

T.C.  
İSTANBUL AYDIN ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



WEB 3.0 VERİ TOPLAMA TEKNİKLERİ VE PANDEMİ İLE DEĞİŞEN  
ONLINE SEPET ANALİZİ

YÜKSEK LİSANS TEZİ

İsmail GÜLER

Bilgisayar Mühendisliği Ana Bilim Dalı

Bilgisayar Mühendisliği Programı

Temmuz, 2021



T.C.  
İSTANBUL AYDIN ÜNİVERSİTESİ  
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



**WEB 3.0 VERİ TOPLAMA TEKNİKLERİ VE PANDEMİ İLE DEĞİŞEN  
ONLINE SEPET ANALİZİ**

**YÜKSEK LİSANS TEZİ**  
**İSMAİL GÜLER**  
**(Y1713.010024)**

**Bilgisayar Mühendisliği Ana Bilim Dalı**  
**Bilgisayar Mühendisliği Programı**

**Tez Danışmanı: Prof. Dr. Zafer ASLAN**

**Temmuz, 2021**

## ONAY BELGESİ

## YEMİN METNİ

"Web 3.0 Veri Toplama Teknikleri ve Pandemi ile Değişen Online Sepet Analizi" konusunda yapmış olduğum çalışmamı bağımsız olarak ve üçüncü şahıslardan yardım almadan tamamladığımı belirtmek isterim. Belirtilen kaynaklar ve yardımcılarından başka kaynak kullanmadığımı, tezin ayrıca elektronik ortamda sunulması halinde yazılı ve elektronik versiyonlarının tamamen aynı olduğunu beyan ederim. (Temmuz, 2021).

**İsmail GÜLER**

## **ÖNSÖZ**

Bu çalışmanın amacı; Big Data ve Web 3.0 gerçek zamanlı verileri kullanarak yapay zekanın önemli bir rol oynadığı satış ve reklam konusunda daha aktif bir sistem önerisidir.

Yüksek lisans eğitimim ve her zaman yol göstericim olduğu için Prof. Dr. Zafer ASLAN'a tüm samimiyetimle teşekkür ederim. Bu uzun süreç boyunca sürekli desteği, bilimsel rehberliği, sürekli iyimserliği ve teşviki için ona çok minnettarım.

Lisans ve Yüksek lisans eğitimi ve tez çalışmalarım süresince yardım ve destekleri için Prof. Dr. Ali GÜNEŞ ve Doç. Dr. Metin ZONTUL'a teşekkür ederim.

Son olarak bana destek veren, cesaretlendiren, her zaman yanımda olduklarını hissettiğim sevgili ailem ve arkadaşlarıma teşekkür ederim.

**Temmuz 2021**

**İsmail GÜLER**

## İÇİNDEKİLER

<b>YEMİN METNİ</b> .....	<b>iii</b>
<b>ÖNSÖZ</b> .....	<b>iv</b>
<b>İÇİNDEKİLER</b> .....	<b>v</b>
<b>KISALTMALAR</b> .....	<b>vii</b>
<b>TANIMLAMALAR</b> .....	<b>viii</b>
<b>ŞEKİL LİSTESİ</b> .....	<b>ix</b>
<b>TABLO LİSTESİ</b> .....	<b>xi</b>
<b>ÖZET</b> .....	<b>xii</b>
<b>ABSTRACT</b> .....	<b>xiii</b>
<b>1. GİRİŞ</b> .....	<b>1</b>
1.1 Araştırma Problemi.....	1
1.2 Araştırmanın Amacı.....	2
1.3 Varsayımlar .....	4
<b>2. WEB TEKNOLOJİSİNİN TARİHSEL GELİŞİMİ</b> .....	<b>5</b>
2.1 Web 1.0 Yapısı.....	5
2.2 Web 2.0 Yapısı.....	6
2.3 Web 1.0 ve Web 2.0 Arasındaki Farklar.....	8
2.4 Web 3.0 .....	9
2.5 Web 3.0 Geçiş Süreci.....	10
<b>3. WEB 3.0 (SEMANTİK WEB)</b> .....	<b>13</b>
3.1 Semantik Web Mimarisi .....	14
3.2 Semantik Web Katman Yapıları .....	16
3.2.1 Hiper Metin Web Teknolojileri .....	16
3.2.2 Anlamsal Web Teknolojileri.....	16
3.2.3 Gerçekleşmemiş Anlamsal Web Teknolojileri .....	17
3.3 Web Servisleri.....	18
3.4 Ajanlar.....	19
3.5 Semantik Teknolojiler ile Uygulamaya Entegrasyon Süreci .....	20
3.5.1 Anlamsal Veri Entegrasyonu .....	20
3.5.2 Heterojen Veri Kümelerini Entegre Etme.....	21
3.6 Ontoloji .....	22
3.6.1 Ontoloji Web Dili (OWL).....	23
3.6.2 Ontoloji Uygulama Senaryoları .....	25
3.7 Kullanılan Programlama Dilleri.....	26
3.7.1 Genişletilebilir İşaretleme Dili (XML) .....	26
3.7.2 Kaynak Tanımlama Çerçevesi (RDF).....	27
3.7.3 Kaynak Tanımlama Çerçevesi Şeması (RDFS).....	28
3.8 Test Süreçleri .....	29
3.9 Veri Tabanı Yapısı.....	30

<b>4. PAZAR SEPET ANALİZİ VE BİRLİKTELİK KURALLARI .....</b>	<b>33</b>
4.1 Temel Kavramlar .....	33
4.2 Pazar Sepeti Analizinin Güçlü ve Zayıf Yönleri .....	34
4.3 Birliktelik Kuralları.....	35
4.3.1 Destek (Support) .....	35
4.3.2 Güven Seviyesi (Confidence) .....	36
4.3.3 Kaldıraç (Lift) .....	36
4.4 Birliktelik Kuralların Belirlenmesinde Kullanılan Algoritmaları.....	37
4.4.1 Apriori Algoritması.....	37
<b>5. VERİ.....</b>	<b>41</b>
5.1 Verilerin Açıklanması ve Hazırlanması.....	41
5.2 Veri Toplama .....	42
<b>6. SONUÇLAR .....</b>	<b>45</b>
6.1 Pazar Sepet Analizi .....	45
6.1.1 Python'da Apriori Algoritmasının Uygulanması .....	46
6.1.2 Apriori Algoritması ile Elde Edilen Birliktelik Kuralları .....	50
6.2 Analiz Sonuçları.....	59
6.3 Gelecek Çalışmalar ve Öneriler .....	63
<b>KAYNAKÇA .....</b>	<b>65</b>
<b>EKLER.....</b>	<b>71</b>
<b>ÖZGEÇMİŞ.....</b>	<b>76</b>



## KISALTMALAR

<b>AI</b>	:Artificial Intelligence (Yapay Zekâ)
<b>AJAX</b>	:Asynchronous JavaScript and XML (Eşzamansız JavaScript ve XML)
<b>HTML</b>	:Hyper Text Markup Language (Hiper Metin İşaretleme Dili)
<b>JSON</b>	:JavaScript Object Notation (JavaScript Nesne Notasyonu)
<b>J2EE</b>	:Java 2 Enterprise Edition (Java 2 Kurumsal Sürüm)
<b>IoT</b>	:Internet of Things (Nesnelerin İnterneti)
<b>IRI</b>	:Internationalized Resource Identifier (Uluslararasılaştırılmış Kaynak Tanımlayıcı)
<b>OWL</b>	:Web Ontology Language (Web Ontoloji Dili)
<b>RDF</b>	:Resource Description Framework (Kaynak Tanımlama Çerçevesi)
<b>RDFS</b>	:Resource Description Framework Schema (Kaynak Tanımlama Çerçeve Şeması)
<b>RDB</b>	:Relational Databases (İlişkisel Veri Tabanları)
<b>RSS</b>	:Rich Site Summary (Zengin Site Özeti)
<b>sn</b>	:Saniye (Second)
<b>SPARQL</b>	:SPARQL Protocol and RDF Query Language (SPARQL Protokolü ve RDF Sorgu Dili)
<b>SWRL</b>	:Semantic Web Rule Language (Semantik Web Kural Dili)
<b>URI</b>	:Uniform Resource Identifier (Standart Kaynak Tanımlayıcı)
<b>URL</b>	:Uniform Resource Loader (Standart Kaynak Yükleyici)
<b>WSDL</b>	:Web Services Description Language (Web Servisleri Tanımlama Dili)
<b>WWW</b>	: World Wide Web (Dünya Çapında Ağ)
<b>XML</b>	:Extensible Markup Language (Genişletilebilir İşaretleme Dili)

## TANIMLAMALAR

<b>Agents</b>	:Ajanlar
<b>Cookie</b>	:Çerez
<b>Confidence</b>	:Güven Seviyesi
<b>Consortium</b>	:Konsorsiyum
<b>Covid-19</b>	:Koronavirüs Hastalığı 2019
<b>Google Search</b>	:Google Arama Motoru
<b>Google Adwords</b>	:Google Online Pazarlama Aracı
<b>Intranet</b>	:Yerel Ağ
<b>Lift</b>	:Kaldıraç
<b>Metatype</b>	:Meta eklentisi
<b>Support</b>	:Destek
<b>Pseudo code</b>	: Sözde Kod
<b>Public</b>	:Açık
<b>Views</b>	:Görüntüleme
<b>World Wide Web Consortium:</b>	Dünya Çapında Ağ Konsorsiyumu
<b>World Wide Web</b>	:Dünya Çapında Ağ

## ŞEKİL LİSTESİ

<b>Şekil 1:</b> Web 1.0 Kullanıcı ile Makine İlişkisi [4].....	5
<b>Şekil 2:</b> Web 2.0 Kullanıcı, Makine ve Toplum İlişkisi [4] .....	7
<b>Şekil 3:</b> Web 3.0 Tasarım İlişkisi [7].....	10
<b>Şekil 4:</b> Web'in Gelişim Süreci [10]. .....	12
<b>Şekil 5:</b> Anlamsal Web Örneği [12] .....	13
<b>Şekil 6:</b> Semantik Web Teknolojisi Katman Yapısı [13]. .....	15
<b>Şekil 7:</b> WSDL Web Servis Yapısı [15].....	18
<b>Şekil 8:</b> Akıllı Kişisel Ajanlar .....	19
<b>Şekil 9:</b> Veri Entegrasyon Yapısı [17].....	20
<b>Şekil 10:</b> OWL Örneği.....	24
<b>Şekil 11:</b> XML Sınıf Etiketleri .....	26
<b>Şekil 12:</b> RDF ve RDFS İlişkisi .....	28
<b>Şekil 13:</b> Semantik Web Veri Tabanı Yapısı [25].....	31
<b>Şekil 14:</b> Apriori Algoritması Pseudo-Code [36].....	38
<b>Şekil 15:</b> Apriori Uygulama Adımları [38] .....	39
<b>Şekil 16:</b> Apriori algoritması tarafından oluşturulan kuralların pandemi öncesi dağılımı (destek=0.0045 ve güven seviyesi =0.6).....	47
<b>Şekil 17:</b> Apriori algoritması tarafından oluşturulan kuralların pandemi sonrası dağılımı (destek=0.0045ve güven seviyesi =0.6).....	48
<b>Şekil 18:</b> Apriori algoritması tarafından oluşturulan kuralların pandemi öncesi dağılımı (destek=0.0045ve güven seviyesi =0.7).....	49
<b>Şekil 19:</b> Apriori algoritması tarafından oluşturulan kuralların pandemi sonrası dağılımı (destek=0.0045 ve güven seviyesi =0.7).....	50
<b>Şekil 20:</b> 0,6 ile 0,7 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kuralların Pandemi Öncesi Grafiği.....	53
<b>Şekil 21:</b> 0,6 ile 0,7 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kuralların Pandemi Sonrası Grafiği.....	55

**Şekil 22:** 0,7 ile 0,8 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kuralların Pandemi Öncesi Grafiği..... 56

**Şekil 23:** 0,7 ile 0,8 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kuralların Pandemi Sonrası Grafiği..... 58

## TABLO LİSTESİ

<b>Tablo 1:</b> Kullanıcı Pazar Sepeti Veri Seti .....	44
<b>Tablo 2:</b> 0.6 ile 0.7 Arasında Güven Seviyesi Olan Pandemi Öncesi Kurallar Tablosu .....	52
<b>Tablo 3:</b> 0.6 ile 0.7 Arasında Güven Seviyesi Olan Pandemi Sonrası Kurallar Tablosu .....	54
<b>Tablo 4:</b> 0.7 ile 0.8 Arasında Güven Seviyesi Olan Pandemi Öncesi Kurallar Tablosu .....	55
<b>Tablo 5:</b> 0.7 ile 0.8 Arasında Güven Seviyesi Olan Pandemi Sonrası Kurallar Tablosu .....	57

## WEB 3.0 VERİ TOPLAMA TEKNİKLERİ VE PANDEMİ İLE DEĞİŞEN ONLINE SEPET ANALİZİ

### ÖZET

Son yıllarda, World Wide Web'in artan popülaritesi, web'de mevcut sayfa sayısında hızlı artışla beraber bir büyümeye yol açmıştır. Çok sayıda web sayfası, kullanıcıların gerekli bilgileri bulmasını giderek zorlaştırmaktadır. Web'de belirli bilgiler aranırken, çok sayıda ilgisiz arama sonucu zaman kaybedilir ve ilgili materyaller gözden kaçırabilmektedir. Mevcut web uygulamaları, içeriği yalnızca kendi dilinde temsil edilen HTML biçiminde web sayfaları sağlar ve bu nedenle içeriğin semantiğine makineler tarafından erişilemez. Makinelerin, kullanıcının bilgi problemlerini çözmeye desteklemesini sağlamak için Web 3.0 geliştirilmiştir. Bu tez ile web sayfalarının semantiklerini işlenebilir hale getiren mevcut web 3.0 araçları ile veri toplama yöntemleri önerilmektedir.

Günümüzde şirketler, büyük verilerin kullanımı açısından zengin, ancak bu verilerden elde edilen bilgilerin kullanımı ve paraya dönüştürülmesi açısından yetersizdir. Büyük veri değerli bir kaynak olarak görülüyor ve veri madenciliği kavramı hala yeni ve gelişmekte olsa da çeşitli endüstrilerdeki şirketler stratejik kararlar almak için bu verilere güvenmektedir. Fark edilmeyebilecek gerçekler, depolanan bilgileri eleyen tekniklerle ortaya çıkarılabilir. Pazar sepeti analizi, tüketici pazar sepetlerinde aynı anda bulunan öğeleri belirlemek için çok kullanışlı bir tekniktir. Bu tür bilgiler, promosyon desteği, envanter kontrolü ve çapraz satış kampanyaları gibi pazarlama faaliyetleriyle ilgili kararlar için bir temel olarak kullanılabilir.

Bir alışveriş sitesinin ürün yelpazesindeki farklı ürünlerin birbiriyle nasıl ilişkili olduğunu ve bu ilişkilerden pazarlama faaliyetleri ile Covid-19 pandemi döneminde nasıl yararlanılacağını bulgular ile göstermek amaçlanmıştır. İşlemsel verilerden madencilik ilişkilendirme kuralları, ürünlerin birlikte ortaya çıkmaları ve ortak satın almaları hakkında bize değerli bilgiler sağlayacaktır. Bu tür bilgiler, promosyon desteği, envanter kontrolü ve satışlar arası kampanyalar gibi pazarlama faaliyetleriyle ilgili Covid-19 pandemisinde verilebilecek satış kararları için bir temel olarak kullanılabilir. Web 3.0 alt yapısının kullanılmasıyla ortaya çıkan verilerin uygun tekniklerle anlamlı hedef satış politikalarına dönüştürülmesi gösterilmektedir. Covid-19 pandemisi sürecinde tüketicilerin bilgisayar donanım ve performans parçalarına ihtiyaçları artmıştır. Bu dolaylı sonuç, insanların alışveriş alışkanlıklarının Bitcoin, Ethereum vb. kripto para işlemlerine yöneldiği belirlenmiştir.

**Anahtar Kelimeler:** *e-business, online alışveriş, semantik web, web 3.0*

## WEB 3.0 DATA COLLECTION TECHNIQUES AND ONLINE BASKET ANALYSIS CHANGED BY THE PANDEMIC

### ABSTRACT

In recent years, the increasing popularity of the World Wide Web has led to growth with the rapid increase in the number of pages available on the web. A large number of web pages make it increasingly difficult for users to find the necessary information. When searching for specific information on the web, too many irrelevant search results waste time, and relevant materials can be missed. Existing web applications only provide web pages in HTML format whose content is represented in its own language, and therefore the semantics of the content is not accessible to machines. Web 3.0 was developed to enable machines to support the user in solving information problems. In this thesis, existing web 3.0 tools and data collection methods that make the semantics of web pages workable are proposed.

Companies today are rich in the use of big data but insufficient in the use and monetization of the information derived from that data. Big data is seen as a valuable resource, and while the concept of data mining is still new and evolving, companies in various industries rely on it to make strategic decisions. Facts that may go unnoticed can be uncovered by techniques that sift through the stored information. Market basket analysis is a very useful technique for identifying concurrent items in consumer market baskets. Such information can be used as a basis for decisions regarding marketing activities such as promotional support, inventory control, and cross-selling campaigns.

It is aimed to show with findings how different products in the product range of a shopping site are related to each other and how these relationships can be benefited from marketing activities during the Covid-19 pandemic period. Mining attribution rules from transactional data will provide us with valuable information about the co-occurrence and joint purchases of products. Such information can be used as a basis for sales decisions that may be made during the Covid-19 pandemic regarding marketing activities such as promotional support, inventory control, and cross-sales campaigns. It is shown that the data generated by the use of Web 3.0 infrastructure are transformed into meaningful target sales policies with appropriate techniques. During the Covid-19 pandemic, consumers' needs for computer hardware and performance parts have increased. This indirect result is that people's shopping habits change to Bitcoin, Ethereum, etc. may indicate that he is heading towards crypto money transactions.

**Keywords:** *e-business, online shopping, semantic web, web 3.0*

## 1. GİRİŞ

İçinde bulunduğumuz tüketim çağında, perakende firmalarının pazarda rekabet avantajı elde edebilmeleri için müşterilerin Covid-19 pandemi döneminde online alışveriş davranışlarını incelemek son derece önemli duruma gelmiştir. Perakende şirketleri için en büyük zorluklardan biri, büyük miktarda satış verisine sahipken önemli bilgilerin nasıl çıkarılacağını ve bu verileri nasıl işleyeceğini bilmektir. Satış verileri, şirketin ürün ilişkilerini keşfetmesi ve pazarlama stratejisini geliştirmesi gerektiğinde yararlı olan müşterilerin alışveriş davranışlarını anlamak için kullanılabilir.

Pazar sepeti analizi, ürünler arasındaki ilişkileri keşfetmek için kullanılan en önemli tekniklerden biridir [1]. Satış işlemleri arasında satın alınan ürünlerle satın alma çakışmasını basitçe analiz eder ve neyin ne ile satın alındığını açıklar. Pazar sepeti analizi, satış işlemlerinde birlikte oluşan ürün kombinasyonlarını arayarak çalışır. Bu kombinasyonlardan en yüksek getiriye sahip olanlar bulunmalıdır. Bulunan bu en yüksek kombinasyonlar bundan sonraki satış stratejisinin en önemli anahtarları olacaktır.

Web 3.0 teknolojileri altyapısı ile verilerini toplayan bir elektronik alışveriş firmasının, online alışverişlerden elde ettiği verileri kullanarak satış ve pazarlama konusunda aktif bir sistem kurma çalışması yapılacaktır.

### 1.1 Araştırma Problemi

İçinde yaşadığımız teknolojik çağ, şirketlerin muazzam miktarda veri toplamasını mümkün kılmıştır. Veri madenciliği, dünya çapında birçok işletme için giderek daha yaygın hale gelmektedir. Günlük olarak toplanan büyük miktarda veri, her işletmenin farklı yönleriyle ilgili faydalı bilgileri yakalar. Oldukça ayrılmış düzeyde veri toplanması, bilginin çıkarılması için bir ham madde olarak görülmektedir. Elde edilen verilen işlenmemesi, bir şablona alınmaması veya kullanılmadan sadece veri tutulması büyük bir problem olarak görülmektedir. Bazı gerçek veriler, doğrudan ayrıştırılmış



verilerden ortaya çıkarılabilsede genellikle gizli kurallar ve kalıplar belirleyerek bunları işleme almak amaçlanmaktadır. Veri madenciliği yoluyla önemsiz olmayan iç görüler oluşturulabilir. Veri madenciliği, verilerin bilinmeyen yönlerini ortaya çıkaran çeşitli istatistiksel analizleri içerir. Madencilik araçları, birçok işletmede önemli bilgileri ortaya çıkarmak ve dolayısıyla yöneticilere karmaşık problemler için çözümler sağlamak için faydalı bulunmuştur.

Pek çok ürün çevrimiçi olarak daha uygun bir fiyata bulunabilmektedir ve alıcılar için kapınıza teslim edilen ürünü almak çok uygun olabilmektedir. Son yıllarda pazar sepetlerini analiz etmek perakendeciler için oldukça çekici hale gelmiştir. İlerleyen teknoloji, müşterileri ve satın aldıkları ürünler hakkında bilgi toplamalarını mümkün kılmıştır. Elektronik satışların online hale getirilmesi, pazar sepeti analizinde işlem verilerinin kullanımını ve uygulanmasını arttırmıştır. Perakende işinde bu tür bilgileri analiz etmek, satın alma davranışını anlamak için oldukça faydalıdır. Veri Madenciliği satın alma kalıplarını, perakendecilerin promosyonları düzenlemesine, ayarları depolamasına ve müşterilere daha iyi hizmet etmesine olanak tanır.

Satın alma kurallarını belirlemek her başarılı işletme için çok önemlidir. İşlem verileri, ortak satın almalarla ilgili yararlı bilgileri araştırmak ve promosyon ile reklamları buna göre ayarlamak için kullanılır. Pazar sepeti analizi yapılması, bir dizi ürün arasındaki ilişkinin veya korelasyonların araştırılması önemlidir ve perakende şirketlerinin pazarda rekabet avantajı elde etmesi için faydalıdır. Covid-19 pandemisi sebebiyle online alışverişlerde müşterilerin satın alma davranışlarını inceleyerek farklı satış opsiyonları sunulabilmektedir. Pazar sepet analizi yaparak müşterilerin diğer ürünlere olan talebi de ölçülebilmektedir.

## **1.2 Araştırmanın Amacı**

Web 3.0 ortamları, veri kaynaklarının, uygulamaların ve araçların entegrasyonu için güçlü bir potansiyel sağlamıştır. Bu kadar yaygın ve dinamik bir alanda, web içeriğine erişmek ve yönetmek için mevcut teknikler aslında kullanıcı ihtiyaçlarını karşılamak için yetersiz kalmaktadır ve web kaynaklarının kullanılabilirliğini geliştirmek için daha fazla bilgi bulma, birleştirme ve paylaşma yollarına ihtiyaç duyulmaktadır.

Gelecekte elde edilecek veriler ve web madenciliği yöntemleri için birkaç önemli zorluk ortaya çıkmaktadır. Bu zorluklar, ek açıklamaları bulmak ve etiketler gibi zayıf

yapılandırılmış bilgilerin analizinden, web kaynaklarının sunduğu bilgileri arama ve entegre etmede kullanıcıları destekleyen akıllı yöntemlerin sağlanmasına kadar birçok alana yayılmaktadır. Bu zorlukların en temel amacı bilgi erişimini arttırmakla sınırlı değildir. Aynı zamanda bir veri setinin diğer bilgi kaynaklarıyla birlikte kullanıldığında edindiği zenginleştirilmiş anlambilimden yararlanmayı da içerir. Semantik web, doğal dilin temel yapılarının araştırılması, makine öğrenmesi, öneri ajanlarının tespiti ve yapay zekâ gibi farklı teknolojilerin sinerjisi bu açıdan özellikle verimli olabilmektedir.

Büyük veri ve nesnelerin interneti çağında, heterojen kaynaklar tarafından üretilen büyük miktarda bilgiyi elde etmek için daha fazla çalışılmaktadır. Her birey, çeşitli mobil cihazlar ve akıllı nesnelerle sensör ağları, bulut hizmetleri ve konumlandırma hizmetleriyle etkileşime girdiğinde dijital izler bırakır. İnsanlar, nesneler ve etkileşimleri hakkında uygulanabilir bilgiler elde etmek için bu tür verileri araştırmak için uygun yaklaşımların tasarımına artan bir ilgi gösterilmektedir. Daha genel olarak, toplanan verilerin boyutsallığı ve karmaşıklığı, hemen hemen tüm uygulama alanlarında hızlı bir şekilde artmakta ve yenilikçi veri analizi yaklaşımlarına ihtiyaç duyulmaktadır.

Araştırmanın temel amacı, Covid-19 pandemisi sebebiyle online alışveriş alanında müşterilerin satın alma davranışlarını inceleyerek bir teknoloji firmasının farklı ürünlerinin birbirleriyle nasıl ilişki kurduğunu ve bu ilişkilerden pazarlama faaliyetleriyle nasıl yararlanılacağını görmektir. İşlemsel verilerden madencilik ilişkilendirme kuralları, bize ürünlerin birlikte ortaya çıkmaları ve ortak satın alımları hakkında değerli bilgiler sağlayacaktır. Alışveriş yapanların bir kısmı bir alışveriş gezisinde meraktan veya can sıkıntısından tek bir ürün alırken, bazıları ise verimlilik nedeniyle birden fazla ürün satın alabilir. Sepette bulunan birden fazla ürün araştırmamızda incelenecek temel konulardan birisidir.

Anlamsal web'in kullanım amacı, veri madenciliği, sorgulama dilleri, semantik analiz, veri görselleştirme, grafik veri tabanları ile verilerin analizi ve kullanımıyla ilgili diğer alanların katkıları da dahil olmak üzere çok geniş bir alanda veri araması yapılmaktadır.

Web 3.0 teknolojileri öncesinde çoğu bilgi zayıf biçimde yapılandırılmış bir formda bulunmaktaydı. Örneğin metin, ses ve videoda mevcuttur. Bilgi yönetimi açısından mevcut teknoloji aşağıda belirtilen alanlardaki kısıtlanmıştır.

- **Bilgi arama:** Şirketler genellikle ana hatlarıyla belirtilen sınırlamalara dayalı olarak, anahtar kelime tabanlı arama motorlarına dayanır.
- **Bilgi çıkarma:** İlgili bilgiler için alınan dokümanlara göz atmak için insani zaman ve çaba gerekmektedir. Mevcut akıllı "Agents" bu görevi tatmin edici bir şekilde gerçekleştirememektedir.
- **Bilgilerin korunması:** Şu anda terminolojideki tutarsızlıklar ve modası geçmiş bilgilerin kaldırılmaması gibi sorunlar var.
- **Bilginin açığa çıkarılması:** Kurumsal veri tabanlarında örtük olarak var olan yeni bilgiler, veri madenciliği kullanılarak çıkarılmaktadır. Ancak, bu görev dağıtılmış, zayıf yapılandırılmış belgeler için hala zordur.
- **Bilgileri görüntüleme:** Çoğu zaman bilgilere erişimin belirli çalışan gruplara kısıtlanması arzu edilir. Belirli bilgileri gizleyen "Views", veri tabanları alanında bilinir, ancak intranet veya web üzerinden gerçekleştirilmesi zordur [2].

### 1.3 Varsayımlar

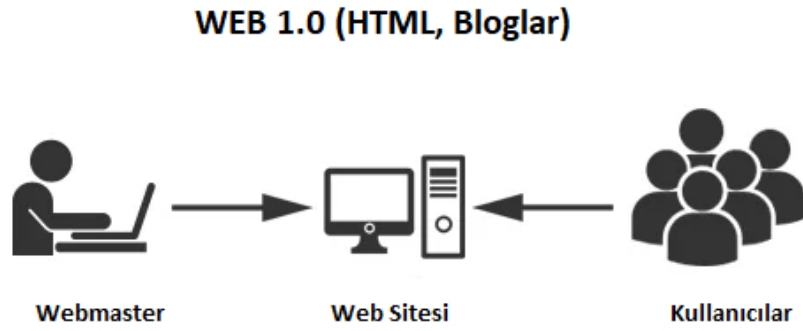
Bu çalışmada, Web 3.0 teknolojileri veri toplama teknikleri ve pandemi ile değişen online sepet analizi Apriori algoritması kullanılarak alışkanlıkların değişimi açıklanması amaçlanmıştır.

## 2. WEB TEKNOLOJİSİNİN TARİHSEL GELİŞİMİ

### 2.1 Web 1.0 Yapısı

Web 1.0, basit statik web siteleri ile karakterize edilen World Wide Web'i geliştirmenin ilk aşamasını ifade eden terimdir.

Web 2.0 terimi 1999 yılında Darcy DiNucci tarafından icat edilene kadar Web 1.0 popülerliğini devam ettirdi. Bu süre zarfında, web büyük bir dönüşüm geçiriyordu. 1990'larda geliştirilmiş olan çoğu web sitesi statik HTML sayfalarıyla ve HTML işaretlemesine gömülü birkaç basit yapıyla oluşturulmuştur. 1990'ların sonuna doğru ve 2000'lerin başında olan gelişim ile etkileşimli web sitesi özellikleri, bir web tarayıcısında neler yapılabileceğini yeniden tanımladı. Bu gelişmeler, web geliştirme dünyasında önemli bir evrim noktası oldu [3].



**Şekil 1:** Web 1.0 Kullanıcı ile Makine İlişkisi [4]

Web 1.0 kullanıcı ile web sitesinin gelişimi arasında yer alan grafik Şekil 1'de yer almaktadır. Web 1.0 olarak bilinen dönemde, web siteleri aşağıdaki tipik Web 1.0 özelliklere sahiptir:

**Statik sayfalar:** Sayfalar, web sitesi ziyaretçisinin davranışına göre değişen etkileşimli özellikler sunmamaktadır. Bu noktada web siteleri büyük ölçüde bilgilendiricidir.

**Dosyalarda depolanan web sitesi içeriği:** Hemen her modern web sitesi, web sitesi içeriğinin çoğunu depolamak için bir veri tabanı kullanır. Web 1.0 sırasında durum böyle değildi ve çoğu web sitesi içeriği ayrı bir veri tabanında değil doğrudan web sitesi dosyalarında saklanmıştır.

**İçerik ve düzenin birleşimi:** Günümüzde iyi bir web tasarım uygulaması, web sayfası işaretlemesi ile stilin ayrılmasını zorunlu kılmaktadır. Hemen hemen her modern web sitesi, web sayfalarının görünümünü ve düzenini belirlemek için harici stil sayfalarından yararlanır. Web 1.0 sırasında çoğu stil, genellikle tablo gibi HTML öğelerini yanlış kullanarak sayfa işaretlemesinin içine yerleştirilmiştir.

**Özel HTML etiketleri:** Web 1.0 sırasında tarayıcılar, özel etiketleri destekleyerek, bu etiketleri kullanan web siteleri ve site ziyaretçilerini, istenmeyen tarayıcılar kullanarak önemli uyumsuzluk sorunları yaratarak göze çarpmaya başlamıştır [3].

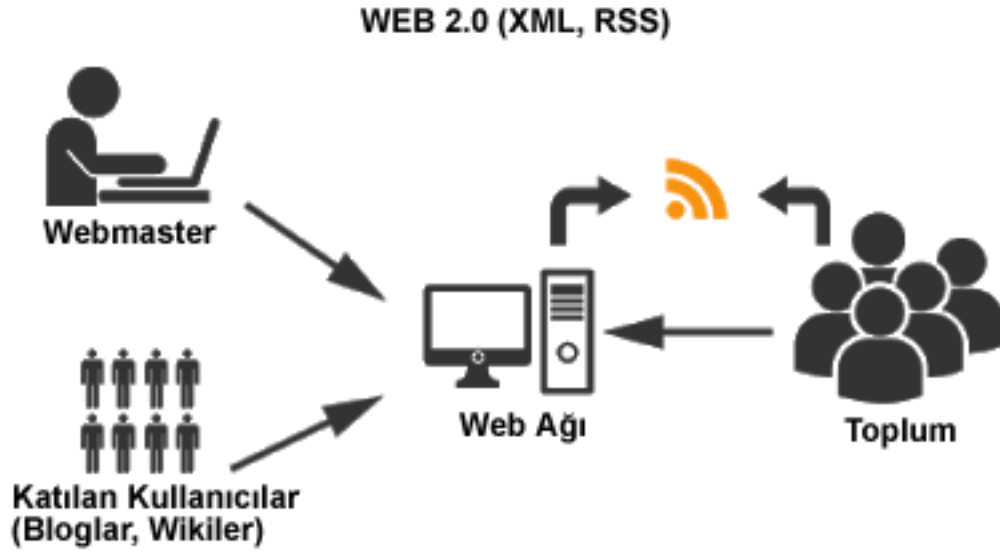
Sunucular yükseltirken, ortalama bağlantı hızı arttıkça Web 1.0'dan 2.0'a geçiş gerçekleşti ve geliştiriciler yeni beceriler ve teknikler öğrenmiştir. Bugün hala web'in bazı yerlerinde Web 1.0'ın izleri bulunmaktadır.

## 2.2 Web 2.0 Yapısı

Web 2.0 terimi 2004 yılında yeni teknolojiler ve ağlarla ilgili yayınlarda uzmanlaşmış bir Amerikan yayıncısı olan O'Reilly Media'dan Tim O'Reilly ve Dale Dougherty tarafından oluşturulmuştur. Web 2.0 terimi, örneğin bloglar, Google, Skype, Facebook, Flickr, Youtube ve Wiki gibi ikinci nesil Internet servislerini de ifade etmektedir. Web 2.0, sosyal ağ siteleri, Wiki ve kullanıcılar arasında iş birliğini ve paylaşımı vurgulayan iletişim araçları gibi kullanımı kolaydır [5]. Web 2.0 yalnızca kullanıcılar arasında içerik paylaşımına, iş birliğine ve iletişime izin vermekle kalmaz, aynı zamanda bilgi ve iletişim teknolojileri konusunda az deneyim sahibi olanların bile içerik üretmesine izin verir. Web 2.0'in öğrenme ortamlarında kullanılması, katılan tüm katılımcıların öğrenme sürecine aktif olarak katılmalarını sağlar, onlara içerik

üretme ve önerme, tartışmaları teşvik etme ve genel olarak gerçek öğrenme toplulukları oluşturma imkânı sunar.

Web 2.0, web'in ilk günleriyle karşılaştırıldığında daha yüksek kullanıcı etkileşimi ve iş birliği sağlayan, daha yaygın ağ bağlantısı ve gelişmiş iletişim kanalları ile karakterize edilen, mevcut kullandığımız çevrimiçi teknolojinin halidir.



**Şekil 2:** Web 2.0 Kullanıcı, Makine ve Toplum İlişkisi [4]

Kullanıcılar, topluluk, katılan kullanıcılar ve makine arasında yer alan ilişki Şekil 2’de yer almaktadır. Web 2.0 ve geleneksel World Wide Web arasındaki en önemli farklılıklardan biri, internet kullanıcıları, içerik sağlayıcılar ve işletmeler arasında daha fazla iş birliği yapmaktadır. Başlangıçta, veriler web sitelerinde yayınlanmakta ve kullanıcılar içeriği basit bir şekilde görüntülemekte veya indirmekteydi. Giderek artan bir şekilde, kullanıcılar web içeriğinin doğası ve kapsamı üzerinde daha fazla girişe sahip oldular ve bazı durumlarda bunun üzerinde gerçek zamanlı kontrol uyguladılar.

Web 2.0'ın sosyal doğası, kendisiyle orijinal statik web arasındaki önemli bir farktır. Web siteleri gittikçe artan bir şekilde topluluğa dayalı girdi, etkileşim, içerik paylaşımı ve iş birliğini mümkün kılmaktadır. Sosyal medya sitelerinin ve uygulamalarının türleri arasında forumlar, mikroblog, sosyal ağ, sosyal yer imi ve wikiler bulunur [5].

## Web 2.0'ın Özellikleri

- **Ziyaretçilerin web sayfalarında değişiklik yapma yeteneği:** Amazon, ziyaretçilerin ürün incelemeleri yayınlamasına olanak tanır. Bir çevrimiçi form kullanarak, bir ziyaretçi gelecek ziyaretçilerin okuyabileceği Amazon sayfalarına bilgi ekleyebilir.
- **İnsanları diğer kullanıcılara bağlamak için web sayfalarını kullanma:** Facebook ve MySpace gibi sosyal ağ siteleri kısmen popülerdir çünkü kullanıcıların birbirlerini bulmasını ve iletişim halinde kalmasını kolaylaştırır [5].
- **İçerik paylaşmanın hızlı ve etkili yolları:** YouTube en iyi örneklerden bir tanesidir. Bir YouTube üyesi bir video oluşturabilir ve başkalarının bir saatten az bir sürede izlemesi için siteye yükleyebilir.

### 2.3 Web 1.0 ve Web 2.0 Arasındaki Farklar

Web 1.0 ve Web 2.0 arasında birçok önemli farklılık bulunmaktadır. Bunlardan bazıları, Web 1.0'ın statik olması ve kullanıcıların yalnızca siteleri ziyaret etmesine izin vermesidir. Web 2.0, kullanıcıların video yükleyerek, blog yazarak veya Wikipedia gibi siteleri düzenleyerek sitelere erişmelerini ve onlarla etkileşime geçmelerini sağlamıştır. Ayrıca Web 2.0 çok daha güvenilir bir bağlantıya sahiptir, daha fazla trafik çekebilmektedir, daha fazla altyapıya sahiptir, XML ile gelişmiş tarayıcıları desteklemektedir ve kullanıcılarının tüm URL'yi girmeden önceki bir sayfaya dönmelerini sağlamaktadır. Aralarındaki diğer önemli farklar ise aşağıda listelenmiştir.

**Bağlanabilirlik:** Web 1.0 ile karşılaştırıldığında, Web 2.0'daki bağlantı daha güvenilirdir ve kullanıcılar için daha fazla önem taşımaktadır [5].

**Kullanım:** Web 1.0 yardımı ile pek çok aktiviteye girilemediğinden, Web 2.0 kullanıcıların görüşlerini paylaşmalarını sağlayan ek özelliklerle birlikte gelir. Web 2.0, Web 1.0'dan daha fazla ticari kullanıma sahiptir [5].

**Gelişmiş Tarayıcılar:** Başlangıçta Web 1.0 tamamen HTML ile ilgiliydi, ancak bugün Web 2.0 daha gelişmiş ve hızlı olan XML'e yükseltilmiştir [5].

**Taşınmazlık:** Web 1.0 web siteleri statik olduğu için, bir kişi bir kere istediği sayfaya geri dönemez. Bu da çok fazla sıkıntıya neden olmaktadır. Web 2.0, kullanıcıya yalnızca bir tıklamayla istediğiniz herhangi bir sayfaya ulaşma imkânı sunar [5].

**Destek:** Web 1.0 ile karşılaştırıldığında, Web 2.0 tabanlı uygulamaları desteklemek için geniş bir altyapı vardır [5].

**Dinamik Sistem Araçları:** Web 2.0 tabanlı uygulamalarda kullanılan yazılım AJAX'dır. J2EE olarak adlandırılan Web 1.0'dan daha güçlüdür [5].

**Etkileşimler:** Web 1.0 sitelerinde etkileşimler mümkün değildi. Kullanıcılar yalnızca bu siteleri ziyaret edebilmekteydi. Bugün, blog yazmak ve insanlarla etkileşimde bulunmak önemli bir pazarlama aracı olduğunda, Wikipedia gibi bazı web siteleri kullanıcılara makaleler hakkında yorum yapma veya bunları değiştirme olanağı sunar. Web 1.0 ve Web 2.0 sınıf etkinlikleri gibidir; burada Web 1.0 ders ve Web 2.0 tartışmalar ve görüşlere benzetilmektedir [5].

## 2.4 Web 3.0

Web 3.0, kullanıcılara yardımcı olmak için akıllı programlar çalıştırabilen AI sistemlerinin gücü aracılığıyla interneti daha akıllı hale getirecek veya bilgiyi insan benzeri zekâ ile işleyecek web evriminin bir sonraki aşamasıdır.

Web 3.0'ın dört temel özelliği aşağıdaki gibidir.

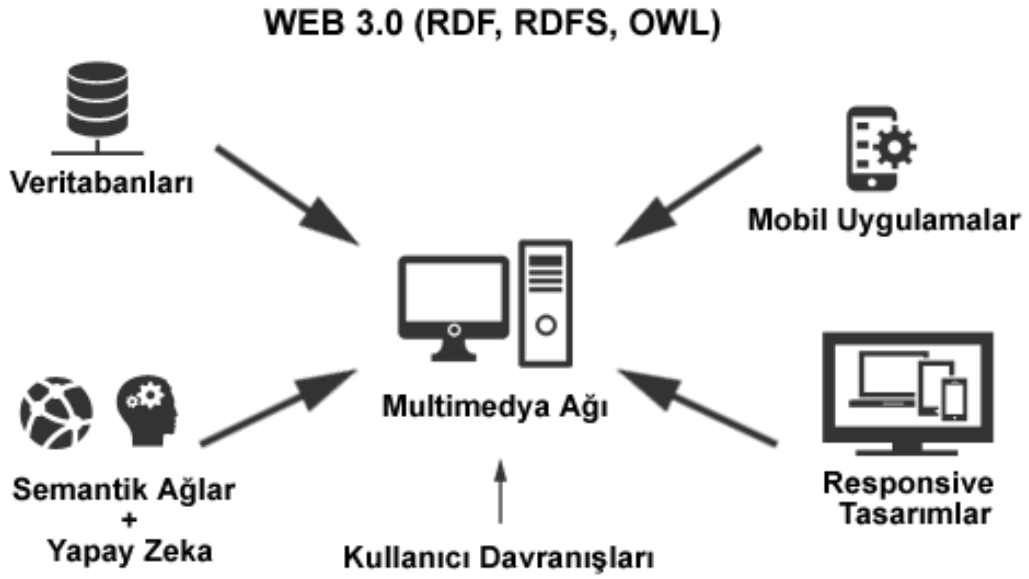
- Erişebilirlik
- Anlamsal ağ
- Yapay zekâ
- 3D Grafikler

Bir Web 3.0 uygulaması için ortak gereksinim, büyük ölçekli bilgileri anlama ve bunu kullanıcılar için gerçek bilgilere ve faydalı uygulamalara dönüştürme yeteneğidir. Bununla birlikte, bu uygulamalar hala erken aşamadır, bu da iyileştirme için çok fazla alana sahip oldukları ve Web 3.0 uygulamalarının potansiyel olarak nasıl çalışabileceğinden çok uzak oldukları anlamına gelir.

İnternet 3.0 uygulamalarına dönüştürdükleri ürünler geliştiren veya sahip olan şirketlerden bazıları Amazon, Apple ve Google'dır. Web 3.0 teknolojilerini kullanan



iki uygulama örneđi, Siri ve Wolfram Alpha'dır [6]. Multimedya ađının merkezde olduđu Web 3.0 tasarım iliřkisi Őekil 3'te yer almaktadır.



**Őekil 3:** Web 3.0 Tasarım İliřkisi [7]

Web 3.0, daha kiřisel ve özelleřtirilmiř bir tarama deneyimi, daha akıllı ve daha insan benzeri bir arama asistanı ve daha adil bir web oluřturmaya yardımcı olacađı beklenmektedir. Her bir kullanıcıyı kendi verileri üzerinden yetkilendirerek ve sistemde çevrimiçi olmasını sađlayarak gelecek sayısız yenilik sayesinde daha zengin bir genel deneyim yaratacaktır.

Web 3.0 kaçınılmaz olarak geldiđinde akıllı cihazların davranıř kalıplarımızı nasıl deđiřtirdiđini düşünmek ne kadar zor olsa da internet günlük yařamlarımıza katlanarak daha fazla entegre olacaktır.

### **2.5 Web 3.0 Geçiř Süreci**

Web 3.0'a geçiř, dijital dünyanın iřleri farklı Őekilde yapması ve kullanıcıların ileriye dönük olarak verilerinin nasıl toplandıđı, kullanıldıđı ve paylařıldıđı konusunda bir fırsattır. Birkaç güçlü teknoloji řirketine ait olmak yerine, kiřisel veriler hem teknolojik hem de yasal olarak daha iyi korunacak ve mülkiyeti hak sahiplerine, yani bireysel kullanıcılara geri dönecektir [8].

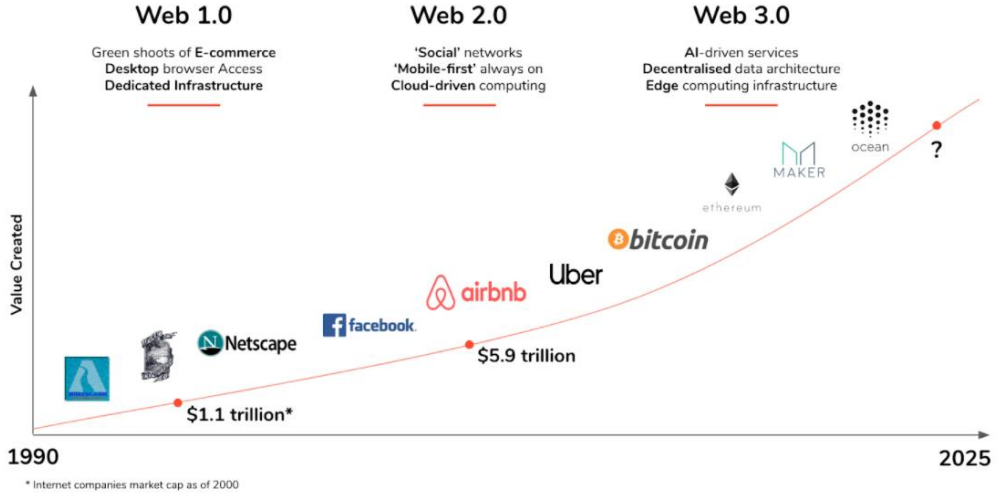
Başlangıçta, web 3.0'ın teknolojik engeller nedeniyle asla gerçekleşmeyen semantik web kavramını hayata geçireceği yaygın olarak bekleniyordu. World Wide Web'in kurucusu Tim Berners-Lee'ye göre web 3.0, neyin, ne zaman ve kim tarafından paylaşılacağına başkanlık eden tek bir otoritenin olmadığı orijinal web kavramına dönüşü ifade eder [8].

Gelişmiş veri egemenliğine ek olarak, web 3.0'ın bugün interneti rahatsız eden diğer bazı büyük sorunları çözmeye adımlar atması bekleniyor: çevrimiçi güvenlik ve güvenlik açıklarının kapatılması, bilgi kaynaklarının doğrulanması, ağ tarafsızlığı. 5G ağları, blok zinciri, IoT ve bağlı cihazlar, AI ve yenilikçi web protokolleri gibi teknolojiler, bu zorlukların üstesinden gelmek için kullanılabilir ve sonunda verilerimizin, zamanımızın ve kullanım davranışlarımızın sahipliğini geri talep etmemizi sağlar [8].

Web 3.0 ile gelen özellikler sayesinde bilgisayarlar, insanlar gibi bilgileri yorumlayabilmektedir ve insanların ihtiyaçlarına göre tasarlanmış faydalı içerikleri yapay zekâ ile üretip dağıtabilmektedir [9]. Bilgi ve verilerin bilgisayar sistemlerinin onları anlayacağı şekilde depolandığı bir teknolojidir. Yeni nesil bir yapay zekâyı doğurabilecek güce sahiptir. Bilgisayar alanında çalışma yapan bilim insanları ve internet uzmanları, web alanındaki bu yeni paradigmanın, insanların çevrimiçi yaşamlarını daha kolay ve daha sezgisel hale getireceğini düşünmektedir. Bunun nedeni de daha iyi arama işlevlerine sahip akıllı uygulamalar, kullanıcılara aradıklarını tam olarak vermektedir [9].

Web 3.0 dosyalarınızı, uygulamalarınızı yönetebileceğiniz ve her yere götürebileceğiniz ve verilerinize herhangi bir cihazdan erişebileceğiniz bir tarayıcı gibi bir web işletim sistemi olabilecektir. Nesnelerin interneti (IoT) yavaş yavaş daha dijital merkezli bir bağlı bir toplum yarattığı gibi, Web 3.0 web üzerindeki kalıcı karmaşıklığı ortadan kaldırmaktadır ve daha fazla sayıda insan için daha erişilebilir hale getirmektedir.

# The Evolution of the Web



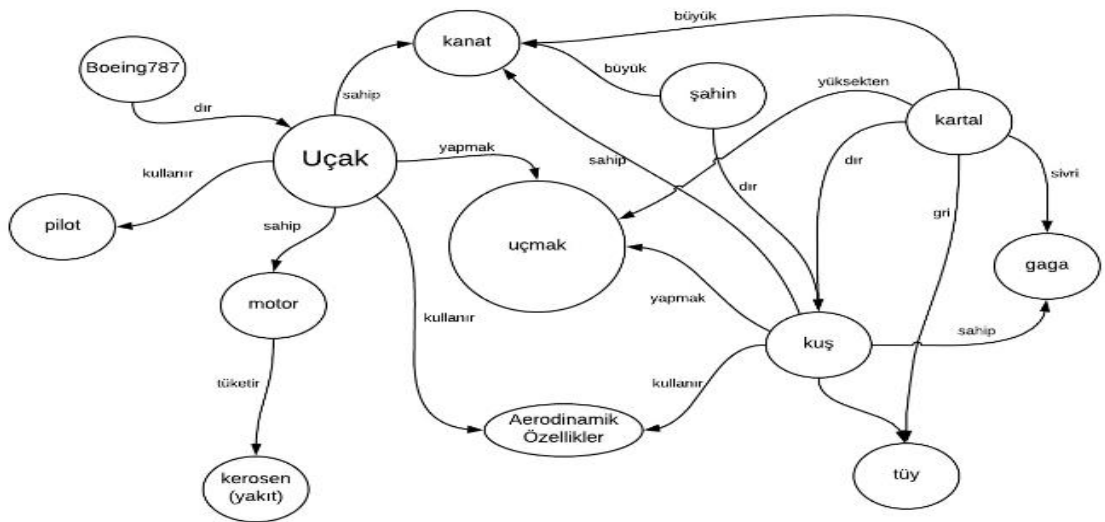
Şekil 4: Web'in Gelişim Süreci [10].

Web 3.0 gelişim sürecinin Web 1.0 ve Web 2.0 dönemiyle birlikte gösterimi Şekil 4'te yer almaktadır.

### 3. WEB 3.0 (SEMANTİK WEB)

Semantik web, insanların yerine makineler tarafından kolayca işlenebilecekleri şekilde ilişkilendirilmiş bir veri ağıdır. Mevcut kullanılan World Wide Web'in genişletilmiş bir versiyonu olarak algılanabilir ve küresel bağlantılı bir veri tabanı şeklinde etkin bir veri sunum aracıdır. Semantik içeriğin web sayfalarına dahil edilmesini destekleyerek, semantik web şu anda mevcut yapılandırılmamış belgelerin web'in bir bilgi web'ine dönüştürülmesini hedefler [11].

Anlamsal web'in temel amacı, kullanıcıların daha az çabayla bilgi arama, keşfetme, paylaşma ve katılmalarını sağlamak için mevcut web'in gelişimini tetiklemektir. İnsanlar, çevrimiçi uçak bileti rezervasyonu yapmak, farklı bilgiler aramak, çevrimiçi sözlükleri kullanmak gibi birden fazla görevi yürütmek için web'i kullanabilirler. Buna rağmen, makineler, bu sayfaların hiçbirini insan müdahalesi olmadan yürütemez, çünkü web sayfaları makineler tarafından değil insanlar tarafından okunmaktadır. Semantik web, verilerin makineler tarafından hızlı bir şekilde yorumlanabileceği, web'de bulunan bilgileri keşfetme, harmanlama ve aksiyon alma ile ilgili sayısız görevleri yerine getirmelerine izin veren bir vizyon olarak düşünülebilir. Şekil 5'te yer alan anlamsal web örneği yer almaktadır.



Şekil 5: Anlamsal Web Örneği [12]

Web'de çok fazla veri bulunmaktadır, ancak bilgisayarlar bu verilerle ilgili herhangi bir karar verememektedir veya anlayamamaktadır. Semantik web yapısı, insanlar ve makinalar arasında sağlam bir iletişim kurabilmek ve bu ilişkiden anlam çıkarmak için verilerin kolayca işlenmesini hizmet etmektedir. Akıllı veri, bilgisayarın insanlar olarak daha fazla çıkarım yapabilmesi yönünden oldukça bağlantılı olan web bilgileri anlamına gelir.

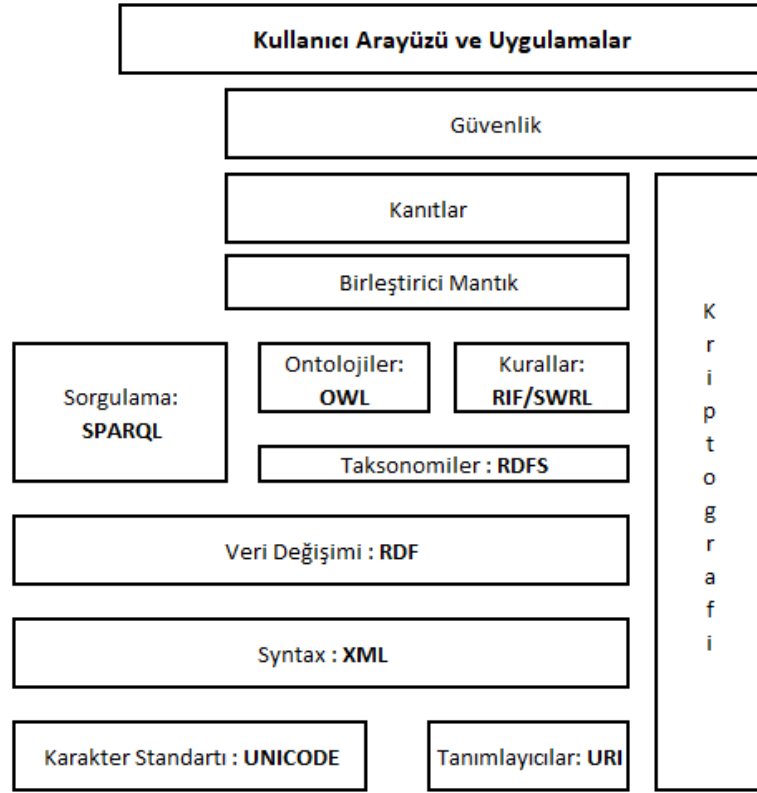
World Wide Web içerisinde birbirine bağlı bulunan birçok web sayfaları vardır. Kullanıcıların İnternetteki günlük aktiviteleri, web sayfalarının aranması, kullanıcıların en kısa sürede ilgili bilgilerin alınması amaçlanır. Web 2.0 tabanlı arama motorları, bilgi yönetimi krizine eğilimli olan makineler ve meta-veri arasında birlikte çalışabilirlikten yoksun olduğu için büyük dezavantajlara sahiptir [11].

Arama motorları optimizasyon yapabilmek için kullanıcılar tarafından verilen kelimeleri ayrıştırarak bazı algoritmaları kullanır. Semantik Web, bilgisayarların birbirleri arasında birlikte çalışabilmesi mantığına dayanmaktadır. İnsanlar gibi verileri kullanıcılar tarafından işlerler. Semantik webin en önemli amacı verileri bütünleştirmek için meta-veri eklemektir. Web kaynakları ve bilgi yönetimi arasındaki ilişkiyi belirlemek için çalışma yapmaktır [11].

Semantik Web, değişken veri alışverişlerinin yorumlanmasıyla ilgili yeni teknolojiler ve standartlar kullanan ve ayrıca bu verilerden faydalı bilgileri otomatik olarak çıkaran World Wide Web'in bir uzantısı olarak tanımlanmaktadır [11].

### **3.1 Semantik Web Mimarisi**

Kullanılan dillere ve çeşitlerine bağlı olarak birçok katman içermektedir. Semantik Web mimarisi Şekil 6'da gösterilmektedir.



**Şekil 6:** Semantik Web Teknolojisi Katman Yapısı [13].

Semantik Web’de bulunan ilk katman URI / IRI (Tekdüzen Kaynak Tanımlayıcısı / Uluslararası Kaynak Tanımlayıcısı), web tanımlarını ve belgelerini içeren için standart bir form yapısıdır. Standart Kaynak Bulucu (URL), web kaynaklarını yerelleştirmek için dahil edilen bir URI'den sonra gelir. IRI uluslararası kabul görmüş bir formdur. XML katmanı, semantik web'de kullanılan ortak sözdiziminin olmasını sağlayan XML yapısı ve XML şemasına içermektedir. XML, ortak bir dil kullanarak platformdur. Veri alışverişi için anahtar konumda bulunmaktadır [13].

Anlamsal ağdaki temel veri temsili Kaynak Tanım Çerçevesi (RDF) çalışması üç temel ögeye dayanır (konu-yüklem-nesne) ve aynı şekilde verilen verilere ait bir grafik oluşturur. XML makineler tarafından yapay zekâ kullanılarak anlaşılakta olan kelimelerin olduğu dokümanın yapısıdır. Ontology Web Language (OWL) açıklama mantığını kullanmaktadır. Semantik webe güçlü bir anlambilim sağlamaktadır. SPARQL, RDF ve OWL ontolojileri sorgu yapmak için kullanılmaktadır. SQL’

benzeyen bir yapıya sahiptir. Bilgi alışverişi sırasında güvenliği sağlamak için dijital imzalar kullanılmaktadır [13].

### **3.2 Semantik Web Katman Yapıları**

Şekil 2.1’de Semantik Web’i oluşturan katman yapıları bulunmaktadır. Bu mimariye Semantik Ağ Yığını denir. Katman yapısı üç gruba ayrılmıştır.

#### **3.2.1 Hiper Metin Web Teknolojileri**

Bunlar, köprü metni web alanında iyi bilinen alt katman teknolojileridir. Bu teknolojiler anlamsal web uygulamalarını uygulamak için kullanılır.

##### **1- URI**

Evrensel Kaynak Tanımlayıcısı (URI), anlamsal web kaynaklarını tanımlamak için kullanılır. Bu tanımlama, üst katmanlardaki kaynaklarla manipülasyon sağlamak için kullanılmaktadır [14].

##### **2- UNICODE**

Metnin çeşitli dillerde temsil edilmesine ve manipüle edilmesine yardımcı olarak insan dilleri ile anlamsal uygulamalar arasındaki boşluğun kapatılmasını sağlar [14].

##### **3- XML**

Genişletilmiş İşaretleme Dili (XML), semantik web belgelerini yapılandırılmış veri biçiminde oluşturmak için kullanılan işaretleme dilidir [14].

#### **3.2.2 Anlamsal Web Teknolojileri**

World Wide Web Consortium tarafından standartlaştırılmış ve anlamsal web uygulamaları oluşturmak için kullanılan orta katman teknolojilerine verilen addır. RIF/SWRL hariç tümü World Wide Web Consortium tarafından standartlaştırılmış bir yapıya sahiptir [14].

##### **1- RDF**

Kaynak Tanımlama Çerçevesi (RDF), verileri anlamlı bir şekilde ifade etmek için kullanılan çerçevedir. Verileri üçlü olarak tanımlayarak bilgiyi bir grafik şeklinde ifade etmek daha kolaydır [14].

## **2- RDFS**

RDF Şeması (RDFS), RDF belgenin uygun yapısını korumasını sağlayan kelime haznesini sağlamaktadır. Düzgün bir sınıf hiyerarşisini ve özelliklerini korumayı sağlamaktadır [14].

## **3- OWL**

Web Ontology Language (OWL), RDF yapısına daha fazla anlam, kısıtlama ve sınırlamalar eklemek için kullanılan dildir. RDF ifadelerinin anlamlarını ifade eder [14].

## **4- SPARQL**

SPARQL Protokolü (SPARQL), bir RDF sorgu dilidir. RDF tarafından temsil edilen veri tabanında sorgulama yapabilmek için kullanılmaktadır. Anlamsal uygulamalar ile bilgi almak için sorgulama yapılmaktadır [14].

### **3.2.3 Gerçekleşmemiş Anlamsal Web Teknolojileri**

Henüz standartlaşmamış veya tamamen anlamsal web uygulamaları oluşturmak için uygulanması gereken uygulamalar olan üst katman teknolojilerine verilen addır.

#### **1- RIF/SWRL**

Kural Değişim Biçimi/Semantik Web Kural Dili (RIF/SWRL), RDF verilerine kural eklemek için kullanılır. OWL tarafından doğrudan ifade edilemeyen bilgileri temsil etmeyi sağlar [14].

#### **2- Şifreleme**

Anlamsal web uygulamalarından gelen ifadelerin doğru kaynaklardan gelmesini sağlamak ve RDF belgelerinin dijital imzaları kullanılarak gerçekleştirilebilmesini sağlamak için kullanılmaktadır [14].

#### **3- Güven**

İfade desteği için güven, kurumlar güvenilir kaynaklardan gelir ve yeni bilgiler almak için resmi dile dayanır [14].



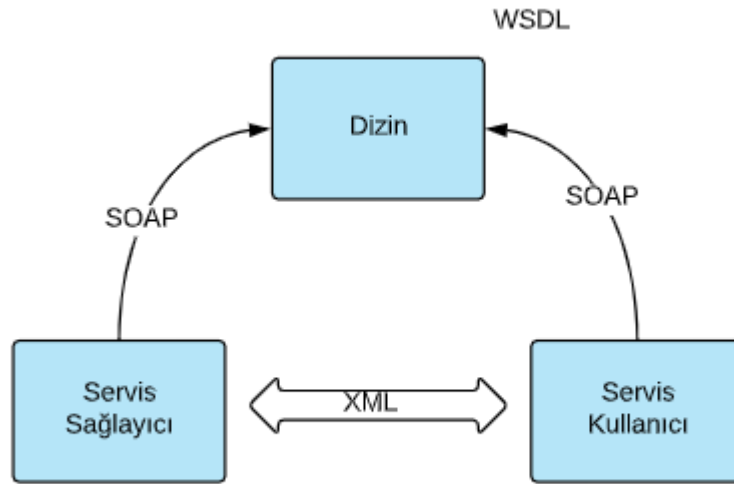
#### 4- Kullanıcı Ara yüzü

İnsanların semantik web uygulamalarını kullanmasını sağlayacak en üst katmandır [14].

### 3.3 Web Servisleri

İnternet üzerinden kendisini sorunsuz erişilebilen ve standart bir mesajlaşma sistemi kullanan herhangi bir yazılım parçalarına verilen addır. İki standart iletişim yapısı kullanılmaktadır: Web Servisleri Açıklama Dili (WSDL) ve JavaScript Nesne Notasyonu (JSON). WSDL, XML mesajları göndermek ve web servislerinin tanımlama ara yüzü amacıyla kullanılmaktadır. JSON, veri nesnelerini iletmek için kullanıcılar tarafından okunabilen metni kullanan açık bir standart formata sahiptir [3].

WSDL (Web Hizmeti Açıklama Dili), XML tabanlı bir tanımlama diline verilen addır. SOAP tabanlı bir web hizmetinin işlevselliğini açıklamak için kullanılmaktadır. WSDL dosyaları, SOAP tabanlı hizmetleri test etmenin merkezinde yer alır. SoapUI, test istekleri ve bazı hizmetler oluşturmak için WSDL dosyalarını kullanılmaktadır [3].



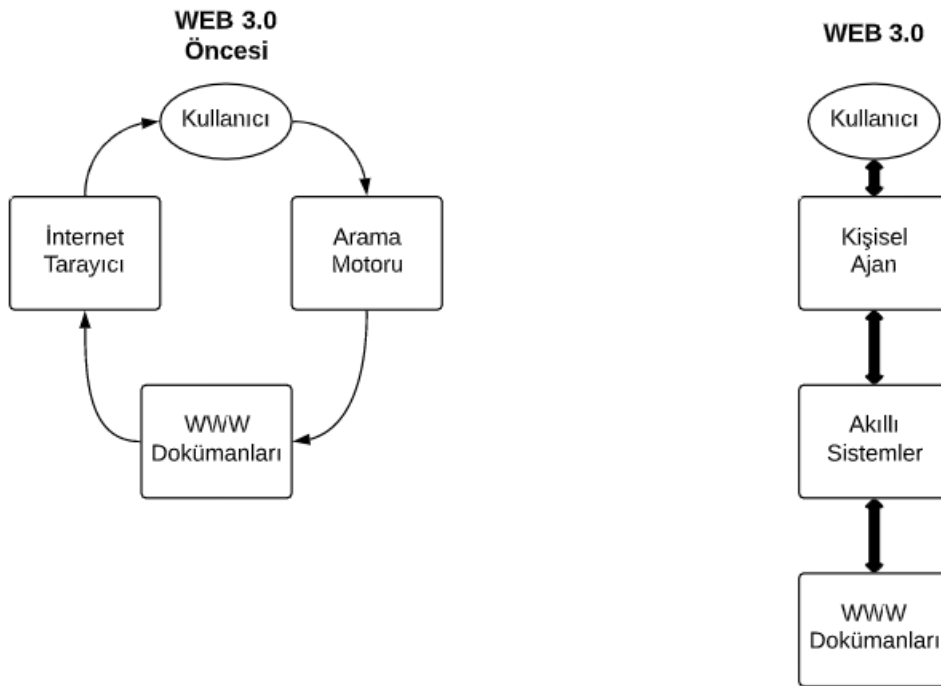
Şekil 7: WSDL Web Servis Yapısı [15]

WSDL web servis yapısı Şekil 7’de yer almaktadır. Optimizasyon sorunlarını çözmek için yazılımcılar tarafından geliştirilen ve internet üzerinden sunulan birçok web servisi vardır. Her web servisi bir problemi çözmek için oluşturulmuştur. Bir veya

daha fazla optimizasyon yöntemi web servis altyapısında kullanılmaktadır. Web servisleri iki sınıfa ayrılmaktadır: ticari amaçla yapılan ve kişisel amaç için kullanılan web servisleri. Ticari amaçla yapılan web hizmetleri, optimizasyon sorunlarını çözmek için profesyonel şirketler tarafından geliştirilmektedir. Kişisel amaç için kullanılan web hizmetleri doktora öğrencileri veya araştırma görevlileri tarafından çözümlenme yöntemlerini doğrulamak için geliştirilmesi amaçlanmıştır. Kişisel amaç için kullanılan web hizmetleri ücretsiz olarak sunulmaktadır. Örneğin, Araç navigasyon yönlendirme sistem sorunlarını bulmak ve çözümlenebilmek için internet üzerinde birçok web servisi kullanıcılar tarafından bulunabilmektedir [3].

### 3.4 Ajanlar

Ajanlar, bir görevin gerçekleştirilmesine katkıda bulunabilen bir yazılım nesnesidir. Semantik Web'deki kişisel bir ajan kullanıcıdan bazı görevler ve tercihler almaktadır. Web kaynaklarından bilgi isteyecek, diğer araçlarla iletişim kuracak, kullanıcı gereksinimleri ve tercihleri hakkındaki bilgileri karşılaştıracak, belirli seçenekleri seçecek ve kullanıcıya cevaplar verecektir [16]. Akıllı kişisel ajanların Web 3.0'da bulunduğu konum Şekil 8'de yer almaktadır.



Şekil 8: Akıllı Kişisel Ajanlar

Ajanlar, iletişim dilleri gibi daha ileri teknolojilere de ihtiyaç duyulacaktır. Ayrıca, gelişmiş uygulamalar için, ajanların inançlarını, arzularını ve niyetlerini resmi olarak temsil etmek ve kullanıcı modelleri oluşturmak ve sürdürmek faydalı olacaktır. Ancak, bu noktalar Semantik Web teknolojilerine esnek değildir.

- Alan bilgisine erişme,
- Görevi hakkında akıl yürütme,
- Anlamli mesajlar oluşturma,
- Bunları diğer ajanlara veya insanlara göndermek,
- Alınan mesajları yorumlama,
- Alan bilgisine ve toplanan bilgilere dayalı kararlar vermek,
- Kararlara göre anlamli bir şekilde hareket etmek,
- Çoklu Ajan Sistemi, sistem görevlerini yerine getirmek için birbirleriyle rekabet eden veya iş birliği yapan ajanlardan oluşan bir yazılım sistemidir [16].

### 3.5 Semantik Teknolojiler ile Uygulamaya Entegrasyon Süreci

#### 3.5.1 Anlamsal Veri Entegrasyonu

Semantik veri entegrasyonu, semantik teknoloji kullanılarak farklı kaynaklardan gelen verileri birleştirme ve anlamli bilgilere dönüştürmek için kullanılmaktadır. Anlamsal veri entegrasyonu Şekil 9'da gösterilmiştir.



Şekil 9: Veri Entegrasyon Yapısı [17]

### 3.5.2 Heterojen Veri Kümelerini Entegre Etme

Firmalar kapasite olarak büyüdükçe, verileri de daha büyük boyutlara ulaşmaktadır. Doğru veri yönetimi stratejisine sahip olmadan, bölüm içi veya uygulamaya özgü veri depoları hızla ortaya çıkar ve üretkenlik ile iş birliğini engellemektedir.

Semantik Veri Entegrasyonu, standart kurumsal uygulama entegrasyon çözümlerinin ötesine geçen bir çözüm sunar. Veri yayınlama ve değiştirme için standart bir model, Kaynak Tanımlama Çerçevesi (RDF) üzerine inşa edilmiş veri merkezli bir mimari kullanır [16].

Bu çerçevede, bir kuruluşun heterojen verileri yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış şekilde ifade edilir, depolanır ve erişilebilir. Veri yapısı, veri içerisindeki bağlantılar aracılığıyla ifade edildiğinden, veri tabanı tarafından dayatılan bir yapı ile sınırlandırılmaz ve verilerin evrimi ile güncelliği geçmez. Veri yapısında değişiklikler meydana geldiğinde, veri içindeki bağlantılardaki değişiklikler yoluyla veri tabanına yansıtılır.

Buna ek olarak, Semantik teknolojinin bel kemiği olan RDF, mevcut verilerden yeni olguların çıkarılmasının yanı sıra Bağlantılı Açık Veri (LOD) kaynaklarına erişerek mevcut bilginin zenginleştirilmesini sağlar [16].

Anlamsal veri entegrasyonu yaparken ana görevler şunlardır:

- Veri kümesinin istenen biçimini tanımlayan bir Uygulama Profili (RDF Şekli) oluşturulmalıdır.
- Mevcut ontolojilerin yeniden kullanılması ve gerektiğinde yeni ontolojilerin tasarlanması gerekmektedir.
- Etki alanındaki kullanılabilir bağlı açık veri kümelerinin tam olarak kullanılması planlanmalıdır.
- Basit, mantıklı ve sürdürülebilir bir URL stratejisi tasarlanmalıdır.
- Entegrasyonu gerçekleştirmek için mevcut dönüşüm ve ETL araçlarının çeşitliliğini kullanmak gerekir.
- Bir veri güncelleme stratejisi tasarlanmalı ve uygulanmalıdır [16].

### 3.6 Ontoloji

Ontoloji kelimesi, Yunanca ontos (varlık) ve logos (çalışma) kelimelerinden gelir ve temel kategoriler kadar varlık ve varoluş konusuna köküne sahiptir. Başka bir deyişle, Ontoloji terimi, "bazı alanlarda var olan veya olabilecek şeylerin kategorilerinin incelenmesi" için kullanılır [18].

Ontoloji, yapay zekâ ve bilgi temsili alanında da iyi bilinen bir kavramdır. Yazılım mühendisleri ontoloji ile bilgiyi, makinelerin akıl yürütebileceği, geçerli çıkarımlar yapacak şekilde temsil etmenin bir yolunu bulmak için tasarlamıştır. Bir ontoloji, kavramların ve bunların ilişkilerinin bir kelime dağarcığı ile tanımlandığı belirli bir alana göre bilginin temsilidir.

Ontoloji, "paylaşılan bir kavramsallaştırmanın resmi, açık bir özelliği" olarak tanımlanabilir. Kavramsallaştırma yapısı içerisinde, insanların buldukları dünyadaki nesnelere için nasıl düşündüklerini ifade etmek için kullanılan, genellikle belirli bir konuya bağlı olan soyut bir modele verilen addır. Soyut modeldeki kavram ve ilişkilere açık isimler verilmektedir. Tanımların verilmesi açık bir ifade anlamına gelir [19].

Bu konuda;

- Tanım terimi bir kurala aittir, kavram ya da ilişkinin anlamına bağlı bir isme aittir.
- Biçim, anlamın tanımını tanımlamak için bir kodlama dilinin kullanılması anlamına gelir. Tutarsız tanımların karmaşıklığa yol açmasından kaçınmak amaçlanır. Kullanılmakta olan diller mantık temelli dillerdir. Yapay zekâ topluluğundaki bilgiyi temsil etmek için uzun süre boyunca kullanılmaktadır.
- Formalite yapısı, doğal dilde bulunan ve diğer gösterimlerde yaygın olan belirsizliği önlemek amacıyla kullanılır. Bunlar aynı zamanda özellik anlamlarından yeni bilgiler elde etmek için otomatik çıkarım için kapıyı açar.
- Farklı uygulamalar ve topluluklar arasında paylaşımlar yapılır. Yeniden kullanılan bir ontolojinin asıl amacı bu paylaşımlarda bulunmaktır.
- Belirli bir konudaki ilgi alanlarına atıfta bulunan terimler sözlüğü domainidir.
- Terimler için bazı anlam özellikleri, ideal olarak bir mantık biçimine dayandırılmıştır [19].

Semantik Web yapısında kullanılan ontolojiler, kavramları ortak niteliklere ve özelliklere dayalı olarak sınıflara ayrılmaktadır. Bir sınıf yapısı bir dizi özellik tarafından tanımlanır ve bir nesnenin bir sınıfa ait olması amaçlanır. Temel koşul, sınıfla ilişkili tüm özelliklere sahip olmalıdır. Özellikler, çıkarım mekanizmalarının üyeliği otomatik olarak tanımlaması için gerekli ve yeterli olarak tanımlanabilir.

Semantik Web ontolojileri, makinelerin web'deki bilgileri yorumlamasını ve işlemlerini sağlayarak hem insanlar hem de bilgisayarlar tarafından anlaşılabilen ortak bir model sağlayarak, amaçlanan anlamlarına göre verileri paylaşmak, değiştirmek ve yeniden kullanmak için kullanılır.

Ontolojiler birden çok farklı nesne yapısında bulunabilir. Bunlar ontolojide kavramlar olarak gösterilebilir ve genellikle sınıfların ve alt sınıfların hiyerarşilerinde düzenlenir. Çoğunlukla her sınıf, özellikleri ve nitelikleri tanımlayan çeşitli özelliklerle ve ayrıca çeşitli kısıtlamalarla ilişkilidir. Bir ontoloji, bir bilgi tabanı oluşturan sınıfların bir dizi örneğini içerebilmektedir.

Semantik web vizyonunun ontolojilerin kullanım amacı, anlamsal bir düzeyde çoklu ve heterojen dijital içeriği entegre ederek anlamsal olarak birlikte çalışabilirliğini sağlamaktır.

### **3.6.1 Ontoloji Web Dili (OWL)**

RDFS'nin kelime dağarcığı yetersizliği ve taksonomiler gibi ontolojileri tanımlamak için basit bir dil olması amaçlanmıştır. Ancak birçok durumda Semantik Web'in taleplerini ele almak için daha fazla ifadeye ihtiyaç vardır. Web Ontoloji Dili, sınırlamalarının üstesinden gelmek için RDFS'yi genişleten bir ontoloji dilidir. OWL, Semantik Web'de ontolojiler yayımlamak ve paylaşmak için kullanılan bir standartta verilen addır [20].

OWL, RDF ve RDFS'yi genişleten birçok yeni dilin temelini oluşturur. OWL, birleşim, kesişim ve tamamlayıcı gibi küme operatörlerini kullanarak sınıfları diğer sınıfların bir kombinasyonu olarak tanımlamaya izin verir. OWL'de iki sınıfın ayrık veya aynı olduğunu belirtmek mümkündür.

Ontoloji web diline ait bir kod örneği Şekil 10'da yer almaktadır.

```

<owl:Class rdf:ID="Pilot">
<rdfs:label>Pilot</rdfs:label>
<rdfs:subClassOf rdf:resource="#Pilot">
</owl:Class>
<owl:Class rdf:ID="Havayolu Sirketi">
<rdfs:label> Havayolu Sirketi Calisani</rdfs:label>
<rdfs:subClassOf rdf:resource="#Calisan">
</owl:Class>
<owl:Class rdf:ID=" Calisan">
<rdfs:label> Calisan </rdfs:label>
<owl:intersectionOf rdf:parseType="Grup">
<owl:Class rdf:about="#Kisi" />
<owl:Restriction >
<owl:onProperty rsd:resource="#calistigiKurum" />
<owl:Class rdf:about="#Kurum" />
</owl:someValuesFrom>
</owl:Restriction>
</owl:intersectionOf>
</owl:Class>
<owl:ObjectProperty rdf:ID="ucusYapar">
<rdfs:label>ucar</rdfs:label>
<rdfs:domain rdf:resource="#Rota">
<rdfs:range rdf:resource="#Ulke">
</owl:ObjectProperty>

```

Şekil 10: OWL Örneği

Web Ontoloji Dili, anlam ve anlambilimi ifade etmek için daha zengin bir altyapı sağlar, ancak bir dil ne kadar ifade ediciyse, dille akıl yürütmek o kadar zordur. Karmaşık dil yapıları daha fazla bilgiyi temsil etmeye izin verse de hesaplama verimsiz hale gelir ve sonunda karar verilemez. Semantik Web için dil seçimi ve ontoloji, tasarlanacak uygulamanın türüne bağlı olarak her zaman ifade edilebilirlik ve verimli akıl yürütme arasında bir denge uyumu olmalıdır. OWL, farklı ifade ve hesaplama özelliklerine sahip üç dilden oluşan bir aileden oluşur: OWL Full, OWL DL ve OWL Lite [20].

OWL, web ontolojileri için önerilen standartta verilen addır. Bilginin semantiğini makine tarafından erişilebilir bir şekilde tanımlamamızı sağlar. OWL, RDF ve RDF Şeması üzerine kurulu bir yapıya sahiptir. XML tabanlı RDF sözdizimini kullanılır. Örnekler, RDF açıklamaları kullanılarak tanımlanır ve çoğu RDFS modelleme ilkeleri kullanılır. Biçimsel anlambilim ve akıl yürütme desteği, OWL'nin mantıklar üzerinde eşlenmesi yoluyla sağlanır. Bu amaçla yüklem mantığı ve açıklama mantığı kullanılmıştır [20].

OWL pratikte kullanılmak için yeterince zengindir fakat uzantılar yapım aşamasındadır. Kurallar da dahil olmak üzere daha fazla mantıksal özellik sağlamaktadır.

### 3.6.2 Ontoloji Uygulama Senaryoları

Semantik web’de veri toplama yöntemlerinde ilk önce sistemdeki ontoloji uygulama senaryoları olarak adlandırılan senaryoları belirlenmelidir. Bu senaryolar, endüstri veya araştırmadan alınan belirli ontoloji uygulamalarının soyutlamalarıdır. Her senaryo, ontolojinin amaçlanan hedefini, ontolojinin rolünü, önemli aktörlerin kimler olduğunu ve destekleyici teknolojileri tanımlayan bir genel bakış içerir. Aşağıda, dört ontoloji uygulaması kategorisi olarak görülen ontoloji senaryolarını ele alınmaktadır. Kategorilerin sahip olduğu her bir uygulama senaryosu tanımlanabilmektedir. Örneğin, objektif olma, bilgiye ortak erişim ve indeksleme temel senaryolar olabilir [21].

**Objektif Olma Yapısı:** Bilgi yapısının altyapısı tek bir dilde yazılmaktadır. Birden çok hedef sistemde kullanım amaçlanır Her hedef sistem bir forma dönüştürülür. Bu yaklaşımın olumlu yanları bilginin yeniden kullanıma açılması, iyileştirilmiş veriler, sürdürülebilir kullanım ve uzun süre boyunca bilgi saklanması yer almaktadır [21].

**Bilgiye Ortak Erişimde Bulunma:** Kullanıcılar veya kullanılan bilgisayar uygulaması tarafından bilgi istenir. Bilinmeyen kelimeler kullanılarak veya erişilemeyen bir biçimde ifade edilmektedir. Ontoloji, terimlerin ortak bir şekilde anlaşılmasını sağlayarak veya terim kümeleri arasında haritalama yaparak bilgiyi anlaşılır hale getirmeye yardımcı olur. Bu yaklaşımın olumlu yanları, birlikte çalışabilirlik durumunu, bilgi kaynaklarının daha etkin kullanımına açılmasını ve yeniden kullanımını içermektedir [21].

**İndeksleme Sağlama:** Bilgi yapılarını indeksleme için bir mekanizma temelinde ontoloji kullanılır. Bu yaklaşımın olumlu yanları, bilgi kaynaklarının daha etkin kullanımına ve yeniden kullanımına yol açan önemli bilgi kaynaklarına daha hızlı erişimin açılmasıdır [21].



## 3.7 Kullanılan Programlama Dilleri

### 3.7.1 Geniřletilebilir İřaretleme Dili (XML)

XML, hızla artan sayıda yazılım geliştirme uygulamalarının kaynağıdır. XML, belirli bir hiper metin belgesi türü için bir biçimlendirme dili olan HTML'in yerine, rastgele belge yapısı için bir biçimlendirme dili olarak geliştirilmiştir. Bir XML belgesi, her etiketin bir dizi öznitelik-değer çiftine sahip olabileceği, uygun şekilde yazılmış bir açık ve kapalı etiketler kümesinden oluşmaktadır [13].

XML için temel olan, etiketlerin ifadelerinin ve izin verilen kombinasyonlarının sabit olmamasıdır. Sadece XML uygulaması başında tanımlanabilir. XML'in temel veri modeli, her etiketin verideki etiketli bir düğüme karşılık geldiği etiketli bir ağaç yapısıdır. Model ve her iç içe alt etiket, Şekil 11'de gösterildiği gibi ağaçtaki bir alt etiketi oluşturur.

```
<class-def>
  <class name="bitki">
    <subclass-of>
      <NOT><class name="hayvan"></NOT>
    </subclass-of>
  </class-def>
<class-def>
  <class name="agac" />
  <subclass-of>
    <NOT><class name="bitki"></NOT>
  </subclass-of>
</class-def>
<class-def>
  <class name="dal" />
  <slot-constraint>
    <slot name="sahiplik">
      <has-value>
        <class name="agac">
      </has-value>
    </slot-constraint>
  </class-def>
```

Şekil 11: XML Sınıf Etiketleri

İç içe etiketleri dengeli bir ağaç oluşturan herhangi bir XML belgesi, iyi biçimlendirilmiş bir XML belgesi olduğunu gösterir. Şekil 11'de XML'e ait bu ağaç yapısı gösterilmiştir. Aynı zamanda, hangi etiketlerin kullanılması gerektiği ve bu

etiketlerin iç içe yerleştirilmesine izin verildiği konusunda kısıtlamalar uygulanabilmektedir.

### 3.7.2 Kaynak Tanımlama Çerçevesi (RDF)

Web uygulamaları arasında meta veri alışverişi için güvenli bir mekanizma sağlayan URL tabanlı bir sözdizimi veri gösterimidir. RDF, üç nesne türüne sahiptir. Bunlar Kaynak, Özellik ve Açıklama- dayalı soyut veri modelleri oluşturarak meta verileri işlemektedir. Kaynak bir ifadedir, Özellik kaynağı açıklayan bir özniteliktir. Açıklama bazı özelliklere ve değerlere sahip bir kaynaktır [22].

World Wide Web Consortium önerisi haline gelen kaynak tanımlama çerçeve modeli, RDF veri modelini ve XML tabanlı bir serileştirme sözdizimini tanımlar. Spesifikasyon, veri modelini bir ağda RDF verilerini taşımak için gereken sözdiziminden ayırarak, RDF'ye duyarlı bir sistemin yeni bir RDF olmayan veri kaynağına erişmesine izin verir. Bunu yapmak için, bir RDF bağdaştırıcısı önce RDF olmayan veri kaynağındaki kaynaklara benzersiz kaynak tanımlayıcıları (URI'ler) atar ve ardından kaynak özelliklerini açıklayan ifadeler oluşturur. RDF veri modelinin nesnelere açık bir şekilde belirtmek için URI'leri ve nesnelere arasındaki ilişkileri tanımlamak için özelliklerini kullanması, onu temelde XML'in ağaç tabanlı veri modelinden ayırır [22].

Kaynakların yerleşik kavramı ve kaynaklar arasındaki ilişki ile RDF, Web'i makine tarafından işlenebilir bilgilerle doldurma hedefini yerine getirmeyi amaçlar. RDF veri modelinin basitliği, verilerin temsil edilmesini basitleştirir ve birleşik modelleme dili veya açıklama mantıkları RDF'nin üstünde tanımlanabilir. RDF, Web uygulamaları arasında birlikte çalışmayı sağlamak için en düşük ortak paydadır. Nesne yönelimli olduğundan bilgi alışverişi için XML'den daha uygun bir veri modeline sahiptir ve yeni sözlükler tanımlamak için son derece esnekler.

RDF, XML'e göre daha kullanışlıdır, çünkü veri modelleme için bağımsız sözdizimi serileştirme ve kısaltma kullanımını sağlar, sözdizimi somutlaştırma ve şema dilinde kullanılan terminolojileri tanımlamada etki alanı bağımsızlığı, kelime dağarcığı ve ayrıcalıklar gibi anlamsal tabanlı özelliklerin kullanılmasını sağlamaktadır. RDF modelleme çeşitli mantıksal ifadeleri çözümlemede yetersiz kalmaktadır [22].

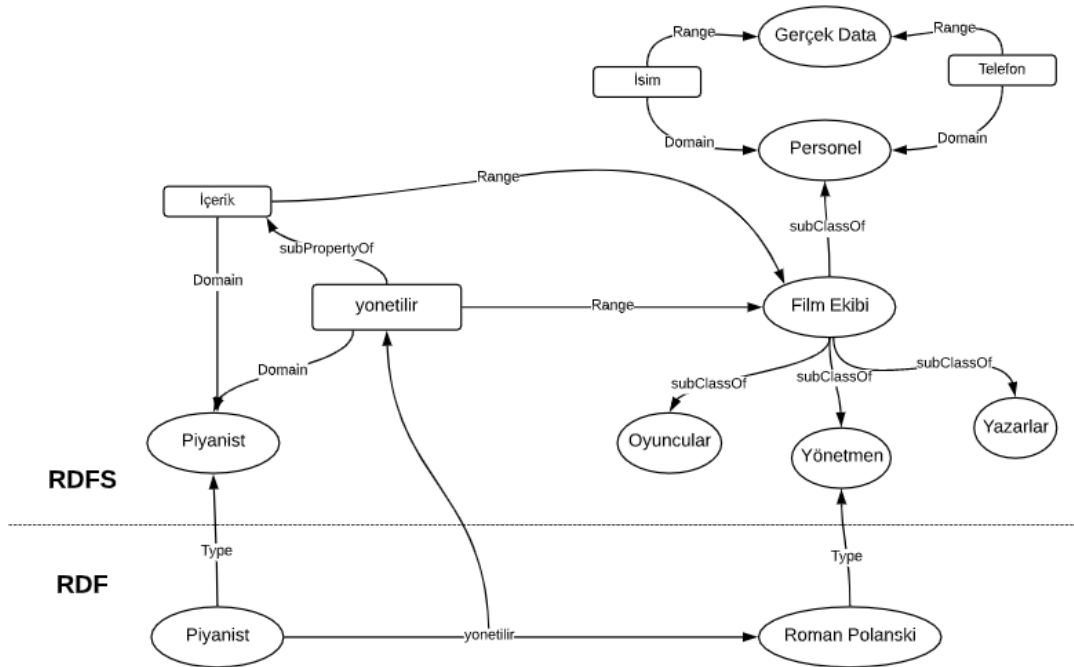
RDF bilgi içeriği ile ilgilenirken, XML serileştirme ile ilgilidir. Birlikte çalışabilirlik sorununun çözümlerini, XML sözdizimi ve veri modeli içeren ağaç yapısına sahip olduğundan RDF, bir XML sözdizimi de dahil olmak üzere farklı sözdizimine sahip URI'leri kullanan bir grafiğe dayalı veri modelidir. Bu sebeple veri toplama yapısını salt XML'den ziyade RDF ile çözmek daha kolaydır [22].

### 3.7.3 Kaynak Tanımlama Çerçevesi Şeması (RDFS)

RDF Şeması, RDF'ye nesne yönelimli, genişletilebilir bir tür sistemi tanıtan bir RDF uygulamasıdır. RDFS, özellik etki alanlarını ile aralıklarını ve sınıf ile alt sınıf hiyerarşilerini tanımlamak için araçlar sağlar.

RDF, özellikler ve değerleri kullanarak kaynaklar hakkında basit açıklamaları ifade etmenin bir yolunu sağlar. Kullanıcıların belirli kaynak türlerini veya sınıflarını tanımladıklarını ve bu açıklamada belirli özellikleri kullanacaklarını belirtmeleri gerekir. RDF şeması, kaynakların ek açıklamaları olarak şema bilgileri sağlar, ancak açıklamaların bir uygulama tarafından nasıl kullanılmasını gerektirmediğini belirlemez [23].

RDF ve RDFS ilişkisi arasındaki fark bir örnek ile Şekil 11'de gösterilmiştir.



Şekil 12: RDF ve RDFS İlişkisi

Yukarıdaki şekilde RDF ve RDF şeması bağlantısı gösterilmektedir. Bloklar özellikleri, RDFS alanındaki elipsler sınıfları ve RDF alanındaki elipslerde örneklerdir.

RDFS'deki bir sınıf, tür veya kategorinin genel kavramına karşılık gelir. Bir sınıf, bütün kategoriye temsil etmek için kullanılabilir. Bir sınıfa ait olan kaynağa onun örneği denir. RDF'de, değeri kaynak rdfs:Class olan rdf:typeproperty ile bir kaynak sınıfı atanır. İki sınıf arasındaki ilişki, bu iki sınıfı ilişkilendirmek için rdfs:subClassOfproperty kullanılarak tanımlanır. Bu ilişkinin anlamı, alt sınıfın herhangi bir örneğinin aynı zamanda sınıfın bir örneği olmasıdır. Bir sınıf, birden fazla sınıfın alt sınıfı olabilir. RDF Şeması, tüm sınıflara ait örnekler kaynak olduğundan, tüm sınıfları rdfs:Resource sınıfının alt sınıfları olarak tanımlar [23].

Belirli sınıfları tanımlamanın yanı sıra, kullanıcıların bu sınıfları karakterize eden özellikleri de tanımlayabilmeleri gerekir. RDF Şemasında tüm özellikler rdf:Property sınıfı ve rdfs:domain, rdfs:range ve rdfs:subPropertyOf RDFS özellikleri kullanılarak tanımlanır. RDF Şeması ayrıca, özelliklerin ve kaynakların nasıl ilişkili olduğunu açıklamak için bir hiyerarşi sağlar. En önemli bilgiler rdfs:domain ve rdfs:range özellikleri kullanılarak sağlanır [23].

rdfs:domain özelliği, belirli bir özelliğin belirlenmiş bir sınıf için geçerli olduğunu gösterir. RDF'de, özellik tanımları varsayılan olarak bağımsızdır ve global kapsamı vardır. RDF Şeması, bir etki alanı belirtilmeden bir özelliği tanımlayabilir, bu da bir özellik tanımının kullanımını farklı bir duruma genişletmek mümkündür. rdfs:range belirli bir özelliğin değerlerinin belirlenmiş bir sınıfın örnekleri olduğunu belirtir. RDFS'de belirli bir özelliği, uygulandığı kaynağın sınıfına bağlı olarak farklı aralıklara tanımlamak mümkün değildir. Bir özellik için tanımlanan herhangi bir aralık, o özelliğin tüm kullanımları için geçerlidir [23].

### **3.8 Test Süreçleri**

Semantik web'de her düğümün www'in tamamındaki birden fazla düğümle ilişkili olabileceği göz önüne alındığında, veri tabanı testiyle ilgili sorunların anlamsal veri tabanları için analiz edilmesi önem arz etmektedir. Anlamsal web, veri tabanlarının test edilmesiyle ilgili tüm sorunlar, veriler sorgulanırken ele alınmaktadır. Anlamsal

web için kullanılan sorgu dili SPARQL'dir [23]. Verileri analiz etmek için bu verilerin daha önce sorgulanması gerektiği ortaya çıkmaktadır.

Semantik web servis testi, geliştirme testi, depolama alanı testi ve son müşteri testi olmak üzere en az üç aşamadan oluşur.

- **Geliştirme testi:** Web servislerinin tasarımcısı tarafından gerçekleştirilir.
- **Depolama alanı testi:** Bir operatör tarafından verilerin kaydedileceği alanda yapılan testtir.
- **Son istemci testi:** Hizmetlerin hala yasal olarak çalıştığını doğrulamak için web hizmetlerinin istemcisi tarafından gerçekleştirilen testtir [23].

Test sürecinde, test senaryosu oluşturma ve test senaryosu yürütmenin tamamen değişken iki adım olduğu bilinmektedir. Hangi test senaryosu seçim metodolojisi kullanılırsa kullanılsın, birkaç kod hatasının hala tespit edilmekte sorunlar mevcuttur.

### 3.9 Veri Tabanı Yapısı

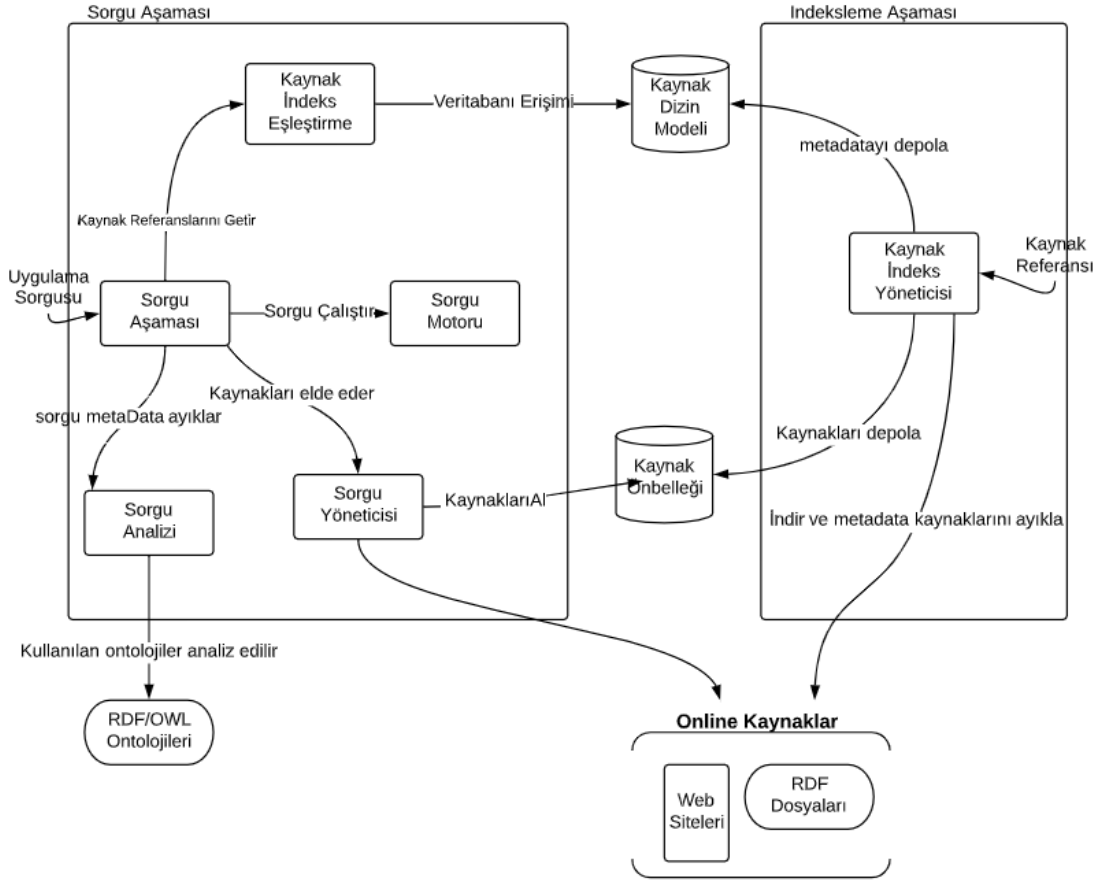
Anlamsal web için kullanılan veri modeli, ilişkisel veri tabanlarının modeliyle doğrudan bağlantı kurmaktadır. İlişkisel bir veri tabanı, satırlardan veya kayıtları içeren tablolardan oluşmaktadır. Her kayıt bir dizi alandan oluşur. Semantik web için temel yapılardan biri, her zaman makineler tarafından işlenebilecek şekilde çok miktarda ilişkisel veri tabanı bilgisinin web üzerinden ulaşılabilir olmasıdır. RDF'nin yapısını formatı ilişkisel veri tabanı bilgilerini ifade etmek için kullanılan bir yapıdır [24].

İlişkisel veri tabanı sistemleri ile RDF verileri yönetilir. Tabloda yer alan, aynı özelliklere sahip birçok kayıt içermektedir. Tek bir hücre genellikle kendi başına işleme alınmaz. SQL sorguları tabloları bir araya getirmektedir ve tablolardan veri dışarıya çıkarılmaktadır. Bu nedenle, ilişkisel veri tabanı modelleriyle hazırlanan yazılımların bazıları çok sayıda öğeye sahip olabilen az sayıda tabloyla işlemler yapmak için optimize edilmiştir [24].

Bir veri tabanı tablosunun temel bir yönü, genellikle bir tablodaki verilerin kesin olabilmesidir. Ne RDF ne de ilişkisel veri tabanı modellerinde bunu ifade etmenin basit yolları yoktur. Örneğin, bir tablodaki bir satır yalnızca İstanbul plakası "34XY21" olan kırmızı bir araba olduğunu göstermekle kalmaz, aynı zamanda tablo,

herhangi bir arabanın İstanbul plakası varsa, o zaman tabloda olması gerektiği yazılı olmayan semantiği de taşıyabilir.

Sorgu ve indeksleme aşamalarının yer aldığı grafik Şekil 13'te yer almaktadır.



Şekil 13: Semantik Web Veri Tabanı Yapısı [25]

Her bir RDB modelinde, her tablonun bir birincil anahtarı vardır. Değeri her satırı benzersiz bir şekilde tanımlamak için kullanılabilen bir sütuna sahip olmaktır. Bazı ürünler bunu içermez, bu da yinelenen satırların öneminde bir belirsizliğe yol açar. Bir satırın kimliğini değiştirmeden birincil anahtarın değiştirilebilmesidir. SQL, tabloların ayarlanmasına izin verir. Böylece bu tür değişiklikler referans bütünlüğünü korumak için yerel sistem boyunca artırılabilir. SQL bunu standart bir şekilde göstermese de birçok sistem aslında kullanır [24].

RDB sistemleri, RDF ve XML'in yapacağı yapılandırılmamış düzeyde veri tiplerine sahiptir. Kombinasyon kuralları, RDB'lerde esnek bir şekilde uygulanma

eğilimindedir, çünkü bir sorgu, anlambilim üzerinde herhangi bir kontrol olmaksızın, veri türüne göre eşleşen herhangi bir sütunla tablolara katılabilir [24].

Bir tabloya erişebilmek ve onun hakkında daha fazla şekilde kullanılmasını sağlayacak ek açıklamalar yapabilmek için, tablonun temel nesnelere web'de birinci sınıf nesnelere olarak dışa aktarılması gerekir. Herhangi bir sistemi web'e eşlerken, URI alanına eşleme çok önemlidir. Burada bu ortak işlemi genel olarak tüm ilişkisel veri tabanları için yapılmaktadır. Bunun birden fazla satıcı arasında tutarlı bir şekilde yapılmasının faydalı olacağı açıktır.

## 4. PAZAR SEPET ANALİZİ VE BİRLİKTELİK KURALLARI

### 4.1 Temel Kavramlar

Pazar sepeti analizi perakende ve online alışverişlerde, yöneticilerin satın alma modellerini vurgulayarak müşterilerini daha iyi anlamalarına ve nihayetinde hizmet etmelerine yardımcı olan bir dizi istatistiksel yakınlık hesaplamasıdır. En basit ifadeyle Pazar sepeti analizi, siparişlerde en sık hangi ürün kombinasyonlarının birlikte gerçekleştiğini ilişkisel olarak bize vermektedir [26]. Bu ilişkiler, çapraz satış, tavsiyeler, promosyonlar ve hatta öğelerin bir menüye veya mağazaya yerleştirilmesi yoluyla karlılığı artırmak için çok sık kullanılmaktadır.

Pazar sepeti analizi ile müşteri satın alma davranışını anlama ve bu davranışları pazara göre yönlendirebilmek mümkündür. İşlemlerde sıklıkla birlikte ortaya çıkan ürün kombinasyonlarını veya ikili ürün öğelerini belirleyerek müşterilerin satın aldıkları ürünler arasındaki ilişkileri ortaya çıkarabilmektedir. Perakendeciler ve firmalar, müşterilerin satın aldığı ürünler arasındaki ilişkileri kendileri belirleyebilir ve bu bilgileri anlamlı hale getirerek yeni ürünler veya yeni gelir elde etmek için fiyatlandırma modellerini kendileri oluşturabilirler.

Pazar sepeti analizi teorik olarak belirli bir ürünü veya grupları satın alan müşterinin başka bir ürünü satın olma olasılığı üzerinden en yüksek olasılık teorisine dayanmaktadır [27]. Örneğin elektronik ürünler satan bir mağazadan cep telefonu satın alan bir müşterinin ekran koruyucu da alma olasılığı her zaman daha yüksektir. Bu bağıntı, cep telefonu ve ekran koruyucu arasındaki ilişkinin daha güçlü olduğunu gösterirse ilişkideki bağıntı daha değerli olmaktadır.

Pazar sepet analizi güvenilir sonuçlar oluşturmak için genellikle büyük miktarda işlem verisine ihtiyaç duymaktadır. Büyük veri kümelerinin, yüksek düzeyde ölçeklenebilir hesaplama kaynakları ve depolama alanı olmadan işlenmesi imkansızdır. Bulut tabanlı mimariler, teorik olarak müşteri davranışlarını test etme veya yeni bir pazarlama stratejileri oluşturmak için daha kullanışlı analitiklere izin verir [26].



Firma sahipleri, pazar sepeti analizinden edinilen bilgileri aşağıdaki örneklerde dahil olmak üzere çeşitli şekillerde kullanabilir:

- **Çapraz Satış:** Müşterilerin en sık satın aldığı ürünler mağazanın ürün yerleştirmesinde bir arada gruplandırılır.
- **Online Alışveriş:** Sıklıkla birlikte satın alınan ilişkili ürünler önerilmektedir. “Bu ürünü alan müşteriler bu ürünü de satın aldı...” mantığı müşterilerin alışveriş alışkanlıklarına göre belirlendi.
- **Pazarlama Promosyonları:** Yapılan kampanyalar ile müşterilere hedeflenir ve yakın zamanda aldıkları ürünlerle ilgili yeni veya benzer ürünler satın alma için önerilmektedir [26].

#### 4.2 Pazar Sepeti Analizinin Güçlü ve Zayıf Yönleri

Pazar sepeti analizinin en önemli avantajlarından biri, yönlendirilmemiş veri madenciliği için bir analiz konusu olmasıdır. Bu teknik, büyük bir veri kümesiyle nereden başlayacağımızı bilmediğimizde kullanılır. Veri madenciliği tekniklerinin çoğu, yönlendirilmemiş veri madenciliği için kullanılmazken, büyük verileri analiz etmek ve kullanıcıya uygun bir başlangıç sağlamak için pazar sepeti analizi kolayca uygulanabilir [26].

İlk olarak, verilerin kaydedilme şekli veri formatı ile başlanmalıdır. Her değişkenin ilgili bir veri türü vardır, bir nesnenin tutabileceği veri türü olmalıdır. Değişken uzunluklu formatta kaydedilen veriler, yerden tasarruf sağladığı için kullanışlıdır. Sabit biçimli ve değişken uzunluktaki veriler arasındaki fark, bir kaydın tutabileceği karakter sayısıdır. Sabit uzunluklu biçim için, her alan en uzun adı tutacak kadar uzun olacak şekilde önceden tanımlanmalıdır. Kısa adlara sahip kayıtlar için yer kaybı olarak görülebilir. Değişken uzunluklu verilerde, her alan kaydının uzunluğu kadar boyutu uzun olabilir. İşlem verileri söz konusu olduğunda, öğeleri temsil etmenin en doğal yolu, bunların değişken uzunluklu bir veri türünde kaydedilmesidir. Pek çok teknik, veri kayıtlarıyla sabit bir formatta çalışırken, pazar sepeti analizi, önemli bilgileri kaybetmeden değişken uzunluktaki verileri işleyebilir. Analizin bir diğer önemli avantajı ve gücü, operasyonel basitliğidir. Sinir ağlarının aksine, pazar sepet analizinde hesaplamalar basittir ve daha küçük problemler oldukça kolay çözümlenebilmektedir.

Ürün ve işlemlerin sayısı arttıkça, birliktelik kuralları oluşturmak için gereken hesaplamalar çok hızlı, hatta katlanarak büyür. Bu sorun için olası bir çözüm, öğelerin sayısını azaltmaktır. Maddeleri genelleştirerek ve daha yüksek bir sınıflandırma düzeyinde toplayarak kolayca yapılabilir. Genelleştirilmiş öğeler her zaman eyleme geçirilebilir olmasa da kural oluşturma sürecini kontrol etmek için bazı yöntemler vardır. Minimum destek filtreleme buna bir örnektir.

Analizin temel sorunu, doğru toplama düzeyini belirlemektir. Genelleme sürecinde bazı bilgiler kaybolabilir ve maddelerin sıklığı orijinal seviyelerden farklı olabilir. Olası çözüm, genelleştirilmiş öğelerden kaybolan bilgileri yakalayabilen sanal öğeler eklemektir.

### 4.3 Birliktelik Kuralları

Büyük veriler üzerinde veri madenciliği tekniklerinin uygulanmaya başlanması, kayıtlardaki ürünler arasındaki korelasyonların bulunmasına ve birliktelik kurallarının oluşturulmasına ihtiyaç duyar. Birliktelik kurallarının temel prensibi, öğeler arasındaki olası tüm kuralları incelemek ve bunları 'if-then' ifadelerine dönüştürmektir [28].

Birliktelik Kuralları, minimum destekten daha büyük desteğe sahip tüm veri kümelerini bulur ve ardından minimum güvenden daha büyük güvene sahip istenen kuralları oluşturmak için büyük veri kümelerini kullanır. Bir birliktelik kuralının ilginçliği, X ve Y bağımsız olduğunda gözlemlenen desteğin beklenene oranıdır. Birliktelik kuralları uygulamasının tipik ve yaygın olarak kullanılan bir örneği, pazar sepeti analizidir. Pazar sepet analizinde kullanılan üç önemli unsur vardır.

#### 4.3.1 Destek (Support)

Destek, öge setinin veri setinde ne sıklıkta görüldüğünün bir göstergesidir [29].

$$Destek (Support) = \frac{frekans(X,Y)}{N}$$

Hem X hem de Y ile işlem sayısının toplam işlem sayısına bölümüdür. Kurallar, düşük destek değerleri için kullanışlı değildir. Destek, bir kuralın verilen veri kümesine ne sıklıkla uygulanabileceğini ölçer. Bu ölçümde miktar dikkate alınmaz. Veri tabanından farklı bir T işleminde öğeyle her karşılaşıldığında destek sayısı bir artar.

### 4.3.2 Güven Seviyesi (Confidence)

Birliktelik kuralının güven seviyesi, XUY içeren işlem sayısının X içeren toplam kayıt sayısına oranı olarak tanımlanmaktadır. Güven seviyesi, birliktelik kurallarının gücünün bir ölçüsüdür ve öğelerin ne sıklıkta kullanıldığını belirlemek için kullanılır. Y öge kümesinden X öge kümesini içeren işlemlerde görünmektedir. Güven seviyesi, X içeren bir işlemde Y bulmanın ne kadar olası olduğunu söylemektedir [29].

$$\text{Güven Seviyesi (Confidence)} = \frac{\text{frekans}(X, Y)}{\text{frekans}(X)}$$

### 4.3.3 Kaldıraç (Lift)

Kaldıraç, bir kuralın önemini ölçer. Kaldıraç değeri, bir kuralın güvenilirliği ile beklenen güven seviyesinin oranı olarak temsil edilir. Lift sıfır ile sonsuz arasındaki değerleri alabilir. Her birliktelik kuralında, buna göre kural gövdesi ve kural başlığı olarak da adlandırılan bir öncül ve bir sonucumuz vardır. X ve Y'nin bağımsız olması durumunda gözlemlenen desteğin beklenene oranıdır [30] ve şu şekilde tanımlanır:

$$\text{Kaldıraç (Lift)} = \frac{\text{Destek}}{\text{Destek}(X) \times \text{Destek}(Y)}$$

Yukarıdaki bölümde, bir kuralın kesinliğini ölçmek için üç önemli metrik tanımlanmaktadır [31, 32].

Birliktelik kuralları genellikle çok büyük veri kümeleri için kullanılmaktadır ve en verimli algoritmalar veri kümelerinin işlenmesinde büyük öneme sahiptir. Bu yöntem, her farklı veri kümesi boyutu için veri kümesinden bir tarama geçişi yapar. Bazen veri kümesi ana belleğe okunamayacak kadar büyük olmaktadır ve diskte tutulması gerekir. Böyle durumlarda, aynı anda iki ardışık boyuttaki veri setlerini kontrol ederek geçiş sayısını azaltma yöntemi denenebilir [33]. Örneğin, iki ögeli kümeler oluşturulduktan sonra, kümelerdeki gerçek veri sayısını saymak için örnek kümesine geçmeden önce üç öğeden oluşan tüm kümeler bunlardan oluşturulabilir. Elde edilebilecek üç öge içeren kümeler listelenecek ve tüm veri kümesinden geçiş sayısı azaltılacaktır.

## 4.4 Birliktelik Kurallarının Belirlenmesinde Kullanılan Algoritmaları

### 4.4.1 Apriori Algoritması

Apriori algoritması, 1994 yılında Agarwal ve Srikant tarafından geliştirilmiştir. Tüm ARM algoritmalarının en ünlülerinden biridir [34]. Apriori, işlemler de dahil olmak üzere veri tabanları üzerinde çalışmak üzere tasarlanmıştır ve sık kullanılan öge kümelerini üretmek için kullanılmaktadır.

Algoritma, minimum destek koşulunu karşılayan öge kümeleri, mümkün olan maksimum öge kombinasyonu sayısına ulaşana kadar, veri kümesi üzerinde birden çok geçiş yürütürken, tüm seyrek öge kümeleri kaldırılır.

Apriori, veri madenciliği birliktelik kurallarında en yaygın kullanılan algoritmadır. İşlemsel veri tabanında sık görülen bireysel öğelerin tanımlanmasıyla başlar ve bunları veri tabanında yeterince sık görünene kadar daha büyük öge kümelerine genişleterek devam eder. Apriori algoritması, minimum destek koşulunu karşılayan başka uzantı bulunmayana kadar çalıştırılır. Python’da Apriori algoritması için apyori kütüphanesi en sık kullanılan kütüphanedir [35].

Algoritmanın ana fikri, veri tabanını sık öge kümelerini tararken, sonraki her adımda seyrek bulunan öğeleri budamaktır. Bu işlemleri yaparken iki çok önemli adım vardır Bunlar, birleştirme ve budama adımdır. Başlangıçta,  $L_k$ 'nin kendisiyle birleştirilmesi,  $C_{k+1}$ 'in üretilmesiyle sonuçlanır. Budama adımında, sık olmayan  $(k+1)$  öge kümesinin bir alt kümesi olamayacağından, seyrek olan herhangi bir  $k$  öge kümesi varsa budanır.

- $C_k$  –  $k$  boyutuna sahip aday öge kümeleri.
- $L_k$  –  $k$  boyutunda sık kullanılan öge kümeleri.

Şekil 14’te Apriori algoritmasının temelini oluşturan pseudo-code bulunmaktadır.

```

1)  $L_1 = \{\text{large 1-itemsets}\};$ 
2) for (  $k = 2; L_{k-1} \neq \emptyset; k++$  ) do begin
3)    $C_k = \text{apriori-gen}(L_{k-1});$  // New candidates
4)   forall transactions  $t \in \mathcal{D}$  do begin
5)      $C_t = \text{subset}(C_k, t);$  // Candidates contained in  $t$ 
6)     forall candidates  $c \in C_t$  do
7)        $c.\text{count}++;$ 
8)   end
9)    $L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\}$ 
10) end
11)  $\text{Answer} = \bigcup_k L_k;$ 

```

**Şekil 14:** Apriori Algoritması Pseudo-Code [36]

Apriori algoritması aşağıdaki adımlarda uygulanmaktadır:

1. Sık kullanılan öğeler bulunmalıdır ve  $L_k$ 'ye ( $k=1$ ) eklenmelidir.
2. Büyüklüğü ( $k+1$ ) olan  $C_{k+1}$  aday öğe kümelerinin bir koleksiyonunu oluşturmak için  $L_k$  kullanılmalıdır.
3.  $C_{k+1}$ 'deki hangi öğelerin sık olduğunu bulmak için veri tabanı taranmalıdır ve bunları  $L_{k+1}$ 'e eklenmelidir.
4.  $L_{k+1}$  boşta değilse:
  - $k:=k+1$
  - Adım 2'ye gidilir [36].

Aşağıdaki örnek, Apriori'nin birkaç basit adımda nasıl uygulandığını göstermektedir. Örnek bir işlem veri tabanını seçilmiştir. Aşağıdaki kümelere oluştuğunu varsaymaktadır:  $\{a, c, d\}$ ,  $\{b, c, e\}$ ,  $\{a, b, c, e\}$ ,  $\{b, e\}$ . Her harf, ürün kataloğunda belirli bir ürüne karşılık gelir. Örneğin  $\{a\}$  bilgisayar,  $\{b\}$  cep telefonudur.

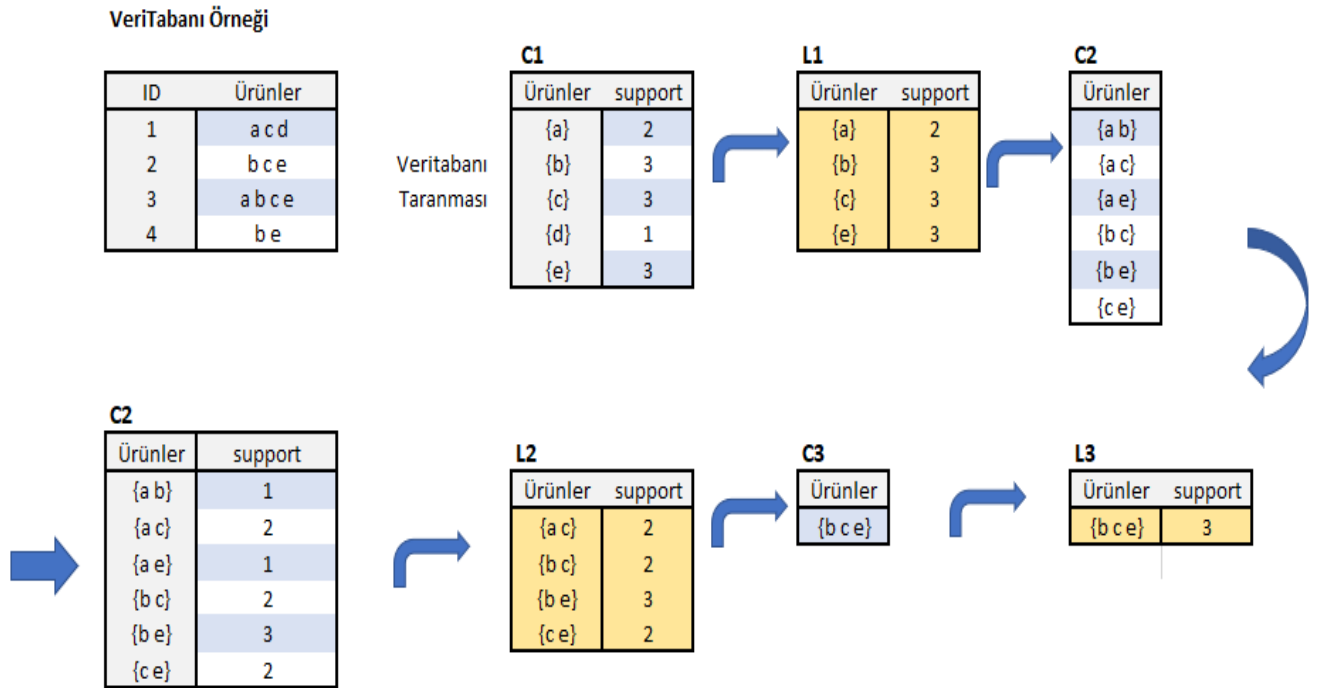
İlk adımda, algoritma destek olarak da adlandırılan her bir öğenin frekanslarını ayrı ayrı hesaplamaktadır. Bir öğenin sık olduğundan emin olmak istiyorsak, minimum destek seviyesini önceden tanımlanabilir. Bu durumda minimum destek 2'dir. Bu nedenle maddelerden dördünün sık olduğu bulunmuştur.

Bir sonraki adımda, bütün ürünlerin yer aldığı 2 çift sık öğenin bir listesi oluşturulur. Dizinde bulunan seyrek öğeler, daha fazla analiz için budanır. Tüm olası öğe çiftlerini

bulmak için, Apriori algoritması tüm olası kombinasyonların kaldırılmasını sağlar [37].

Son adımda, sık bir çifti sık tek bir öğeye bağlayarak, sık öğelerin üç üçlüsünün tümünün bir listesi oluşturulur. Algoritma bu adımda sona erer, çünkü bir sonraki adımda oluşturulan dört öğe çifti gerekli minimum desteği karşılamaz [37].

Veri setinde, işlemlerin boyutu arada bir yerde olma eğilimine sahiptir. Müşteriler tek bir alışveriş gezisinde az ürün satın almayı planlar ve dört veya daha fazla ürün içeren kurallar arayanlar yalnızca birkaç işlem için geçerli olabilme ihtimali vardır. Bu nedenle, bir müşterinin tek bir işlemde satın aldığı ortalama ürün sayısı 2 civarında olduğu için analizde oldukça karmaşık hesaplamalar beklenmemektedir. Tek bir işlemde daha fazla öğe varsa, kural oluşturma süreci daha uzun sürer. Süpermarket ya da buna benzer yerlerde durum böyledir. Ortalama işlem, daha büyük olacaktır [37].



**Şekil 15:** Apriori Uygulama Adımları [38]

Şekil 15'te sık öğelerin bulunması ve sık olmayan üyelerin budanması gösterilmiştir. Bu budama adımları aşağıda yer alan iki adımda yapılmaktadır.

## A

- Apriori veri tabanını tarar ve her bir öge için C desteğini hesaplar.
- Minimum önceden tanımlanmış destek eşiğinden daha büyük C desteğine sahip tüm ögeler için, sık kullanılan 1 öge kümeleri L1 oluşturulur.
- k-frekans aday öge kümeleri, k-1 boyutundaki Lk-1 sık öge kümeleri kullanılarak ve öge üst kümeleri oluşturmak için ögelerinin birleştirilmesiyle oluşturulur.
- Her yeni k-frekans öge kümesinin desteği hesaplanır. Destek  $\leq$  minimum olan öge kümeleri budanır. Bu işlem, başka aday öge kümeleri oluşturulamayana kadar tekrarlanır.
- Ardından Apriori, B adımına geçer [39].

## B

- Her öge kümesi L için, tüm boş olmayan alt kümeler oluşturulur.
- Her bir J alt kümesi için verinin güven seviyesi hesaplanır
- Önceden tanımlanmış minimum güvenden daha büyük güven seviyesine sahip tüm alt kümeler, güçlü kurallara eklenir [39]

## 5. VERİ

### 5.1 Verilerin Açıklanması ve Hazırlanması

Araştırmada kullanılan veri seti, ev aletleri ve elektronik ürün satan büyük bir mağazanın veri tabanındaki satış kayıtlarının Ekim 2019 ile Ekim 2020 arasındaki satış verilerini içermektedir. 1578817 satır veri toplanmıştır [40]. 13 Kategori, 54 ürün alt kategorisi, 79 ürün bulunmaktadır.

#### Veriler hakkında bilinmesi gereken bazı detaylar:

- Belirli bir zaman aralığında toplanan veriler, farklı dönemlerde farklı sonuçlar üretmektedir.
- Ürünlerin kampanya ve promosyonları, verilerin elde edildiği dönemdeki satın alımları direkt etkilemektedir.
- Alışveriş alışkanlıkları, çalışmanın yapıldığı pazarların hedef kitlesine ve bölgesine göre değişiklik göstermektedir.

Veri tabanında yer alan veriler aşağıdaki bilgilerden oluşur:

- Tarih – satın alma tarihi;
- Sipariş ID;
- Ürün ID;
- Kategori ID;
- Kategori;
- Ürün Kategorisi;
- Ürün;
- Kullanıcı ID;



## 5.2 Veri Toplama

Pazar sepeti analizinin en önemli aşaması, verilerin nasıl toplanacağına karar verilmesidir. Doğru ayrıntı düzeyini seçmek araştırma için kritik bir noktadır. Araştırma sorusuna bağlı olarak, olası farklı toplama seviyeleri vardır. Bunlar ürün kategorileri, markalar, marka uzantıları ve alanlar olarak sınıflandırılabilir.

Web 3.0 teknolojisi, arama motorlarından, kullanıcıların bilgisayarlarındaki cookielerden, web sitelerin metatype'larından faydalanarak kendi sunucularında büyük veri havuzlarını oluşturmaktadır. Arama motorlarında yapılan ürün aramaları, kendi veri tabanlarında tutulmaktadır. Her arama anahtar kelimeler ve arama kriterleri detaylı bir şekilde tutulmaktadır. Örneğin Google Adwords aracında hangi kelimenin, hangi zaman diliminde, hangi bölgeden kaç kez aratıldığı analizi de yapılabilmektedir [41].

Web 3.0, Google Search ten ulaşılan anahtar kelimelerden bağımsız olarak kullanıcıların cookielerdeki dataları da çekerek veri tabanına kaydederek her bilgiyi sunucularında barındırmaktadır [41]. Aynı şekilde web sitelerinin metatype'larından tutulan web sitesi içeriği bilgilerini karşılaştırmalı olarak kullanabilmek için yine kendi sunucularında barındırmaktadır.

Veri kümesinde yer alan veriler, tek tek öğeler halinde ürün kategorilerinden toplanmıştır. Bu tür bir toplama, ürünlerin genelleştirilmesine sebep olur, böylece tek bir ürün kategorisi birkaç farklı ürünü hesaplamaya katmaktadır. Genelleştirilmiş öğeler, kategoriler arasında kurulacak olan ilişkileri ortaya çıkarma avantajına sahip olacaktır [42]. EK B'de Python programına verilerin hangi kütüphane ile alındığı açıklanmıştır.

Kategori kodlarına sahip ürünler, EK C'deki gibi taksonomi adı verilen hiyerarşik kategorilere ayrılır. İşleme alınacak verilerden sonuçlar çıkarmak istiyorsak, öğeleri daha ayrıntılı bir düzeyde yalın bir şekilde tanımlamak daha iyi olacaktır.

Veri seti, Python programı kullanarak büyük veri analitiğini tam olarak kavramak için çeşitli veri madenciliği yaklaşımları geliştirme bakış açısıyla çalışmanın yapılması hedeflenmiştir. Verilerin toplanmasından sonra, iyi yönlendirilmiş bir şekilde temizlenmesi en büyük öncelik olmalıdır. Verilerin temizlenmesi sırasında karşılaşılan çeşitli adımlar aşağıda açıklanmıştır:

**Ham verilerin dönüştürülmesi:** Bu adımda, canlı kaynaklardan toplanan araştırma verilerinde bulunan belirsizlikler önce ele alınır ve daha sonra her birini dikkate almak için Python programının gerektirdiği uygun bir biçime dönüştürmek için çalışmalar yapılır. Veri seti yapılandırılmış bir biçimde programa sunulur. Öneriyi farklı bir belge olarak değerlendirir ve buna göre analiz eder. Böylece araştırmacılardan ve sektör profesyonellerinden alınan veri setleri, sonuçların karşılaştırmasının uygun şekilde yapılabilmesi için ayrı ayrı istenen formata dönüştürülmüştür [43].

**Token veri:** Hesaplamalı programlamada "Token" olarak adlandırılan ilgili veriler, stringler, anahtar kelimeler vb. şeklinde küçük alt parçalara string ayırma işlemidir. Bunlar, ihtiyaçlara ve özellikle konuya göre büyük ölçüde boşluk veya noktalama işaretleriyle ayrılır. Sözcük vektörleriyle ayrılan bu belirteçler, ilgili kısıtlamalar uygulanarak kullanılabilir verileri filtrelemek için konumlandırılır [43].

**Satışı durdurulan veya istenmeyen ürünleri kaldırma:** Veri madenciliği algoritmalarında önemli ürünlerin önemini vurgulamak amacıyla, daha sonraki çalışmalar için belli bazı ürünlerin kaldırılması çok önemli hale gelmektedir. Kullanılabilirlik alanlarına göre bunlar birlikte karşılaştırılır ve son olarak istenmeyen ürünler çıkarılır [43].

**Ayrıştırma:** Ayrıştırma, verilerin temel kök biçimine indirgeme işlemidir. Kökten ayırmanın en güçlü kullanımı, nerede uygulanabileceğine bağlı olarak küçük bir veriyi en tam biçimiyle yeniden tanımlamaya yardımcı olmadaki önemini içerir [43].

**Tablo 1:** Kullanıcı Pazar Sepeti Veri Seti

	siparis_tarihi	siparis_id	urun_id	kategori_kodu	marka	kullanici_id
0	2020-04-24 11:50:39 UTC	22943599320 54530000	15159662235 09080000	electronic s.tablet	samsung	15159160000 00000000
1	2020-04-24 11:50:39 UTC	22943599320 54530000	15159662235 09080000	electronic s.tablet	samsung	15159160000 00000000
2	2020-04-24 14:37:43 UTC	22944440240 58080000	22739483190 57180000	electronic s.audio.he adphone	huawei	15159160000 00000000
3	2020-04-24 14:37:43 UTC	22944440240 58080000	22739483190 57180000	electronic s.audio.he adphone	huawei	15159160000 00000000
4	2020-04-24 19:16:21 UTC	22945842631 54070000	22739483168 17420000	electronic s.audio.mi crophone	karcher	15159160000 00000000

Veri toplama ve temizleme işlemleri uygulanarak bir kullanıcıya ait pazar sepeti Tablo 1’de ifade edilmiştir. Bu veri seti, Python’da yazmış olduğumuz programda yer alan Apriori algoritmasına entegre edilecektir. EK E’de verilerin listelenmesi yer almaktadır.

## 6. SONUÇLAR

### 6.1 Pazar Sepet Analizi

Pazar sepeti analizi, satın alma işlemlerini analiz etmeye ve satış fırsatlarını belirlemeye yardımcı olan etkili bir araçtır. Pazar sepeti analizinin işlemleri, Michael Kechinov'un paylaştığı internet sitesinden bir elektronik alışveriş mağazasına ait gerçek online alışveriş veri seti alınmıştır [40]. Google sunucularında veri seti Python programında çalıştırılmıştır ve birliktelik kuralları belirlenmiştir. Bu gerçek değerler baz alınarak sonuçlar üretilmiştir. Kural oluşturma prosedürü için Apriori algoritması ile temel destek-güven çerçevesi kullanılmıştır. Pazar sepeti analizinin işlenmesi için destek ve güven seviyesi kriterleri araştırmacı tarafından önceden veri setine göre hesaplanmıştır. Minimum destek 0.0045 (%0.0045) ve minimum güven seviyesi 0.7 (%70) olarak hesaplanarak veri analizi yapılmıştır. Elektronik mağazasından gelen tarayıcı verileri toplanmıştır ve mağaza için market sepeti analizi yapılmıştır.

Kullanılan model, işlemlerde hangi öğelerin sıklıkla birlikte görüldüğünü bulmak için verileri analiz etmektedir. Bu prosedür, biri İşlem Kimliği ve diğeri Ürün Adı için sınırlı miktarda değişken gerektirir. Bu nedenle, bu model pazar sepeti araştırmalarında oldukça tercih edilmektedir.

Teorik olarak, her birliktelik kuralı [destek, güven seviyesi] ile karakterize edilir. Birliktelik kuralları tablosu, öğe kümelerindeki çeşitli öğeler arasındaki ilişki yüzdesini gösterir. Bu tabloda oluşturulan birliktelik kuralları iki ölçü ile karakterize edilir. Bunlar, olasılık ve önem. Olasılıklar, bir tüketicinin belirli ürünleri (Bilgisayar, Cep telefonu) satın alması durumunda sonucu (Kulaklık) satın alma şansıdır.

Ürünler arasında oluşan bağımlılık ağı, öğeler arasındaki ilişkilerin grafiksel bir temsilini sağlamaktadır. Sadece satın alınan ürünlerin ne sıklıkta birbirine bağlandığını ve hangi ürünün diğerine bağlı olduğunu gösterir [44].

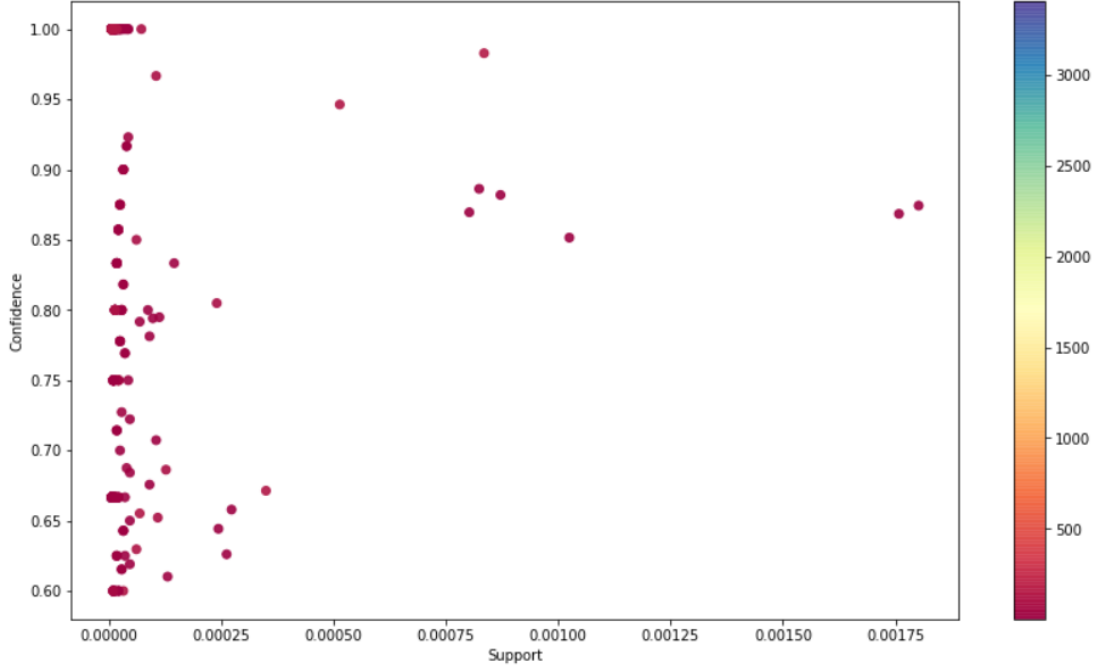
### 6.1.1 Python'da Apriori Algoritmasının Uygulanması

Python'da Apriori algoritmasını uygulamak için algoritma kullanılmadan önce verilerin işlemlere dönüştürülmesi gerekir. Birlikte satın alınan tüm ürünler tek satırda gösterilmektedir. Bu dönüşüm için; Sipariş ID ve Ürün Adı sütunları yeterlidir. Veri seti Covid-19 pandemisi öncesi ve sonrası diye ikiye bölünmüştür. Covid-19 pandemisinin başlangıç tarihi Dünya Sağlık Örgütü'nün ilan ettiği 11 Mart 2020 tarihi belirlenmiştir [45, 46]. Verilen veri seti 1578817 satır içerirken, bu dönüşüm 432,41 saniye sürer ve 1367354 işlem üretilir. Ardından işlemi basitleştirmek için; bu işlem seti sonuc.csv dosyasına yazdırılır. Yazma işlemi 2,32 saniye sürer. Daha sonra bu dosya 185,64 saniyede okunur. Veri yoğunluğu 0,0045'tir. Yoğunluk, satın alınan toplam öge sayısının, o matristeki olası toplam öge sayısına bölünmesiyle elde edilen değeri temsil eder. Apriori algoritmasında destek eşiğini seçmek için yoğunluğu dikkate almak gerekmektedir. İşlemlerin bir özeti olarak, 826571 işlemin yalnızca 1 ürün içermekte olup, 1215 işlem 2 ürün, 738 işlem 3 ürün, 458 işlem 4 ürün, 158 işlem ise 5 ürün içermektedir.

İşlem seti Covid-19 pandemisi öncesi ve sonrası olmak üzere iki bölüm halinde Apriori algoritmasında işlenmektedir. İşlem setini Covid-19 pandemisi öncesi yapılan online alışveriş verileriyle oluşturduktan sonra, Apriori algoritması şu parametrelerle yürütülür: support=0.0045 ve confidence=0.6. Bu parametre değerleri, beklenen çıktı ilişkilendirme kuralları için eşikleri gösterir. 1.32 saniye sürmektedir ve en az %0.0045 desteği ve en az %60 güven seviyesi olan 981 kural döndürmektedir. EK C'de Python programındaki gruplanmalar gösterilmektedir. Kurallar lift değerine göre sıralanır. Daha yüksek bir lift değeri, ilişkinin gücünü gösterir. Şekil 16'da gösterilen Apriori tarafından oluşturulan birliktelik kurallarının bir özeti bulunmaktadır. Sepette 2 ürün bulunan kural sayısı 463'tür. 3 ürün bulunan 281 kural, 4 ürün bulunan 175 kural ve 5 ürün bulunan 55 kural bulunmaktadır.

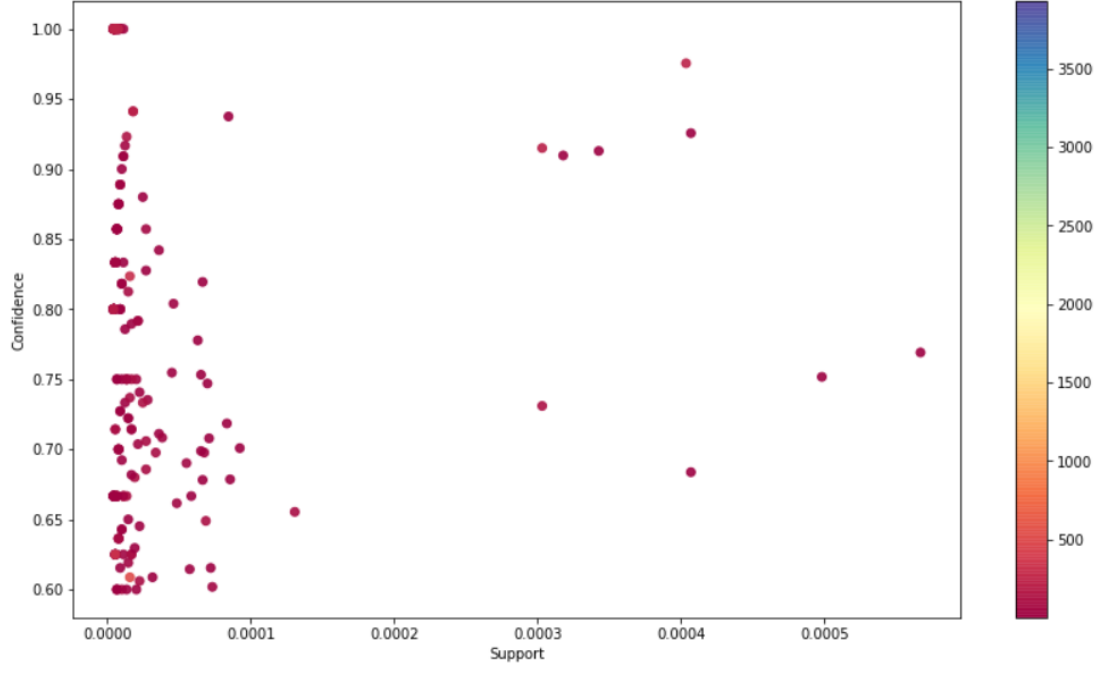
Şekil 16, güven seviyesi, destek ve lift değerlerine göre kuralların özetini göstermektedir. Bu şekil, farklı metrikler arasındaki ilişkiyi göstermektedir. EK F ve G'de Python programında nasıl sonuçlara ulaşıldığı gösterilmektedir. Optimal kuralların destek-güven seviyesi sınırı oluşturduğu gösterilmiştir. Optimal kurallar, desteğin, güven seviyesinin veya her iki ölçümün en üst düzeye çıkarıldığı yerleri gösterir. Grafikte sınır olmasına rağmen 463'ün üzerindeki ürün sayısı ile

kıyaslandığında daha az kuralın olduğu görülmektedir. Bu nedenle Apriori ile farklı güven seviyesi ve destek değerleri daha fazla kural keşfedilmeye çalışılmaktadır.



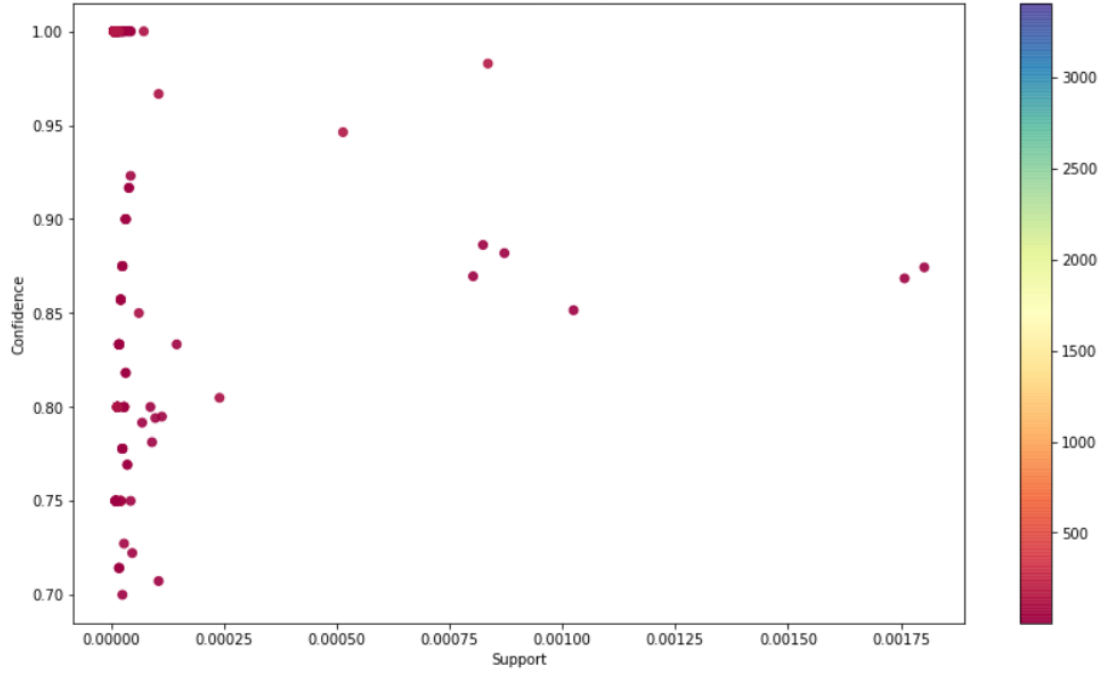
**Şekil 16:** Apriori algoritması tarafından oluşturulan kuralların pandemi öncesi dağılımı (destek=0.0045 ve güven seviyesi =0.6)

Bir sonraki yapılacak analizde işlem setini Covid-19 pandemisi sonrası yapılan online alışveriş verilerini içermektedir. Apriori algoritması şu parametrelerle yürütülür: support=0.0045 ve confidence=0.6. Bu parametre değerleri, beklenen çıktı ilişkilendirme kuralları için eşikleri gösterir. 1.17 saniye sürmektedir ve en az %0.0045 desteği ve en az %60 güven seviyesi olan 530 kural döndürmektedir. Kurallar lift değerine göre sıralanır. Daha yüksek bir lift değeri, ilişkinin gücünü gösterir. Şekil 17’de gösterilen Apriori tarafından oluşturulan birliktelik kurallarının bir özeti bulunmaktadır. Sepette 2 ürün bulunan kural sayısı 255’dir. 3 ürün bulunan 153 kural, 4 ürün bulunan 87 kural ve 5 ürün bulunan 28 kural bulunmaktadır.



**Şekil 17:** Apriori algoritması tarafından oluşturulan kuralların pandemi sonrası dağılımı (destek=0.0045ve güven seviyesi =0.6)

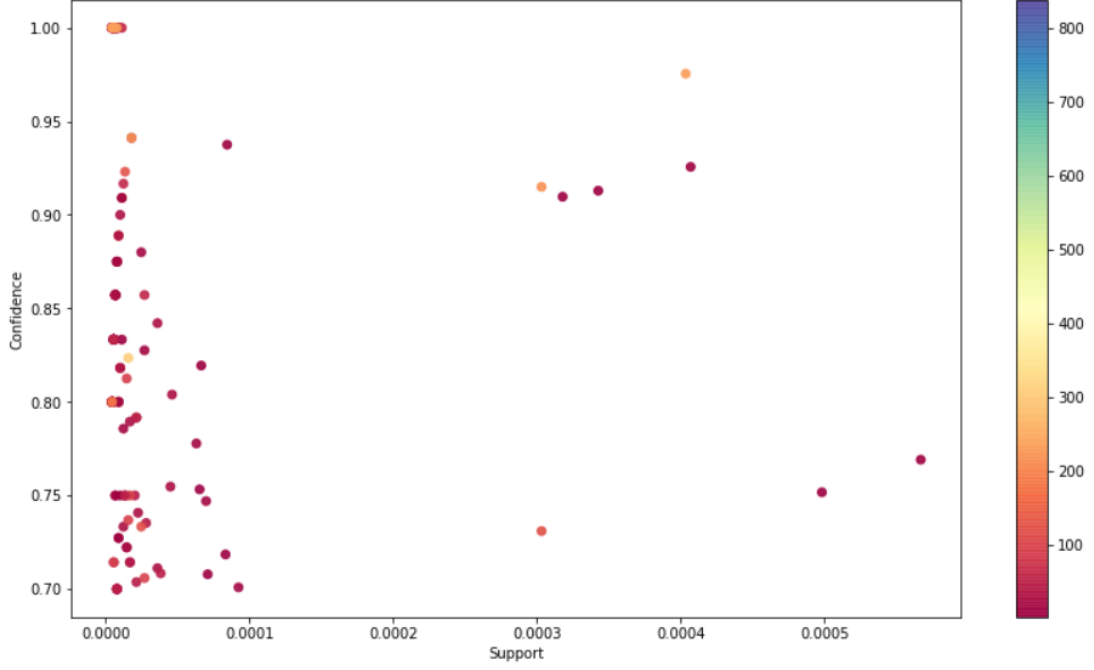
Girilen değerler: destek=0.0045 ve güven seviyesi =0.6. Covid-19 pandemisi öncesi işlem seti 2,08 saniye sürer ve en az %0,0045 desteği ve en az %60 güven seviyesi olan 981 kuralı döndürür. Pandemi sonrası işlem seti 2,08 saniye sürer ve en az %0,0045 desteği ve en az %60 güven seviyesi olan 530 kuralı döndürür. Daha sonra, Apriori'ye girdi olarak destek=0.0045 ve güven seviyesi =0.7 verilir. Bu değerler, 2,16 saniyede 1083 kural üretir.



**Şekil 18:** Apriori algoritması tarafından oluşturulan kuralların pandemi öncesi dağılımı (destek=0.0045ve güven seviyesi =0.7)

Şekil 18’deki işlem seti Covid-19 pandemisi öncesi verilerinin bulunduğu, 712 kuralın destek ve güven seviyesi arasındaki dağılımını göstermektedir. Artan destek ve güven seviyesi değeri, daha mutlak ve daha anlaşılır destek güven seviyesi sınırına neden olur. Şekilde, kuralların çoğunun [0.0001,0.0005] destek aralığına yayıldığı görülmektedir. Sepette 2 ürün bulunan kural sayısı 322’dir. 3 ürün bulunan 202 kural, 4 ürün bulunan 133 kural ve 5 ürün bulunan 50 kural bulunmaktadır.





**Şekil 19:** Apriori algoritması tarafından oluşturulan kuralların pandemi sonrası dağılımı (destek=0.0045 ve güven seviyesi =0.7)

Şekil 19, işlem setini Covid-19 pandemisi sonrası yapılan online alışveriş verilerini içermektedir. Apriori algoritması şu parametrelerle yürütülür: destek=0.0045 ve güven seviyesi=0.7. Bu parametre değerleri, beklenen çıktı ilişkilendirme kuralları için eşikleri gösterir. 1.17 saniye sürmektedir ve en az %0.0045 desteği ve en az %70 güven seviyesi olan 371 kural döndürmektedir. Kurallar lift değerine göre sıralanır. Şekil 19’da gösterilen Apriori tarafından oluşturulan birliktelik kurallarının bir özeti bulunmaktadır. Sepette 2 ürün bulunan kural sayısı 175’dir. 3 ürün bulunan 102 kural, 4 ürün bulunan 63 kural ve 5 ürün bulunan 25 kural bulunmaktadır.

### 6.1.2 Apriori Algoritması ile Elde Edilen Birliktelik Kuralları

Bu analizin amacı, iki veya daha fazla ürünü birbirine bağlayan bir dizi kural oluşturmaktır. Covid-19 pandemisi öncesi ve sonrası yapılan online satışlar arasındaki farklar ortaya konulması amaçlanmıştır. Veri seti iki dönemi de içeren verilere sahiptir. Apriori'nin çıktısını yorumlamak için üç önemli ölçüt vardır; destek (support), güven seviyesi (confidence) ve kaldıraç (lift). Destek, bir öğenin görüldüğü işlemlerin oranıyla ölçülen o öğenin ne kadar popüler olduğunu gösterir. Ayrıca güven seviyesi, sol taraftaki bir öğe satın alındığında kuralın sağ tarafındaki bir öğenin satın alınma

olasılığını ifade eder. Güven seviyesi, sağ taraftaki ögesinin de yer aldığı sol taraftaki ürün ile yapılan işlemlerin oranı ile ölçülür. Bu nedenle, daha yüksek güven seviyesi kuralları, sol taraftaki öğelerin varlığı göz önüne alındığında, sağ taraftaki öğelerin işlemin bir parçası olma olasılığının daha yüksek olduğu kurallardır. Artış, sağ taraftaki öğenin ne kadar popüler olduğunu kontrol ederken, aynı zamanda sol taraftaki öğe satın alındığında sağ taraftaki ögesinin ne kadar satın alınacağını açıklar. Oluşturulan kuralların her biri, tamamlayıcı ürünler için birden fazla artışa sahip olmalıdır. Artış ise ürünler arasındaki ilişkinin ne kadar güçlü olduğunu açıklar. Kaldıraç 1'e eşitse ürünler ilişkili değildir ve 1'den küçükse ürünler ikamedir.

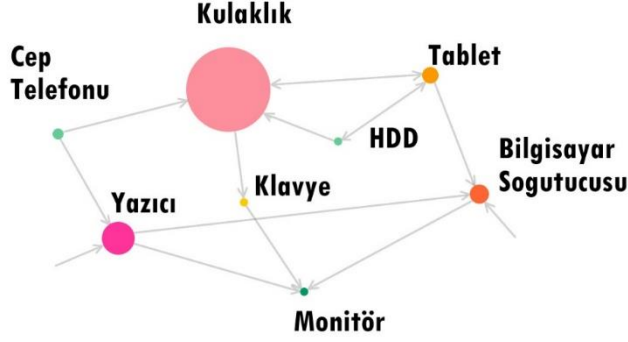
Apriori'nin parametrelerle son yürütmesi: support=0.0045 ve güven seviyesi=0.6. Covid-19 pandemisi öncesi online satın alınan ürünler için 1,32 saniyede 981 kural oluşturur. Yürütme süresi, bu kadar büyük bir veri kümesi için normal kabul edilmektedir. Ayrıca metriklerin tanımına dayalı olarak; destek ve güven seviyesi eşiği normal çalışmaktadır. Veri kümesi yalnızca altı aylık siparişleri (sepetteki ürünleri) içerir. Bu nedenle, veri kümesindekiler gibi elektronik öğeler için bu süre zarfında bir öğenin veya öğe kümesinin oluşma sıklığı düşük olabilir. Destek eşiği (0.0045) ve güven seviyesi eşiği (0.7), desteğin bir öğenin ne kadar popüler olduğunu gösterdiği ve güven seviyesinin, kuralın sağ tarafındaki bir öğenin sol taraftaki öğe olduğunda satın alınma olasılığını ifade ettiği durumlarda kabul edilebilir. Özetle; üretilen kural sayısı, destek ve güven seviyesi eşiği ve yürütme süreleri dikkate alındığında, bu Apriori uygulamasının çıktısı aşağıdaki gibi yorumlanacaktır.

**Tablo 2:** 0.6 ile 0.7 Arasında Güven Seviyesi Olan Pandemi Öncesi Kurallar Tablosu

KURALLAR (Öncül => Sonuç)	Destek	Güven Seviyesi	Kaldıraç
MASAÜSTÜ BİLGİSAYAR, KLAVYE => KULAKLIK	0.00073	0.66667	401.66739
JOYSTICK => TELEVİZYON	0.00019	0.65000	312.46732
KLAVYE, CEP TELEFONU => MONİTÖR	0.00073	0.66666	225.07225
YAZICI => A4 KAĞIT, BİLGİSAYAR SOĞUTUCUSU	0.00255	0.67500	110.94854
TABLET => HDD HARİCİ DİSK	0.01277	0.68627	89.45978
HDD HARİCİ DİSK => RAM	0.00073	0.66666	86.90378
YAZICI => BİLGİSAYAR ÇANTASI, BİLGİSAYAR SOĞUTUCUSU	0.00073	0.66666	84.53222
YAZICI => BİLGİSAYAR SOĞUTUCUSU, MONİTÖR	0.00073	0.66666	84.53222
ROUTER, YAZICI => BİLGİSAYAR SOĞUTUCUSU	0.00073	0.66666	79.94692
TABLET => HDD HARİCİ DİSK, KULAKLIK	0.00182	0.63333	75.29608
İŞLEMÇİ, BUZDOLABI => BİLGİSAYAR SOĞUTUCUSU	0.00073	0.66666	61.00089
HDD HARİCİ DİSK => TABLET	0.00073	0.66666	60.23687
KULAKLIK, HDD HARİCİ DİSK => TABLET	0.00073	0.66666	60.23687
ROUTER => BİLGİSAYAR MASASI	0.00073	0.66666	57.81672
MOUSE, NOTEBOOK, BİLGİSAYAR ÇANTASI => HDD HARİCİ DİSK	0.00109	0.70000	53.05896
HDD HARİCİ DİSK, CEP TELEFONU, TABLET => KULAKLIK	0.00073	0.66666	47.68032

Covid-19 pandemisi öncesi yapılan online alışverişler ile oluşturulan birliktelik kuralları Tablo 2'de gösterilmiştir. Destek eşiği (0.0045) ve güven eşiği (0.6) olduğunda minimum kaldıraç değeri 47.68 ve maksimum değer 401.66 ve ortalama 118.17'dir. Oluşturulan kural kümesinin herhangi bir kuralı için 1'den çok daha büyük bir kaldıraç değeri vardır. Öncül (kuralın solundaki ürünleri) ve sonucun (kuralın sağındaki ürünleri) beklenenden daha sık birlikte görüldüğünü göstermektedir. Öncül oluşumunun bir sonucun ortaya çıkması üzerinde olumlu bir etkiye sahiptir. Bunun bir sonucu olarak, Apriori tarafından oluşturulan tüm kurallar, birlikte satın alınan ürünleri, yani tamamlayıcı ürünleri gösterir. Bu analiz, paket ürünler yerine tamamlayıcı ürünleri bulmayı amaçlamaktadır. Bu nedenle oluşturulan birliktelik kuralları temel olarak güvenlerine dayalı olarak incelenir.

## 8 Kural



Support Boyutu: 0 - 0.12

Lift: 47.68 - 401.66

Şekil 20: 0,6 ile 0,7 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kurallarının Pandemi Öncesi Grafiği

Şekil 20'ye göre, kuralların güvenilirliğinin 0,6 ile 0,7 arasında olduğu artışa dayalı öğeler ve ilk on kuralı içeren grafikdir. Bu grafikte, dairelerin boyutu desteği, rengi ise kaldıraç değerini temsil eder.

Okun giden yönü, kuralın sol tarafındaki üründen kuralın sağ tarafındaki ürüne doğru olduğunu göstermektedir. Bu güven seviyesi aralığında birliktelik kuralları, sol tarafta ağırlıklı olarak "Yazıcı" ve sağ tarafta "Bilgisayar soğutucusu" ürünlerini içermektedir. Grafik, bir "Yazıcı" satın alındığında "Bilgisayar soğutucusu" veya "Tablet" satın alındığında "Hdd harici disk" alındığını göstermektedir.

**Tablo 3:** 0.6 ile 0.7 Arasında Güven Seviyesi Olan Pandemi Sonrası Kurallar Tablosu

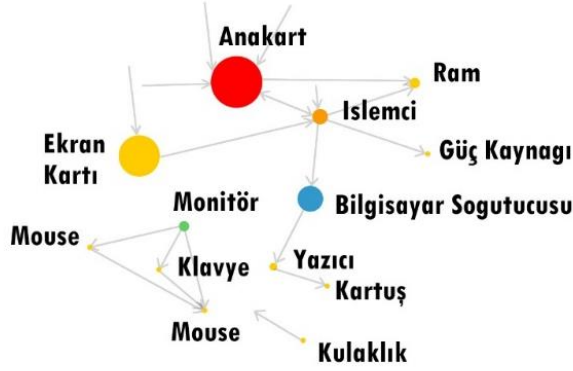
KURALLAR (Öncül => Sonuç)	Destek	Güven Seviyesi	Kaldıraç
ANAKART => İŞLEMÇİ, RAM	0.00068	0.66666	1335.16629
EKRAN KARTI => İŞLEMÇİ	0.00056	0.63333	786.54078
MONİTÖR => MOUSE, KLAVYE	0.00068	0.66666	125.57933
YAZICI => BİLGİSAYAR SOĞUTUCUSU, A4 KAĞIT	0.00135	0.62307	125.35901
MONİTÖR => MOUSE, KLAVYE, NOTEBOOK	0.00056	0.62500	117.73062
İŞLEMÇİ => BİLGİSAYAR SOĞUTUCUSU	0.00045	0.60000	101.90587
İŞLEMÇİ => BİLGİSAYAR SOĞUTUCUSU, ANAKART	0.00045	0.60000	101.90587
HDD HARİCİ DİSK => RAM, NOTEBOOK	0.00056	0.63333	101.78253
YAZICI => RENKLİ KARTUŞ	0.00056	0.61428	97.00400
İŞLEMÇİ => ANAKART, RAM	0.00169	0.65000	95.53675
MONİTÖR, MOUSE => KLAVYE	0.00271	0.60588	92.58893
BİLGİSAYAR SOĞUTUCUSU => YAZICI	0.00079	0.63636	85.96121
İŞLEMÇİ => EKRAN KARTI, GÜÇ KAYNAĞI	0.00045	0.66666	84.92156
HDD HARİCİ DİSK => TABLET	0.00124	0.61666	57.05648
KULAKLIK, KLAVYE => MOUSE	0.00056	0.63333	27.79505
MONİTÖR, KLAVYE, KULAKLIK => MOUSE	0.00045	0.70000	26.68324

Covid-19 pandemisi sonrası yapılan online alışverişler ile oluşturulan birliktelik kuralları Tablo 3'te gösterilmiştir. Destek eşiği (0.0045) ve güven seviyesi eşiği (0.6) olduğunda minimum kaldıraç değeri 26.68 ve maksimum değer 1335.16 ve ortalama 210.21'dir.

Bu güven seviyesi aralığında birliktelik kuralları, sol tarafta ağırlıklı olarak “Anakart” ve sağ tarafta “İşlemci” ürünlerini içermektedir. Grafik, bir “Anakart” satın alındığında “İşlemci” veya “İşlemci” satın alındığında “Bilgisayar soğutucusu” ya da “Anakart” alındığını göstermektedir.

## 13 Kural

Support Boyutu: 0.0045  
Lift: 26.68 - 1335.16



**Şekil 21:** 0,6 ile 0,7 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kurallarının Pandemi Sonrası Grafiği

Covid-19 pandemisi öncesi ve sonrası arasında online alışverişlerin periyodik olarak değişime uğradığı gözlenmektedir. Şekil 21'e göre, elektronik ürünler arasında en çok artışa geçen ürünler bilgisayar parçaları olmuştur. Covid-19 pandemisi sebebiyle evde geçirilen sürenin artmasıyla performans ve güç sorunu olan elektronik ürünlerde de artış görülmüştür.

**Tablo 4:** 0.7 ile 0.8 Arasında Güven Seviyesi Olan Pandemi Öncesi Kurallar Tablosu

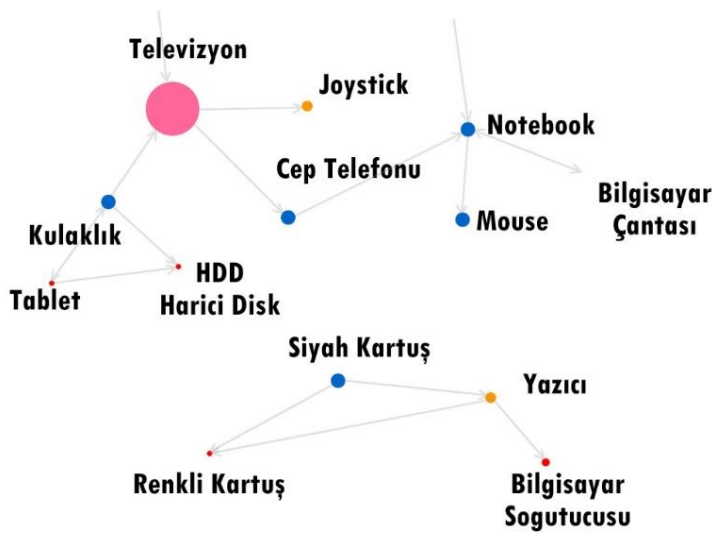
KURALLAR (Öncül => Sonuç)	Destek	Güven Seviyesi	Kaldıraç
JOYSTICK => TELEVİZYON	0.00002	0.75000	312.46733
RENKLİ KARTUŞ=> SİYAH KARTUŞ	0.00084	0.78283	138.09932
SİYAH KARTUŞ=> RENKLİ KARTUŞ	0.00051	0.74631	132.96729
RENKLİ KARTUŞ=> YAZICI, SİYAH KARTUŞ	0.00001	0.80000	112.40923
YAZICI => RENKLİ KARTUŞ, BİLGİSAYAR SOĞUTUCUSU	0.00003	0.77500	110.94854
TABLET => HDD HARİCİ DİSK, KULAKLIK	0.00002	0.73333	75.29609
MOUSE, NOTEBOOK => BİLGİSAYAR ÇANTASI	0.00001	0.75000	53.05897
NOTEBOOK => MOUSE	0.00001	0.75000	29.44343
NOTEBOOK => MOUSE, BİLGİSAYAR ÇANTASI	0.00015	0.73333	25.03543
MOUSE, CEP TELEFONU => NOTEBOOK	0.00007	0.79167	23.78366
TELEVİZYON => CEP TELEFONU	0.00103	0.75152	15.64982
TELEVİZYON => CEP TELEFONU, KULAKLIK	0.00001	0.80000	14.70303
TELEVİZYON => CEP TELEFONU, BİLGİSAYAR ÇANTASI	0.00009	0.80000	14.70303
CEP TELEFONU => NOTEBOOK, MOUSE	0.00001	0.75000	13.78409
CEP TELEFONU => TELEVİZYON	0.00004	0.71667	3.58703

Covid-19 pandemisi öncesi yapılan online alışverişler ile oluşturulan birliktelik kuralları Tablo 4'te gösterilmiştir. Destek eşiği (0.0045) ve güven seviyesi (0.7) olduğunda minimum kaldıraç değeri 3.58 ve maksimum değer 321.46 ve ortalama 71.72'dir.

Tablo 4'te göre 0,7 ile 0,8 arasında güven seviyesi değerine sahip kuralların dağılımı bir önceki güven seviyesi aralığındaki gibidir. Şekil 21'de görüldüğü gibi bu güven seviyesi aralığında yüksek destek değerlerine göre bazı aykırı değerler mevcuttur. Bu aykırı değerler, o dönemdeki bir promosyon kampanyası tarafından oluşturulduğu tespit edilmiştir. Bu nedenle, promosyon kampanyaları nedeniyle bazı paket ürün satışları mevcuttur. Örneğin belirli bir kartuş ürün grubunun hem siyah hem de renkli olanları satın alınmıştır. Bu kuralın daire boyutundan dolayı destek değeri ve daire renginden dolayı kaldıraç değeri yüksektir. Bu nedenle, bu kampanya döneminde her iki kartuş ürünleri birlikte alınmıştır.

Bu güven seviyesi aralığında birliktelik kuralları, sol tarafta ağırlıklı olarak "Televizyon" ve sağ tarafta "Cep Telefonu" ürünlerini içermektedir. Grafik, bir "Televizyon" satın alındığında "Cep Telefonu" veya "Notebook" satın alındığında "Mouse" alındığını göstermektedir.

## 13 Kural



Şekil 22: 0,7 ile 0,8 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kuralların Pandemi Öncesi Grafiği

Şekil 22’de, 0,7 ile 0,8 arasında güven seviyesine sahip pandemi öncesi satılan ürünlerin birbirleriyle olan ilişkileri ortaya konmuştur.

**Tablo 5:** 0.7 ile 0.8 Arasında Güven Seviyesi Olan Pandemi Sonrası Kurallar Tablosu

KURALLAR (Öncül => Sonuç)	Destek	Güven Seviyesi	Kaldıraç
EKRAN KARTI => İŞLEMÇİ	0.00001	0.73333	786.54078
SİYAH KARTUŞ => RENKLİ KARTUŞ	0.00040	0.77548	231.34472
YAZICI => BİLGİSAYAR SOĞUTUCUSU, A4 KAĞIT	0.00001	0.72308	125.35902
İŞLEMÇİ => BİLGİSAYAR SOĞUTUCUSU	0.00000	0.80000	101.90587
İŞLEMÇİ => ANAKART, BİLGİSAYAR SOĞUTUCUSU	0.00000	0.80000	101.90587
HDD HARİCİ DİSK => RAM, NOTEBOOK	0.00001	0.73333	101.78253
YAZICI => BİLGİSAYAR SOĞUTUCUSU, SİYAHKARTUŞ	0.00001	0.71429	97.00400
İŞLEMÇİ => ANAKART, RAM	0.00002	0.75000	95.53676
İŞLEMÇİ => İŞLEMÇİ SOĞUTMA MACUNU	0.00002	0.73684	93.86067
MONİTÖR, MOUSE => KLAVYE	0.00003	0.70588	92.58893
HDD HARİCİ DİSK => TABLET	0.00001	0.71667	57.05649
KULAKLIK, TABLET, KLAVYE => MOUSE	0.00001	0.73333	27.79505
MONİTÖR, KLAVYE, KULAKLIK => MOUSE	0.00000	0.80000	26.68325
TELEVİZYON => CEP TELEFONU	0.00001	0.77500	18.73213
NOTEBOOK => MOUSE, KULAKLIK	0.00001	0.75714	15.88189
CEP TELEFONU => ÇAMAŞIR MAKİNASI	0.00002	0.74118	3.63589
CEP TELEFONU => TELEVİZYON	0.00041	0.72564	3.57587

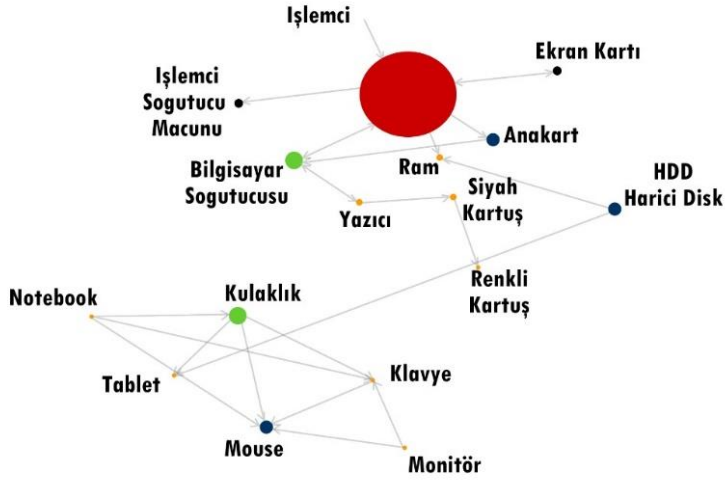
Covid-19 pandemisi sonrası yapılan online alışverişler ile oluşturulan birliktelik kuralları Tablo 5’te gösterilmiştir. Destek eşiği (0.0045) ve güven seviyesi eşiği (0.7) olduğunda minimum kaldıraç değeri 3.57 ve maksimum değer 786.54 ve ortalama 116.54’tür.

Bu güven seviyesi aralığında birliktelik kuralları, sol tarafta ağırlıklı olarak “İşlemci” ile “Yazıcı” ürünleri ve sağ tarafta “Mouse” ile “Bilgisayar Soğutucusu” ürünlerini içermektedir. Grafik, bir “İşlemci” satın alındığında “Anakart” veya “Yazıcı” satın alındığında “Bilgisayar soğutucusu” alındığını göstermektedir.



## 15 Kural

Support Boyutu: 0 - 0.004  
Lift: 3.57 - 786.54



**Şekil 23:** 0,7 ile 0,8 arasında Güven Seviyesine Sahip Oluşturulan Birliktelik Kurallarının Pandemi Sonrası Grafiği

Şekil 23'te, 0,7 ile 0,8 arasında güven seviyesine sahip pandemi öncesi satılan ürünlerin birbirleriyle olan ilişkileri ortaya konmuştur.

Covid-19 pandemisi öncesi ve sonrası arasında online alışverişlerin bir önceki güven seviyesi aralığındaki gibi değişime uğradığı gözlenmektedir. Elektronik ürünler arasında en çok artışa geçen ürünler yine masaüstü bilgisayar parçaları olmuştur. Elektronik ürünlerindeki güç ve performans sorunlarına yönelik online alışverişler yapılmıştır. Pandemi sürecinde tüketicilerin bilgisayar donanım ve performans ihtiyaçlarına olan ilgisiyle kripto paraya ilginin arttığı döneme denk gelmiştir. Bu dolaylı sonuç, insanların alışveriş alışkanlıklarının Bitcoin, Ethereum vb. kripto para işlemlerine yöneldiğini göstermektedir.

Bu kuralların sonuçları şu şekilde sıralanabilir:

- Sepette “Yazıcı”, “Anakart” veya “İşlemci” varsa; “Bilgisayar soğutucusu” sunulan en iyi seçimdir.
- Sepette “İşlemci” varsa; “İşlemci soğutucusu” sunulmalıdır.
- Sepette “Klavye” ve “Kulaklık” birlikte ise bilgisayar için “Mouse” sunulmalıdır.

- Sepette “Monitör” varsa; müşteriler “Mouse” ya da “Klavye” almaktadır, buna göre seçimler sunulmalıdır.
- Sepette “Renkli kartuş” var ise; “Siyah kartuş” sunulmalıdır.
- Sepette “Yazıcı” varsa “Siyah kartuş” teklif edilmelidir.
- Sepette “Televizyon” var ise “Cep telefonu” teklif edilmelidir.

Anakart, İşlemci, Ekran Kartı, Bilgisayar ve İşlemci Soğutucusu gibi ürünlerin pandemi döneminde satın alınması kripto para madenciliği ile ilişkilendirilebilir. Aynı zamanda elektronik donanımdaki belirli bir dönemde artan satışlar güncel bilişim alanındaki yeni teknolojilerin yaygınlaştığına işaret edebilir.

## **6.2 Analiz Sonuçları**

Bu araştırmada, Covid-19 pandemisi döneminde online satış sürecini ve çapraz satışı iyileştirmek için ürünler arasındaki birliktelik kuralları araştırılmıştır. 1578817 sipariş ve yaklaşık 79 ürünü içeren online satış veri seti Apriori Algoritması ile analiz edilmiştir. Apriori Algoritması, veriyi çalışma süresine ve oluşturulan birliktelik kuralı sayısına göre en iyi sonuçları bulmak için parametrelerin farklı değerleriyle yürütülmüştür. 1,32 sn'de 981 kural oluşturan en iyi sonucu aldıktan sonra oluşturulan kurallar ayrı ayrı incelenir. Oluşturulan kurallar önce güven seviyesi değerlerine, ardından kaldıraç değerlerine göre incelenir. Kaldıraç değerlerinden dolayı birliktelik kurallarının yorumlanması, ürünler arasında farklı ilişkilendirme türlerinin olduğunu göstermektedir. Oluşturulan kuralları ilişkilendirme türlerine göre bölen bir artış eşiği vardır. Artış eşiğinden yüksek olan kurallar paket ürünleri, kaldıraç eşiğinden düşük olan kurallar ise tamamlayıcı ürünleri gösterir. Paket ürünler bu projenin parçası değildir çünkü satın alınmaya zorlanan ürünlerdir. Böylece, hangi ürünün hangi ürünle bilinçli olarak satın alındığını tespit etmek için kaldıraç eşiğinden daha az yükseliş gösteren kurallar ayrı ayrı incelenir. Tamamlayıcı ürünleri belirten kurallara göre hangi ürünlerin birlikte satın alındığı keşfedilir ve bu çıktı otomatik bir teklif sistemi geliştirmek için kullanılabilir.

Literatür taraması sonuçlarına göre pazar sepet analizi üzerine yapılan çalışmalardan bazıları aşağıda yer almaktadır.

Gancheva (2013), bir gzellik salonu rn katalogundaki farklı rnlerin birbirleriyle nasıl iliŐki kurduĐunu ve bu iliŐkilerden pazarlama faaliyetleriyle nasıl yararlanılacaĐını gstermektedir. Veri madenciliĐi iliŐkilendirme kuralları, rnlerin birliktelik kurallarını ortaya koymaktadır. Bulgaristan'ın Sofya Őhrinde yer alan bir kozmetik maĐaza zincirinin 4 Őubesinin yer aldıĐı veri seti kullanılmıŐtır. 4 Őubenin yer aldıĐı veri setinde aylık ortalama 19,068 adet veri bulunmaktadır. Apriori algoritması ile birliktelik kurallarına ulaŐılmıŐtır. Her maĐaza iŐin ayrı bir analiz sonucu retilmiŐtir. Birinci maĐaza iŐin bir mŐteri aynı anda gece iŐin yz kremi ile el kremi satın alırsa gndz iŐin yz kremi alma olasılıĐı %92 olduĐu tespit edilmiŐtir. Aynı maĐazada bir mŐteri DuŐ kremi ve SaŐ kremi birlikte satın alırsa, Őampuan da alma olasılıĐı %79'dur. Aynı ihtimal ikinci maĐazada %44 olduĐu tespit edilmiŐtir. MaĐazalar arasında da karŐılaŐtırma yapılabilidiĐi gzlemlenmiŐtir [47].

Ozgormus (2015), marketler iŐin modellenen yerleŐim problemlerini Őzmek iŐin stok yerleŐimlerinin ve optimizasyonun entegre edilmesi amacıyla bu doktora projesini yapmıŐtır. Capitol Migros Store'un market veri seti ve Tabu Search algoritması kullanılmıŐtır. İki ana blmde analiz yapılmıŐtır. Birinci blmde market veri seti ile anlık satın almaları tahmin etmek iŐin metodoloji geliŐtirilmiŐtir. İkinici blmde de market blok yerleŐimlerinin en iyi tasarımı ve dzenlenmesi iŐin etkili Őzm sunulmuŐtur. Bu analizde bulunan en gŐl iliŐki balık ve meyve arasındadır. Balık satın alan mŐterilerin yzde 63' aynı zamanda meyve de satın almıŐtır. MaĐazanın departman konumlarını seŐerek ve departmanları nceden belirlenmiŐ sınırlar iŐinde boyutlandırarak elde edilen geliri en st dzeye Őıkarmak amaŐlanmıŐtır [48].

Yazgan (2016), 6 aylık periyotta 36 rn grubunda 170810 fiŐten toplanan veriler zerinden bir ŐalıŐma yapmıŐtır. Trkiye'de yer alan bir perakende satıŐ firmasının verileri zerinde birliktelik kuralı analizi yapmıŐtır ve rnlerin birbirleriyle olan iliŐkilerini Apriori algoritması ile bulmuŐtur. Analiz ŐalıŐmasını KNIME programında yapmıŐtır. Yapılan araŐtırma sonucunda meyve, sebze ve atıŐtırmalık grubundaki rnlerin daha sık satın alındıĐı tespit edildi. SonuŐlara gre soĐuk iŐecekler ve atıŐtırmalıklar, st rnleri ve meyve ile sebzeler, st rnleri ve atıŐtırmalıklar grubundaki rnler birbirine yakın raflara yerleŐtirildi [49].

Gndz (2017), Samsun'da hizmet veren bir alıŐveriŐ maĐazasından aldıĐı veri setleri zerinden SPSS programında Apriori algoritmasını kullanarak analiz ŐalıŐması yapmıŐ

ve bulduğu birliktelik kurallarıyla çözüm sunmuştur. Veri madenciliği ve Birliktelik kurallarını kullanmıştır. Çalışma sonucunda ev ve mutfak eşyaları ile temizlik ürün kategorilerinde en çok güven seviyesi değerine sahip birliktelik kurallarını tespit etmiştir [26].

Attal (2018), Kuveyt'teki bir süpermarketten, özellikle Al-Shuhada süpermarketinden belirli bir periyottaki POS verilerini incelemiştir. Hatalı kayıtların tespit edilip çıkarılması, aynı işlemde tekrarlanan öğelerin çıkarılması ve Arapça işlemlerin İngilizceye çevrilmesini içeren POS verilerinin temizlenme işlemlerini yaparak Apriori ve Eclat algoritmalarında analizi gerçekleştirmişlerdir. 6426 kaydın yer aldığı projede çips ile çikolata alma ihtimalinin %72, çikolata-meyve-sebze ürünlerinin birlikte alma ihtimalini de %66 olarak tespit etmiştir [50].

Kurniawan (2018), BC UIN Malang süpermarketinin verilerini kullanarak piyasa analizi yapmıştır. Kullanılan yöntem, veri madenciliği tekniği ile mevcut adımların takip edilmesi, Apriori algoritması kullanılarak birliktelik kurallarının belirlenmesi ve pazar sepet analizi uygulama analizinin geliştirilmesidir. %46,69 güven seviyesi değeri ve %1,78 destek değeri ortalamasıyla 30 kural olmuştur. [51].

Karademir (2021), 8230 market verisini hem Weka hem de Python'da kullanarak algoritma performans analizi yapmıştır. Apriori ve Fp-growth algoritmalarını karşılaştırmıştır. Veri madenciliği uygulama tekniklerini kullanarak verilerin hangi algoritmada daha iyi performans verdiği araştırılmıştır. Araştırmanın sonucu olarak en iyi performansı Apriori algoritması vermiştir [52].

Alanlar (2021), Amerika Birleşik Devletleri'nin Massachusetts eyaletinde hizmet veren bir işletmenin 10 aylık sürede toplanan 213.455 adet market alışverişi verisini kullanmıştır. Veri seti için kullanılan tarih aralığı 20.10.2019-10.03.2020'dir. Araştırmanın amacı, veri madenciliğinin temel kavramlarının incelenmesi, Apriori ve GRI algoritmalarını kullanarak müşterilerin Covid-19 pandemisi öncesi ve sonrası dönemlerinde alışveriş alışkanlıklarındaki eğilimin incelenmesidir. Pandemi öncesi sonuçlarına göre deniz ürünleri, tavuk ürünleri ve sebzeler; unlu mamuller, tavuk ürünleri, meyveler ve sebzeler; peynir, kurabiye, kek, sos ve yemeklik malzemeler birlikte satın alındığı tespit edilmiştir. Pandemi sonrası sonuçlarına göre mısır gevrekleri, baharatlar, çeşniler, tavuk ürünleri ve sebzeler; yumurta, atıştırmalıklar,

sirke ve sebzeler; mısır gevrekleri, tavuk ürünleri, kahve ve sebzeler birlikte satın alındığı tespit edilmiştir [53].

Apriori algoritmasının kullanılması, yeni raf düzenlerinin belirlenmesi, birliktelik kurallarının bulunması, pazar sepet analizinin yapılması ve ürün ilişkilerinin ortaya çıkarılması yukarı belirtilen çalışmalar ile yaptığımız çalışmanın benzer yönleridir. Web 3.0 tekniklerinin ilk defa pazar sepet analizi ile kullanılması, Web 3.0 veri toplama tekniklerinin nasıl müşteri bilgilerini topladığının belirlenmesi, Covid-19 pandemisi dönemi öncesi ve sonrası satış politikalarının ortaya konması, elektronik mağazası ürün veri setinin kullanılması ile ortaya çıkan bulguların kripto para madenciliğini desteklemesi ve Covid-19 pandemisi döneminde ürünlerin ilişkisine göre online satış önerilerinin ortaya çıkarılması yukarıda yer alan benzer kaynaklar ile yaptığımız çalışmanın farklı yönleridir.

Bu çalışmanın amacı, Covid-19 pandemisi sebebiyle online alışveriş alanında müşterilerin satın alma davranışlarını inceleyerek bir teknoloji firmasının farklı ürünlerinin birbirleriyle nasıl ilişki kurduğunu ve bu ilişkilerden pazarlama faaliyetleriyle nasıl yararlanılacağını görmektir. Web 3.0 veri toplama tekniklerinin kullanıcı verilerinin elde edilmesinde izlendiği yol, işlemsel verilerden madencilik ilişkilendirme kuralları, ürünlerin birlikte ortaya çıkmalarını sağlayan birliktelik kuralları ve ortak satın alımları hakkında değerli bilgiler sağlayacaktır.

Bir veri kümesinde yer alan ürünler arasında ilk bakışta tanımlanamayan ince ilişkileri tespit etmektir. Ürün satış verilerinin incelenmesi, veri madenciliği ilişkilerinin gerekli olduğu önemli bir örnektir. Pazar sepetindeki farklı ürünler arasındaki ilişkiler, veri madenciliği yöntemlerinin birliktelik kuralları kullanılarak bulunabilir. Bu nedenle bu çalışmanın amaçları şu şekilde ifade edilebilir;

- Bu çalışmaya konu olan Web 3.0 teknikleri ile veri toplama araçları kullanılarak birbiriyle ilişkili ürün veri setlerini bulma,
- Covid-19 pandemisi döneminde ürünlerin ilişkisine göre online satış önerilerini düzenleyerek satış miktarlarını ve satış gelirlerini artırmak,
- Satış süreci için ürünlere uygulanan kampanya ve promosyonlarda bu ilişkilerden yararlanmak.

Bu çalışmayla birlikte Covid-19 pandemi öncesi yapılan online alışverişler ve etkilediği ürün grubu ortaya konmuştur. Covid-19 pandemisiyle birlikte evlerinde karantina sürecine giren insanların yapmış olduğu online elektronik ürünlerdeki alışverişlerin davranışları üzerindeki etkileri bulgularla belirlenmiştir.

### **6.3 Gelecek Çalışmalar ve Öneriler**

Bu araştırmada, Web 3.0 teknikleri ile veri toplama yöntemleri ve sıklıkla kullanılan birliktelik kuralları detaylı olarak ele alınmıştır. Ayrıca birliktelik kuralları için kullanılan temel algoritmalar ayrıntılı olarak verilmiştir. Çeşitli veri toplama yöntemleri özetlenmiş ve her bir yaklaşımın artıları ve eksilerinin vurgulandığı kapsamlı bir literatür taraması verilmiştir. İnceleme bölümünde elde edilen bulgular dikkate alınarak, pazar sepet analizi için en uygun araç olarak Apriori algoritması seçilmiştir.

Dünyada perakende elektronik ürün sektöründe faaliyet gösteren bir mağaza zincirinde veri seti alınmıştır. Tezin amacı, ürünler arasında anlamlı ilişkiler bulmaktır. Bu ilişkiden yararlanılarak Covid-19 pandemisi sürecinde müşterilerin online alışverişlerindeki alışkanlıklarının değişimleri belirlenmiş, etkili reklam kampanyalarının ve promosyonların belirlenmesi ile ürünlerin satış stratejileri planlanmıştır. Bu aynı zamanda perakende mağazanın satış hacmini ve satış gelirini de artırabilmektedir.

Birliktelik kurallarına uygun olarak piyasadan 1 yıl boyunca toplanan veri seti analiz için Google sunucularında bir bilgisayar programına aktarılmıştır. Analiz için Apriori algoritması Python programında çalıştırılmıştır.

Ürünlerin dağılımı incelendikten sonra veri seti Apriori Algoritması ile taranarak birliktelik kuralları bulunmuştur. Ürün grupları üzerinde işlem yapılmış ve elde edilen kurallar tablolar halinde sunulurken ilgili açıklamalar yapılmıştır. Online mağaza yöneticileri, elde edilen ilişkileri kuran bu ürünlerin stoklarını daha sık kontrol edebilir ve uygun seviyede elinde ürün bulundurabilir.

Araştırmanın yapıldığı veri setindeki ürünlerle bu çalışmanın yapılması, belirli bir zaman aralığında veri toplanması, farklı dönemlerde farklı sonuçlar vermesi, belirli bir

zaman diliminde satın almaları etkilemesi gibi bu çalışmanın kısıtları göz ardı edilmemelidir.

Gelecekteki arařtırmalarda ise her müşteri için ayrı bir pazarlama stratejisi oluşturulabilmesi planlanabilir. Bunun için ürün grupları ve müşteri profillerini içeren bir veri seti incelenebilir. Ürünler için elde edilen birliktelik kuralları müşteri profiline göre gruplandırılabilir. Bu sayede farklı ürünlerle ilgilenen her müşteri, farklı kampanyalar oluşturularak ürün satın alması hedeflenir.

## KAYNAKÇA

- [1] **Linoff, G. ve Berry, M. J.** (2011). Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. In *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*, pp. 535-542. Indiana: Wiley Publishing.
- [2] **Liu, C. ve Forsythe, S.** (2010). Online shopping continuance: a focus on post-adoption purchase behaviors. *International Journal of Retail & Distribution Management* 38 (2), pp. 97–114.
- [3] **Blank, G. ve Reisdorf, B.C.** (2012). The Participatory Web. *Information, Communication & Society*, 15:4, pp. 537-554.
- [4] **Zhao, J., Lu, X., Wang, X. ve Ma, Z.** (2015). Web Information Credibility: From Web 1.0 to Web 2.0. *International Journal of u-and e-Service, Science and Technology*, 8(7), 161-172.
- [5] **O'Reilly.** (2005). What is Web2.0. Design patterns and business models for the next generation of software. O'Reilly Media. Adres: [www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-isweb-20.html](http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-isweb-20.html), (05.06.2021).
- [6] **Rudolph, S.** (2011). Foundations of description logics. Reasoning Web.Semantic Technologies for the Web of Data. Springer, pp 76–136.
- [7] **WEB 3.0** (t.y.)., Adres: <https://theemme.wordpress.com/2016/08/23/web-3-0/>, (10.09.2021)
- [8] **Kiesel, J., Kneist, F., Meyer, L., Komlossy, K., Stein, B. ve Potthast, M.** (2020). Web Page Segmentation Revisited: Evaluation Framework and Dataset. *CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pp. 3047–3054. DOI:10.1145/ 3340531.3412782
- [9] **Silva, J. M., Mahfujur, Rahman, A. S. M., ve El Saddik, A.** (2008). Web 3.0: A vision for bridging the gap between real and virtual. *Proceedings of the 1st ACM International Workshop on Communicability Design and Evaluation in Cultural and Ecological Multimedia Systems* (pp. 9—14). New York: ACM. DOI:10.1145/1462039.1462042
- [10] **Mersh, M. ve Muirhead, R.** (2019). What Is Web 3.0 & Why It Matters. Adres: <https://medium.com/fabric-ventures/what-is-web-3-0-why-it-matters-934eb07f3d2b>, (05.06.2021).



- [11] **Zhao, Y., Dong, J. ve Peng, T.** (2009). Ontology Classification for Semantic-Web-Based Software Engineering. In: IEEE Transactions on Services Computing, vol. 2, Issue 4, October 2009, pp. 303-317, DOI:10.1109/TSC.2009.20.
- [12] **Siahaan, A. P. U., Ikhwan, A., ve Aryza, S.** (2018). A novelty of data mining for promoting education based on FP-growth algorithm.
- [13] **El-Seoud, S. A., El-Sofany, H., ve Karam, O.** (2015). Semantic Web Architecture and its Impact on E-learning Systems Development. International Journal of Emerging Technologies in Learning (iJET), 10(5), pp. 29-34. DOI: 10.3991/ijet.v10i5.4754.
- [14] **Dutta, B.** (2006). Semantic Web Based E-learning. DRTC Conference on ICT for Digital Learning Environment 11th – 13th January, 2006, DRTC, Bangalore.
- [15] **Belqasmi, F., Glitho, R., ve Fu, C.** (2011). RESTful web services for service provisioning in next-generation networks: a survey. IEEE Communications Magazine, 49(12), 66-73.
- [16] **Cappuzzo, R., Papotti, P. ve Thirumuruganathan, S.** (2020). Creating embeddings of heterogeneous relational datasets for data integration tasks. SIGMOD, pp. 1335–1349.
- [17] **What is Semantic Data Integration?** (t.y.)  
Adres:  
<https://www.ontotext.com/knowledgehub/fundamentals/semantic-data-integration/>, (10.09.2021)
- [18] **Meng, K., Cui, C. ve Li, H.** (2020). An Ontology Framework for Pile Integrity Evaluation Based on Analytical Methodology, in IEEE Access, vol. 8, pp. 72158-72168, 2020, DOI: 10.1109/ACCESS.2020.2986229.
- [19] **Heflin, J.** (2007). An Introduction to the OWL Web Ontology Language. Lehigh University. National Science Foundation (NSF).
- [20] **Zangeneh, P. ve McCabe, B.** (2020). Ontology-based knowledge representation for industrial megaprojects analytics using linked data and the semantic web. Advanced Engineering Informatics, 46, 101164. DOI: 10.1016/j.aei.2020.101164.
- [21] **Jasper, R. ve Uschold, M.** (1999). A framework for understanding and classifying ontology applications. Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW. pp. 16–21.
- [22] **Van Woensel, W., Casteleyn, S., Paret, E. ve De Troyer, O.** (2011). Transparent mobile querying of online RDF sources using semantic indexing and caching, in: Proceedings of the 12th International Conference on Web Information System Engineering, Springer-Verlag, Sydney, Australia, sayfa 185–198.

- [23] **Allemang, D. ve Hendler, J.** (2011). *Semantic Web for the Working Ontologist, Second Edition: Effective Modeling in RDFS and OWL*, 2nd ed. Morgan Kaufmann.
- [24] **Peelar, S. ve Frost, R.** (2020). A Compositional Semantics for a Wide-Coverage Natural-Language Query Interface to a Semantic Web Triplestore, 2020 IEEE 14th International Conference on Semantic Computing (ICSC), 2020, pp. 257-262, DOI: 10.1109/ICSC.2020.00054.
- [25] **Nebot, V. ve Berlanga, R.** (2012). Building data warehouses with semantic web data. *Decision Support Systems*, 52(4), 853-868.
- [26] **Gündüz, S.** (2017). Veri Madenciliğinde Kullanılan Birliktelik Analizi ve Market Sepet Analizi: Bir Uygulama, Yüksek Lisans Tezi, On Dokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun. Sayfa: 19-27.
- [27] **Karthiyayini, R. ve Balasubramanian, Dr. R.** (2016). Affinity Analysis and Association Rule Mining using Apriori Algorithm in Market Basket Analysis, Volume 6, Issue 10, October 2016, ISSN: 2277 128X.
- [28] **Hwang, Y.** (2019). *Hands-On Data Science for Marketing: Improve your marketing strategies with machine learning using Python and R*. Packt Publishing Ltd.
- [29] **Krishna, B ve Amarawat, G.** (2019). Data Mining in Frequent Pattern Matching Using Improved Apriori Algorithm. *Emerging Technologies in Data Mining and Information Security*. Springer, Singapore.
- [30] **Mukherjee, A., Sundarraj, R. P., ve Dutta, K.** (2015). Online programmatic ad-placement for supply side platform of mobile advertisement: An apriori rule-generation approach. In *Proceedings of the 25th International Conference on Workshop on Information Technology and Systems (WITS'15)*, Dallas, TX. 1--17.
- [31] **Li, C., Li, Y., Yang, Y. ve Deng, Y.** (2020) Analysis of Web Log Mining Based on Association Rule. In: Yang CT., Pei Y., Chang JW. (eds) *Innovative Computing. Lecture Notes in Electrical Engineering*, vol 675. Springer, Singapore. DOI: 10.1007/978-981-15-5959-4\_81
- [32] **Liu, F., Zhang, S., Ge, J., Lu, F., ve Zou, J.** (2016). Agricultural major Courses recommendation using Apriori Algorithm applied in China Open University system, pp. 0–4.
- [33] **Segura-Delgado, A., Gacto, M.J., Alcalá, R., ve Alcalá-Fdez, J.** (2020). Temporal association rule mining: an overview considering the time variable as an integral or implied component. *WIREs Data Mining and Knowledge Discovery*, vol. 10, issue. 4, e1367.
- [34] **Chapman, C., ve Feit, E. M.** (2016). In *R for Marketing Research and Analytics*. Springer, 339-360.
- [35] **Radhakrishna, V., Aljawarneh, S. A., Kumar, P. V., ve Janaki, V.** (2019). ASTRA – A novel interest measure for unearthing latent temporal associations and trends through extending basic Gaussian membership function. *Multimedia Tools and Applications*, 78(4), 4217–4265.

- [36] **Vyas, S. ve Prasad, A.** (2021) Analysis of Algorithms K-Means and Apriori for Data Mining. In: Rathore V.S., Dey N., Piuri V., Babo R., Polkowski Z., Tavares J.M.R.S. (eds) Rising Threats in Expert Applications and Solutions. Advances in Intelligent Systems and Computing, vol 1187. Springer, Singapore. DOI: 10.1007/978-981-15-6014-9\_22
- [37] **Jahangir, I., Basit, A., Hannan, A., ve Javed, S.** (2018). Prediction of Dengue Disease through Data Mining by using Modified Apriori Algorithm. In Proceedings of the 4th ACM International Conference of Computing for Engineering and Sciences (ICCES'18). Association for Computing Machinery, New York, NY, USA, Article 5, 1–4. DOI: 10.1145/3213187.3287612
- [38] **Chee, C. H., Jaafar, J., Aziz, I.A, Hasan, M.H ve Yeoh, W.** (2019). Algorithms for frequent itemset mining: a literature review. *Artif Intell Rev* 52, 2603–2621, Springer. DOI: 10.1007/s10462-018-9629-z
- [39] **Borgelt, C.** (2002). The apriori algorithm for finding association rules. Adres: <http://www.borgelt.net/docs/apriori.pdf>, (05.06.2021).
- [40] **Kechinov, M.** eCommerce behavior data from multi category store. Adres: <https://www.kaggle.com/mkechinov/e-commerce-behavior-data-from-multi-category-store>, (28.11.2020).
- [41] **Abay, K. A., Tafere, K. ve Woldemichael, A.** (2020). Winners and Losers from COVID-19: Global Evidence from Google Search. World Bank Policy Research Working Paper.
- [42] **Singh, P., Singh, S., Mishra, P. ve Garg, R.** (2019). A data structure perspective to the rdd-based apriori algorithm on spark. *International Journal of Information Technology*, pages 1—10.
- [43] **Chen, Y.-L., Tang, K., Shen, R.-J. ve Hu, Y.-H.** (2005). Market Basket Analysis in a Multiple Store Environment. National Central University.
- [44] **Szymkowiak, M., Klimanek, T., ve Józefowski, T.** (2018). Applying Market Basket Analysis To Official Statistical Data. *Econometrics, Advances in Applied Data Analysis*. DOI: 10.15611/eada.2018.1.03
- [45] **Fernandes, N.** (2020). Economic Effects of Coronavirus Outbreak (Covid-19) on the World Economy, Ssrn, Preliminary and Subject to Revisions as New Data is Released, Ver.2.
- [46] **World Health Organization, WHO Coronavirus Disease (COVID-19) Dashboard** (t.y.). Adres: <https://covid19.who.int>, (05.06.2021).
- [47] **Gancheva, V.** (2013). Market Basket Analysis of Beauty Products Doctoral dissertation. Erasmus University Rotterdam. Masters Thesis.
- [48] **Ozgormus, E.** (2015). Optimization of Block Layout for Grocery Stores. PhD thesis. Auburn, Alabama.
- [49] **Yazgan, P.** (2016). Association Rules And Market Basket Analysis: A Case Study In Retail Sector, Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

- [50] **Dalal Al Attal, M. N., AlBaghli, N., Al Muhaimed, N. ve Al, S.** Redesigning a Retail Store Based on Association Rule Mining. Proceedings of the International Conference on Industrial Engineering and Operations Management, July26-2.
- [51] **Kurniawan, F., Hammad, J., Nugroho, S. M., ve Hariadi, M.** (2018). Market Basket Analysis to Identify Customer Behaviors by Way of Transaction Data. Knowledge Engineering and Data Science, 1, 20-25. DOI: 10.17977/um018v1i12018p20-25
- [52] **Karademir, M.** (2021). Veri Madenciliği Market Sepeti Analizi: Apriori ve Fpgrowth Algoritmaları Karşılaştırması, Yüksek Lisans Tezi, İstanbul Medeniyet Üniversitesi Lisansüstü Eğitim Enstitüsü, İstanbul.
- [53] **Alanlar, E.** (2021). Pazar Sepeti Analizi ile Birlikte Kuralların Belirlenmesi: Perakende Sektöründe Covid-19 Etkisi, Yüksek Lisans Tezi, Karabük Üniversitesi Lisansüstü Eğitim Enstitüsü, Karabük.
- [54] **Wang, J, Li, L. ve Cai, H.** (2020). Exploring how deprecated Python library APIs are (not) handled. ESEC/FSE 2020: Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, 233-244. DOI: 10.1145/3368089.3409735
- [55] **Lakshmanan, V. ve Tigani, J.** (2019). Google BigQuery: The Definitive Guide. Publisher: O'Reilly Media, Inc. ISBN: 9781492044468.
- [56] **DataFrame** (t.y.).  
Adres: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.dropna.html>, (21.06.2021).
- [57] **Pandas filter string data based on its string length** (t.y.).  
Adres: <https://stackoverflow.com/questions/19937362/python-pandas-filter-string-data-based-on-its-string-length>, (21.06.2021).
- [58] **Beckham, C., Hall, M. ve Frank, E.** (2016). WekaPyScript: Classification, Regression, and Filter Schemes for WEKA Implemented in Python. Journal of open research software. DOI: 10.5334/jors.108
- [59] **McKinney, W.** (2011). pandas: a foundational Python library for data analysis and statistics. Python High Perform Sci Comput. 2011;14(9):1-9
- [60] **Mondal, K. C., Nandy, B. D. ve Baidya, A.** (2018). A Factual Analysis of Improved Python Implementation of Apriori Algorithm. Methodologies and Application Issues of Contemporary Computing Framework, 1-17. DOI: 10.1007/978-981-13-2345-4\_11
- [61] **Ji, X., Tong, W., Ning, B., Mason, C. E., Kreil, D. P., Labaj, P. P., Chen, G. ve Shi, T.** (2019). QuaPra: efficient transcript assembly and quantification using quadratic programming with Apriori algorithm. Science China Life Sciences, 62(7), 937-946. DOI: 10.1007/s11427-018-9433-3

- [62] **Li, C., Ding, N., Zhang, G. ve Li, L.** (2019). Association Analysis of Serial Cases Based on Apriori Algorithm. Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence, 137–140. DOI: 10.1145/3325730.3326340
- [63] **Zhong, R.** (2020). Huiping Wang. Research of Commonly Used Association Rules Mining Algorithm in Data Mining. IEEE-International Conference on Internet Computing and Information Services-2011.
- [64] **Zong, Y.** (2021). A Study on Cultivating Students' Cultural Self-confidence Under Apriori Algorithm. In: Xu Z., Parizi R.M., Loyola-González O., Zhang X. Cyber Security Intelligence and Analytics. CSIA 2021. Advances in Intelligent Systems and Computing, vol 1342. Springer, Cham. DOI: 10.1007/978-3-030-70042-3\_28
- [65] **Pyplot** (t.y.).  
Adres: [https://matplotlib.org/2.0.2/users/pyplot\\_tutorial.html](https://matplotlib.org/2.0.2/users/pyplot_tutorial.html), (21.06.2021).
- [66] **Hunt, J.** (2019). Graphing with Matplotlib pyplot. In: Advanced Guide to Python 3 Programming. Undergraduate Topics in Computer Science. Springer, Cham. DOI: 10.1007/978-3-030-25943-3\_6
- [67] **Rathore, M. ve Gupta, S.** (2021) A Comparative Analysis of R and Weka for Market Basket Analysis Using Apriori Algorithm. In: Purohit S., Singh Jat D., Poonia R., Kumar S., Hiranwal S. (eds) Proceedings of International Conference on Communication and Computational Technologies. Algorithms for Intelligent Systems. Springer, Singapore. DOI: 10.1007/978-981-15-5077-5\_19

## EKLER

### A. Gerekli kütüphane tanımlamaları [54]

```
# PyDrive kütüphanesinin yüklenmesi ve içeriye aktarılması.
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials

# PyDrive istemcisinin kimliğinin doğrulanması ve
oluşturulması.
auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

file_id = ''
downloaded = drive.CreateFile({'id':file_id})
downloaded.FetchMetadata(fetch_all=True)
downloaded.GetContentFile(download.metadata['title'])
```

### B. Dosyanın verilerinin df değişkeni içine aktarılması [55]

```
import pandas as pd
df = pd.read_csv("/content/kz - Copy.csv")
```

### C. Verilerin Normalizasyonu

#### Gereksiz Alanların Silinmesi [56]

```
del df['user_id']
del df['brand']
del df['category_id']
del df['product_id']
del df['price']
df = df.dropna(axis=0, subset=['category_code'])
```

### Tanımlı olan kategorilerin filtrelenmesi [57]

```
mask = (df['category_code'].str.len() > 6)
df = df.loc[mask]
```

### Tanımsız tarihlerin filtrelenmesi [57]

```
df['event_time'] = pd.to_datetime(df['event_time'])
df = df[df['event_time'] > '1970-01-02']
```

### Pandemi öncesi ve sonrası verilerin değişkenlere aktarılması [57]

```
pandemi_starting_date = "2020-03-11"
df_before_pandemic = df[(df['event_time'] < pandemi_starting_date)]
df_after_pandemic = df[(df['event_time'] > pandemi_starting_date)]
```

### İhtiyaç duyulmayan alanların kaldırılması

```
del df_before_pandemic['event_time']
del df_after_pandemic['event_time']
```

### Verilerin kategorilerine göre gruplanması [58]

```
df_before_pandemic_grouped = (df_before_pandemic.sort_values(['order_id'])
    .groupby('order_id')['category_code']
    .agg(', '.join)
    .reindex(df_before_pandemic['order_id'].unique())
    .reset_index())

df_after_pandemic_grouped = (df_after_pandemic.sort_values(['order_id'])
    .groupby('order_id')['category_code']
    .agg(', '.join)
    .reindex(df_after_pandemic['order_id'].unique())
    .reset_index())
```

```
transactions_before_pandemic = df_before_pandemic_grouped["category_code"].str.split(",", n = 1, expand = False)
transactions_after_pandemic = df_after_pandemic_grouped["category_code"].str.split(",", n = 1, expand = False)
```

### D. Apriori kütüphanesinin eklenmesi [59]

```
!pip install apyori
#apyori kütüphanesi kullanılmıştır
```

### E. Listeleme yöntemi [60]

```
def count(list):
    u = 0
    for i in range(0, len(list)):
        u = u+1
    return u
from apyori import apriori
```

### F. Analiz [61]

#### Confidence 0,6 değerine göre [62]

```
confi = .6
lift = 1
support = 0.0045
transactions_before_pandemic_rules = apriori(transactions_befor
e_pandemic, min_support=support, min_confidence=confi, min_lift=
lift, min_length=2)
transactions_after_pandemic_rules = apriori( transactions_after
_pandemic, min_support=support, min_confidence=confi, min_lift=
lift, min_length=2)
```

#### Confidence 0.7 değerine göre [62]

```
confi = .7
lift = 1
support = 0.0045
transactions_before_pandemic_rules = apriori(transactions_befor
e_pandemic, min_support=support, min_confidence=confi, min_lift=
lift, min_length=2)
transactions_after_pandemic_rules = apriori( transactions_after
_pandemic, min_support=support, min_confidence=confi, min_lift=
lift, min_length=2)
```

#### Kuralların belirlenmesi [63]

```
transactions_before_pandemic_rules = list(transactions_before_p
andemic_rules)
transactions_after_pandemic_rules = list(transactions_after_pan
demic_rules)
```

#### Belirlenen kuralların sayısı [63]

```
print('Before: ',count(transactions_before_pandemic_rules))
print('After: ', count(transactions_after_pandemic_rules))
```

#### CSV uzantılı sonuç üretmek fonksiyon tanımlanması ve kullanılması [64]

```
import csv

def to_excel(list, name):
    with open(name + '.csv', 'w', newline='') as file:
```



```

writer = csv.writer(file)
writer.writerow(["Rule1", "Rule2", "Support", "Confidence", "Lift"])
for item in list:
    pair = item[0]
    items = [x for x in pair]
    writer.writerow([
        items[0],
        items[1],
        str(item[1]),
        str(item[2][0][2]),
        str(item[2][0][3]),
    ])
to_excel(transactions_before_pandemic_rules, 'Before_6')
to_excel(transactions_after_pandemic_rules, 'After_6')

to_excel(transactions_before_pandemic_rules, 'Before_7')
to_excel(transactions_after_pandemic_rules, 'After_7')

```

## G. Grafik olarak gösterilmesi

### Grafik kütüphanesinin eklenmesi [65, 66]

```
import matplotlib.pyplot as plt
```

### Confidence 0.6 için pandemi öncesi [67]

```

rules = pd.read_csv('/content/Before_6.csv')
# Visualizing the rules distribution color mapped by Lift
plt.figure(figsize=(14, 8))
plt.scatter(rules['Support'], rules['Confidence'], c=rules['Lift'], alpha=0.9, cmap='Spectral');
plt.title('Rules distribution color mapped by lift - Before 6')
;
plt.xlabel('Support')
plt.ylabel('Confidence')
plt.colorbar();

```

### Confidence 0.6 için pandemi sonrası [67]

```

rules = pd.read_csv('/content/After_6.csv')
# Visualizing the rules distribution color mapped by Lift
plt.figure(figsize=(14, 8))
plt.scatter(rules['Support'], rules['Confidence'], c=rules['Lift'], alpha=0.9, cmap='Spectral');
plt.title('Rules distribution color mapped by lift - After 6');
plt.xlabel('Support')
plt.ylabel('Confidence')
plt.colorbar();

```

### Confidence 0.7 için pandemi öncesi [67]

```
rules = pd.read_csv('/content/Before_7.csv')
# Visualizing the rules distribution color mapped by Lift
plt.figure(figsize=(14, 8))
plt.scatter(rules['Support'], rules['Confidence'], c=rules['Lift'], alpha=0.9, cmap='Spectral');
plt.title('Rules distribution color mapped by lift - Before 7')
;
plt.xlabel('Support')
plt.ylabel('Confidence')
plt.colorbar();
```

### Confidence 0.7 için pandemi sonrası [67]

```
rules = pd.read_csv('/content/After_7.csv')
# Visualizing the rules distribution color mapped by Lift
plt.figure(figsize=(14, 8))
plt.scatter(rules['Support'], rules['Confidence'], c=rules['Lift'], alpha=0.9, cmap='Spectral');
plt.title('Rules distribution color mapped by lift - After 7');
plt.xlabel('Support')
plt.ylabel('Confidence')
plt.colorbar();
```

## ÖZGEÇMİŞ