



**APRIORI ALGORİTMASININ MICROSOFT EXCEL ORTAMINA EKLENTİ
OLARAK KODLANMASI**

Halil KAYACI

MAYIS 2019

DENİZLİ

**APRİORİ ALGORİTMAİNİN MICROSOFT EXCEL ORTAMINA EKLENTİ
OLARAK KODLANMASI**

Pamukkale Üniversitesi

İktisadi ve İdari Bilimler Fakültesi

Yönetim Bilişim Sistemleri Bölümü

Lisans Tezi

Halil KAYACI

Danışman: Prof. Dr. Selçuk Burak HAŞİLOĞLU

MAYIS 2019

DENİZLİ

ÖNSÖZ VE ERİŞİM

Bu satırları yazma aşamasına gelmek, üniversite maceramızın da son düzlüğüne geldiğimizin göstergesidir. Oysa daha dün gibi hatırimda üniversiteye başlanılan ilk gün, girilen ilk ders, tanışılan ilk hoca, acemilik ve “üniversiteli olmak” hissi... Aslında her tecrübe sahibinin daha o günlerde söylediği gibi oldu her şey. Bu günler çok çabuk geçecek ve nasıl geçtiğini anlamayacaksınız, kıymetini bilin çünkü bir daha gelmeyecek bu zamanlar diyorlardı. Öyle de oldu... Nasıl geldi, nasıl geçti anlayamadım. Şimdilerde ben de alt sınıflara aynı şeyleri söyler oldum tıpkı bana söylendiği gibi. Ama biliyorum ki onlar da bu yazıyı yazma aşamasına gelene kadar hiç anlamayacaklar ne demek istediğimi, tıpkı bir zamanlar benim anlamadığım gibi. Hatta İbrahim AKSEL hocam 4 yıllık bir tatildesiniz hayat asıl bu 4 yıldan sonra başlayacak demişti. Bundan sonraki hayatımızda bu kadar boş zamanı bulamayacağımızı da eklemişti. Bu 4 yılda heybenize ne koyarsanız bundan sonraki yıllarda tercih edeceğiniz yollarda sizlere onlar eşlik edecek demek istiyordu. Hocamın ne kadar haklı olduğunu şimdi çok daha iyi anlıyorum. Bölümümüz ilk açıldığı yıla göre çok gelişti. Ben mezun olduktan sonra eminim ki daha da gelişecek. Biz biraz cefa çektik bizden sonra gelenler sefa sürecekler gibi. Ama onların da belki en büyük dezavantajları sınıfların giderek artan öğrenci nüfusundan dolayı samimi ortamı kaybetmeleri olacaktır. Benim dönemimin en büyük avantajı sayımızın az olması. Bu sayede hem daha iyi bir eğitim aldık hem de hocalarımızla daha samimi bir ortamda geçti üniversite hayatımız. Güzel hatırlayacağım ve sonraki yaşamımda bana farklı bakış açıları kazandıran zamanlar... Bu çalışma ve bu başlık vesilesi ile güzel zamanlarımın mimarı, beni hayata hazırlayan başta Yönetim Bilişim Sistemleri Bölümü Bölüm Başkanı Selçuk Burak HAŞILOĞLU hocama ve tecrübeleriyle hayatıma yön veren diğer hocalarıma çok teşekkür ediyorum. Üzerimde emeğiniz çok büyüktür, bu sorumluluk altında her ne kadar zorlansam da emeğinize layık olmaya çalışacağım.

E-Posta:halilkayaci@gmail.com

Telefon: 0539 469 0957

ÖZET

Gelişen teknoloji ile birlikte büyük miktarda veriler toplanabilmekte ve bu veriler saklanabilmektedir. Fakat bu veriler tek başlarına anlamsızdırlar. Ancak belirli bir amaç ve yöntem ile işlendikten sonra bu veriler anlamlı hale gelir. Bu amaçla veri madenciliğinde birden fazla algoritma geliştirilmiştir. Bu algoritmalarından biri de Apriori algoritmasıdır.

Apriori algoritması, bir veri kümesinde sık tekrarlanan öğelerin birlikteliklerini minimum destek değeri ile analiz eden algoritmadır.

Bu çalışmada, veri madenciliği tekniklerinde kullanılan ve özellikle market sepet analizinde kullanılmak üzere birliktelik kuralları üreten Apriori Algoritmasının, Microsoft Excel ortamına eklenti olarak kodlanması gerçekleştirilmiştir. Amaç, veri madenciliğinde kullanılan tekniklerin daha kolay bir ortamda ve daha kolay bir biçimde çalışmasını sağlamaktır. Bu sayede çalışmanın içerisinde belirtilen biçimdeki veri seti Excel'e aktarılarak, tek bir tıklama ile verilerin analizi sağlanmaktadır. Girilen veriler, istenilen destek değerleri sonucunda ikili olarak analiz edilip bir sonuç tablosu kullanıcıya geri döndürülmektedir. Visual Studio Geliştirme Ortamı(IDE) ve C# programlama dili kullanılarak algoritma adımları gerçekleştirilmiştir. Sonuç olarak geliştirilen bu yazılım sayesinde, karmaşık ve büyük olan veri, kolay ve pratik bir biçimde analiz edilerek, anlamlı bir şekilde kullanıcıya sunulmaktadır.

İÇİNDEKİLER

ÖNSÖZ VE ERİŞİM	i
ÖZET.....	ii
İÇİNDEKİLER	iii
ŞEKİLLER DİZİNİ.....	iv
TABLolar DİZİNİ	v
GİRİŞ	1
VERİ MADENCİLİĞİ.....	3
VERİ MADENCİLİĞİ TEKNİKLERİ	3
Verilerin Sınıflandırılması.....	3
Verilerin Kümelenmesi	4
BİRLİKTELİK KURALI.....	6
Birliktelik Kuralının Matematiksel Gösterimi.....	6
Destek ve Güven Değerleri	6
AMAÇ VE PROBLEMİN TANIMLANMASI.....	7
SİSTEM ANALİZİ	8
Problemin Tanımı.....	8
İstenilenler	8
Gereksinimlerin Tanımlanması	8
Fonksiyonel Gereksinimler	8
YÖNTEM.....	11
Kullanılan Teknolojiler	11
Apriori Algoritması	11
Destek Değeri	11
Güven Değeri	11
İkili İlişkinin Tersine Güven Değeri.....	12
Lift Değeri	12
ÖRNEK UYGULAMA.....	12
YAZILIMIN RAPORU	16
Yazılımın Kaynak Kodları	16
KULLANIM KILAVUZU.....	17
SONUÇ	23
KAYNAKLAR	24
ÖZ GEÇMİŞ	25

ŞEKİLLER DİZİNİ

Şekil 1: Apriori Algoritmasının klasik özet kodu (pseudo code).....	9
Şekil 2: Uygulamanın akış diyagramı	10
Şekil 3: Eklenti Kurulumu Adım 1	17
Şekil 4: Eklenti Kurulumu Adım 2	17
Şekil 5: Eklenti Kurulumu Adım 3	18
Şekil 6: Eklenti Kurulumu Adım 4	18
Şekil 7: Kullanım Kılavuzu Adım 1.....	18
Şekil 8: Kullanım Kılavuzu Adım 2.....	19
Şekil 9: Kullanım Kılavuzu Adım 3.....	19
Şekil 10: Kullanım Kılavuzu Adım 4.....	20
Şekil 11: Kullanım Kılavuzu Adım 5.....	20
Şekil 12: Kullanım Kılavuzu Adım 6.....	21
Şekil 13: Kullanım Kılavuzu Adım 7.....	21
Şekil 14: Kullanım Kılavuzu Adım 8.....	22
Şekil 15: Kullanım Kılavuzu Adım 9.....	22

TABLÖLAR DİZİNİ

Tablo 1: Örnek Uygulama Veri Seti	12
Tablo 2: Ürünlerin Destek Değerleri	13
Tablo 3: İkili İlişkilerin Destek Değerleri	13
Tablo 4: İkili İlişkilerin Güven Değerleri	13
Tablo 5: İkili İlişkilerin Tersine Güven Değerleri	14
Tablo 6: İkili İlişkilerin Lift Değerleri	14
Tablo 7: Örnek Uygulamanın Sonuçları	14

GİRİŞ

Büyük bir hızla değişen ve gelişen teknoloji ile birlikte birçok sektörde olduğu gibi bilişim sektöründe de olağanüstü gelişmeler meydana gelmektedir. Bu gelişim, bilgisayar sistemlerinin güçleri ve kapasiteleri konusunda oldukça olumlu bir etkiye yol açmıştır. Sistemdeki bu artışla beraber teknolojinin ulaşılabilirliği ve maliyetleri konusunda da azalmalar meydana gelmektedir. Ayrıca güçleri ve kapasiteleri artan bilgisayarların fiziki boyutlarında da ciddi oranda azalmalar meydana gelmektedir. Bunun neticesinde bilgi sistemleri büyük verileri daha kısa sürelerde işleyip, istemciye sunup ve bu verileri depolayabilmektedirler. Maliyetlerinde meydana gelen azalma neticesinde de bilgisayarlar kullanımlarının yaygınlığı artmıştır. Meydana gelen bu yaygınlaşma, sayısal olarak verilerin bilgisayar ortamında toplanmasına, saklanmasına ve kullanıcılarının hizmetine sunulmasına olanak sağlamıştır.

Gündelik yaşamda kullanılan birçok araç ve gereç de bu gelişmelerden etkilenmiş ve nitelik bakımından oldukça gelişmiş aygıtlara dönüşmüşlerdir. Günümüzde bu gelişmiş aygıtlar ile birçok verinin belirlenen bir formatta saklanması mümkün hale gelmiştir. Örneğin eski tip yazarkasalar ile bir marketten yapılan alışverişlerde sadece temel matematik işlemleri yapılabilirken, güncel teknoloji ile donatılmış ve satış işlemlerinde kullanılan terminaller ile yapılan alışveriş işlemlerinin bütün detayları veri olarak depolanıp, tekrar kullanılabilir hale gelmiştir. Bu işlem; satılan ürünlerin, zaman içindeki hareketleri, bu zaman zarfı içerisindeki müşteri hareketliliklerini takip edilebilir hale getirmektedir. Böylece işletme sahibi hangi zamanlarda, hangi müşterilere, hangi ürünlerle kampanya yapabileceği bilgisine sahip olabilmektedir. Doğru yapılan kampanya ve avantajlı satışlar neticesinde işletme, sadık müşteri kitlesine ulaşabilir veya yeni bir müşteri kitlesi kazanabilir.

İşlenmemiş, ham haldeki veri tek başına bir anlam ifade edemeyip, değersiz olabilir. Fakat bir amaca uygun olarak işlenip anlamlı hale getirilen veri, artık değersiz olmaktan çıkıp, anlamlı bir ifade haline gelerek bilgiye dönüşür. Veriyi bilgiye çevirmeye veri analizi denmektedir(Akpınar, 2000: 1 – 22).

Veri analizi yapılarak birçok konuda anlamlı bilgilere ulaşabilmek oldukça kolaydır. Bu analiz yapılması ve büyük verinin anlamlandırılması için analizlerin yazılımlarla desteklenmesi gerekmektedir. Tam da bu noktada veri madenciliği teknikleri devreye girmektedir. Veri madenciliği teknikleri büyük veri içinden gelecek ile ilgili

tahmin yapmamızı ve bu verinin anlamlandırılması sağlayan bağıntı ve kuralların bilgisayar programları kullanarak aranmasıdır(Akpınar, 2000: 1 – 22).

Apriori algoritması, veri madenciliği tekniklerinden biridir. Bu algoritma genel olarak market sepeti analizleri için kullanılsa da oldukça yaygın bir kullanım alanı da mevcuttur. Birçok en iyileme işleminde de apriori algoritması kullanılmaktadır. Apriori algoritması, market sepeti analizlerinde ürünlerin arasındaki birliktelik ilişkilerini bulmak için kullanılmaktadır. Veri madenciliği tekniklerinin kullanımı yaygınlaştıkça bu tekniklerden faydalanarak verileri anlamlı hale getirmeyi amaçlayan birçok bilgisayar yazılımı da geliştirilmiştir. Fakat geliştirilen bu yazılımlar biraz kullanıcı deneyimi ve bilgisi gerektirmektedir. Daha önce bu bilgiye sahip olmayan bir kullanıcının bu yazılımlar ile veri analizi yapması kısmen zor bir durumdur. İşte bu durum, bu çalışmanın neticesinde ortaya çıkan yazılım fikrinin ortaya çıkmasına sebep olmuştur. Analiz edilecek veriye sahip fakat veri analizi konusunda geliştirilen yazılımları kullanabilecek bilgiye sahip olmayan kullanıcılar hedeflenmiştir.

Hedeflenen çalışma neticesinde Microsoft Excel ortamı için bir eklenti kodlanmıştır. Böylece kullanıcının daha aşina olduğu bir ortamda, kullanıcıya daha basit bir kullanım ile verisini analiz etme olanağı sağlanmış olmaktadır. Basit bir kurulum aşaması sonrasında eklentiye yükleyen kullanıcı, çalışmada belirlenen formatta verisini Excel ortamına aktararak analizini gerçekleştirebilecektir. Eklentiye çalıştırmak için tek bir butona basması yeterli olacaktır. Nihai olarak kullanıcının girdi olarak vermiş olduğu veriden, analiz süreci bittikten sonra bir tablo olarak ürünlerin ikili birliktelik analizleri kullanıcıya sunulacaktır.

VERİ MADENCİLİĞİ

Gelişen teknoloji ile birlikte bilgi sistemleri hemen her sektörde hayatımızın bir parçası haline gelmiştir. Bu sayede, yaptığımız her işlem kayıt altına alınıp, depolanmaya hatta bu işlemler paylaşımına sunulmaya başlandı. Hemen her sektörde olduğu gibi bu gelişme ticaret sektöründe de etkisini gösterdi. Yapılan alışverişlerde ürünlerin ve müşterilerin hareketlilikleri de kayıt altına alınmaya başlandı. Trafik kontrol ve denetleme kameraları ile geçiş yollarımız dahi kayıt altında tutulabiliyor. Bu gibi verilere erişimin kolay olması bir avantaj olsa da bu verilerin tek başlarına bir anlam ifade etmemesi de olağandır. Tıpkı işlenmeden önce bir anlam ifade etmeyen fakat işlendikten sonra değer ve anlam kazanan maden cevherleri gibi.

Veri madenciliği tam da bu noktada devreye girerek anlamsız verinin işlenerek anlamlı bir hale getirilmesinde, yani verinin bilgiye dönüştürülmesinde bizlere yardım etmektedir. Teknolojik gelişmeler ve buna paralel olarak gelişen veri toplama ve saklama işlemleri, veri madenciliğinin de hızla gelişmesine neden olmuştur.

Veri madenciliği, temel olarak bilgisayar destekli bir bilgi çözümleme işlemidir(Nazife, 2003). Veri madenciliği daha önceden bilinmeyen geçerli ve uygulanabilir bilgilerin geniş veri tabanlarından elde edilmesi ve bu bilgilerin işletme kararları verirken kullanılmasıdır(Silahtaroglu, 2014). Veri madenciliği, ayrı sorgular vererek büyük miktarda olan veriden yararlı bilgi, desenler ve eğilimler (genelde önceden belli olmayan) çıkarabilmektir(Thuraisingham, 2003: 34 – 35).

VERİ MADENCİLİĞİ TEKNİKLERİ

Veri madenciliği yöntemleri, kullanılacak konu ve veri türüne göre farklılık göstermektedir.

Bu yöntemler işlevlerine göre 3 temel başlık altında gruplanabilirler. Bu başlıklar;

- Verilerin sınıflandırılması
- Verilerin kümelenmesi
- Birliktelik kurallarının çıkarılması

Bu başlıklar kendi içerlerinde çok farklı algoritmalar içermektedirler.

Verilerin Sınıflandırılması

Analiz için kullanılan verilerin daha önceden belirlenen sonuçlara uygun olarak ayrıştırılmasıdır. Çok bilinen veri madenciliği tekniklerinden olan bu teknik kategorize

sonuçları tahmin etmek için kullanılmaktadır. Bu yöntem için önceden bir durum-sonuç ilişkisi ve bu durumlar neticesinde değişkenlerin gösterdiği değişkenliklerin bilinmesi gereklidir. Analiz neticesinde veriler belirli bir oran ve bir sonuç olarak geriye döndürülür. Sonuçlar bir oran ile döndüğünden dolayı diğer olayların olma ihtimalleri de bilinir.

Verilerin Kümelenmesi

Analiz için kullanılacak olan veri kümesinde birbirlerine benzeyen elemanlarından yeni bir küme oluşturulmasıdır. Veri kümelemesinin amacı birbirine benzeyen fakat özellik bakımından birbirinden çok farklı olan kümelerin bulunmasını sağlamaktır. Böylece büyük veri yığınının amaca hizmet edecek verilerin ayıklanması sağlanmış olacaktır. Kümeleme algoritmalarının temeli; oluşturulacak olan küme içerisindeki elemanlarının benzerliklerinin en çok olmasına fakat oluşturulan kümenin diğer kümeler ile olan benzerlik oranlarının en az olmasına dayanır.

Verilerin kümelenmesi ile verilerin sınıflandırılması arasında ki en önemli fark, kümelendirmede tanımlı değerlerin bulunmamasıdır. Sınıflandırma, daha önceden tanımlanmış değişkenler ve bu değişkenlerin gösterdikleri değişkenlikler ile yapılır.

Veri madenciliğinde verilerin kümelenmesi için farklı algoritmalar kullanılmaktadır. Başlıca kümeleme yöntemleri:

- Hiyerarşik Yöntemler
- Model Tabanlı Yöntemler
- Bölümleme Yöntemleri
- Izgara Tabanlı Yöntemler
- Yoğunluk Tabanlı Yöntemler
- Birliktelik Kurallarının Çıkarılması

Birliktelik kuralları, bir veri setindeki veriler arasındaki ilişkileri tanımlamayı amaçlayan model olarak tanımlanabilir.

Birliktelik kuralları anlaşılabilir ve basit olduğundan finans, iletişim, pazarlama, perakende satış ve elektronik ticaret gibi birçok alana yayılmıştır(Chen Y, Chen J, 2006: 1503 – 1520).

Birliktelik kuralların yaygın olarak uygulandığı bir diğer alan da alışveriş merkezleri ve süpermarketlerdir. Günümüz teknolojisindeki gelişme ile birlikte küçük

ölçekli marketlerden büyük ölçekli süper market zincirlerine kadar bütün alışveriş noktalarında, yapılan satışların tarihi, saati ve satılan ürün ile ilgili olan bilgiler(ürünün numarası, ürünün adı, ürünün adedi vs.) kayıt altında tutulmaktadır. Literatürde market sepet analizi(market basket analysis) olarak geçen uygulama, kayıt altında tutulan bu bilgiler sayesinde sektörde oldukça fazla kullanılan bir uygulamadır. Bu analiz ürünlerin birlikte satın alınma eğilimlerini gösterir. Aynı zamanda bu eğilimler neticesinde müşterilerin de satın alma eğilimleri modellenabilir. Elde edilen birliktelik kuralları ışığında müşterilerin satın alma esnasında hani ürünleri birlikte aldıkları bilgisi ortaya çıkar. Bu bilgi, işletme yöneticileri tarafından çeşitli satış stratejilerinin geliştirilmesine yardımcı olur. Örnek olarak alışverişleri sırasında kola alan müşteriler genel olarak patates cipsini de satın alıyorsa satın alınan bu iki ürün arasında kuvvetli bir ilişkinin olduğu bilgisi ortaya çıkar.

Birliktelik kurallarının çıkarılmasında çok bilindik ve yaygın olarak kullanılan algoritmaları şöyle sıralayabiliriz:

- Apriori Algoritması
- AIS Algoritması
- AprioriTid Algoritması
- SETM Algoritması

Apriori algoritmasının temel felsefesi; eğer bir küme minimum destek değerini sağlıyorsa bu kümenin alt kümeleri de bu destek değerini sağlar şeklindedir. Birlikteliği olan ürünler aranırken veri seti çok sık bir şekilde taranır. Zaten algoritmanın amacı, sık geçen eleman kümelerini tespit etmektir (Eker, 2016: 20).

Apriori algoritmasında ilk adım, veri seti üzerinde minimum destek değeri şartını sağlayan ve sık geçen tek elemanlı verilerin kümelenmesi adımıdır. Sonraki adımlarda, önceki adımlardan elde edilen eleman kümeleri ve yeni, daha geniş olan kümelerin üretilmesi sağlanır. Kümeleme işlemi sonrasında, minimum destek değeri şartını sağlayan kümeler ayrılır. Sık geçen bu kümeler bir sonraki adımın nesne kümesi olurlar ve bu işlem sık geçen eleman kümesi bulunmaya kadar tekrarlanır.

BİRLİKTELİK KURALI

Birliktelik Kuralının Matematiksel Gösterimi

Agrawal ve Srikant yayınlarında birliktelik kurallarını matematiksel olarak aşağıdaki şekilde ifade etmiştir (Agrawal & Srikant, 1993: 487 – 499):

$I = \{I_1, I_2, \dots, I_m\}$ bir dizi – nesne kümesi olsun.

$T = \{t_1, t_2, \dots, t_n\}$ ise veri tabanındaki işlemleri (alışverişi) göstereyin.

Her bir t_k 'nın alacağı değer 0 veya 1'dir. Eğer $T_k = 0$ ise I_k satın alınmamış, eğer $t_k = 1$ ise I_k satın alınmış demektir. Her bir işlem için veri tabanında ayrı bir kayıt vardır. Şimdi x için X 'teki her bir I_k 'ya karşılık gelen t_k değeri $t_k = 1$ 'dir. Bu birliktelik kuralı şu şekilde ifade edilmektedir: $X \Rightarrow I_j$, X, I 'nın bir alt kümesidir. I_j ise I içindeki herhangi bir elemandır ve bu eleman X içinde yer almamaktadır. $X \Rightarrow I_j$ kuralının T için uygun olduğunun söylenebilmesi için belli bir güven seviyesinden söz etmek gerekecektir. Dolayısıyla, T için deki tüm X 'lerin ne kadarının I_k 'yı sağladığı %c değeriyle ifade edilmelidir. Bu durumda, birliktelik kuralı güven seviyesiyle birlikte aşağıdaki gibi ifade edilebilmektedir(Silahtaroglu, 2014).

$$X \Rightarrow I_j \mid c$$

Destek ve Güven Değerleri

Birliktelik kurallarının çıkarılması sırasında bazı değerlerin temel alınması gerekir. Bunlar destek ve güven değerleridir. Destek seviyesi, T içindeki işlemlerin ne kadarını X 'in sağladığıdır. Mevcut veri tabanından birçok kural çıkarılabilir. Ancak bu kuralların tamamı önem arz etmeyebilir. Hazırlanan algoritma, tanımlanacak destek ve güven seviyeleri ile elde edilmeye başlanan kurallarda bu seviyeleri aşma durumlarına göre filtrelenir. Bu şekilde belirli bir seviyenin üzerinde kalan daha anlamlı kurallar elde edilmiş olur. Bu şekilde en küçük destek seviyesini sağlayan nesne kümelerine geniş, diğerlerine küçük nesne kümesi denilir (Agrawal R, Srikant R,1993: 487 – 499).

AMAÇ VE PROBLEMİN TANIMLANMASI

Bu çalışmanın amacı, veri madenciliğinde kullanılan tekniklerden biri olan ve market sepet analizlerinde kullanılan apriori algoritmasının Microsoft Excel ortamına eklenti olarak kodlanmasıdır. Bu çalışma ile veri madenciliği tekniklerini kullanarak veri analizi yapmak, analiz sonucunda anlamlı veriyi kullanıcıya sunmak, diğer veri madenciliği yazılımlardan daha kolay, daha pratik bir kullanıcı deneyimi sağlamak hedeflenmektedir. Bu sayede kullanıcılar, bilgi ve deneyim gerektiren diğer veri madenciliği yazılımlarının aksine daha kolay ve aşına oldukları bir ortamda veri analizlerini gerçekleştirebileceklerdir. Yazılımı kullanmak için sadece eklenti Excel ortamına kurmak adına basit bir kurulum yapmaları yeterli olacaktır. Geliştirilen yazılım için kullanıcıların bir ön bilgi ve tecrübe sahibi olmalarına gerek yoktur. Yalnızca veri setini belirtilen formatta Excel'e girdi olarak eklemeleri ve aktif olan eklenti sayesinde algoritmanın uygula butonuna basmaları yeterli olacaktır. Nihai olarak analiz sonucu kullanıcıya, yeni bir Excel sayfasında bir tablo halinde gösterilecektir. Böylece sepet analizi yapmak isteyen, büyük ve analiz gerektiren veriye sahip her kullanıcı, bu yazılım sayesinde kolay bir şekilde verisini anlamlandırıp kullanabilecektir.

Geliştirilen bu yazılım sayesinde şu problemlerin çözülmesi beklenmektedir:

- Kullanıcılara ön bilgi ve deneyim gerektirmeden veri analizi imkânı sağlama
- Veri sahiplerinin, ürünlerin birliktelikleri ile ilgili bilgi sahibi olmasını sağlama
- Kampanya yapılabilecek ikili ürünler, ürünlerin raflardaki yerlerinin optimize edilmesi, müşteri analizi gibi konularda kullanıcıların bilgi sahibi olmasını sağlama
- Bilimsel çalışmalarda apriori algoritmasını kullanmak ya da algoritmanın nasıl çalıştığını anlamak isteyen bireylere hız kazandırma
- Veri madenciliği tekniklerinin sadece bilinen yazılımlarla değil farklı ortamlarda da kullanılabilirliğinin ispatını ortaya koyma

SİSTEM ANALİZİ

Problemin Tanımı

Verilerin toplanmasını ve saklanması kolaylaştıran teknolojiler sayesinde kullanıcılar büyük veriler elde etmektedir. Elde edilen büyük veriler beraberinde bu verilerden anlamlı bir sonuç çıkarma sorununu da beraberinde getiriyor. Verilerin işlenip anlamlı hale getirilerek bilgiye çevrilmesi oldukça önemli ve zor bir konudur. Veri madenciliği teknikleri bu zorlukların üstesinden gelmeyi sağlamaktadır. Veri madenciliği teknikleri sayesinde ham halde olan büyük veri işlenerek istenilen amaç doğrultusunda anlamlı bir bilgiye dönüşmektedir. Bu çalışmada hazırlanan uygulamanın amacı da market fiş verilerinden anlamlı bir bilgi elde etmektir.

İstenilenler

1. Kullanıcı verilerini Excel ortamına aktararak işlemlerini orada yapmalıdır.
2. Kullanıcı bir minimum destek değeri(minimum support value) belirlemelidir.
3. Veri seti; tarih, saat, fiş numarası, ürün adı şeklinde olmalıdır.
4. Tek bir buton ile sonuçlar ve işlem adımları program tarafından otomatik olarak yapılmalıdır.
5. Sonuçlar kullanıcıya bir tablo halinde sunulmalıdır.
6. Tüm süreci kullanıcıya açıklayan bir kullanım kılavuzu da uygulama içerisinde bulunmalıdır.

Gereksinimlerin Tanımlanması

Fonksiyonel Gereksinimler

Kullanıcının istediklerini yerine getirmek için fonksiyonlar yazılacaktır. Algoritmanın adımları fonksiyonel olarak tanımlanmıştır. Kullanıcı isteklerini yerine getirmek amacıyla algoritma adımları dinamik bir şekilde yerine getiren fonksiyonlar bulunmaktadır.

- **textToColumn:** Kullanıcının tek satır olarak girmiş olduğu veri setini yeni bir Excel çalışma sayfasına virgül karakterini referans alarak dört sütuna ayıran fonksiyon.
- **createPivotTable:** textToColumn() fonksiyonun sütunlara ayırdığı veri setini fiş numarası ve ürün ismini referans alarak Excel pivot tablo oluşturan fonksiyon.
- **writeColumn:** Pivot tablo içerisindeki her bir ürünün destek değerinin hesaplanıp yeni bir sayfada listelenmesi için Excel sütun isimlerini (A...Z, AA...

AZ, BA... BZ gibi) ürün sayısı kadar oluşturan ve geriye sütun isimlerinin olduğu bir string dizi döndüren fonksiyon.

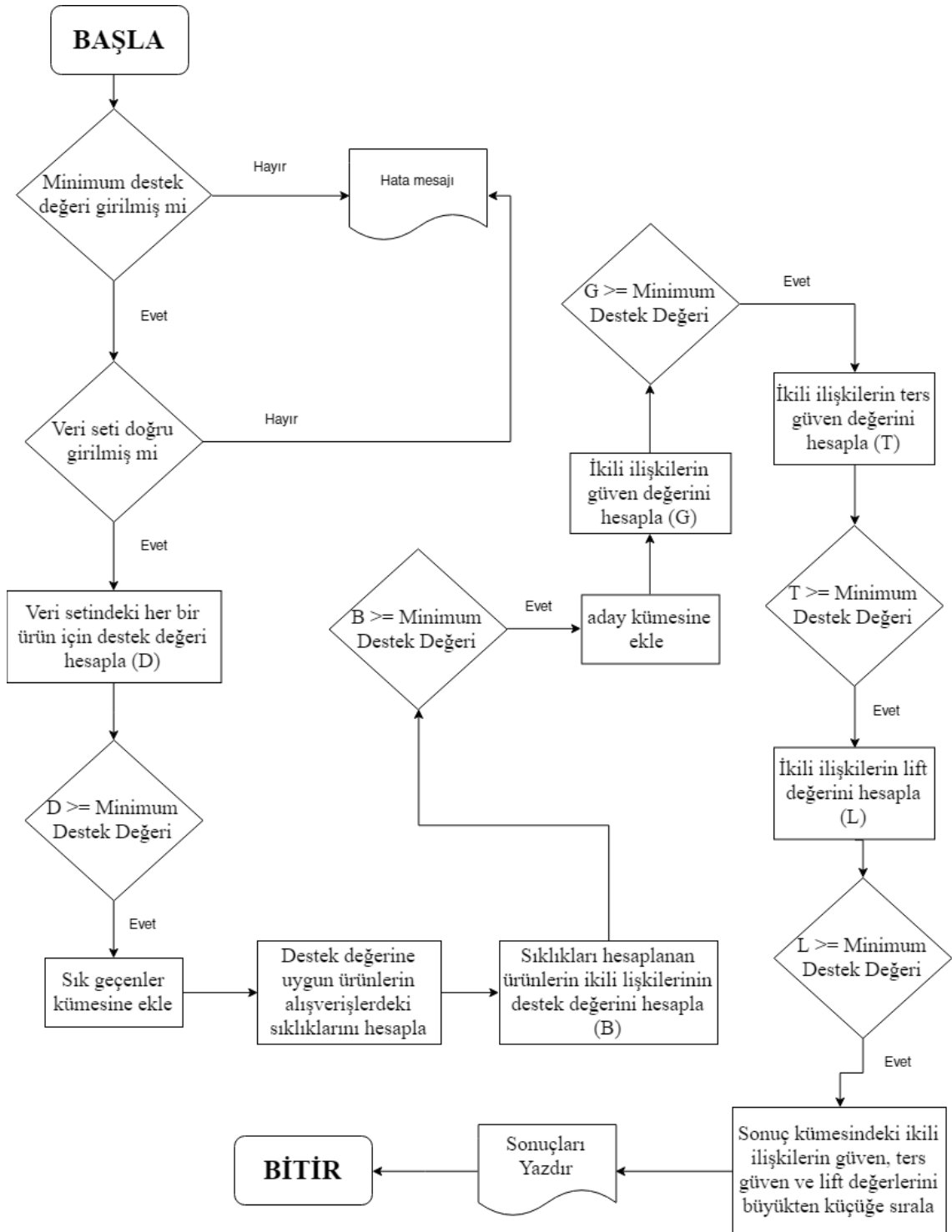
- **calcCombination:** Minimum destek değerine uygun olan ürünler arasındaki ikili birlikteliklerin hesaplanması ve bu değerlerin bir dizide tutulması için dizi boyutunun ürün sayısının ikili kombinasyonunu hesaplanarak belirlenmesini sağlayan fonksiyon.
- **swap:** Analiz sonucunda elde edilen sonuç değerleri ve bu değerlerin sahip olduğu ikili birlikteliklerin bilgilerinin büyükten küçüğe doğru sıralayan fonksiyon.
- **resultPrint:** swap() fonksiyonunun sonuçları sıralamasından sonra bu sonuçları yeni bir Excel çalışma sayfası açarak kullanıcıya gösteren fonksiyon.
- **createTable:** resultPrint() fonksiyonunun yazdırdığı sonuçları bir tablo haline getiren fonksiyon.
- **changeAllLabelColor:** Kullanım kılavuzunun kullanıcı gösterildiği pencerede seçili olan adımın dışındaki bütün adımların renklerini siyah yapan fonksiyon.

```

 $L_1 = \{\text{sık geçen 1-öge kümesi}\};$ 
for ( $k=2; L_{k-1} \neq \emptyset; k++$ ) do begin
     $C_k = \text{apriori-gen}(L_{k-1});$  // Yeni adaylar
    forall transactions-hareketler  $t \in D$  do begin
         $C_t = \text{subset}(C_k, t);$  // Adaylar t içindedir
        forall candidates – adaylar  $c \in C_t$  do
             $c.\text{count}++;$ 
        end
    end
     $L_k = \{c \in C_k \mid c.\text{count} \geq \text{minsup}\}$ 
end
Answer =  $\cup_k L_k;$ 

```

Şekil 1: Apriori Algoritmasının klasik özet kodu (pseudo code)



Şekil 2: Uygulamanın akış diyagramı

YÖNTEM

Geliştirilen yazılım bir marketin fiş verileri kullanılarak test edilmiştir. Bu veri setine Apriori algoritması uygulanarak verilerden anlamlı bir kural çıkarmak amaçlanmıştır. Fiş verileri; tarih, saat, fiş numarası ve ürün adından oluşmaktadır.

Kullanılan Teknolojiler

Yazılım; Microsoft .NET çatısı altında, C# programlama dili kullanılarak, Visual Studio geliştirme ortamında geliştirilmiştir. Microsoft Excel'e bir eklenti olarak geliştirilen bu yazılım basit bir kurulum ile kullanılabilir.

Apriori Algoritması

Bu çalışmada, veri setinden anlamlı bir bilgi elde etmek için Apriori algoritması kullanılmıştır. Algoritma öncelikle her bir ürünün destek seviyesini hesaplamaktadır. Bu seviyeler yazılım içerisinde tanımlanan destek değeri(Support) ile karşılaştırılır. Destek değerinin altında kalan ürünler bir sonraki aşamaya dâhil edilmezler. Destek seviyesini aşan ürünler ile algoritma adımlarına devam edilir. Destek değerini aşan elemanlar kümesi içindeki ürünlerin ikili olarak ilişkilerine bakılır. Bunun neticesinde ikili olarak değerlendirilen ürünlerin de destek değerleri hesaplanır. İkili olarak analiz edilen bu ürünlerin destek değerleri, tanımlı destek değerinin altında kalıyorsa bu birliktelik de ihmal edilir. Destek değerini aşan birlikteliklerin güven(Confidence) değerleri, ters ikili güven değerleri ve kaldırma(lift) değerleri hesaplanarak kullanıcıya bir tablo şeklinde sunulur.

Destek Değeri

X ve Y ürün gruplarını birlikte içeren alışveriş sayısını, $sayı(X, Y)$ şeklinde gösterilir.

X ürün grubunu alan müşterilerin Y ürün grubunu da alma durumları; yani birliktelik kuralı $(X \rightarrow Y)$ şeklinde gösterilir.

$destek(X \rightarrow Y) = sayı(X, Y) / N$ olarak formüle edilir. (N = Tüm alışveriş sayısı)

Güven Değeri

X ürün grubu ürünlerinden alan müşterilerin Y ürün grubu ürünlerinden alma olasılığı güven değerini ifade etmektedir.

$güven(X \rightarrow Y) = destek(X \rightarrow Y) / destek(X)$

İkili İlişkinin Tersine Güven Değeri

Y ürün grubu ürünlerinden alan müşterilerin X ürün grubu ürünlerinden alma olasılığı ters güven değerini ifade etmektedir.

$$\text{tersgüven}(Y \rightarrow X) = \text{destek}(X \rightarrow Y) / \text{destek}(Y)$$

Lift Değeri

$$\text{lift}(X \rightarrow Y) = \text{destek}(X \rightarrow Y) / (\text{destek}(X) * \text{destek}(Y))$$

Olacak şekilde hesaplanmıştır.

ÖRNEK UYGULAMA

Tablo 1’ de bir marketin fiş verileri verilmiştir. Bu verilere göre ürünlerin birliktelikleri şu şekilde hesaplanır:

Fiş Numarası	Satın Alınan Ürünler
1	Ekmek, Şeker, Süt, Çay
2	Yumurta, Süt, Ekmek
3	Tereyağı, Ekmek, Süt
4	Şeker, Süt
5	Çay, Şeker

Tablo 1: Örnek Uygulama Veri Seti

Bu uygulamada yapılan toplam alışveriş sayısı 5’tir. Minimum destek değeri de 0,4 olarak alınmıştır. Bir ürünün alışverişlerdeki alınma sıklığı toplam alışverişin %40’ından az ise o ürün aday kümeye alınmayıp ihmal edilecek demektir. Tablo 2’de her bir ürünün destek değeri aşağıdaki formüle göre hesaplanmıştır:

$\text{destek}(X) = \text{sayı}(X) / N$ ($N = \text{Tüm alışveriş sayısı}$, $\text{sayı}(X) = \text{ürünün alışverişlerin tamamında geçme sıklığı}$)

Ürün Adı	Tüm Alışverişlerde Geçme Sıklığı (sayı(X))	Toplam Alışveriş Sayısı (N)	Destek(X)
Ekmek	3	5	0,6
Şeker	3	5	0,6
Süt	4	5	0,8
Çay	2	5	0,4

Yumurta	1	5	0,2
Tereyağı	1	5	0,2

Tablo 2: Ürünlerin Destek Değerleri

Minimum destek değerine uygun olmayan yumurta ve tereyağı ihmal edilerek, minimum destek şartını sağlayan ekmek, şeker, süt ve çay ile algoritma adımlarına devam edilmiştir. Tablo 3'te minimum destek değerini sağlayan ürünlerin birlikte alınma sıklıklarının destek değerleri aşağıdaki formüle göre hesaplanmıştır:

$\text{destek}(X \rightarrow Y) = \text{sayı}(X, Y) / N$ ($N = \text{Tüm alışveriş sayısı}$, $\text{sayı}(X, Y)$ ürünlerin alışverişlerde birlikte alınma sıklığı)

Ürün adı	Birlikte Alınma Sıklığı (sayı(X, Y))	Toplam Alışveriş Sayısı	Destek(X -> Y)
{Ekmek, Şeker}	1	5	0,2
{Ekmek, Süt}	3	5	0,6
{Ekmek, Çay}	1	5	0,2
{Şeker, Süt}	2	5	0,4
{Şeker, Çay}	2	5	0,4
{Süt, Çay}	1	5	0,2

Tablo 3: İkili İlişkilerin Destek Değerleri

Tablo 3' e bakıldığında ürünlerin birlikteliklerinin destek değerlerinde minimum destek değerini sağlamayan değerler vardır. Bu birliktelikler ihmal edilerek algoritma adımlarına {Ekmek, Süt}, {Şeker, Süt} ve {Şeker, Çay} birliktelikleri ile devam edilmiştir. Tablo 4' te ikili ilişkilerin güven değerleri aşağıdaki formüle göre hesaplanmıştır:

$$\text{güven}(X \rightarrow Y) = \text{destek}(X \rightarrow Y) / \text{destek}(X)$$

Birliktelikler	Destek (X -> Y)	Destek(X)	Güven(X -> Y)
{Ekmek, Süt}	0,6	0,6	1
{Şeker, Süt}	0,4	0,6	0,6666666666666667
{Şeker, Çay}	0,4	0,6	0,6666666666666667

Tablo 4: İkili İlişkilerin Güven Değerleri

Tablo 4'ten de anlaşılacağı gibi ikili ilişkilerin güven değerlerinden minimum destek değerini sağlamayan değer yoktur. Bu sebeple aynı ilişkiler ile algoritma adımlarına devam edilmiştir. Tablo 5' te ikili ilişkilerin tersine güven değerleri aşağıdaki formüle göre hesaplanmıştır:

$$\text{tersgüven}(Y \rightarrow X) = \text{destek}(X \rightarrow Y) / \text{destek}(Y)$$

Birliktelikler	Destek(X -> Y)	Destek(Y)	Ters Güven(Y -> X)
{Süt, Ekmek}	0,6	0,8	0,75
{Süt, Şeker}	0,4	0,8	0,5
{Çay, Şeker}	0,4	0,4	1

Tablo 5: İkili İlişkilerin Tersine Güven Değerleri

Tersine ikili ilişkilerin güven değerlerinden de minimum destek değerini sağlamayan değer olmadığından aynı ilişkiler ile algoritma adımlarına devam edilmiştir. Tablo 6' da ikili ilişkilerin lift değerleri aşağıdaki formüle göre hesaplanmıştır:

$$\text{lift}(X \rightarrow Y) = \text{destek}(X \rightarrow Y) / (\text{destek}(X) * \text{destek}(Y))$$

Birliktelikler	Destek(X -> Y)	Destek(X) * Destek(Y)	Lift(X -> Y)
{Ekmek, Süt}	0,6	0,6 * 0,8	1,25
{Şeker, Süt}	0,4	0,6 * 0,8	0,83333333333333
{Şeker, Çay}	0,4	0,6 * 0,4	1,66666666666666

Tablo 6: İkili İlişkilerin Lift Değerleri

Birliktelikler	Güven(X->Y)	Tersgüven(X->Y)	Lift(X->Y)
{Ekmek, Süt}	1	0,75	1,25
{Şeker, Süt}	0,666666666666667	0,5	0,83333333333333
{Şeker, Çay}	0,666666666666667	1	1,66666666666666

Tablo 7: Örnek Uygulamanın Sonuçları

Tablo 7'deki sonuçlar neticesinde;

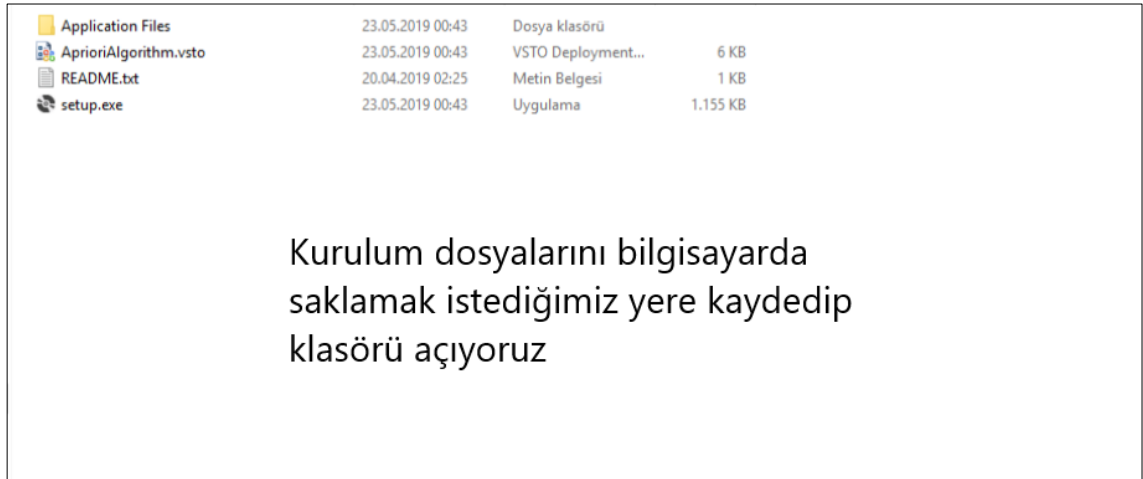
- Alışverişlerinde ekmek alan müşterilerin %100' ünün süt de aldığı, alışverişlerinde süt alan müşterilerin ise %75' inin ekmek de aldığı görülmektedir.
- Alışverişlerinde şeker alan müşterilerin yaklaşık olarak %67' sinin süt de aldığı, alışverişlerinde süt alan müşterilerin ise %50' sinin ekmek de aldığı görülmektedir.
- Alışverişlerinde şeker alan müşterilerin yaklaşık olarak %67' sinin çay da aldığı, alışverişlerinde çay alan müşterilerin ise %100' ünün şeker de aldığı görülmektedir.

YAZILIMIN RAPORU

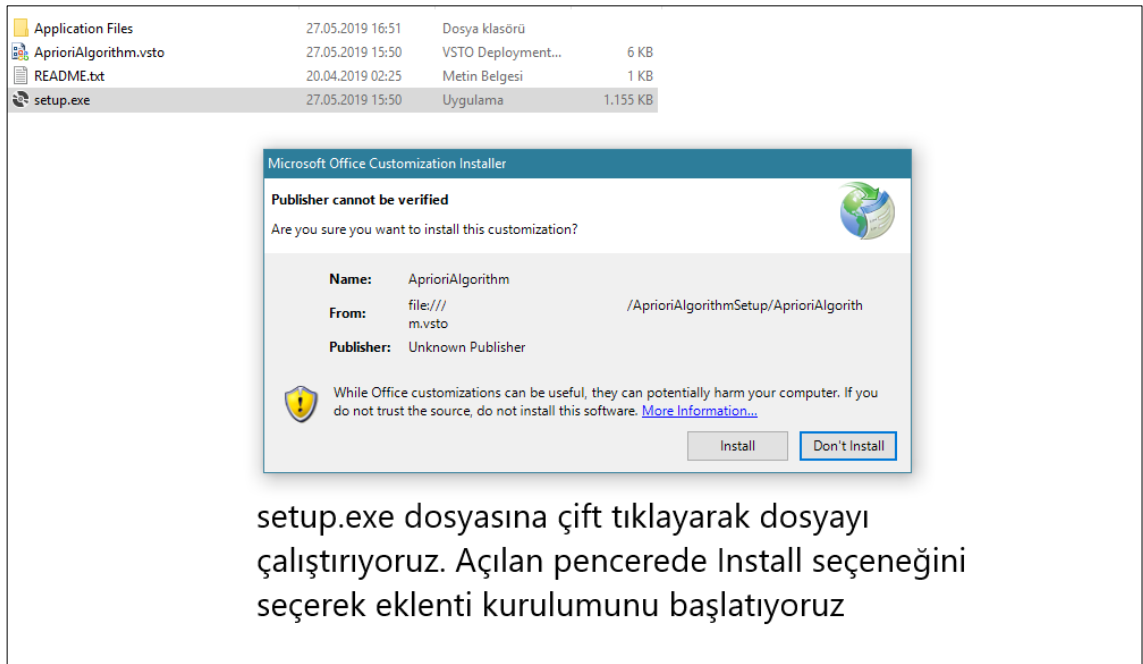
Yazılımın Kaynak Kodları

Bu çalışma ile gerçekleştirilen uygulamanın kaynak kodlarına <https://github.com/halilkayaci/AprioriAlgorithm> adresinden veya Pamukkale Üniversitesi İktisadi ve İdari bilimler Fakültesi Yönetim Bilişim Sistemleri Bölümünün internet sayfasından ulaşabilirsiniz.

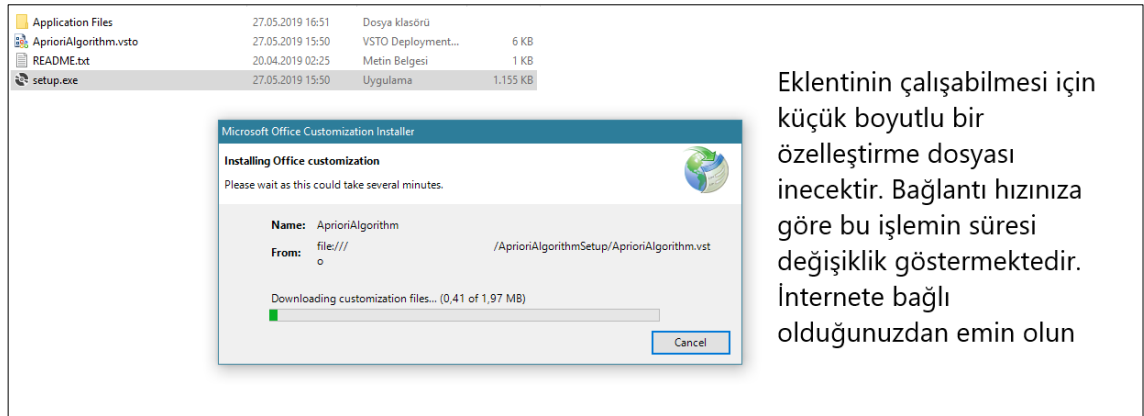
KULLANIM KILAVUZU



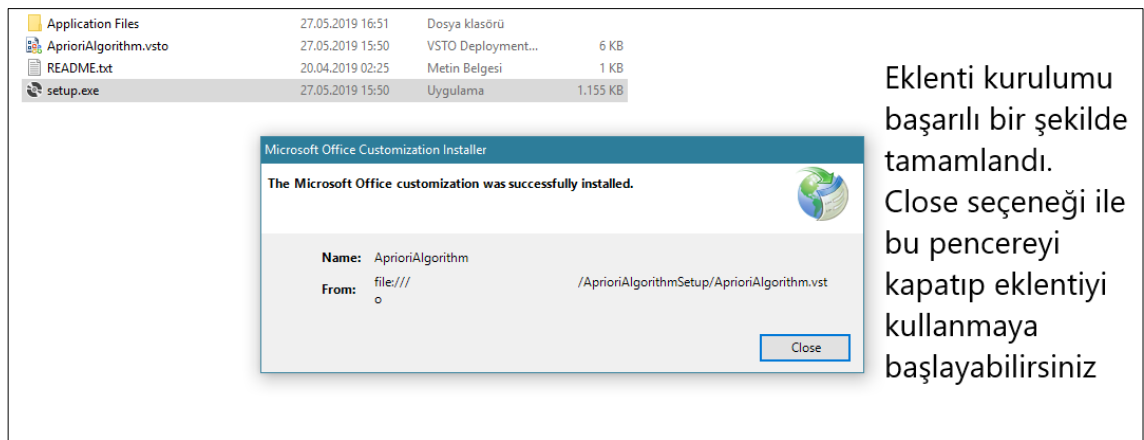
Şekil 3: Eklenti Kurulumu Adım 1



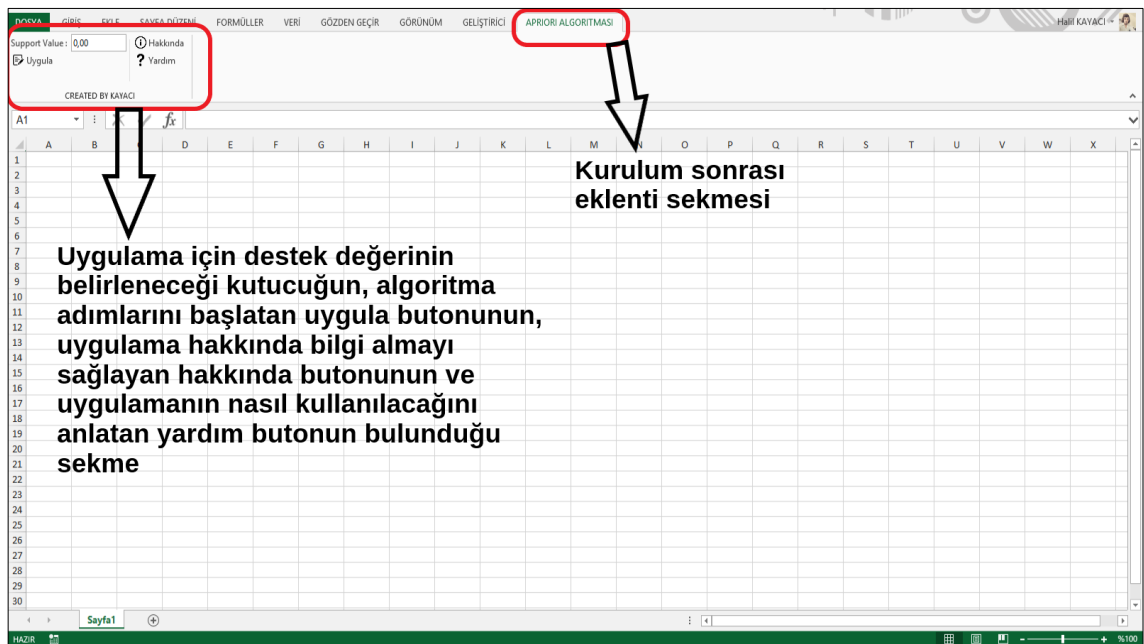
Şekil 4: Eklenti Kurulumu Adım 2



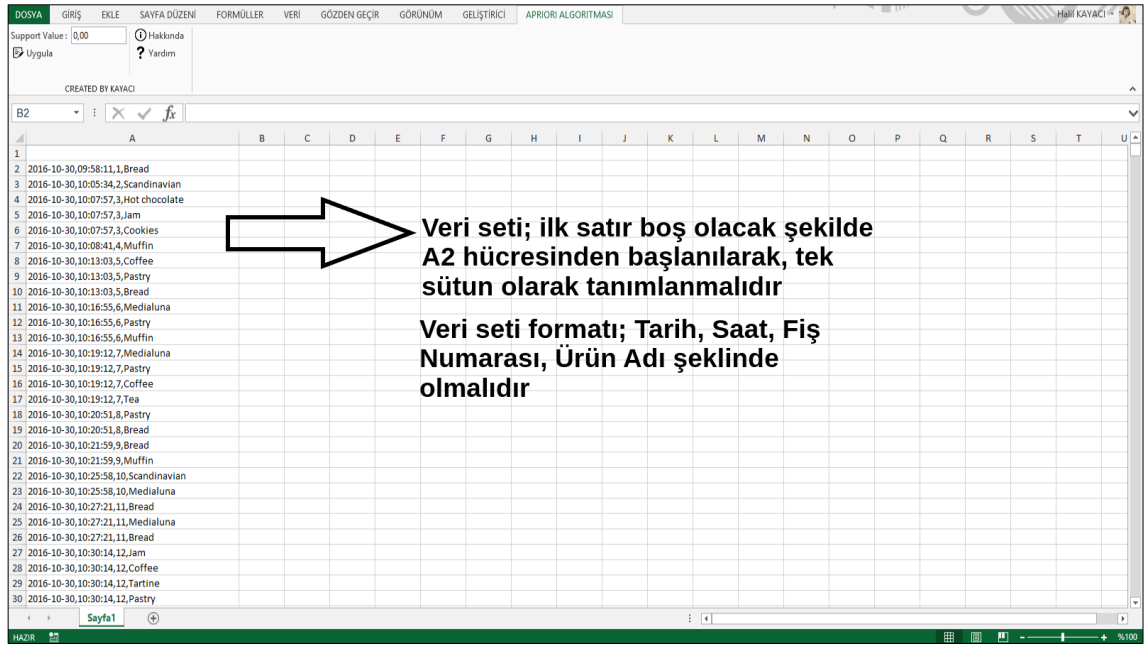
Şekil 5: Eklenti Kurulumu Adım 3



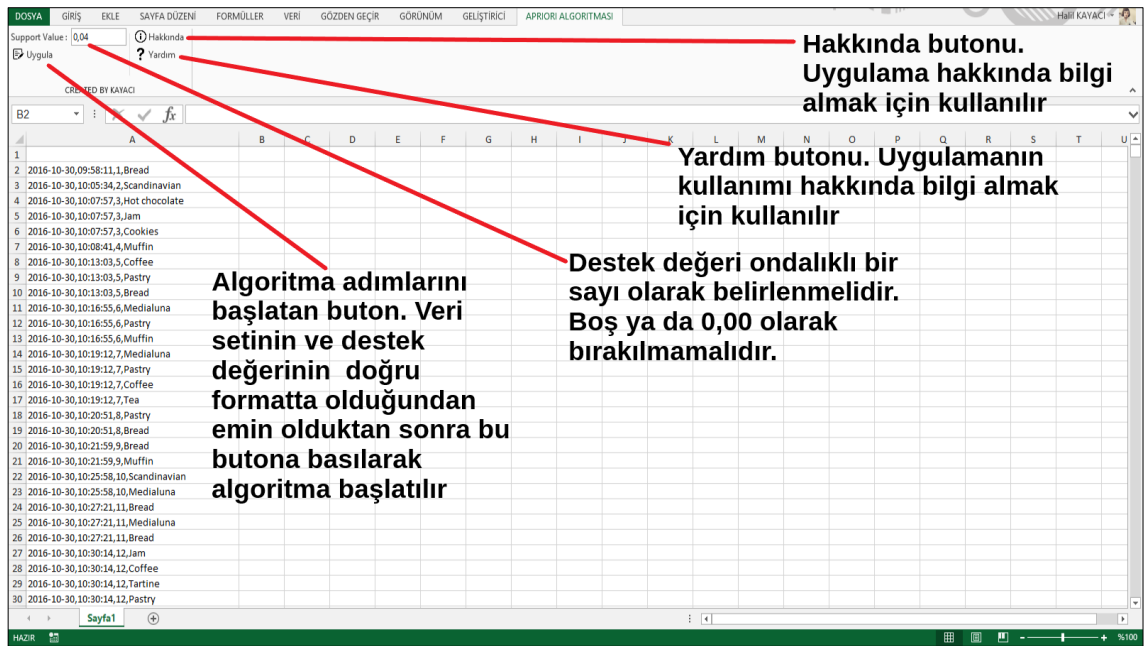
Şekil 6: Eklenti Kurulumu Adım 4



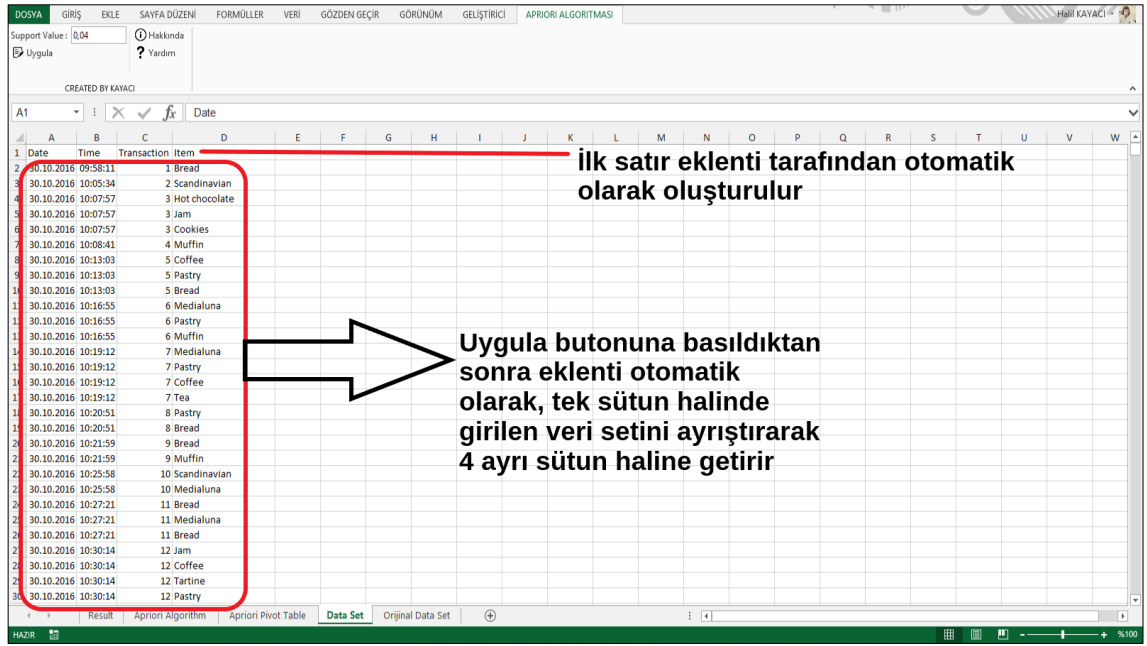
Şekil 7: Kullanım Kılavuzu Adım 1



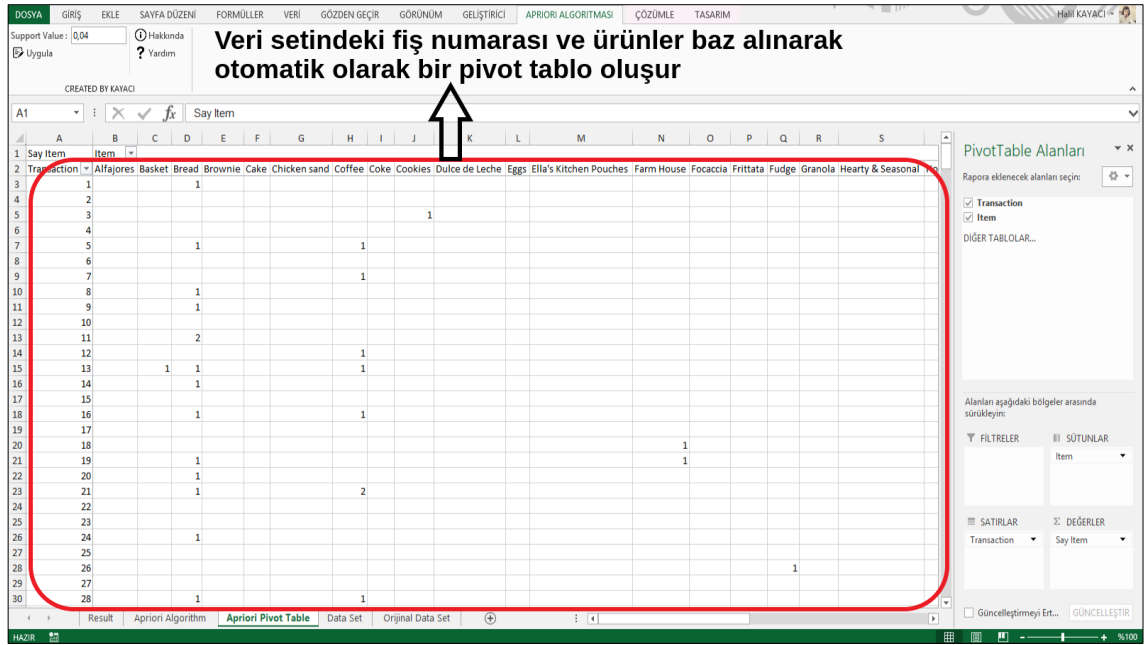
Şekil 8: Kullanım Kılavuzu Adım 2



Şekil 9: Kullanım Kılavuzu Adım 3



Şekil 10: Kullanım Kılavuzu Adım 4



Şekil 11: Kullanım Kılavuzu Adım 5

Tüm ürünlerin destek değeri hesaplanır **Belirlenen destek değerine uygun olan ürünler bu satırda gösterilir** **Destek değerine uygun ürünlerin beraber alınma sıklıkları**

Urunkler	Alfajores	Basket	Bread	Brownie
Urunklerin Destek Degerleri	0,05513308		0,003802281	0,2832
Destek Degerine Uygun Urunkler	Alfajores	Bread	Brownie	
İkili İlişkiler	{Alfajores,Bread} = 7 {Alfajores,Brownie} = 5 {Alfajores,Coffee} = 17 {Alfajores,Cookies} = 4 {Alfajores,Farm House} = 3 {Alfajores,Hearty & Seasonal} = 1 {Alfajores,Jam} = 2 {Alfajores,Juice} = 8 {Alfajores,Medialuna} = 5 {Alfajores,Muffin} = 1 {Alfajores,Pastry} = 2 {Alfajores,Scandinavian} = 3 {Alfajores,Soup} = 1 {Alfajores,Tea} = 4	{Bread,Brownie} = 3 {Bread,Coffee} = 37 {Bread,Cookies} = 9 {Bread,Farm House} = 3 {Bread,Hearty & Seasonal} = 7 {Bread,Jam} = 9 {Bread,Juice} = 4 {Bread,Medialuna} = 10 {Bread,Muffin} = 6 {Bread,Pastry} = 22 {Bread,Scandinavian} = 5 {Bread,Soup} = 5 {Bread,Tea} = 14	{Brownie,Coffee} = 13 {Brownie,Cookies} = 4 {Brownie,Farm House} = 2 {Brownie,Hearty & Seasonal} = 2 {Brownie,Jam} = 1 {Brownie,Juice} = 5 {Brownie,Medialuna} = 1 {Brownie,Muffin} = 1 {Brownie,Pastry} = 4 {Brownie,Scandinavian} = 1 {Brownie,Soup} = 1 {Brownie,Tea} = 4	
İkili İlişkilerin Destek Degerleri	{Alfajores,Bread} = Bu iliski destek degerine uygun degildir {Alfajores,Brownie} = Bu iliski destek degerine uygun degildir {Alfajores,Coffee} = Bu iliski destek degerine uygun degildir {Alfajores,Cookies} = Bu iliski destek degerine uygun degildir {Alfajores,Farm House} = Bu iliski destek degerine uygun degildir {Alfajores,Hearty & Seasonal} = Bu iliski destek degerine uygun degildir {Alfajores,Jam} = Bu iliski destek degerine uygun degildir {Alfajores,Juice} = Bu iliski destek degerine uygun degildir {Alfajores,Medialuna} = Bu iliski destek degerine uygun degildir {Alfajores,Muffin} = Bu iliski destek degerine uygun degildir	{Bread,Brownie} = Bu iliski destek degerine uygun degildir {Bread,Coffee} = 0,0703422053231939 {Bread,Cookies} = Bu iliski destek degerine uygun degildir {Bread,Farm House} = Bu iliski destek degerine uygun degildir {Bread,Hearty & Seasonal} = Bu iliski destek degerine uygun degildir {Bread,Jam} = Bu iliski destek degerine uygun degildir {Bread,Juice} = Bu iliski destek degerine uygun degildir {Bread,Medialuna} = Bu iliski destek degerine uygun degildir {Bread,Muffin} = Bu iliski destek degerine uygun degildir {Bread,Pastry} = 0,0418250950570342	{Brownie,Coffee} = Bu iliski destek degerine uygun degildir {Brownie,Cookies} = Bu iliski destek degerine uygun degildir {Brownie,Farm House} = Bu iliski destek degerine uygun degildir {Brownie,Hearty & Seasonal} = Bu iliski destek degerine uygun degildir {Brownie,Jam} = Bu iliski destek degerine uygun degildir {Brownie,Juice} = Bu iliski destek degerine uygun degildir {Brownie,Medialuna} = Bu iliski destek degerine uygun degildir {Brownie,Muffin} = Bu iliski destek degerine uygun degildir {Brownie,Pastry} = Bu iliski destek degerine uygun degildir {Brownie,Scandinavian} = Bu iliski destek degerine uygun degildir	

Şekil 12: Kullanım Kılavuzu Adım 6

DOSYA

GİRİŞ

EKLE

SAĞFAH DÜZENİ

FÖRMLER

VERİ

GÖZDEN GEÇİR

GÖRÜM

GELİŞTİR

APRİORİ ALGORİTMA

YENİ

YARDIM

HAZIR KAYACI

Support Value : 0.04

Hakkında

Yükle

?

CREATED BY KATAY

A50

<

Şekil 13: Kullanım Kılavuzu Adım 7

DOSYA GİRİŞ EKLE SAYFA DÜZENİ FORMÜLLER VERİ GÖZDEN GEÇİR GÖRÜNÜM GELİŞTİRİCİ APRIORI ALGORİTMA HAKKI KAYACI

Support Value: 0,04

Uygula

CREATED BY KAYACI

A50

A	B	C	D
66 İkili İlişkilerin Lift Değerleri	{Alfajores,Bread} = Bu ilişki destek değerine uygun değildir	{Bread,Brownie} = Bu ilişki destek değerine uygun değildir	{Brownie,Coffee} = Bu ilişki destek değerine uygun değildir
67	{Alfajores,Brownie} = Bu ilişki destek değerine uygun değildir	{Bread,Coffee} = 0,492896036469545	{Brownie,Cookies} = Bu ilişki destek değerine uygun değildir
68	{Alfajores,Coffee} = Bu ilişki destek değerine uygun değildir	{Bread,Cookies} = Bu ilişki destek değerine uygun değildir	{Brownie,Farm House} = Bu ilişki destek değerine uygun değildir
69	{Alfajores,Cookies} = Bu ilişki destek değerine uygun değildir	{Bread,Farm House} = Bu ilişki destek değerine uygun değildir	{Brownie,Hearty & Seasonal} = Bu ilişki destek değerine uygun değildir
70	{Alfajores,Farm House} = Bu ilişki destek değerine uygun değildir	{Bread,Hearty & Seasonal} = Bu ilişki destek değerine uygun değildir	{Brownie,Jam} = Bu ilişki destek değerine uygun değildir
71	{Alfajores,Hearty & Seasonal} = Bu ilişki destek değerine uygun değildir	{Bread,Jam} = Bu ilişki destek değerine uygun değildir	{Brownie,Juice} = Bu ilişki destek değerine uygun değildir
72	{Alfajores,Jam} = Bu ilişki destek değerine uygun değildir	{Bread,Juice} = Bu ilişki destek değerine uygun değildir	{Brownie,Medialuna} = Bu ilişki destek değerine uygun değildir
73	{Alfajores,Juice} = Bu ilişki destek değerine uygun değildir	{Bread,Medialuna} = Bu ilişki destek değerine uygun değildir	{Brownie,Muffin} = Bu ilişki destek değerine uygun değildir
74	{Alfajores,Medialuna} = Bu ilişki destek değerine uygun değildir	{Bread,Muffin} = Bu ilişki destek değerine uygun değildir	{Brownie,Pastry} = Bu ilişki destek değerine uygun değildir
75	{Alfajores,Muffin} = Bu ilişki destek değerine uygun değildir	{Bread,Pastry} = 1,04951931797569	{Brownie,Scandinavian} = Bu ilişki destek değerine uygun değildir
76	{Alfajores,Pastry} = Bu ilişki destek değerine uygun değildir	{Bread,Scandinavian} = Bu ilişki destek değerine uygun değildir	{Brownie,Soup} = Bu ilişki destek değerine uygun değildir
77	{Alfajores,Scandinavian} = Bu ilişki destek değerine uygun değildir	{Bread,Soup} = Bu ilişki destek değerine uygun değildir	{Brownie,Tea} = Bu ilişki destek değerine uygun değildir
78	{Alfajores,Soup} = Bu ilişki destek değerine uygun değildir	{Bread,Tea} = Bu ilişki destek değerine uygun değildir	
79	{Alfajores,Tea} = Bu ilişki destek değerine uygun değildir		
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			

HAZIR

Result Apriori Algorithm Apriori Pivot Table Data Set Original Data Set

Birlikteliklerin lift değerleri

Şekil 14: Kullanım Kılavuzu Adım 8

DOSYA GİRİŞ EKLE SAYFA DÜZENİ FORMÜLLER VERİ GÖZDEN GEÇİR GÖRÜNÜM GELİŞTİRİCİ APRIORI ALGORİTMA HAKKI KAYACI

Support Value: 0,04

Uygula

CREATED BY KAYACI

A1

Confidence

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1 Confidence	RConfidence	Lift													
2 {Bread,Coffee} = 0,248322147651007	{Pastry,Coffee} = 0,472972972972973	{Bread,Pastry} = 1,04951931797569													
3 {Bread,Pastry} = 0,147651006711409	{Pastry,Bread} = 0,297297297297297	{Coffee,Pastry} = 0,938806731259561													
4 {Coffee,Pastry} = 0,132075471698113	{Tea,Coffee} = 0,28735632183908	{Coffee,Tea} = 0,570375189763609													
5 {Coffee,Tea} = 0,0943396226413094	{Coffee,Bread} = 0,139622641509434	{Bread,Coffee} = 0,492896036469545													
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															
26															
27															
28															
29															
30															

HAZIR

Result Apriori Algorithm Apriori Pivot Table Data Set Original Data Set

Sonuçlar, büyükten küçüğe doğru sıralanarak bir tablo halinde kullanıcıya sunulur

Şekil 15: Kullanım Kılavuzu Adım 9

SONUÇ

Geliştirilen bu yazılım ile veri madenciliği tekniklerinden birliktelik analizinin, daha kolay ve pratik bir biçimde Microsoft Excel ortamında uygulanması sağlanmıştır. Diğer veri madenciliği tekniklerini kullanarak veri analizi yapan platformların aksine, kullanıcının önceden programla ilgili bilgi sahibi olmasına ya da deneyim sahibi olmasına gerek yoktur. Kullanıcı yararına olan bu durum, kullanıcının daha fazla aşına olduğu Excel ortamında küçük bir eklenti ekleyerek veri madenciliği yapmasına olanak vermiştir. Bu yazılım sayesinde kullanıcılar ellerindeki anlamsız ve büyük olan veri setinden anlamlı bir bilgi elde edeceklerdir. Bu bilgiler ışığında daha doğru ve daha stratejik kararlar almaları kolaylaşacaktır. Yazılımın uzun vadedeki hedefi; kullanıcıların ellerindeki anlamsız veriden anlamlı veriyi elde etmeleri için, bir uzmana ya da özellikli bilgilere ihtiyaç duymadan, kişisel olarak veri analizi yapabilmesidir.

KAYNAKLAR

- Akpınar H. (2000). *Veri Tabanlarında Bilgi Keşfi Ve Veri Madenciliği*, İstanbul Üniversitesi İşletme Fakültesi Dergisi, Cilt: 29.
- Agrawal R. & Srikant R. (1993). *Fast Algorithms For Mining Association Rules*, Conference On Very Large Databases.
- Ay D. (2009). *Veri Madenciliği Ve Apriori Algoritması İle Süpermarket Analizi*.
- Chen M. & Lin C. (2007). *Mining Approach To Product Assortment And Shelf Space Allocation*, Expert Systems with Applications, Sayı: 4, Cilt: 32.
- Chen Y. & Chen J. (2006). *A Data Mining Approach For Retail Knowledge Discovery With Consideration Of The Effect Of ShelfSpace AdjacencyOn Sales*, Decisions Support Systems, On Sales, Decisions Support Systems, Sayı: 3, Cilt: 42.
- Yang T. & Lai H. (2006). *Comparison Of Product Bundling Strategies On Different Online Shopping Behaviors*, Electronic Commerce Research and Applications, Sayı: 4, Cilt: 5.
- Eker H. (2016). *Veri madenciliğinde Apriori Algoritmasının Sınav Yerleri Üzerinde Uygulanması*.
- Liao S. & Hsieh C. & Huang, S. (2008). *Mining Product Maps For New Product Development*, Expert Systems with Applications, Sayı: 1, Cilt: 34.
- Liao S & Chen Y. (2004). *Mining Customer Knowledge For Electronic Catalog Marketing*, Expert Systems with Applications, Sayı: 4, Cilt: 27.
- Nazife B. (2003). *Veri Madenciliği Sunumu*, Orta Doğu Teknik Üniversitesi.
- Silahtaroglu G.(2014). *Veri Madenciliği*, Papatya Yayıncılık, ISBN: 978-975-6797-81-5.
- Thuraisingham B. (2003). *Web Data Mining And Applications in Businness Intelligence And Counter Terrorism*

ÖZ GEÇMİŞ

KİMLİK BİLGİLERİ

Adı Soyadı : Halil KAYACI
Doğum Yeri : KONYA
Doğum Tarihi : 02.08.1994
E-posta : halilkayaci@gmail.com

EĞİTİM BİLGİLERİ

Lise : Adil Karaağaç Anadolu Teknik Lisesi
Lisans : Pamukkale Üniversitesi Yönetim Bilişim Sistemleri Bölümü

İŞ DENEYİMİ

Delta Proje ve Danışmanlık – Proje Asistanı

Netrevart Bilişim Ar-Ge – Stajyer