



T.C.

DÜZCE ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÖNETİM BİLİŞİM SİSTEMLERİ ANABİLİM DALI

**BİRLİKTELİK KURALLARIYA KÜTÜPHANE ÖDÜNÇ
ALMA KAYITLARININ ANALİZİ: DÜZCE
ÜNİVERSİTESİ ÖRNEK UYGULAMASI**

SÜLEYMAN ÖZLÜK

YÜKSEK LİSANS TEZİ

DANIŞMAN

DR. ÖĞR.ÜYESİ FATİH KAYAALP

DÜZCE, 2022

TEŞEKKÜRLER

Tez yazma sürecinde yardım ve katkılarını esirgemeyen, ilgi, alaka ve rehberliğiyle bana yol gösteren danışmanım Dr. Öğr. Üyesi Fatih KAYAALP'a,

Tez çalışmam boyunca değerli katkılarını esirgemeyen Dr. Öğr. Üyesi Mehmet RAMAZANOĞLU'na ve diğer hocalarıma şükranlarımı sunarım.

Tez çalışmam boyunca yardımlarını ve desteklerini esirgemeyen sevgili aileme ve arkadaşlarıma sonsuz teşekkürlerimi sunarım.



Süleyman ÖZLÜK

ÖZET

BİRLİKTELİK ANALİZİYLE KÜTÜPHANE ÖDÜNC ALMA KAYITLARININ ANALİZİ: DÜZCE ÜNİVERSİTESİ ÖRNEK UYGULAMASI

Süleyman ÖZLÜK

Yüksek Lisans, Yönetim Bilişim Sistemleri Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Fatih KAYAALP

Kasım 2022, 86 Sayfa

Bilişim Teknolojileri, maliyetlerinin ucuzlaması dolayısıyla birçok kurum ve şirkette günümüzde yaygın olarak kullanılır hale gelmiştir. Bu sayede çeşitli sistemler içerisinde ortaya çıkan ve depolanan verilerin miktarı da çok büyük hacimlere ulaşmıştır. Bu verilerin içerisinde anlamlı bilgi çıkarma süreçleri olarak da tarif edilen Veri Madenciliği oldukça önem kazanmıştır. Yapılan bu çalışmada Düzce Üniversitesi Kütüphane Kayıtları veri kümesi üzerinde Veri Madenciliği yöntemlerinden biri olan Birliktelik Kuralları algoritmalarından Apriori, Fp-Growth ve Eclat algoritmalarının çeşitli destek ve güven değerlerine göre çalıştırılmalarıyla tekrar eden ilişkilerin/kuralların keşfedilmesi çalışması gerçekleştirilmiştir. Kitap seviyesinde tekrar eden kural örnekleri bulunamamıştır. Fakat kütüphanede bulunan kitap kategorileri seviyesinde beraber ödünç alınma örüntüleri görülmüş ve bunlara ilişkin bulunan kurallar ortaya konulmuştur. Aynı zamanda okuyucuların cinsiyetlerinin, kitap ödünç alma davranışlarına olan etkisi de araştırılmış ve bulunan sonuçlar değerlendirilmiştir. Bu kuralların bulunmasının yanı sıra belirtilen Birliktelik Kuralları algoritmalarının Python, R ve SPMF platformlarındaki performansları da deneyler sırasındaki çalışma süreleri (ms) ve işgal ettikleri bellek miktarları (MB) ölçütleri üzerinden karşılaştırılmıştır. Elde edilen sonuçlara göre SPMF programının erkekler veri setinde en az bellek işgaliyeti ve en düşük çalışma süresi değerleriyle göre öne çıktığı görülmüştür.

Anahtar kelimeler: Birliktelik Kuralları, Market Sepet Analizi, Apriori, SPMF, Python, R

ABSTRACT

ANALYSIS OF LIBRARY LOAN RECORDS BY ASSOCIATION RULES: DUZCE UNIVERSITY CASE STUDY

Süleyman ÖZLÜK

Master's Degree, Department of Management Information Systems

Thesis Advisor: Assist.Prof. Fatih KAYAALP

November, 2022, 86 Pages

Information Technologies has become widely used in many institutions and companies due to the cheapening of its costs. In this way, the amount of data generated and stored in various systems has reached very large volumes. Data Mining, which can be described as the processes of extracting meaningful information from these raw data, has gained importance in recent years. In this study, the discovery of association rules was carried out by running Apriori, Fp-Growth and Eclat algorithms, which are one of the Data Mining methods, on the Düzce University Library Records dataset, according to various support and confidence values. No examples of repetitive rules were found at the book level. However, at the level of the book categories in the library, patterns of borrowing together were discovered and the rules related to them were revealed. At the same time, the effect of the gender of the readers on their book loan behavior was also investigated and the results were evaluated. In addition to finding these rules, the performances of the specified Association Rules algorithms on Python, R and SPMF platforms were compared on the basis of their running time (ms) and the amount of memory they occupied (MB) during the experiments. According to the results obtained, it was seen that the SPMF program stood out in the men's data set with the least memory occupation and the lowest running time values.

Keywords: Association Rules, Market Basket Analysis, Apriori, SPMF, Python, R

İÇİNDEKİLER

JÜRİ ÜYELERİNİN İMZA SAYFASI	ii
TEŞEKKÜRLER	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKİLER	vi
TABLO LİSTESİ	ix
ŞEKİL LİSTESİ	x
KISALTMALAR	xii
1. GİRİŞ	1
1.1. Araştırmanın Problemleri	3
1.2. Araştırmanın Amacı	3
1.3. Araştırmanın Önemi	4
1.4. Araştırmanın Sınırlılıkları	5
2. LİTERATÜR	6
3. VERİ MADENCİLİĞİ	9
3.1. Veri Tabanlarında Bilgi Keşfi Süreci	9
3.2. Veri Tabanlarında Bilgi Keşfi Süreci Aşamaları	10
3.2.1. Veri Seçimi	10
3.2.2. Veri Ön İşleme ve Temizleme	10
3.2.3. Veri İndirgeme	11
3.2.4. Veri Madenciliği	11
3.2.5. Yorumlama ve Değerlendirme	11
3.3. Veri Madenciliği Nedir	11
3.4. Veri Madenciliği Süreci	12
3.4.1. İş Anlayışı	13
3.4.2. Veriyi Anlama	13
3.4.3. Veriyi Hazırlama	14
3.4.4. Modelleme	14
3.4.5. Değerlendirme	14
3.4.6. Uygulama	14
3.5. Veri Madenciliği Kullanım Alanları	14
3.6. Veri Madenciliği Modelleri	16
3.6.1. Tahmine Dayalı Modeller	17
3.6.1.1. Sınıflandırma Modeli	17
3.6.1.2. Regresyon Modeli	17
3.6.1.3. Sapma Tespiti Analizi	18
3.6.2. Tanımlayıcı Modeller	18

3.6.2.1.	Kümeleme Modeli	18
3.6.2.2.	Birliktelik Kuralları Modeli	19
4.	BİRLİKTELİK KURALLARI	20
4.1.	Birliktelik Kuralları Modeli	20
4.2.	Birliktelik Kurallarının Kullanıldığı Alanlar.....	20
4.2.1.	Pazar Sepeti Analizi.....	20
4.2.2.	Tıbbi Teşhis	21
4.2.3.	Nüfus Sayımı Verileri.....	21
4.2.4.	Protein Dizisi	21
4.3.	Birliktelik Kuralları Terimleri	21
4.3.1.	Destek Değeri	22
4.3.2.	Güven Değeri.....	22
4.3.3.	Kaldıraç Değeri.....	23
4.4.	Birliktelik Kuralları Algoritmaları	24
4.4.1.	Sıralı Algoritmalar	24
4.4.1.1.	AIS Algoritması.....	24
4.4.1.2.	SETM Algoritması.....	25
4.4.1.3.	Apriori Algoritması	26
4.4.1.4.	Apriori-TID Algoritması	30
4.4.1.5.	Apriori-Hybrid Algoritması.....	31
4.4.1.6.	OCD Algoritması.....	31
4.4.1.7.	Carma Algoritması.....	32
4.4.1.8.	FP-Growth Algoritması	32
4.4.2.	Paralel Algoritmalar.....	33
4.4.2.1.	CD Algoritmalar	33
4.4.2.2.	PDM Algoritmalar	34
4.4.2.3.	DD Algoritmalar	34
4.4.2.4.	IDD Algoritmalar.....	35
4.4.2.5.	Eclat Algoritmalar	35
5.	VERİ MADENCİLİĞİ PLATFORMLARI	37
5.1.	Python.....	37
5.2.	R Dili ve RStudio	38
5.3.	SPMF.....	39
6.	DENEY ÖN ÇALIŞMALARI.....	42
6.1.	Veri Kümesi	42
6.2.	Veri Ön İşleme	42
6.3.	Birliktelik Kuralları Deneyleri	42
6.4.	Performans Kriterleri.....	43
6.4.1.	Çalışma Süresi	43
6.4.2.	Kullanılan Bellek Boyutu	44

7. VERİ SETLERİNİN ANALİZİ	45
7.1. Veri Setlerin Yazılıma Hazırlanması	55
7.2. Arff Dosya Uzantısı	58
8. YÖNTEM.....	61
8.1. Evren ve Örneklem.....	61
8.2. Verilerin Toplanması.....	61
8.3. Verilerin Analizi.....	61
9. BULGULAR VE YORUM	62
9.1. Çalışma Sürelerinin Karşılaştırılması.....	62
9.2. Bellek Miktarlarının Karşılaştırılması.....	68
9.3. Deney Sonuçlarının Değerlendirilmesi	73
10. SONUÇLAR VE ÖNERİLER	78
11. KAYNAKÇA	80
12. EKLER.....	87



TABLO LİSTESİ

Tablo 1. Bazı Sektörlerin VM Kullanım Alanları.....	15
Tablo 2. Destek Değeri Örneği	22
Tablo 3. Güven Değeri Örneği.....	23
Tablo 4. Apriori Algoritması 1. Aşama	29
Tablo 5. Apriori Algoritması 2. Aşama	29
Tablo 6. Apriori Algoritması 3. Aşama	30
Tablo 7. Deneylerde Kullanılan Destek ve Güven Değerleri Tablosu.....	43
Tablo 8. Tüm Veri Seti Kitap Kategori Oranları	45
Tablo 9. Erkekler Veri Seti Kitap Kategori Oranları	48
Tablo 10. Bayanlar Veri Seti Kitap Kategori Oranları	51

ŞEKİL LİSTESİ

Şekil 1. Veri Madenciliği Bilgi Keşfi Süreci.	10
Şekil 2. Veri Madenciliğinin Kapsam Bileşenleri.	12
Şekil 3. Veri Madenciliği Süreci.	13
Şekil 4. Veri Madenciliği Modelleri.	16
Şekil 5. Kümeleme Modeli.	18
Şekil 6. Birliktelik Kuralları Algoritmaları.	24
Şekil 7. AIS Algoritması Sözde Kodu.	25
Şekil 8. SETM Algoritması Sözde Kodu.	26
Şekil 9. Apriori Algoritması Sözde Kodu.	27
Şekil 10. Apriori Algoritması Akış Diyagramı	28
Şekil 11. Apriori-TID Algoritması Sözde Kodu	31
Şekil 12. FP-Growth Algoritması Sözde Kodu.....	33
Şekil 13. CD Algoritması Sözde Kodu.	34
Şekil 14. DD Algoritması Sözde Kodu.	35
Şekil 15. Eclat Algoritması Sözde Kodu.....	36
Şekil 16. Veri Biliminde, Makine Öğrenimi Yazılımları 2017- 2019 KDnuggets Anketi.....	37
Şekil 17. Python Apriori Algoritması Kullanıcı Arayüzü.....	38
Şekil 18. RStudio Kullanıcı Arayüzü.....	39
Şekil 19. SPMF Kullanıcı Arayüzü.....	40
Şekil 20. SPMF Apriori Algoritması	40
Şekil 21. Kategorilerin Veri Setinden Çekilmesi İşlemi.....	55
Şekil 22. Tüm Veri Seti Düzenlenmiş Hali	56
Şekil 23. Veri Setine “t” Eklenmesi İşlemi.....	56
Şekil 24. Veri Setine “t” Eklenmesi İşlemi 1.....	57
Şekil 25. Veri Setine “?” Eklenmesi İşlemi	57
Şekil 26. Veri Setine “t” Eklenmesi İşlemi 1.....	58
Şekil 27. Tüm Veri Setinin. Arff Dosyası Olarak Düzenlenmiş Hali-1.....	59
Şekil 28. Tüm Veri Setinin. Arff Dosyası Olarak Düzenlenmiş Hali-2.....	60
Şekil 29. Tüm Veri Setinin. Arff Dosyası Olarak Düzenlenmiş Hali-3.....	60

Şekil 30. Tüm Veri Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması	62
Şekil 31. Tüm Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması	63
Şekil 32. Tüm Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması	63
Şekil 33. Erkekler Veri Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması	64
Şekil 34. Erkekler Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması	65
Şekil 35. Erkekler Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması	65
Şekil 36. Bayanlar Veri Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması	66
Şekil 37. Bayanlar Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması	67
Şekil 38. Bayanlar Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması	67
Şekil 39. Tüm Veri Setinde Apriori Algoritmasının Bellek Miktarlarının Karşılaştırılması	68
Şekil 40. Tüm Veri Setinde FP-Growth Algoritmasının Bellek Miktarlarının Karşılaştırılması	69
Şekil 41. Tüm Veri Setinde Eclat Algoritmasının Bellek Miktarlarının Karşılaştırılması	69
Şekil 42. Erkekler Veri Setinde Apriori Algoritmasının Bellek Miktarlarının Karşılaştırılması	70
Şekil 43. Erkekler Veri Setinde FP-Growth Algoritmasının Bellek Miktarlarının Karşılaştırılması	70
Şekil 44. Bayanlar Veri Setinde Eclat Algoritmasının Bellek Miktarlarının Karşılaştırılması	71
Şekil 45. Bayanlar Veri Setinde Apriori Algoritmasının Bellek Miktarlarının Karşılaştırılması	71
Şekil 46. Bayanlar Veri Setinde FP-Growth Algoritmasının Bellek Miktarlarının Karşılaştırılması	72
Şekil 47. Bayanlar Veri Setinde Eclat Algoritmasının Bellek Miktarlarının Karşılaştırılması	72
Şekil 48. SPMF Apriori Algoritması Destek: 0,01, Güven: 0,1 Değeri İçin Kural Sonuçları	77

KISALTMALAR

AIS	Agrawal İmielinski Swami
ARFF	Attribute Relationship File Format
CD	Collison Detection
DD	Data Distribution
IDD	Intelligent Data Distribution
MSA	Market Sepet Analizi
OCD	Offline Nomination
PDM	Paralel Data Mining
SETM	Set Oriented Mining
VM	Veri Madenciliđi
VTBK	Veri Tabanlarında Bilgi Keşfi
WEKA	Waikato Environment for Knowledge Analysis

1- GİRİŞ

Bilgi kaynaklarının hacmi, iş süreci yönetimi gibi yaşamın tüm alanlarını geniş ölçüde etkileyen veritabanı teknolojisi, ağ teknolojisi ve dijital kütüphane teknolojisinin sürekli gelişimi ve kapsamlı uygulamasıyla birlikte giderek daha da büyümektedir (Liu vd., 2017). Akademik kütüphane yönetimi, tıbbi sağlık, bankacılık, pazarlama, E-Ticaret, sigortacılık alanında bilgi hacminin küresel olarak 20 ayda bir ikiye katlandığı tahmin edilmektedir (Kim, 2022). Ancak bu verilerin depolanması ve etkin bir şekilde nasıl kullanacağı sorun haline gelmiştir. İnsanlar genellikle bu kadar büyük bilgi kaynakları karşısında kaybolurlar ve ihtiyaç duyulan bilgiyi hızlı ve etkili bir şekilde bulmakta zorlanırlar. Bu da bol veriye sahip olunmasına rağmen bilgi kıtlığına neden olur.

Günümüzde, birçok kütüphanede Kütüphane Otomasyon Sistemleri kullanılmakta ve okuyucuların kütüphane kaynaklarını ziyaret geçmişleri hakkında çok sayıda veri toplanmaktadır. Bu tür verilerin arkasında pek çok önemli bilgi gizlidir ve bu gizli bilgileri ortaya çıkarmak için Veri Madenciliği (VM) kullanılmaktadır. Bu süreçte VM tekniklerinden olan Birliktelik Kuralları, büyük miktarlardaki veriler içerisinden tanımlayıcı veri öğeleri arasındaki karşılıklı ilişkileri çıkarmak için gerekli bilgiyi içeren önemli teknolojilerden biridir (Alpaydın, 2000). VM modelleri tahmine dayalı ve tanımlayıcı olmak üzere ikiye ayrılır. Tahmine dayalı modellerden Sınıflandırma, Regresyon ve Sapma Tespiti Analizi; tanımlayıcı modellerden ise Kümeleme, Birliktelik Kuralları ve Sıralı Örüntü Keşfi en çok kullanılan veri madenciliği yöntemlerindendir.

Birliktelik Kuralları yöntemi, bir veri kümesinde sıklıkla tekrar eden öğelerin saptanması ve veri kümesindeki öğeler arasındaki bağlantıların ve ilişkilerin (örüntü) tespit edilmesi için kullanılmaktadır. Birliktelik kuralları, verilerin davranışlarının analizi ile geleceğe yönelik çalışmalar yapılmasını destekleyen tanımlayıcı bir modeldir. Birliktelik Kurallarının en tipik örneği Pazar Sepet Analizidir (Market Basket Analysis). Verilerin çoğu barkod tarayıcılar kullanılarak toplanır. Pazar sepeti veri tabanı olarak bilinen bu veri tabanı, geçmiş işlemlerle ilişkin çok sayıda kayıttan oluşmaktadır. Her bir kayıta, bir müşteri tarafından bir satışta satın alınan tüm ürünler listelenir. Hangi müşteri gruplarının hangi ürün grubuna meyilli olduğunu bilmek,

marketlerin market düzeninin ve market kataloğunun birbiriyle en uygun şekilde yerleştirilmesi için gerçek hayatta da sıklıkla uygulanmakta olan bir yöntemdir.

Birliktelik analizinin kendi içerisinde birçok algoritma bulunmaktadır. Bu algoritmalar Sıralı algoritmalar ve Paralel algoritmalar olarak iki kategoriye ayrılır. Sıralı algoritmalar AIS, SETM, Apriori, Apriori-TID, Apriori-Hybrid, OCD, Carma, FP-Growth algoritmalarıdır. Paralel algoritmalar ise CD, PDM, DD, IDD, Eclat algoritmalarıdır. Birliktelik kuralları algoritmalarında en çok kullanılan algoritmalar Apriori, FP-Growth ve Eclat algoritmalarıdır (Knatardzic, 2011). Birliktelik kuralları mühendislik, tıp, eğitim, bankacılık, pazarlama, sigortacılık, finans, nüfus sayım verilerinde, protein dizisinde, perakende gibi alanlarda sıklıkla kullanılmaktadır (Olson ve Delen, 2008).

Bu tez çalışmasında, Düzce Üniversitesi lisans ve lisansüstü öğrencileriyle, akademik ve idari personellerinin kütüphaneden kitap ödünç alma kayıtları içerisinde sıklıkla tekrar eden ödünç alma desenlerini keşfetmek için Birliktelik Analizi algoritmalarından olan Apriori, FP-Growth ve Eclat algoritmaları farklı destek ve güven değerleriyle R, Python ve SPMF VM platformlarında uygulanmıştır. Yapılan deneyler neticesinde Apriori, FP-Growth ve Eclat algoritmalarının her 3 platformdaki çalışma süreleri ve deneyler sırasında kullandıkları bellek miktarlarının ortalamaları grafiklerle karşılaştırılmıştır. Aynı zamanda bu süreçte kitap ödünç alma kuralları da keşfedilmeye çalışılmış fakat kitap/lar seviyesinde beraber ödünç alınma kuralları bulunamamıştır. Yapılan detaylı incelemelerde kütüphanede bulunan kitap kategorileri arasında tekrar eden beraber ödünç alma kuralları görülmüştür. Bu kurallar aynı zamanda okuyucu cinsiyetleri açısından da değerlendirilerek kitap okuma davranışları değerlendirilmiştir.

Bu tezin ikinci bölümünde literatür araştırması kısmında Birliktelik analiziyle ilgili IEEE, Science Direct vb. popüler veritabanlarında son 5 yılda yayınlanmış yüksek lisans/doktora tezi, makale ve konferans bildirisi şeklindeki çalışmalar incelenmiştir.

Üçüncü bölümde ise veri tabanlarında bilgi keşfi sürecinin aşamaları detaylı bir şekilde incelenmiştir. Sonrasında da Veri madenciliği, veri madenciliği süreci, kullanım alanları, modellerinden bahsedilmiştir.

Dördüncü bölümde ise VM modellerinden en çok kullanılan Birliktelik Kuralları incelenmiştir. Birliktelik kuralları terimlerinden destek değeri, güven değeri ve kaldıraç değeri örnekler ile açıklanmıştır.

Beşinci bölümde, popüler VM platformlarından Python, R Dili- RStudio ve SPMF ile ilgili bilgiler verilip örnek bir Apriori algoritması ile incelenmiştir.

Altıncı bölümde tez çalışmasında kullanılan veri seti hakkında bilgiler (temin edilen yer, kayıt sayısı, öznitelikler ve açıklamaları, veri türleri vb.), önışleme işlemleri ve deneylerde kullanılan performans kriterleriyle (çalışma süresi ve kullanılan bellek miktarı) ilgili bilgiler verilmiştir.

Yedinci bölümde veri setlerinin VM platformlarında kullanılabilmesi için hazırlanması aşamalarıyla ilgili bilgiler verilmiştir

Sekizinci bölümde tez çalışmamızın yönteminden, araştırma modelinden, evren ve örneklemden, verilerin toplanması ve analizinden bahsedilmiştir.

Dokuzuncu bölümde ise bahsedilen veri setine Python, RStudio ve SPMF platformları aracılığıyla Apriori, FP-Growth ve Eclat algoritmalarının uygulanmasıyla yapılan deneylerden elde edilen sonuçlar tablolar ve grafikler aracılığıyla incelenmiş ve en sonda ise deney sonuçları değerlendirilmiştir.

Onuncu bölümde ise sonuçlar ve öneriler detaylı bir şekilde sunulmuştur.

Araştırmanın Problemleri

- Düzce Üniversitesi kütüphanesi kitap ödünç alma veri setinden farklı Birliktelik Kuralları algoritmalarıyla kural keşif süreçleri, çalışma süresi ve kullanılan bellek miktarı bakımından farklılık göstermekte midir?
- Düzce Üniversitesi lisans ve lisansüstü öğrencileriyle akademik ve idari personellerinin kütüphaneden kitap ödünç alma tercihlerinde, cinsiyete göre farklılık gözlemlenmekte midir?
- Kullanımına karar verilmiş olan veri kümesi üzerindeki belirtilen Birliktelik Kuralları algoritmalarının uygulanması, işlem süresi ve bellek işgalietisi açısından kullanılan platforma göre farklılık göstermekte midir?

Araştırmanın Amacı

Bu çalışmanın amacı Düzce Üniversitesi lisans ve lisansüstü öğrencileriyle akademik ve idari personellerinin kütüphaneden kitap ödünç alma kayıtları içerisinde sıklıkla tekrar eden ödünç alma kurallarının Python, R ve SPMF platformlarında Birliktelik Analizi algoritmalarından Apriori, FP-Growth ve Eclat algoritmalarıyla tespit edilmesi ve bu süreçte hem algoritmaların hem de platformların performanslarının çalışma süreleri (ms) ve kullanılan bellek miktarları (MB) açısından karşılaştırmalı olarak değerlendirilmesidir. Bu işlemler eldeki veri kümesinin tümü, erkeklere dair kayıtlar ve kadınlara dair kayıtlar şeklinde 3 ayrı veri kümesi üzerinde de uygulanarak aynı zamanda erkeklerin ve kadınların kitap ödünç alma davranışları da değerlendirilmiştir. Elde edilecek kurallara göre kütüphanedeki kitapların dizilişi konusunda tavsiyelerde bulunulabilecektir.

VM 'de Birliktelik Kuralları algoritmalarını detaylı bir şekilde incelemek ve istatistiksel çözümlemelere ağırlık vererek bir kütüphane uygulaması gerçekleştirip çıkan sonuçları değerlendirmektir.

Araştırmanın Önemi

Bu çalışmanın önemi, kütüphane kitap ödünç alma kayıtları veri kümesine birliktelik kuralları algoritmalarından Apriori, FP-Growth ve Eclat'ın Python, RStudio ve SPMF platformlarında uygulanarak çıkan kuralların tespit edilip bu kuralların cinsiyete göre farklılık gösterip göstermediğinin değerlendirilmesinin yanı sıra algoritmalar ve platformların çalışma süresi (ms) ve kullanılan bellek miktarları (MB) ölçütleri ışığında farklılıklar gösterip göstermediğinin ortaya konmasıdır. Diğer bir önemi ise önümüzdeki yıllarda çalışmamızda kullandığımız Apriori, FP-Growth ve Eclat algoritmalarından herhangi birisinin kullanımına ihtiyaç duyulması halinde öğrenciler tarafından literatürde yer alan algoritma ve programlardan, doğru algoritmanın ve doğru programın seçilmesine katkı sağlamaktır. Çalışmanın sonucunda elde edilen sonuçların Düzce Üniversitesi Lisans/ Lisansüstü, akademik ve idari personele kitap karar verme aşamalarında yol göstermesi veya kütüphanedeki kitapların dizilişleri konusunda fikir verme faydaları hedeflenerek örnek bir kaynak olması amaçlanmıştır.

Araştırmanın Sınırlılıkları

- Deneyler Düzce Üniversitesi kütüphanesi kitap ödünç alma veri kümesi ile yapılmıştır.
- Birlikte kurallarında yapılan deneyler ilerleyen bölümlerde yazılımsal ve donanımsal özellikleri verilmiş olan sadece tek bir bilgisayardan yapılmıştır.
- Sadece Apriori, Fp-Growth ve Eclat algoritmaları uygulanmıştır.
- Sadece Python, R/RStudio ve SPMF platformları kullanılmıştır.



2- LİTERATÜR

Birliktelik Analiziyle ilgili YÖK Tez Bankasında ve IEEE, Science Direct vb. popüler veritabanlarında yayınlanmış yüksek lisans/doktora tezi, makale ve konferans bildirisi şeklindeki çalışmalar incelenmiş ve bunlarla ilgili derleyici bilgiler aşağıda sunulmuştur.

2019 yılında Tokyürek tarafından yapılan yüksek lisans çalışmasında, birliktelik kuralları algoritmalarını kullanarak bir mağazaya ait 4625 kayıt ve 106 özellikten oluşan verilerin analizini gerçekleştirmiştir. Mağaza deposu verileri, FP-Growth ve Apriori algoritmalarının farklı destek ve güven değerleriyle Weka programında çalıştırılmalarıyla analiz edilmiş ve çıkan sonuçlar karşılaştırılarak değerlendirilmiştir (Tokyürek, 2019).

Diğer bir çalışmada, Türk Hava Yollarına ait 2016 yılındaki yaz ve kış dönemlerine ait veriler analiz edilmiştir (Uluyardımcı ve Zontul, 2019). 2036113 kayıttan oluşan veri kümesi, Apriori algoritması ile analiz edilmiş ve 824 yaygın öge kural sayısı çıkarılmıştır. Sonuç olarak çıkan yaygın öge kuralları değerlendirilerek müşteriye/kuruma olumlu fayda sağlayabilecek değerlendirmelerde bulunulmuştur.

2021 yılında Karademir tarafından yapılan yüksek lisans çalışmasında, içerisinde 82300 kayıt bulunan veri kümesine Apriori ve FP-Growth algoritmalarını uygulayarak bulduğu sonuçları karşılaştıran bir market sepet analizi gerçekleştirmiştir. Apriori ve FP-Growth algoritmaları Python ve Weka yazılımlarında çalıştırılmıştır. Elde edilen algoritma sonuçları ile market işletmecisine hangi ürünlerin ne sıklıkla alındığı sunulmuştur (Karademir, 2021).

2019 yılında Nair tarafından yapılan yüksek lisans çalışmasında birliktelik kuralları ile bir otomotiv yedek parça şirketinin 6 aylık (167,334 kayıt), 12 aylık (203,753 kayıt) ve 22 aylık (543,316 kayıt) verilerini Weka ve SPMF yazılımları ile analiz etmiştir. Bu çalışmada ilk önce Weka ve SPMF yazılımları karşılaştırılıp SPMF yazılımının daha verimli olduğu analizler sonucunda bulunmuş; sonraki aşamada ise SPMF yazılımı üzerinde Apriori, AprioriClose, AprioriRare, AprioriTID, Charm_bilset, Eclat, Eclat_bilset, FPClose, FP-Growth, dEclat, dEclat_bilset algoritmaları uygulanarak incelenen performans parametreleri (çalışma süresi ve

bellek işgaliyeti) doğrultusunda dEclat_bilset algoritmasının performans değerlendirmesinin en verimli olduğu değerlendirilmiştir (Nair, 2019)

Birlikte kurallarında Market Sepet Analizi (MSA) bir VM tekniğidir. VM, pazarlama, eğitim alanı, biyoinformatik, nükleer bilim gibi alanlarda kullanılmaktadır. Sunulan çalışmada Extended Bakery veri kümesi kullanılmıştır. Veri kümesi 2000 kayıt ve 26 sütundan oluşmaktadır. Sonuç olarak VM tekniği tartışılıp elde edilen kurallara göre satış oranlarının artırılmasına yönelik değerlendirmeler yapılmıştır (Kaur ve Kang, 2016).

Diğer bir çalışmada kalp hastalığını tahmin etmek için tıbbi bir veri kümesi üzerine VM birliktelik kurallarından Apriori algoritması, Miner (SDM) platformu kullanarak uygulanmıştır. Veri kümesi 76 öznitelikten oluşmaktadır. Fakat 76 öznitelikten kalp hastalığına etkisinin daha fazla olduğu analiz edilmiş olan 14 öznitelik seçilerek işleme alınmış ve elde edilen kurallara göre yorumlar yapılmıştır. Hem erkeklerde hem de kadınlarda kalp hastalığı için önemli risk faktörleri ortaya konulmuştur (Said vd., 2015).

2018 yılında Taş tarafından yapılan yüksek lisans çalışmasında, perakende satış işlemi yapan bir mağazaya ait 10000 kayıtlı bir veri kümesine WEKA platformunda FP-Growth algoritmasını uygulayıp sonuçları analiz etmiş ve çalışmanın sonucunda elde edilen sonuçlara göre perakende mağaza üzerine satış stratejileri önerilerinde bulunmuştur (Taş, 2018).

Çoban ve arkadaşları tarafından yapılan çalışmada Atatürk Üniversitesi Araştırma Hastanesi Acil Servis hastalarından elde edilen ve kişisel bilgileri çıkarılmış 20961 kayıtlık veri kümesine Apriori algoritması uygulanmıştır. Sonuç olarak acil servise gelen hastaların profilleriyle ilgili anlamlı bilgiler ortaya çıkarılmıştır (Çoban vd., 2015).

Diğer bir çalışmada Romanya KOBİ'lerinin sürdürülebilirlik için karşılaştıkları temel engelleri ve kolaylaştırıcıları tespit etmek için 71 firmanın veri kümesi RapidMiner programında FP-Growth algoritması ile analiz edilmiştir. Elde edilen analizler sonucunda KOBİ'lerin, profesyonellerin ve politika yapımcıların sorunlu uygulamaları etkileyen faktörleri daha iyi anlamalarına yardımcı olunmuştur (Costache vd., 2021).

Krisnanto ve arkadaşları 2022’de yapmış oldukları çalışmada VM tekniği kullanarak ürün kalemleri arasındaki uygulamaların oluşumuna yardımcı olmak için 2020 yılındaki ürün satış veri kümesi üzerine RapidMiner programı kullanarak Apriori algoritmasını uygulamışlardır. Analizler sonucunda promosyon yoluyla satış cirosunu artırma stratejisinin yürütülmesinde ürün önerilerinde referans olarak kullanılmıştır (Krisnanto vd., 2022)

Literatür taramalarında görüldüğü üzere Birliktelik Kuralları çoğunlukla Pazar Sepeti analizinde kullanıldığı görülmüştür. Pazar sepet analizinde Apriori, FP-Growth algoritmaları kullanılmıştır, Eclat algoritmaları da azda olsa kullanılmıştır. Birliktelik kurallarının kütüphane ödünç alma verileri üzerine uygulanarak ödünç alma davranışlarının analizinin yapımı ve bu işlemlerin birden fazla algoritma/platform üzerinde uygulanarak birbirleriyle karşılaştırılması tarzında açık olduğu görülerek bu çalışmanın iskeleti planlanmıştır.

3. VERİ MADENCİLİĞİ

Veri madenciliği süreçlerine dair bilgiler aşağıdaki alt başlıklar ile verilmiştir.

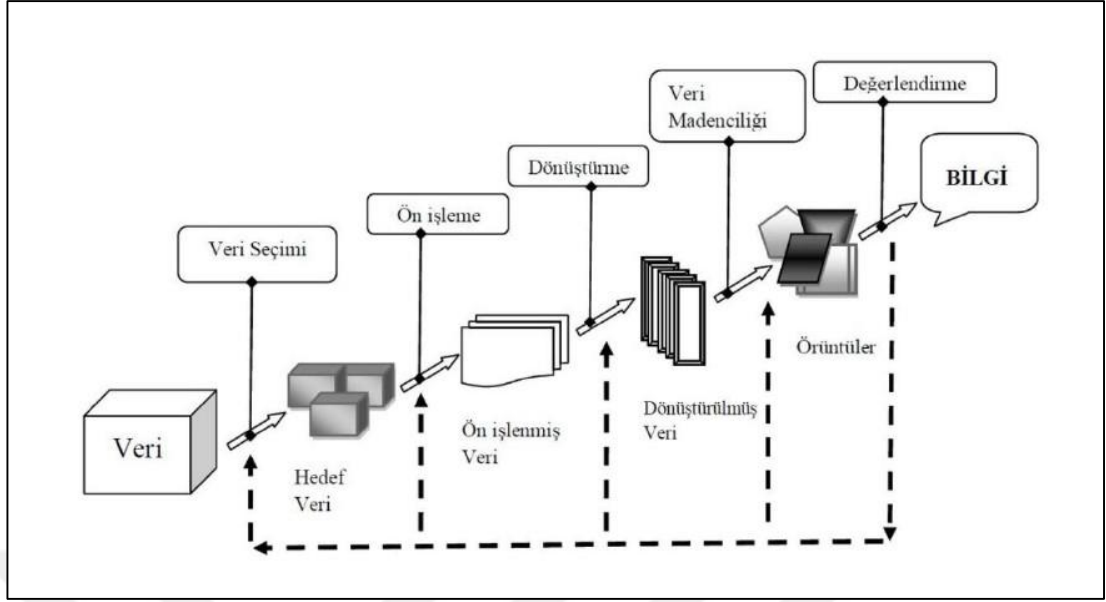
3.1. Veri Tabanlarında Bilgi Keşfi Süreci

Veri tek başına ve ham halde bir anlam ifade etmez. Verinin bilgiye dönüştürülebilmesi için verinin işlenip anlamlı ve işe yarar hale getirilmesi gerekir. Bu işlemler organize bir tanımlama sürecidir. Bu süreç Veritabanlarında Bilgi Keşfi (VTBK) olarak isimlendirilir. VTBK, büyük ve karmaşık veri kümelerinden geçerli, yeni, kullanışlı ve anlaşılabilir desenlerin elde edilmesi süreci olarak tanımlanabilir (Şen, 2008).

1996'da popüler hale gelen VTBK modeli, hacimli veri kümelerinden bilgi keşfetmek için hesaplama tekniklerini ve araçlarını kullanma sürecini tanımlar. Süreç, analistin hedeflerini belirlemesi ve gerekli alan edinmesiyle başlar. Verileri dönüştürmek ve yeni bilgileri ortaya çıkarması beklenen kalıpları çıkarmak için VTBK'nin diğer 5 aşaması boyunca devam eder. Keşfedilen bilginin görselleştirilmesi ve izleyiciye sunulmasıyla sona erer (Fayyad 1996).

Günümüzde verilerin erişilebilirliği ve bolluğu, bilgi keşfinin ve Veri Madenciliğini oldukça önemli ve gerekli bir konu haline getirmektedir. Alanın son zamanlardaki büyümesi göz önüne alındığında, araştırmacılar ve uygulayıcılar için artık çok çeşitli yöntemlerin mevcut olması şaşırtıcı değildir. Bilgi çağından itibaren verinin toplanması kolaylaşmış ve saklanması ucuz hale gelmiştir. Depolanan bilgi miktarının her yirmi ayda bir ikiye katlandığı öngörülmektedir (Rokachi, Maimon 2008). Şekil 1'de veri madenciliği bilgi keşfi süreci aşamaları verilmiştir.

Şekil 1. Veri Madenciliği Bilgi Keşfi Süreci (Hudairy, 2004)



3.2. Veri Tabanlarında Bilgi Keşfi Süreci Aşamaları

Veritabanlarında bilgi keşfi aşamalarıyla ilgili bilgiler aşağıdaki alt başlıklar ile verilmiştir.

3.2.1. Veri Seçimi

VTBK'nin ilk sürecidir. Bu süreçte veriye karar verilerek başlanır. Hatalı ve veri sürecini yanlış yöne taşıyacak veriler temizlenir. Eksik ve hatalı veriler var ise tahmin yolu ile bulmaya çalışılır. Bu aşamanın sonucunda çıkarılacak bilgiyi son derece önemli ölçüde etkileyeceği için kullanıcının isteğine göre verilerin seçilmesine dikkat edilmesi gerekir (Akpınar, 2000).

3.2.2. Veri Ön İşleme ve Temizleme:

VTBK'nin en çok zaman alan sürecidir. VM 'de kullanılacak veri kümesi belirlendikten sonra verilerin temizlenmesi ve istenilen formata uygun olarak biçimlendirilmesi gerekir. Eksik veriler doğrudan kullanılırsa prosedürlerde karışıklığa neden olabileceği ve hatalı/yanlış sonuçlar üretebileceği için önem taşımaktadır. Veri ön işleme ve temizleme aşamasının doğru ve başarılı olması kesin sonuçlara ulaşmaya imkân sağlamaktadır (Özdemir, Yalçın, Çam, 2009).

3.2.3. Veri İndirgeme:

Veri seti üzerinde doğru ve faydalı sonuçlar elde etmek için verilerin indirgenmesi gerekir. Seçilmiş olan veri setinden tekrar eden ve gereksiz verilerin indirgendiği aşamadır. Yapılan bu işlem ile veri setinin boyutu indirgendiğinde veri madenciliğine uygulanacak sorguların daha hızlı sonuç üretmeleri sağlanabilir (Özdemir, Yalçın, Çam, 2009).

3.2.4. Veri Madenciliği:

VM, VTBK sürecinin en önemli aşamasıdır. Veri seti bu aşamada artık veri madenciliği yöntemleri ve algoritmaları ile uygulanabilir haldedir. Veri Madenciliği, büyük miktarda veriden ilginç kalıpları ve bilgileri belirleme sürecidir. Bu adımlarda, veri kalıplarını çıkarmak için akıllı kalıplar uygulanır. Veriler, örüntüler şeklinde temsil edilir ve modeller, sınıflandırma ve kümeleme teknikleri kullanılarak yapılandırılmaktadır (Aytaç, Bilge, 2013).

3.2.5. Yorumlama ve Değerlendirme:

Bilgi keşfi sürecinin son aşamasıdır. Verilerin kullanıcı tarafından anlaşılır hale getirilmesi için veri özetleme ve görselleştirme yöntemleri kullanarak yorumlandığı ve değerlendirildiği bir aşamadır. Elde edilen sonuçların günlük hayatta kullanılmasına odaklanır (Fayyad, 1996).

3.3. Veri madenciliği Nedir

Günümüzde bilgi teknolojilerinin etkin ve verimli bir şekilde kullanılması birçok sektör tarafından kaçınılmaz bir ihtiyaç haline gelmiştir. Bu süreçte depolanan verinin büyüklüğü de artmıştır. Depolanmış olan bu verilerden çıkarım yapma çalışmalarına verilen bir isim olarak VM oldukça önem kazanmıştır. VM en genel tanımı ile büyük boyutlu verilerin toplanması, analiz edilmesi ve aralarından anlamlı bilgilerin çıkarımı olarak tanımlanabilir (Morais vd., 2017). VM kapsam bileşenleri şekil 2’de verilmiştir.

Şekil 2. Veri Madenciliğinin Kapsam Bileşenleri (Baykasoğlu, 2005)



Şekil 2’de görüldüğü gibi VM, verileri analiz etmek için makine öğrenimi, istatistik, yapay zekâ ve veritabanı teknolojileri gibi farklı disiplinlerden yararlanır.

VM, büyük veri kümelerinin içerisinde keşfedilmeyen bilgilerin, ön tahminleri ortaya çıkaramadığımız, ayrıca keşfedilmesi halinde ihtiyacı olan birime büyük fayda sağlayarak faydalı bilginin ortaya çıkarma sürecidir (Silahtaroglu, 2016)

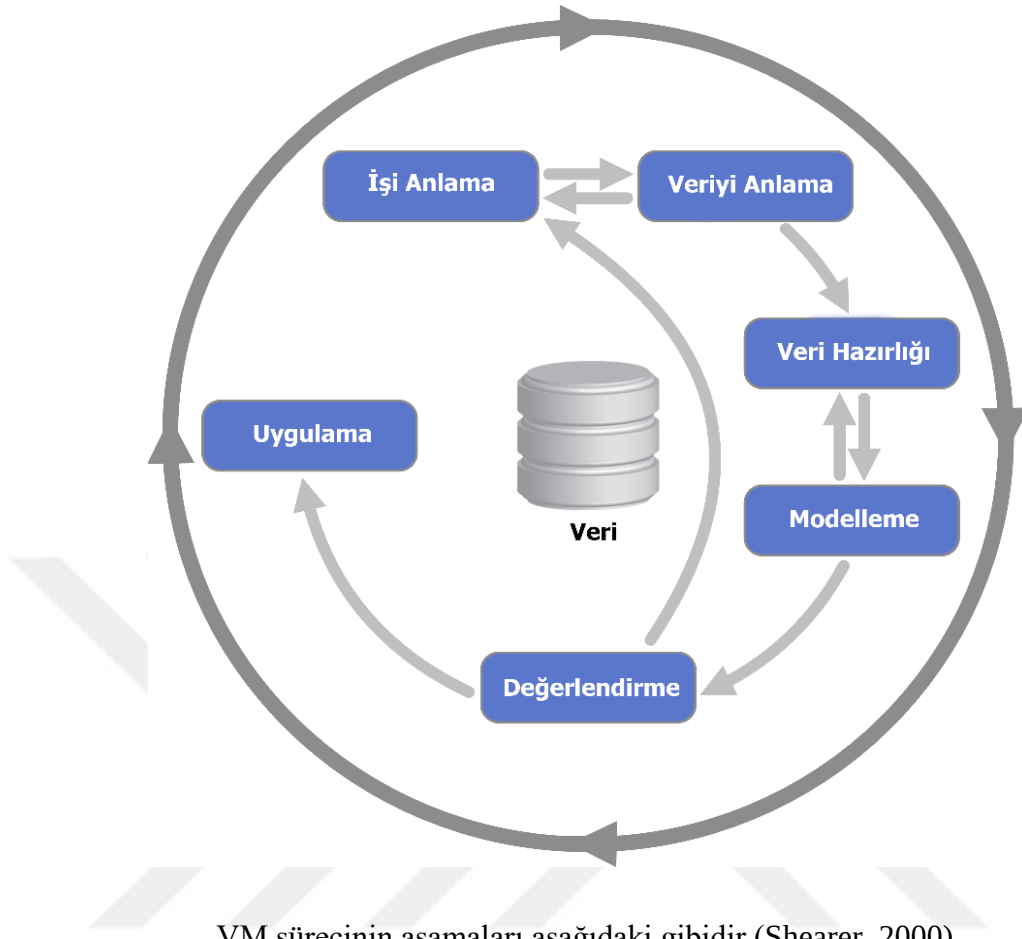
İçerisinde milyonlarca kayıt olan bir veri setinin manuel olarak analiz edilmesi imkânsız bir iştir. Bu işlemin daha hızlı, daha az maliyetli ve otomatik olarak yapılabilmesi daha mantıklıdır (Nair, 2019).

VM, basit bir deyişle, büyük bir veri setinden kullanılabilir bilgileri çıkarma süreci olarak tanımlanır. VM pazarlama, sağlık, E-Ticaret, bankacılık, eğitim gibi birçok alanda uygulanmaktadır (Nair, 2019).

3.4. Veri Madenciliği Süreci

CRISP-DM olarak da adlandırılan metodolojiye göre bir VM süreci şekil 4’de de görüldüğü üzere döngüsel olarak 6 adımdan meydana gelir. Bu adımlar bir iş sorununu anlamak, verileri anlamak, verileri toplayıp hazırlamak, madencilik modellerini oluşturmak, elde edilen sonuçları değerlendirmek ve sonuçların başarılı/verimli olması durumunda ilgili modellemenin uygulamaya alınmasıdır. VM süreci şekil 3’te verilmiştir.

Şekil 3. Veri Madenciliği Süreci (Nair, 2019)



3.5.1. İş Anlayışı

Bir VM çalışmasında, işletmelerin hedefleri belirlenir ve bu hedefe ulaşmada yardımcı olacak önemli faktörler keşfedilir.

3.4.2. Veriyi Anlama:

Bu adım, tüm verileri toplayacak ve (herhangi bir araç kullanılıyorsa) araçtaki verileri dolduracaktır. Veriler, veri kaynağı, konumu, nasıl elde edildiği ve herhangi bir sorunla karşılaşıp karşılaşılmadığı ile listelenir. Veriler görselleştirilir ve eksiksizliğini kontrol etmek için sorgulanmaktadır.

3.4.3. Veriyi Hazırlama

Bu adım, uygun verilerin seçilmesini, temizlenmesini, verilerden özniteliklerin oluşturulmasını ve birden çok veri tabanından verilerin bütünleştirilmesini içermektedir.

3.4.4. Modelleme

Karar ağacı, seçilen modeli değerlendirmek için test tasarımı oluşturma, veri setinden modeller oluşturma ve oluşturulan modelin uzmanlarla sonucu tartışmak için değerlendirilmesi gibi veri madenciliği tekniğinin seçimi bu adımda yapılmaktadır.

3.4.5. Değerlendirme

Bu adım, ortaya çıkan modelin iş gereksinimlerini karşılama derecesini belirleyecektir. Model gerçek uygulamalar üzerinde test edilerek değerlendirme yapılabilir. Model, tekrarlanması gereken herhangi bir hata veya adım için gözden geçirilmektedir.

3.4.6. Uygulama

Bu adımda bir dağıtım planı yapılır. Veri madenciliği modeli sonuçlarını izleme ve sürdürme stratejisi, yararlılığını kontrol etmek için oluşturulur, nihai raporlar yapılır ve herhangi bir hatayı kontrol etmek ve görmek için tüm sürecin gözden geçirilmesi yapılmaktadır.

3.5. Veri Madenciliği Kullanım Alanları

VM kullanım alanları, son yıllarda değişen dünya piyasasında finansal ve ekonomik koşulların sonucunda rekabetin artmasıyla, pazarlama, bankacılık, perakendecilik, telekomünikasyon, bilim ve mühendislik, sağlık, kriminoloji, eğitim, e-ticaret, spor, turizm, web, işletmeler, kütüphanecilik ve daha birçok alanda rastlanmıştır. Tablo 1’de bazı sektörlerin VM kullanım alanları verilmiştir.

Tablo 1. Bazı Sektörlerin VM Kullanım Alanları (Baykal, 2006)

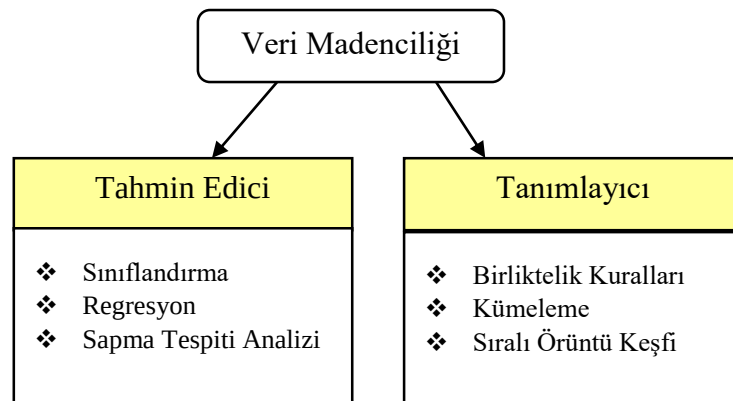
Sektörler	VM 'nin Kullanım Amacı
Pazarlama	• Müşteri ilişkileri yönetimi • Satış tahmini, müşteri dağılımı • Müşterilerin Mail kampanyalarına katılımın artırılması • Pazar sepet analizi • Satış tahmininin değerlendirilmesi • Pazarlama kampanya ve stratejilerinin oluşturulması (İnan, 2003)
Bankacılık, Sigortacılık ve Borsa	• Kredi kartı dolandırıcılıklarının belirlenmesi • Hileli hesapların ortaya çıkarılması • Risk analizi • Risk yönetimi • Kredi taleplerinin değerlendirilmesi (Albayrak, 2008)
Perakendecilik	• Satış noktası verilerinin analizleri • Pazar sepeti analizi • Genel piyasa analizleri
Telekomünikasyon	• Web site ziyaretçilerinin profil analizi • Hatların yoğunluk kestirimleri
Bilim ve Mühendislik	• Uzay ve hücre analizleri
Sağlık	• Hastalanma süreci tahmini • Hasta tanısı tahmini yapabilmek • Uygun tahlillerin yapılması • Bölgesel salgınlar hakkında analizler
Kriminoloji	• Suç eğilimlerinin tanımlanması

	İstihbarat birimleri
Eğitim	- E-Okul gibi uygulamalarda - Sınava hazırlık aşamasındaki öğrencilerin hangi konularda eksikliklerini tespit etmede - Öğrenci başarılarını belirlemede - Başarı ve başarısızlık nedenlerinin tahmin edilmesinde
E- Ticaret	- Pazarlama stratejisi analizi - Ziyaretçi ile müşteri davranışlarının analizi sağlamak - Siteye giren kullanıcıların davranışlarının analizi
Spor	Oyuncu – takım eşleşmesi analizleri
Turizm	- Müşterinin mekân ve tarih tahmini analizleri - Tatil bölgelerinin fiyat tahmini analizleri - En çok seçilen tatil paketinin belirlenmesi
Kütüphanecilik	- Kişi ile kitap türü tahmini analizi - Kişilerin kitaplarının ödünç verme süresi analizi

3.6. Veri Madenciliği Modelleri

VM, fonksiyonel olarak Şekil 4’de de görüldüğü üzere tahmine edici ve tanımlayıcı olmak üzere ikiye ayrılıyor (Akpınar, 2000).

Şekil 4. Veri Madenciliği Modelleri (online, 2022 [veri madenciliği modelleri])



3.6.1. Tahmine Dayalı Veri Madenciliği Modelleri

Tahmine dayalı veri madenciliği modelleri, farklı veri kümelerinden toplanan bilinen sonuçları kullanarak yeni verilecek olan veri girdilerine karşılık çıktı değerlerini tahmin eder. Tahmine dayalı veri madenciliği modellerinin temel amacı, her zaman olmasa da genel olarak istatistiksel modellemeye dayalı olarak geçmiş verilere dayanarak geleceği tahmin etmektedir. Tahmine dayalı modelleme, sağlık sektörlerinde kalp yetmezliği, yüksek tansiyon, diyabet, enfeksiyon, kanser vb. olan yüksek riskli hastaları belirlemek için kullanılır. Kaza riskini poliçe sahibine atamak için araç sigorta şirketinde de kullanılmaktadır. Bir veri madenciliği görevinin tahmine dayalı modeli, sınıflandırma, regresyon, sapma tespiti analizi içermektedir (Alaybey, 2022).

3.6.1.1. Sınıflandırma Modeli

Veri madenciliğinde sınıflandırma, bir makine öğrenme modelinin yeni bir gözleme/kayıta belirli bir kategorik çıktı etiketi/durumu atadığı bir veri analizi biçimini ifade eder. Modelin veri setlerinden öğrendiklerine dayanır. Başka bir deyişle, sınıflandırma, nesneleri önceden tanımlanmış birçok kategoriye atama eylemidir (Şen, 2008).

Bankacılık ve finansal hizmetler endüstrisindeki bir sınıflandırma örneği, yapılan bankacılık işlemlerinin hileli olup olmadığının belirlenmesidir. Aynı şekilde makine öğrenimi de bir kredi başvurusunun onaylanıp onaylanmayacağını tahmin etmek için kullanılabilir.

3.6.1.2. Regresyon Modeli

Regresyon, bir işlev için verilerin değerini doğrulayan bir yöntemi ifade eder. Genellikle uygun veriler için kullanılır. Makine öğrenimi veya istatistik bağlamında doğrusal bir regresyon modeli, temelde sonuç olarak bilinen bağımlı değişken ile özellikler olarak bilinen bağımsız değişkenler arasındaki ilişkileri modellemek için doğrusal bir yaklaşımdır. (Çelik, 2019)

3.6.1.3. Sapma Tespiti Analizi

Tahmine dayalı veri madenciliği modellerinden sapma tespiti analizinin genel olarak karmaşık verilerde görülen kural dışı özelliklerin ve davranışların uygun algoritmalar sayesinde tespit edilmesidir. Karmaşık veriler, sistem hatalarından veya veritabanları analizinden kaynaklanan hatalardan meydana gelmektedir. Sapma tespiti analizi birçok sektörde kullanılmaktadır bunlar, tıp, eğitim, bankacılık gibi sektörel alanlarda veri tabanlarındaki verileri işleyip kullanılabilir hale getirmektir (Arslan, 2022).

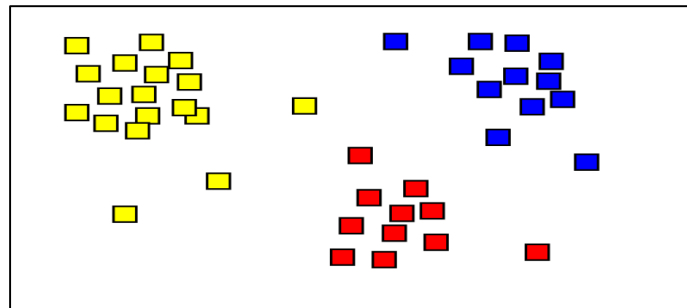
3.6.2. Tanımlayıcı Modeller

Tanımlayıcı modeller verilerin en genel özelliklerini göstermektedir. Tanımlayıcı modelde karar verme süreçlerinden rehberlik etmek kullanılmaktadır. Tanımlayıcı modeller çok boyutlu veriler arasında verilerin analiz edilerek işlenmesi gereken veriler arasında belirli kurallar oluşmasını amaçlamaktadır. Tanımlayıcı modelde kümeleme, birliktelik kuralları ve sıralı örüntü keşfi modelleri tercih edilmektedir (Arslan, 2022)

3.6.2.1. Kümeleme Modeli

Kümeleme, bir küme adı verilen aynı gruptaki nesnelerin diğer grup kümelerindeki nesnelere göre daha benzer olması için bir nesne kümesini gruplandırmaktır (Sarman, 2011). Kümeleme modeli ilk olarak 1854 yılında kolera salgınının yaşandığı Londra’da ölen hastaların yerlerinin belirlenmesinde kolera salgını için önlem alınmasında keşfedilmiştir (Coşlu, 2013). Şekil 5’te kümeleme modeli örneği verilmiştir.

Şekil 5. Kümeleme Modeli



3.6.2.2. Birliktelik Kuralları Modeli

Birliktelik Analizi, VM 'de en çok kullanılan yöntemlerden biridir. Büyük veri tabanlarındaki verilerin birbiriyle ilişkilerinin büyüklüğünü bulmak için kullanılan bir modeldir. Birliktelik Analizi, bir veri kümesindeki sık tekrar eden öğelerin saptanması ve öğeler arasındaki bağlantıları ve ilişkileri tespit etmek için kullanılmaktadır. Birliktelik kuralları davranışlarının analizi ile geleceğe yönelik çalışmalar yapılmasını destekleyen bir tanımlayıcı modeldir. (Zawaidah, 2011). Apriori, Eclat (Denklik Sınıfı Dönüşümü) ve FP-Growth (Sık Model) gibi algoritmaları mevcuttur.

4. BİRLİKTELİK KURALLARI

Birliktelik kurallarına dair bilgiler aşağıdaki alt başlıklar ile verilmektedir.

4.1. Birliktelik Kuralları Modeli

Birliktelik kuralları, VM 'nin en önemli tekniklerinden biridir. İşlemsel veritabanlarında ve diğer bilgi havuzlarında öge veya nesne kümeleri arasında sık örüntüler, ilişkiler, korelasyonları bulmaya çalışır (Tokyürek, 2019). Günlük faaliyetlerden elde edilen verilerle birlikte veri hacmi çarpıcı bir şekilde artmaktadır. Bu nedenle, veri tabanındaki büyük miktarda veriden birliktelik kuralları madenciliği, herhangi bir ticari karar verme sürecinde yardımcı olabilecek birçok endüstri için ilgi çekicidir. Teknik verilerden birliktelik kurallarını keşfetmek için geleneksel olarak insan davranışının bazı yönlerini gösteren öğeler arasındaki ilişkileri tanımlamaya odaklanmıştır, ör. Müşterilerin birlikte satın alacakları ürünleri belirlemek için satın alma davranışı (WEBB, 2003).

Birliktelik kuralı öğrenme, makine öğreniminin çok önemli kavramlarından biridir ve Pazar Sepeti analizinde, Web kullanımı madenciliğinde, sürekli üretimde vb. kullanılır. Burada pazar sepeti analizi, çeşitli büyük perakendeciler tarafından ilişkileri keşfetmek için kullanılan bir tekniktir (Coşlu, 2013).

4.2. Birliktelik Kuralları Modelinin Kullanıldığı Alanlar:

Birliktelik kurallarında kullanılan alanlar aşağıdaki alt başlıklar ile verilmiştir.

4.2.1. Pazar Sepeti Analizi

Pazar sepet analizi, birliktelik madenciliğinin en tipik örneğidir. Verilerin çoğu süpermarkette barkod tarayıcılar kullanılarak toplanır. “Market sepeti” veri tabanı olarak bilinen bu veri tabanı, geçmiş işlemlere ilişkin çok sayıda kayıttan oluşmaktadır. Tek bir kayıt, bir müşteri tarafından bir satışta satın alınan tüm ürünleri listeler. Hangi grupların hangi ürün grubuna meyilli olduğunu bilmek, bu mağazalara mağaza düzenini ve mağaza kataloğunu birbiriyle en uygun şekilde yerleştirmek için ayarlama özgürlüğü vermektedir (Nair, 2019).

4.2.2. Tıbbi Teşhis

Tıbbi teşhiste birliktelik kuralları, hekimlerin hastaları iyileştirmesine yardımcı olmak için yararlı olabilir. Teşhis kolay bir süreç değildir ve güvenilir olmayan sonuçlara yol açabilecek bir dizi hata içerir. İlişkisel birliktelik kuralı madenciliğini kullanarak, çeşitli faktörler ve semptomlarla ilgili olarak hastalığın ortaya çıkma olasılığını belirlemede kullanılmaktadır (Serban, Campan ve Czibula, 2006)

4.2.3. Nüfus Sayımı Verileri

Her hükümetin tonlarca nüfus sayımı verisi vardır. Bu veriler, etkin kamu hizmetleri (eğitim, sağlık, ulaşım) planlamak ve kamu işletmelerine (yeni fabrikalar, alışveriş merkezleri kurmak ve hatta belirli ürünleri pazarlamak için) yardımcı olmak için kullanılabilir. Birliktelik kuralı madenciliği ve veri madenciliğinin bu uygulaması, sağlam kamu politikasını destekleme ve demokratik bir toplumun verimli işleyişini sağlama konusunda muazzam bir potansiyele sahiptir (Malerba, vd. 2001).

4.2.4. Protein Dizisi

Proteinler, yirmi çeşit amino asitten oluşan dizilerdir. Her protein, bu amino asitlerin dizisine bağlı olan benzersiz bir 3B yapıya sahiptir. Dizideki küçük bir değişiklik, proteinin işleyişini değiştirebilecek yapıda bir değişikliğe neden olabilir. Amino asit dizisi üzerinde çalışan proteinin bu bağımlılığı, büyük bir araştırma konusu olmuştur. Daha önce bu dizilerin rastgele olduğu düşünülüyordu, ancak şimdi olmadıkları görülmektedir. Bir proteinde bulunan farklı amino asitler arasındaki ilişkilerin doğasını deşifre ettiler. Bu birliktelik kurallarının bilgisi ve anlaşılması, yapay proteinlerin sentezi sırasında son derece yardımcı olmaktadır (Gupta, Mangal, Tiwari ve Mitra, 2006)

4.3. Birliktelik Kuralları Terimleri

Pazar sepet analizlerinde yaygın olarak kullanılan terimler ve kelimeler kullanılmaktadır. Bu kelimelerin ve terimlerin tanımları aşağıda verilmiştir.

4.3.1. Destek Değeri

Destek (support) değeri, A ve B öğelerinin birlikte olduğu kayıt sayısının, veri setinde saklanan tüm kayıtların oranıdır. (Özçalıcı, 2017)

$$S = \frac{\sigma(A \cup B)}{\text{Toplam Kayıt Sayısı}}$$

Tablo 2’de belirli bir $A \rightarrow B$ kuralının destek değerinin hesaplanabileceğini göstermek için basit bir veri seti kullanılmıştır.

Tablo 2. Destek Değeri Örneği

Sıra	Sepet	Destek / Toplam Kayıt Sayısı
1	Ekmek, Süt, Meyve Suyu	Toplam Destek = 5 Destek {Ekmek, Süt} = $2/5 = \%40$ Destek {Süt, Meyve Suyu} = $3/5 = \%60$ Destek {Ekmek, Süt, Meyve Suyu} = $1/5 = \%20$
2	Ekmek, Süt, Bez	
3	Süt, Meyve Suyu	
4	Ekmek, Meyve Suyu	
5	Süt, Meyve Suyu, Bez	

4.3.2. Güven Değeri

Güven (confidence) değeri, A ve B öğelerinin birlikte veri setinde yer aldığı kayıt sayısının, veri setinde yer alan A öğesinin tekrar ettiği kayıt sayısına oranıdır. (Özçalıcı, 2017)

Destek ve güven değeri, 0 ile 1 değerleri arasında yer almaktadır.

$$S = \frac{\sigma(A \cup B)}{\sigma(A)}$$

Tablo 3’te belirli bir $A \rightarrow B$ kuralının güven değerinin hesaplanabileceğini göstermek için basit bir veri seti kullanılmıştır.

Tablo 3. Güven Değeri Örneği

Sıra	Sepet	Güven
1	Ekmek, Süt, Meyve Suyu	Toplam Destek = 5
2	Ekmek, Süt, Bez	Güven {Ekmek, Süt} = $2/3 = \%66$
3	Süt, Meyve Suyu	Güven {Süt, Meyve Suyu} = $3/4 = \%75$
4	Ekmek, Meyve Suyu	
5	Süt, Meyve Suyu, Bez	Güven {Ekmek, Süt, Meyve Suyu} = $1/2 = \%50$

4.3.3. Kaldıraç (Lift) Değeri

A ve B'nin birbirlerinde bağımsız olması halinde, veri setinde ne kadar geçtiğini hesaplamaktadır.

Kaldıraç değeri 1'den büyük ve küçük olma durumlarına göre açıklanmaktadır. Kaldıraç değeri 1'den büyük ise kurallar arasında anlamlı kural olduğunu belirtmektedir, 1'den küçük ise kurallar arasında anlamsız kural olduğunu belirtmektedir (Karaibrahimoğlu, 2014).

$$Kaldıraç A \rightarrow B = \frac{\text{destek}(A \rightarrow B)}{\text{destek}(A) * \text{destek}(B)}$$

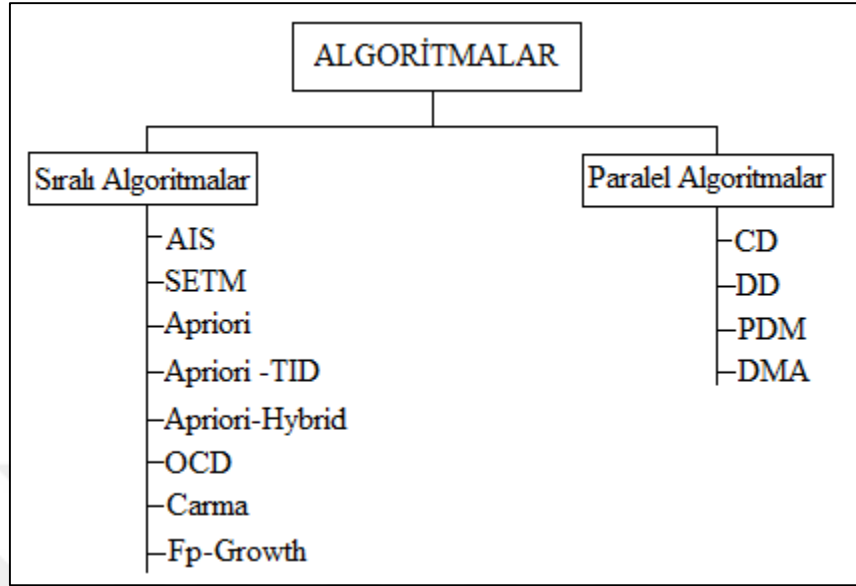
Tablo 2 ve 3'teki örnekler üzerinde çıkan kurallar şu şekilde açıklanmıştır:

Süt $\Rightarrow$ Meyve Suyu (Destek= $\%60$, Güven= $\%75$, Kaldıraç=0,93)

- Destek: Müşterilerin $\%60$ süt ve meyve suyunu birlikte almıştır.
- Güven: Süt alan müşterilerin $\%75$ 'i meyve suyu 'da almıştır.
- Kaldıraç: Kaldıraç değeri 0,93 ve bu değer 1'den küçük olduğu için süt ve meyve suyu arasında ilişkiye anlamsız kural denilebilir.

4.4. Birliktelik Kuralları Algoritmaları

Şekil 6. Birliktelik Kuralları Algoritmaları (Söylemez, Doğan ve Özcan, 2016)



4.4.1. Sıralı Algoritmalar

Aşağıda sıralı algoritma çeşitlerinden olan AIS, SETM, Apriori, Apriori-TID, Apriori-Hybrid, OCD, Carma ve FP-Growth algoritmaları ile ilgili bilgiler verilmektedir.

4.4.1.1. AIS Algoritması

AIS algoritması, bir veri setindeki tüm büyük öge kümelerini oluşturmak için geliştirilen ilk birliktelik kuralı algoritmasıdır (Zhao ve Bhowmick, 2003). AIS'deki sık öge kümelerini elde etmek için veritabanları birçok kez taranmaktadır. AIS algoritmasında, veritabanları birkaç kez taranarak sık öge kümeleri oluşturmaktadır. Her bir ögenin destek sayısı, veritabanı üzerinden ilk geçiş sırasında toplanıp, minimum destek sayısına bağlı olarak, desteği minimum değerinden daha az olan ögeler listeden elenmektedir. AIS algoritması, tüm veritabanı üzerinde birden çok geçiş yapar. Her geçiş sırasında tüm işlemleri taramaktadır. İlk geçişte, bireysel ögelerin desteğini sayar ve veri tabanında hangilerinin büyük veya sık olduğunu belirler. Her geçişin büyük öge kümeleri, aday öge kümelerini oluşturmak için genişletilmektedir. Daha sonra, her aday kümenin desteği hesaplanır. Minimum desteğe eşit veya daha büyük desteklere sahip aday kümeler, büyük öge kümeleri olarak seçilir. Bu büyük öge kümeleri, bir sonraki geçiş için aday kümeler oluşturmak

üzere genişletilmektedir. Bu süreç, daha fazla büyük öge kümesi bulunmadığında sona ermektedir. AIS algoritmanın dezavantajı, sonunda küçük olduğu ortaya çıkan çok sayıda aday öge kümesinin üretilmesidir (Gürgen, 2008). Şekil 7’de AIS algoritmasının sözde kodu verilmiştir.

Şekil 7. AIS Algoritmasının Sözde Kodu (Zerman, 2018)

```

L1 = {nesneküme}
for (k=2; Lk-1 ≠ 0 ; k++)
{
    Ck = Boş küme;
    for (t=0; t<öge sayısı; t++)
    {
        Lt = Altküme (Lk-1)
        Ct = İlişkili ürünler;
        for (c=0; c<Ct; c++)
        {
            if (c ∈ Ck)
                c++;
            else
                Ck ya Ct'yi ekle;
                C++;
        }
        if (ürünün destek değeri > girilen destek değeri)
            Lk = Ck;
    }
}
Return Uk Lk;

```

4.4.1.2. SETM Algoritması

SETM algoritması, Houtsma tarafından 1995’te büyük öge kümelerini hesaplamak için önerilmiştir. SETM algoritması büyük öge kümelerinin her bir üyesi <TID, öge kümesi> biçimindedir. Burada TID, bir işlemin benzersiz tanımlayıcısıdır. SETM algoritması AIS algoritmasına benzer şekilde veritabanı üzerinden çoklu geçişler yapar. İlk geçişte, bireysel öğelerin desteğini sayar ve veri tabanında hangilerinin büyük veya sık olduğunu belirlemektedir. Ardından, önceki geçişin büyük öge kümelerini genişleterek aday öge kümelerini oluşturmaktadır (Girotra, vd. 2013). Aday kümeler üreten SETM algoritması, aday öge kümelerinin bir kopyasını, oluşturan işlemin TID’si ile birlikte sıralı bir şekilde kaydeder. Daha sonra, aday öge kümeleri öge kümelerine göre sıralanır ve bir toplama işlevi kullanılarak küçük öge kümeleri silinmektedir. Veritabanı TID algoritması temelinde sıralanmışsa, bir sonraki geçişte bir işlemde bulunan büyük öge kümeleri TID’e göre sıralanarak elde edilir. Bu şekilde, veri tabanında birkaç geçiş yapılmaktadır. Daha fazla büyük öge kümesi

bulunmadığında, algoritma sona erer (Girotra, vd. 2013). Şekil 8’de SETM algoritmasının sözde kodu verilmiştir.

Şekil 8. SETM Algoritmasının Sözde Kodu (Zerman, 2018)

```

1)  $P_1 = \{\text{Büyük } p - \text{öge setleri}\}$ 
2)  $P_1 = \{\text{Büyük } p - \text{öge setleri TID'e göre uygun şekilde sıralaması sağlanmaktadır.}\}$ 
3) For  $\{t=2; P_{k-1} \neq \emptyset; t++\}$  [Başla]
4)  $D_t = \emptyset$ 
5) Bütün hepsi için işlemler  $f \in B$  [Başla]
6)  $P_t = p \in P_{p-1}, p, TID = f.TID // f'de yer alan p-öge setleri$ 
7) Hepsi için büyük veri öğeleri  $p_f \in P_f$  [Başla]
8)  $D_f = p - \text{uzantıları } p_f \text{ içerdiği } f, f \text{ deki adaylar}$ 
9)  $D_t += \{ \langle f.TID, e \rangle ; d.count < minsup \}$ 
10) Son
11) Son
12) çeşit  $D_t$  üzerindeki veri setleri
13)  $d \in D_t$ ’deki  $d.count < minsup$  olsun veri setleri sil ve gönder  $P_t$ ;
14)  $P_t = \{ \langle p - \text{veri setleri, miktarın } p \text{ içinde } pt > p \in P_t \} // \text{Birleştirme Adımı}$ 
15) Çeşit  $P_t$  üzerinde TID
16) Son
17) Cevap  $U_k P_t$ 

```

4.4.1.3. Apriori Algoritması

Agrawal ve Srikant (1994) ilk olarak Apriori algoritmasını önermişlerdir. Bu algoritma, “sık kullanılan k Öğelerinin her alt $(k-1)$ öğelerinin sık olması gerektiğini” belirten Apriori özelliğine dayanmaktadır. Apriori algoritmasında iki ana işlem yürütülür: biri, karşılık gelen öğelerinin destek sayısının işlem veritabanını tarayarak hesaplandığı aday oluşturma işlemidir ve ikincisi, destek sayısı minimum eşikten küçük olan aday öge kümelerinin budamasıyla oluşturulan büyük öge kümesi oluşturma işlemidir. Bu işlemler, aday öge kümeleri veya büyük öge kümeleri boşalana kadar yinelemeli olarak yinelemektedir. Orijinal veritabanı, aday kümesi için ilk kez taranır, bir sensör ögesinden oluşur ve orada destek sayılır, daha sonra bu 1 Öge Kümesi adayları, kullanıcı tarafından belirtilen eşikten daha az öge sayısına sahip olan öğeleri kaldırarak budama işlemi yapılır. Apriori özelliğine göre, 2 sık öge kümesinden oluşan her alt 1 öge kümesi sık olmalıdır (Agrawal ve Srikant, 1994). Şekil 9’da Apriori algoritmasının sözde kodu verilmiştir.

Şekil 9. Apriori Algoritmasının Sözde Kodu (Agrawal, Srikant, 1993)

```

1)  $T_1 = t - \text{sık bulunan veri öğeleri } (A)$ ;
2) for ( $s = 2$  ;  $T_{s-1} \neq \emptyset$  ;  $s++$ ) {
3)  $K_s = \text{Apriori-gen}(T_{s-1})$ ;
4) Her işlem için  $z \in A$  { //Sayım için A taraması
5)  $K_s = \text{Altküme}(K_s, z)$  ; Adayların  $z$  altkümeleri elde edilmektedir.
6) Adayların her biri için  $k \in k_1$ 
7)  $k.\text{count}++$  ;
8) }
9)  $T_s = \{k \in K_s, k.\text{count} > \text{min.}\}$ 
10) }
11) Return  $T = \bigcup_{s} T_s$ 

```

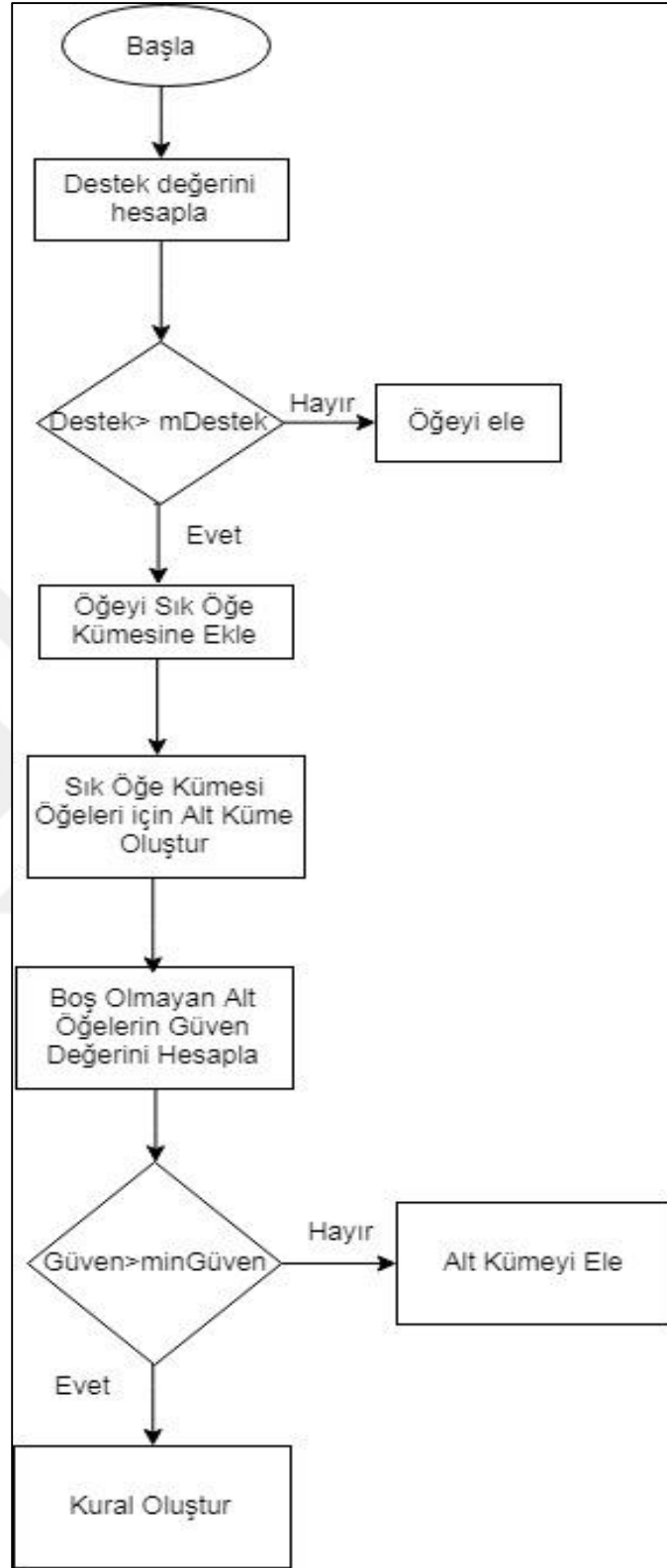
Prosedür apriori-gen (T_{s-1} ; sık ($s-1$) veri öğeleri

```

1) Her biri için  $t_1 \in T_{s-1}$ 
2) Her biri için  $t_2 \in T_{s-1}$ 
3) if ( $t_1[1] = t_2[1]$ )  $\wedge$  ( $t_1[2] = t_2[2]$ )
    $\wedge \dots \wedge (t_1[s-2] = t_2[s-2]) \wedge (t_1[s-1] < t_2[s-1])$  sonra {
4)  $k = t_1 \bowtie t_2$  ; //Birleştirme adımı : Sık olmayan adayları sil.
5) if seyrek altkümesi varsa ( $k, T_{s-1}$ ) sonra
6) Sil  $k$  ; //Budama Adımı: Sık olmayan adayları sil.
7) Else ekle  $k$  to  $K_s$  ;
8) }
9) Return  $K_s$  ;
   Yordamın sık olmayan altkümesi var. ( $k$ : Aday,  $s$ -veri öğeleri )
    $T_{s-1}$ : sık ( $s-1$ ) veri öğeleri ; Ön bilgileri kullan listede var mı ?
1) Her biri için ( $s-1$ ) altküme  $z$  of  $k$ 
2) if  $z \notin T_{s-1}$  sonra
3) Return TRUE;
4) Return FALSE;

```

Şekil 10. Apriori Algoritmasının Akış Diyagramı



Apriori algoritmasının Tablo 4’de örnek veri seti ile anlatılmıştır.

Tablo 4. Apriori algoritması 1. Aşama

Sıra	Öğeler
1	Ekmek, Süt, Meyve Suyu
2	Bez, Süt, Zeytin
3	Ekmek, Limon, Süt, Zeytin
4	Limon, Zeytin
5	Süt, Meyve Suyu
6	Bez, Limon
7	Bez, Zeytin, Süt

Yukarıda ifade edilen {Ekmek, Süt, Meyve Suyu, Bez, Zeytin, Limon} bir markette satılan ürünler belirtilmiştir. Apriori algoritmasının ilk adımı her ürünün aynı anda kaç işlemde satıldığını bulmaktır.

Tablo 5. Apriori Algoritması 2. Aşama

Ürün Kodu	Frekans
Ekmek	2
Süt	5
Meyve Suyu	2
Bez	3
Zeytin	4
Limon	3

Tablo 5'te Apriori algoritmasına ait destek değerleri belirlenmiştir. Bir sonraki aşamada ise 2'li olarak incelenecektir. Minimum destek değeri 2 olarak tanımlansın.

Tablo 6. Apriori Algoritması 3. Aşama

Ürün Kodu	Frekans
Ekmek, Süt	2
Bez, Süt	2
Meyve Suyu, Süt	2
Limon, Zeytin	2
Bez, Zeytin	2
Çikolata, Süt	2

Tablo 6'e göre örneğin {Ekmek, Süt} uygulama veri setinde 2 yerde geçmektedir.

Bunlar: TID 1 = {Ekmek, Süt, Meyve Suyu}, TID 3 = {Ekmek, Limon, Süt, Zeytin} işlemleridir. Apriori algoritmasının sonrasında 3 ve 4 elemanlı olarak inceleme yapılabilir fakat minimum değeri 2 olarak belirlendiği için devam etmeye gerek yoktur.

4.4.1.4. Apriori – TID Algoritması

Apriori algoritması destek değerini saymak için her geçişte tüm veritabanını tarar. Tüm geçişlerde tüm veri tabanını taranması gerekmeyebilir. Bu hipotezden esinlenerek Agrawal ve Srikant, Apriori-TID adlı başka bir algoritma önermişlerdir. Apriori algoritmasına benzer şekilde, Apriori-TID, bir geçişin başlangıcından önce aday öge kümelerini belirlemek için Apriori 'nin aday oluşturma işlevini kullanır. Apriori 'den temel farkı, ilk geçişten sonra desteği saymak için veritabanını kullanmamasıdır (Zerman, 2018). Şekil 11'de Apriori-TID algoritmasının sözde kodu verilmiştir.

Şekil 11. Apriori-TID Algoritması Sözde Kodu (Zerman, 2018)

```

1.  $A_1 = \{\text{Büyük A - veri ögeleri}\}$ 
2.  $F_1 = \text{Veritabanı (D)}$ ;
3. for ( $z=2$ ;  $A_{z-1} \neq \emptyset$ ;  $z++$ ) (Başla)
4.  $F_z = \text{apriori - gen } (A_{z-1})$ ; Yeni adaylar
5.  $F_z = \emptyset$ ;
6. Tüm Girişler  $t \in F_{z-1}$  (Başla)
7.  $\forall$  Tanımlayıcı  $t$ .TID ile işlemin içerdiği  $F_k$  içindeki aday ögeleri belirler.

    $F_t = \{f \in F_z, (f - f[z]) \in t.\text{öge kümesi} \wedge (f - f[z-1]) \in t.\text{öge kümeleri};$ 

8. Hepsini için adaylar  $f \in F_t$  yapalım;
9. C.Miktar ++;
10. İF ( $F_1 \neq \emptyset$ ) sonra  $F_z += < t.TID, F_t >$ ;
11. Son
12.  $A_z = \{f \in F_z, f.\text{count} \geq \text{minsup}\}$ 
13. Son
14. Cevap =  $U_z A_z$ ;

```

4.4.1.5. Apriori-Hybrid Algoritması

Apriori-Hybrid algoritması, tüm veri geçişlerinde aynı algoritmayı kullanmanın gerekli olmadığı fikrine dayanmaktadır. Agrawal 1994'te bahsettiği gibi, Apriori önceki geçişlerde daha iyi performansa sahiptir ve Apriori-TID sonraki geçişlerde Apriori' den daha iyi performans göstermektedir. Deneysel gözlemlere dayalı olarak, ilk geçişlerde Apriori'yi kullanan ve geçiş sonundaki setin belleğe sığması beklendiğinde Apriori-TID'ye geçen Apriori-Hybrid tekniği geliştirilmiştir. Bu nedenle, her geçişin sonunda bir tahmin gereklidir. Apriori-Hybrid algoritmasının performansı, büyük veri setleri için deneyler ve analizler yapılarak da değerlendirilmiştir. Geçişlerin en sonunda meydana geldiği durum dışında, Apriori-Hybrid algoritması Apriori algoritmasından daha iyi performans gösterdiği gözlemlenmiştir (Srikant, 1996).

4.4.1.6. OCD Algoritması

Çevrimdışı Aday Belirleme (OCD) algoritması, Mannila ve arkadaşları, 1994'te küçük örneklerin genellikle büyük öge kümelerini bulmak için oldukça iyi olduğu fikrine dayalı olarak önerilmiştir. OCD algoritması, gereksiz aday kümeleri ortadan kaldırmak için önceki geçişlerden elde edilen bilgilerin analizinin sonuçlarını

kullanır. OCD algoritması, aday kümeleri belirlemek için AIS algoritmasından farklı bir yaklaşım izler. OCD algoritması, geçişi olabildiğince basit tutarak geçişler arasındaki aday kümeleri budamak için önceki geçişlerden elde edilen tüm bilgileri kullanır. OCD algoritmasının avantajı, daha düşük bir destek eşiği ile artar (Mannila ve arkadaşları, 1994) AIS algoritmasında oluşturulan adaylar, OCD algoritmasının adaylardan önemli ölçüde daha yüksektir. AIS algoritması, geçiş sırasında yinelenen adaylar oluşturabilirken, OCD algoritması herhangi bir adayı bir kez oluşturur ve veri tabanına karşı değerlendirmeden önce alt kümelerinin büyük olup olmadığını kontrol etmektedir.

4.4.1.7. Carma Algoritması

CARMA (Sürekli İlişkilendirme Kuralı Madenciliği Algoritması) (Hidber, 1999), büyük öge kümelerinin hesaplanmasını çevrimiçi hale getirmektedir. Çevrimiçi olan CARMA, mevcut ilişkilendirme kurallarını kullanıcıya gösterir ve kullanıcının veri tabanının ilk taraması sırasında herhangi bir işlemde parametreleri, minimum desteği ve minimum güveni değiştirmesine izin vermektedir. Carma algoritması en fazla 2 veritabanı taramasına ihtiyaç duymaktadır. CARMA ilk taramada öge kümelerini oluşturur ve ikinci taramada tüm öge kümelerini saymayı bitirir. CARMA, işlemlerden anında ürün setleri oluşturur. Her işlemi okuduktan sonra, önce işlemin alt kümeleri olan öge kümelerinin sayısını artırır. Ardından, öge kümelerinin tüm acil alt kümeleri, mevcut minimum destek ve okunan veritabanı parçasına göre potansiyel olarak büyükse, işlemde yeni öge kümeleri oluşturmaktadır. Bir öge kümesinin potansiyel olarak büyük olup olmadığının daha doğru bir şekilde tahmin edilmesi için, öge kümesinin sayısı için bir üst sınır hesaplar; bu, mevcut sayısının toplamı ve öge kümesi oluşturulmadan önceki oluşum sayısının bir tahminidir (Hidber, 1999).

4.4.1.8. FP-Growth Algoritması

FP-Growth, bir veri setinde en sık görülen öge setini belirlemek için kullanılabilecek alternatif bir algoritmadır. FP-Growth, Apriori algoritmasında kullanılan paradigmadan farklı bir yaklaşım kullanır. FP-Growth, bir veri kümesindeki en sık veri kümesi kümesini belirlemek için kullanılabilecek alternatif bir algoritmadır. FP-Growth algoritması, Apriori algoritmasının geliştirilmiş halidir (Tokyürek, 2019). FP-Growth, genellikle bir aday Nesil olmadan bir madencilik öge kümesi olan bir

yöntemdir. Orijinal işlem veritabanını sıkıştırmak için çok yoğun bir veri yapısı (FP-tree) oluşturur. Şekil 12’de FP-Growth algoritmasının sözde kodu verilmiştir.

Şekil 12. FP-Growth Algoritmasının Sözde Kodu (Han, Pei ve Yin, 2000)

```

Girdi: FPTree
Çıktı: Tüm sık öge kümeleri
Metod: FP-Growth (Fp-tree, null)
FP-Growth (Tree, a)
1. {
2.   if Tree tek bir yol P içeriyorsa then
3.     for each  $\beta$  (P yolu içerisindeki düğümlerin kombinasyonu) do
4.        $Sup \leftarrow \min \{b. support \mid b \text{ in } \beta\}$ 
5.       Sık öge kümesi  $aU\{b.item \mid b \text{ in } \beta\}$  y1 sup destek değeri ile üret
6.   else
7.     For each item a (a in Tree.header)
8.        $Sup \leftarrow a. support$ 
9.       aUa için sık öge kümelerini sup destek değeri ile üret
10.       $\beta = aUa$  için şartlı örüntü temellerini ve sonra
11.       $\beta$  için şartlı örüntü ağacını  $Tree_\beta$  oluştur
12.      if  $Tree_\beta \neq \emptyset$  then
13.        FP-Growth ( $Tree_\beta$ )
14. }
```

4.4.2. Paralel Algoritmalar

Aşağıda paralel algoritmalarından CD, PDM, DD, IDD ve Eclat algoritmaları ile ilgili bilgiler verilmektedir.

4.4.2.1. CD (Collision Detection – Çarpışma Algılama) Algoritması

Çarpışma algılama (CD), paralel olarak fazlalık hesaplamalar yoluyla sıralı bir Apriori tabanlı aday oluşturma yaklaşımıdır. Bir nesnenin diğerine ne zaman girdiğinin belirlenmesini içerir. Özellikle çok sayıda nesne söz konusu olduğunda ve nesneler karmaşık şekillere sahip olduğunda, kullanılan bir paralel algoritma çeşididir (Agrawal, Srikant, 1993). Şekil 13’te CD algoritmasının sözde kodu verilmiştir.

Şekil 13. CD Algoritması Sözde Kodu (Dunham ve Diğerleri, 2000)

```

1. Input
2.  $I, s, D^1, D^2, \dots, D^p$ 
3. Output
4.  $L$ 
5. Algorithm:
6.  $C_1 = I$ 
7. for  $k=1; C_k \neq \emptyset; k++$  do begin
    // step one : counting to get the local counts
8.    count( $C_k, D^1$ ; // local processor is  $\bar{I}$ 
    // step two: exchanging the local counts with other processors.
9.    // to obtain the global counts in the whole database.
10.   Forall itemset  $X \in C_k$  do begin
11.        $X.count = \sum_{j=1}^p \{X_j.count\}$ ;
12.   End
13.   //step thee: identifying the large itemsets and
14.   //generating the candidates of size  $k+1$ 
15.    $L_k = \{c \in C_k \mid c.count \geq x \mid D^1 \cup D^2 \cup \dots \cup D^p\}$ ;
16.    $C_{k+1} = \text{apriori\_gen}(L_k)$ ;
17. End
18. Return  $L = L_1 \cup L_2 \cup \dots \cup L_k$ ;

```

4.4.2.2. PDM (Paralel Data Mining – Paralel Veri Madenciliği) Algoritması

Temel olarak, paralel veri madenciliği (PDM), aday öge kümelerinin paralel oluşturulmasından ve büyük öge kümelerinin paralel belirlenmesinden oluşmaktadır. Algoritma, erken geçişlerde aday öge kümeleri oluşturmak için bir karma tablo kullanır ve daha sonra, sonraki geçişlerde doğrudan önceki büyük öge kümelerinden aday öge kümeleri oluşturmak için bir yöntem kullanmaktadır (Gürgen, 2008).

4.4.2.3. DD (Data Distribution – Veri Dağıtımı) Algoritması

DD, aday öge kümelerini işlemciler arasında sıralı bir biçimde bölümlere ayırır. Her işlemci, veri tabanındaki tüm işlemler için aday öge kümelerinin yerel olarak depolanan alt kümesinin sayısını hesaplamaktan sorumludur. Her işlemciye aday öge kümelerinin belirli bir alt kümesi atandığından, yerel olarak atanan işlemlere ek olarak diğer işlemcilerin belleğinde depolanan işlemlerin geri kalanının taranması artık gereklidir. Aday öge kümelerinin sayısını hesapladıktan sonra, her işlemci yerel aday öge kümelerinden sık görülen öge kümelerini bulur ve bunları diğer tüm işlemcilere göndermektedir (Veli, 2020). Şekil 14'te DD algoritmasının sözde kodu verilmiştir.

Şekil 14. DD Algoritmasının Sözde Kodu (Agrawal 1996)

```

1. Input
2.  $I, s, D^1, D^2, \dots, D^p$ 
3. Output
4.  $L$ 
5. Algorithm:
6.    $C_1 \leftarrow C I;$ 
7.   for  $k=1; C_k \neq \emptyset; k++$  do begin
8.     // step one : counting to get the local counts
9.     count( $C_k^i, D^i$ ; // local processor is  $i$ 
10.    // step two: broadcast the local database partition to others,
11.    // receive the remote database partitions from others,
12.    // scan  $D^j (1 \leq j \leq p, j \neq i)$  to get the global counts.
13.    broadcast( $D^i$ );
14.    for ( $j=1; (j \leq p \text{ and } j \neq i); j++$ ) do begin
15.      Receive ( $D^j$ ) from processor  $j$ ;
16.      Count( $C_k^i, D^j$ );
17.    End
18.    //step three: identify the large itemsets in  $C_k^i$ .
19.    //Exchange with other processors to get all large itemsets  $C_k$ .
20.    //generate the candidates of size  $k+1$ 
21.    //partition the candidates and distribute over all processors.
22.     $L_k^i = \{c | c \in C_k^i \mid c.\text{count} \geq x \cdot s * |D^1 \cup D^2 \cup \dots \cup D^p|\};$ 
23.     $L_k = \bigcup_{i=1}^p L_k^i;$ 
24.     $C_{k+1} = \text{apriori\_gen}(L_k);$ 
25.     $C_{k+1} \leftarrow C C_{k+1};$  // partition the candidate itemsets across the processors
26.  End
27.  Return  $L = L_1 \cup L_2 \cup \dots \cup L_k;$ 

```

4.4.2.4. IDD Algoritması

IDD (Akıllı Veri Dağıtımı), DD ‘ye göre önemli bir gelişmedir. Adayları, adayların ilk maddesine göre işlemciler arasında bölümlere ayırır, yani aynı ilk öğeye sahip adaylar aynı bölüme bölünmektedir. Bu nedenle, her işlemcinin yalnızca işlemciye atanan öğelerden biriyle başlayan alt kümeleri kontrol etmesi gerekmektedir. Bu işlem, DD ‘deki fazla hesaplamayı azaltır, çünkü DD için her işlemcinin her işlemin tüm alt kümelerini kontrol etmesi gerekir, bu da çok fazla hesaplama sağlamaktadır (Han, Karypis, Kumar, 1997).

4.4.2.5. Eclat Algoritması:

Eclat algoritması, Apriori algoritmasının daha verimli ve ölçeklenebilir bir versiyonudur. Apriori algoritması, bir grafiğin Genişlik-İlk Aramasını taklit ederek yatay bir anlamda çalışırken; Eclat algoritması, bir grafiğin Derinlik-İlk Araması gibi dikey bir şekilde çalışır. Eclat algoritmasının bu dikey yaklaşımı, onu Apriori

algoritmasından daha hızlı bir algoritma haline getirir ((Han, Karypis, Kumar, 2000)).
Şekil 15'te Eclat Algoritmasının sözde kodu verilmiştir.

Şekil 15: Eclat Algoritmasının Sözde Kodu (Eclat, 4 Mart 2022)

```

Input: D, a,  $I \subseteq T$ 
Output: F[I] (D,a)

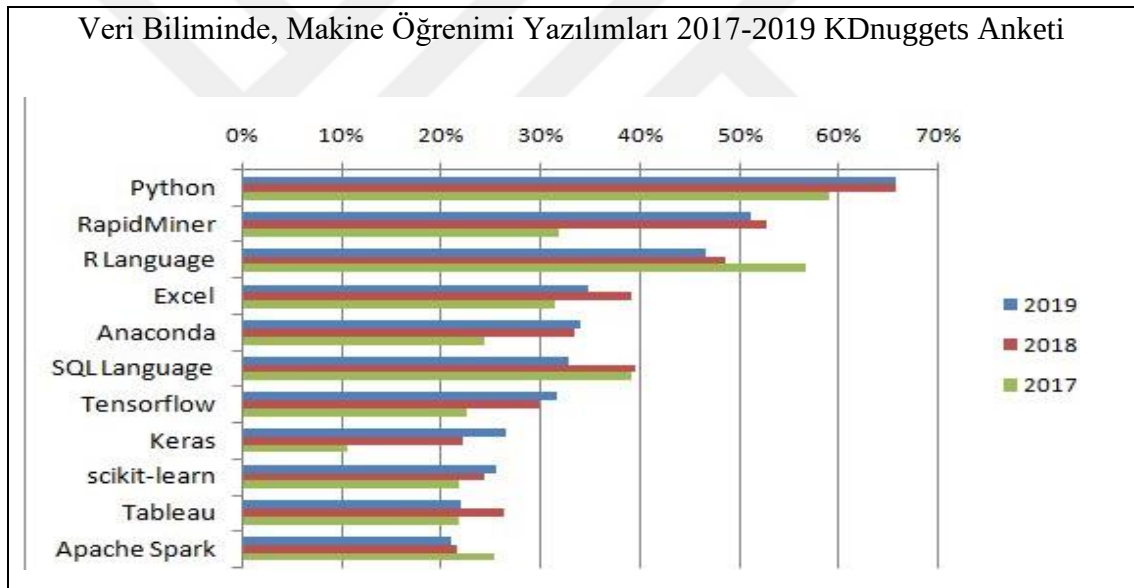
1. F[I]: = {}
2. for all  $i \in T$  occurring in D do
3.   F[I]: = F[I]  $\cup$  {I  $\cup$  {i}}
4.   // Create  $D^i$ 
5.    $D^i$ : = {}
6.   For all  $j \in T$  occurring in D such that  $j > i$  do
7.     C: = cover({i})  $\cap$  cover({j})
8.     if  $|C| \geq a$  then
9.        $D^i$ : =  $D^i \cup \{j.C\}$ 
10.    end if
11.  end for
12.  // Depth-first recursion
13.  Compute F [I  $\cup$  {i}] ( $D^i$ .a)
14.  F[I]: = F[I]  $\cup$  F [I  $\cup$  {i}]
15. end for

```

5. VERİ MADENCİLİĞİ PLATFORMLARI

Veri madenciliği yazılımları ile ilgili 2017-2019 arasında yapılmış olan KDnuggets anketinde en çok kullanılan yazılımlar aşağıdaki Şekil 16’da gösterilmiştir. Ankete göre 2019 yılında en çok kullanılan yazılımlar sırasıyla Python, RapidMiner ve R dili olarak seçilmiştir. Biz de anketten esinlenerek VM çalışmalarında son yıllarda popüler hale gelmiş olan bu platformlardan Python ve R ile; kullanım oranı daha düşük olsa da hızlı çalıştığı yönünde bildirimlerde bulunan bir platform olan SPMF yazılımlarını tercih edip birbirleriyle performans incelemesi yapılmaya karar verilmiştir.

Şekil 16. Veri Biliminde, Makine Öğrenimi Yazılımları 2017-2019 KDnuggets Anketi (Kdnuggets, 4 Mart 2022)



5.1. Python

Python, Object Oriented (nesne yönelimli) bir programlama dilidir ve oyunlardan siber güvenliğe, web sitelerinden uygulamalara kadar her alanda kullanılabilir. Özellikle veri madenciliği, makine öğrenmesi ve derin öğrenme gibi uygulamalarda en çok kullanılan programlama dillerindendir.

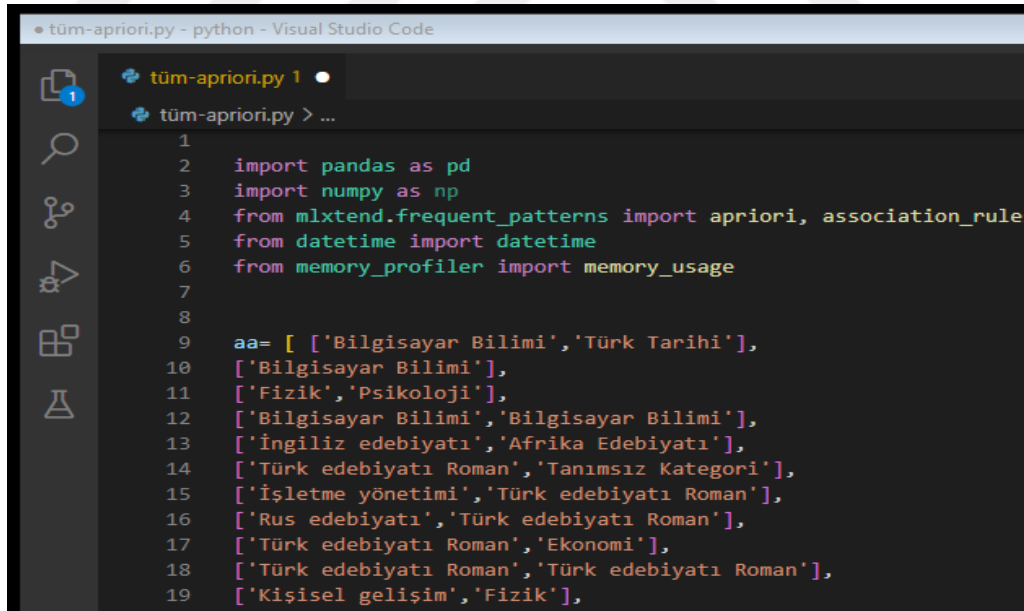
Büyük miktarda veri ile uğraşırken, hız anahtardır. Yavaş bir dil, işleri inanılmaz derecede yavaşlatabilir. Python çok yönlü ve dinamik bir programlama dilidir. Python NumPy, SciPy, PyBrain, Pandas vb. gibi çok fonksiyonel komutlar

barındıran zengin bir kütüphane altyapısına sahiptir. Python'ı öğrenmeyi ve de anlamayı kolaylaştıran bir diğer şey ise arkasında güçlü bir geliştirici topluluğunun olmasıdır (Krause, Lindemann, 2014).

Python'da programımızı yazmak ve yürütmek için kod editörlerinden yararlanmaktayız. Çeşitli Python kod editörleri bulunmaktadır. Bunlar PyCharm, Eclipse, IDLE, Visual Studio Code, Sublime Text, Atom, Rodeo kod editörleri bulunmaktadır. Bu tez çalışmasında Python'da uyguladığımız Apriori, FP-Growth ve Eclat algoritmalarını Visual Studio Code editöründe uygulamalar yapılmıştır. Visual Studio Code, Windows, Linux ve Mac OS'ta kullanılabilen bir kod düzenleyicidir. Visual Studio Code açık kaynak kodludur. Şekil 17'de Python algoritmasının kullanıcı arayüzü verilmiştir.

Tez çalışmasında uyguladığımız Apriori, FP-Growth ve Eclat algoritmalarını Python 3.10.8 versiyonu kullanılmıştır. Bununla birlikte Pandas kütüphanesi 1.5.1 versiyonu ve Numpy kütüphanesi 1.23.0 versiyonu kullanılmıştır.

Şekil 17. Python Apriori Algoritması Kullanıcı Arayüzü



```

1
2 import pandas as pd
3 import numpy as np
4 from mlxtend.frequent_patterns import apriori, association_rules
5 from datetime import datetime
6 from memory_profiler import memory_usage
7
8
9 aa= [ ['Bilgisayar Bilimi','Türk Tarihi'],
10       ['Bilgisayar Bilimi'],
11       ['Fizik','Psikoloji'],
12       ['Bilgisayar Bilimi','Bilgisayar Bilimi'],
13       ['İngiliz edebiyatı','Afrika Edebiyatı'],
14       ['Türk edebiyatı Roman','Tanımsız Kategori'],
15       ['İşletme yönetimi','Türk edebiyatı Roman'],
16       ['Rus edebiyatı','Türk edebiyatı Roman'],
17       ['Türk edebiyatı Roman','Ekonomi'],
18       ['Türk edebiyatı Roman','Türk edebiyatı Roman'],
19       ['Kişisel gelişim','Fizik'],

```

5.2. R Dili ve RStudio

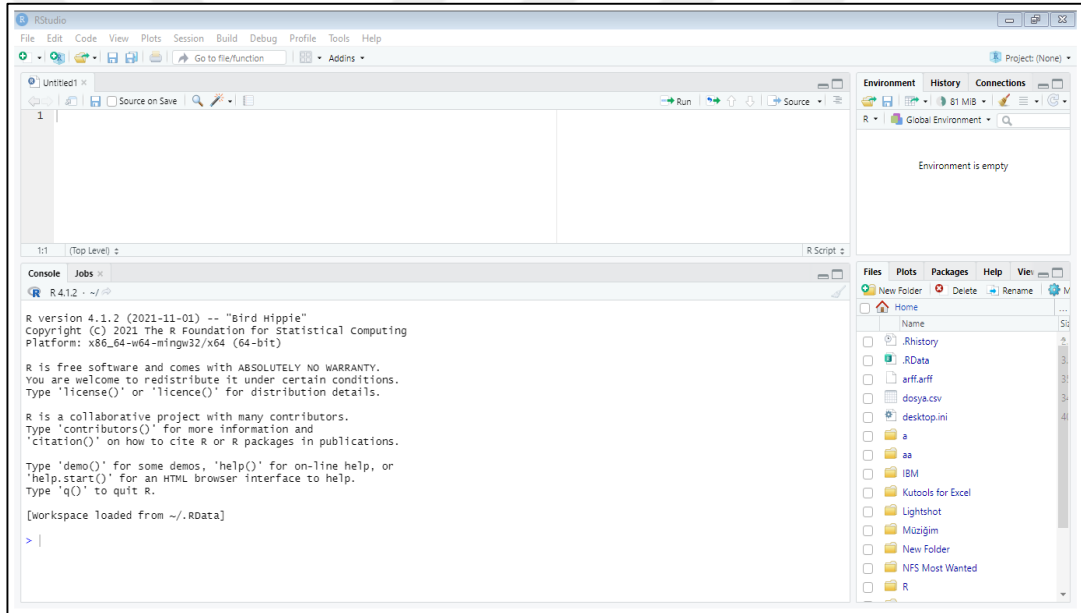
R dili, istatistiksel hesaplama, veri analizi ve bilimsel araştırmalarda verileri temizlemek, analiz etmek, görselleştirmek ve anlamlı hale getirmek için istatistikçiler,

veri bilimcileri, veri analistleri, araştırmacılar ve pazarlamacılar tarafından yaygın olarak kullanılan bir programlama dilidir.

R kodları, açık kaynak ve ücretsiz bir IDE olan RStudio içinde yazılabilir. RStudio, Windows, Linux, Mac OS gibi işletim sistemleri ile uyumludur. RStudio, çok sayıda fonksiyon ve pakete sahiptir (R Core Team, 2021). Şekil 18’de RStudio kullanıcı arayüzü verilmiştir.

Tez çalışmasında uyguladığımız Apriori, FP-Growth ve Eclat algoritmalarını R 4.2.1 versiyonu kullanılmıştır. Bununla birlikte RStudio 3.3.0 versiyonu, ArulesViz kütüphanesi 1.5-1 versiyonu ve Arules kütüphanesi 1.7-5 versiyonu kullanılmıştır.

Şekil 18. RStudio Kullanıcı Arayüzü

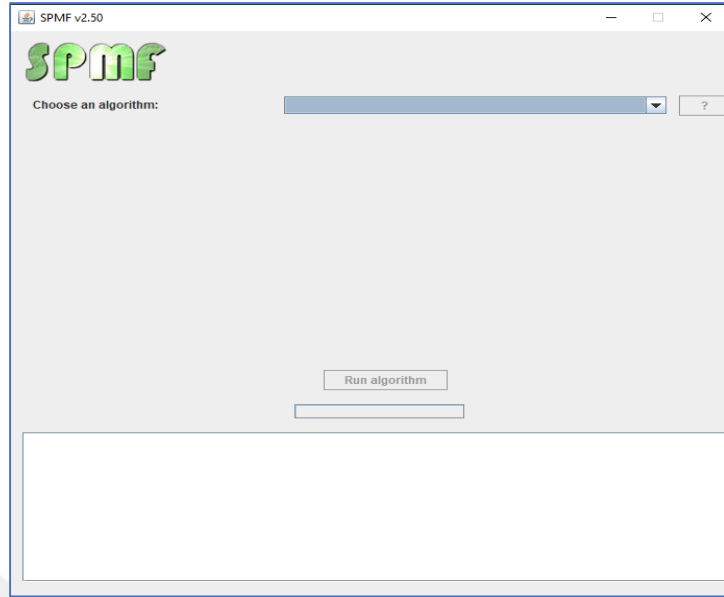


5.3. SPMF

SPMF, Java ile yazılmış, kalıp madenciliği (verilerdeki kalıpların keşfi) konusunda uzmanlaşmış, açık kaynaklı bir yazılım ve veri madenciliği kütüphanesidir. SPMF, basit bir kullanıcı arabirimiyle veya komut satırından bağımsız bir program olarak kullanılabilir. Ayrıca her bir algoritmanın kaynak kodu diğer Java yazılımlarına kolaylıkla entegre edilebilir.

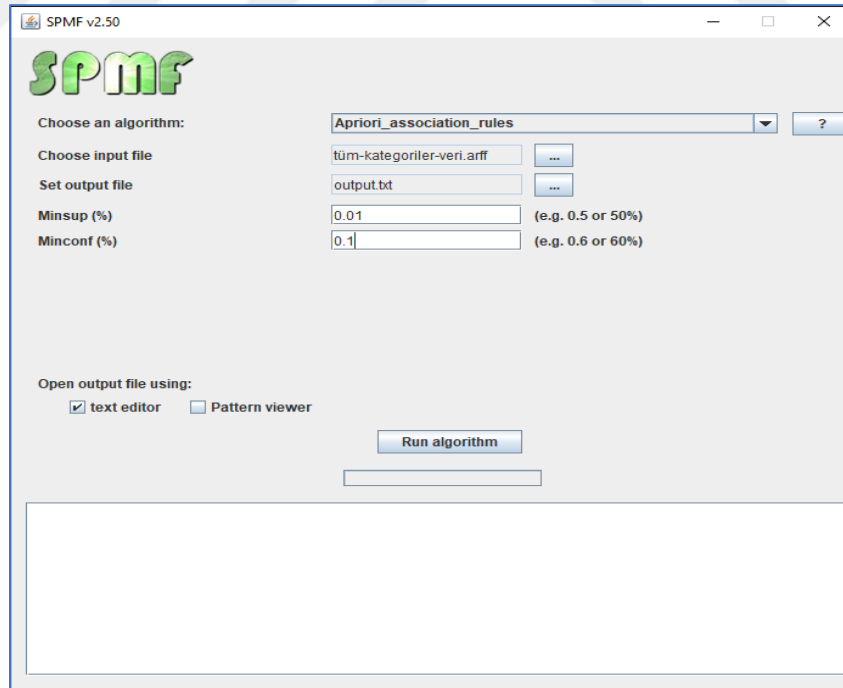
SPMF hızlıdır ve kullanımı kolaydır. Diğer kütüphanelere bağımlılığı yoktur (Anonim, 12 Şubat 2022). Şekil 19’da SPMF kullanıcı arayüzü verilmiştir.

Şekil 19. SPMF Kullanıcı Arayüzü



SPMF kullanımı, programın grafiksel ara yüzünden kullanılacak algoritmanın seçimi ve seçilen algoritmanın parametre değerlerini girerek kolaylıkla yapılabilir.

Şekil 20. SPMF Apriori Algoritması



SPMF’de Apriori algoritmasını seçtikten sonra kullanacağımız “Choose input file” alanına veri setini “.arff” uzantılı olarak yüklememiz gerekmektedir. “Set output file” alanına çıktı dosyasını ayarlamamız gerekmektedir. “Minsup” alanı destek

değerini girdiğimiz alandır, “Minconf” alanına güven değerini girdiğimiz alandır. “Run algorithm” butonuna tıkladığımızda Apriori algoritmamız çalışmaktadır.

Tez çalışmasında uyguladığımız Apriori, FP-Growth ve Eclat algoritmalarını SPMF 2.57 versiyonu kullanılmıştır.



6. DENEY ÖN ÇALIŞMALARI

Bu bölümde veri kümesinin tanıtımı, veri ön işleme ve dönüştürme işlemleri, yapılan deneyler ve performans kriterleri hakkında bilgiler verilmektedir.

6.1. Veri Kümesi

Bu tez çalışmasında, Düzce Üniversitesi lisans/lisansüstü öğrencileriyle akademik/idari personellerinin 2008 ile 2021 yılları arasında üniversite kütüphanesinden yapmış oldukları kitap ödünç işlemlerine dair kayıtların bulunduğu bir veri kümesi kullanılmıştır. Kullanılan veri kümesi 64872 kayıttan ve 27 öznitelikten oluşmaktadır.

6.2. Veri Ön İşleme

Veri ön işleme: aşamasında verilerin temizlenmesi ve dönüştürülmesi işlemleri yapılmıştır.

Veri temizleme: Veri kümesinde bulunan kayıtların bir kısmının bazı sütunlarında eksik veriler vardı. Bu satırların silinerek veri kaybı yaşanmaması adına bu kayıtların ilgili sütunlarına “Tanımsız Kategori” yazılarak doldurmalar yapılmıştır.

Veri dönüştürme: Veri madenciliği işlemlerinin düzgün çalışması için verilerin uygun formata dönüştürülmesi gerekmektedir. Örneğin SPMF programındaki algoritmaların veri kümesi üzerinde çalıştırılabilmesi için veri kümesinde uzantısının arff dosya uzantısına dönüştürülmesi yapılmıştır.

6.3. Birliktelik Kuralları Deneyleri

Literatürde bulunan Birliktelik Kurallarıyla ilgili çalışmalar incelenmiş ve çalışmaların deneylerinde tercih edilmiş olan Destek ve Güven değerleri aşağıdaki Tablo 7’de verilmiştir. Tablo 7’de verilmiş olan Destek ve Güven değerlerinin Python, R ve SPMF platformlarında yapılacak deneylerde kullanılmasına karar verilmiştir.

Tablo 7: Deneylerde kullanılan Destek ve Güven Değerler Tablosu

Platform	Destek Değerleri	Güven Değerleri	Kullanılan Algoritmalar
Python	0,01	0,1, 0,2, 0,3	Apriori
	0,02		
	0,05		
	0,07		
R	0,01	0,1, 0,2, 0,3	Eclat
	0,02		
	0,05		Fp-Growth
	0,07		
SMPF	0,01	0,1, 0,2, 0,3	
	0,02		
	0,05		
	0,07		

6.4. Performans Kriterleri

6.4.1. Çalışma Süresi:

Çalıştırılan programın/kodun, işlemleri ne kadar zamanda tamamladığını ölçmek için kullanılmış ve ms cinsinden ölçülmüştür. SPMF programında Apriori, FP-Growth ve Eclat algoritmaları çalıştırıldığında çalışma süresi “ms” cinsinden çıktı olarak verilmektedir. R programında Apriori, FP-Growth ve Eclat algoritmaların çalışma sürelerini algoritmaların ilk satır ve son satırına “format (Sys.time(),

"%H:%M:%OS3")" R süre kodu eklenerek arasındaki fark “ms” cinsinden hesaplanmıştır. R programında Apriori, FP-Growth ve Eclat algoritmaların çalışma sürelerini algoritmaların ilk satır ve son satırına “datetime.utcnow().isoformat(sep=' ', timespec='milliseconds)' ” Python süre kodu eklenerek arasındaki fark “ms” cinsinden hesaplanmıştır.

6.4.2. Kullanılan Bellek Boyutu

Çalıştırılan programın/kodun çalışmaya başladığı andan bitinceye-bitinceye kadar sistemde ne kadar bellek işgal ettiğini görmek için kullanılmış ve MB cinsinden ölçülmüştür. SPMF programında Apriori, FP-Growth ve Eclat algoritmaları çalıştırıldığında kullanılan bellek boyutunu “MB” cinsinden çıktı olarak vermektedir. R programında Apriori, FP-Growth ve Eclat algoritmaları çalıştırıldığında kullanılan bellek boyutunu “memory.size (max=TRUE)” fonksiyonu ile “MB” cinsinden hesaplanmıştır. Python programında Apriori, FP-Growth ve Eclat algoritmaları çalıştırıldığında kullanılan bellek boyutunu “memory usage” fonksiyonu ile “MB” cinsinden hesaplanmıştır.

Algoritmaların performanslarını değerlendirmek ve karşılaştırmak için yapılan deneyler Intel Core İ7-2670QM CPU @ 2.20Ghz (8 CPUs) işlemci ve 8 GB DDR3 RAM donanım özelliklerine sahip bir bilgisayar üzerinde çalıştırılmıştır. Güvenilir zaman ölçümü ve bellek boyutu için, deneyler devam ederken bilgisayarda eş zamanlı olarak başka bir uygulama çalıştırılmamıştır.

7. VERİ SETLERİNİN ANALİZİ

Bu tez çalışmasında kullanılan Düzce Üniversitesi Kütüphane ödünç alma kayıtlarının düzenlenmesi sonucunda en çok tercih edilen kitap kategorilerinin listesi aşağıda verilmektedir.

Tablo 8. Tüm Veri Seti Kitap Kategori Oranları

Kitap Kategorisi	Kitap Tercih Oranı	Kitap Tercih Sayısı
Türk Edebiyatı Roman	%50,04	7187
Türk Tarihi	%31,58	3099
İşletme Yönetimi	%15,66	2249
Amerikan Edebiyatı	%15,05	2162
Fizik	%9,98	1434
Bilgisayar Bilimi	%9,62	1382
Tıp	%8,43	1211
İngiliz Edebiyatı	%7,92	1138
Psikoloji	%7,56	1086
Felsefe	%6,04	865
Matematik	%5,47	786
Din	%5	718
Fransız Edebiyatı	%4,91	706
Rus Edebiyatı	%4,35	626
Kişisel Gelişim	%3,99	574
Alman Edebiyatı	%3,74	537
Ekonomi	%3,64	523

Tablo 8. Tüm Veri Seti Kitap Kategori Oranları Devamı

Mühendislik	%3,51	504
Araştırma Yöntemleri	%2,86	412
Tanımsız Kategori	%2,71	390
Toplum Bilimleri	%2,68	385
Sanat Tarihi	%2,41	347
Kimya	%1,97	283
Mimari – Mimarlık	%1,50	216
Portekiz Edebiyatı	%1,46	210
Hukuk	%1,31	189
Aile	%1,30	187
Çocuk Psikolojisi	%1,02	147
Yapı Endüstrisi	%1,21	147
İtalyan Edebiyatı	%0,98	142
English Language	%0,91	132
Avustralya Edebiyatı	%0,66	95
Yunan Edebiyatı	%0,65	94
İran Edebiyatı	%0,59	86
Peyzaj Mimarlığı	%0,55	80
İrlanda Edebiyatı	%0,55	79
Çocuk Edebiyatı	%0,52	76
İspanyol Edebiyatı	%0,47	68

Tablo 8. Tüm Veri Seti Kitap Kategori Oranları Devamı

İskoç Edebiyatı	%0,4	66
Arap Edebiyatı	%0,44	64
Norveç Edebiyatı	%0,44	64
Kırgız Edebiyatı	%0,43	62
Tarım	%0,41	59
Japon Edebiyatı	%0,39	57
Afrika Edebiyatı	%0,30	44
Çevre Eğitimi	%0,28	41
Spor	%0,26	38
Çin Edebiyatı	%0,22	32
ABD Tarih	%0,21	31
Elektrik	%0,21	31
Kanada Edebiyatı	%0,18	27
Afgan Edebiyatı	%0,16	23
Hint Edebiyatı	%0,16	23
Kürt Edebiyatı	%0,16	23
Biyoloji	%0,11	17
Doğa Bilimleri	%0,11	17
İsrail Edebiyatı	%0,08	12
Müzik Eğitimi	%0,06	9
Coğrafya ve Seyahat	%0,04	7

Tablo 8. Tüm Veri Seti Kitap Kategori Oranları Devamı

Bulgar Edebiyatı	%0,04	6
Fars Edebiyatı	%0,04	6
Arjantin Edebiyatı	%0,03	5
Botanik	%0,03	5
İzlanda Edebiyatı	%0,01	2
Macar Romanı	%0,007	1
Toplam	%100	14360

Tablo 9. Erkekler Veri Seti Kitap Kategori Oranları

Kitap Kategorisi	Kitap Tercih Oranı	Kitap Tercih Sayısı
Türk Edebiyatı Roman	%50,95	2908
Türk Tarihi	%21,62	1234
İşletme Yönetimi	%16,05	916
Amerikan Edebiyatı	%15,36	877
Bilgisayar Bilimi	%9,77	558
Fizik	%9,70	554
Tıp	%8,56	489
İngiliz Edebiyatı	%8,11	463
Psikoloji	%7,92	452
Matematik	%5,97	341
Felsefe	%5,90	337

Tablo 9. Erkekler Veri Seti Kitap Kategori Oranları Devam

Fransız Edebiyatı	%5,37	307
Din	%4,67	267
Rus Edebiyatı	%4,50	257
Kişisel Gelişim	%4,29	245
Ekonomi	%3,92	224
Alman Edebiyatı	%3,64	208
Mühendislik	%3,38	193
Tanımsız Kategori	%2,92	167
Toplum Bilimleri	%2,75	157
Araştırma Yöntemleri	%2,71	155
Sanat Tarihi	%2,41	138
Kimya	%1,82	104
Portekiz Edebiyatı	%1,59	91
Mimari – Mimarlık	%1,45	83
Hukuk	%1,27	73
Aile	%1,22	70
Çocuk Psikolojisi	%1,08	62
Yapı Endüstrisi	%1,01	58
İtalyan Edebiyatı	%0,96	55
English Language	%0,85	49
Yunan Edebiyatı	%0,75	43
Peyzaj Mimarlığı	%0,68	39

Tablo 9. Erkekler Veri Seti Kitap Kategori Oranları Devam

İran Edebiyatı	%0,63	36
İrlanda Edebiyatı	%0,59	34
Avustralya Edebiyatı	%0,56	32
İskoç Edebiyatı	%0,50	29
Çocuk Edebiyatı	%0,49	28
Kırgız Edebiyatı	%0,49	28
Tarım	%0,49	28
Arap Edebiyatı	%0,47	27
İspanyol Edebiyatı	%0,47	27
Norveç Edebiyatı	%0,42	24
Çevre Eğitimi	%0,29	17
Afrika Edebiyatı	%0,28	16
Çin Edebiyatı	%0,21	12
Kanada Edebiyatı	%0,21	12
Elektrik	%0,19	11
Spor	%0,19	11
Afgan Edebiyatı	%0,17	10
Japon Edebiyatı	%0,17	10
Biyoloji	%0,12	7
Hint Edebiyatı	%0,12	7
Kürt Edebiyatı	%0,12	7
Doğa Bilimleri	%0,08	5

Tablo 9. Erkekler Veri Seti Kitap Kategori Oranları Devam

İsrail Edebiyatı	%0,08	5
Coğrafya ve Seyahat	%0,07	4
Fars Edebiyatı	%0,07	4
Botanik	%0,03	2
İzlanda Edebiyatı	%0,03	2
Müzik Eğitimi	%0,03	2
Arjantin Edebiyatı	%0,01	1
Pakistan Romanı	%0,01	1
Toplam	%100	57007

Tablo 10. Bayanlar Veri Seti Kitap Kategori Oranları

Kitap Kategorisi	Kitap Tercih Oranı	Kitap Tercih Sayısı
Türk Edebiyatı Roman	%49,45	4279
Türk Tarihi	%21,55	1865
İşletme Yönetimi	%15,40	1333
Amerikan Edebiyatı	%14,85	1285
Fizik	%10,17	880
Bilgisayar Bilimi	%9,52	824
Tıp	%8,34	722
İngiliz Edebiyatı	%7,80	675
Psikoloji	%7,32	634
Felsefe	%6,10	528

Tablo 10. Bayanlar Veri Seti Kitap Kategori Oranları Devam

Din	%5,21	451
Matematik	%5,14	445
Fransız Edebiyatı	%4,61	399
Rus Edebiyatı	%4,26	369
Alman Edebiyatı	%3,80	329
Kişisel Gelişim	%3,80	329
Mühendislik	%3,59	311
Ekonomi	%3,45	299
Araştırma Yöntemleri	%2,97	257
Toplum Bilimleri	%2,69	233
Tanımsız Kategori	%2,57	223
Sanat Tarihi	%2,41	209
Kimya	%2,06	179
Mimari – Mimarlık	%1,53	133
Portekiz Edebiyatı	%1,37	119
Aile	%1,35	117
Hukuk	%1,34	116
Yapı Endüstrisi	%1,02	89
İtalyan Edebiyatı	%1,00	87
Çocuk Psikolojisi	%0,98	85
English Language	%0,95	83
Avustralya Edebiyatı	%0,72	63

Tablo 10. Bayanlar Veri Seti Kitap Kategori Oranları Devam

Yunan Edebiyatı	%0,58	51
İran Edebiyatı	%0,57	50
Çocuk Edebiyatı	%0,55	48
Japon Edebiyatı	%0,54	47
İrlanda Edebiyatı	%0,52	45
İspanyol Edebiyatı	%0,47	41
Peyzaj Mimarlığı	%0,47	41
Norveç Edebiyatı	%0,46	40
Arap Edebiyatı	%0,42	37
İskoç Edebiyatı	%0,42	37
Kırgız Edebiyatı	%0,39	34
Tarım	%0,35	31
Afrika Edebiyatı	%0,32	28
Spor	%0,31	27
Çevre Eğitimi	%0,27	24
Çin Edebiyatı	%0,23	20
Elektrik	%0,23	20
Hint Edebiyatı	%0,18	16
Kürt Edebiyatı	%0,18	16
Kanada Edebiyatı	%0,17	15
Doğa Bilimleri	%0,13	12
Biyoloji	%0,11	10

Tablo 10. Bayanlar Veri Seti Kitap Kategori Oranları Devam

İsrail Edebiyatı	%0,08	7
Müzik Eğitimi	%0,08	7
Bulgar Edebiyatı	%0,06	6
Arjantin Edebiyatı	%0,04	4
Botanik	%0,03	3
Coğrafya ve Seyahat	%0,03	3
Fars Edebiyatı	%0,02	2
Macar Romanı	%0,01	1
Toplam	%100	8653

Tüm veri setinde 65 kitap kategorisi elde edilmiştir. Toplam kitap kaydı 14360'tır. En çok tercih edilen kitap kategorisi, 7187 ödünç kayıt sayısı ve %50,04 oranı ile Türk Edebiyatı Roman olarak görülmektedir. En az tercih edilen kitap kategorisi, 1 ödünç kayıt sayısı ve %0,007 oranı ile Macar Romanı olarak görülmektedir.

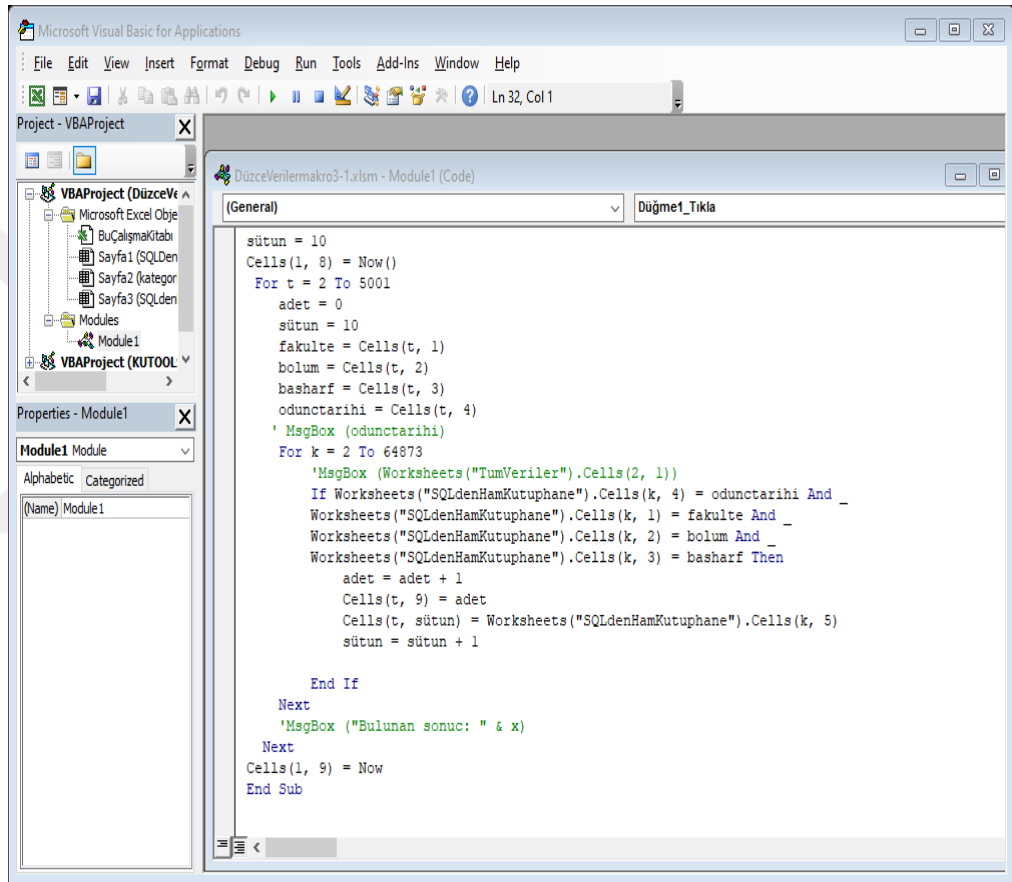
Erkekler veri setinde 63 kitap kategorisi elde edilmiştir. Toplam kitap kaydı 5707'dir. En çok tercih edilen kitap kategorisi, 2908 ödünç kayıt sayısı ve %50,95 oranı ile Türk Edebiyatı Roman olarak görülmektedir. En az tercih edilen kitap kategorisi, 1 ödünç kayıt sayısı ve %0,001 oranı ile Pakistan Romanı olarak görülmektedir.

Bayanlar veri setinde 62 kitap kategorisi elde edilmiştir. Toplam kitap kaydı 8653'tür. En çok tercih edilen kitap kategorisi, 4279 ödünç kayıt sayısı ve %49,45 oranı ile Türk Edebiyatı Roman olarak görülmektedir. En az tercih edilen kitap kategorisi, 1 ödünç kayıt sayısı ve %0,001 oranı ile Macar Romanı olarak görülmektedir.

7.1. Veri Setlerinin Yazılıma Hazırlanması

2008 – 2021 yılına ait olan Düzce Üniversitesi Kütüphane kitap ödünç alma kayıtları, Excel ortamında her satırda bir öğrencinin ödünç aldığı kitap kategorileri olarak düzenlenmiştir. Fakat bu işlem yapılırken beraber alımı tercih edilen kitap kategorilerini tespit edebilmek adına, tek kitap ödünç alma kayıtları göz ardı edilmiş ve 2 veya daha fazla kitabın beraber ödünç alındığı kayıtlar dikkate alınmıştır.

Şekil 21. Kategorilerin Veri Setinden Çekilmesi İşlemi



Şekil 21’de Excel makro kodu sayesinde Düzce Üniversitesi lisans ve lisansüstü öğrencileriyle akademik ve idari personellerinin kütüphaneden aldıkları kitapları listeledikten sonra kategoriye göre listelenmesi aşağıdaki şekilde bir kısmı gösterilmektedir.

Şekil 22. Tüm Veri Seti Düzenlenmiş Hali

	A	B	C	D	E	F	G
1	İngiliz edeb	İngiliz edeb	Türk edeb	Alman Edebiyatı			
2	Bilgisayar	Türk Tarihi					
3	Bilgisayar	Fizik					
4	Fransız Ed	Türk edeb	Türk Tarihi				
5	Fizik	Psikoloji					
6	Bilgisayar	Bilgisayar Bilimi					
7	İngiliz edeb	Afrika Edebiyatı					
8	Türk edeb	Tanımsız Kategori					
9	Türk edeb	Matemati	Matematik				
10	İşletme y	Türk edebiyatı	Roman				
11	Türk edeb	Amerikan	Tanımsız K	Fransız Ed	İtalyan edebiyatı		
12	Rus edebi	Türk edebiyatı	Roman				
13	Türk edeb	Mühendis	Türk edeb	Türk Tarih	Türk edebiyatı	Roman	
14	Türk edeb	Ekonomi					
15	English lar	Türk edeb	Türk edeb	English lar	Mühendislik		
16	Türk edeb	Türk edeb	İngiliz edebiyatı				
17	Türk edeb	Türk edebiyatı	Roman				
18	Kişisel gel	Fizik					

Yukarıda verilmiş olan Şekil 22’de kategorilerin düzenlenmiş halinin bir kısmı gösterilmektedir.

Şekil 23. Veri Setine “t” Eklenmesi İşlemi

```

Sub Düğme2_Tıkla()
    For k = 1 To 5707
        m = Cells(k, 1)
        For t = 2 To m + 1
            For y = 1 To 64
                If Cells(k, t) = Worksheets("erkekler2").Cells(1, y) Then
                    Worksheets("erkekler2").Cells(k + 1, y) = "t"
                End If
            Next y
        Next t
    Next k
End Sub

```

Şekil 22’de kategoriler düzenlendikten sonra yeni bir sayfada kaç tane kategorinin olduğu bulunur ve ilk satıra eklenmektedir. Daha sonra şekil 23’teki makro kodu sayesinde her satırda hangi kategori var ise o hücreye “t” yazılır. Beraber ödünç alınmış olan kitap kategorileri aynı satırda “t” karakterlerinin eklenmesiyle

Şekil 26. Veri Setine “?” Eklenmesi İşlemi 1

	A	B	C	D	E	F	G	H	I	J	K	L
1	ABD Tarih	ABD Tarih	Afgan Ede	Afrika Edebi	Aile	Alman Ede	Amerikan Ede	Arap Ede	Araştırma	Arjantin e	Avustralya	Bilgisayar
2	?	?	?	?	?	?	?	?	?	?	?	?
3	?	?	?	?	?	?	?	?	?	?	?	?
4	?	?	?	?	?	?	?	?	?	?	?	?
5	?	?	?	?	?	?	?	?	?	?	?	?
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	?	?	?	?	?	?	?	?	?	?	?	?
31	?	?	?	?	?	?	?	?	?	?	?	?
32	?	?	?	?	?	?	?	?	?	?	?	?
33	?	?	?	?	?	?	?	?	?	?	?	?
34	?	?	?	?	?	?	?	?	?	?	?	?
35	?	?	?	?	?	?	?	?	?	?	?	?
36	?	?	?	?	?	?	?	?	?	?	?	?
37	?	?	?	?	?	?	?	?	?	?	?	?
38	?	?	?	?	?	?	?	?	?	?	?	?
39	?	?	?	?	?	?	?	?	?	?	?	?

7.2. Arff Dosya Uzantısı

Veri setinin SPMF platformunda düzgün çalıştırılabilmesi için arff dosya uzantısına çevrilmesi gerekmektedir. Bu işlemi yapmak için Weka programı sayesinde kolayca yapılmaktadır. Şekil 26’da ki A1 sütunundaki veriler seçildikten sonra farklı bir Excel sayfasına kopyalanıp “csv” uzantılı dosya formatına çevrilmesi gerekmektedir. Weka programında “csv” dosyasını seçip farklı kaydet seçeneği ile “.arff” dosya uzantılı çevrilmiştir. Çevrilen veri seti aşağıdaki şekil 27’de gösterilmiştir.

Şekil 27. Tüm Veri Setinin. Arff Dosyası Olarak Düzenlenmiş Hali-1

```

@relation 'Tüm Kategoriler'

@attribute 'ABD Tarih' {t}
@attribute 'Afgan Edebiyatı' {t}
@attribute 'Afrika Edebiyatı' {t}
@attribute 'Aile ' {t}
@attribute 'Alman Edebiyatı' {t}
@attribute 'Amerikan Edebiyatı' {t}
@attribute 'Arap Edebiyatı' {t}
@attribute 'Araştırma yöntemleri' {t}
@attribute 'Arjantin edebiyatı' {t}
@attribute 'Avustralya edebiyatı' {t}
@attribute 'Bilgisayar Bilimi' {t}
@attribute Biyoloji {t}
@attribute Botanik {t}
@attribute 'Bulgar edebiyatı' {t}
@attribute 'Coğrafya ve seyahat' {t}
@attribute 'Çevre eğitimi' {t}
@attribute 'Çin Edebiyatı' {t}
@attribute 'Çocuk edebiyatı' {t}
@attribute 'Çocuk Psikolojisi' {t}
@attribute Din {t}
@attribute 'Doğa Bilimleri' {t}|
@attribute Ekonomi {t}
@attribute 'Elektrik ' {t}
@attribute 'English language' {t}
@attribute 'Fars edebiyatı' {t}
@attribute Felsefe {t}
@attribute Fizik {t}
@attribute 'Fransız Edebiyatı' {t}
@attribute 'Hint Edebiyatı' {t}
@attribute Hukuk {t}
@attribute 'İngiliz edebiyatı' {t}
@attribute 'İran edebiyatı' {t}
@attribute 'İrlanda edebiyatı' {t}
@attribute 'İskoç Edebiyat' {t}
@attribute 'İspanyol Edebiyatı' {t}
@attribute 'İsrail Edebiyatı' {t}
@attribute 'İşletme yönetimi' {t}
@attribute 'İtalyan edebiyatı' {t}
@attribute 'İzlanda edebiyatı' {t}
@attribute 'Japon Edebiyatı' {t}
@attribute 'Kanada Edebiyatı' {t}
@attribute 'Kırgız edebiyatı' {t}
@attribute Kimya {t}
@attribute 'Kişisel gelişim' {t}
@attribute 'Kürt edebiyatı' {t}
@attribute 'Macar romanı' {t}
@attribute Matematik {t}
@attribute 'Mimari - Mimarlık' {t}
@attribute Mühendislik {t}
@attribute 'Müzik eğitimi' {t}
@attribute 'Norveç Edebiyatı' {t}

```


8. YÖNTEM

Bu tez çalışmasında Düzce Üniversitesi kütüphane kitap ödünç alma kayıtları veri setine birliktelik kuralları algoritmalarından Apriori, FP-Growth ve Eclat algoritmalarının Python, RStudio ve SPMF platformlarında analiz edilip beraber ödünç alımı tercih edilmiş kitap kategorilerinin tümü ve cinsiyetlere göre dağılımları analiz edilmiş; yapılan bu analiz işlemlerine dair deneylerin çalışma süreleri ve kullanılan bellek miktarları bakımından karşılaştırmaları yapılmıştır.

8.1. Evren ve Örneklem

Yazılan bu tezin ana kütlesi, Düzce Üniversitesi kütüphane kitap ödünç alma kayıtları veri setinden oluşmaktadır. Bu araştırmanın evreni 14360 kayıttan ve 65 kitap kategorisinden oluşmaktadır.

8.2. Verilerin Toplanması

Verilerin toplanmasında Düzce Üniversitesi kütüphanesinde kitap ödünç alma kayıtları Excel dosyası olarak alınmıştır.

8.3. Verilerin Analizi

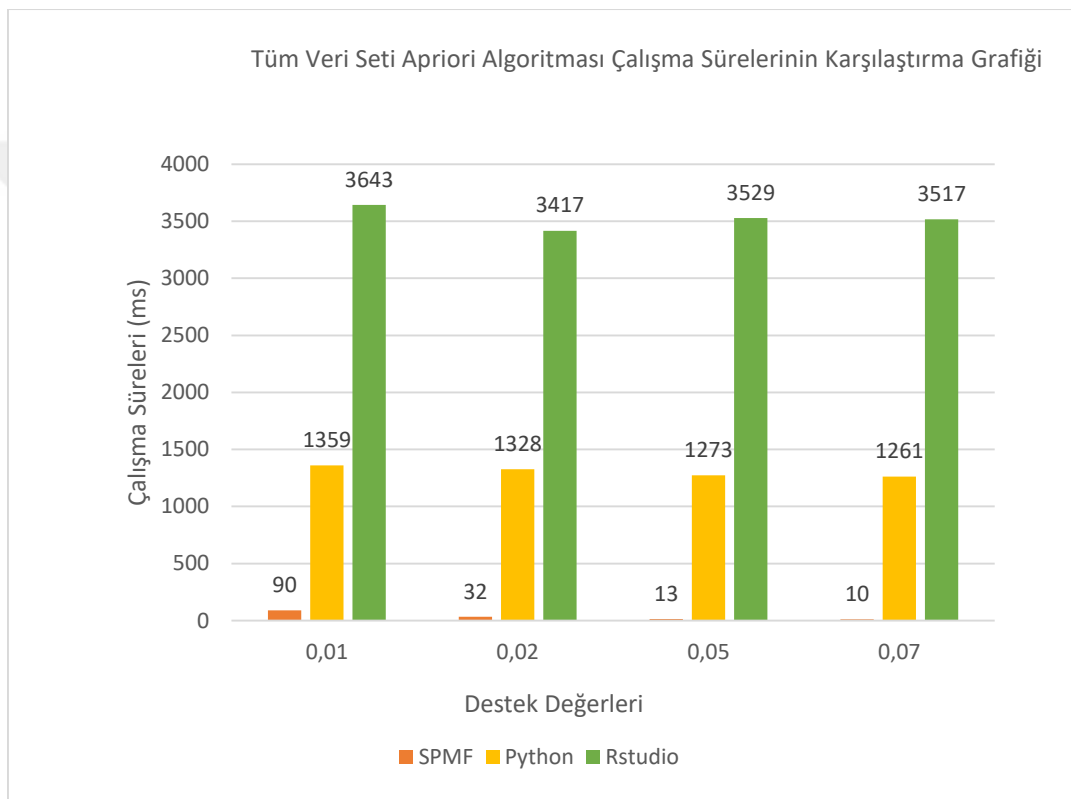
Veriler R, Python ve SPMF platformlarında birliktelik kurallarından Apriori, FP-Growth ve Eclat algoritmaları kullanılarak kitap kategorileri ve cinsiyet yönünden analiz edilmiştir.

9. BULGULAR VE YORUM

Bu bölümde yapılan Birliktelik Kuralları deneylerinden elde edilen sonuçlar ve bu sonuçlara ilişkin yorumlar verilmektedir.

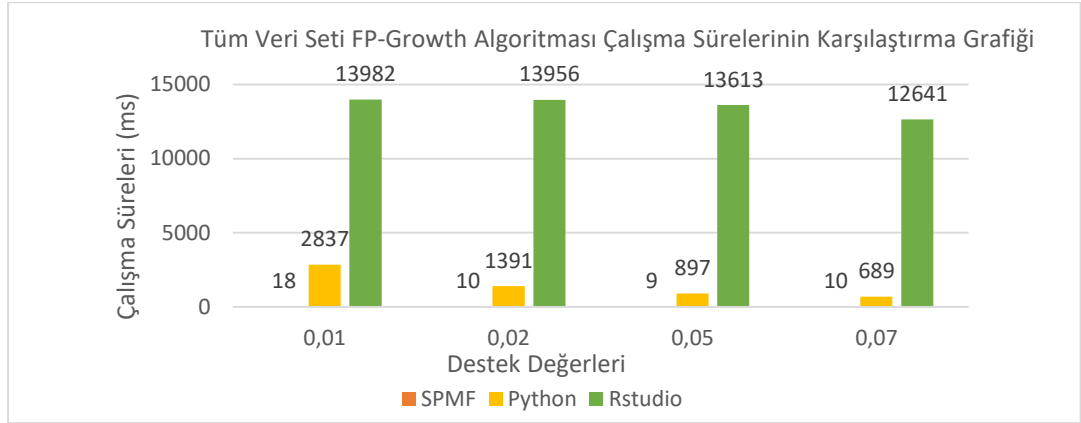
9.1. Çalışma Sürelerinin Karşılaştırılması

Şekil 30. Tüm Veri Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması



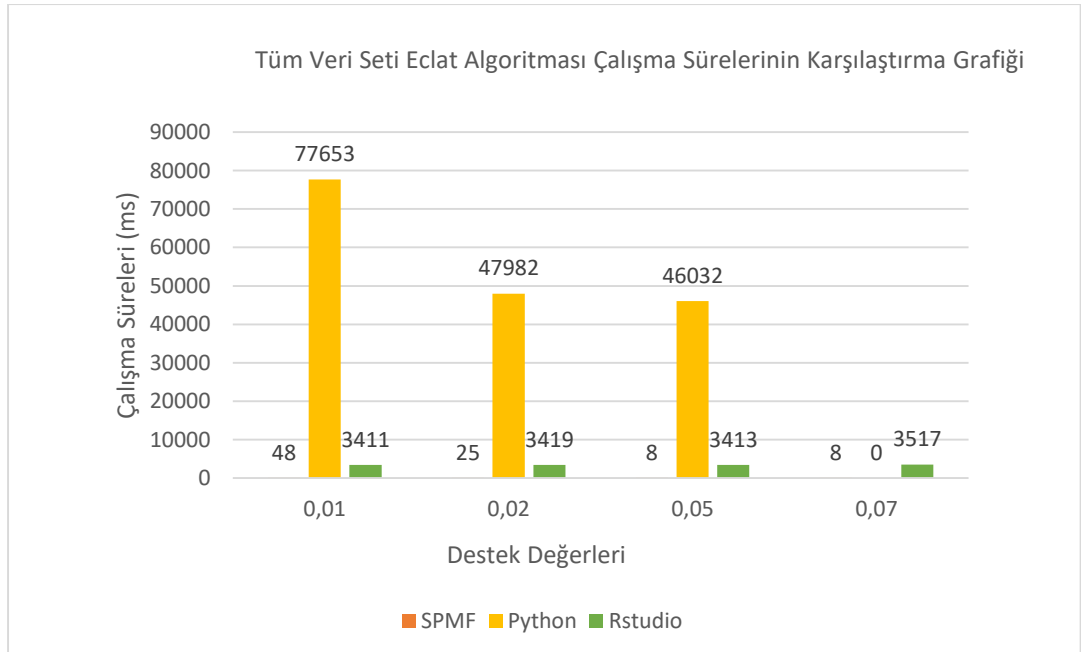
Şekil 30'da SPMF, Python ve RStudio programlarında tüm veri setinde Apriori algoritmasının 0,01, 0,02, 0,05, 0,07 destek değerlerine 0,1, 0,2, 0,3 güven değerleri verilerek çalışma sürelerinin (ms) ortalamaları alınarak grafik çizilmiştir. Grafiğe göre Apriori algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu destek değeri 0,07'dir. RStudio programının 0,01 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 31. Tüm Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması



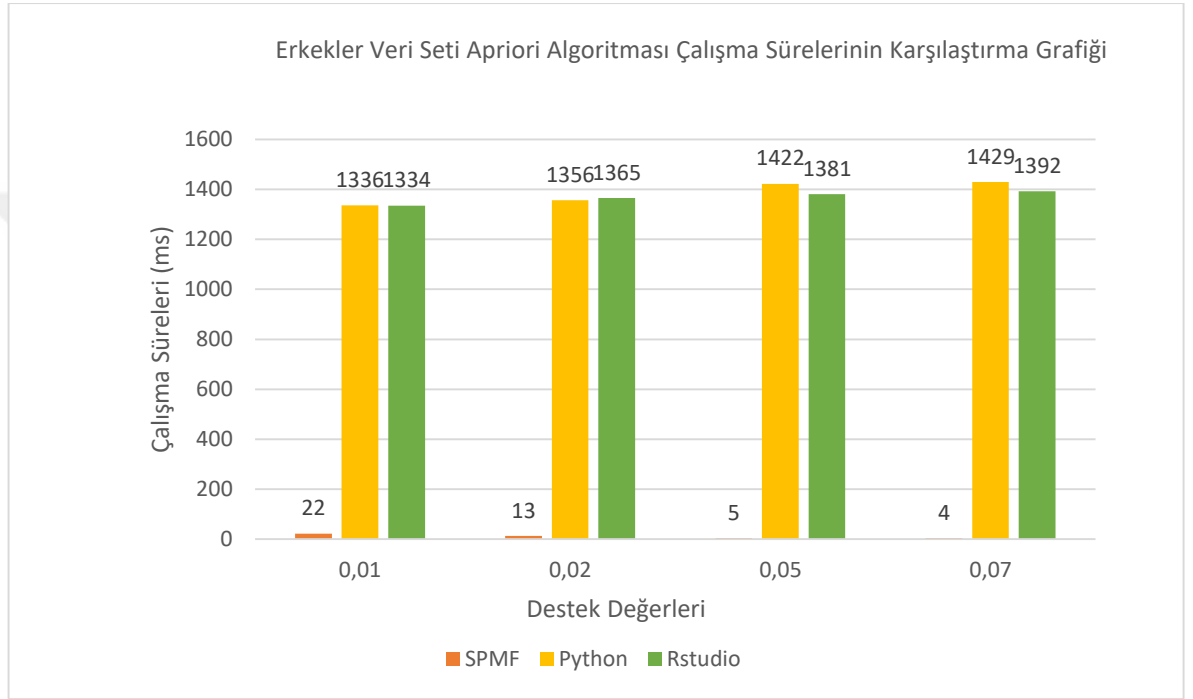
Şekil 31’de FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 destek değeridir. RStudio programının 0,01 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 32. Tüm Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması



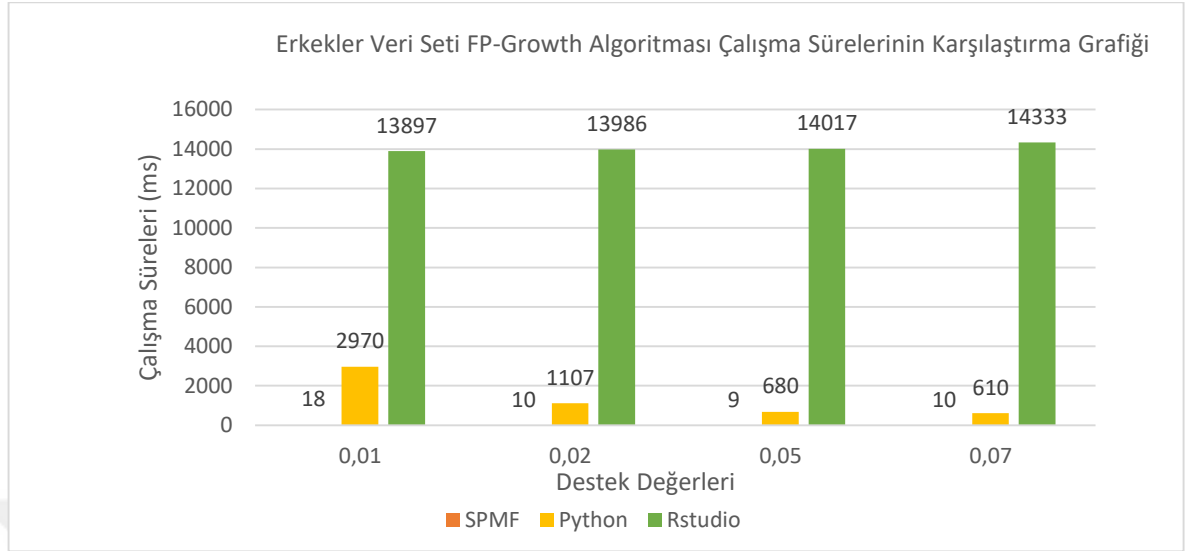
Şekil 32’de Eclat algoritmasının tüm platformlarda karar verilmiş destek değerlerinin çalışma süreleri alınarak analiz edilmiştir. Grafiğe göre Eclat algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 ve 0,07 destek değeridir. Çalışma süresinin en çok olduğu program ise Python programıdır.

Şekil 33. Erkekler Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması



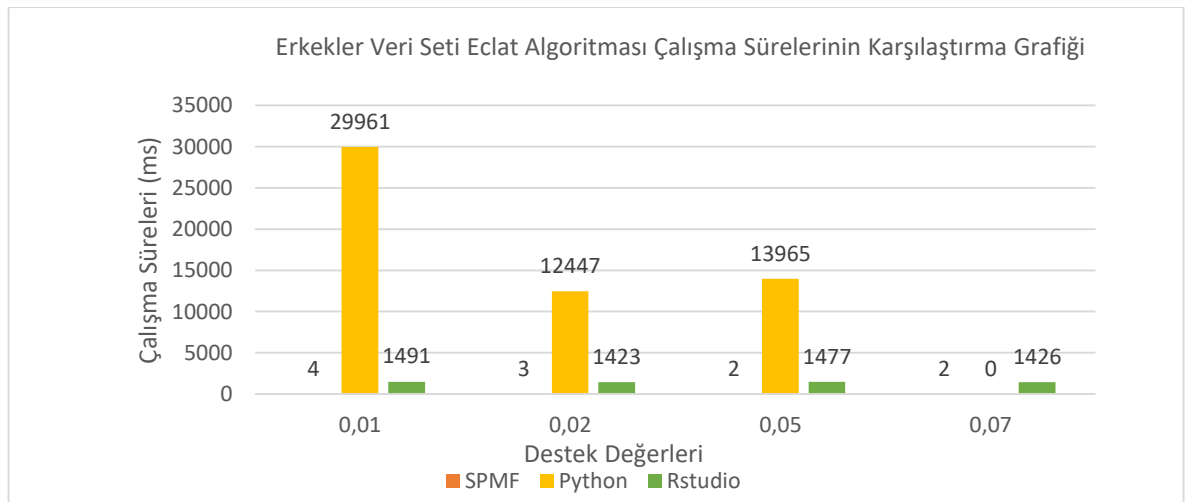
Şekil 33’te Apriori algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre Apriori algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,07 destek değeridir. Python programının 0,07 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 34. Erkekler Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması



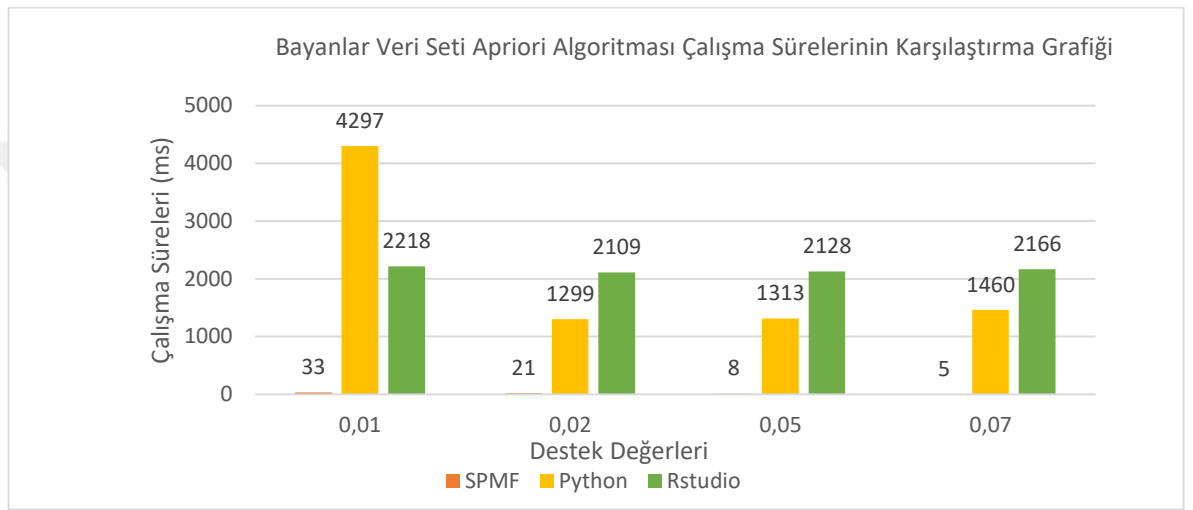
Şekil 34'te FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 destek değeridir. RStudio programının 0,07 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 35. Erkekler Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması



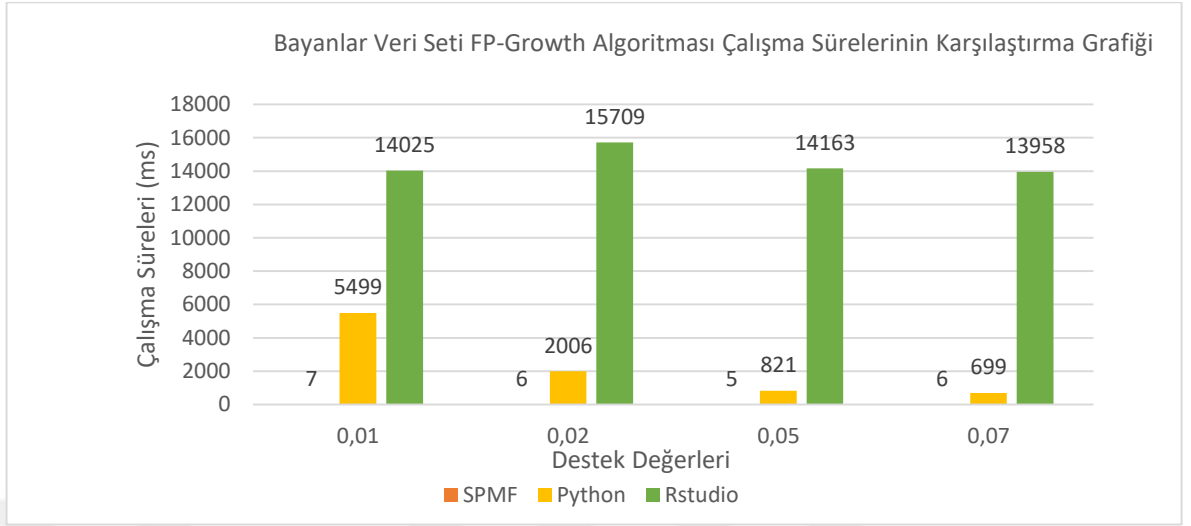
Şekil 35'te Eclat algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre Eclat algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 ve 0,07 destek değeridir. Python programının 0,01 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 36. Bayanlar Veri Setinde Apriori Algoritmasının Çalışma Sürelerinin Karşılaştırılması



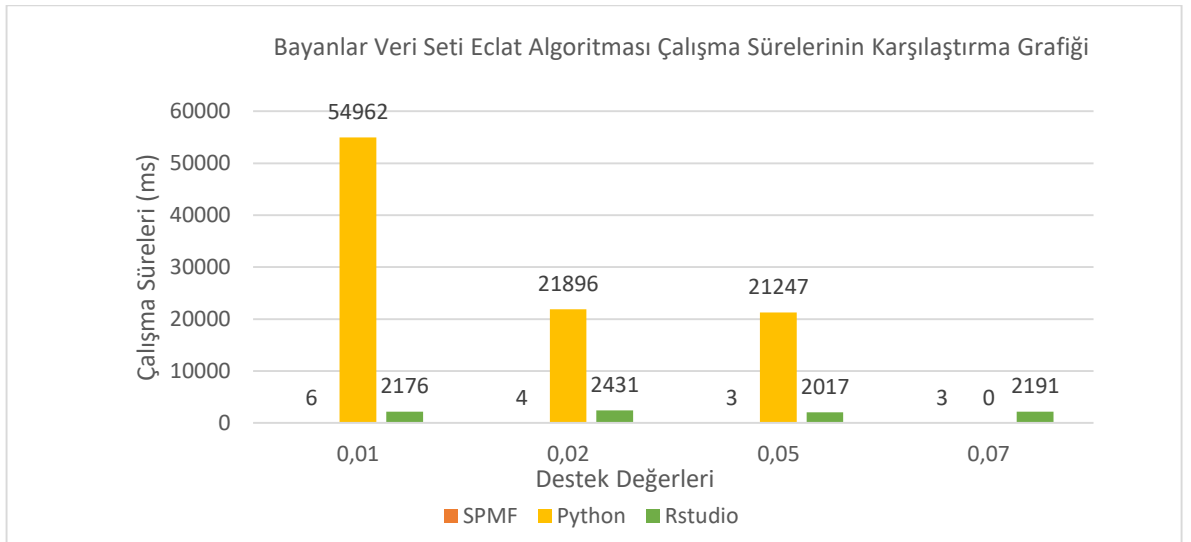
Şekil 36'da Apriori algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre Apriori algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,07 destek değeridir. Python programının 0,01 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

Şekil 37. Bayanlar Veri Setinde FP-Growth Algoritmasının Çalışma Sürelerinin Karşılaştırılması



Şekil 37’de FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 destek değeridir. RStudio programının 0,02 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

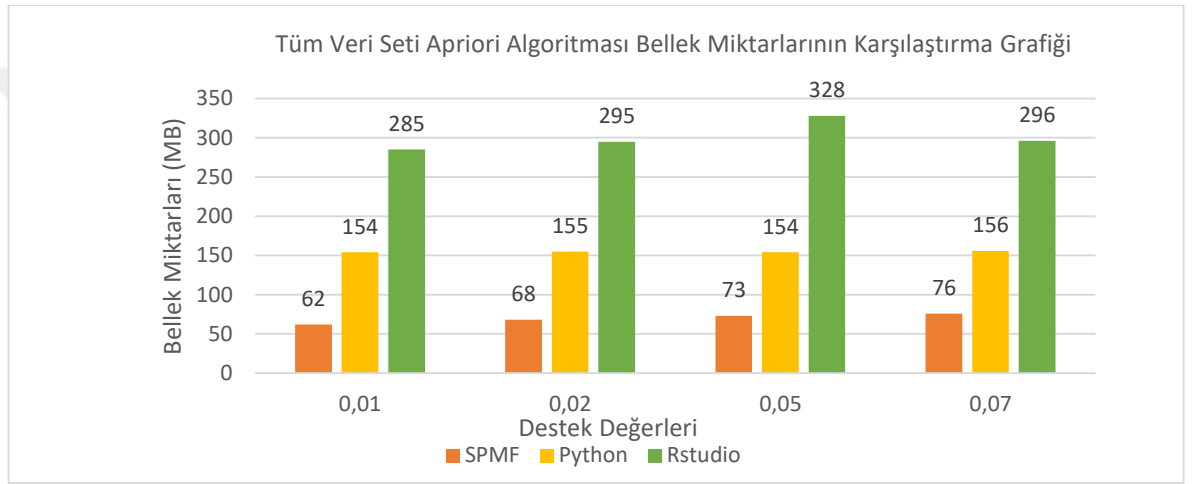
Şekil 38. Bayanlar Veri Setinde Eclat Algoritmasının Çalışma Sürelerinin Karşılaştırılması



Şekil 38’de Eclat algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle çalışma sürelerinin ortalamaları alınarak analiz edilmiştir. Grafiğe göre Eclat algoritmasının çalışma süresinin en az olduğu program SPMF programıdır. SPMF programının çalışma süresinin en az olduğu 0,05 ve 0,07 destek değeridir. Python programının 0,01 destek değerinde çalışma süresinin en çok olduğu gözlemlenmiştir.

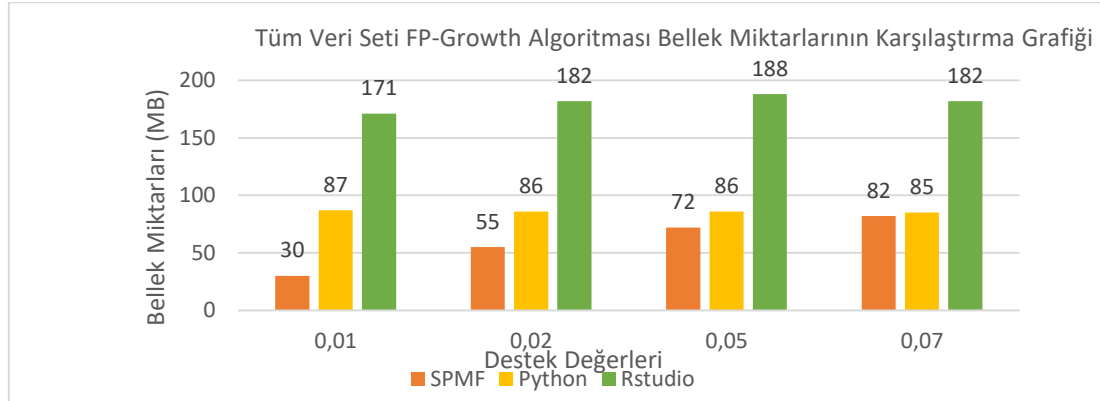
9.2. Bellek Miktarları Karşılaştırılması

Şekil 39. Tüm Veri Setinde Apriori Algoritmasının Bellek Miktarının Karşılaştırılması



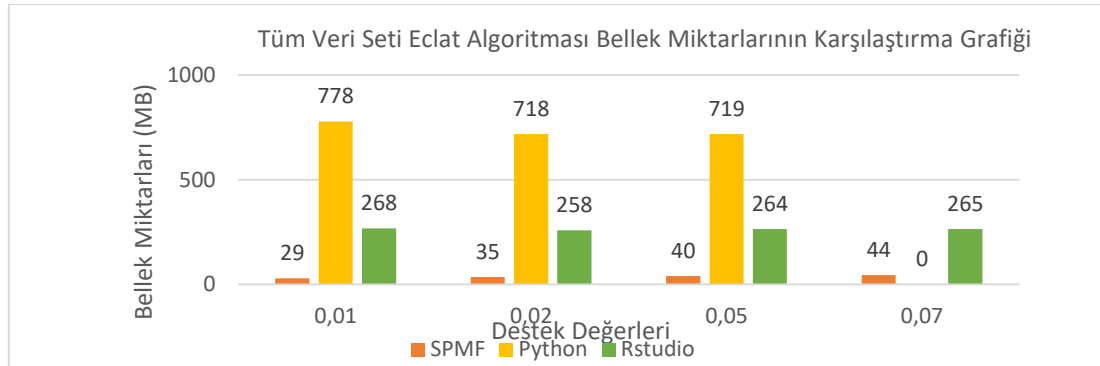
Şekil 39’da Apriori algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre Apriori algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 destek değeridir. RStudio programının 0,05 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 40. Tüm Veri Setinde FP-Growth Algoritmasının Bellek Miktarının Karşılaştırılması



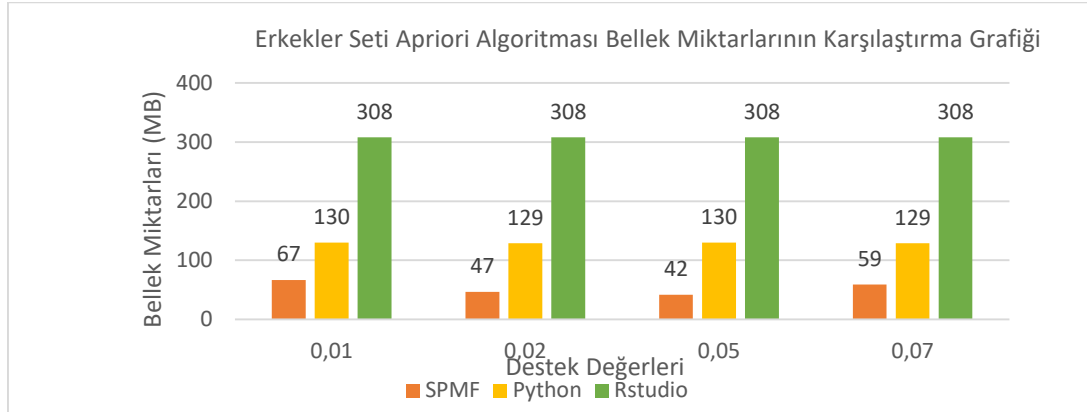
Şekil 40'ta FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 destek değeridir. RStudio programının 0,05 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 41. Tüm Veri Setinde Eclat Algoritmasının Bellek Miktarının Karşılaştırılması



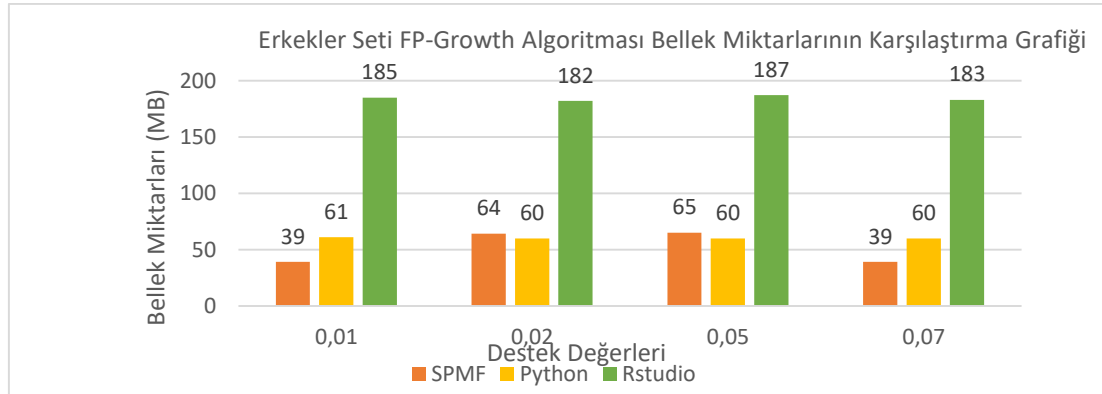
Şekil 41'de Eclat algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre Eclat algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 destek değeridir. Python programının 0,01 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 42. Erkekler Veri Setinde Apriori Algoritmasının Bellek Miktarının Karşılaştırılması



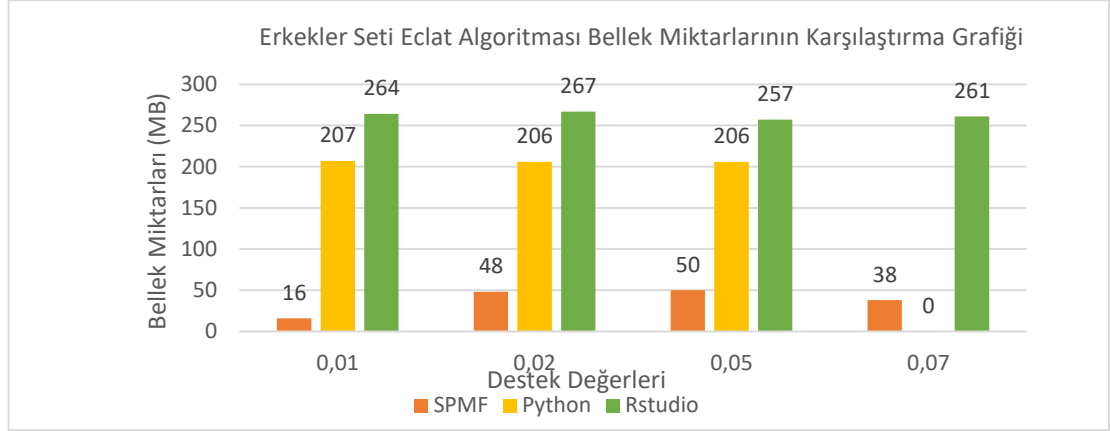
Şekil 42’de Apriori algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre Apriori algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,05 destek değeridir. RStudio programının bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 43. Erkekler Veri Setinde FP-Growth Algoritmasının Bellek Miktarının Karşılaştırılması



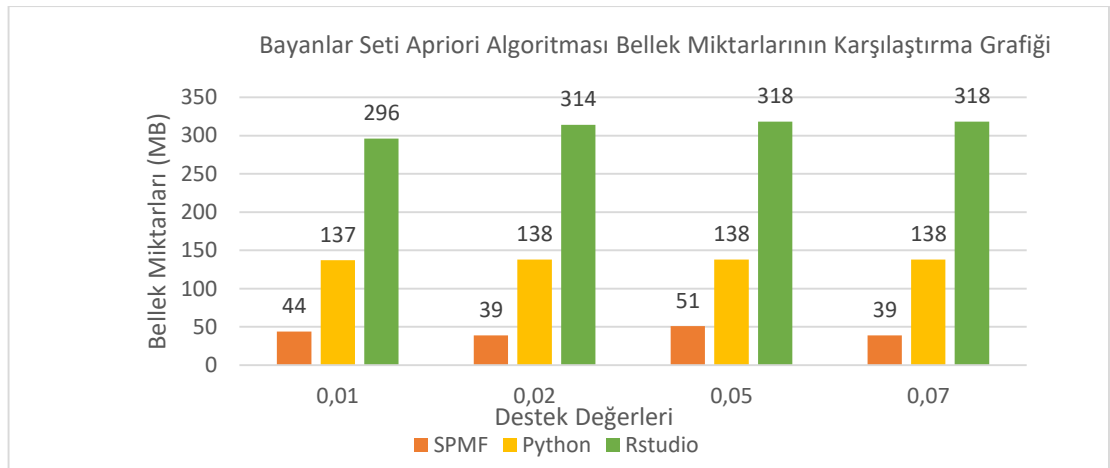
Şekil 43’te FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 ve 0,07 destek değeridir. RStudio programının 0,05 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 44. Erkekler Veri Setinde Eclat Algoritmasının Bellek Miktarının Karşılaştırılması



Şekil 44'te Eclat algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre Eclat algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 destek değeridir. RStudio programının 0,02 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

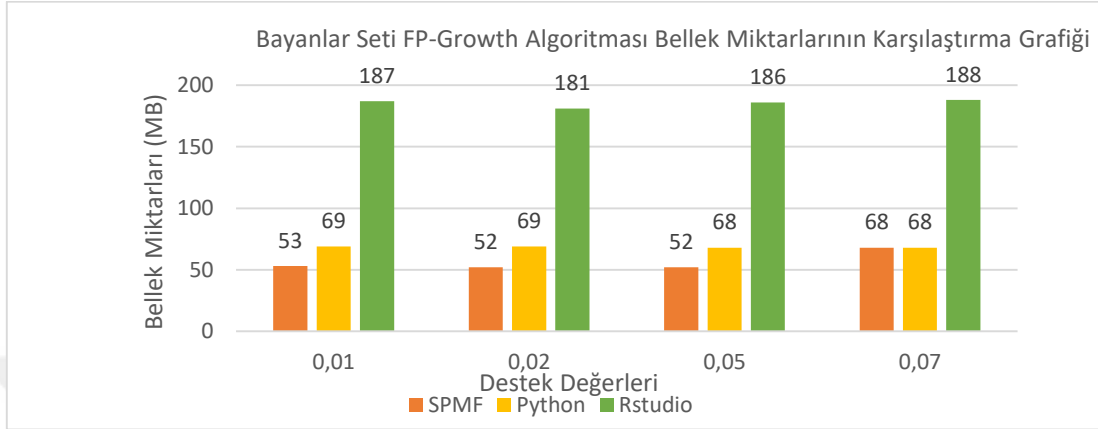
Şekil 45. Bayanlar Veri Setinde Apriori Algoritmasının Bellek Miktarının Karşılaştırılması



Şekil 45'te Apriori algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre Apriori algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,02 ve 0,07

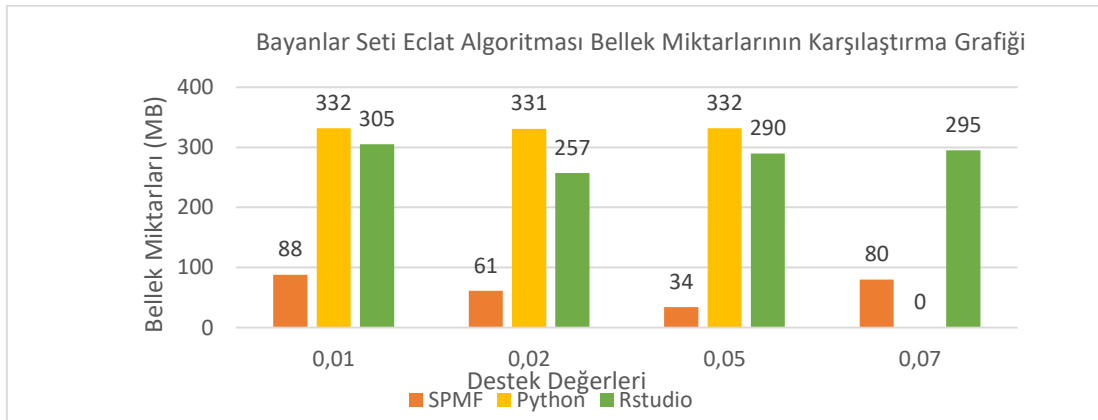
destek değeridir. RStudio programının 0,05 ve 0,07 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 46. Bayanlar Veri Setinde FP-Growth Algoritmasının Bellek Miktarının Karşılaştırılması



Şekil 46’da FP-Growth algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz edilmiştir. Grafiğe göre FP-Growth algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,01 destek değeridir. RStudio programının 0,07 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

Şekil 47. Bayanlar Veri Setinde Eclat Algoritmasının Bellek Miktarının Karşılaştırılması



Şekil 47’de Eclat algoritmasının tüm platformlarda karar verilmiş destek ve güven değerleriyle kullanılan bellek miktarlarının ortalamaları alınarak analiz

edilmiştir. Grafiğe göre Eclat algoritmasının bellek miktarının en az olduğu program SPMF programıdır. SPMF programının bellek miktarının en az olduğu 0,05 destek değeridir. RStudio programının 0,01 ve 0,05 destek değerinde bellek miktarının en çok olduğu gözlemlenmiştir.

9.3. Deney Sonuçlarının Değerlendirilmesi

Bu çalışmada, Düzce Üniversitesi lisans ve lisansüstü öğrencileriyle akademik ve idari personellerinin kütüphaneden kitap ödünç alma kayıtlarındaki örüntülerin keşfedilmesi için Birliktelik Analizi algoritmalarından Apriori, FP-Growth ve Eclat algoritmalarının Python, R ve SPMF platformlarında çeşitli destek ve güven değerleriyle analizlerinin yapımı gerçekleştirilmiştir. Bu amaçla yapılan deneylere göre hem tekrar eden örüntüler/kurallar tespit edilmiş hem de Apriori, FP-Growth ve Eclat algoritmalarının her 3 veri madenciliği platformunda, Tüm veri seti, Erkekler veri seti ve Bayanlar veri seti üzerinde çalışma süreleri ve bellek miktarlarının karşılaştırmaları yapılmıştır. Şekil 30, Şekil 31, Şekil 32, Şekil 33, Şekil 34, Şekil 35, Şekil 36, Şekil 37 ve Şekil 38’de algoritmaların çalışma sürelerinin grafikleri; Şekil 39, Şekil 40, Şekil 41, Şekil 42, Şekil 43, Şekil 44, Şekil 45, Şekil 46 ve Şekil 47’de de algoritmaların deneyler sırasındaki bellek kullanım miktar grafikleri verilmiştir.

Şekil 30’daki grafiğe göre tüm veri setine Apriori algoritmasının SPMF programında tüm destek değerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduğu destek değerleri sıralaması 0,07, 0,05, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 48’deki grafiğe göre tüm veri setinin Apriori algoritmasının SPMF programında tüm destek değerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 48’deki grafiğe göre 0,01 destek değeri diğer tüm destek değerlerine göre daha az bellek işgal ettiği de ölçülmüştür. Şekil 30 ve Şekil 39 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 31’deki grafiğe göre tüm veri setine FP-Growth algoritmasının SPMF programında tüm destek değerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduğu destek değerleri sıralaması 0,05, 0,07, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 40’taki grafiğe göre tüm veri setinin FP-Growth algoritmasının SPMF programında tüm

destek deęerleri iin bellek kullanım miktarları bakımından en az bellek iřgal ederek iřlemleri tamamladıęı gzlemlenmiřtir. řekil 40'taki grafięe gre 0,01 destek deęeri dięer tm destek deęerlerine gre daha az bellek iřgal ettięi de llmřtr. řekil 31 ve řekil 40 beraber incelendięinde en verimli programın SPMF programı gzlemlenmiřtir.

řekil 32'deki grafięe gre tm veri setine Eclat algoritmasının SPMF programında tm destek deęerleri iin alıřma sreleri bakımından en hızlı alıřtıęı sylenebilir. SPMF programı alıřma sresi bakımından en verimli olduęu destek deęerleri sıralaması 0,05, 0,07, 0,02 ve 0,01 olarak gzlemlenmiřtir. řekil 41'deki grafięe gre tm veri setinin Eclat algoritmasının SPMF programında tm destek deęerleri iin bellek kullanım miktarları bakımından en az bellek iřgal ederek iřlemleri tamamladıęı gzlemlenmiřtir. řekil 50'deki grafięe gre 0,01 destek deęeri dięer tm destek deęerlerine gre daha az bellek iřgal ettięi de llmřtr. řekil 32 ve řekil 41'de bayanlar veri setine Eclat algoritmasının alıřma srelerinin ve bellek miktarlarının karřılařtırılması grafięinde Python programının 0,07 destek deęerinde kural vermemesi sonucunda alıřma sresi ve bellek miktarı 0 olarak ıkmıřtır. řekil 32 ve řekil 41 beraber incelendięinde en verimli programın SPMF programı gzlemlenmiřtir.

řekil 33'teki grafięe gre Erkekler veri setine Apriori algoritmasının SPMF programında tm destek deęerleri iin alıřma sreleri bakımından en hızlı alıřtıęı sylenebilir. SPMF programı alıřma sresi bakımından en verimli olduęu destek deęerleri sıralaması 0,07, 0,05, 0,02 ve 0,01 olarak gzlemlenmiřtir. řekil 42'deki grafięe gre Erkekler veri setinin Apriori algoritmasının SPMF programında tm destek deęerleri iin bellek kullanım miktarları bakımından en az bellek iřgal ederek iřlemleri tamamladıęı gzlemlenmiřtir. řekil 42'deki grafięe gre 0,05 destek deęeri dięer tm destek deęerlerine gre daha az bellek iřgal ettięi de llmřtr. řekil 33 ve řekil 42 beraber incelendięinde en verimli programın SPMF programı gzlemlenmiřtir.

řekil 34'deki grafięe gre Erkekler veri setine FP-Growth algoritmasının SPMF programında tm destek deęerleri iin alıřma sreleri bakımından en hızlı alıřtıęı sylenebilir. SPMF programı alıřma sresi bakımından en verimli olduęu destek deęerleri sıralaması 0,05, 0,07, 0,02 ve 0,01 olarak gzlemlenmiřtir. řekil

43'teki grafiğe göre Erkekler veri setinin FP-Growth algoritmasının SPMF programında tüm destek değerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 43'teki grafiğe göre 0,07 ve 0,01 destek değeri diğer tüm destek değerlerine göre daha az bellek işgal ettiği de ölçülmüştür. Şekil 34 ve Şekil 43 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 35'deki grafiğe göre Erkekler veri setine Eclat algoritmasının SPMF programında tüm destek değerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduğu destek değerleri sıralaması 0,07, 0,05, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 44'teki grafiğe göre Erkekler veri setinin Eclat algoritmasının SPMF programında tüm destek değerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 44'deki grafiğe göre 0,01 destek değeri diğer tüm destek değerlerine göre daha az bellek işgal ettiği de ölçülmüştür. Şekil 35 ve Şekil 44'te bayanlar veri setine Eclat algoritmasının çalışma sürelerinin ve bellek miktarlarının karşılaştırılması grafiğinde Python programının 0,07 destek değerinde kural vermemesi sonucunda çalışma süresi ve bellek miktarı 0 olarak çıkmıştır. Şekil 35 ve Şekil 44 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 36'daki grafiğe göre Bayanlar veri setine Apriori algoritmasının SPMF programında tüm destek değerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduğu destek değerleri sıralaması 0,07, 0,05, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 45'teki grafiğe göre Bayanlar veri setinin Apriori algoritmasının SPMF programında tüm destek değerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 45'teki grafiğe göre 0,02 ve 0,07 destek değeri diğer tüm destek değerlerine göre daha az bellek işgal ettiği de ölçülmüştür. Şekil 36 ve Şekil 45 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 37'deki grafiğe göre Bayanlar veri setine FP-Growth algoritmasının SPMF programında tüm destek değerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduğu

destek deęerleri sıralaması 0,05, 0,07, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 46'daki grafięe göre Bayanlar veri setinin FP-Growth algoritmasının SPMF programında tüm destek deęerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 46'daki grafięe göre 0,02 ve 0,05 destek deęeri dięer tüm destek deęerlerine göre daha az bellek işgal ettięi de ölçülmüştür. Şekil 37 ve Şekil 46 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 38'deki grafięe göre Bayanlar veri setine Eclat algoritmasının SPMF programında tüm destek deęerleri için çalışma süreleri bakımından en hızlı çalıştığı söylenebilir. SPMF programı çalışma süresi bakımından en verimli olduęu destek deęerleri sıralaması 0,07, 0,05, 0,02 ve 0,01 olarak gözlemlenmiştir. Şekil 47'deki grafięe göre Bayanlar veri setinin Eclat algoritmasının SPMF programında tüm destek deęerleri için bellek kullanım miktarları bakımından en az bellek işgal ederek işlemleri tamamladığı gözlemlenmiştir. Şekil 47'deki grafięe göre 0,05 destek deęeri dięer tüm destek deęerlerine göre daha az bellek işgal ettięi de ölçülmüştür. Şekil 38 ve Şekil 47'de bayanlar veri setine Eclat algoritmasının çalışma sürelerinin ve bellek miktarlarının karşılaştırılması grafięinde Python programının 0,07 destek deęerinde kural vermemesi sonucunda çalışma süresi ve bellek miktarı 0 olarak çıkmıştır. Şekil 38 ve Şekil 47 beraber incelendiğinde en verimli programın SPMF programı gözlemlenmiştir.

Şekil 30, Şekil 31, Şekil 32, Şekil 33, Şekil 34, Şekil 35, Şekil 36, Şekil 37 ve Şekil 38'deki Tüm veri seti, Erkekler veri seti ve Bayanlar veri setine Apriori, FP-Growth ve Eclat algoritmalarının Python, RStudio ve SPMF programlarının çalışma sürelerine bakıldığında bütün destek deęerlerinde SPMF programının en az çalışma süresi harcadığı gözlemlenmiştir.

Şekil 39, Şekil 40, Şekil 41, Şekil 42, Şekil 43, Şekil 44, Şekil 45, Şekil 46 ve Şekil 47'deki Tüm veri seti, Erkekler veri seti ve Bayanlar veri setine Apriori, FP-Growth ve Eclat algoritmalarının Python, RStudio ve SPMF programlarının bellek miktarlarına bakıldığında bütün destek deęerlerinde SPMF programının en az bellek harcadığı gözlemlenmiştir.

Aşağıdaki Şekil 48’te SPMF programında Apriori algoritmasına Destek: 0,01 ve Güven: 0,1 değerleri verilerek kural sonuçlarından bir kısmı verilmiştir.

Şekil 48. SPMF Apriori Algoritması Destek: 0,01, Güven: 0,1 Değeri için Kural Sonuçları

Alman Edebiyatı=t ==> Türk edebiyatı Roman=t	#SUP: 221 #CONF: 0.41154562383612664
İngiliz edebiyatı=t ==> Türk edebiyatı Roman=t	#SUP: 498 #CONF: 0.437609841827768
İngiliz edebiyatı=t ==> Türk Tarihi=t	#SUP: 185 #CONF: 0.1625659050966608
İngiliz edebiyatı=t ==> Amerikan Edebiyatı=t	#SUP: 199 #CONF: 0.17486818980667837
Bilgisayar Bilimi=t ==> Türk edebiyatı Roman=t	#SUP: 498 #CONF: 0.36034732272069464
Türk Tarihi=t ==> Türk edebiyatı Roman=t	#SUP: 1256 #CONF: 0.4052920296869958
Türk edebiyatı Roman=t ==> Türk Tarihi=t	#SUP: 1256 #CONF: 0.17475998330318632
Fizik=t ==> Türk edebiyatı Roman=t	#SUP: 517 #CONF: 0.36052998605299863
Fransız Edebiyatı=t ==> Türk edebiyatı Roman=t	#SUP: 306 #CONF: 0.43342776203966005
Psikoloji=t ==> Türk edebiyatı Roman=t	#SUP: 414 #CONF: 0.3812154696132597
Tanımsız Kategori=t ==> Türk edebiyatı Roman=t	#SUP: 182 #CONF: 0.4666666666666667
Matematik=t ==> Türk edebiyatı Roman=t	#SUP: 281 #CONF: 0.3575063613231552
İşletme yönetimi=t ==> Türk edebiyatı Roman=t	#SUP: 756 #CONF: 0.3361493997332148
Türk edebiyatı Roman=t ==> İşletme yönetimi=t	#SUP: 756 #CONF: 0.10518992625573953
Amerikan Edebiyatı=t ==> Türk edebiyatı Roman=t	#SUP: 930 #CONF: 0.4301572617946346
Türk edebiyatı Roman=t ==> Amerikan Edebiyatı=t	#SUP: 930 #CONF: 0.129400306108251

Kural sonuçlarına göre Alman Edebiyatı kategorisinde kitap alan lisans/lisansüstü, akademik ve idari personellerin Türk Edebiyatı Roman kategorisinde kitap alma olasılığı %41 olarak görülmektedir. Toplam birlikte alınan kitap kategori sayısı 221’dir.

2’inci kuralda ise İngiliz Edebiyatı kategorisinde kitap alan lisans/lisansüstü, akademik ve idari personellerin Türk Edebiyatı Roman kategorisinde kitap alma olasılığı %43 olarak görülmektedir. Toplam birlikte alınan kitap kategori sayısı 498’dir.

3’üncü kuralda ise İngiliz Edebiyatı kategorisinde kitap alan lisans/lisansüstü, akademik ve idari personellerin Türk Tarihi kategorisinde kitap alma olasılığı %16 olarak görülmektedir. Toplam birlikte alınan kitap kategori sayısı 185’tir.

10. SONUÇLAR VE ÖNERİLER

Birliktelik Kuralları, veri madenciliğinin en popüler yöntemlerinden biridir. Büyük veri tabanlarındaki verilerin birbiriyle ilişkilerinin büyüklüğünü bulmak için kullanılan bir modeldir. Birliktelik Analizi, bir veri kümesindeki sıklıkla tekrar eden öğelerin saptanması ve veri kümesindeki öğeler arasındaki bağlantılar/ilişkilerin tespit edilmesi için kullanılmaktadır. Bu tez çalışmasında Düzce Üniversitesi Kütüphanesi Kitap Ödünç Alma kayıtları içinde tekrar eden kuralların tespit edilmesine yönelik bir Birliktelik Analizi gerçekleştirilmiştir. Bu amaçla üniversite kütüphanesinden alınan kayıt veri kümesi üzerinde popüler Birliktelik Kuralları algoritmalarından olan Apriori, FP-Growth ve Eclat algoritmaları çok tercih edilen veri madenciliği platformlarından olan Python, RStudio ve SPMF ortamlarında Tablo 7’de listelenmiş olan destek ve güven değerleri verilerek analiz edilmiş ve bu deneyler sırasında programların çalışma süreleri ve bellek kullanım miktarları bakımından performans analizleri gerçekleştirilmiştir. Bu çalışmalar sırasında her bir kitap seviyesinde aralarında tekrar eden beraber ödünç alma kuralları görülememiştir. Bu sebeple iki veya daha fazla kitabın ödünç alındığı kayıtlardaki kitap kategorileri üzerinden deneylere devam edilerek kategoriler arasındaki beraber ödünç almaya ilişkin tekrar eden örüntüler araştırılmış ve çeşitli kurallar elde edilmiştir. Bu analizlerin yanı sıra kitap ödünç alma davranışlarının cinsiyete göre değişimi de incelenmiştir.

Yapılan deney sonuçları incelendiğinde, SPMF programının bütün algoritmaların tüm destek ve güven değerleri için yapılmış deneylerinde çalışma süreleri açısından en hızlı olduğu görülmüştür. SPMF programı Apriori, FP-Growth ve Eclat algoritmalarında erkekler veri setinde en hızlı çalışma sürelerini vermiştir. Python ile R’nin sıralaması ise kullanılan algoritmanın destek ve güven değerlerine göre zaman zaman yer değiştirmektedir.

Deney sonuçları, bellek kullanım miktarları açısından değerlendirildiğinde SPMF programının Apriori, Fp-Growth ve Eclat algoritmalarındaki tüm destek ve güven değerleri için en az bellek kullanımı yaptığı görülmüştür. SPMF programının Apriori algoritmasıyla Bayanlar veri seti, FP-Growth algoritmasıyla Tüm veri seti ve

Eclat algoritmasıyla yine bayanlar veri seti üzerinde en az bellek kullanımı yaptığı görülmüştür.

Sonuç olarak SPMF programı Apriori, FP-Growth ve Eclat algoritmalarında çalışma süreleri bakımından 0,05 destek değeri için en verimli programdır. SPMF programının tüm veri seti üzerindeki tüm algoritmaların bellek miktarı bakımından 0,01 destek değeri için, erkekler veri seti üzerindeki tüm algoritmaların bellek miktarı bakımından 0,01 destek değeri için ve bayanlar veri seti üzerindeki tüm algoritmaların bellek miktarı bakımından 0,02 destek değeri için en verimli program olduğu görülmüştür.

Deney sonuçlarına göre cinsiyetler arasında çok belirgin kitap ödünç alma farklılıkları görülmemiştir. Fakat kitap kategorileri arasında sıklıkla tekrar eden kurallar bulunmuştur. Örneğin şekil 48’de görüldüğü üzere Alman Edebiyatı veya İngiliz Edebiyatı kitapları alanların sıklıkla Türk Edebiyatı Roman kategorisinden de kitaplar ödünç aldığı veya Bilgisayar Bilimi kategorisinden kitap ödünç alanların sıklıkla Türk Edebiyatı Roman kategorisinden kitaplar ödünç aldıkları görülmüştür. Kütüphanedeki kitapların dizilişlerine karar verilmesi aşamasında ilgili görevlilerin bu gibi kuralları dikkate almasının faydalı olabileceği düşünülmektedir.

Sonraki çalışmalarda birliktelik kuralları ile sıklıkla tekrar eden öğelerin saptanması ve veri kümesindeki öğeler arasındaki bağlantılar/ilişkilerin tespit edilmesinde, çalışma süresi ve bellek miktarları bakımında verimli olduğu için SPMF programının kullanılması tercih olacaktır.

11. KAYNAKÇA

- Agrawal, R. and Srikant, R. (1993). Fast Algorithms for Mining Association Rules, 20. VLDB Conference, s.487-499
- Akpınar, H. (2000). Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, c.29, s.1-22, Nisan
- Alaybey, F. (2022, 10 Temmuz) Veri Madenciliği Giriş Yöntemleri ve Metodolojileri. Erişim: <https://medium.com/@furkanalaybeg/veri-madenciligi-ve-yontemleri-d0e2fd238e44>.
- Albayrak, M., (2008). *EEG sinyallerindeki epileptiform aktivitenin veri madenciliği süreci ile tespiti*, Doktora Tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya, 56–70
- Alpaydın, E. (2000). Zeki Veri Madenciliği: Ham Veriden Altın Bilgiye Ulaşma Yöntemleri, *Bilişim 2000 Eğitim Semineri*
- Anonim, (2022. 12 Şubat). <https://www.philippe-fournier-viger.com/spmf/Applications>. Singapore: World Scientific
- Arslan, E. (2022). Veri Madenciliği Yöntemleri: <https://emraharslanbm.wordpress.com/2012/07/17/veri-madenciligi-yontemleri/>.
- Aydoğan, F. (2003). *E-Ticarette Veri Madenciliği yaklaşımlarıyla Müşteriye Hizmet Sunan Akıllı Modüllerin Tasarımı ve Gerçekleştirimi*. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- Baykal, A. (2006). Veri Madenciliği Uygulama Alanları, *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, c. 107, sayı 7, ss. 95-107
- Baykasoğlu, A. (2005). *Veri Madenciliği ve Çimento Sektöründe Bir Uygulama*, Endüstri Mühendisliği Bölümü, Gaziantep Üniversitesi, Gaziantep
- Bigus, J. P. (1996). Data Mining With Neural Networks: Solving Business Problems from Application Development to Decision Support, McGraw Hill

- Costache, C., Dumitrascu, D. D., & Maniu, I. (2021). Facilitators of and barriers to sustainable development in small and medium-sized enterprises: A descriptive exploratory study in Romania. *Sustainability (Switzerland)*, 13(6).
- Coşlu, E. (2013). Veri Madenciliği, Akademik Bilişim – XV. *Akademik Bilişim Konferansı Bildirileri 23-25 Ocak- Akdeniz Üniversitesi, Antalya*, s. 615-619
- Çelik, B. (2019). *Mekanik arızaların veri madenciliği Apriori algoritması ile analiz edilmesi*, Yüksek Lisans Tezi, Endüstri Mühendisliği Bölümü, Kocaeli Üniversitesi, Kocaeli
- Çoban, Ö. Aksakallı, I. Günay, F. (2015). Acil Servis Verilerinden Birliktelik Kuralı (Apriori) Yöntemi ile Hasta Profili Çıkarımı, 2. *Ulusal Yönetim Bilişim Sistemleri Kongresi*, Erzurum
- Dunham, M. H. Xiao, Y., Gruenwald L. Ve Hossain, Z., (2000), A Survey of association Rules, Teknik Rapor, Southern Methodist University, *Department of Computer Science*
- Fayyad, U, M. Piatetsky-Shapiro, G., Smyth, (1996). *Knowledge discovery and data mining: Towards a unifying framework. in KDD*, vol. 96, 1996, pp. 82–88.
- Girotra, M. Nagpal, K. Inocha, S. Sharma, N. (2013) *Comparative Survey on Association Rule Mining Algorithms*, International Journal of Computer Applications (0975 – 8887)
- Gupta, N. Mangal, N. Tiwari, K ve Mitra, P. (2006). *Mining quantitative association rules in protein sequences*. Data Mining: Theory, Methodology, Texhniques, and Applications, c.3755, ss. 273-281.
- Gürgen, G. (2008). *Birliktelik Kuralları ile Sepet Analizi Uygulaması*, Yüksek Lisans Tezi, Sosyal Bilimler Enstitüsü, Marmara Üniversitesi, İstanbul
- Han, E. Karypis, G. and Kumar, V. (1997). *Scalable Parallel Data Mining For Association Rules*, Proceedings of the ACM SIGMOD Conference, pp. 277-288

- Han, E. Karypis, G. and Kumar, V. (2000). *Scalable paralel data mining for association rules*. IEEE Transactions on Knowledge and Data Engineering, 12(3):337–352
- Han, J. & Kamber, M. & Pei, J. (2012). *Data mining: Concepts and Techniques*. Morgan Kaufmann Publishers 3th ed., Waltham, 5-8.
- Han, J. Pei, J. ve Yin, Y. (2000). *Mining Frequent Patterns Without Candidate Generation*, Proc. Of the 2000 ACM SIGMOD International Conference on Management of Data, Dallas, Texas, United States, ss. 1-12.
- Hidber, C. (1999). *Online Association Rule Mining*, SIGMOD 1999, Proceedings ACM SIGMOD International Conference on Management of Data, June 1-3, Philadelphia, Pennsylvania, pp.145-156.
- Hudairy, H. (2004). *Data mining and decision making support in the governmental sector*, Master Thesis, Faculty of Graduate School of The University of Louisville, Kentucky, 1-5.
- İnan, O. (2003). *Veri madenciliği*, Yüksek Lisans Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Kantardzic, M. (2011). *Data Minig: Consepts, Models, Methods, and Algorithms*, John Wiley & Sons
- Karademir, M. (2021). *Veri madenciliği market sepeti analizi: Apriori ve FP-Growth algoritmaları karşılaştırılması*, Yüksek Lisans Tezi, Bilgisayar Mühendisliği Bölümü, İstanbul Medeniyet Üniversitesi, İstanbul.
- Karaibrahimoğlu, K. (2014). *Veri Madenciliğinden Birliktelik Kuralı ile Onkoloji Verilerinin Analiz Edilmesi: Meram Tıp Fakültesi Onkoloji Örneği*, Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- Kaur, M. ve Kang, S. (2016). *Market Basket Analysis: Identify the changing trends of market data*.
- Kdnuggets, Erişim: 4 Mart (2022). <https://www.kdnuggets.com/2020/06/data-science-tools-popularity-animated.html>

- Kim, K. (2002), *A model of digital library information seeking process as a frame for classifying usability problems*, dissertation for the degree of philosophy doctor, The State University of New Jersey, New Brunswick.
- Krause, F. Lindemann, O. (2014). *Expyriment: a Python library for cognitive and neuroscientific experiments*, Behav Res Methods
- Krisnanto, U., Juharsah, J., Putra, P., Achmad, A. D., & Timotius, E. (2022). *Utilizing Apriori Data Mining Techniques on Sales Transactions*. *Webology*, 19(1), 5581–5590.
- Li, R. (2016). History of data mining. Retrieved from <https://www.linkedin.com/pulse/history-data-mining-gregory-piatetsky-shapiro>
- Liu, S., Wang, L., & Huang, W. W. (2017). Effects of process and outcome controls on business process outsourcing performance: *Moderating roles of vendor and client capability risks*. *European Journal of Operational Research*, 260(3), 1115–1128.
- Malerba, D., Esposito, F., Lanza, A., ve Lisi, (2001). *First-order rule induction for the recognition of morphological patterns in topographic maps*, Berlin, Germany.
- Mannila, H. Toivonen, H. and Verkamo, A. I. (1994) *Efficient Algorithms for Discovering Association Rules*, *Proceedings of the AAAI Workshop on Knowledge Discovery in Databases (KDD-94)*, pp. 181-192
- Morais, A. Peixoto, H. Coimbra, C. Abelha, A. Machado, J. (2017). *Predicting the need of Neonatal Resuscitation using Data Mining*. *Procedia Computer Science*, c.113, ss.571-576
- Nair, M. (2019). *Birliktelik kuralları algoritmalarının otomotiv sektörü verileri üzerinde SPMF ve Weka ile performans analizi*, Yüksek Lisans tezi, Bilgisayar Mühendisliği Bölümü, Düzce Üniversitesi, Düzce.
- Olson, D. Delen, D. (2008). *Advanced Data Mining Techniques*, ss. 11-53

Özçalıcı, M. (2017). Veri Madenciliğinde Birliktelik Kuralları ve İkinci El Otomobil Piyasası Üzerine Bir Uygulama, *Ordu Üniversitesi Sosyal Bilimler Araştırmaları Dergisi*, 7(1), 45-58.

Özdemir, A., Yalçın, F. A., Çam, H. (2009). Veri Tabanında Bilgi Keşfi Süreci, *Gümüşhane Devlet Hastanesi, Sosyal Ekonomik Araştırmalar Dergisi*. 1(20), 347-366

Performance Comparison of Association Rule Algorithms with SPMF on Automotive Industry Data - Scientific Figure on ResearchGate.Erişim: 4 Mart 2022.
https://www.researchgate.net/figure/Pseudo-code-for-Eclat-algorithm_fig3_334825343

R Core Team, (2021). R Language Definition, <https://cran.r-project.org/doc/manuals/r-release/R-lang.html>

Rokachi, L., Maimon, O. (2008) Data Mining With Decision Trees Theory

Said, I. U. Adam, A. H ve Garko. A. B. (2015). Association rule mining on medical data to predict heart disease, *International journal of science technology and management*

Sarman, G. (2011). Veri Madenciliğinde Kümeleme Teknikleri Üzerine Bir Çalışma: K-Means ve K-Medoids Kümeleme Algoritmalarının Karşılaştırılması,” *Süleyman Demiral Üniversitesi, Fen Bilimleri Enstitüsü Dergisi*, c. 15, sayı 3, ss. 192-202

Serban, G. Campan, A. Ve Czibula, I. G (2006). A programming interface for finding relational association rules. *Journal of Computers Communications & Control*, C.1, ss,439-444

Shearer, C. (2000). The Crisp-DM Model: The new blueprint for data mining. *Journal of Data Warehousing*. c. 5. s. 4. Ss.13-23

Silahtaroglu, G. (2016). *Veri Madenciliği: Kavram ve Algoritmaları*. Papatya Bilim, İstanbul, 9-11

- Söylemez, İ. Doğan, A. Özcan, U. (2016). Trafik Kazalarında Birliktelik Kuralı Analizi: Ankara İli Örneği, *Ege Akademik Bakış*, 16, 11-20.
- Srikant, R. (1996) Fast Algorithms for Mining Association Rules and Sequential Patterns, *Ph.D Dissertation, University of Wisconsin, Madison*.
- Şen, F. (2008). *Veri Madenciliği ile Birliktelik Kurallarının Bulunması*, Yüksek Lisans Tezi, Bilgisayar Mühendisliği Bölümü, Sakarya Üniversitesi, Sakarya, Türkiye
- Taş, Y. (2018). Birliktelik kuralları madenciliği ve bir uygulama”, Yüksek Lisans Tezi, Yönetim Bilişim Sistemleri Bölümü, Cumhuriyet Üniversitesi, Sivas, Türkiye
- Tokyürek, E. (2019). *Birliktelik kural çıkarım algoritmaları kullanılarak market sepet analizi*, Yüksek lisans tezi, Bilgisayar Mühendisliği, Fen Bilimleri Enstitüsü, Şeyh Edebalı Üniversitesi, Bilecik.
- Uluyardımcı, M. M. Ve Zontul, M. (2019). Veri madenciliği yöntemleri ile uçuş biletleme analizi, *Aurum mühendislik sistemleri ve mimarlık dergisi*, c. 3, s. 2, ss. 153-168
- Veli, A. (2020). *Farklı Sektörlere Ait Büyük Verinin Birliktelik Kuralları Yaklaşımıyla İncelenmesi*, Yüksek Lisans Tezi, Elektrik-Elektronik ve Bilgisayar Mühendisliği Bölümü, Düzce Üniversitesi, Düzce, Türkiye
- Vm modelleri: [https://www.researchgate.net/publication/318043710_A_bio-inspired_approach_for_Association_Rules_Mining?channel=doi&linkId=5956d0070f7e9ba95e0caef3&showFulltext=true]
- Zawaidah, F. H. Jbara, Y. H. Zanona, M. (2011). *An improved algorithm for mining association rules in large databases*, vol. 1, no. 7, pp. 311-316
- Zerman, M. (2018). *Birliktelik kuralları algoritmaları ile büyük veriler üzerinde analitik analizler- havaalanı örneği*, Yüksek lisans tezi, Bilgisayar Mühendisliği, Fen Bilimleri Enstitüsü, Haliç Üniversitesi, İstanbul, Türkiye
- Zhao, O. Bhowmick, S.S. (2003). *Association Rule Mining: A Survey*, Technical Report, AIS, Nanyang Technological University, Singapore

Webb, G. I. (2003). *Association Rules*, The Handbook Of Data Mining, s. 26-39



12. EKLER

12.1. Python Apriori Destek: 0.01 Güven. 0.1 Tüm Veri Seti

Antecedents	consequents	antecedent support	consequent support	support
confidence	lift	leverage	conviction	
0	(Alman Edebiyatı)	(Türk edebiyatı Roman)		0.039228
0.511922	0.016992	0.433155	0.846135 -0.003090	0.861043
1	(Fizik)	(Amerikan Edebiyatı)		0.101951
0.011048	0.108368	0.687257	-0.005028	0.944693
2	(Türk Tarihi)	(Amerikan Edebiyatı)		0.222362
0.023285	0.104717	0.664105	-0.011777	0.940841
3	(Amerikan Edebiyatı)	(Türk Tarihi)		0.157681
0.023285	0.147672	0.664105	-0.011777	0.912369
4	(Türk edebiyatı Roman)	(Amerikan Edebiyatı)		0.511922
0.157681	0.070764	0.138232	0.876658 -0.009956	0.977432
5	(Amerikan Edebiyatı)	(Türk edebiyatı Roman)		0.157681
0.511922	0.070764	0.448780	0.876658 -0.009956	0.885451
6	(Tıp)	(Amerikan Edebiyatı)		0.087267
0.010209	0.116987	0.741922	-0.003551	0.953914
7	(İngiliz edebiyatı)	(Amerikan Edebiyatı)		0.080694
0.157681	0.015104	0.187175	1.187047	0.002380
8	(İşletme yönetimi)	(Amerikan Edebiyatı)		0.160129
0.157681	0.017062	0.106550	0.675732 -0.008188	0.942771
9	(Amerikan Edebiyatı)	(İşletme yönetimi)		0.157681
0.160129	0.017062	0.108204	0.675732 -0.008188	0.941775
10	(Araştırma yöntemleri)	(Türk edebiyatı Roman)		0.029578
0.511922	0.012097	0.408983	0.798917 -0.003045	0.825828
11	(Fizik)	(Bilgisayar Bilimi)		0.101951
0.012097	0.118656	1.209476	0.002095	1.023317
12	(Bilgisayar Bilimi)	(Fizik)		0.098105
0.012097	0.123307	1.209476	0.002095	1.024360
13	(Bilgisayar Bilimi)	(Türk Tarihi)		0.098105
0.017761	0.181041	0.814170	-0.004054	0.949544
14	(Bilgisayar Bilimi)	(Türk edebiyatı Roman)		0.098105
0.511922	0.036781	0.374911	0.732359 -0.013442	0.780813
15	(Bilgisayar Bilimi)	(İşletme yönetimi)		0.098105
0.013496	0.137562	0.859074	-0.002214	0.973834
16	(Din)	(Türk Tarihi)		0.051675
0.011678	0.225981	1.016275	0.000187	1.004676
17	(Din)	(Türk edebiyatı Roman)		0.051675
0.022935	0.443843	0.867013	-0.003518	0.877590
18	(Ekonomi)	(Türk edebiyatı Roman)		0.037200
0.511922	0.013076	0.351504	0.686635 -0.005968	0.752630
19	(Felsefe)	(Türk Tarihi)		0.061814
0.012517	0.202489	0.910626	-0.001228	0.975081
20	(Felsefe)	(Türk edebiyatı Roman)		0.061814
0.024334	0.393665	0.768994	-0.007310	0.804964
21	(Fizik)	(Türk Tarihi)		0.101951
0.018530	0.181756	0.817387	-0.004140	0.950374

22	(Fizik) (Türk edebiyatı Roman)	0.101951	0.511922
0.037760	0.370370 0.723490 -0.014431 0.775182		
23	(Fizik) (İşletme yönetimi)	0.101951	0.160129
0.013356	0.131001 0.818101 -0.002970 0.966482		
24	(Fransız Edebiyatı) (Türk edebiyatı Roman)		0.051535
0.511922	0.023355 0.453189 0.885268 -0.003027 0.892589		
25	(Kişisel gelişim) (Türk edebiyatı Roman)		0.041396
0.511922	0.017272 0.417230 0.815026 -0.003920 0.837513		
26	(Matematik) (Türk edebiyatı Roman)		0.056639
0.511922	0.021397 0.377778 0.737959 -0.007598 0.784411		
27	(Mühendislik) (Türk edebiyatı Roman)		0.036081
0.511922	0.012447 0.344961 0.673855 -0.006024 0.745113		
28	(Psikoloji) (Türk Tarihi)	0.077687	0.222362
0.013705	0.176418 0.793380 -0.003569 0.944214		
29	(Psikoloji) (Türk edebiyatı Roman)	0.077687	0.511922
0.030907	0.397840 0.777149 -0.008863 0.810545		
30	(Rus edebiyatı) (Türk edebiyatı Roman)		0.044962
0.511922	0.020069 0.446345 0.871900 -0.002948 0.881556		
31	(Sanat Tarihi) (Türk edebiyatı Roman)		0.024963
0.511922	0.010279 0.411765 0.804350 -0.002500 0.829732		
32	(Tanımsız Kategori) (Türk edebiyatı Roman)		0.028040
0.511922	0.013705 0.488778 0.954790 -0.000649 0.954728		
33	(Toplum bilimleri) (Türk edebiyatı Roman)		0.027620
0.511922	0.011188 0.405063 0.791259 -0.002951 0.820386		
34	(Türk Tarihi) (Türk edebiyatı Roman)		0.222362
0.511922	0.093700 0.421384 0.823140 -0.020132 0.843526		
35	(Türk edebiyatı Roman) (Türk Tarihi)		0.511922
0.222362	0.093700 0.183035 0.823140 -0.020132 0.951862		
36	(Tıp) (Türk Tarihi)	0.087267	0.222362
0.012866	0.147436 0.663044 -0.006539 0.912117		
37	(İngiliz edebiyatı) (Türk Tarihi)	0.080694	0.222362
0.013565	0.168111 0.756023 -0.004378 0.934786		
38	(Türk Tarihi) (İşletme yönetimi)	0.222362	0.160129
0.031956	0.143711 0.897470 -0.003651 0.980827		
39	(İşletme yönetimi) (Türk Tarihi)	0.160129	0.222362
0.031956	0.199563 0.897470 -0.003651 0.971517		
40	(Tıp) (Türk edebiyatı Roman)	0.087267	0.511922
0.035522	0.407051 0.795143 -0.009152 0.823137		
41	(İngiliz edebiyatı) (Türk edebiyatı Roman)		0.080694
0.511922	0.036501 0.452340 0.883610 -0.004808 0.891205		
42	(Türk edebiyatı Roman) (İşletme yönetimi)		0.511922
0.160129	0.056150 0.109684 0.684977 -0.025824 0.943341		
43	(İşletme yönetimi) (Türk edebiyatı Roman)		0.160129
0.511922	0.056150 0.350655 0.684977 -0.025824 0.751646		
44	(Türk Tarihi, İşletme yönetimi) (Türk edebiyatı Roman)		0.031956
0.511922	0.010629 0.332604 0.649716 -0.005730 0.731317		
45	(Türk Tarihi, Türk edebiyatı Roman) (İşletme yönetimi)		0.093700
0.160129	0.010629 0.113433 0.708386 -0.004375 0.947330		
46	(Türk edebiyatı Roman, İşletme yönetimi) (Türk Tarihi)		0.056150
0.222362	0.010629 0.189290 0.851270 -0.001857 0.959206		

12.2. Python FP-Growth Destek: 0.01 Güven. 0.1 Tüm Veri Seti

[{ 'İtalyan edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.5657894736842105], [{ 'Portekiz edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.4583333333333333], [{ 'Mimari - Mimarlık' }, { 'Türk edebiyatı Roman' }, 0.37668161434977576], [{ 'Kimya' }, { 'Türk edebiyatı Roman' }, 0.366666666666666664], [{ 'Tanımsız Kategori' }, { 'Türk edebiyatı Roman' }, 0.48877805486284287], [{ 'Sanat Tarihi' }, { 'Türk edebiyatı Roman' }, 0.4117647058823529], [{ 'Toplum bilimleri' }, { 'Türk edebiyatı Roman' }, 0.4050632911392405], [{ 'Araştırma yöntemleri' }, { 'Türk edebiyatı Roman' }, 0.408983451536643], [{ 'Ekonomi' }, { 'İşletme yönetimi' }, 0.2575187969924812], [{ 'Ekonomi' }, { 'Türk Tarihi' }, 0.2424812030075188], [{ 'Ekonomi' }, { 'Türk edebiyatı Roman' }, 0.35150375939849626], [{ 'Mühendislik' }, { 'Türk edebiyatı Roman' }, 0.3449612403100775], [{ 'Alman Edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.43315508021390375], [{ 'Kişisel gelişim' }, { 'Türk edebiyatı Roman' }, 0.4172297297297297], [{ 'Rus edebiyatı' }, { 'Türk Tarihi' }, 0.17729393468118196], [{ 'Rus edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.4463452566096423], [{ 'Rus edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.4463452566096423], [{ 'Fransız Edebiyatı' }, { 'Amerikan Edebiyatı' }, 0.1926729986431479], [{ 'Fransız Edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.4531886024423338], [{ 'Fransız Edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.4531886024423338], [{ 'Din' }, { 'Türk Tarihi' }, 0.2259810554803789], [{ 'Din' }, { 'Türk Tarihi' }, { 'Türk edebiyatı Roman' }, 0.41317365269461076], [{ 'Türk edebiyatı Roman' }, { 'Din' }, { 'Türk Tarihi' }, 0.21036585365853658], [{ 'Din' }, { 'Türk edebiyatı Roman' }, 0.4438430311231394], [{ 'Din' }, { 'Türk edebiyatı Roman' }, 0.4438430311231394], [{ 'Felsefe' }, { 'Amerikan Edebiyatı' }, 0.12782805429864252], [{ 'Felsefe' }, { 'İşletme yönetimi' }, 0.1346153846153846], [{ 'Felsefe' }, { 'Türk Tarihi' }, 0.20248868778280543], [{ 'Türk edebiyatı Roman' }, { 'Felsefe' }, { 'Türk Tarihi' }, 0.1752873563218391], [{ 'Felsefe' }, { 'Türk Tarihi' }, { 'Türk edebiyatı Roman' }, 0.3407821229050279], [{ 'Felsefe' }, { 'Türk edebiyatı Roman' }, 0.3936651583710407], [{ 'Felsefe' }, { 'Türk edebiyatı Roman' }, 0.3936651583710407], [{ 'Matematik' }, { 'Bilgisayar Bilimi' }, 0.14444444444444443], [{ 'Matematik' }, { 'Türk Tarihi' }, 0.15432098765432098], [{ 'Matematik' }, { 'Türk edebiyatı Roman' }, 0.37777777777777777], [{ 'Matematik' }, { 'Türk edebiyatı Roman' }, 0.37777777777777777], [{ 'Psikoloji' }, { 'Amerikan Edebiyatı' }, 0.12061206120612061], [{ 'Psikoloji' }, { 'Türk edebiyatı Roman' }, { 'Amerikan Edebiyatı' }, 0.15610859728506787], [{ 'Psikoloji' }, { 'Amerikan Edebiyatı' }, { 'Türk edebiyatı Roman' }, 0.5149253731343284], [{ 'Psikoloji' }, { 'İşletme yönetimi' }, 0.1269126912691269], [{ 'Psikoloji' }, { 'Türk Tarihi' }, 0.17641764176417643], [{ 'Psikoloji' }, { 'Türk edebiyatı Roman' }, { 'Türk Tarihi' }, 0.15384615384615385], [{ 'Psikoloji' }, { 'Türk Tarihi' }, { 'Türk edebiyatı Roman' }, 0.3469387755102041], [{ 'Psikoloji' }, { 'Türk edebiyatı Roman' }, 0.39783978397839787], [{ 'Psikoloji' }, { 'Türk edebiyatı Roman' }, 0.39783978397839787], [{ 'Psikoloji' }, { 'Türk edebiyatı Roman' }, 0.39783978397839787], [{ 'İngiliz edebiyatı' }, { 'İşletme yönetimi' },

0.1048526863084922], [{}İngiliz edebiyatı'}, {}Türk Tarihi'}, 0.1681109185441941],
 [{}Türk edebiyatı Roman', {}İngiliz edebiyatı'}, {}Türk Tarihi'},
 0.16666666666666666], [{}İngiliz edebiyatı', {}Türk Tarihi'}, {}Türk edebiyatı
 Roman'}, 0.4484536082474227], [{}İngiliz edebiyatı'}, {}Amerikan Edebiyatı'},
 0.18717504332755633], [{}Türk edebiyatı Roman', {}İngiliz edebiyatı'}, {}Amerikan
 Edebiyatı'}, 0.15708812260536398], [{}İngiliz edebiyatı', {}Amerikan Edebiyatı'},
 {}Türk edebiyatı Roman'}, 0.37962962962962965], [{}İngiliz edebiyatı'}, {}Türk
 edebiyatı Roman'}, 0.45233968804159447], [{}İngiliz edebiyatı'}, {}Türk edebiyatı
 Roman'}, 0.45233968804159447], [{}İngiliz edebiyatı'}, {}Türk edebiyatı Roman'},
 0.45233968804159447], [{}Tıp'}, {}Amerikan Edebiyatı'}, 0.11698717948717949],
 [{}Türk edebiyatı Roman', {}Tıp'}, {}Amerikan Edebiyatı'}, 0.11614173228346457],
 [{}Amerikan Edebiyatı', {}Tıp'}, {}Türk edebiyatı Roman'}, 0.4041095890410959],
 [{}Tıp'}, {}Türk Tarihi'}, 0.14743589743589744], [{}Türk edebiyatı Roman', {}Tıp'},
 {}Türk Tarihi'}, 0.13582677165354332], [{}Tıp', {}Türk Tarihi'}, {}Türk edebiyatı
 Roman'}, 0.375], [{}Tıp'}, {}Türk edebiyatı Roman'}, 0.40705128205128205],
 [{}Tıp'}, {}Türk edebiyatı Roman'}, 0.40705128205128205], [{}Tıp'}, {}Türk edebiyatı
 Roman'}, 0.40705128205128205], [{}Tıp'}, {}Türk edebiyatı Roman'},
 0.40705128205128205], [{}Bilgisayar Bilimi'}, {}Amerikan Edebiyatı'},
 0.10192444761225944], [{}Fizik'}, {}Bilgisayar Bilimi'}, 0.11865569272976681],
 [{}Bilgisayar Bilimi'}, {}Fizik'}, 0.1233071988595866], [{}Bilgisayar Bilimi'},
 {}İşletme yönetimi'}, 0.1375623663578047], [{}Bilgisayar Bilimi'}, {}Türk Tarihi'},
 0.1810406272273699], [{}Türk edebiyatı Roman', {}Bilgisayar Bilimi'}, {}Türk
 Tarihi'}, 0.16730038022813687], [{}Bilgisayar Bilimi', {}Türk Tarihi'}, {}Türk
 edebiyatı Roman'}, 0.3464566929133858], [{}Bilgisayar Bilimi'}, {}Türk edebiyatı
 Roman'}, 0.37491090520313614], [{}Bilgisayar Bilimi'}, {}Türk edebiyatı Roman'},
 0.37491090520313614], [{}Fizik'}, {}Amerikan Edebiyatı'}, 0.1083676268861454],
 [{}Fizik'}, {}İşletme yönetimi'}, 0.1310013717421125], [{}Fizik'}, {}Türk Tarihi'},
 0.18175582990397804], [{}Fizik', {}Türk edebiyatı Roman'}, {}Türk Tarihi'},
 0.17407407407407408], [{}Fizik', {}Türk Tarihi'}, {}Türk edebiyatı Roman'},
 0.35471698113207545], [{}Fizik'}, {}Türk edebiyatı Roman'},
 0.37037037037037035], [{}Fizik'}, {}Türk edebiyatı Roman'},
 0.37037037037037035], [{}Fizik'}, {}Türk edebiyatı Roman'},
 0.37037037037037035], [{}İşletme yönetimi'}, {}Amerikan Edebiyatı'},
 0.10655021834061135], [{}Amerikan Edebiyatı'}, {}İşletme yönetimi'},

0.1082039911308204], [{'İşletme yönetimi', 'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.11581569115815692], [{'İşletme yönetimi', 'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.38114754098360654], [{'İşletme yönetimi', 'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.11581569115815692], [{'İşletme yönetimi', 'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.38114754098360654], [{'Amerikan Edebiyatı'}, {'Türk Tarihi'}, 0.1476718403547672], [{'Türk Tarihi'}, {'Amerikan Edebiyatı'}, 0.10471698113207548], [{'Türk edebiyatı Roman', 'Amerikan Edebiyatı'}, {'Türk Tarihi'}, 0.1383399209486166], [{'Türk edebiyatı Roman', 'Türk Tarihi'}, {'Amerikan Edebiyatı'}, 0.1044776119402985], [{'Amerikan Edebiyatı', 'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.4204204204204204], [{'Türk edebiyatı Roman', 'Amerikan Edebiyatı'}, {'Türk Tarihi'}, 0.1383399209486166], [{'Türk edebiyatı Roman', 'Türk Tarihi'}, {'Amerikan Edebiyatı'}, 0.1044776119402985], [{'Amerikan Edebiyatı', 'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.4204204204204204], [{'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.1382324819013796], [{'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.44878048780487806], [{'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.1382324819013796], [{'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.44878048780487806], [{'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.1382324819013796], [{'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.44878048780487806], [{'Türk edebiyatı Roman'}, {'Amerikan Edebiyatı'}, 0.1382324819013796], [{'Amerikan Edebiyatı'}, {'Türk edebiyatı Roman'}, 0.44878048780487806], [{'İşletme yönetimi'}, {'Türk Tarihi'}, 0.19956331877729258], [{'Türk Tarihi'}, {'İşletme yönetimi'}, 0.14371069182389937], [{'İşletme yönetimi'}, {'Türk Tarihi'}, 0.19956331877729258], [{'Türk Tarihi'}, {'İşletme yönetimi'}, 0.14371069182389937], [{'İşletme yönetimi'}, {'Türk Tarihi'}, 0.19956331877729258], [{'Türk Tarihi'}, {'İşletme yönetimi'}, 0.14371069182389937], [{'İşletme yönetimi', 'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.18929016189290163], [{'İşletme yönetimi', 'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.33260393873085337], [{'Türk edebiyatı Roman', 'Türk Tarihi'}, {'İşletme yönetimi'}, 0.11343283582089553], [{'İşletme yönetimi', 'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.18929016189290163], [{'İşletme yönetimi', 'Türk Tarihi'}, {'Türk

edebiyatı Roman'}, 0.33260393873085337], [{'Türk edebiyatı Roman', 'Türk Tarihi'}, {'İşletme yönetimi'}, 0.11343283582089553], [{'İşletme yönetimi', 'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.18929016189290163], [{'İşletme yönetimi', 'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.33260393873085337], [{'Türk edebiyatı Roman', 'Türk Tarihi'}, {'İşletme yönetimi'}, 0.11343283582089553], [{'İşletme yönetimi'}, {'Türk edebiyatı Roman'}, 0.3506550218340611], [{'Türk edebiyatı Roman'}, {'İşletme yönetimi'}, 0.10968446933479033], [{'İşletme yönetimi'}, {'Türk edebiyatı Roman'}, 0.3506550218340611], [{'Türk edebiyatı Roman'}, {'İşletme yönetimi'}, 0.10968446933479033], [{'İşletme yönetimi'}, {'Türk edebiyatı Roman'}, 0.3506550218340611], [{'Türk edebiyatı Roman'}, {'İşletme yönetimi'}, 0.10968446933479033], [{'İşletme yönetimi'}, {'Türk edebiyatı Roman'}, 0.3506550218340611], [{'Türk edebiyatı Roman'}, {'İşletme yönetimi'}, 0.10968446933479033], [{'İşletme yönetimi'}, {'Türk edebiyatı Roman'}, 0.3506550218340611], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216], [{'Türk edebiyatı Roman'}, {'Türk Tarihi'}, 0.1830351044939216], [{'Türk Tarihi'}, {'Türk edebiyatı Roman'}, 0.42138364779874216]]

12.3. Python Eclat Destek: 0.01 Tüm Veri Seti

{'Türk edebiyatı Roman,Türk Tarihi': 0.011835143414090782, 'Türk Tarihi,Türk Tarihi': 0.014132553606237816, 'Türk edebiyatı Roman,Türk edebiyatı Roman,Türk edebiyatı Roman': 0.01176552492341966, 'Türk Tarihi,Türk edebiyatı Roman':

0.011974380395433026, 'Türk edebiyatı Roman,Türk edebiyatı Roman':
 0.0657198551935394, 'İşletme yönetimi,İşletme yönetimi': 0.013297131718184349}

12.4. RStudio Apriori Destek: 0.01 Güven. 0.1 Tüm Veri Seti

lhs	rhs	support	confidence	coverage	lift
count					
[1] {}	=> {Fizik}	0.10221418	0.1022142	1.00000000	
1.00000000 1468					
[2] {}	=> {İşletme yönetimi}	0.16132851	0.1613285		
1.00000000 1.00000000 2317					
[3] {}	=> {Amerikan Edebiyatı}	0.15819524	0.1581952		
1.00000000 1.00000000 2272					
[4] {}	=> {Türk Tarihi}	0.22322796	0.2232280		
1.00000000 1.00000000 3206					
[5] {}	=> {Türk edebiyatı Roman}	0.51302047	0.5130205		
1.00000000 1.00000000 7368					
[6] {Sanat Tarihi}	=> {Türk edebiyatı Roman}	0.01030497	0.4111111		
0.02506615 0.8013542 148					
[7] {Tanımsız Kategori}	=> {Türk edebiyatı Roman}	0.01364712			
0.4875622 0.02799053 0.9503757 196					
[8] {Toplum bilimleri}	=> {Türk edebiyatı Roman}	0.01141902			
0.4100000 0.02785127 0.7991884 164					
[9] {Araştırma yöntemleri}	=> {Türk edebiyatı Roman}	0.01211530			
0.4084507 0.02966161 0.7961684 174					
[10] {Mühendislik}	=> {Türk edebiyatı Roman}	0.01239382			
0.3436293 0.03606740 0.6698160 178					
[11] {Ekonomi}	=> {Türk edebiyatı Roman}	0.01343824	0.3580705		
0.03752959 0.6979653 193					
[12] {Alman Edebiyatı}	=> {Türk edebiyatı Roman}	0.01698928			
0.4333925 0.03920067 0.8447861 244					
[13] {Kişisel gelişim}	=> {Türk edebiyatı Roman}	0.01740705			
0.4187605 0.04156803 0.8162646 250					
[14] {Rus edebiyatı}	=> {Türk edebiyatı Roman}	0.02019217			
0.4489164 0.04497981 0.8750458 290					
[15] {Fransız Edebiyatı}	=> {Türk edebiyatı Roman}	0.02353433			
0.4555256 0.05166411 0.8879287 338					
[16] {Din}	=> {Türk Tarihi}	0.01190642	0.2301480		
0.05173374 1.0310001 171					
[17] {Din}	=> {Türk edebiyatı Roman}	0.02290767	0.4427995		
0.05173374 0.8631224 329					
[18] {Matematik}	=> {Türk edebiyatı Roman}	0.02165437	0.3815951		
0.05674697 0.7438204 311					
[19] {Felsefe}	=> {Türk Tarihi}	0.01260270	0.2031425		
0.06203871 0.9100228 181					
[20] {Felsefe}	=> {Türk edebiyatı Roman}	0.02443949	0.3939394		
0.06203871 0.7678824 351					
[21] {Psikoloji}	=> {İşletme yönetimi}	0.01023534	0.1303191		
0.07854059 0.8077875 147					

[22] {Psikoloji}	=> {Türk Tarihi}	0.01399527 0.1781915
0.07854059 0.7982490 201		
[23] {Psikoloji}	=> {Türk edebiyatı Roman}	0.03140231 0.3998227
0.07854059 0.7793504 451		
[24] {Tıp}	=> {Amerikan Edebiyatı}	0.01023534 0.1165741
0.08780114 0.7369005 147		
[25] {Tıp}	=> {Türk Tarihi}	0.01315973 0.1498810
0.08780114 0.6714259 189		
[26] {Tıp}	=> {Türk edebiyatı Roman}	0.03613703 0.4115781
0.08780114 0.8022645 519		
[27] {İngiliz edebiyatı}	=> {Amerikan Edebiyatı}	0.01587523 0.1947054
0.08153461 1.2307917 228		
[28] {Amerikan Edebiyatı}	=> {İngiliz edebiyatı}	0.01587523 0.1003521
0.15819524 1.2307917 228		
[29] {İngiliz edebiyatı}	=> {Türk Tarihi}	0.01406489 0.1725021
0.08153461 0.7727622 202		
[30] {İngiliz edebiyatı}	=> {Türk edebiyatı Roman}	0.03718145
0.4560205 0.08153461 0.8888934 534		
[31] {Bilgisayar Bilimi}	=> {Fizik}	0.01246345 0.1261452
0.09880240 1.2341260 179		
[32] {Fizik}	=> {Bilgisayar Bilimi}	0.01246345 0.1219346
0.10221418 1.2341260 179		
[33] {Bilgisayar Bilimi}	=> {İşletme yönetimi}	0.01413452 0.1430585
0.09880240 0.8867527 203		
[34] {Bilgisayar Bilimi}	=> {Amerikan Edebiyatı}	0.01023534
0.1035941 0.09880240 0.6548496 147		
[35] {Bilgisayar Bilimi}	=> {Türk Tarihi}	0.01810333 0.1832276
0.09880240 0.8208095 260		
[36] {Bilgisayar Bilimi}	=> {Türk edebiyatı Roman}	0.03739034
0.3784355 0.09880240 0.7376616 537		
[37] {Fizik}	=> {İşletme yönetimi}	0.01357750 0.1328338
0.10221418 0.8233746 195		
[38] {Fizik}	=> {Amerikan Edebiyatı}	0.01093163 0.1069482
0.10221418 0.6760521 157		
[39] {Fizik}	=> {Türk Tarihi}	0.01872998 0.1832425
0.10221418 0.8208761 269		
[40] {Fizik}	=> {Türk edebiyatı Roman}	0.03794736 0.3712534
0.10221418 0.7236620 545		
[41] {İşletme yönetimi}	=> {Amerikan Edebiyatı}	0.01733742
0.1074666 0.16132851 0.6793286 249		
[42] {Amerikan Edebiyatı}	=> {İşletme yönetimi}	0.01733742
0.1095951 0.15819524 0.6793286 249		
[43] {İşletme yönetimi}	=> {Türk Tarihi}	0.03279488 0.2032801
0.16132851 0.9106391 471		
[44] {Türk Tarihi}	=> {İşletme yönetimi}	0.03279488 0.1469120
0.22322796 0.9106391 471		
[45] {İşletme yönetimi}	=> {Türk edebiyatı Roman}	0.05730400
0.3552007 0.16132851 0.6923714 823		
[46] {Türk edebiyatı Roman}	=> {İşletme yönetimi}	0.05730400
0.1116992 0.51302047 0.6923714 823		

[47] {Amerikan Edebiyatı} => {Türk Tarihi} 0.02367358 0.1496479
0.15819524 0.6703815 340
[48] {Türk Tarihi} => {Amerikan Edebiyatı} 0.02367358 0.1060512
0.22322796 0.6703815 340
[49] {Amerikan Edebiyatı} => {Türk edebiyatı Roman} 0.07150815
0.4520246 0.15819524 0.8811045 1027
[50] {Türk edebiyatı Roman} => {Amerikan Edebiyatı} 0.07150815
0.1393865 0.51302047 0.8811045 1027
[51] {Türk Tarihi} => {Türk edebiyatı Roman} 0.09483359 0.4248284
0.22322796 0.8280926 1362
[52] {Türk edebiyatı Roman} => {Türk Tarihi} 0.09483359 0.1848534
0.51302047 0.8280926 1362
[53] {İşletme yönetimi,Türk Tarihi} => {Türk edebiyatı Roman} 0.01127977
0.3439490 0.03279488 0.6704392 162
[54] {İşletme yönetimi,Türk edebiyatı Roman} => {Türk Tarihi} 0.01127977
0.1968408 0.05730400 0.8817929 162
[55] {Türk edebiyatı Roman,Türk Tarihi} => {İşletme yönetimi} 0.01127977
0.1189427 0.09483359 0.7372704 162
[56] {Amerikan Edebiyatı,Türk Tarihi} => {Türk edebiyatı Roman} 0.01009609
0.4264706 0.02367358 0.8312935 145
[57] {Amerikan Edebiyatı,Türk edebiyatı Roman} => {Türk Tarihi} 0.01009609
0.1411879 0.07150815 0.6324832 145
[58] {Türk edebiyatı Roman,Türk Tarihi} => {Amerikan Edebiyatı} 0.01009609
0.1064611 0.09483359 0.6729728 145

12.5. RStudio Eclat Destek: 0.01 Tüm Veri Seti

items	support	transIdenticalToItemsets	count
[1] {Sanat Tarihi,Türk edebiyatı Roman}	0.01030497	148	148
[2] {Tanımsız Kategori,Türk edebiyatı Roman}	0.01364712	196	196
[3] {Toplum bilimleri,Türk edebiyatı Roman}	0.01141902	164	164
[4] {Araştırma yöntemleri,Türk edebiyatı Roman}	0.01211530	174	174
[5] {Mühendislik,Türk edebiyatı Roman}	0.01239382	178	178
[6] {Ekonomi,Türk edebiyatı Roman}	0.01343824	193	193
[7] {Alman Edebiyatı,Türk edebiyatı Roman}	0.01698928	244	244
[8] {Kişisel gelişim,Türk edebiyatı Roman}	0.01740705	250	250
[9] {Rus edebiyatı,Türk edebiyatı Roman}	0.02019217	290	290
[10] {Fransız Edebiyatı,Türk edebiyatı Roman}	0.02353433	338	338
[11] {Din,Türk edebiyatı Roman}	0.02290767	329	329
[12] {Din,Türk Tarihi}	0.01190642	171	171
[13] {Matematik,Türk edebiyatı Roman}	0.02165437	311	311
[14] {Felsefe,Türk edebiyatı Roman}	0.02443949	351	351
[15] {Felsefe,Türk Tarihi}	0.01260270	181	181
[16] {Psikoloji,Türk edebiyatı Roman}	0.03140231	451	451
[17] {Psikoloji,Türk Tarihi}	0.01399527	201	201
[18] {İşletme yönetimi,Psikoloji}	0.01023534	147	147
[19] {Tıp,Türk edebiyatı Roman}	0.03613703	519	519

[20]	{Tıp,Türk Tarihi}	0.01315973	189	189
[21]	{Amerikan Edebiyatı,Tıp}	0.01023534	147	147
[22]	{İngiliz edebiyatı,Türk edebiyatı Roman}	0.03718145	534	534
[23]	{İngiliz edebiyatı,Türk Tarihi}	0.01406489	202	202
[24]	{Amerikan Edebiyatı,İngiliz edebiyatı}	0.01587523	228	228
[25]	{Bilgisayar Bilimi,Türk edebiyatı Roman}	0.03739034	537	537
[26]	{Bilgisayar Bilimi,Türk Tarihi}	0.01810333	260	260
[27]	{Amerikan Edebiyatı,Bilgisayar Bilimi}	0.01023534	147	147
[28]	{Bilgisayar Bilimi,İşletme yönetimi}	0.01413452	203	203
[29]	{Bilgisayar Bilimi,Fizik}	0.01246345	179	179
[30]	{Fizik,Türk edebiyatı Roman}	0.03794736	545	545
[31]	{Fizik,Türk Tarihi}	0.01872998	269	269
[32]	{Amerikan Edebiyatı,Fizik}	0.01093163	157	157
[33]	{Fizik,İşletme yönetimi}	0.01357750	195	195
[34]	{İşletme yönetimi,Türk edebiyatı Roman,Türk Tarihi}	0.01127977	162	162
[35]	{İşletme yönetimi,Türk edebiyatı Roman}	0.05730400	823	823
[36]	{İşletme yönetimi,Türk Tarihi}	0.03279488	471	471
[37]	{Amerikan Edebiyatı,İşletme yönetimi}	0.01733742	249	249
[38]	{Amerikan Edebiyatı,Türk edebiyatı Roman,Türk Tarihi}	0.01009609	145	145
[39]	{Amerikan Edebiyatı,Türk edebiyatı Roman}	0.07150815	1027	1027
[40]	{Amerikan Edebiyatı,Türk Tarihi}	0.02367358	340	340
[41]	{Türk edebiyatı Roman,Türk Tarihi}	0.09483359	1362	1362
[42]	{Türk edebiyatı Roman}	0.51302047	7368	7368
[43]	{Türk Tarihi}	0.22322796	3206	3206
[44]	{Amerikan Edebiyatı}	0.15819524	2272	2272
[45]	{İşletme yönetimi}	0.16132851	2317	2317
[46]	{Fizik}	0.10221418	1468	1468
[47]	{Bilgisayar Bilimi}	0.09880240	1419	1419
[48]	{İngiliz edebiyatı}	0.08153461	1171	1171
[49]	{Tıp}	0.08780114	1261	1261
[50]	{Psikoloji}	0.07854059	1128	1128
[51]	{Felsefe}	0.06203871	891	891
[52]	{Matematik}	0.05674697	815	815
[53]	{Din}	0.05173374	743	743
[54]	{Fransız Edebiyatı}	0.05166411	742	742
[55]	{Rus edebiyatı}	0.04497981	646	646
[56]	{Kişisel gelişim}	0.04156803	597	597
[57]	{Alman Edebiyatı}	0.03920067	563	563
[58]	{Ekonomi}	0.03752959	539	539
[59]	{Mühendislik}	0.03606740	518	518
[60]	{Araştırma yöntemleri}	0.02966161	426	426
[61]	{Toplum bilimleri}	0.02785127	400	400
[62]	{Tanımsız Kategori}	0.02799053	402	402
[63]	{Sanat Tarihi}	0.02506615	360	360
[64]	{Kimya}	0.02088846	300	300
[65]	{Portekiz edebiyatı}	0.01517894	218	218
[66]	{Mimari - Mimarlık}	0.01552709	223	223

[67] {Hukuk}	0.01371675 197	197
[68] {Aile}	0.01371675 197	197
[69] {İtalyan edebiyatı}	0.01065311 153	153
[70] {Yapı endüstrisi}	0.01044423 150	150
[71] {Çocuk Psikolojisi}	0.01086200 156	156

12.6. SMPF Apriori Destek: 0.01 Güven. 0.1 Tüm Veri Seti

Alman Edebiyatı=t ==> Türk edebiyatı Roman=t	221.0	0.41154562383612664
İngiliz edebiyatı=t ==> Türk edebiyatı Roman=t	498.0	0.437609841827768
İngiliz edebiyatı=t ==> Türk Tarihi=t	185.0	0.1625659050966608
İngiliz edebiyatı=t ==> Amerikan Edebiyatı=t	199.0	0.17486818980667837
Bilgisayar Bilimi=t ==> Türk edebiyatı Roman=t	498.0	0.36034732272069464
Türk Tarihi=t ==> Türk edebiyatı Roman=t	1256.0	0.4052920296869958
Türk edebiyatı Roman=t ==> Türk Tarihi=t	1256.0	0.17475998330318632
Fizik=t ==> Türk edebiyatı Roman=t	517.0	0.36052998605299863
Fransız Edebiyatı=t ==> Türk edebiyatı Roman=t	306.0	0.43342776203966005
Psikoloji=t ==> Türk edebiyatı Roman=t	414.0	0.3812154696132597
Tanımsız Kategori=t ==> Türk edebiyatı Roman=t	182.0	0.46666666666666667
Matematik=t ==> Türk edebiyatı Roman=t	281.0	0.3575063613231552
İşletme yönetimi=t ==> Türk edebiyatı Roman=t	756.0	0.3361493997332148
Türk edebiyatı Roman=t ==> İşletme yönetimi=t	756.0	0.10518992625573953
Amerikan Edebiyatı=t ==> Türk edebiyatı Roman=t	930.0	0.4301572617946346
Türk edebiyatı Roman=t ==> Amerikan Edebiyatı=t	930.0	0.129400306108251
Rus edebiyatı=t ==> Türk edebiyatı Roman=t	266.0	0.4249201277955272
Mühendislik=t ==> Türk edebiyatı Roman=t	162.0	0.32142857142857145
Ekonomi=t ==> Türk edebiyatı Roman=t	176.0	0.3365200764818356
Kişisel gelişim=t ==> Türk edebiyatı Roman=t	231.0	0.4024390243902439
Araştırma yöntemleri=t ==> Türk edebiyatı Roman=t	161.0	0.3907766990291262
Din=t ==> Türk edebiyatı Roman=t	304.0	0.4233983286908078
Tıp=t ==> Türk edebiyatı Roman=t	463.0	0.3823286540049546
Felsefe=t ==> Türk edebiyatı Roman=t	327.0	0.37803468208092483
Toplum bilimleri=t ==> Türk edebiyatı Roman=t	145.0	0.37662337662337664
Bilgisayar Bilimi=t ==> Türk Tarihi=t	248.0	0.17945007235890015
Fizik=t ==> Bilgisayar Bilimi=t	174.0	0.12133891213389121
Bilgisayar Bilimi=t ==> Fizik=t	174.0	0.12590448625180897
Bilgisayar Bilimi=t ==> İşletme yönetimi=t	195.0	0.1410998552821997
Fizik=t ==> Türk Tarihi=t	254.0	0.17712691771269176
Psikoloji=t ==> Türk Tarihi=t	185.0	0.1703499079189687
İşletme yönetimi=t ==> Türk Tarihi=t	446.0	0.19831036016007114
Türk Tarihi=t ==> İşletme yönetimi=t	446.0	0.14391739270732495

Amerikan Edebiyatı=t ==> Türk Tarihi=t	307.0	0.1419981498612396
Din=t ==> Türk Tarihi=t	158.0	0.2200557103064067
Tıp=t ==> Türk Tarihi=t	166.0	0.1370767960363336
Felsefe=t ==> Türk Tarihi=t	168.0	0.19421965317919074
Fizik=t ==> İşletme yönetimi=t	181.0	0.12622036262203626
Fizik=t ==> Amerikan Edebiyatı=t	146.0	0.10181311018131102
Amerikan Edebiyatı=t ==> İşletme yönetimi=t	230.0	0.10638297872340426
İşletme yönetimi=t ==> Amerikan Edebiyatı=t	230.0	0.10226767452200979
Türk Tarihi=t İşletme yönetimi=t ==> Türk edebiyatı Roman=t	147.0	0.3295964125560538
Türk edebiyatı Roman=t İşletme yönetimi=t ==> Türk Tarihi=t	147.0	0.19444444444444445
Türk edebiyatı Roman=t Türk Tarihi=t ==> İşletme yönetimi=t	147.0	0.11703821656050956

12.7. SMPF FP-Growth Destek: 0.01 Güven. 0.1 Tüm Veri Seti

Alman Edebiyatı=t ==> Türk edebiyatı Roman=t	221.0	0.41154562383612664
İngiliz edebiyatı=t ==> Türk edebiyatı Roman=t	498.0	0.437609841827768
İngiliz edebiyatı=t ==> Türk Tarihi=t	185.0	0.1625659050966608
İngiliz edebiyatı=t ==> Amerikan Edebiyatı=t	199.0	0.17486818980667837
Bilgisayar Bilimi=t ==> Türk edebiyatı Roman=t	498.0	0.36034732272069464
Türk Tarihi=t ==> Türk edebiyatı Roman=t	1256.0	0.4052920296869958
Türk edebiyatı Roman=t ==> Türk Tarihi=t	1256.0	0.17475998330318632
Fizik=t ==> Türk edebiyatı Roman=t	517.0	0.36052998605299863
Fransız Edebiyatı=t ==> Türk edebiyatı Roman=t	306.0	0.43342776203966005
Psikoloji=t ==> Türk edebiyatı Roman=t	414.0	0.3812154696132597
Tanımsız Kategori=t ==> Türk edebiyatı Roman=t	182.0	0.46666666666666667
Matematik=t ==> Türk edebiyatı Roman=t	281.0	0.3575063613231552
İşletme yönetimi=t ==> Türk edebiyatı Roman=t	756.0	0.3361493997332148
Türk edebiyatı Roman=t ==> İşletme yönetimi=t	756.0	0.10518992625573953
Amerikan Edebiyatı=t ==> Türk edebiyatı Roman=t	930.0	0.4301572617946346
Türk edebiyatı Roman=t ==> Amerikan Edebiyatı=t	930.0	0.129400306108251
Rus edebiyatı=t ==> Türk edebiyatı Roman=t	266.0	0.4249201277955272
Mühendislik=t ==> Türk edebiyatı Roman=t	162.0	0.32142857142857145
Ekonomi=t ==> Türk edebiyatı Roman=t	176.0	0.3365200764818356
Kişisel gelişim=t ==> Türk edebiyatı Roman=t	231.0	0.4024390243902439
Araştırma yöntemleri=t ==> Türk edebiyatı Roman=t	161.0	0.3907766990291262
Din=t ==> Türk edebiyatı Roman=t	304.0	0.4233983286908078
Tıp=t ==> Türk edebiyatı Roman=t	463.0	0.3823286540049546

Felsefe=t ==> Türk edebiyatı Roman=t	327.0	0.37803468208092483
Toplum bilimleri=t ==> Türk edebiyatı Roman=t	145.0	0.37662337662337664
Bilgisayar Bilimi=t ==> Türk Tarihi=t	248.0	0.17945007235890015
Fizik=t ==> Bilgisayar Bilimi=t	174.0	0.12133891213389121
Bilgisayar Bilimi=t ==> Fizik=t	174.0	0.12590448625180897
Bilgisayar Bilimi=t ==> İşletme yönetimi=t	195.0	0.1410998552821997
Fizik=t ==> Türk Tarihi=t	254.0	0.17712691771269176
Psikoloji=t ==> Türk Tarihi=t	185.0	0.1703499079189687
İşletme yönetimi=t ==> Türk Tarihi=t	446.0	0.19831036016007114
Türk Tarihi=t ==> İşletme yönetimi=t	446.0	0.14391739270732495
Amerikan Edebiyatı=t ==> Türk Tarihi=t	307.0	0.1419981498612396
Din=t ==> Türk Tarihi=t	158.0	0.2200557103064067
Tıp=t ==> Türk Tarihi=t	166.0	0.1370767960363336
Felsefe=t ==> Türk Tarihi=t	168.0	0.19421965317919074
Fizik=t ==> İşletme yönetimi=t	181.0	0.12622036262203626
Fizik=t ==> Amerikan Edebiyatı=t	146.0	0.10181311018131102
Amerikan Edebiyatı=t ==> İşletme yönetimi=t	230.0	0.10638297872340426
İşletme yönetimi=t ==> Amerikan Edebiyatı=t	230.0	0.10226767452200979
Türk Tarihi=t İşletme yönetimi=t ==> Türk edebiyatı Roman=t	147.0	0.3295964125560538
Türk edebiyatı Roman=t İşletme yönetimi=t ==> Türk Tarihi=t	147.0	0.19444444444444445
Türk edebiyatı Roman=t Türk Tarihi=t ==> İşletme yönetimi=t	147.0	0.11703821656050956

12.8. SMPF Eclat Destek: 0.01 Tüm Veri Seti

Alman Edebiyatı=t	537.0
İngiliz edebiyatı=t	1138.0
Türk edebiyatı Roman=t	7187.0
Bilgisayar Bilimi=t	1382.0
Türk Tarihi=t	3099.0
Fizik=t	1434.0
Fransız Edebiyatı=t	706.0
Psikoloji=t	1086.0
Tanımsız Kategori=t	390.0
Matematik=t	786.0
İşletme yönetimi=t	2249.0
Amerikan Edebiyatı=t	2162.0
Rus edebiyatı=t	626.0
Mühendislik=t	504.0

Ekonomi=t	523.0
Kişisel gelişim=t	574.0
Araştırma yöntemleri=t	412.0
Portekiz edebiyatı=t	210.0
Din=t	718.0
Sanat Tarihi=t	347.0
Mimari - Mimarlık=t	216.0
Tıp=t	1211.0
Felsefe=t	865.0
Hukuk=t	189.0
Yapı endüstrisi=t	147.0
Çocuk Psikolojisi=t	147.0
Toplum bilimleri=t	385.0
Aile =t	187.0
Kimya=t	283.0
Toplum bilimleri=t Türk edebiyatı Roman=t	145.0
Tanımsız Kategori=t Türk edebiyatı Roman=t	182.0
Araştırma yöntemleri=t Türk edebiyatı Roman=t	161.0
Mühendislik=t Türk edebiyatı Roman=t	162.0
Ekonomi=t Türk edebiyatı Roman=t	176.0
Alman Edebiyatı=t Türk edebiyatı Roman=t	221.0
Kişisel gelişim=t Türk edebiyatı Roman=t	231.0
Rus edebiyatı=t Türk edebiyatı Roman=t	266.0
Fransız Edebiyatı=t Türk edebiyatı Roman=t	306.0
Din=t Türk Tarihi=t	158.0
Din=t Türk edebiyatı Roman=t	304.0
Matematik=t Türk edebiyatı Roman=t	281.0
Felsefe=t Türk Tarihi=t	168.0
Felsefe=t Türk edebiyatı Roman=t	327.0
Psikoloji=t Türk Tarihi=t	185.0
Psikoloji=t Türk edebiyatı Roman=t	414.0
İngiliz edebiyatı=t Amerikan Edebiyatı=t	199.0
İngiliz edebiyatı=t Türk Tarihi=t	185.0
İngiliz edebiyatı=t Türk edebiyatı Roman=t	498.0
Tıp=t Türk Tarihi=t	166.0
Tıp=t Türk edebiyatı Roman=t	463.0
Bilgisayar Bilimi=t Fizik=t	174.0
Bilgisayar Bilimi=t İşletme yönetimi=t	195.0
Bilgisayar Bilimi=t Türk Tarihi=t	248.0
Bilgisayar Bilimi=t Türk edebiyatı Roman=t	498.0
Fizik=t Amerikan Edebiyatı=t	146.0

Fizik=t İşletme yönetimi=t	181.0
Fizik=t Türk Tarihi=t	254.0
Fizik=t Türk edebiyatı Roman=t	517.0
Amerikan Edebiyatı=t İşletme yönetimi=t	230.0
Amerikan Edebiyatı=t Türk Tarihi=t	307.0
Amerikan Edebiyatı=t Türk edebiyatı Roman=t	930.0

