pineda-Bravo, Pérez-Alfonso, ve Pérez-García | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tanımlayıcı/Sıra Örüntüleri | Veri tabanı olay günlükleri | Retos Turísticos | 2016 | Küba |
| 34 | Chang, Chen, Kuo, Hsu ve Cheng | Turist memnuniyeti-müdavimlik | Tahmin edici/Sınıflama (Karar ağaçları) | Anket | Anatolia: An International Journal of Tourism & Hospitality Research | 2016 | Çin |
| 35 | Mastorakis, Trihas, Perakakis ve Kopanakis | Turist memnuniyeti ve müdavimlik | Tanımlayıcı/Kümeleme (metin madencilik) | Veri tabanı (Web uygulaması e-ticaret) | Anatolia: An International Journal of Tourism & Hospitality Research | 2015 | Yunanistan |
| 36 | Zhang ve Huang | Turist seyahat davranışlarını, kararlarını belirleme/öneri | Tahmin Edici/Sınıflama (Naive Bayesian, Yapay Sinir Ağları, Karar Ağaçları) | Anket (Web) | Applied Artificial Intelligence | 2015 | Çin |
| 37 | Höpken, Fuchs, Keil, ve Lexhagen | Turizm yönetiminde VM uygulaması | Tanımlayıcı/Özetleme | Veri tabanı | Information Technology and Tourism | 2015 | İsveç |
| 38 | Vu, Li, Law ve Ye | Turizm yönetiminde VM uygulaması | Tanımlayıcı/Kümeleme | Coğrafi etiketli Sosyal Ağ Verisi | Tourism Management | 2015 | Çin |
| 39 | Versichele, Claeys Bouuaert, Neutens, Moerman ve Van de Weghe | Turistler için gezinti, tur rotası belirleme/öneri | Tanımlayıcı/Birliktelik Kuralları | Veri tabanı | Tourism Management | 2015 | Belçika |
| 40 | Li ve Law Vu ve Rong | Turizm yönetiminde VM uygulaması | Tanımlayıcı/Kümeleme | Veri Tabanı | Tourism Management | 2015 | Çin |

Yapılan çalışmaların arasında 18 tanesinde tanımlayıcı yöntemler, 19 tanesinde tahmin edici yöntemler 1 tanesinde hem tanımlayıcı hem tahmin edici yöntemler kullanılmıştır ve 2 çalışmada literatür kullanılarak gerçekleştirilmiştir. Araştırmanın sonuçlarına göre en çok makale, 2019 yılında yayımlanmıştır. Son iki yılda bu çalışmaların arttığı söylenebilir. En çok araştırma yapılan ülke Çin olduğu görülmektedir.

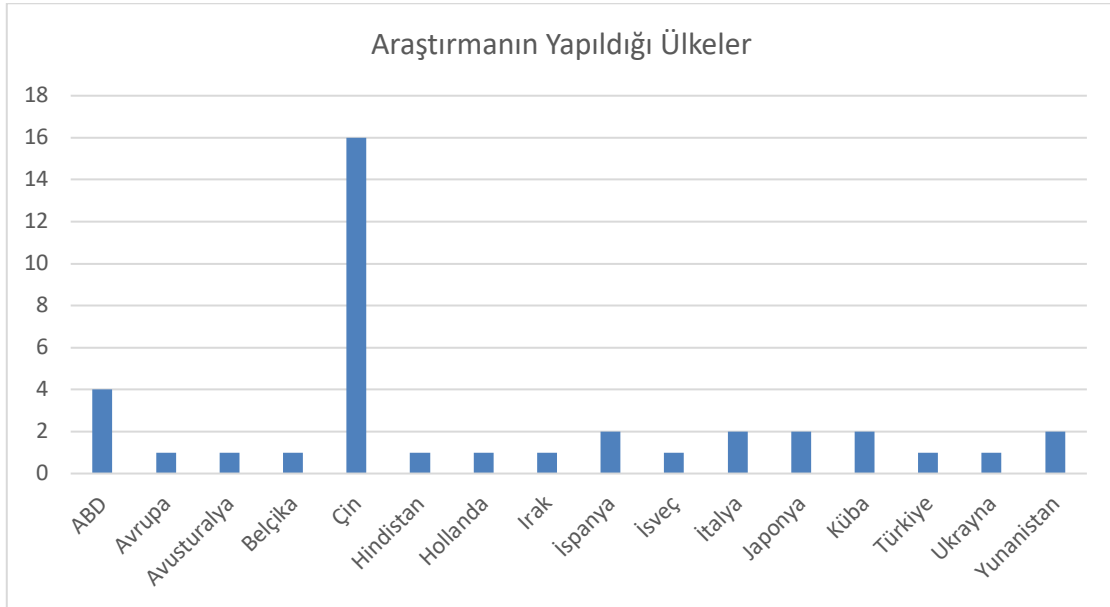
Şekil 3'e bakıldığında 2015 yılına ait verilerin incelenmesiyle başlayan bu analizde, 2016 yılındaki çalışma sayısında düşüş görülmektedir. 2017'de ise düşüş daha da artmaktadır. 2018'de eğri yükselişe geçmiştir. 2019 yılında ise en yüksek değere ulaşan eğride en çok çalışma yapılan yıl,

2019 olarak yorumlanmaktadır. Yıllara göre çalışmalar incelendiğinde en çok çalışmanın, (n=18) 2019 yılında yapıldığı görülmektedir.



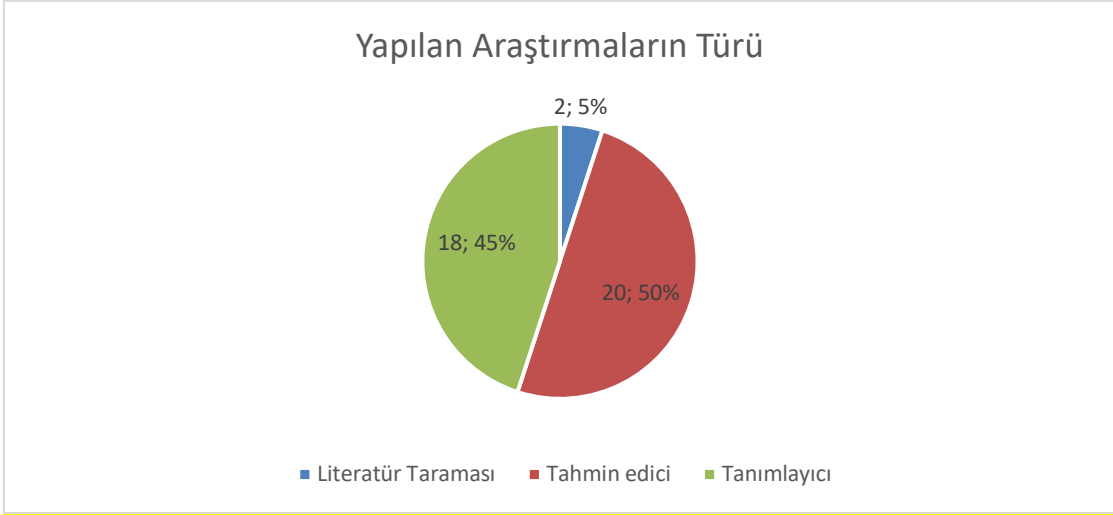
Şekil 3. Yıllara Göre Araştırma Sayısı

Araştırmanın yapıldığı ülkelere göre incelendiğinde (Şekil 4) en çok araştırmanın yapıldığı ülkenin Çin (16) olduğu görülmektedir. Çin'i sırasıyla ABD, İspanya, İtalya, Japonya, Küba ve Yunanistan takip etmektedir.



Şekil 4. Araştırmanın Yapıldığı Ülkeler

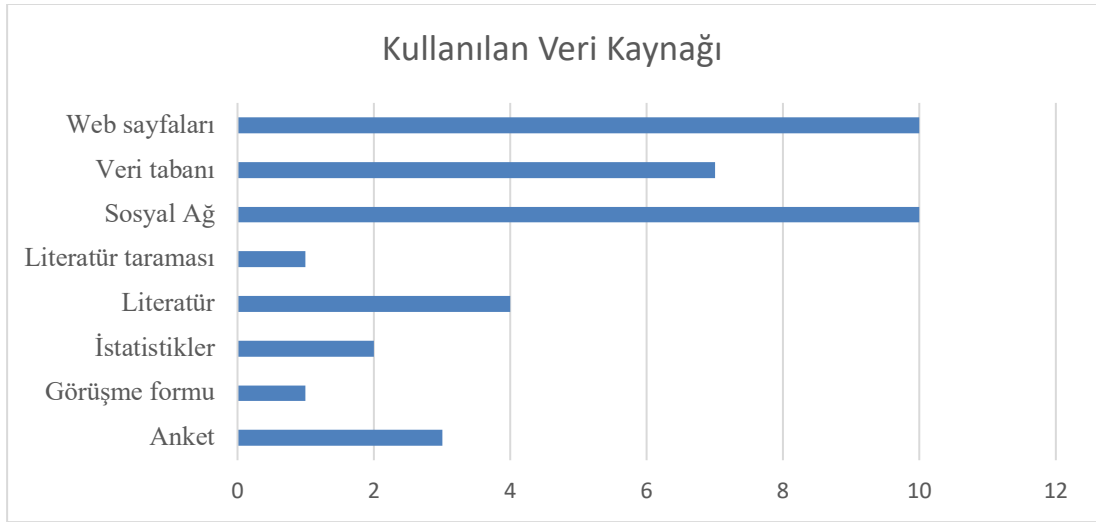
Yapılan araştırmaların türüne bakıldığında (Şekil 5) en çok, tahmin edici modellerin (n=20) geliştirildiği çalışmaların yapıldığı, sonrasında ise tanımlayıcı modellere ait (n=18) çalışmaların yer aldığı görülmektedir. En son sırada ise literatür taramasına ait çalışmaların yapıldığı görülmektedir.



Şekil 5. Yapılan Araştırmaların Türü

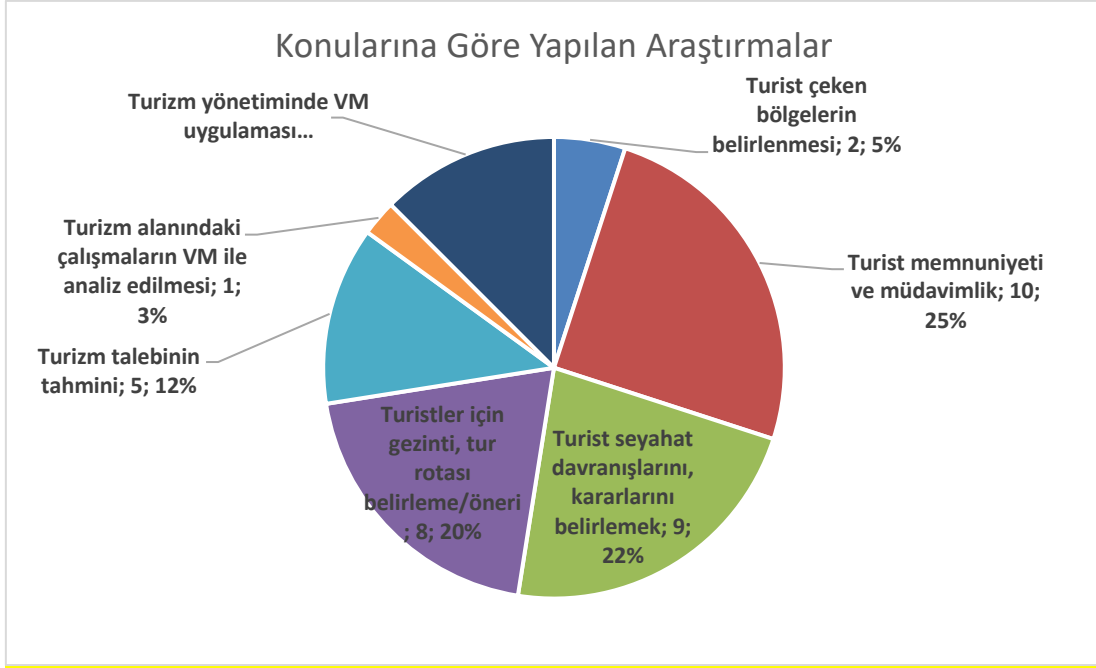
*Çalışmada birden fazla yöntem kullanılmışsa baskın olan yöntem dikkate alınmıştır.

Yapılan araştırmalar kullanılan veri kaynaklarına göre incelendiğinde (Şekil 6) araştırmalarda en çok, web sayfaları ve sosyal ağ verisinden ve veri tabanlarından yararlandığı görülmektedir.



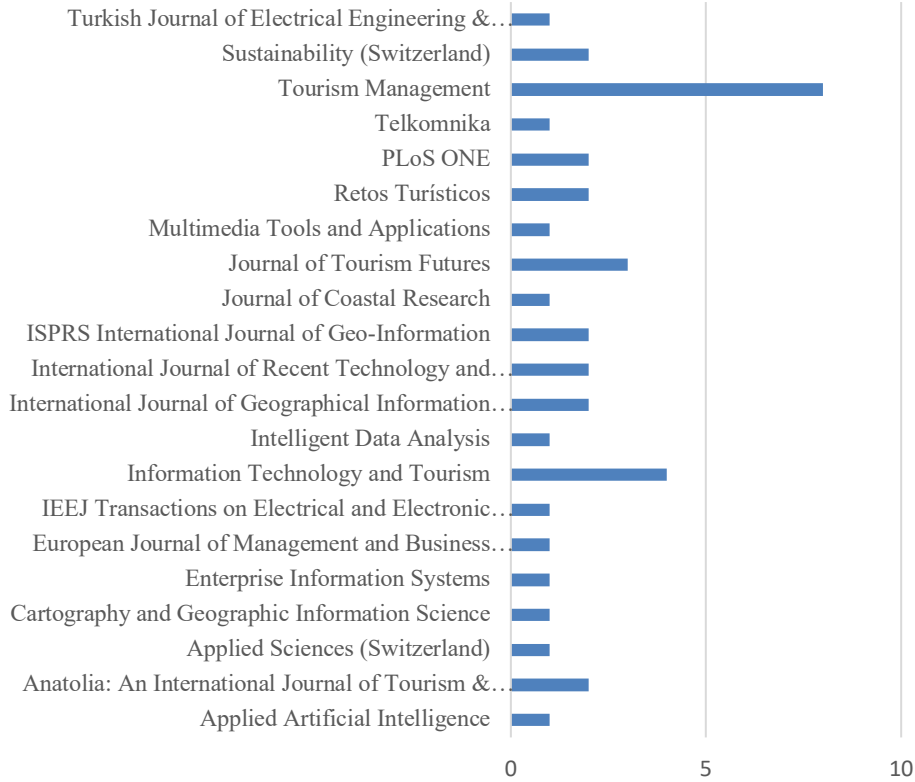
Şekil 6. Kullanılan veri kaynağı *birden fazla veri kaynağı kullanan çalışmalar olabilir.

Konularına göre incelendiğinde ise (Şekil 7) en çok çalışılan konuların, turist memnuniyeti ve müdavimlik, turist seyahat davranışlarını, kararlarını belirleme ve turistler için gezinti, tur rotası belirleme/öneri olduğu görülmektedir.



Şekil 7. Konularına Göre Araştırmalar

Araştırmaların Yayımlandığı Dergiler



Şekil 8. Araştırmanın Yayımlandığı Dergiler

Yapılan araştırmaların yayımlandığı dergilere bakıldığında, (Şekil 8) çalışmaların en çok Tourism Management (n=8) dergisinde olduğu görülmektedir. Tüm tablolar ve grafikler incelendiğinde en çok yayın yapılan yıl 2019 yılı, en çok yayın yapılan ülke Çin olduğu tespit edilmiştir.

TARTIŞMA, SONUÇ ve ÖNERİLER

Her alanda yığına verinin depolandığı günümüzde veriyi kullanarak yararlı bilgiler elde etmek çok önemli hale gelmiştir. Verideki gizli bilgileri ortaya çıkarmak ve buna göre stratejiler oluşturmak, işletmelerin ve yöneticilerin rekabet gücünü artırmaktadır.

Günümüzde sağlıktan, turizme hemen her alanda veri madenciliği çalışmaları yapılmaktadır. Veri madenciliği, işletmelerin maliyetlerini azaltmada, müşteri memnuniyetini ve kârlarını artırmada kullandıkları etkili bir yaklaşımdır. Elde edilen bilgiler, yöneticilere doğru kararlar verebilmeleri için rehberlik etmekte, “veriye dayalı” bir yönetim anlayışını da şekillendirmektedir. Ortaya çıkan modeller, veriler üzerinde test edilerek, gerçek hayata uygulanmadan önce test edilme (simulation) imkânı bulmaktadır.

Yapılan çalışmada turizmin araştırma sınırlılıkları içinde ve 2015-2019 yılları arasında taranan veri tabanlarında turizmdeki veri madenciliği sorgusuyla aratılan makalelere ilişkin aşağıdaki sonuçlar elde edilmiştir.

2015-2019 yılları arasında en çok makale, 2019 yılında yayımlanmıştır. 2018 yılı da dikkate alındığında son iki yılda çalışmaların arttığı söylenebilir. En çok araştırma yapılan ülke Çin’dir. Uygulama çalışmalarında tahmin edici ve tanımlayıcı modellerin oranı birbirine yakındır. İlgili makaleler, en çok Tourism Management dergisinde yayımlanmıştır. En çok, web sayfaları ve sosyal medya veri kaynağı olarak kullanılmaktadır. En çok çalışılan konular, turist memnuniyeti ve müdavimlik, turist seyahat davranışlarını, kararlarını belirlemek ve turistler için gezinti, tur rotası belirleme ve önerilerdir.

İncelenen çalışmalar, birçok özelliği doğrultusunda analiz edilerek bu alanda çalışma yapmak isteyen araştırmacılara ışık tutmaktadır. Çalışmanın içerisinde yer alan makalelerdeki çeşitlilik ve özelliklerde yine bu alanda çalışma yapmak isteyen araştırmacılar için rehber niteliği taşımaktadır. Bu çalışma, veri madenciliği alanında çalışmalarına yeni başlayan araştırmacılar ve öğrenciler için, bu alanda yol gösterici bir niteliğe sahiptir.

Bu çalışma Veri Madenciliği tekniklerinin turizm sektöründe ne kadar etkin bir şekilde kullanıldığını ortaya koymaktadır. Aynı zamanda bu çalışma birçok sektörde olduğu gibi turizm sektöründe de Veri Madenciliği uygulamalarının kullanılabilceğini ortaya koymuştur.

KAYNAKÇA

- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied System Analysis*, 16, 3-9.
- Akgöbek, Ö. ve Çakır, F. (2009). Veri Madenciliğinde Uzman Bir Sistem Tasarımı. *Akademik Bilişim Konferansları*. İnternet Teknolojileri Derneği, Şanlıurfa.
- Akgöbek, Ö. ve Kaya, S. (2011). Veri Madenciliği Teknikleri ile Veri Kümlelerinden Bilgi Keşfi: Medikal Veri Madenciliği Uygulaması. *e-Journal of New World Sciences Academy*, 6(1), 237-245.
- Akgün, A. ve Çizel, B. (2017). Günlük Tur Programları Oluşturmada Veri Madenciliği: A Grubu Seyahat Acentesi Örneği. *Turizm & Araştırma Dergisi*. 6(1): 73-87.
- Aktan, C. C. ve Vural, İ. Y. (2005). *Bilgi Çağı, Bilgi Yönetimi ve Bilgi Sistemleri*. Konya: Çizgi Kitabevi.

- Alaei, A. R., Becken, S., and Stantic, B. (2019). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research*, 58(2), 175–191.
- Alpaydın, E. (2000). Zeki Veri Madenciliği: Ham Bilgiden Altın Bilgiye Ulaşma Yöntemleri. *Bilişim 2000 Eğitimi Semineri*. Altun, M., Kayıkçı, K., Irmak S. (2019). Sınıf Öğretmenliği Öğrencilerinin Mezuniyet Notlarının Regresyon Analizi ve Yapay Sinir Ağları Yöntemleriyle Tahmini, *E-Uluslararası Eğitim Araştırmaları Dergisi*, 10(3) 2019, 29-43.
- Amjad, R. and Croock, M. S. (2019). Dominated Destinations of Tourist Inside İraq Using Personal Information and Frequency of Travel. *Telkonnika*, 17(4), 1723–1730.
- Argüden, Y. ve Erşahin, B. (2008). *Veri Madenciliği Veriden Bilgiye Masraftan Değere*. ARGE Danışmanlık. İstanbul.
- Aydın, S. (2007). *Veri Madenciliği ve Anadolu Üniversitesi Uzaktan Eğitim Sisteminde Bir Uygulama*. Eskişehir Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.
- Bigné, E., Oltra, E., and Andreu, L. (2019) Harnessing Stakeholder Input on Twitter: A Case Study of Short Breaks in Spanish Tourist Cities. *Tourism Management*, 71, 490-503.
- Bin, C., Gu, T., Sun, Y., and Chang, L. (2019) A Personalized POI Route Recommendation System Based on Heterogeneous Tourism Data and Sequential Pattern Mining. *Multimedia Tools and Applications*, 78 (24), 35135-35156.
- Cankurt, S. and Subaşı, A. (2016). Tourism Demand Modelling and Forecasting Using Data Mining Techniques in Multivariate Time Series: A Case Study in Turkey. *Turkish Journal of Electrical Engineering & Computer Sciences*, 24(5), 3388–3404.
- Cavagnaro, E., Staffieri, S. and Postma, A. (2018). Understanding Millennials' Tourism Experience: Values and Meaning to Travel as a Key for Identifying Target Clusters for Youth (Sustainable) Tourism, *Journal of Tourism Futures*, 4(1), 31-42.
- Chang, K.-C., Chen, M.-C., Kuo, N.-T., Hsu, C.-L., and Cheng, Y.-S. (2016). Applying Data Mining Methods to Tourist Loyalty Intentions in The International Tourist Hotel Sector. *Anatolia. An International Journal of Tourism & Hospitality Research*, 27(2), 271–274.
- Chua, A., Servillo, L., Marcheggiani, E., and Moere, A.V. (2016). Mapping Cilento: Using Geotagged Social Media Data to Characterize Tourist Flows in Southern Italy. *Tourism Management*, 57, 295-310.
- Chen H., Dai Z., Tang J., Li M. (2018). Web Evaluation Analysis of Tourism Destinations Based on Data Mining. *IEEE 4th ICC*. Chengdu, China.
- Çuhadar, M. (2013). *Turizmde Veri Madenciliği Alanında Yapılan Akademik Çalışmaların İncelenmesi: Türkiye ve Dünya Karşılaştırması*. 14. Ulusal Turizm Kongresi Bildiriler Kitabı. Kayseri.
- Çüllüoğlu Gülen, Ö. (2014) *Veri Madenciliği Teknikleri ile Üstün Yetenekli Öğrencilerin İlgili Alanlarının Analizi*. Gazi Üniversitesi Bilişim Enstitüsü, Ankara.
- Dalkılıç, G. ve Türkmen, F. (2002). Karınca Kolonisi Optimizasyonu. *YPBS2002–Yüksek Performanslı Bilişim Sempozyumu*, Kocaeli.
- Doğan, O. (2017). Türkiye’de Veri Madenciliği Konusunda Yapılan Lisansüstü Tezler Üzerine Bir Araştırma. *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19 (3), 929-951.
- Dunham, M. H. (2003). *Data Mining Introductory and Advanced Topics*. Pearson Education. NJ.
- Fazzolari M., and Petrocchi M. (2018). A Study on Online Travel Reviews Through Intelligent Data Analysis. *Information Technology and Tourism*, 20 (1-4), 37-58.

- Hu F., Li Z., Yang C., and Jiang Y. (2019). A Graph-Based Approach to Detecting Tourist Movement Patterns Using Social Media Data, *Cartography and Geographic Information Science*, 46:4, 368-382.
- Fonseka, P., Kaluarachchi, S.J, Jayaneththi D., and Thennakoon D. (2017). Tripper Tourism Management System Using Data Mining. *International Journal of Scientific and Research Publications*, 7(7), 374-378.
- Frاند, J. (2016). *Data Mining: What is The Data Mining?* <http://www.anderson.ucla.edu/faculty/jason.frاند/teacher/technologies/palace/datamining.htm> [Erişim Tarihi: 26.12.2016].
- Ghalehkhondabi, I., Ardjmand, E., Young, W.A., and Weckman, G.R. (2019), A Review of Demand Forecasting Models and Methodological Developments Within Tourism and Passenger Transportation Industry, *Journal of Tourism Futures*, 5(1), 75-93.
- Gu, Z., Zhang, Y., Chen, Y., and Chang, X. (2016). Analysis of Attraction Features of Tourism Destinations in A Mega-City Based on Check-in Data Mining—A Case Study of Shenzhen. *ISPRS Int. J. Geo-Inf.* 5, 11-210.
- Zhang H. and Zhaojun S. (2019) Intelligent Travelling Visitor Estimation Model with Big Data Mining. *Enterprise Information Systems*. 1-14.
- Höpken, W., Fuchs, M., Keil, D., and Lexhagen, M. (2015) Business Intelligence for Cross-Process Knowledge Extraction at Tourism Destinations. *Information Technology and Tourism*, 15 (2), 101-130.
- Huang, C., Wang, Q., Yang, D., and Xu, F. (2018). Topic Mining of Tourist Attractions Based on A Seasonal Context Aware LDA Model. *Intelligent Data Analysis*, 22(2), 383–405.
- IDC. (2016). *IDC Global Research*. <http://www.idc.com> [Erişim Tarihi: 26.12.2016].
- Karaöz Akın, B., Gürsoy, U. T., (2018). Sosyal Medya Analitiği ile Değer Yaratma: Duygu Analizi ile Geleceğe Yönelim, *Journal of Mehmet Akif Ersoy University Economics And Administrative Sciences Faculty*, 5, 797-811.
- Karampela, S., Papapanos, G., and Kizos, T. (2019). Perceptions of Agritourism And Cooperation: Comparisons Between an Island and A Mountain Region in Greece. *Sustainability (Switzerland)*, 11 (3), 680-698.
- Kozak, M., Rita, P. and Bigné, E. (2018), New Frontiers In Tourism: Destinations, Resources, and Managerial Perspectives. *European Journal of Management and Business Economics*, 27 (1), 2-5.
- Kirilenko, A. P., and Stepchenkova, S. (2018). Tourism Research From its Inception to Present Day: Subject Area, Geography, and Gender Distributions. *PLoS ONE*, 13(11), 1–20.
- Lenormand, M., Luque, S., Langemeyer, J., Tenerelli, P., Zulian, G., Aalders, I., Chivulescu, S., Clemente, P., Dick, J., van Dijk, J., van Eupen, M., Giuca, R. C., Kopperoinen, L., Lellei-Kovács, E., Leone, M., Lieskovský, J., Schirpke, U., Smith, A. C., Tappeiner, U., and Woods, H. (2018). Multiscale Socio-Ecological Networks in The Age of Information. *PLoS ONE*, 13(11), 1–16.
- Li, G., Law, R., Vu, H.Q., Rong, J. (2015) Discovering the Hotel Selection Preferences of Hong Kong Inbound Travelers Using the Choquet. *Integral Tourism Management*, 36, 321-330.
- Li, Q., Li, S., Zhang, S., Hu, J., and Hu, J. (2019). A Review of Text Corpus-Based Tourism Big Data Mining. *Applied Sciences (Switzerland)*, 9 (16), 3300.
- Li, X., and Song, H. (2016). Intelligent Planning of Tourism Scenic Routes Based on Genetic Algorithm in Coal WANBEI Mining Subsidence. *Telkomnika*, 14(2A), 69–76.

Lin, W., Yuming, H., Zhou H., Xia, P. and Ran, L. (2018) A Hybrid Ensemble Learning Method for Tourist Route Recommendations Based on Geo-Tagged Social Networks. *International Journal of Geographical Information Science*, 32:11, 2225-2246.

Lucas, J. P., Luz, N., Moreno, M. N., Anacleto, R., Almeida Figueiredo, A., and Martins, C. (2013). A Hybrid Recommendation Approach for A Tourism System. *Expert Systems with Applications*, 40(9), 3532-3550.

Li, Y., Xiao, L., Ye, Y., Xu, W., and Law, A. (2016) Understanding Tourist Space at A Historic Site Through Space Syntax Analysis: The case of Gulangyu, China. *Tourism Management*, 52, 30-43.

Mastorakis, G., Trihas, N., Perakakis, E., and Kopanakis, I. (2015). E-CRM in Tourism Exploiting Emerging Information and Communication Technologies. *Anatolia. An International Journal of Tourism & Hospitality Research*, 26(1), 32-44.

Mirjana, P. B., Schatten, M. and Zrinka, M. (2013). Data Mining Applications in Tourism: A Keyword Analysis. *Conference: Central European Conference on Information and Intelligent Systems*. Varaždin, Croatia.

Oracle. (2020). *Data Mining Concepts*. https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm#CHDFGCIJ [Erişim Tarihi: 23.04.2020].

Özcan, C. (2014). *Veri Madenciliğinin Güvenlik Uygulama Alanları ve Veri Madenciliği ile Sahtekarlık Analizi*. Yüksek Lisans Tezi. İstanbul Bilgi Üniversitesi, İstanbul.

Özen, Ü. (2014). *Bilgi Sistemlerine Giriş: Temel Kavramlar*. Atatürk Üniversitesi AOF Yayınevi. Erzurum.

Padmaja, N., and Sudha, T. (2019). Predicting Travel Behaviour of International and Domestic Tourists Using Big Data. *International Journal of Recent Technology and Engineering*, 8 (3), 1572-1580

Pineda-Bravo, F., Pérez-Alfonso, D., and Pérez-García, W. (2016). Aplicación De La Minería De Procesos En El Hotel Club Amigo Atlántico Guardalavaca. *Retos Turísticos*, 15(3), 12-20.

Pineda-Bravo, F., Pérez-García, W., Cristóbal Martínez-Martínez, C. (2018). Procedimiento Para El Diagnóstico De Procesos Turístico, Con Integración De Técnicas De Minería De Procesos. *Retos Turísticos*, 17(3), 32-40.

Richard, B. (2017), Hotel Chains: Survival Strategies for A Dynamic Future, *Journal of Tourism Futures*, 3(1), 56-65.

Sağsan, M. (2014). *Bilgi Yönetimi Bakış Açısından Bilgi Süreçleri ve Bir Model Önerisi*. Dokuz Eylül Üniversitesi Strateji Dairesi Başkanlığı: <http://strateji.deu.edu.tr/> [Erişim Tarihi: 12.25.2016].

Savaş, S., Topaloğlu, N., ve Yılmaz, M. (2011). Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, (21), 1-23.

SDN. (2016). Shift Delete Net. <http://shiftdelete.net/dunyada-ne-kadar-bilgi-var-27154> [Erişim Tarihi: 31.12.2016].

Sever, H., Oğuz, B. (2002). Veri Tabanlarında Bilgi Keşfine Formel Bir Yaklaşım Kısım 1: Eşleştirme Sorguları ve Algoritmalar. *Bilgi Dünyası*, 3(2), 1-33.

Sevimli Deniz, S. (2019) Veri Madenciliği Araçları Kullanılarak Türkiye'nin Turizm Gelirlerinin Aylara Göre Yapay Sınır Ağları ile Tahminlenmesi. *Van Yüziüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 19, 227- 233.

- Song, S., Kawamura, H., Uchida, J., and Saito, H. (2019). Determining Tourist Satisfaction from Travel Reviews. *Information Technology and Tourism*, 21 (3), 337-367.
- Su, L., Stepchenkova, S., and Kirilenko, A.P. (2019). Online Public Response to A Service Failure Incident: Implications for Crisis Communications. *Tourism Management*, 73, 1-12.
- Şeker, Ş. E. (2014). MIS Sözlük. <http://mis.sadievrenseker.com/2014/06/buyuk-veri-big-data/> [Erişim Tarihi: 04.13.2020].
- Qi, S., Wong, C.U.I., Chen, N., Rong, J., and Du, J. (2018). Profiling Macaultural Tourists by Using User-Generated Content from Online Social Media. *Information Technology and Tourism*, 20 (1-4), 217-236.
- Tao, Y., Zhang, F., Shi, C., and Chen, Y. (2019). Social Media Data-Based Sentiment Analysis of Tourists' Air Quality Perceptions. *Sustainability (Switzerland)*, 11 (18), 5070.
- TDK. (2020). Türk Dil Kurumu sözlükler: <https://tdk.gov.tr>. [Erişim Tarihi: 04.16.2020].
- Terzi, Ö. (2012). Monthly Rainfall Estimation Using Data-Mining Process. *Applied Computational Intelligence and Soft Computing*, 2012, 1-6.
- Tkach, V., Pavlenchyk, A., Sadchenko, O., Nikola, S., Drozdova, V., and Davydenko, I. (2019) Modelling Buying Demand In The Tourism Industry Based on Machine Training Methods. *International Journal of Recent Technology and Engineering*, 8 (2), 744-747.
- Versichele, M., de Groote, L., Claeys Bouuaert, M., Neutens, T., Moerman, I., and Van de Weghe, N. (2015). Pattern Mining in Tourist Attraction Visits Through Association Rule Learning on Bluetooth Tracking Data: A Case Study of Ghent, *Tourism Management*, 44, 67-81.
- Vu, H.Q., Li, G., Law, R., and Ye, B.H. (2015) Exploring the Travel Behaviors of Inbound Tourists to Hong Kong Using Geotagged Photos. *Tourism Management*, 46, 222-232.
- Wan, L., Hong, Y., Huang, Z., Peng, X., and Li, R. (2018). A Hybrid Ensemble Learning Method for Tourist Route Recommendations Based on Geo-Tagged Social Networks. *International Journal of Geographical Information Science*, 32(11), 2225–2246.
- Wang, Z. and Liu, B. (2019). Tourism Recommendation System Based on Data Mining. *Journal of Physics: Conference Series*. 1345. 022027. (https://www.researchgate.net/publication/337598685_Tourism_recommendation_system_based_on_data_mining/citation/download).
- Yılmaz, Y. (2011). Blogcu. <http://yyilmaz091.blogcu.com/bilgi-turleri-veri-enformasyon-bilgi-bilgelik-kavramlari/10342406> [Erişim Tarihi: 31.12.2016].
- Zeng, C., Nakatoh, T., Hirokawa, S., and Eguchi, M. (2019). Text Mining of Tourism Preference in A Multilingual Site. *IEEE Transactions on Electrical and Electronic Engineering*, 14(4), 590-596
- Zhang, K., Chen, Y., and Li, C. (2019). Discovering the Tourists' Behaviors and Perceptions in A Tourism Destination by Analyzing Photos' Visual Content with A Computer Deep Learning Model: The Case of Beijing. *Tourism Management*, 75, 595-608.
- Zhang, W. (2019). Application of an Improved Ant Colony Algorithm in Coastal Tourism Route Optimization. *Journal of Coastal Research*, 98, 84–87.
- Zhang, C., and Huang, Z. (2015). Mining Tourist Motive for Marketing Development via Twice-Learning. *Applied Artificial Intelligence*, 29(2), 119–133.

Zhou, X., Zhan, Y., Feng, G., Zhang, D., and Li, S. (2019). Individualized Tour Route Plan Algorithm Based on Tourist Sight Spatial Interest Field. *ISPRS International Journal of Geo-Information*, 8 (4), 192.

Zhu, X., Sun, T., Yuan, H., Hu, Z., and Miao, J. (2019) Exploring Group Movement Pattern Through Cellular Data: A Case Study of Tourists In Hainan. *ISPRS International Journal of Geo-Information*, 8 (2),
https://www.researchgate.net/publication/330890029_Exploring_Group_Movement_Pattern_through_Cellular_Data_A_Case_Study_of_Tourists_in_Hainan/citation/download