



Web Kullanıcılarının Bilgi Erişim ve Ziyaret Desenlerinin Web Madenciliği ile Keşfi: Kırklareli Üniversitesi Örneği*

Veli Özcan BUDAK, Kırklareli Üniversitesi, Bilgi İşlem Daire Başkanlığı, Öğretim Görevlisi, veliozcanbudak@gmail.com,  0000-0002-0960-0542

Çiğdem Selçukcan EROL, İstanbul Üniversitesi, Enformatik Bölümü, Doçent, cigdemsel@gmail.com,  0000-0002-5057-7145

ÖZ

Web siteleri, kurumsal ya da bireysel açıdan hitap edilen kitleyle ilk temasın sağlandığı bir etkileşim aracıdır. Bu araç, yoğun bir bilgi erişim ve ziyaret trafiğinin bulunduğu süreçlerde, kullanıcı davranışlarındaki farklı desenlerin tespit edilebileceği önemli bir potansiyeli içinde barındırmaktadır. Bu desenler, kullanıcı ihtiyaçlarının daha belirginleştirilmesi ve site geliştiricilerinin bu ihtiyaçlar doğrultusunda güncellemeler yapabilmesi açısından oldukça kritik görevler üstlenebilir.

Bu çalışmanın amacı, dünya genelinde yaşanan Covid-19 pandemisinin ülkemizde etkinliğini arttırdığı süreçte, Kırklareli Üniversitesi web sitelerindeki kullanıcıların bilgi erişim ihtiyaçlarındaki değişimin belirlenmesidir. Bu amaç doğrultusunda, kullanıcıların bilgi erişim ve ziyaret davranışları, apriori algoritmasıyla bağımsız ve birlikte olacak şekilde incelenerek, aralarındaki ilişkilerin ortaya çıkarılması hedeflenmiştir. Bilgi erişim kavramı açısından çalışma sonuçları, kullanıcıların “tez yazımı”na yönelik çeşitli arama terimleriyle bilgi ihtiyaçlarını karşılamaya çalıştıklarını göstermiştir. Bu sonuç, özellikle lisansüstü öğrencilerin ilgili süreçte aktif olduklarına işaret etmektedir. Ziyaret davranışları açısından, “uzaktan eğitim”, “koronavirüs” ve “tatil” temalı sayfaların ağırlıklı olarak ziyaret edildiği ortaya çıkmıştır. Bilgi erişim davranışları sonrasında sergilenen ziyaret davranışları açısından, “tez yazımı”, “tatil” ve “eğitim öğretimin ertelenmesi” temalı ziyaretlerin birliktelikleri göze çarpmıştır. Çalışma sonucunda ortaya çıkarılmış olan davranış desenleri ve bu desenlerden nasıl faydalanılabileceğine yönelik öneriler çalışma kapsamında detaylı bir şekilde açıklanmıştır.

Anahtar Kelimeler : **Bilgi Erişim, Apriori, Birliktelik Kuralları, Web Kullanım Madenciliği**

* Bu çalışma, İstanbul Üniversitesi’nde yürütülmekte olan “Bilgiye Erişimde Kullanılabilirliğin Yeni Bir Yapay Zekâ Yöntemiyle Geliştirilmesi” başlıklı doktora tezi kapsamında üretilmiştir.

Discovering Web Users' Information Retrieval and Visit Patterns with Web Mining: Kırklareli University Sample*

ABSTRACT

Websites are an interaction tool in which the first contact is made with target users, institutionally, or individually. This tool has an important potential in which different patterns of user behaviors are able to be detected in the process of excessive information retrieval and visit traffic. These patterns might play an important role, in terms of making the user needs more clear and to enable site developers to make updates in line with these needs.

The purpose of this study is to determine the differentiation of users' information retrieval needs on the websites of Kırklareli University, in the period of the Covid-19 pandemic that experienced worldwide increases its effectiveness in our country. In the direction of this purpose, it is aimed to reveal the relationships between the information retrieval behavior of the users and the visiting behavior of them by analyzing them with the apriori algorithm, both independently and together. In terms of the concept of information retrieval, the results have shown that the users tried to meet their information needs with various searching terms for "thesis writing". This result indicates that the postgraduate students were active in the related period. In terms of the visiting behavior, it has been revealed that the pages with the theme of "distance education", "coronavirus" and "holiday" were mostly visited ones. In terms of the visiting behaviors shown after the information retrieval behaviors, the coexistence of the visits with the theme of "thesis writing", "holiday" and "postponing the education" has been observed. The behavior patterns which have been revealed as a result of the study and the suggestions about how to use these patterns are explained in detail within the scope of the study.

Keywords : **Information Retrieval, Apriori, Association Rules, Web Usage Mining**

1. GİRİŞ

Web, farklı bilgi ihtiyacı bulunan bireylerin yararlanabileceği çok çeşitli türde uygulamaları içinde barındıran günümüzün önemli dijital platformlarından birisidir. Bu platform üzerinde bilgi ihtiyacının giderilmesine yönelik olarak hizmet veren uygulamalardan birisi web siteleridir. Web siteleri, hitap ettiği kitleye ve sunduğu bilgiye bağlı olarak oldukça fazla kullanım trafiğine sahip olabilmektedir. Bu trafik, site geliştiricileri açısından gözle görülemeyecek çeşitli kullanıcı ihtiyaçlarının belirlenmesi noktasında önemli görevler üstlenebilir. Yüksek kullanım trafiğine sahip sitelerin bile zamanla popülerliğini yitirebilme olasılığı her zaman bulunmaktadır. Bu riskin giderilebilmesi/indirgenebilmesi amacıyla web

* This study has been produced from the doctoral thesis titled "Improving Usability in Access to Information by a New Artificial Intelligence Method" at İstanbul University.

kullanıcılarına ait davranışların incelenerek, ihtiyaçlara hızlı cevap verebilecek bir ortamın sağlanmasının önemli olduğu düşünülmektedir.

Web üzerindeki kullanım trafiğine çeşitli kullanıcı etkileşimleri sebep olabilmektedir. Bunlara; ziyaretler, bilgi erişim, içerik paylaşımı ve hatta kötü niyetli saldırılar örnek gösterilebilir. Bu çalışma kapsamında temel olarak, “bilgi erişim” ve “ziyaret” olmak üzere iki farklı etkileşim türü odak alınmıştır. “Ziyaret” kavramı adından da anlaşılacağı üzere, bir web kullanıcısının ihtiyacını gidermek amacıyla bir web sitesinde yer alan ilgili sayfayı/sayfaları ziyaret etmesi şeklinde tanımlanabilir. Bilgi erişim (information retrieval), Türk Dil Kurumu (2020) tarafından “bellekte saklı verilerden belli bir konuda bilgi alma yöntem ve yordamları”; Britannica (2007) tarafından da “özellikle bir bilgisayarda yer alan veritabanından bilginin elde edilmesi” şeklinde ifade edilmiştir. Bir başka ifadeyle, “kullanıcıların ilgi duydukları bilgiye hızlı erişimlerine odaklanmış bir alan” olarak tanımlanmıştır (Baeza-Yates ve Ribeiro-Neto, 2011). Bilgi erişim davranışı, aslında günlük yaşamımızda sürekli olarak gerçekleştirdiğimiz bir eylemdir. Bu davranışa verilebilecek en güzel örnek, arama motorları üzerinde gerçekleştirilen aramalardır. Bir bilgiye ihtiyaç duyan kişi, kullandığı arama motoru arayüzünde bulunan metin yazma kısmına ihtiyacının genel çerçevesini tanımlayan bir terimi girerek sorgulamalar gerçekleştirmekte ve bu sorgulamalar doğrultusunda arama motorlarının arka planında çalıştırılan çeşitli kodlar sayesinde en uygun sonuçlar sırayla kullanıcıya listelenmektedir. Bu eylem, bilgi erişim kavramının özünü oluşturmaktadır.

Web kullanıcılarının bilgi erişim ve ziyaret davranışlarının incelenmesi, web sitelerinin devamlılığı açısından kritik noktalara ışık tutabilir. Bu sebeple, kullanıcı davranışı odak alınarak yapılacak çalışmaların, site geliştiricilerine önemli avantajlar getireceği söylenebilir. Bu çalışma kapsamında, kullanıcı davranışlarının analiz edilmesi amacıyla veri madenciliğinin bir alt alanı ya da web’e odaklanmış alanı olarak ifade edilen web madenciliğinden faydalanılmıştır. Veri madenciliği, “veri toplama, temizleme, işleme, analiz etme ve verilerden yararlı bilgiler edinme çalışması” şeklinde tanımlanmıştır (Aggarwal, 2015). Web madenciliği ise, web köprü yapısı (hyperlink), sayfa içeriği ve kullanım verilerinden faydalı bilgiler veya bilgiler bulmayı amaçlayan (Liu, 2011) bir veri madenciliği alt alanıdır. “Web’deki değerli bilginin otomatikleştirilmiş keşfi” olarak da tanımlanmıştır (Chakrabarti, 2002). Web madenciliği, “web yapı madenciliği”, “web içerik madenciliği” ve “web kullanım madenciliği” olmak üzere üç farklı kategoriye sahiptir (Gezer, Erol ve Gülseçen, 2007; Markov ve Larose, 2007; Liu, 2011; Grace, Maheswari ve Dhinakaran, 2011). Web yapı madenciliği, arama motorları mimarisinde anahtar role sahip olan ve bağlantılar üzerinden önemli web sayfalarının tespit edilmesinde kullanılan bir web madenciliği çeşitidir (Liu, 2011). Web içerik madenciliği, web sayfalarından kullanışlı bilginin ortaya çıkarılmasında kullanılmaktadır (Liu, 2011). Web kullanım madenciliği ise, web günlük dosyalarından kullanıcı erişim desenlerinin tespit edilmesi amacıyla kullanılan (Liu, 2011) ve bu çalışmada da faydalanılmış olan web madenciliği çeşitidir. Kullanıcıların web sayfalarında

gerçekleştirdikleri ziyaretler günlük dosyalarında önemli verinin depolanmasına sebep olmaktadır (Chauhan ve Tarar, 2016). Dolayısıyla web kullanım madenciliği sayesinde, web kullanıcılarının ne tür bilgi elde edinme davranışında buldukları keşfedilebilir. Bu sayede, gözle görülemeyen önemli bilginin ortaya çıkarılarak, web sitelerinin geliştirilmesi ve kullanıcı ihtiyaçlarına hitap edecek bir yapıda hizmet sunulması sağlanabilir.

Bu açıklamalar doğrultusunda, web kullanım madenciliği yapılarak Kırklareli Üniversitesi (KLU) web sitelerindeki kullanıcı bilgi erişim ve ziyaret davranışlarının, apriori algoritmasıyla incelenmesi bu çalışmanın amacını oluşturmuştur. Bu amaç için, dünya genelinde yaşanan Covid-19 pandemisinin ülkemizde etkinliğini arttırdığı süreç temel alınarak 2020 yılının Mart ayına odaklanılmıştır. Bu sürecin seçilmesinin temel sebebi, ülkemizde örgün eğitime ara verilme kararı sonrasında KLU web kullanıcılarının bilgi edinme ihtiyaçlarının daha belirgin olacağı varsayımdır. Öyle ki, çalışmanın 3. bölümünde de aktarılacağı üzere, bu varsayımın istatistikî açıdan gerçekleştiği tespit edilmiştir. Diğer taraftan, çalışmada belirlenen amaç kapsamında, KLU web kullanıcılarının bilgi erişim ve ziyaret davranışları bağımsız ve birlikte olacak şekilde incelenerek, çeşitli açılardan davranış desenleri arasındaki ilişkilerin ortaya çıkarılması hedeflenmiştir. Belirlenen hedefler genel olarak aşağıdaki şekilde üç kategoriye ayrılmıştır:

- Kullanıcıların “bilgi erişim” açısından yaptığı aramalardaki birlikteliklerin keşfedilmesi
- Kullanıcıların “ziyaretler” açısından sergilediği davranışlardaki birlikteliklerin keşfedilmesi
- Kullanıcıların “bilgi erişim” doğrultusunda sergilediği “ziyaret” davranışlarındaki birlikteliklerin keşfedilmesi

Çalışmanın 2. bölümünde literatürde gerçekleştirilmiş olan benzer çalışmalara, 3. bölümünde çalışmada kullanılan yöntem, 4. bölümde yapılan analizler doğrultusunda ortaya çıkarılan davranış desenlerine ait bulgulara, 5. bölümde ise çalışma bulgularının değerlendirilmesi doğrultusunda, site geliştiricilere yönelik olarak ortaya çıkarılmış çeşitli önerilere yer verilmiştir.

2. LİTERATÜR TARAMASI

Bu bölümde, literatürde yapılmış benzer çalışmalar aktarılmıştır.

Tam bağlantı (Leaders Complete Linkage: LCL) kümeleme algoritmasıyla Apriori algoritmasının birlikte kullanıldığı bir çalışma, Ilampiray (2012) tarafından bir üniversiteden sağlanan web günlük dosyaları ile gerçekleştirilmiştir. Kümeleme algoritmasıyla gruplanmış kullanıcıların Apriori algoritmasıyla birlikteliklerinin incelendiği bu çalışmada, yazarlar,

önerdikleri yaklaşım sayesinde eğitim kurumlarında ağın verimli kullanılmasının sağlanabileceğini belirtmişlerdir.

Bir öğrenme yönetim sistemi (Learning Management System: LMS) üzerinden elde edilmiş web günlük dosyası verisinin Apriori ve Eclat algoritmalarıyla incelendiği bir diğer çalışma Kotiyal ve diğ. (2013) tarafından yapılmıştır. Her iki algoritmanın performans olarak karşılaştırılmasının da yapıldığı çalışmada, ilişkili web sayfalarına daha verimli bir şekilde erişim için, kişiselleştirilmiş bir web hizmetinin sağlanabilmesine yönelik bir teknik üzerine odaklanılmıştır. Çalışma sonuçlarında, geçmiş ziyaret davranışları kullanılarak web kullanıcılarının gezinme davranışlarının tespit edilebileceği belirtilmiştir.

Çınar ve Bilge (2016), web günlük dosyalarından faydalanarak, saldırı girişim ve türlerinin tespitine yönelik bir web madenciliği çalışması gerçekleştirmiştir. Apriori ve K-means algoritmalarından faydalanılmış olan bu çalışmada, yazarlar, ortaya çıkardıkları desenler yardımıyla saldırı girişimlerini %88.7 oranında azalttıklarını ifade etmişlerdir.

Çelik (2017), bir e-ticaret firmasına ait web günlük dosyaları üzerinden gerçekleştirdiği çalışmada, tüketicilerin satın alma davranışlarının eğilimini belirlemeye çalışmıştır. Destek vektör makineleri ile lojistik regresyon karşılaştırmasının yapıldığı bu çalışma sonuçlarında, sayfa/ürün görüntüleme sayılarıyla, ürün satışları arasında büyük farklar olmadığı ve destek vektör makinelerinin sınıflandırmada daha başarılı olduğu ifade edilmiştir.

Bir üniversitenin web günlük dosyasından faydalanılarak yapılmış olan bir başka çalışmada, kurumsal ve kurumsal-dışı açılardan kritik web trafiği Apriori algoritması kullanılarak tespit edilmeye çalışılmıştır (Kindie, Mamuye ve Tilahun, 2018). Akademik çalışanların eğitim sitelerine, yönetici konumundaki çalışanların sosyal medya ve eğlence temalı sitelere erişim sağladıklarının belirlendiği bu çalışmada, sistem performansına ve bant genişliği yönetimine yönelik olarak yararlanılabilecek kullanım desenleri ortaya çıkarılmıştır.

Apriori algoritması kullanılarak gerçekleştirilmiş bir web kullanım madenciliği çalışması Sathya ve Devi (2018) tarafından gerçekleştirilmiştir. İlgili çalışmada, web günlük dosyaları üzerinden kullanıcıların ziyaret ettikleri sayfalarda sıklıkla tekrarlanmış desenlerin tespiti yapılmış olup, sistem yöneticilerine yönelik olarak öneride bulunulmuştur.

Budak, Kartal ve Gülseçen (2018) tarafından, web günlük dosyaları kullanılarak yapılmış olan bir diğer çalışmada, Apriori algoritmasıyla web kullanıcılarının site-içi arama araçlarında kullandıkları kelime/kelime gruplarının birliktelikleri tespit edilmeye çalışılmıştır. Çalışma sonuçlarında ortaya çıkarılan desenler üzerinden web sitelerinin daha iyi hizmet sunmasına yönelik dönemsel ve sürekli olacak şekilde öneriler getirilmiştir. Buna ek olarak, periyodik olarak kullanıcı davranışlarının takip edilmesinin önemine vurgu yapılmıştır.

Bu çalışmanın kapsamı web kullanım madenciliği temeline oturtulmasına rağmen, kullanıcı davranış desenleri üzerinden çeşitli eksikliklerin tespit edilmesi ya da davranış benzerliklerinin belirlenmesi açısından düşünüldüğünde, aşağıda aktarılan çalışmaların da bu çalışmanın arka planıyla ortak noktalar taşıdığı söylenebilir.

Sağın ve Ayvaz (2018), perakende sektöründe hizmet veren bir donanım firmasına ait veri seti üzerinde Apriori ve FP-Growth algoritmalarını kullanarak, birlikte satılan ürün gruplarını belirlemeye çalışmıştır. Çalışma sonuçlarında, firmanın, ortaya çıkarılmış olan satış desenlerini ilgili dönemde fark edemediği belirtilmiş ve bu desenlerin firma satışlarını artırma olasılığından bahsedilmiştir.

Web günlük dosyaları yerine 214 kişiden anket yardımıyla toplanan veri setinin kullanıldığı bir çalışma Şimşek Gürsoy, Akçay Kasapoğlu ve Atalay (2019) tarafından gerçekleştirilmiştir. E-ticaret alışkanlıkları üzerinden kişi profillerinin ve e-ticaret sitelerinin kullanım birlikteliklerinin tespit edilmeye çalışıldığı bu çalışmada, Apriori ve Eclat algoritmalarından faydalanılmıştır. Çalışma sonuçlarında, e-ticaret sitelerinin pazarlama stratejilerine yönelik olarak çeşitli öneriler sunulmuştur.

Bilgiç (2019), bir süpermarkete ait veri seti üzerinden Apriori algoritmasını kullanarak, tüketici tercih ve satın alma davranışlarını tespit etmeye çalışmıştır. Çalışma sonuçlarında, süpermarketin satış yaptığı ürünler arasındaki birliktelikler tespit edilmiştir. Buna ek olarak, firmaların, ürünlerini daha ayrıntılı kodladıkları takdirde, çalışmada izlenen analiz yöntemi ile daha detaylı tüketici davranış desenlerine ulaşabilecekleri belirtilmiştir.

Literatürde yapılmış olan bu çalışmalar incelendiğinde, kullanılan veri setleri farklı olmasına rağmen, ortak noktaları, kullanıcı ya da tüketici davranış desenlerinin tespit edilmeye çalışılmasıdır. Çınar ve Bilge (2016), yaptıkları çalışmada saldırı tespitine yönelik bir araştırma gerçekleştirmelerine rağmen, bu saldırıların engellenerek hedef kullanıcılara daha iyi bir hizmetin sağlanması, dolaylı bir hedef şeklinde düşünülebilir. Genel olarak özetlemek gerekirse, kullanıcı ihtiyaçlarının tespit edilmesi doğrultusunda yapılacak geliştirmelerin daha olumlu bir hizmet ortamının sağlanması açısından önemli olduğu düşünülmektedir. Bu çalışma da, aynı doğrultuda ilerleyen bir çalışma özelliği taşımaktadır.

3. YÖNTEM

Bu çalışma, Fayyad, Piatetsky-Shapiro ve Smyth (1996) tarafından veritabanlarında bilgi keşfi (VBK) süreci olarak ifade edilmiş adımlar takip edilerek gerçekleştirilmiştir. VBK temel olarak “Veri Seçimi”, “Ön İşleme”, “Dönüştürme”, “Veri Madenciliği” ve “Uygulama/Değerlendirme” olmak üzere beş aşamadan oluşmaktadır (Fayyad, Piatetsky-Shapiro ve Smyth, 1996). “Veri Seçimi”, “Ön İşleme”, “Dönüştürme” ve “Veri Madenciliği”

aşamaları bu bölümde, “Uygulama/Değerlendirme” aşaması ise çalışmanın 4. ve 5. bölümlerinde aktarılmıştır.

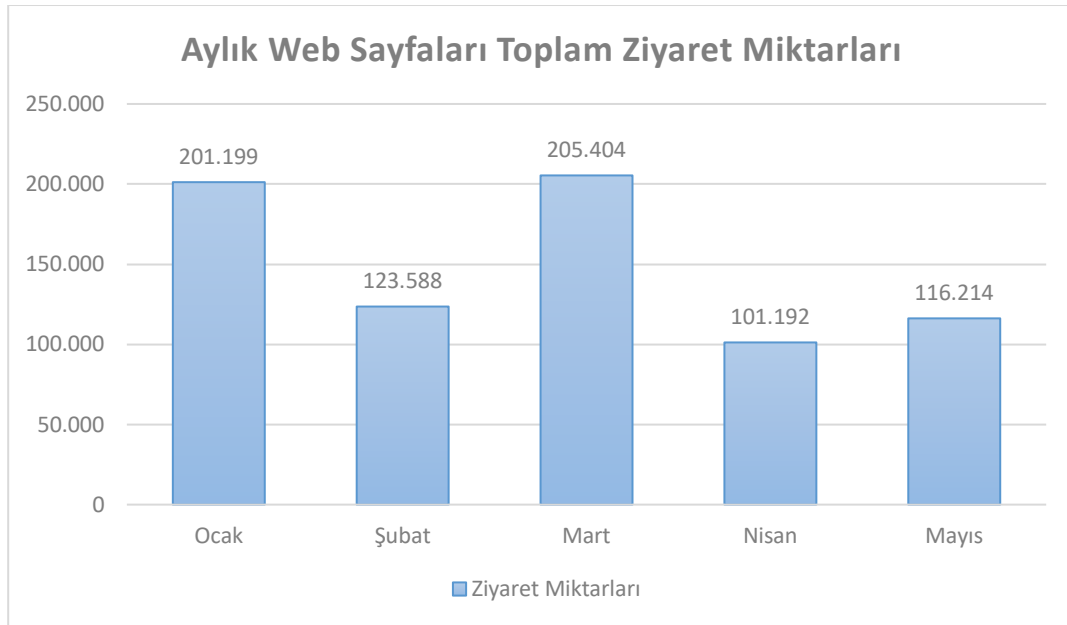
Veri Seçimi

KLU web kullanıcılarının “bilgi erişim” ve “ziyaret” olmak üzere iki farklı türde oluşturdukları web trafiğinin incelenmesi amacıyla, üniversitede yer alan tüm web siteleri üzerinden veri sağlanmaya çalışılmıştır. Web kullanım madenciliğinde çoğunlukla web günlük dosyaları tercih edilmesine rağmen, tıpkı bir günlük dosyası gibi kullanılması amacıyla bir veritabanı (MySQL) oluşturulmuş ve kullanıcı etkileşimleri bu veritabanı üzerinde kaydedilmiştir (veri depolama işlemleri PHP yazılım dili kullanılarak gerçekleştirilmiştir). Bu çalışma için veri depolama süreci 2020 Ocak-Mayıs ayları arasındaki zaman dilimini kapsamıştır. Bu süreçte veri depolanmasının nasıl gerçekleştirildiğine yönelik işlemler şu şekildedir:

“Bilgi erişim” kavramına yönelik olarak veri elde etmek amacıyla, üniversitede yer alan tüm web siteleri içinde bir modül olarak kullanılan site-içi arama araçlarından faydalanılmıştır. Bu işlem için, *id*, *arama_terimi*, *ip* ve *tarih* niteliklerini barındıran bir tablo, oluşturulmuş olan veritabanına dahil edilmiştir. Bu tablo sayesinde, “Kullanıcıların “bilgi erişim” açısından yaptığı aramalardaki birlikteliklerin keşfedilmesi” şeklinde belirlenen birinci hedefe yönelik olarak veri toplanmıştır. “Ziyaret” davranışlarının kaydedilmesi amacıyla da, *id*, *bilgi_erişim_id*, *sayfa_id*, *ip* ve *tarih* niteliklerinden oluşan bir tablo veritabanına dahil edilmiş olup, kullanıcıların, tüm web sitelerindeki web sayfası ziyareti davranışları bu tabloda depolanmıştır. 2. tablo sayesinde, çalışmada belirlenmiş olan diğer iki hedefe yönelik olarak veri elde edilmiştir. Daha detaylı ifade edilecek olunursa, *bilgi_erişim_id* haricinde kalan nitelikler üzerinden elde edilmiş olan veri kullanılarak, “Kullanıcıların “ziyaretler” açısından sergilediği davranışlardaki birlikteliklerin keşfedilmesi” şeklindeki ikinci hedefe yönelik veri seti oluşturulmuştur. *bilgi_erişim_id* niteliği, 3. hedefte ifade edilen “bilgi erişim doğrultusunda sergilenen ziyaret davranışları” verisinin elde edilmesi amacıyla 2. tabloya eklenmiştir. Bu durumu bir senaryo olarak ifade etmek gerekirse; bir kullanıcının herhangi bir KLU web sitesinde yer alan site-içi arama aracında arama yaptığı anda 1. tabloya bir *id* değeriyle birlikte kayıt gerçekleştirilmiştir. Kullanıcının, arama sonuçlarında önerilmiş olan herhangi bir sayfaya tıklayarak ulaşması durumundaysa, 1. tabloda oluşturulan *id* değeri, *bilgi_erişim_id* niteliğine karşılık olacak şekilde 2. tabloya kaydedilmiştir. Böylelikle, kullanıcıların bilgi erişim davranışı doğrultusunda sergilediği ziyaret davranışlarına ait veri seti oluşturulmuştur. Bu durumda, 2. tabloda *bilgi_erişim_id* değeri boş olan satırlar ikinci hedefe, dolu olanlar ise üçüncü hedefe yönelik analiz imkânı sağlamıştır.

Çalışmada analiz edilecek dönemin belirlenmesi için dünya genelinde yaşanan Covid-19 pandemisinin ülkemizde etkinliğini arttırdığı ve örgün eğitime ara verilme kararının verildiği (12 Mart 2020) döneme odaklanılmıştır. Bu döneme odaklanılmasının temel sebebi, KLU web

kullanıcılarının bilgi edinme ihtiyaçlarının daha belirgin olacağı varsayımdır. Şekil 1' de görülen veri depolama sürecindeki kullanıcı ziyaretlerine ait istatistikler bu varsayımı desteklemektedir. Kullanıcı ziyaretleri, KLU web sitelerinde yer alan farklı web sayfalarına kullanıcılar tarafından gerçekleştirilen ziyaretleri ifade etmektedir. Bu istatistiklerde, iki türlü ziyaret davranışı birlikte hesaplanmıştır. Birinci tür, kullanıcıların doğrudan (arama yapılmadan) gerçekleştirdikleri sayfa ziyaretlerini, ikincisi ise, kullanıcıların bilgi erişim davranışı (arama yapılarak) sonrasında karşlarına getirilen sayfa önerileri listesinde, herhangi bir sayfaya tıklayarak gerçekleştirdikleri sayfa ziyaretlerini ifade etmektedir. Aynı IP üzerinden farklı zamanlarda yapılan her bir ziyaret bu istatistiklerde bağımsız olarak işlenmiştir.

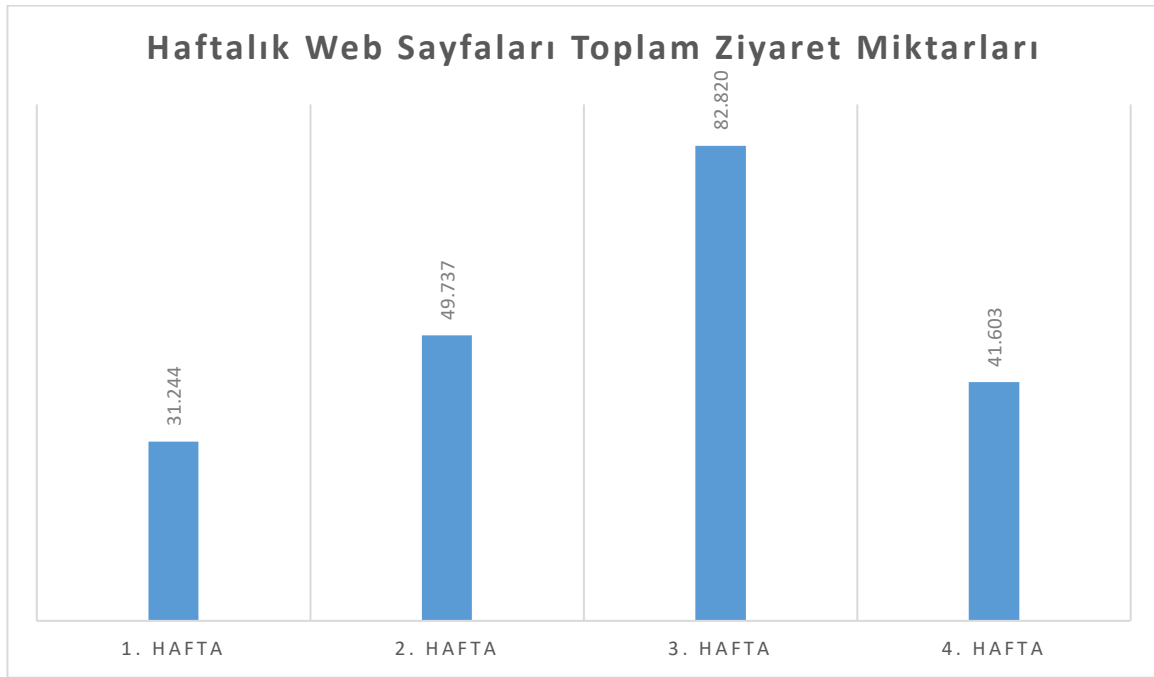


Şekil 1: 2020 Ocak-Mayıs ayları arasındaki web sayfaları toplam ziyaret miktarları.

Ön İşleme

Web günlük dosyaları üzerinden yapılan analizlerde ön işleme aşaması oldukça önemlidir. Bunun temel sebebi, ön işleme aşamasında yapılması muhtemel hataların analiz sonuçlarını olumsuz etkileyebilme olasılığı ve bu doğrultuda da sürecin tekrar başa sarılmasının söz konusu olmasıdır. Bu durum, zaman, emek ve mali açılarından olumsuz etkilere sebep olabilir. Web günlük dosyaları özelinde düşünüldüğünde, web kullanıcılarının davranışları dışında analiz için gerekmeyecek verinin de günlük dosyalarında yer alması muhtemeldir. Buna örnek olarak, js ve css gibi kullanıcı davranışından bağımsız tamamıyla kullanılan web uygulamasının arka planında çalıştırılan dosyaların da günlük dosyalarında tutuluyor olmasıdır. Bu ve benzeri gereksiz veri satırlarının arındırılmaması çalışma sonuçlarında yanlış bulgulara sebep olabilir. Dolayısıyla, ön işleme aşamasının web madenciliğinde en kritik aşamalardan birisi olduğu söylenebilir. Bu çalışmada, verinin, MySQL veritabanı içinde daha

önceden nasıl tutulacağı belirlendiğinden dolayı ön işleme aşaması web günlük dosyalarına nazaran daha hızlı ve kolay gerçekleştirilmiştir. Ön işleme aşamasında genel itibariyle “Veri Temizleme”, “Veri Bütünleştirme”, “Veri İndirgeme” ve “Veri Dönüştürme” gibi işlemler yapılırken (Han, Kamber ve Pei, 2011; Özkan, 2013), bu çalışmada sadece odaklanılan dönemdeki verinin belirli zaman aralığında olacak şekilde indirgenme işlemi gerçekleştirilmiştir. İndirgeme işleminin yapılmasındaki temel amaç ise, örgün eğitime ara verilen sürece daha fazla odaklanılmasıdır. İndirgeme sürecine başlamadan önce ilgili ayda gerçekleşmiş olan haftalık web sayfası ziyaret miktarları istatistikleri incelenmiştir (Şekil 2).



Şekil 2: 2020 Mart ayı haftalık web sayfaları toplam ziyaret miktarları istatistiği.

Yapılan incelemeler doğrultusunda, örgün eğitime ara verme kararının alındığı tarihe yakın olması amacıyla ilgili ayın ilk haftası elenerek, son üç haftası veri setinde bırakılmıştır. Sonuç itibariyle, toplam 170.148 satırdan oluşan “ziyaret” veri seti ile birlikte ilgili süreçte toplanmış olan 12.869 satırlık “bilgi erişim” veri seti çalışmada kullanılmıştır. Buna ek olarak, “ziyaret” veri setinin tutulduğu tablo içinde çalışmanın üçüncü hedefinin analizinde kullanılacak veri seti ise toplam 4.012 satırdan oluşmuştur.

Dönüştürme

Dönüştürme aşaması, Han, Kamber ve Pei (2011) ve Özkan (2013) tarafından veri ön işleme adımlarından biri olarak belirtilmiş olan “Veri Dönüştürme” aşamasına denk gelmektedir. Bu aşama, Fayyad, Piatetsky-Shapiro ve Smyth (1996) tarafından belirtilen VBK sürecinde ayrı bir adım olarak belirtilmesinden dolayı, bu çalışmada, ön işleme aşaması haricinde ayrı bir bölüm olarak aktarılmıştır.

Bu aşamada, ilgili süreçte oluşturulan veri setlerindeki kayıtların, çalışmada yararlanılan Apriori algoritmasının işleyebileceği bir biçime dönüştürülme işlemi gerçekleştirilmiştir. Bu işlemin nasıl gerçekleştirildiğine yönelik bir örnek Tablo 1’de aktarılmıştır.

Tablo 1: Veri setlerinin dönüştürülme formatı.

ZİYARET DAVRANIŞLARININ DÖNÜŞÜMÜ			
Ham Veri		Dönüştürülmüş Veri	
<i>ip</i>	<i>sayfa_id</i>	<i>ip</i>	<i>sayfa_id_birlikte</i>
1.1.1.1	A	1.1.1.1	A, E
2.2.2.2	B	2.2.2.2	B, C, F
2.2.2.2	C	3.3.3.3	D
3.3.3.3	D		
1.1.1.1	E		
2.2.2.2	F		

BİLGİ ERİŞİM DAVRANIŞLARININ DÖNÜŞÜMÜ			
Ham Veri		Dönüştürülmüş Veri	
<i>ip</i>	<i>arama_terimi</i>	<i>ip</i>	<i>arama_terimi_birlikte</i>
1.1.1.1	T1	1.1.1.1	T1, T5
2.2.2.2	T2	2.2.2.2	T2, T3, T6
2.2.2.2	T3	3.3.3.3	T4
3.3.3.3	T4		
1.1.1.1	T5		
2.2.2.2	T6		

Tablo 1’den anlaşılacağı üzere, her bir *ip* adresinin gerçekleştirmiş olduğu farklı davranışlar tek bir satırda toplanarak yeni veri setleri oluşturulmuştur. Bu durumda, her bir *ip* adresi bağımsız bir kullanıcıyı temsil etmiştir. Tablo 1’de aktarılmayan üçüncü hedefe yönelik veri seti dönüşümü, temel olarak “ziyaret” davranışlarındaki dönüşümün aynısıdır. Üçüncü hedefe yönelik dönüşüm, 2. tabloda yer alan *bilgi_erisim_id* niteliği dolu olan satırlar kullanılarak “ziyaret”inki ile aynı şekilde gerçekleştirilmiştir.

Veri Madenciliği

Dönüşüm işlemi yapılmış son veri setlerinin analizinde, Agrawal ve Srikant (1994) tarafından geliştirilen Apriori birliktelik algoritmasından faydalanılmıştır. Apriori ve benzeri birliktelik algoritmaları genellikle market sepet analizlerinde tüketici satın alma eğilimlerinin tespit edilmesi amacıyla kullanılan algoritmalarlardır. Bu çalışmada, web kullanıcılarının “bilgi erişim” ve “ziyaret” eğilimlerinin ya da davranış desenlerinin keşfedilmesi için Apriori algoritması tercih edilmiştir. Bir birliktelik algoritmasının temel kapsamı aşağıdaki şekilde ifade edilmiştir (Agrawal, Imieliński ve Swami, 1993; Cios ve diğ., 2007):

$I = \{i_1, i_2, \dots, i_m\}$ bir öğeler kümesi olsun. T , kullanıcılar tarafından gerçekleştirilmiş her bir işlemi (bu çalışma için “bilgi erişim” veya “ziyaret” davranışdır) ifade eder ve kendine özgü bir işlem numarasına (TID) sahiptir. İşlem numarası bu çalışma kapsamında kullanıcıların ip değerleridir. D , her bir işlemin bir TID değeriyle $T \subseteq I$ olacak şekilde birleştirildiği tüm işlemlerden oluşan veri setini ifade eder. A ve B 'nin iki öğe kümesi olduğu düşünüldüğünde, sadece $A \subseteq T$ şartının sağlanması durumunda T işleminin A 'yı içerdiği söylenir. $A \subset B$, $A \subset I$, $B \subset I$ ve $A \cap B = \emptyset$ şartlarının sağlandığı durumdaysa, birliktelik kuralı $A \Rightarrow B$ biçiminde ifade edilir.

Birliktelik kuralları oluşturulurken destek (support) ve güven (confidence) olmak üzere iki parametre kullanılmaktadır. Destek parametresi, yukarıdaki tanım dikkate alındığında, A ve B 'yi içeren işlem sayısının toplam işlem sayısına oranıdır (Denklem 1). Bir başka ifadeyle hem A hem de B 'nin D işlem kümesinde birlikte bulunma olasılığıdır (Cios ve diğ., 2007).

$$\text{destek}(A \Rightarrow B) = P(A \cup B) = \frac{||\{T \in D \mid A \cup B \subseteq T\}||}{||D||} \quad (1)$$

Güven parametresi ise, A ve B 'yi içeren işlem sayısının A 'yı içeren işlem sayısına oranı olup (Denklem 2), birliktelik kuralının gücünü ifade eder (Cios ve diğ., 2007).

$$\text{güven}(A \Rightarrow B) = P(B|A) = \frac{||\{T \in D \mid A \cup B \subseteq T\}||}{||\{T \in D \mid A \subseteq T\}||} \quad (2)$$

Veri dönüştürme ve madencilik işlemleri Python yazılım dilinin 3.7 versiyonu kullanılarak gerçekleştirilmiştir. Apriori algoritmasının uygulanabilmesi amacıyla da, pandas (Pandas, 2020) ve apyori (Apyori, 2020) yazılım kütüphanelerinden faydalanılmıştır.

Çalışma Sınırlılıkları

IP adresleri belirli zaman aralığında sadece bir kişiye özel olarak sabit tutulmakta ve bu zaman aralığı geçildiğinde değişebilmektedir. Analiz edilecek son veri seti hazırlanırken, her bir IP değeri, bağımsız olarak o IP değerine ait davranışların birlikte gruplanmasında kullanılmıştır. Bu noktada, aynı IP değerinin farklı bir zaman diliminde farklı bir kullanıcıyı temsil etme olasılığı bulunmaktadır. Bu değişkenlik, çalışma için bir sınırlılık olarak ortaya çıkmıştır.

4. BULGULAR

Çalışmada yapılmış olan analizler doğrultusunda ortaya çıkarılan bulgular, çalışmanın hedefleri doğrultusunda bu bölümde kategorize edilmiştir. Yapılan analizlerde, hem “bilgi erişim” hem de “ziyaret” bakımından farklı destek ve güven kriterleri denenmiştir. Her üç hedef için daha düşük destek ve güven eşikleri denendiğinde, birbirine benzer desenlerin ortaya çıktığı ve kural miktarlarının çok fazla miktarlara ulaştığı gözlemlenmiştir. Tam tersi şekilde daha büyük değerli eşik değerlerinde de çalışmada paylaşılan desenlerin ortaya çıkmadığı belirlenmiştir. Bu doğrultuda, en anlamlı olacak şekilde desenleri çıkaran eşik değerleri kullanılmış ve ortaya çıkan davranış desenleri bu bölüm altında paylaşılmıştır. Ek olarak, kullanıcılar tarafından ziyaret edilmiş olan sayfalara ait başlıkların daha sade görünmesi amacıyla Tablo 2’deki kodlama biçimleri hem ikinci hem de üçüncü hedef için temel alınmıştır.

Tablo 2: Kullanıcılar tarafından ziyaret edilmiş sayfalara ait kodlamalar.

Sayfa Kodu	Ziyaret Edilmiş Sayfa Başlıkları
S1	Koronavirüs İçin Alınan Tedbirler Kapsamında Üniversitemiz Personeli ve Öğrencilerine Bilgilendirme
S2	Öğrencilerimize Tatil Duyurusu
S3	Koronavirüs Riskine Karşı 14 Kural
S4	Uzaktan Öğretim Duyurusu
S5	Rektörümüz Şengörür'den Öğrencilerimize Uzaktan Öğretim Mesajı
S6	“Yök Dersleri Platformu” Öğrencilerin Erişimine Açıldı
S7	Kırklareli Üniversitesi Koronavirüs Bilgilendirme Komisyonu Web Sayfası
S8	Bahar Yarıyılında Yüz Yüze Eğitim Yapılmayacaktır
S9	Tez Yazım Kılavuzu ve Şablonu Güncellendi!
S10	Tez Yazım Kılavuzu
S11	Yaz Stajı Yapan Öğrencilerimizin Dikkatine
S12	2019-2020 Eğitim Öğretim Yılı Yaz Döneminde Staj Yapacak Öğrencilerimizin Dikkatine
S13	Tez İşlemleri
S14	Sosyal Bilimler Enstitüsü Tez Yazım Kılavuzu Yenilendi
S15	30 Eylül 2019 Tarihli Öğretim Elemanı İlanının Ön Değerlendirme Sonuçları
S16	04 Aralık 2019 Tarihli Araştırma Görevlisi ve Öğretim Görevlisi İlanının Ön Değerlendirme Sonuçları
S17	Yatay Geçiş Başvuru Sonuçları
S18	2019-2020 Eğitim-Öğretim Yılı Bahar Yarıyılı Yatay Geçiş Başvuru Sonuçları
S19	Öğrencilerimizin Dikkatine
S20	Eğitim Öğretime Ara Verilmesi Hakkında
S21	Öğrencilerimize Tatil Duyurusu
S22	Eğitime 3 Hafta Ara Verilmesi Hakkında Duyuru
S23	Koronavirüs İçin Alınan Tedbirler Kapsamında Üniversitemiz Personeli ve Öğrencilerine Bilgilendirme

Kullanıcıların "bilgi erişim" açısından yaptığı aramalardaki birlikteliklerin keşfedilmesi

KLU web kullanıcıları tarafından web sitelerinde yer alan site-içi arama araçlarında yapılmış aramalar üzerinden ortaya çıkarılmış desenler Tablo 3'de yer almaktadır.

Tablo 3: KLU web kullanıcılarının "bilgi erişim" davranış desenleri.

Kural	Birliktelik Deseni	Destek*	Güven**
1	tez yazım -> kılavuz	0.00608	0.69565
2	kılavuz -> tez yazım kılavuzu	0.00456	0.52173
3	lisans -> sınav programı	0.00171	0.9
4	lisansüstü başvuru -> yüksek lisans	0.00152	0.42105
5	mezuniyet töreni -> mezuniyet	0.00209	0.42307
6	vize sınav -> sınav programı	0.00133	0.43749
7	tez -> tez yazım kılavuzu	0.00855	0.62500
8	tez yazım -> tez yazım kılavuzu	0.00418	0.47826
9	şablon -> tez yazım kılavuzu	0.00114	1.0
10	tezsiz yüksek lisans -> yüksek lisans	0.00152	0.47058
11	yazokulu -> yaz okulu	0.00228	0.8
12	tez, kılavuz -> tez yazım	0.00247	0.76470
13	tez, kılavuz -> tez yazım kılavuzu	0.00266	0.82352
14	tez yazım kılavuzu, kılavuz -> tez yazım	0.00323	0.70833
15	tez yazım, tez -> tez yazım kılavuzu	0.00209	0.73333
16	tez, kılavuz -> tez yazım kılavuzu, tez yazım	0.00190	0.58823

*: Minimum destek kriteri 0.001 olarak ayarlanmıştır.
**: Minimum güven kriteri 0.4 olarak ayarlanmıştır.

Tablo 3 incelendiğinde, çoğunlukla "tez yazım" temalı aramaların farklı terimler kullanılarak gerçekleştirildiği göze çarpmaktadır. Bu noktada, özellikle lisansüstü öğrencilerin tez yazımına yönelik "bilgi erişim" davranışlarının ağır bastığı görülmektedir. Farklı terimler kullanılması, öğrencilerin ihtiyaçlarını giderecek bilgiye erişmekte zorlandıklarını ifade edebilir. Buna ek olarak, eğitime ara verilme kararının öğrenciler üzerinde bir stres oluşturma durumu da söz konusu olabilir. "tez yazım" dışında yapılan aramalara bakıldığında, 3. ve 6. kural, lisans öğrencilerinin vize sınavlarının nasıl yapılacağına yönelik bir araştırma eğiliminde olduklarını göstermektedir. 4. ve 10. kural birlikte değerlendirildiğinde, lisansüstü eğitim görmek isteyen öğrencilerin de ilgili dönemde bilgi edinme ihtiyaçlarının olduğunu göstermiştir. 5. ve 11. kurallara odaklanıldığında, KLU öğrencilerinin eğitime ara verilen bu süreçte "mezuniyet" ve "yaz okulu" temasıyla bilgi ihtiyaçlarını gidermeye çalıştıkları tespit edilmiştir.

Kullanıcıların "ziyaretler" açısından sergilediği davranışlardaki birlikteliklerin keşfedilmesi

KLU web kullanıcılarının, “bilgi erişim” davranışı göstermeden çeşitli web sayfalarına yaptıkları ziyaretler üzerinden ortaya çıkarılan desenler Tablo 4’de aktarılmıştır.

Tablo 4: KLU web kullanıcılarının web sayfalarına yaptıkları "ziyaret" desenleri.

Kural	Birliktelik Deseni		Destek*	Güven**	
1	S2	->	S1	0.03786	0.35001
2	S3	->	S1	0.02169	0.82886
3	S1	->	S4	0.05780	0.28674
4	S5	->	S1	0.02826	0.31357
5	S6	->	S1	0.01277	0.25940
6	S5	->	S4	0.04137	0.45908
7	S7	->	S4	0.01131	0.45778
8	S6	->	S4	0.01804	0.36645
9	S6	->	S5	0.01638	0.33284
10	S7	->	S6	0.01523	0.61644
11	S6	->	S8	0.01030	0.20935
12	S5, S1	->	S4	0.02004	0.70910
13	S6, S5	->	S4	0.01038	0.63336

*: Minimum destek kriteri 0.01 olarak ayarlanmıştır.

**: Minimum güven kriteri 0.2 olarak ayarlanmıştır.

Tablo 4 incelendiğinde, “uzaktan öğretim”, “koronavirüs” ve “tatil” temalı sayfa ziyaretlerinin birlikte gerçekleştirildiği tespit edilmiştir. Bununla birlikte, koronavirüs’e yönelik bilgi içeren sayfaların birlikte ziyaret edildiği görülmektedir (2. kural). Ek olarak, “Yök Dersleri Platformu” ve “uzaktan öğretim” temalı ziyaretlerin birlikte yapıldığı da belirlenmiştir. Farklı sayfalar üzerinden ziyaretçilerin ortaya çıkardığı bu desen, üniversitenin uzaktan öğretime yönelik öğrencilere alternatif sunma durumunun takip edildiğini göstermiştir.

Kullanıcıların “bilgi erişim” doğrultusunda sergilediği “ziyaret” davranışlarındaki birlikteliklerin keşfedilmesi

“bilgi erişim” davranışı sonrasında, KLU web kullanıcılarının karşısına aradıkları terimle ilgili web sayfalarının bulunduğu bir liste getirilmiştir. Bu liste üzerinde herhangi bir web sayfasının tıklanması doğrultusunda gerçekleşen “ziyaret” davranışlarına ait birliktelik desenleri Tablo 5’deki şekilde ortaya çıkmıştır.

Tablo 5: KLU web kullanıcılarının "bilgi erişim" doğrultusunda sergilediği "ziyaret" davranış desenleri.

Kural	Birliktelik Deseni		Destek*	Güven**	
1	S10	->	S9	0.00244	0.83333
2	S11	->	S12	0.00244	0.71428
3	S13	->	S14	0.00244	0.41666

4	S13	->	S9	0.00488	0.83333
5	S15	->	S16	0.00342	1.0
6	S17	->	S18	0.00244	0.5
7	S19	->	S20	0.00244	0.45454
8	S21	->	S20	0.00244	0.83333
9	S21	->	S22	0.00244	0.41666
10	S23	->	S22	0.00488	0.5

*: Minimum destek kriteri 0.002 olarak ayarlanmıştır.

***: Minimum güven kriteri 0.4 olarak ayarlanmıştır.

Tablo 5 incelendiğinde, tıpkı Tablo 3’deki gibi “tez yazım” temalı aramalarla ilgili ziyaretlerin çoğunlukla gerçekleştiği belirlenmiştir. Bu bulguya ek olarak, “tatil” ve “eğitim öğretimin ertelenmesi” temalı ziyaretlerin birliktelikleri tespit edilmiştir. “bilgi erişim” davranışlarında ortaya çıkmamış olmasına rağmen, “yaz stajı” temalı ziyaretler dikkat çekicidir. Yine, “yatay geçiş” temalı ve “kadro ilanı” temalı ziyaretler Tablo 3’de bulunmamasına rağmen bir arama sonucunda gerçekleşen sayfa ziyaretlerini ifade etmektedir. “bilgi erişim” çerçevesinde bu davranışların Tablo 3’de gözlemlenmemesinin sebebi, bu bilgiye ihtiyaç duyan kitlenin az olmasından kaynaklanıyor olabilir. “koronavirüs” temalı ziyaretlerin sıklıkla gözlemlendiği Tablo 3’e nazaran, Tablo 5’de sadece 10. satır bu bilgi ihtiyacını temsil etmiştir. Bu noktada, kullanıcıların koronavirüs ile ilgili arama yapmaya fazla ihtiyaç duymadıkları ve ilgili sayfaların zaten rahatlıkla erişilebilir olduğu belirtilebilir.

5. TARTIŞMA VE SONUÇ

Dünya genelinde yaşanan Covid-19 pandemisinin ülkemizde etkinliğini arttırdığı süreç temel alınarak 2020 yılının Mart ayına odaklanılan bu çalışmada, Kırklareli Üniversitesi web kullanıcılarının “bilgi erişim” ve “ziyaret” davranışlarındaki birliktelik desenlerinin incelenmesi amaçlanmıştır. Bu amaç doğrultusunda, web kullanıcılarının bilgi erişim ve ziyaret davranışları bağımsız ve birlikte olacak şekilde incelenerek, çeşitli açılardan davranış desenleri tespit edilmiştir. Apriori birliktelik algoritması sayesinde ortaya çıkarılmış olan bu desenlerin, kullanıcıların bilgi ihtiyaçlarıyla alakalı önemli noktaları gözler önüne serdiği düşünülmektedir. Birliktelik algoritmaları, çoğunlukla market sepet analizlerinde tüketici satın alma eğilimlerinin tespit edilmesi amacıyla kullanılan algoritmalarlardır. Ortaya çıkarılan tüketici satın alma desenlerinin keşfi doğrultusunda daha verimli satış stratejileri geliştirilebilmektedir. Bu durum, bu çalışma için farklı bir şekilde değerlendirilmektedir. Web kullanıcılarının, aynı amaç için farklı davranışlar ortaya çıkarması, sistem geliştiriciler açısından bir dezavantaj olarak nitelendirilebilir. Daha açık bir ifadeyle, kullanıcıların, bilgi ihtiyaçlarını giderme aşamasında daha fazla zaman ve emek harcadıkları anlamına gelebilir. Bu noktada, ortaya çıkan desenlerden yola çıkılarak, aynı amaca işaret eden farklı davranışların azaltılmasının, kritik derecede önemli olduğu düşünülmektedir.

Çalışma bulguları göstermektedir ki, ilgili dönemde lisansüstü öğrenciler çok fazla miktarda “bilgi erişim” davranışı sergilemiş ve bu davranışların merkezindeyse “tez yazım” teması bulunmuştur. Bu davranışlar, farklı terimler kullanılmasına rağmen, genel itibariyle aynı amaç için gerçekleştirilmiş davranışlardır. Bu noktadan hareketle, lisansüstü öğrencilerin “tez yazım” konusunda bilgi edinme sıkıntısı yaşadıkları oldukça belirgindir. Bu sıkıntının giderilmesi ve öğrencilerin bu konuda “bilgi erişim” davranışlarının indirgenmesi amacıyla, lisansüstü eğitim veren birimlerdeki sistem yöneticilerinin, ilgili konuda detaylı bir rehber web sayfası hazırlayarak web sitelerinde daha belirgin bir şekilde sunmaları önerilmektedir. Buna ek olarak, örgün eğitimin tekrar başlatılması sonrasında lisansüstü öğrencilerin (özellikle de tez dönemindeki öğrenciler) bu konudaki bilgi eksikliklerinin giderilmesi amacıyla, her dönem başında gerekli bir eğitimin ya da bilgilendirmenin yüzyüze bir ortamda yapılması sağlanabilir.

“ziyaret” davranışları üzerinden ortaya çıkmış olan desenlerde, “uzaktan öğretim”, “koronavirüs” ve “tatil” temalı sayfa ziyaretlerinin çoğunlukta olduğu gözlemlenmiştir. Farklı temalar olmasına rağmen, ilgili dönemde yaşanan bilgi ihtiyacı genel olarak eğitim-öğretimin nasıl sürdürüleceğine yöneliktir. Bu doğrultuda, farklı sayfalar oluşturularak sunulan ancak yine de aynı ortak amaca sahip bu web sayfalarının tek bir noktada buluşturulması önemlidir. Her ne kadar bahsedilen bu dönem geçmişte kalmış olsa da, aynı bilgi ihtiyacının, örgün eğitime geçiş yapıldığında da tekrar gündem olması muhtemeldir. İlgili dönemde bu ihtiyaca yönelik olarak “Koronavirüs Bilgilendirme Komisyonu” ismiyle üniversitede yeni bir web sayfası (kovkom.klu.edu.tr) açılmıştır. Aynı amaca hitap eden web sitelerinin ilgili dönemde tüm üniversiteler tarafından da oluşturulduğu bilinmektedir. Bu yöntemin olumlu olduğu düşünülmekle beraber, web sitelerinin ilerleyen süreçte ortaya çıkması muhtemel bilgi ihtiyacını karşılayacak şekilde güncellenmesi önerilmektedir. Bu güncelleştirmeler, çalışmada sunulan yöntemin periyodik olarak izlenmesi ve güncel bilgi ihtiyaçlarının ortaya çıkarılması doğrultusunda yapılabilir. Bunlara ek olarak, ortaya çıkarılan ihtiyaçlardan yola çıkılarak “Sıkça Sorulan Sorular (SSS)” adı altında bir sayfanın hazırlanması da sağlanabilir.

“bilgi erişim” doğrultusunda sergilenen “ziyaret” davranışları incelendiğinde, “tez yazım”, “tatil” ve “eğitim öğretimin ertelenmesi” temalı ziyaretlerin birliktelikleri ön plana çıkmıştır. Bu noktada, hem “bilgi erişim” davranış desenlerinde ortaya çıkan hem de çalışmanın üçüncü hedefinde ortaya çıkan bilgi ihtiyaçları, birlikte değerlendirilerek önlemler alınabilir. Bunun için, “Tez Yazımı ve Tez Yazım Kılavuzu”, “Eğitim Öğretimin Ertelenmesi Hakkında” ve “Uzaktan Eğitim Süreci” olmak üzere üç farklı görsel hazırlanarak, ilgili bilgiyi sunan sayfalara yönlenecek bir bağlantıyla birlikte kullanıcıların arama yaptıkları sayfalara eklenebilir. Bu öneri, ilgili dönemi kapsamasına rağmen, daha güncel desenlerin ortaya çıkarılmasıyla birlikte farklı başlıklar üzerinden gerçekleştirilebilir.

Sonuç olarak, bu çalışmada aktarılmış olan yöntem izlenerek gözle görülemeyen dönemsel bilgi ihtiyaçlarının ortaya çıkarılması mümkündür. Ortaya çıkan bilgi ihtiyaçları temel alınarak yapılacak güncellemeler sayesinde, web sitelerinin geliştirilmesi ve kullanıcı ihtiyaçlarına daha iyi hitap edecek bir hizmetin sunulması sağlanabilir. Dönemsel olarak belirginleşen ve genellenebilir bilgi ihtiyaçlarının giderilmesi doğrultusunda da ilerleyen süreçte aynı desenlerin ortaya çıkma ihtimali düşürülebilir.

KAYNAKÇA

- Aggarwal, C. C. (2015). *Data Mining: The Textbook* (2015 edition). Springer.
- Agrawal, R., Imieliński, T., ve Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, 207–216. <https://doi.org/10.1145/170035.170072>
- Agrawal, R., ve Srikant, R. (1994). Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, 487–499.
- Baeza-Yates, R., ve Ribeiro-Neto, B. (2011). *Modern Information Retrieval: The concepts and technology behind search*, 2. Baskı. Addison-Wesley Publishing Company.
- Bilgiç, E. (2019). R Programlama Dili İle Pazar Sepet Analizi: Muş İl Merkezindeki Bir Süpermarkette Tüketicilerin Satın Alma Davranışlarının Tespiti Üzerine Bir Uygulama. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 7(3), 89-97. <https://doi.org/10.18506/anemon.462998>
- Britannica, E. (2007). *Britannica Concise Encyclopedia*. Encyclopedia Britannica, Inc.
- Budak, V. Ö., Kartal, E., ve Gülseçen, S. (2018). Site-içi Aramalar ve Apriori Algoritması Kullanılarak Web Sitesi Ziyaretçilerinin İhtiyaç Tespitine Yönelik Bir Örnek Olay İncelemesi. *Bilişim Teknolojileri Dergisi*, 11(2), 211-222. <https://doi.org/10.17671/gazibtd.397142>
- Chakrabarti, S. (2002). *Mining the Web: Discovering Knowledge from Hypertext Data* (1 edition). Morgan Kaufmann.
- Chauhan, A., ve Tarar, S. (2016). Prediction of User Browsing Behavior Using Web Log Data. *International Journal of Scientific Research in Science, Engineering and Technology*, 2(1), 419-422. <https://doi.org/10.32628/IJSRSET1621108>
- Cios, K. J., Pedrycz, W., Swiniarski, R. W., ve Kurgan, L. A. (2007). *Data Mining: A Knowledge Discovery Approach* (2007 edition). Springer.
- Çelik, S. (2017). Web Günlük Dosyalarının Analizi için Web Kullanım Madenciliğinin Uygulanması. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 46(1), 62-75.
- Çınar, I., ve Bilge, H. Ş. (2016). Web Madenciliği Yöntemleri ile Web Loglarının İstatistiksel Analizi ve Saldırı Tespiti. *Bilişim Teknolojileri Dergisi*, 9(2), 125-0.

- Fayyad, U., Piatetsky-Shapiro, G., ve Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3), 37-37. <https://doi.org/10.1609/aimag.v17i3.1230>
- Gezer, M., Erol, Ç., ve Gulsecen, S. (2007, Ocak 1). *Bir Web Sayfasının Veri Madenciliği İle Analizi*. Akademik Bilişim, Kütahya. <https://doi.org/10.13140/2.1.5004.7687>
- Grace, J., V.Maheswari, ve Dhinaharan, N. (2011). Analysis of Web Logs And Web User In Web Mining. *International Journal of Network Security & Its Applications*, 3. <https://doi.org/10.5121/ijnsa.2011.3107>
- Han, J., Kamber, M., ve Pei, J. (2011). *Data Mining: Concepts and Techniques* (3 edition). Morgan Kaufmann.
- Ilampiray, P. (2012). Efficient resource utilization of web using data clustering and association rule mining. *Journal of Theoretical and Applied Information Technology*, 37(2), 211-216. Scopus.
- Kindie, A., Mamuye, A., ve Tilahun, B. (2018). Web usage characterization for system performance improvement. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 244, 238-245. Scopus. https://doi.org/10.1007/978-3-319-95153-9_23
- Kotiyal, B., Kumar, A., Pant, B., Goudar, R. H., Chauhan, S., ve Juneja, S. (2013). User behavior analysis in web log through comparative study of Eclat and Apriori. *2013 7th International Conference on Intelligent Systems and Control (ISCO)*, 421-426. <https://doi.org/10.1109/ISCO.2013.6481192>
- Liu, B. (2011). *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, 2. Baskı. Springer-Verlag. <https://doi.org/10.1007/978-3-642-19460-3>
- Markov, Z., ve Larose, D. T. (2007). *Data Mining the Web: Uncovering Patterns in Web Content, Structure, and Usage* (1 edition). Wiley-Interscience.
- Apyori (2020). *apyori: Simple Apriori algorithm Implementation*. (1.1.2). Ziyaret Tarihi: 13 Haziran 2020, Adres: <https://github.com/ymoch/apyori>
- Özkan, Y. (2013). *Veri madenciliği yöntemleri*, 2. Baskı. Papatya Yayıncılık.
- Pandas (2020). *pandas: Powerful data structures for data analysis, time series, and statistics* (1.0.4). Erişim Tarihi: 13 Haziran 2020, Adres: <https://pandas.pydata.org>
- Sağın, A. N., ve Ayvaz, B. (2018). Determination of Association Rules with Market Basket Analysis: Application in the Retail Sector. *Southeast Europe Journal of Soft Computing*, 7(1), Article 1. <https://doi.org/10.21533/scjournal.v7i1.149>
- Sathya, M., ve Devi, P. I. (2018). *Apriori algorithm on web logs for mining frequent link*. 2018-February, 1-5. Scopus. <https://doi.org/10.1109/ITCOSP.2017.8303127>
- Şimşek Gürsoy, U. T., Akçay Kasapoğlu, Ö., ve Atalay, K. (2019). R Programlama İle Birliktelik Kuralları Analizi: Tüketicilerin İnternet Üzerinden Yaptıkları Alışveriş Verisinin Apriori ve Eclat

Algoritmalarıyla İncelenmesi. *Alphanumeric Journal*, 7(2), 357-368.
<http://dx.doi.org/10.17093/alphanumeric.585663>

TDK (2020). *Türk Dil Kurumu | Sözlük*. Eriřim Tarihi: 10 Haziran 2020, Adres: <https://sozluk.gov.tr/>