



İnternette Yapılan Alışverişlerin Veri Madenciliği Teknikleri ile Analizi ve Depo Süreçlerinin İyileştirilmesi

Alper KİRAZ^{1*}, İrem DELİİSMAİL²

¹Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Sakarya, Türkiye

¹Sakarya Üniversitesi, Yapay Zeka Uygulama ve Araştırma Merkezi, Sakarya, Türkiye

²Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Yüksek Lisans Öğrencisi

Anahtar kelimeler:

E-ticaret
Birliktelik kuralları
Apriori
Carma
Frequent Pattern
Growth

Özet

Bilişim teknolojilerinin gelişmesiyle verilerin toplandığı kaynaklar çeşitlilik kazanmış ve elde edilen verilerin çokluğu nedeniyle anlamlandırılması önem kazanmıştır. Bu çalışmada e-ticaret sektöründe faaliyet gösteren bir firmada en çok satılan ürünlerin analizinin yapılması, beraber satılan ürünlerin tespit edilerek depo içi operasyonel iyileştirilmesi, maliyetlerin azaltılması, çıkış süreçlerinin hızlandırılması amaçlanmıştır. Apriori algoritması birliktelik kurallarından en yaygın olanı olduğu, Frequent Pattern Growth algoritması yüksek performans gösterdiği ve Carma algoritması da daha fazla iyileştirme olanağı sunduğu için tercih edilmiştir. Söz konusu algoritmalar ile firmanın 6 aylık verileri kullanılarak, 10000 adet sipariş içindeki en çok satılan ürünler tespit edilmiştir. Analiz sonuçları 9.ayın 1200 siparişinin satış verileri ile karşılaştırılmış. En iyi sonucu veren Apriori algoritmasına göre belirlenen ilk ürün grubu olan X40-X39 kodlu ürünler tüm siparişlerin %5'ini kapsamaktadır. Siparişlerin %5 ile %1 aralığını kapsayan ilk 10 ürün grubu için, birlikte depo raflarına yerleşmesinin maliyete etkisi zaman etüdü yapılarak analiz edilmiştir. Analiz sonuçları, önerilen sistemin süreç iyileştirmeye olumlu etkisini doğrulamaktadır.

Analysis of Online Shopping with Data Mining Techniques and Improving Warehouse Processes

Keywords:

E-commerce
Association rules
Apriori
Carma
Frequent Pattern
Growth

Abstract

With the development of information technologies, the sources where data collected have diversified and it has gained importance because of the multiplicity of data obtained. This study aims to analyze the top-selling products in a company operating in e-commerce sector, besides identifying the products sold together, improving operational performance in the warehouse, reducing the costs and expediting the release processes. The reasons of preferring algorithms used in the study were; Apriori algorithm was the most common of the association rules, the Frequent Pattern Growth algorithm showed high performance and the Carma algorithm provided further improvement. These algorithms were used to identify the best-selling products in 10000 orders of using the company's 6-month data. The results of the analysis were compared with the sales data of the 9th month, consisting of 1200 order. The X40-X39 coded products, which are the first product group determined according to the Apriori algorithm which gives the best result, cover 5% of all orders. For the first 10 product groups covering 5% to 1% of the orders, the cost effect in bundle of products was analyzed by time study. The analysis results confirm the positive effect of the proposed system on process improvement.

1. GİRİŞ

Teknolojinin hızla değişmesi ve pazardaki rekabet ortamı firmaların süreçlerini değiştirmesine sebep olmaktadır. Firmalar hem maliyetlerini azaltmaya hem de daha kaliteli hizmet vermek için çalışmalar yapmaktadır. Gelecek günler için firmalar operasyon süreçlerini doğru planlamak için farklı yollara başvurmaktadır. Bunlardan biri de veri madenciliği yöntemleridir. Veri madenciliği verilerdeki unsurların arasındaki ilişkiyi incelemeye yarayan bir analiz tekniğidir [1]. Teknolojinin ilerlemesi ve internetin bu kadar önem kazandığı bu günlerde sanal ortamlar her gün artmaktadır ve rekabet ortamı oluşturmaktadır. Yeni bir alışveriş ortamı olan internet son kullanıcıya çeşitli alternatifler sunmaktadır. Geleneksel alışveriş ortamlarından farklı bir hizmet vermektedir. Son kullanıcıya sunulan alternatifler arttıkça rekabet artmaktadır. Bu durumda hizmet veren firmalar, ellerinde bulunan bu elektronik verileri doğru yorumlama yöntemleri aramaktadır.

Birliktelik kuralları kapsamında çoğunlukla Apriori algoritmasının kullanıldığı çalışmalar literatürde yer almaktadır. Genel olarak hizmet sektöründe (çoğunlukla pazar analizi) bu algoritmadan yararlanılmıştır.

Borgelt 2005 yılındaki çalışmasında; FP-Growth Algoritmasını Apriori, Eclat ve Relim algoritmalarıyla karşılaştırılmıştır. C programı üzerinde yapılan bu çalışmada FP-Growth algoritmasının diğer algoritmalara göre daha iyi bir performans gösterdiği sonucuna varılmıştır [2]. Erpolat Türkiye’de otomotiv sektöründe faaliyet gösteren bir yetkili servisin müşterilerine ait satış verilerini Apriori ve FP-Growth Algoritmaları kullanılarak analiz etmiştir. Çalışmada müşterilerin hangi ürünleri beraber satın aldığı belirlenmiş ve karı arttırmak için kampanya ve promosyonlar uygulanmasına ağırlık verilmiştir. Apriori algoritması uygulandığında iki ürün grubu elde edilmiş iken aynı veriler FP-Growth algoritması uygulandığında tek ürün grubu elde edilmiştir. FP-Growth algoritması daha net bir şekilde değerlendirmeyi sağlamıştır [3]. Huang ve arkadaşları Carma algoritmasının verimliliğini değerlendirmek için Apriori algoritması ile karşılaştırmıştır. Destek eşik değerinin aynı olması durumunda Carma algoritması ile Apriori algoritmasının aynı sonuçları verdiği gözlenmiştir. Carma algoritmasının ürettiği grupların Apriori algoritmasının ürettiği veri gruplarının alt kümesi olduğu sonucuna varılmıştır [4]. Lin yaptığı çalışmada, tedarik zinciri yönetiminde önemli problemlerden olan, tedarikçi seçimi problemi için FP-Growth algoritmasını kullanmıştır. Çalışmanın amacı esas ve yedek tedarikçiler seçimindeki karmaşıklığı indirmektedir. FP-Growth algoritmasının tedarikçi elimine edilmesinde ve karmaşıklığı azaltmasında etkili olduğu belirlenmiştir [5]. Mayilvaganan ve arkadaşları 2018 yılında yaptıkları çalışmada Apriori, FP-Growth ve Bulanık FP-Growth algoritmalarını karşılaştırmıştır. Bu çalışmada amaç bir veri tabanının kategorik özneliklere göre yorumlanmasıdır. Çalışma sonucunda Bulanık FP-Growth algoritmasının zaman açısından Apriori ve FP-Growth algoritmalarından daha verimli olduğu ortaya çıkmıştır [6]. Mostafaei ve arkadaşları “Benzin Rasyonunu Analiz Etmede Yeni Yaklaşımlar” adlı çalışmasında 2005 ile 2011 yılları arasındaki karayolu taşımacılığı sektöründe meydana gelen değişiklikleri araştırmışlardır. Benzin tüketiminin durumu Apriori ve Carma algoritmaları incelenmiştir. Benzin tüketiminin doğalgaz, toplu taşıma ve metro kullanımı ile ilişkili olduğu tanımlanmıştır. Sonuçlara göre ulaşım problemlerinin kısa ya da uzun vadede çözümler oluşturması yorumlanmıştır. Ayrıca Apriori ve Carma algoritmalarından elde edilen sonuçları karşılaştırarak, Carma algoritmasının aynı zamanda Apriori algoritmasından elde edilen sonuçlarının aynı olduğu gözlemlenmiştir [7]. Perakende sektöründe Sağın ve arkadaşları 2018 yılında bir firmanın beş buçuk yıllık verilerini iki veri set halinde kullanarak Apriori ve FP-Growth algoritmalarını karşılaştırmıştır. Kurallarının belirlenmesi için hem Apriori hem de FP-Growth algoritmaları ayrı ayrı çalıştırılmıştır. İlk veri setinde FP-Growth algoritması güven değerleri çok yüksek olduğu için Apriori algoritması daha doğru sonucu vermiş, ikinci veri setinde ise her iki algorithmada aynı sonucu vermiştir. Mevsimsel satışlar için kullanılması amaçlanmıştır. Birinci veri seti, ikinci veri setinde de aynı sonuçları gösterdiği için gerçeğe daha yakın olduğu sonucuna varılmıştır. [8]

Bu çalışmada, piyasanın önde gelen markalarının internet siparişlerini son kullanıcıya ulaştırmak için ilgili markaların ürünlerini depolayan bir lojistik firmasının depo süreçlerinin iyileştirilmesi amaçlanmıştır. Ürünlerin stok yerlerinde bulunamaması, siparişlerin zamanında ulaştırılamaması problemine sebep olmaktadır. Bunun için ürünlerin birlikte paketlenerek raflarda tutulmasının, bu problemin çözümündeki etkisi araştırılmaktadır. Çalışmada en çok satılan ürünlerin analizinin yapılarak, beraber satılan ürünleri tespit etmek, depo içi operasyonel iyileştirme yapmak, maliyet azaltmak ve çıkış süreçlerini hızlandırmak amaçlanmıştır. Operasyonel süreçlerin hızlanması için birlikte satılan ürünler belirlenmiştir. Bu ürünlerden, ürün grupları oluşturulmuştur.

Birlikte satılan ürün gruplarının bulunması için de veri madenciliği yöntemlerinden biri olan birliktelik kurallarından yararlanılmıştır. Birliktelik kural yönteminin seçilmesinin sebebi ise mevcut olan durumun incelenmesidir. Birliktelik kurallarından en yaygın olan Apriori algoritması ve diğer algoritmalara göre yüksek performans gösteren, hızlı çalışan FP-Growth algoritması kullanılmıştır. Carma algoritması ise daha fazla iyileştirme olanağı sunduğu ve en iyi çevrimiçi algoritmalarından biri olduğu için seçilmiştir. Çalışmada söz konusu algoritmaların gerçeği temsil kabiliyetleri analiz edilmiştir. Gerçeği temsil kabiliyeti yüksek olarak belirlenen

Apriori algoritmasının sonuçları baz alınarak, zaman etüdü çalışması yapılmış ve birliktelik kurallarına dayalı paketleme sürecindeki iyileştirmenin maliyeti de analiz edilmiştir. Elde edilen analiz sonuçları, depoda yapılacak bu süreç iyileştirmesinin faydasını doğrulamaktadır.

2. MATERYAL ve METOT

Birliktelik kurallarında önemli olan nokta belirli bir zamanda veriler arasındaki birlikteliklerin incelenmesidir [9]. Tekrarı fazla olan nesnelere bulunmasında kullanılan en yaygın yöntem Apriori algoritmasıdır. Bu çalışmada birliktelik kural yönteminin seçilmesinin sebebi ise mevcut olan durumun incelenmesidir. Birliktelik kurallarından en yaygın olan Apriori algoritması ve diğer algoritmalarla göre yüksek performans gösteren, hızlı çalışan FP-Growth algoritması kullanılmıştır. Carma algoritması ise daha fazla iyileştirme olanağı sunduğu ve en iyi çevrimiçi algoritmalarından biri olduğu için seçilmiştir. Önde gelen birliktelik kuralları yönetimi algoritmaları aşağıda açıklanmıştır.

2.1. Apriori Algoritması

Apriori algoritması öğelerin arasındaki ilişkinin nasıl olduğunu analiz eden bir tekniktir. Ürünlerin aralarındaki ilişkinin analizini ölçmek için; destek, güven ve ilgi olmak üzere 3 ana unsur kullanılmıştır

Destek: Öğelerin aynı anda bulunması durumudur. Örneğin; X ve Y ürünleri için $P(X \cap Y)$ olarak olarak düşünülebilir. Veri içindeki X ve Y ürünlerinin bir arada bulunduğu değeri. Bu değer eşik değer olarak kullanılmaktadır. Destek değerlerinin bulunmasında aşağıda listelenen denklemler kullanılır.

$$\text{Destek} = P(X \cap Y) \quad (1)$$

$$\text{Destek}(X) = |X|/|D| \quad (2)$$

$|X| \rightarrow$ Verinin içinde X ürününün bulunma sayısı

$|D| \rightarrow$ Toplam veri sayısı

Güven: $X \Rightarrow Y$ kuralın sol tarafının sağlanması durumunda sağ tarafının da sağlanması olasılığıdır. Yani X ürününün alındığında Y ürününün de $(P(Y|X))$ alınma olasılığıdır. Bunun için aşağıda gösterilen denklem hesaplamada kullanılır.

$$\text{Güven}(Y|X) = \text{Destek}(X \cap Y) / \text{Destek}(X) \quad (3)$$

Güven ölçüsünün bir dezavantajı, bir birliğin öneminin yanlış yorumlanabilmesidir. Bu konuyu bir örnek ile açıklamak gerekirse X ürünü alındığında Y ürününün de alınma olasılığı yüksek olabilir ama bu durum Y ürününün popüler bir ürün olduğunu göstermemektedir. Bunun içinde ilgi değerine bakılmalıdır.

İlgi: Ürünlerin popülerliğini kontrol etmek için hesaplanan değerdir. X ürünü satın alındığında Y ürününün de alınma olasılığının yüksek olması Y ürününün popüler olduğunu göstermemektedir. Bunun için de iki ürün arasında ilgi değerine ihtiyaç duyulmaktadır. İlgi değeri Denklem 4 kullanılarak bulunur.

İlgi değeri=1 \rightarrow X ve Y ürünlerinin arasında bir ilişkinin olmadığını göstermektedir.

İlgi değeri>1 \rightarrow X ürünü alındığında Y ürününün de alınma ihtimalinin yüksek olduğunu göstermektedir.

İlgi değeri<1 \rightarrow X ürünü alındığında Y ürününün alınma olasılığının düşük olduğunu göstermektedir.

$$\text{İlgi}(Y|X) = \text{Güven}(Y|X) / \text{Destek}(X) \quad (4)$$

2.2. Frequent Pattern Growth (FP-Growth) Algoritması

Birliktelik kurallarından biri olan FP-Growth algoritması diğer algoritmalarından yüksek performans göstermektedir. FP-Growth Algoritmasının en büyük avantajı ise büyük veriler için hızlı çalışmasıdır ve sistem kaynaklarının verimli kullanabilmesidir. Tüm verileri Frequent Pattern Tree (FP-Tree) adında sıkıştırılmış bir ağaç

yapısında tutmaktadır. Bu algoritmanın bir özelliği de veri tabanını sadece iki kez taramasıdır. Birinci taramada tüm nesnelerin destek değerlerini bulmaktadır, ikincisinde ise ağaç yapısı oluşturmaktadır. FP-Growth algoritmasında her nesnenin destek değeri bulunur ve kullanıcının vermiş olduğu eşik değerinin altında kalanlar elenmektedir. Elenmeyen değerler büyükten küçüğe sıralanır ve saklanır. Destek değerlerine göre sıralanan nesnelere destek değeri büyük olanlar köke yakın olanlardır. İşlem kayıtlarında olan bir nesnenin ağacın içinde olup olmadığına bakılır. Ağacın içinde yok ise o nesne için yeni bir düğüm oluşturur ve destek sayısını 1 artırır, varsa da olduğu düğümün destek değerini 1 arttırmaktadır. Ağaç oluştuktan sonra da nesnelerin geçtiği dallar belirlenmektedir. Belirlenen dal tek ise nesnelerin kombinasyonudur. Birden fazla dal var ise destek değeri o daldaki minimum destek değerine eşitlenir. Bu şekilde FP-Growth Algoritması böl ve yönet kuralını uygulayarak büyük bir nesne kümesini daha küçük nesne kümelerine bölmektedir. Sonuçta oluşan ağaç yapısı (FP-Tree) asıl veri kümesinden büyük olmamaktadır [3].

FP Growth algoritmasının adımları aşağıdaki gibidir;

1. Minimum destek değeri hesaplanır.
2. Sıklık değerleri bulunur. (Yani toplamda kaç tane satıldıkları bulunur.)
3. Ürünler önem derecelerine göre sıralanır.
4. Siparişin içindeki ürünler önem derecelerine göre sıralanır.
5. FP-Tree çizilir.

2.3. Carma Algoritması

Carma algoritması nesne kümelerinin hesaplama işlemini çevrimiçi olarak yapmaktadır. Algoritma çevrimiçi çalıştığı için kullanıcıya mevcut durumdaki birliktelik kurallarını göstermektedir. Bu algoritmanın artı yönlerinden biri de veri tabanının ilk taramasında herhangi bir işlemde kullanıcı minimum destek ve minimum güven değerlerini değiştirmesini sağlamaktadır. Carma algoritması veritabanını ilk tararken nesne kümelerini oluşturmaktadır. İkinci taramada ise minimum destek seviyesini sağlayan nesne kümelerini saymaktadır. Analiz işlemi 2 adımda bitmektedir, veritabanı en fazla 2 kez taranmaktadır [10,11].

Destek, öncül nesnenin veritabanında geçme sayısının toplam veri sayısına oranıdır. Destek değeri nesne kümesinin öncelik yüzdesidir.

$$\text{Destek} = \frac{\text{Öncül Nesnenin Geçme Sayısı}}{\text{Toplam Veri Sayısı}} \quad (5)$$

Kural Desteği ise nesne grupların toplam veride geçme sıklık oranlarıdır.

$$\text{Kural Desteği} = \frac{P(X \cap Y)}{\text{Toplam Veri Sayısı}} \quad (6)$$

Güven değeri kural destek değerinin öncül destek değerine oranıdır. Öncül nesne alındığında onu izleyen nesnenin alınma olasılığıdır.

$$\text{Güven} = \frac{\text{Kural Desteği}}{\text{Destek}} \quad (7)$$

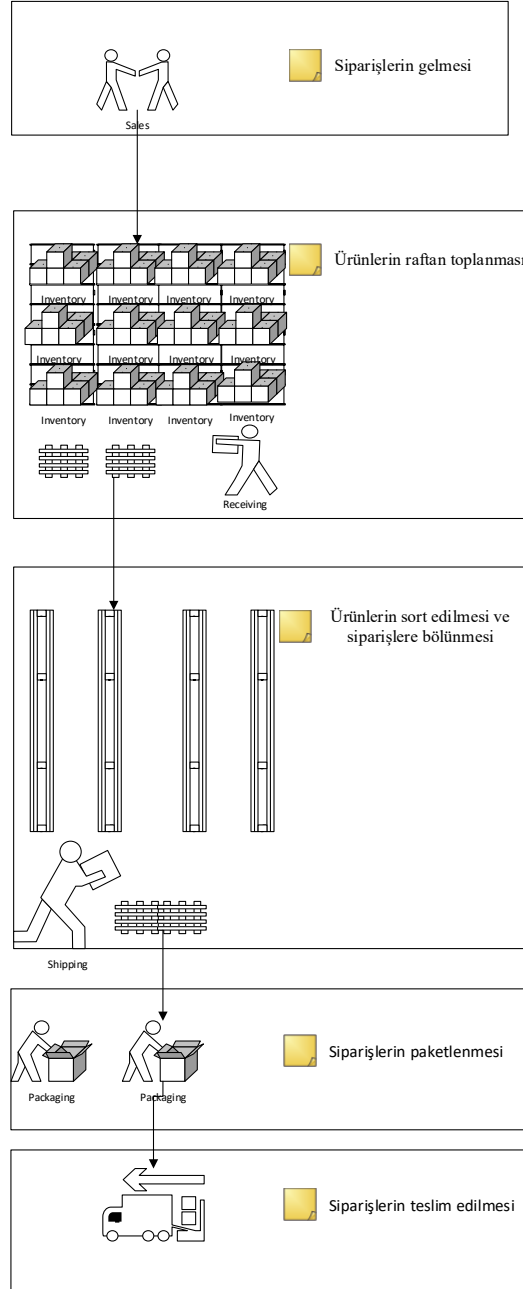
İlgi değeri ise güven değerinin nesne kümesinde izleyen nesnenin veride bulunma sıklığına oranıdır. İlgi değerinin 1'den fazla olması nesne kümesinin doğruluğunu, nesnelerin arasındaki ilişkinin olduğunu göstermektedir.

$$\text{İlgi} = \frac{\text{Güven}}{\text{Destek}(X)} \quad (8)$$

3. BULGULAR ve TARTIŞMA

Bu çalışmada bir lojistik firmasına internet üzerinden gelen siparişlerin analizi yapılmıştır. Firmanın yöneticileri tarafından ikiden fazla ürünler birleştirilmesi ergonomik olarak doğru bulunmamıştır. Bu yüzden ürün grupları ikili olarak kısıtlanmıştır. Deponun mevcut durumda Şekil 1'de görüldüğü gibi internet üzerinden gelen siparişler

ürünlerin raflardan toplanıp paketlenmesi için iş listesine bağlanır ve operatörlere verilir. Ortak bir plastik paletle raflardan toplanan ürünler sipariş bazlı ayrılır. Ayrılan siparişler sırasıyla paketlenir ve son müşteriye iletmek üzere kargoya verilir.



Şekil 1. Lojistik firmasında mevcut operasyonel süreç

Çalışma, piyasanın önde gelen markalarının internet siparişlerini son kullanıcıya ulaştırmak için hizmet veren bir lojistik firmasının yüksek sipariş hacmine sahip müşterisinin depo süreçlerini iyileştirmeyi amaçlamıştır. Ürünlerin stok yerlerinde bulunamaması da probleme oluşturmaktadır. Gelen siparişlerin günlük olarak depodan çıkarılması hedeflenmiştir. Bu sebeple ürünlerin beraber demet halinde raflanarak operasyonel sürecin takip edileceği bir pilot çalışma ile bu durumun incelenecektir. Bu bölümde yukarıda bahsedilen birliktelik kuralları yönetimi algoritmalarına yönelik uygulama çalışmalarından bahsedilmiştir. Firmada uzman kişiler en az 100 siparişte beraber satılan ürün gruplarının dikkate alınması kuralını belirlemiştir.

3.1. Apriori Algoritması Uygulaması

İlk olarak ürün grupların belirlenmesinde birliktelik kuralları algoritmalarından en yaygın olan Apriori algoritması kullanılmıştır. Apriori algoritması için SPSS Modeler programı kullanılmıştır. Satış verilerinden %10'u test verisi olarak kullanılmıştır. Apriori algoritmasının uygulanması için birinci aşamada veri seti makro ile Tablo 1'de sunulduğu üzere düzenlenmiştir. Satırlar siparişleri, sütunlar ise ürünleri temsil etmektedir. Siparişin içinde yer alan ürün için 1, olmayan ürün için 0 olarak tanımlaması yapılmıştır.

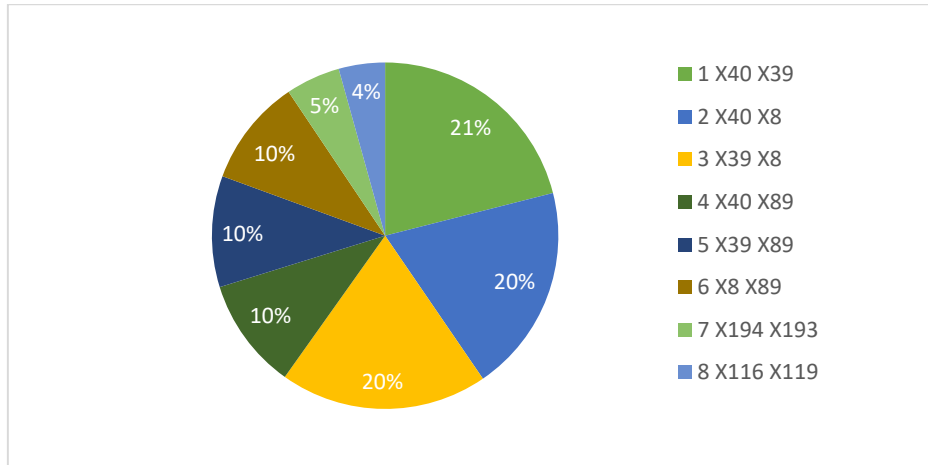
Tablo 1. Apriori algoritması için örnek veri

X40	X3	X8	X89	X117	X194	X193	X217	X116	X20	X37	X460	X119	X3	X183	X216
0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1

10000 tane sipariş içinde minimum destek seviyesi düzeyinde 268 adet ürün grubu bulunmuştur. Tablo 2'de sıralanan ilk sekiz grubu örnek olarak inceleyelim;

Tablo 2. Apriori ilk 8 ürün grupları

Ürün Grupları	Ürün 1	Ürün 2
1	X40	X39
2	X40	X8
3	X39	X8
4	X40	X89
5	X39	X89
6	X8	X89
7	X194	X193
8	X116	X119



Şekil 2. Apriori algoritması ürünlerin beraber satış yüzdesi

Belirlenen kurala göre %21 ile en yüksek ürün grubu X40-X39, %20 ile ikinci grup olarak belirlenen X40 ve X8 ürünleridir (Şekil 2). Bu sonuçlar ile yapılacak bir yorum ise X40 ürününün müşteri portföyünde popüler bir yere sahip olmasıdır.

3.2. FP-Growth Algoritması Uygulaması

FP-Growth algoritması için çalışmada Rapidminer programı kullanılmıştır. Hazırlanan MS Excel'de sonuç olarak tanımlanan değerler seçilmiş değer yani "selected attributes" olarak tanımlanmıştır. Veri setimizin sonuç değerleri ürünlerdir. Bu sebeple de ürünler selected attributes olarak tanımlanmıştır. Minimum destek seviyesi de 0,05 olarak tanımlanmıştır. FP-Growth algoritması ağaç mantığında olduğu için çıkan sonuçlar tekli, ikili, üçlüdür. Apriori algoritmasından farklı sonuçlar elde edilmiştir.

FP-Growth algoritmasına göre en yüksek alınma durumu X3 ürünüdür. İkili olarak ürünleri grupladığımızda da en yüksek birlikte alınma olasılığı ürünler Tablo 3’de bulunmaktadır.

Tablo 3. FP-Growth ilk 10 ikili ürün grupları

Ürün Grupları	Ürün 1	Ürün 2
1	X3	X40
2	X3	X39
3	X3	X8
4	X3	X89
5	X3	X117
6	X3	X194
7	X3	X217
8	X3	X193
9	X3	X116
10	X3	X20

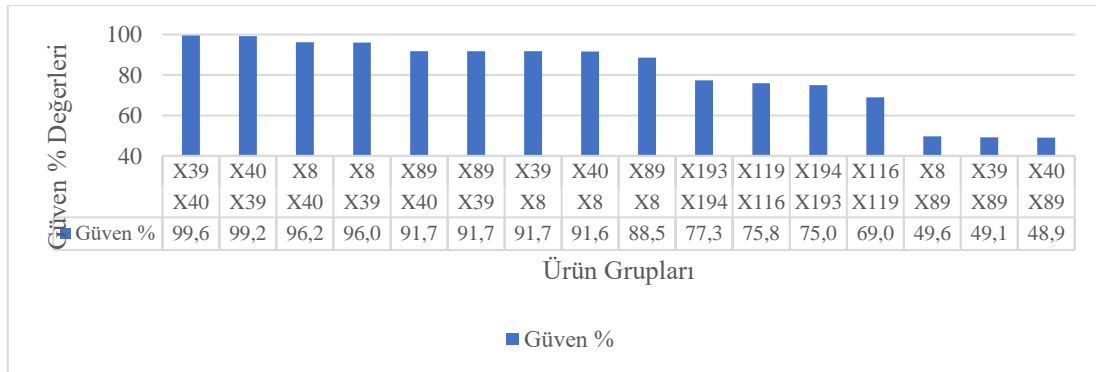
Rapidminer programına göre algoritmanın FP-Tree diyagramı Şekil 3’te sunulmuştur.



Şekil 3. FP-Tree

3.3. Carma Algoritması Uygulaması

Çalışmada üçüncü olarak Carma algoritması ile analiz edilmiştir. Carma algoritmasının uygulanması için birinci aşamada veri seti makro ile Tablo 1’de ki gibi düzenlenmiştir. Carma algoritması için SPSS Modeler programı kullanılmıştır. Carma algoritmasının sonucuna göre en çok beraber satılan ürünler X39-X40,X8-X40,X8-X39 ürünleridir. Ürün gruplarının güven değerleri Şekil 4 ‘de gösterilmiştir.

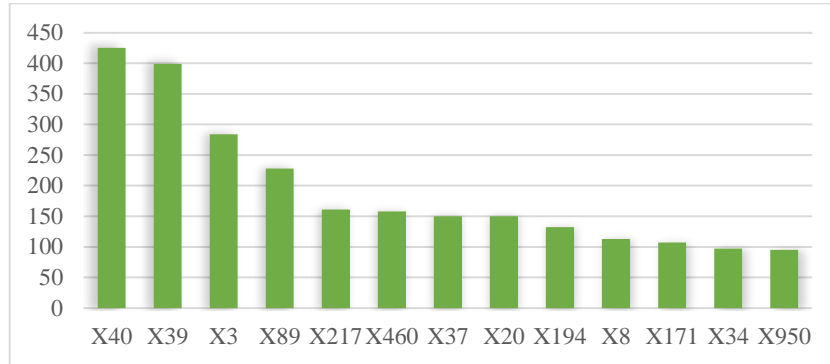


Şekil 4. Carma algoritması ürün gruplarının güven değerleri

Carma ve Apriori algoritmaları aynı eşik güven değeri için hesaplandığından aynı sonuç çıkmaktadır. Bunun için alternatif olarak FP-Growth algoritması ile çalışma yapılmıştır.

3.4. Algoritmaların Karşılaştırılması

Bulunan bu algoritma sonuçlarının karşılaştırılmasının yapılması için yeni bir sipariş verisi üzerinde analiz yapılmıştır. Bu analiz içinde makro kullanılmıştır. Makro ile en çok satılan ürünler ve ikili kombinasyonları üzerinde yorumlama yapılacaktır. Böylece yeni ve gerçek bir veri seti ile hangi algoritmanın sonucuna daha yakın olacağını tespiti yapılacaktır.



Şekil 5. Son satış verisi ürünlerin tekli satış miktarları

Ürünlerin tekli satışlarına bakıldığında en çok satılan ürünler X40-X39-X3-X89-X217-X460 ürünleri olmuştur (Şekil 5). Tablo 4' de ürünlerin satış miktarı listelenmiştir.

Tablo 4. Son satış verisindeki en çok satılan tekli ürünler

Ürünler	Satış Miktarları
X40	425
X39	399
X3	284
X89	228
X217	161
X460	158
X37	150
X20	150
X194	132
X8	113
X171	107

Örneklem olarak satış miktarı 100'den fazla olanlar eklenmiştir. İkili olarak satış yapılan ürünlerin kombinasyonları da Tablo 5 ve Şekil 6'da sunulmaktadır. En çok beraber satılan ürünler X40 ve X39 ürünleridir.

Tablo 5. En çok satılan ikili ürün grupları ve satış miktarları

1.ürün	2.ürün	Beraber Satış Miktarı
X40	X39	391
X39	X89	208
X40	X89	207
X39	X8	103
X40	X8	99
X217	X460	71
X89	X8	64
X460	X20	59
X194	X193	52
X117	X326	50
X568	X567	46
X3	X50	46
X37	X34	44
X37	X183	42

**Şekil 6.** Son satış verisi ürünlerin çiftli satış miktarları

Apriori ve FP-Growth algoritmalarıyla yapılan hesaplamalar ile yeni satış verisine göre en çok satılan tekli ve beraber satılan ilk 10 ürün/ürünler Tablo 6 ve Tablo 7’de gösterilmiştir.

Tablo 6. Son satış verisi-Apriori-FP-Growth algoritmaları tekli ürün karşılaştırması

Son Satış Verisi	Apriori Algoritması	FP-Growth Algoritması
X40	X40	X3
X39	X39	X40
X3	X8	X39
X89	X89	X8
X217	X117	X89
X460	X194	X117
X37	X193	X194
X20	X217	X217
X194	X116	X193
X8	X20	X116

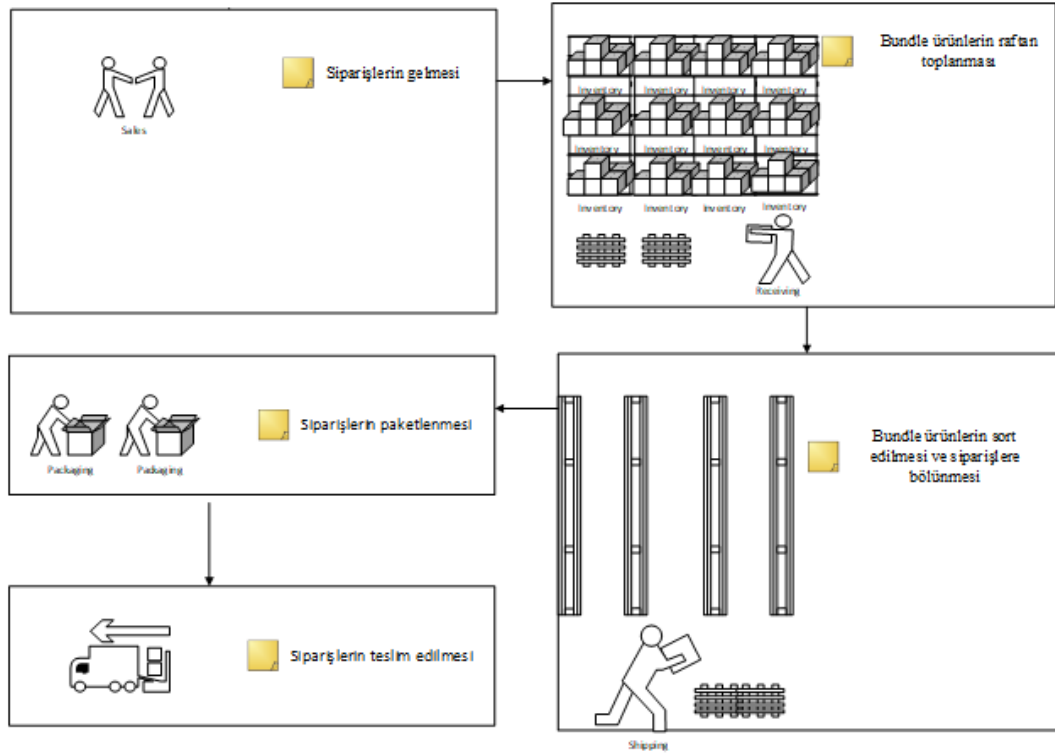
Tablo 7. Son satış verisi-Apriori-FP-Growth algoritmaları çiftli ürün karşılaştırması

Son Satış Verisi		Apriori		FP-Growth	
X40	X39	X40	X39	X3	X40
X39	X89	X40	X8	X3	X39
X40	X89	X39	X8	X3	X8
X39	X8	X40	X89	X3	X89
X40	X8	X39	X89	X3	X117
X217	X460	X8	X89	X3	X194
X89	X8	X194	X193	X3	X217
X460	X20	X116	X119	X3	X193
X194	X193	X117	X114	X3	X116
X117	X326	X117	X326	X3	X20

İncelenen son satış verileri apriori algoritmasını sonuçlarına daha yakındır. İkili ürün satış durumlarında Apriori ile son satış verileri %70 oranında uymaktadır. FP Growth algoritması ile uyumluluk bulunmamaktadır. Bunun yanında tekli olarak ürünleri incelediğimizde hem Apriori hem FP-Growth algoritması %70 oranında uymaktadır. Tekli olarak ürünlerin değerlendirilmesi maliyet açısından bir fayda sağlamayacağından ikili satılan ürünleri baz almak maliyet açısından daha faydalı olacaktır.

3.5. Zaman Etüdü ve Maliyet Analizi

Önerilen depo süreci olarak Şekil 7’de gösterildiği gibi beraber en çok satılan ikili ürünler demet halinde paketlenerek raflarda stoklanacaktır. Siparişler geldiğinde toplama, ayrıştırma ve paketleme süreçlerinde geçerek son kullanıcıya ulaştırılacaktır. Pilot çalışma olduğu için ilk etapta raflardan ürünler tekli toplanıp paketlenerek yerleştirilmiştir. Kabul görülmesi durumunda mal kabul esnasında paketlenerek ürünler konulacaktır. Şekil 1’de bahsedilen mevcut sistemden farklı olarak beraber satılan ürünler demet hâlinde toplanıp paketlenecektir. Tek tek alındığında toplama ve paketleme sürelerinin daha uzun sürdüğü düşünülmektedir ve bunun doğrulanması için de zaman etüdü yöntemine başvurulmuştur.

**Şekil 7.** Lojistik firmasında önerilen operasyonel süreç

Bu yeni durumun incelenmesi için zaman etüdüde örneklem olarak toplamda 2000 ürün içeren siparişler için 23 gözlem yapılmıştır. Yapılan 23 gözlemin sonuçları ve yeterlilik hesaplamaları Tablo 8’de listelenmiştir. Firmada

birlikçiler tarafından tempo %95, tolerans %7,40 olarak belirlenmiştir. İlk olarak her bir süreç için yapılan gözlemlerin yeterli olup olmadığına bakılmıştır. Bunun için de aşağıda gösterilen denklemden yararlanılmıştır.

$$N = \frac{40 * \sqrt{n' * \sum_{i=1}^{n'} x_i^2 - (\sum_{i=1}^{n'} x_i)^2}}{\sum_{i=1}^{n'} x_i} \quad (9)$$

N: Yeterli gözlem sayısı

n': Örneklem olarak yapılan gözlem sayısı

x_i : Yapılan i. gözlemin süresi

x_i^2 :Yapılan i. gözlemin süresi karesi

Tablo 8'de gösterildiği gibi her bir süreç için yapılan 23 gözlem yeterli olduğu belirlenmiştir. Gözlemlerle ilgili verilerin tamamı Tablo 9'da sunulmaktadır. Gözlem sayısının yeterli olduğunun belirlenmesinden sonra süreçlerin normal ve standart zamanları hesaplanmıştır. Bulunan bu standart zamanlar maliyet tahminlenmesinde kullanılmıştır.

Tablo 8. Zaman etüdünde gözlem sayısının yeterlilik kontrolü

Operasyonel Süreçler	Gözlemlerin Toplamı	Gözlemlerin Toplamının Karesi	N
Bundle Ürünlerin Toplama Süreci	429,75	8102,46	14,48
Bundle Ürünlerin Ön Paketleme Süreci	448,43	8799,96	10,42
Bundle Ürünlerin Paketleme Süreci	440,16	8464,61	7,81
Paketleme Süreci	690,13	20928,29	17,04
Toplama Süreci	707,52	22077,29	22,99
Tüm süreçler için $N < n'$ olduğu için gözlemler yeterlidir. ($n'=23$)			

$$\text{Ortalama} = \frac{\sum_{n=1}^N \text{Gözlem (n)}}{N} \quad (10)$$

$$\text{Normal Zaman} = \text{Ortalama} * \text{Tempo} \quad (11)$$

$$\text{Standart Zaman} = \frac{\text{Normal Zaman}}{1 - \text{Tolerans}} \quad (12)$$

Maliyet hesaplamasında kullanılacak olan standart zamanın hesaplanması için yukarıdaki denklemlerden yararlanılmıştır. Her bir sürecin ortalama değeri hesaplanarak normal zaman, normal zaman kullanılarak da maliyette kullanılmak üzere standart zaman hesaplanmıştır.

Tablo 9. Firmanın yeni süreci zaman etüdüleri-23 gözlem

Operasyonel Süreçler	Gözlem No.																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	Gözlem Süreleri																						
Bundle Ürünlerin Toplama Süreci	18	17	16,5	20,3	17,1	17	17,5	18,1	20,1	21,2	18,1	17,1	16	17	19	20	21	17,2	18,2	22	19,2	20	22
Bundle Ürünlerin Ön Paketleme Süreci	18,5	19,4	13,5	19	19,5	20,2	19,5	20,2	21,3	22,2	19,1	18,3	20,2	20,3	19,6	20	21	19,4	18,2	20	19,3	20	20
Bundle Ürünlerin Paketleme Süreci	15,5	16,3	18,5	18,5	19,1	19,5	18,5	20	21,1	20,5	18,5	19	21	19,2	20	18	19	20,1	19,1	21	18,4	19,2	20,1
Paketleme Süreci	30,6	28,4	32,1	31,3	22,4	23,5	29,5	36,6	30,2	26,4	30,2	30,1	32,5	32,6	32	34,2	30,3	31,4	30,5	26,3	28,2	30	31
Toplama Süreci	32,1	38,3	32,3	30,1	38,2	33,1	22,5	27	22,4	32,1	32,2	33,1	32,5	32,6	31,4	28,3	28,3	28,3	30,5	30,3	32,2	30,4	29,6

Operasyonel maliyet gideri olarak operatörün günlük maliyeti ve taşıma maliyetidir. Taşıma maliyeti de sabit olarak günlük 20 TL'dir. Pilot çalışma olarak yapılan 2000 ürün için $2000 \times 1,41 = 2.820,00$ TL kazanç sağlamaktadır. Maliyetler Denklem 13'den yararlanılarak hesaplanmıştır. Detaylı maliyet hesaplaması Tablo 10'da gösterilmiştir.

$$\text{Operasyonel maliyet: } \frac{\text{Standart Zamanı} \times \text{Operasyonel Maliyet Gideri}}{\text{Hedef Verimlilik(\%)} \times \text{Günlük Çalışma Süresi(dk)}} \quad (13)$$

Tablo 10. Maliyet analizi

Maliyet Parametreleri	Toplama Süreci	Paketleme Süreci	Bundle Ürünlerin Toplama Süreci	Bundle Ürünlerin Ön Paketleme Süreci	Bundle Ürünlerin Paketleme Süreci
Standart zaman (dk)	31,56	30,78	19,17	20	19,63
Hedef verimlilik(%)	80	80	80	90	90
Her bir operasyonel için maliyet gideri (TL)	60	44,44	60	44,44	44,44
Günlük çalışma süresi (saat)	8	8	8	8	8
Bir ürün için operasyonel maliyeti (TL)	4,93	3,56	3	2,06	2,02
Toplam maliyet (TL)		8,49		7,08	

E-ticaret firmalarının günlük olarak siparişlerinin depodan çıkması beklenen bir hedeftir. Fakat beklenen hedefin tutturulamaması sebeplerinin başında stokların kaybolması ve sipariş hacminin yüksek olması gelebilir. Firmalar verimliliklerini, karlılıklarını arttırabilmesi için veri madenciliği yöntemlerine başvurulabilir. Veri madenciliği mevcut durumun analizi ya da gelecek tahmini yapılması için kullanılabilir. Örnek olarak bir otomotiv sektöründe bir yetkili servisin karını artırması için müşterilerin satış verileri birliktelik kuralları ile incelenerek firmanın promosyonlara yönelmesine yardımcı olmuştur [3]. Bu çalışmada e-ticaret sektöründe satışı fazla olan alçak raf ürünlerinin, ikili gruplar halinde bulunması için, birliktelik kuralları uygulanmış ve belirlenen ürün grupları ön paket sürecinden geçirilerek raflara yerleştirilmiştir. Bu tarz çalışmalarda ürün tipinin doğru seçilmesi gerekir. Örneğin bir temizlik malzemesi ile bir ev eşyası beraber ön paketleme yapılması doğru olmaz. Bu yüzden de çalışmada alçak raf ürünleri olarak tanımlanan küçük ev eşyaları olarak nitelendirilebilecek (mum, askılık, bebek eşyaları vs.) ürünler seçilmiştir. Apriori, FP-Growth ve Carma algoritmaları sonuçlarının birbirlerine karşı üstünlükleri yerine hangisinin gerçeğe daha yakın sonuçlar verdiği karşılaştırılmıştır. Literatürde genellikle teorik

olarak birbirleri ile olan üstünlükleri karşılaştırılmıştır [1,5,7]. Yapılan çalışmada Apriori ve Carma algoritmaları aynı sonucu vermiştir. Bu çerçevede eşik değerinin etkili olduğu gösterilmiş ve literatürde de 2009 yılında Huang ve arkadaşları tarafından yapılan çalışma ile de kanıtlanmıştır. [4]. Literatürdeki çalışmalar genellikle hizmet sektöründe ya da teorik olarak yapılan çalışmalar olup bu çalışma gerçek bir proje olarak bir firmanın operasyonel iyileştirme için birliktelik kurallarının kullanılabilmesi gösterilmiş ve maliyet/zaman hesaplamaları ile kanıtlanmıştır. Literatürdeki diğer çalışmalar ile ortak olarak müşterilerin satın aldığı ürünlerin baz alınarak sorunların çözülmesidir. Örneğin market analizinde beraber satılan ürünlerin aynı raflara ya da yakın lokasyonlara konulması müşterilerin faydasına olabilecek iken bu çalışmada son kullanıcıyı etkileyen faktör konusunda kesinlik bulunmamaktadır.

4. SONUÇLAR

Teknolojinin hızla değişmesi ve pazardaki rekabet ortamı firmaların süreçlerini değiştirmesine sebep olmaktadır. Firmalar hem maliyetlerini azaltmaya hem de daha kaliteli hizmet vermek için çalışmalar yapmaktadır. Perakende ve e-ticaret hizmeti veren firmalar veri madenciliği yöntemlerini kullanarak satış tahminlemesi yapmaktadır. Yapılan bu tahminler maliyet çalışmalarında kullanılmaktadır. Verimli bir sonuç çıkması sonucunda operasyonel iyileşmeler yapılmaktadır. Firmaya önerilen sistemle maliyetin ürün başına 1,41 TL düşmesi beklenmektedir. Aynı şekilde operasyonel sürecinde toplamda 3,54 dk. azalma olması beklenmektedir. Ürün grupları bulunurken birliktelik kuralları algoritma sonuçları 9. ay satış verileri ile karşılaştırılmıştır. En iyi sonucu veren Apriori algoritması ile oluşan ürün grupları bir arada demet halinde raflara yerleştirildi. Bu önerilen sistem yeni yıla kadar pilot çalışma olarak düşünülecek ve sonrasında kalıcı bir operasyon olarak hayata geçirilecektir. Literatürdeki çalışmalarda genellikle sentetik veri ile çalışılmış olup, bu çalışmada gerçek proje verileri kullanılarak bir firmanın depo sürecini iyileştirmek için birliktelik kurallarının kullanılabilmesini göstermiştir. En iyi sonucu veren algoritmanın doğruluğu zaman etüdü ve maliyete etkisi ile kanıtlanmıştır. Buna ek belirlenen eşik değer algoritmalar için önemi gösterilmiştir. Birliktelik kuralları firmanın diğer bölümlerinde uygulanırsa, operasyon süreçlerinde daha iyi bir performans alınabileceği öngörülmektedir. Önerilen sistem ile raflar arasında bölmeli araçlar toplamada kullanılabilir böylelikle belirli siparişler için ayrıştırma süreci ortadan kalkabilir. Bu analiz sonuçlarına göre en çok satılan ürünlerin daha alt katlara ve yakın raflara konulabilir. Personellerin ergonomik sorunlarının azalması için erişimi daha kolay olan raflara ürünler koyulabilir. E-ticaret sektörü dinamik olduğu için belirli aralıklarla bu analizin yenilenmesi gerektiği unutulmamalıdır.

Kaynaklar

- [1] Akpınar Haldun, “Veri Tabanlarında Bilgi Keşfi Ve Veri Madenciliği”,İ.Ü. İşletme Fakültesi Dergisi, cilt: 29, S:1,p 1-22 2000.
- [2] Borgelt Christian, “An Implementation Of The Fp-Growth Algorithm”, Proceedings Of The 1st International Workshop On Open Source Data Mining: Frequent Pattern Mining Implementations, p 1-5,2015.
- [3] Erpolat Yrd. Doç.Dr Semra, “Otomobil Yetkili Servislerinde Birliktelik Kurallarının Belirlenmesinde Apriori Ve Fp-Growth Algoritmalarının Karşılaştırılması”, Anadolu Üniversitesi Sosyal Bilimler Dergisi, 137,Cilt 12,Sayı1,2012.
- [4] Huang Yuan, Wang Xing, Shia Ben- Chang,” Efficiency And Consistency Study On Carma”, 2009 Fifth International Joint Conference On Inc, Ims And Idc,2009.
- [5] Lin R.H, “Potential Use Of Fp-Growth Algorithm For Identifying Competitive Suppliers In Scm”, Journal Of The Operational Research Society, p 1135-1141,2009.
- [6] Mayılvaganan Dr.M, Kalpanadevi D. “Comparison Of Aprior, Fp-Growth And Fuzzy Fp-Growth Algorithm For Generating Association Rule Mining Of Cognitive Skill”, International Journal Of Engineering Research And General Science,p 2091-2730,2018.
- [7] Mostafaei S., Shakouri H.,Ghodsi R.,”New Approaches To Analyze Gasoline Rationing”, Journal Of AI And Data Mining,Vol 6,No 1, p 177-190, 2017.
- [8] Sağın Ayşe Nur, Ayvaz Berk, “Determination of Association Rules with Market Basket Analysis: An Application in the Retail Sector “, Southeast Europe Journal of Soft Computing Vol.7 No.1, p 10-19, 2018.

- [9] Goebel Michael, Gruenwald Le, “A Survey Of Data Mining And Knowledge Discovery Software Tools”, Acm Sıgkdd Explorations Newsletter, Syf. 20-33, 1999.
- [10] Döşlü Ayhan, “Veri Madenciliğinde Market Sepet Analizi Ve Birlikte Kurallarının Belirlenmesi”, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, İstanbul, 2008.
- [11] Yıldırım Muhammet, “İldeki Kurumlar Arasında Çalışma Performansının Artırılması Ve Veri Madenciliği Tekniklerinin Kullanılması”, Fırat Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, Elazığ, 2006.