VERİ MADENCİLİĞİNDE BİRLİKTELİK KURALI YÖNTEMİNİ KREDİ KARTI KULLANAN MÜŞTERİLER ÜZERİNDE UYGULANMASI VE POWER BI'DA RAPORLANMASI

Dilay ŞİMŞEK
211403201

YÜKSEK LİSANS TEZİ

Endüstri Mühendisliği Anabilim Dalı
Endüstri Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Hüseyin YENER

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Aralık, 2023

VERİ MADENCİLİĞİNDE BİRLİKTELİK KURALI YÖNTEMİNİ KREDİ KARTI KULLANAN MÜŞTERİLER ÜZERİNDE UYGULANMASI VE POWER BI'DA RAPORLANMASI

Dilay ŞİMŞEK
211403201
ORCID: 0000-0003-2859-0148

YÜKSEK LİSANS TEZİ

Endüstri Mühendisliği Anabilim Dalı
Endüstri Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Hüseyin YENER

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Aralık, 2023



JÜRİ VE ENSTİTÜ ONAYI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli "Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge" ile bildirilen 6698 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.

ETİK İLKE VE KURALLARA UYUM BEYANI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli "Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge" ile bildirilen 6698 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.

TEŞEKKÜR

Çalışmamın gerçekleşmesinde kıymetli bilgilerini benimle paylaşan, ne zaman çalışmam ile ilgili danışsam bana o değerli zamanını ayırıp sabırla ve ilgiyle bana yararlı olabilmek için her bilgisini paylaşan, güler yüzünü ve samimiyetini benden hiç esirgemeyen, kıymetli ve saygıdeğer danışman hocam Dr. Öğr. Üyesi Hüseyin YENER'e teşekkürü bir borç biliyor ve şükranlarımı sunuyorum ve beni bu günlere sevgi ve saygı kelimelerinin anlamlarını bilecek şekilde yetiştirerek getiren ve benden hiçbir zaman desteğini esirgemeyen bu hayattaki en büyük şansım olan aileme sonsuz teşekkürler.

Dilay ŞİMŞEK

Aralık, 2023

ÖZET

VERİ MADENCİLİĞİNDE BİRLİKTELİK KURALI YÖNTEMİNİ KREDİ KARTI KULLANAN MÜŞTERİLER ÜZERİNDE UYGULANMASI VE POWER BI'DA RAPORLANMASI

Dilay Şimşek
Yüksek Lisans Tezi
Endüstri Mühendisliği Anabilim Dalı
Endüstri Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Hüseyin Yener
Maltepe Üniversitesi Lisansüstü Eğitim Enstitüsü, 2023

Günümüzde banka sektöründeki kredi kartı kullanan müşteri sayısı giderek artmaktadır. Bunun nedeni kişilerin değişen hayat şartlarına ve taleplerine uyum sağlayabilmeleri için bu yola ihtiyaç duymalarıdır. Giderek artan müşteri sayısından kaynaklı müşterileri analiz etmek de gittikçe zorlaşmaktadır. Bankaların kredi kartı verdikten sonra odaklandıkları en önemli faktörlerden biri kredi kartı ödeme vadelerinin durumlarının belirlenmesidir. Kredi durumlarına göre de müşteri niteliklerinin analizinin yapılmasıdır. Buradaki büyük veri yığınları arasındaki müşteriyi analiz etmek için veri madenciliği yöntemi kullanarak ham veri yığınları arasındaki ilişkileri ve kuralların ortaya çıkarılmasını sağlamaktır.

Bu çalışmada veri madenciliği yöntemlerinden olan birliktelik kuralları yöntemi sayesinde kredi kartı ödeme vadelerine göre müşterilerin hangi özelliklere sahip olduğunu belirlemektir.

Çalışmada banka sektörüne ait müşteri bilgilerini kapsayan bir aylık veri kümesi kullanılmıştır. Veri kümesi üzerinde SQL dili sayesinde veri analizi yapılarak 25.134 adet müşterinin bilgileri algoritmaya hazır hale getirilmiştir. Veri seti içerisinde müşterinin cinsiyeti, arabası olup/olmama durumu, mülkü olup/olmama durumu, çocuk sayısı, yıllık gelir tutarı, eğitim durumu, medeni durumu, yaş durumu ve bunlara bağlı olarak kredi kartı ödeme vadeleri bazında değerlendirilmiştir. Bununla birlikte veri kümesinde belirlenen alanlar için belli kategorilerde kategorize edilerek kullanılmıştır. Çalışmada birliktelik kuralı için müşteri bilgileri mantıksal yapıya çevrilmiştir ve kredi kartı ödeme vadelerindeki durumlar için ayrı ayrı destek ve güven eşik değerleri belirlenmiştir.

Çalışmanın çözümü için apriori algoritması kullanılarak Python yardımıyla birliktelik kuralları oluşturulmuştur. Bunun yanında Power BI sayesinde de veriler görselleştirilmiştir. Bu sayede kredi kartı ödeme vadelerine göre müşterilerin hangi özelliklere sahip olduğu belirlenmiştir. Çalışmanın sonucunda; kredi kartı sahibi olmayan, kredi kartını zamanında ödeyen ve kredi kartının vadesinin borcunu 3 ay geçiren yetişkin müşterilerin cinsiyeti kadındır. Hiç çocuğu olmayan yetişkin müşteriler ise kredi kartının borcunun vadesini geçirmiş ve kötü borca sahip kişilerdir. Ayrıca mülk sahibi olmayan ve eğitim durumu ilköğretim/ortaöğretimden mezun olan müşteriler ise kredi kartının borcunun 6 ay geçmesine rağmen ödemeyen müşterilerdir gibi kurallar oluşturulmuştur. Bu sonuçlar neticesinde mevcut ve potansiyel müşterilerin kredi kartı ödeme vadelerinin belirlenmesini amaçlamaktadır.

Anahtar Sözcükler: Veri madenciliği, birliktelik kuralları, apriori algoritması, banka sektörü, Power BI, Python, SQL

ABSTRACT

APPLİCATİON OF ASSOCİATİON RULE METHOD İN DATA MİNİNG ON CUSTOMERS USİNG CREDİT CARDS AND REPORTİNG İN POWER BI

Dilay Şimşek
Master's Thesis
Department of Industrial Engineering
Industrial Engineering Master's with Thesis Program
Thesis Advisor: Dr. Öğr. Üyesi Hüseyin Yener
Maltepe University Graduate School, 2023

Nowadays, the number of customers using credit cards in the banking sector is increasing. The reason for this is that people need this way to adapt to changing life conditions and demands. It is becoming increasingly difficult to analyze customers due to the increasing number of customers. One of the most important factors that banks focus on after issuing a credit card is determining the status of credit card payment terms. It is the analysis of customer qualifications according to their credit status. The aim here is to reveal the relationships and rules between raw data piles by using the data mining method to analyze the customer among large data piles.

In this study, the association rules method, one of the data mining methods, aims to determine the characteristics of customers according to their credit card payment terms.

A monthly data set covering customer information of the banking sector was used in the study. By performing data analysis on the dataset using SQL language, the information of 25,134 customers was made ready for the algorithm. Within the data set, the customer's gender, whether they have a car or not, whether they have a property or not, number of children, annual income amount, education level, marital status, age status and, accordingly, credit card payment terms were evaluated. In addition, it was categorized and used in certain categories for the areas determined in the data set. In the study, customer information was converted into a logical structure for the association rule and separate support and trust threshold values were determined for situations in credit card payment terms.

To solve the study, association rules were created with the help of Python using the a priori algorithm. In addition, the data was visualized thanks to Power BI. In this way, it was determined which characteristics the customers had according to their credit card payment terms. As a result of the study; The gender of adult customers who do not have a credit card, who pay their credit card on time and whose credit card is due within 3 months is female. Adult customers who do not have any children are those who have overdue credit card debt and have bad debt. In addition, rules have been created such as customers who do not own property and whose education level is primary/secondary education are those who have not paid their credit card debt even after 6 months. As a result of these results, it aims to determine the credit card payment terms of existing and potential customers.

Keywords: Data mining, association rules, apriori algorithm, banking industry, Power BI, Python, SQL

İÇİNDEKİLER

JÜRİ VE ENSTİTÜ ONAYI ii
ETİK İLKE VE KURALLARA UYUM BEYANI iii
TEŞEKKÜR iv
ÖZET v
ABSTRACT vii
İÇİNDEKİLER ix
TABLOLAR LİSTESİ xi
ŞEKİLLER LİSTESİ xii
1. GİRİŞ 1
1.1 Problem 1
1.2 Amaç 1
1.3 Önem 1
1.4 Varsayımlar 2
1.5 Sınırlıklar 2
1.6 Tanımlar 2
2. GENEL BİLGİLER 4
2.1 Veri Madenciliği 4
2.1.1 Birliktelik kuralları 5
2.1.1.1 Birliktelik kuralındaki tanımlamalar 7
2.1.1.2 Apriori algoritması 9
2.1.1.3 Birliktelik kuralları literatür özeti 10
3. YÖNTEM 14

3.1 Araştırma Modeli 14
3.2 Evren ve Örneklem 16
3.3 Veriler ve Toplanması 16
3.4 Verilerin Çözümlenmesi ve Yorumlanması 21
4. BULGULAR VE TARTIŞMA 29
4.1 Bulgular 29
4.2 Tartışma 48
5. SONUÇ ve ÖNERİLER 50
KAYNAKLAR 53
ÖZGEÇMİŞ 57

TABLOLAR LİSTESİ

Tablo 1. Rastgele seçilen 10 müşteriye ait müşteri bilgileri 17

Tablo 2. Veri kümesi içerisindeki tüm durumlar ve frekansları 17

Tablo 3. Kredi durumu vadesi 6 ay geçen borca sahip müşterilerin bilgileri 21

Tablo 4. Müşteri bilgileri ve frekans sayıları 22

Tablo 5. Destek değerinin üzerinde olan müşteri bilgileri ve frekansları 23

Tablo 6. İkili gruplandırma sonucunda müşteri bilgileri ve frekansları..... 23

Tablo 7. Destek değerinin üzerinde olan müşteri bilgileri ve frekansları 23

Tablo 8. Üçlü gruplandırma sonucunda müşteri bilgileri ve frekansları 24

Tablo 9. Tüm gruplandırma sonucunda müşteri bilgileri ve frekansları..... 24

Tablo 10. Kredi kartı borcunu zamanında ödeyen müşterilerin Python Jupyter Notebook uygulaması üzerinde algoritmanın çözümü 25

Tablo 11. Tüm sık geçen müşteri bilgileri 29

Tablo 12. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar 30

Tablo 13. Tüm sık geçen müşteri bilgileri 33

Tablo 14. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar 34

Tablo 15. Tüm sık geçen müşteri bilgileri 35

Tablo 16. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar 35

Tablo 17. Tüm sık geçen müşteri bilgileri 37

Tablo 18. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar 38

Tablo 19. Tüm sık geçen müşteri bilgileri 39

Tablo 20. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar 40

ŞEKİLLER LİSTESİ

Şekil 1. Veri Madenciliği Modelleri 5

Şekil 2. Birliktelik kuralı algoritmaları 6

Şekil 3. Çalışmanın süreci 14

Şekil 4. Apriori algoritması akış diyagramı 14

Şekil 5. 2^N kümesi Gösterilişi 15

Şekil 6. Apriori algoritmasının matematiksel modeli 16

Şekil 7. Veri setindeki kredi kartı durumu vadesi geçmiş ve kötü borca sahip kişilerin bilgileri 20

Şekil 8. Kredi kartını zamanında ödeyen müşterilerin Power BI'da görselleştirilmesi . 33

Şekil 9. Kredi kartı ödeme vadesi geçmiş veya kötü borca sahip müşterilerin Power BI'da görselleştirilmesi 34

Şekil 10. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi 36

Şekil 11. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi 36

Şekil 12. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi 36

Şekil 13. Kredi kartı borcunun vadesi 3 ay geçen müşterilerin Power BI'da görselleştirilmesi 39

Şekil 14. Kredi kartına sahip olmayan müşterilerin Power BI'da görselleştirilmesi 41

Şekil 15. Kredi durumlarına göre müşteri sayısı 42

Şekil 16. Kredi kartını zamanında ödeyen müşterilerin tüm özelliklerini birlikte kapsayan durumların gösterimi 42

Şekil 17. Vadesi geçmiş veya kötü borca sahip müşterilerin tüm özelliklerini birlikte kapsayan durumların gösterimi 43

Şekil 18. Cinsiyete göre kredi durumları 43

Şekil 19. Çocuk sayısına göre kredi durumları 44

Şekil 20. Eğitim durumuna göre kredi kartı ödeme durumları 44

Şekil 21. Medeni durumuna göre kredi durumları 45

Şekil 22. Mülk durumuna göre kredi durumları 45

Şekil 23. Yıllık gelir durumuna göre kredi durumları 46

Şekil 24. Yaş durumuna göre kredi kartı ödeme durumları 46

Şekil 25. Yıllık gelir durumu ve eğitim durumu arasındaki ilişkiyi kredi kartı ödeme durumlarına göre gösterilmesi 47

Şekil 26. Çocuk sayısı ile mülk durumu arasındaki ilişkiyi kredi kartı ödeme durumlarına göre gösterilmesi 47

Şekil 27. Arabası olup/olmama durumuna göre kredi kartı ödeme durumları 48

1. GİRİŞ

1.1 Problem

Günümüzde banka sektöründe kredi kartı kullanan müşteri sayısı giderek artmaktadır. Giderek artan müşteri sayısından kaynaklı, veriyi analiz etmek gittikçe zorlaşmaktadır. Bu çalışmadaki problem, işletmeler kredi kartı ödeme vadelerinde önemli faktörlerinden biri olan müşteri niteliklerinin belirlenmesi problemine odaklanmaktadırlar. Büyük veri yığınları arasından müşteri analizi yapabilmek için veri madenciliğine ihtiyaç vardır. Çünkü veri madenciliği sayesinde ham bilgi yığınları arasındaki ilişkilerin ve kuralların meydana gelmesi önemlidir. Bu anlamda veri madenciliği yöntemlerinden olan birliktelik kuralları sayesinde kredi kartı ödeme vadelerine göre müşterinin hangi özelliklere sahip olduğunu saptanarak yeni gelecek müşteriler için kararları doğru almak açısından önemlidir.

1.2 Amaç

Çalışmanın amacı, birliktelik kuralları sayesinde kredi kartı ödeme vadelerine göre müşterilerin hangi özelliklere sahip olduğunu saptayarak, mevcut veya potansiyel müşteriler için kredi kartının ödeme aşamasında firmalara önemli katkılar sunabilmektedir.

1.3 Önem

Birliktelik kuralları sayesinde firmalar satış stratejilerini değiştirebilirler. Aynı zamanda çalışmanın sonucuna göre müşterilerin sahip olduğu niteliklerin incelenmesine de olanak tanırken hangi müşteri özelliğine sahip olan kişiler kredi kartlarını zamanında veya geç ödedikleri konusunda da fikir vermektedir. Algoritma sayesinde tahmini yapılamayan durumlarda analiz yapılabilmesi karar verme konusunda çok önemli bir yere sahiptir. Ayrıca yeni gelecek müşteriler içinde kredi kartı satış politikasını yönetmemizde yardımcı olabilir.

1.4 Varsayımlar

Birliktelik kuralları, veri yığınları içerisindeki bilgilerin birlikte görülme olasılıklarının ortaya çıkarılmasında kullanılmaktadır. Birliktelik kuralında beraber gerçekleşen muhtemel olayların kuralların belirlenmesini hedeflenmektedir. Bu anlamda bağlantı kurallarında ilk olarak nesne sayısının belirtilmesi gerekmektedir. İkinci adımda ise algoritmaya uygulanacak veri setinin içerisinde kaç adetin nesne seti olduğunun tespit edilerek kural sayısının bilinmesi gerekmektedir. Çözümlemeye bazı varsayımlarla başlanır. Destek ve güven ölçütleri için eşik değerler belirlenmektedir. Bu yüzden çıkarım yapılması istenen kuralların elde edilen kümelerin çoğu tarafından destek görmesi gerekmektedir. Sonuç olarak ilişkilendirme kuralı için dışarıdan minimum güven ve destek eşik değerlerini sağlayacak şeklinde üretilmelidir.

1.5 Sınırlıklar

Çalışmada kullanılacak algoritmanın getirdiği bazı sınırlılıklar vardır. Bunlar; problem için veri büyüklüğü artığında sonuçlarda üstel bir şekilde artmaktadır. Bunun yanında veri kümesindeki nesnelerde çok az rastlanan durumları yok saymaktadır. Başlangıçta belirlenen destek ve güven değerleri algoritma sonucu üretilen ilişki kuralları sayısına kısıtlamalar getirmektedir. Bundan dolayı dışardan çok küçük olacak şekilde belirlenen eşik değerler, kullanıcının gerçekten istediği sonucu kaybetme riskini de getirmektedir. Bunlara ek olarak da doğru olarak özellik sayısına karar vermek zor olmaktadır.

1.6 Tanımlar

Bu çalışmada kullanılan tanımlamalar şu şekilde açıklanmaktadır:

Veri madenciliği, işe yaramayan veriden anlamlı bir veri yaratma yöntemidir.

Algoritma, büyük veri nesnelerini alarak belirlenen bir çıktıyla sonuçlanan kurallardır.

Birliktelik kuralları, veri yığınları içerisindeki bilgilerin birlikte görülme olasılıklarının ortaya çıkarılmasında kullanılmaktadır. Yani verilerin potansiyel ilişkilerinin ortaya çıkarılma işlemidir.

Apriori algoritması, alt yapısında minimum destek ve minimum güven eşik değerleri ele alınarak veri tabanında birliktelik kurallarının belirlenmesini amaçlamaktadır.

Destek değeri, temsil eden kuralın asgari destek eşik değerini (X ve Y nesnelerini kapsayan işlemlerin toplam işlemlerin içerisinde en az %D oranında olduğunu) ifade etmektedir.

Güven değeri, ilgili kuralın minimum güvenirliğini göstermesine güven eşik değeri denilmektedir. Yani X nesnesini içeren işlemlerin en az %G oranında Y nesnesini de içeren işlemler kümesinde yer aldığını göstermektedir.

Lift Değeri, kuralın ilginçliği veya farklılığını ifade eder.

Confidence değeri, 0 ile 1 arasındadır. Yüksek değer kuralımızın çok güvenli veya doğru olduğunu belirtmektedir.

Conviction (Kanaat) değeri, X elemanının, Y elemanı olmaksızın görülme olasılığını ifade etmektedir.

Leveage (Kaldıraç) değeri, X ve Y nesnelerinin birlikte görülmesinin, X ve Y'nin ayrı ayrı görülmesinden ne kadar fazla olduğunu belirtmektedir.

2. GENEL BİLGİLER

Günümüzde gittikçe zorlaşan hayat koşulları sayesinde banka sektöründe kredi kartı kullanan müşteri sayısı giderek artmaktadır. Giderek artan müşteri sayısından kaynaklı, veriyi analiz etmek gittikçe zorlaşmaktadır. İşletmeler kredi kartı verme politikalarını belirlerken önemli faktörlerinden biri müşteri niteliklerinin belirlenmesidir. Gelişen teknoloji ve bilgisayar sayesinde bu veriler depolanabildiğinden verilerin analizi için bazı istatistiksel yöntemler geliştirilmiştir. Veri kavramı, üzerinde yorum yapmaya olanak sağlamayan sistemleştirilmiş ham bilgidir ve veri madenciliği kavramının da ortaya çıkmasını sağlamaktadır. Veri madenciliği sayesinde ham bilgi yığınları arasındaki ilişkilerin ve kuralların meydana gelmesinde önemlidir. Bu anlamda veri madenciliği yöntemlerinden olan birliktelik kuralları sayesinde kredi kartı ödeme vadelerine göre müşterilerin hangi özelliklere sahip olduğunu saptamaktır.

2.1 Veri Madenciliği

Veri madenciliği, ham bilgi yığınları arasındaki bağlantıların analiz edilerek aralarındaki ilişkilerin ortaya çıkarılmasını sağlayan ve veri tabanlarında depolanan verilerin yorum yapılacak bir şekilde sistemleştirilmesinde kullanılan kavramdır. Sonuç olarak işe yaramayan veriden anlamlı bir veri yaratma yöntemine veri madenciliği denilmektedir. Veri madenciliği teknikleri sayesinde büyük veri yığınları kolay bir şekilde analiz edilip, raporlanmasına olanak sağlamaktadır (Gündüz, 2017).

Algoritma ise, büyük veri nesnelerini alarak belirlenen bir çıktıyla sonuçlanan kurallardır. Veri madenciliğinin temelinde üç temel algoritma bulunmaktadır. Bunlar birliktelik algoritmaları, sınıflandırıcı algoritmalar ve kümeleme algoritmalarıdır (Bayram ve Dündar, 2021).

Şekil 1'deki modele göre veri madenciliğinde kesif esaslı yöntemler, tahmin edici ve tanımlayıcı yöntemler olarak iki bölüme ayrılmaktadır.



Şekil 1. Veri Madenciliği Modelleri (Bayram, 2014)

Eğer kesin bir hedef olduğunda ve iyi bir şekilde tanımlanmışsa tahmin edici veya denetimli, elde edilmek istenen değer için belirli bir tanımlama yapılmamışsa bir belirsizlik varsa tanımlayıcı veya denetimsiz yöntemler denilmektedir. Diğer bir deyişle; tanımlayıcı yöntemler verinin ne anlama geldiğini, hangi tür verilerden olduğunu yani veriyi keşfetmektir. Tahmin edici yöntemler ise veri üzerinden bir değer çıkartmaktır.

Veri madenciliği süreci şu şekilde ifade edilmektedir (Bayram ve Dündar, 2021):

- Veriyi anlama süreci

-Verinin hazırlanma süreci

-Verinin modelleme süreci

-Veriyi değerlendirme ve sonuçları kullanma sürecinden oluşmaktadır.

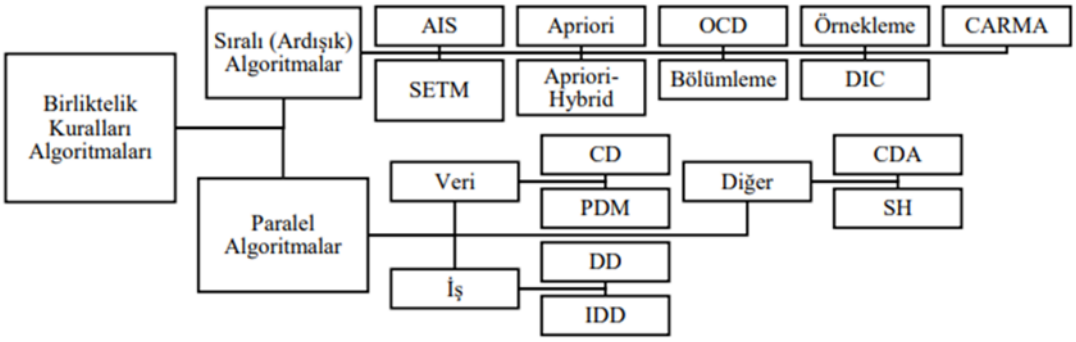
2.1.1 Birliktelik kuralları

Bilgi dediğimiz kavram her alanda gittikçe büyüyerek daha da yaygınlaşmaktadır. Veri tabanlarındaki bilgiler de arttığından firmalar mevcuttaki bilgiler arasındaki ilişkileri belirlemek istemektedir. İlişkiyi analiz etmede veri tabanlarında bulunan verinin diğer verilerle arasındaki ilişkiler karşılaştırılmaktadır. Bu ilişkiler firmalar için karar verme süreçlerinde önemli bir yere sahiptir (Bilgin, 2013).

En yaygın kullanılan veri madenciliği yöntemleri arasında birliktelik kuralları yer almaktadır. Birliktelik kuralları, veri yığınları içerisindeki bilgilerin birlikte görülme olasılıklarının ortaya çıkarılmasında kullanılmaktadır. Yani verilerin potansiyel ilişkilerinin ortaya çıkarılma işlemidir. Birlikte gerçekleşen durumların veya olayların kurallarının belirlenmesini amaçlanmaktadır. Veri tabanlarında depolanan bilgiler pazar analizinin alt yapısının kurulmasında önemli bir yer almaktadır (Öztemiz, 2017) .Sektörde market sepet analizi olarak birliktelik kuralları yöntemleri uygulanmaktadır. Market sepeti analizinde kullanılan birliktelik kuralları ile müşterilerin hangi ürünleri birlikte satın aldıkları belirlenerek alışkanlık analizi yapılmaktadır (Bayram ve Dündar, 2021).

1993 yılında Agrawal ve arkadaşları tarafından ilk defa kullanılan ve çalışmalarında pazar sepetini örnekleyerek birliktelik kurallarının açıklanmasını hedeflemişlerdir. Müşterilerin alışverişlerde aldıkları ürünler arasındaki ilişkiyi saptayarak, müşterilerin yani tüketicilerin alışveriş alışkanlıklarının belirlenmesini amaçlamaktadırlar (Rencüzoğulları, 2019)

Şekil 2'de birliktelik algoritmaları sıralı ve paralel olmak üzere ikiye ayrılmaktadır.



Şekil 2. Birliktelik kuralı algoritmaları (Karademir, 2021)

Paralel algoritmalar, tekrar eden veri kümelerinin belirlenerek, veri kümelerinin paralelleştirilmesi yöntemi olarak büyük öğe kümelerinin keşfedilmesi amaçlanmaktadır. Sıralı algoritmalar ise verilerin kategorisine bağlı olarak veri kümelerinin alfabetik olarak sıralanma işlemidir. Veri kümelerinin belirlenmesi ve sayımı aşamasında basit varsayımlar kullanması en büyük avantajlarındandır (Tunç, 2022) , (Tanyeli, 2021).

2.1.1.1 Birliktelik kuralındaki tanımlamalar

Birliktelik kuralında nesneler küme olacak şeklinde gösterilmektedir. I = { I_1, I_2, I_3,…, I_m} nesnelerin kümesini ifade etmektedir. D ise işlemler kümesidir. Buradaki her m, bir veriyi yani nesneyi göstermektedir. D adı verilen veritabanındaki her taramaya T denildiğinde, nesnelerin kümesi T ⊆ I olacak biçimde kapsamaktadır. Yapılan her adımın tanımı TID olarak isimlendirilmektedir (Bayram, 2014).

Örnek olarak göstermek istenildiğinde; I kümesinde N sayıda nesne olduğunu varsayalım. X ve Y yapılması gereken taramaları temsil ettiğini düşünelim. Birliktelik kuralı aşağıdaki gibi olmaktadır:

$$I = \{ I_1, I_2, I_3,\ldots, I_m \} \quad (2.1)$$

$$X_1, X_2, X,\ldots, X_n \Longrightarrow Y_1, Y_2, Y_3,\ldots, Y_n \quad (2.2)$$

Denkleme göre, X_1, X_2, X_3,…, X_n işlemi yapıldığında aynı zamanda Y_1, Y_2, Y_3,…, Y_n işleminin de yapılmış olduğunu göstermektedir (Rencüzoğulları, 2019).

X ve Y nesnelerin kümesini gösterdiğinden, T'nin alt kümesi X'i kapsamaktadır ve X ⊆ T şeklinde gösterilmektedir. Sonuç olarak X ve Y kümelerinin birliktelik kuralı 'X ⟹ Y' olacak şekilde açıklanmaktadır. Burada X önce, Y sonuç olarak ifade edilmektedir. Dolayısıyla; X ⊆ I, Y ⊆ I ve X ∩ Y = ∅ olacak biçimde gösterilmektedir ve nesneler arasındaki bağlantı ilişkileri %100 doğru sonucu göstermesi beklenmemektedir (Bayram O. , 2014).

Elde bulunan kümelerin çıkarım yapılacak kuralları desteklenmesi gerekmektedir. Bu yüzden X ⇒ Y ilişki kuralları belirlenen destek(D) ve güven (G) eşik değerlerinin üretilmesi gerekmektedir (Bayram, 2014).Bu yargı sonucunda X ve Y kuralı için destek (D) ifadesi (Rencüzoğulları, 2019)

D(X ⇒ Y) =(X ve Y'nin bulunduğu satır sayısı) / (Toplam satır sayısı) (2.3)

şeklinde ifade edilmektedir.

Destek değeri, temsil eden kuralın asgari destek eşik değerini (X ve Y nesnelerini kapsayan işlemlerin toplam işlemlerin içerisinde en az %D oranında olduğunu) ifade etmektedir (Gündüz, 2015)

X nesnesini içeren işlemlerin Y nesnesini de içermesi güven eşik değerini göstermektedir. İlgili kuralın minimum güvenirliğini göstermesine güven eşik değeri denilmektedir. Yani X nesnesini içeren işlemlerin en az %G oranında Y nesnesini de içeren işlemler kümesinde yer aldığını göstermektedir (Gündüz, 2015)

Örneğin, güven eşik değeri %95 olarak belirlenmişse X nesnesinde bir işlemin Y nesnesini de kapsama olasılığı %95 olacaktır anlamına gelmektedir. $X \Longrightarrow Y$ için güven (G) (Rencüzoğulları, 2019):

$$G(X \Longrightarrow Y) = (X \text{ ve } Y\text{'nin bulunduğu satır sayısı}) / (X\text{'in bulunduğu satır sayısı}) \quad (2.4)$$

şeklinde ifade edilmektedir.

Güven değerindeki artış demek X ve Y nesnelerinin birlikte olma olasılığını da arttırmaktadır. Eğer güven değeri %100 ise hareketler tam anlamıyla açıklanabilmektedir (Rencüzoğulları, 2019).

0 ve 1 arasında güven ve destek eşik değerleri değişmektedir ve 1'e yakınlaştıkça aralarındaki güçlü birliktelik de artmaktadır. Bu yüzden aralarındaki ilişki için her iki değer içinde yüksek tutulması gerekmektedir (Bayram, 2014).

Bunların dışında algoritma sonucunda kullanılan başka terimlerde bulunmaktadır (Yüksel ve Zontul, 2019).

Lift Değeri, kuralın ilginçliği veya farklılığını ifade eder. Lift değeri 0 ile ∞ (sonsuz) arasındadır. Yüksek değer, kuralımızın ilginç olduğunu göstermektedir.

$$\text{Lift değeri} = \text{Frekans}(X,Y) / (\text{Frekans}(X) * \text{Frekans}(Y)) \quad (2.5)$$

şeklinde ifade edilmektedir.

Confidence değeri ise 0 ile 1 arasındadır. Yüksek değer kuralımızın çok güvenli veya doğru olduğunu belirtmektedir.

Confidence değeri= (X ⟹ Y) = Destek(X,Y)/ Destek(X) (2.6)

şeklinde ifade edilmektedir.

Birliktelik kuralı oluşturmak için Conviction (Kanaat) ve Leveage (Kaldıraç) değerleri de kullanılmaktadır. Conviction değeri X elemanının, Y elemanı olmaksızın görülme olasılığını ifade etmektedir. Eğer conviction değeri 1 ise X ve Y nesneleri birbirinden bağımsızdır. Conviction değeri 1'den uzak ise ilişki kuralı oluşturulabilmektedir.

Conviction = (X ⟹ Y) = (1- Destek(Y) / (1- Güven(X⟹Y)) (2.7)

şeklinde ifade edilmektedir.

Leverage değeri ise bir satış verisi üzerindeki X ve Y ürünlerinin birlikte satılmasının X ve Y'nin ayrı ayrı satılmasından ne kadar fazla olduğunu belirmektedir.

Leveage = (X⟹ Y) = P(X ve Y) – (P(X)*P(Y)) (2.8)

şeklinde ifade edilmektedir.

2.1.1.2 Apriori algoritması

Birliktelik algoritmasında en yaygın kullanılan teknik olan Apriori algoritmasıdır. Apriori algoritması 1994 yılında Agrawal ve Srikant "Fast Algorithms for Mining Association Rules" adlı çalışmalarında ortaya çıkarmışlardır. Algoritmasının alt yapısında minimum destek ve minimum güven eşik değerleri ele alınarak veri tabanında birliktelik kurallarının belirlenmesini amaçlamaktadır (Çelik, 2019).

Apriori Algoritması, tekrar eden unsurların bulunmasında kullanılan bir algoritmadır ve alt yapısında yinelemeli bir davranış sunmaktadır. Tekrar eden veri kümelerinin belirlenmesinde veri tabanında birden fazla taranma işlemi yapılmaktadır. İlk tarama aşamasında tek elemanlı minimum destek eşit değerini sağlayan tekrar eden veri kümeleri bulunur. Destek eşit değeri, veri kümelerinin veri tabanında kaç satırda olduğunu veya veri tabanının toplam satır sayısının yüzde olarak kaçına denk geldiğini gösterir. Apriori algoritmasının kuralında, k nesne kümesi en düşük destek değerini veriyorsa, bulunan

kümenin alt kümeleri de en düşük destek değerini sağlamaktadır (Bayram ve Dündar, 2021).

Tekrar eden aramalarda bir önceki aramada bulunan tekrar eden veri kümelerine aday kümeler denilmektedir. Bu küme ile yeni kümeler belirlenmek için kullanılmaktadır. Arama sırasında aday kümelerin destek eşik değeri hesaplanmaktadır ve belirlenen minimum destek eşik değerini sağlayan kümeler o tarama esnasında yeni aday küme olmaktadır. Bu süreç yeni tekrar eden veri kümesi buluncaya denk sürmektedir (Alan, 2016).

Büyük sayıdaki veri tabanına sahip verilerin birliktelik kuralları bulunurken iki adımdan oluşmaktadır:

Birinci aşamada, tarama sırasında sık tekrar eden nesneler bulunmaktadır. Bu aşamada her bir nesnenin en az minimum destek sayısı kadar tekrar etmesi gerekmektedir. Minimum destek sayısı veri sayısına ve karar vericiye göre değişebilmektedir (Alan, 2016).

İkinci aşamada sık tekrar eden nesnelerden güçlü birliktelik kuralları belirlenmektedir. Bu kuralların minimum destek ve güven değerlerini sağlaması gerekmektedir. Minimum destek eşik değeri, bir bağlantının hangi oranda tekrar ettiğini ortaya koyarken, güven eşik değeri ise bağlantının doğru olarak sonuçlanmasının olasılığını belirtmektedir (Alan, 2016).

Veri kümesinde olan bir nesnenin kendisinden sonra gelen ve aday olarak belirlenen nesneler ile karşılaştırılması ve aralarındaki ilişkilerin ortaya çıkartılmasını sağlamaktadır. Bununla birlikte bazı aday nesnelerde çıkartılarak nihai olarak elde edilmek istenen sonuca ulaşılmak istenmektedir (Bilgin, 2013).

2.1.1.3 Birliktelik kuralları literatür özeti

Banka sektörüne ait literatürde bazı çalışmalar mevcuttur. Bunlar:

Bayram ve Dündar, 2021 yılında "Türkiye'de Banka Şubesi Lokasyonunun Veri Madenciliği ile Analizi" isimli makalelerinde, Türkiye'deki banka şubelerini birliktelik analizine göre inceleyerek, apriori algoritmasıyla çözmüşlerdir. Çıkan kurallara göre

banka şubelerinin lokasyonlarından kamu bankaları ve İş bankası en güçlü ilişkisi olan bankalar sonucuna varmışlardır (Bayram ve Dündar, 2021).

Meltem Karaatlı, Ece Yıldız (2021) çalışmalarında bamkacılık sektöründe mevduat bankalarına ait finansal oranlarla bankaların finansal yapılarını kümeleme analizi ile incelemişlerdir. Kullandıkları algoritma Beklenti Maksimizasyonu Algoritması'dır. Sonuç olarak finans anlamında benzer bankaları belirlemişlerdir (Karaatlı ve Yıldız, 2021).

İldaş, 2021 yılında "Kredi Riski Ölçüm Modellerinin Değerlendirilmesi" adlı makalesindeki amacı bankacılık sektöründe şirketin maruz kaldığı kredi riskini ölçmeye yönelik uygun bir model belirlemektir. Klasik ve modern kredi riski ölçüm yöntemlerinin yanı sıra portföy kredi riskinin değerlendirilmesi amacıyla uluslararası finansal kuruluşlar tarafından üretilen modellere yer vermiştir (İldaş, 2021).

Karaatlı ve arkadaşları 2021 yılında "Bıst-30 Endeksinde Yer Alan Payların Ortak Hareketlerinin Veri Madenciliği Kapsamında Birliktelik Kuralı İle İncelenmesi" adlı makalelerin de amaçları bir portföy oluşturarak makroekonomik değişkenler ile birlikte hareket eden payları veri madenciliği ile tespit etmişlerdir. Çalışmalarında FP-Growwth algoritmasını kullanarak uygulamalarını WEKA programında göstermişlerdir. Sonuçta Borsa İstanbul 30 endeksindeki payların mali endeks içerisinde bankacılık endeksi paylarının büyük oranda birlikte hareket ettiği sonucuna varmışlardır (Karaatlı vd., 2021).

Özdem ve Akcayol, 2021 çalışmalarında bankacılık sektöründe kampanya yönetimi ve müşterilerin kampanyaya yönelik davranışlarını tahmin etmek için bir model geliştirmişlerdir. Çalışmalarında birliktelik analizinde yararlanmışlarıdır. Sonuç olarak müşteri davranışını etkileyen en önemli davranışın medeni hal ve kredi durumu etkilediğini gözlemlemişlerdir ve müşterilerin kampanyaya katılım tahmininde ise %87 oranında doğru sonuç vermiştir (Özdem ve Akcayol, 2021).

Gürsoy ve Bilgin, 2016 yılında yüksek lisans tezinde banka müşterilerinin internet bankacılığı kullanan müşterilerin davranışlarını birliktelik kuralları ve kümeleme analizi yardımıyla incelemiştir (Gürsoy ve Bilgin, 2016).

Demir, 2018 yılında yapmış olduğu "Veri Madenciliği Yöntemleri ile Banka Çağrı Merkezi Müşterilerine Uygulanan Pazarlama Stratejilerinin Belirlenmesi" adlı yüksek lisans tezinde sınıflandırma yöntemini kullanarak, bankacılık sektöründe ürün kullanımında müşterilerin hangi özelliklere sahip olduğunun belirlemesini amaçlamaktadır (Demir, 2018).

Çalış ve Baynal, 2016 yılında yaptıkları çalışmalarında, bankacılık sektöründe Türkiye'deki bir bankaya ait verilerde müşteri profiline göre satış stratejilerinin geliştirilmesini amaçlamışlardır. Kümeleme yöntemlerinden olan k-ortalamalar yöntemini SPSS yazılında yazarak Clementine programında sonuçları yorumlamışlardır (Çalış ve Baynal, 2016).

Kök ve Aksu, 2013 yılında Muhasebe ve Finansman Dergisi'nde yayınladıkları "Müşteri Kredi Değerliliğinin Belirlenmesinde Analitik Ağ Süreci Kullanımı: Bir Model Önerisi" makalelerinde işletmelerin müşteri kredi değerliliğinin belirlenmesi problemi üzerine durmuşlardır. Çalışmalarında karar verme tekniklerinden Analitik Ağ Süreci (AAS) yaklaşımını ele almışlardır. Sonuçta, müşteri kredi değerleme modeli geliştirmişlerdir. Modele göre mevcut veya potansiyel müşterilerin doğru bir şekilde kredi verilmesi sürecine katkı sağlamışlardır. Geliştirdikleri modelde kredi değerliliğini %92 oranında doğru tahmin ettiği sonucuna varmışlardır (Kök ve Aksu, 2013).

Apriori algoritmasını kullanılarak da literatürde farklı alanlarda yapılan çalışmalar da mevcuttur (Çelik, 2019):

Hanguang ve Yu'nun yaptıkları çalışmada; İzinsiz Giriş Tespit Sistemi'ni geliştirmek için Apriori algoritmasını kullanarak, farklı saldırı biçimlerini algılayarak tespit sisteminin performansını geliştirmek için belirlenen kurallardan faydalanmışlardır.

Keleş ve Kaya, çalışmalarında inşaat sektöründe çalışanların verimliliklerinin ölçümünde belirledikleri faktörlere bağlı olarak edindikleri bilgiler ile Apriori algoritmasıyla verimliliklerinin arttırılması için kuralların ortaya çıkarılmasını amaçlamışlardır.

Tanna ve Ghodasara, çalışmalarında tekrar eden nesnelerin arama esnasında arama sayısını düşürmeyi hedeflemişlerdir.

Doğan vd., sigorta şirketindeki verilerden hareketle Apriori algoritması kullanarak birlikte satın alınan sigorta türlerinin belirlenmesinde kullanmışlardır ve pazarlama analizinde kullanılabilecek kurallar belirlemişlerdir.

Karaibrahimoğlu, hastane verilerine ait göğüs kanseri verilerini kullanarak Apriori algoritması sonucunda verilerin birliktelik ilişkilerinin saptanmasında kullanılmıştır.

Sarkan ve Çataltepe, geçmiş verilerden hareketle şebeke alarm verilerini algoritmaya uyarlayarak otomatik alarm için aralarında ilişki kuralları oluşturmayı amaçlayan yeni bir bakış açısı getirmişlerdir.

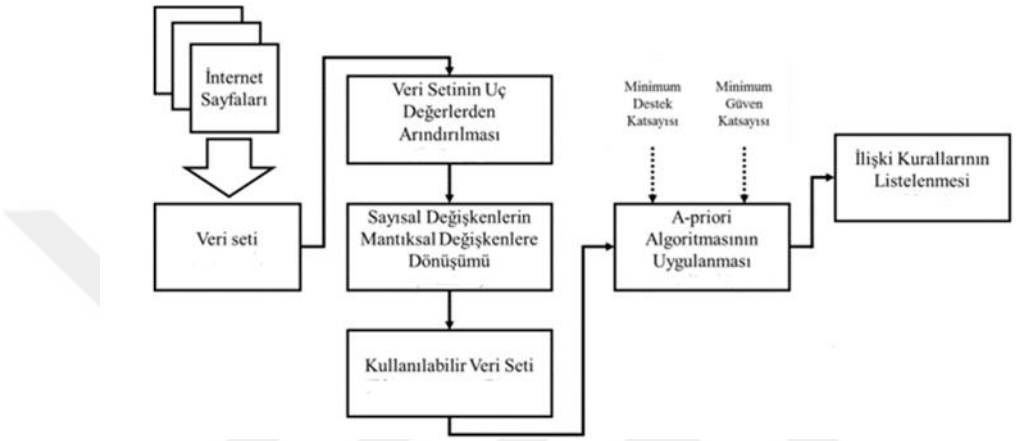
Suryavanshi ve Ingle, tekstil alanındaki verileri kullanarak apriori algoritmasını uygulayarak tekrar eden veri setlerinin ortaya çıkmasını sağlamışlardır. Sonuç olarak veri setini çok fazla taraması sonucunda uygulamalarını kaldırmışlardır.

Doğrul vd., çalışmalarında trafik kaza verilerini kullanarak kazaların en çok hangi durumlarda ve bölgelerde meydana geldiğini belirlemişlerdir.

Bunların dışında uygulanan alanlar da vardır. Öğrencilerin temel özellikleri ile sosyal ağlarındaki aktiflik durumları arasındaki ilişkinin ortaya çıkarılması, pazarlama analistlerinin müşterilerine olan davranışlarının ortaya çıkarılması için kurallar oluşturulması, trafik kazalarından hareketle alınan verileri kullanılarak algoritma sonucuna göre ortaya çıkarılan kurallar ile farklı hava koşullarında, kazanın meydana geldiği yol türüne göre insanların hangi oranda zarar görmediği durumların ortaya çıkarılması, müşterilerin mobilya seçerken hangi kriterleri dikkate aldıklarından hareketle kuralların oluşturulması, hastane ve eczane verilerini kullanarak astım tanısı konan kişilerin hangi hastalıkları bir arada görüldüğünün analizi, ikinci el otomotiv sektöründe bulanan bilgilere göre algoritma sonucunda hangi araç hangi özelliğe sahip gibi kurallar oluşturulması gibi farklı alanlarda apriori algoritması uygulanarak birçok konuya çözüm sağlamıştır.

3. YÖNTEM

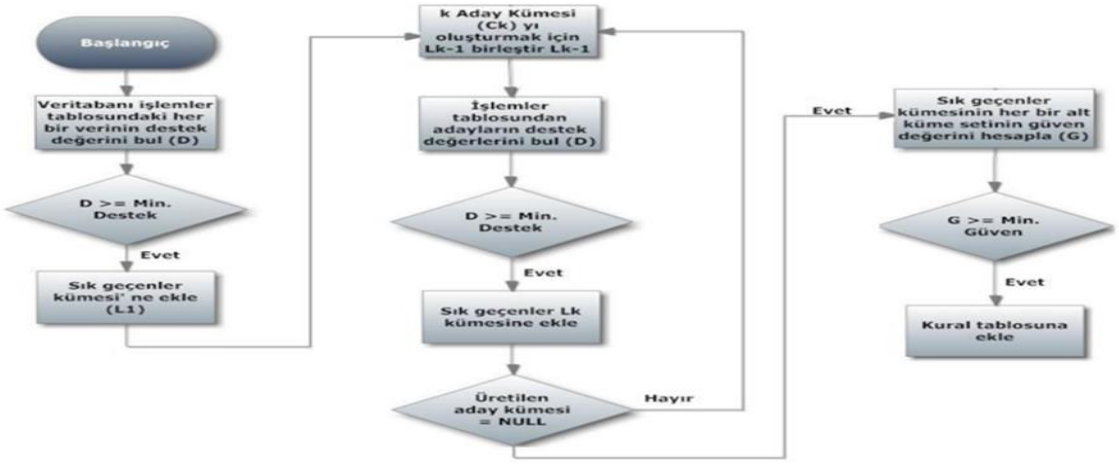
Çalışmanın süreci Şekil 3'te gösterilmiştir. Süreç üç adımdan meydana gelmektedir. Birinci adımda veriler toplanarak analiz etmeye hazır hale getirilmektedir. İkinci adımda, analiz elde edilmektedir. Üçüncü adımda ise çıkarımdaki kurallar belirlenmektedir.



Şekil 3. Çalışmanın süreci

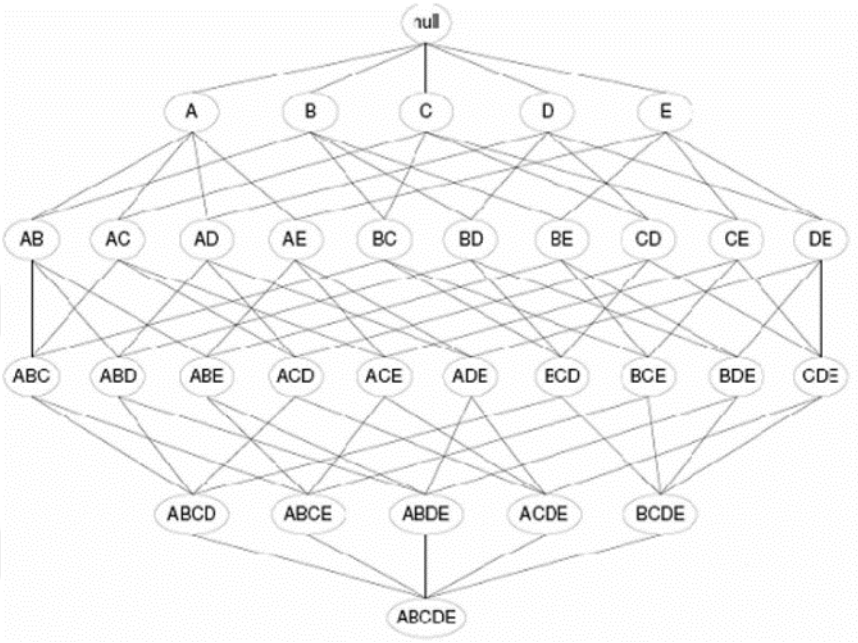
3.1 Araştırma Modeli

Bu çalışmanın çözümü için birliktelik kuralı algoritması yöntemlerinden olan Apriori Algoritması yöntemi kullanılacaktır. Apriori Algoritmasının işleyişinin ayrıntıları Şekil 4'teki gibidir (Alan, 2016):



Şekil 4. Apriori algoritması akış diyagramı (Gülce, 2010)

Büyük veritabanlarında birliktelik kuralları bulunurken, iki aşama ele alınmaktadır. Bunlar: Birinci aşamada, sık tekrarlanan öğeler bulunmaktadır. Bu öğelerin her biri en az önceden belirlenen minimum destek sayısı kadar sık tekrarlanmaktadır. Bu kümeler sık geçen öğe kümesi olarak adlandırılmaktadır. Verilen örneklemde N adet nesne (öğe) söz konusu ise, potansiyel olarak 2^N adet sık geçen öğe kümesi olabilmektedir. Şekil 5'de 2^N adet sık geçen öğe kümesi gösterilmiştir.



Şekil 5. 2^N kümesi Gösterilişi (Bilgin, 2013)

İkinci adımda ise sık tekrarlanan öğelerden güçlü birliktelik kuralları oluşturulmaktadır. Bu kurallar ile minimum destek ve minimum güven değerlerini elde edilmelidir (Bilgin, 2013).

Veri tabanında ilk tarama esnasında fazla sayıda nesne kümeleri belirlemek için veri setinin tamamı sayılmaktadır. Bundan sonraki tarama işlemleri için k'ıncı tarama iki aşamadan meydana gelmektedir. Birinci aşamada, Apriori-gen fonksiyonu uygulanarak (k-1)'inci taramada varılan, $L_k - 1$ nesne kümeleriyle, C_k aday nesne kümeleri oluşturulmaktadır. İkinci adımda ise veritabanı taranarak, C_k'daki aday nesne kümelerin destek eşik değeri sayılmaktadır. Şekil 6'da apriori algoritmasının matematiksel modeli gösterilmiştir.

```
1)  L1= {sık geçen 1-öğe kümesi};
2)  for (k=2; Lk-1≠ Ø; k++) do begin
3)      Ck=apriori-gen (Lk-1); // Yeni adaylar
4)      forall transactions-hareketler t ∈ D do begin
5)              Ct = subset (Ck, t); // Adaylar t içindedir
6)              forall candidates – adaylar c ∈ Ct do
7)                      c.count++;
8)      end
9)      Lk = {c ∈ Ck | c.count ≥ minsup}
10) end
11) Answer = ∪k Lk;
```

Şekil 6. Apriori algoritmasının matematiksel modeli (Gülce, 2010)

3.2 Evren ve Örneklem

Çalışma için banka sektöründe müşteri bilgilerine ait verilere ihtiyaç duyulmaktadır. Bu çalışmada ilk adımda tüm verileri kapsayan yani evreni oluşturacak veri seti kümesi ele alınarak Oracle SQL'de veri analizi yapılmıştır. Tüm veri kümesi 438.557 adet müşteri verisinden oluşmaktadır. Bu veri kümesindeki bazı nitelikler çalışma için daha az önemli olduğundan veri setinden çıkarılmıştır. Çıkartılan niteliklerden bazıları şunlardır; müşteri numarası, müşterinin telefonu olup olmaması, müşterinin iş telefonunun olup olmaması, müşterinin mesleği, müşterinin e-maili olup olmaması, müşterinin doğum günü gibi bilgiler veriden çıkartılmıştır. Algoritma için boş olan değerler olmaması adına herhangi bir alanında boş değer olan veri satırları veri kümesinden çıkartılmıştır. Bu çalışma için sadece 1 aylık bir veri kümesi kullanılacaktır. Sonuç olarak, tüm veri setinden yapılan analizler sonucunda bu çalışma için kullanılacak veri kümesi yani örneklemi temsil edecek veri kümesi 25.134 adet veriden oluşmaktadır.

3.3 Veriler ve Toplanması

Toplam veri seti 25.134 adet veriden oluşmaktadır. Bu veri seti için rastgele seçilen 10 müşteriye ait bilgiler ve her niteliği belirten değerler aşağıdaki Tablo 1'de gösterilmektedir.

Tablo 1. Rastgele seçilen 10 müşteriye ait müşteri bilgileri

Müşteri sayısı	Cinsiyet	Arabası var mı?	Mülkü var mı?	Çocuk sayısı	Yıllık gelir tutarı	Eğitim durumu	Medeni durumu	Yaş durumu	Kredi durumu
1	Kadın	Arabası var	Mülkü var	0 çocuk	350.000-450.000 TL	Lisans	Bekar	Genç	Kredi yok
2	Erkek	Arabası var	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Genç	Ödendi
3	Erkek	Arabası var	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi 3 ay geçen borçlar
4	Kadın	Arabası var	Mülkü yok	2 çocuk ve üzeri	200.000-350.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi 6 ay geçen borçlar
5	Kadın	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
6	Erkek	Arabası yok	Mülkü var	1 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Evli	Genç	Kredi yok
7	Erkek	Arabası var	Mülkü yok	0 çocuk	550.000 TL ve üzeri	Lisans	Evli	Yetişkin	Ödendi
8	Kadın	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	Lisans	Evli	Yetişkin	Ödendi
9	Erkek	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	Akademik derece	Evli	Yetişkin	Vadesi 3 ay geçen borçlar
10	Kadın	Arabası yok	Mülkü var	0 çocuk	200.000-350.000 TL	İlköğretim/Ortaöğretim	Evli	Yaşlı	Ödendi

Çalışma için banka sektöründe müşteri bilgilerine ait verilere ihtiyaç duyulmaktadır. Bu veri seti piyasada hazır olarak bulunmaktadır. Kaggle'nin sunmuş olduğu veri kümesinden yararlanılmıştır. Kaggle, veri bilimi alanında yarışmalar düzenleyen Google tarafından oluşturulmuş bir web sitesidir. Veri seti içerinde Tablo 1'den hareketle bir müşterinin cinsiyeti, arabası olup olmama durumu, mülkü olup olmama durumu, çocuk sayısı, yıllık gelir tutarı, eğitim durumu, medeni durumu, yaş durumuna göre kredi durumları belirtilmiştir. Algoritmaya başlamadan önce değerler string ifadeye çevrilmiştir. Veri kümesi içerisindeki tüm durumlar ve sayıları aşağıdaki Tablo 2'de gösterilmiştir.

Tablo 2. Veri kümesi içerisindeki tüm durumlar ve frekansları

Cinsiyet		Arabası var mı?		Mülkü var mı?		Çocuk sayısı		Yıllık gelir tutarı		Eğitim durumu		Medeni durumu		Yaş durumu		Kredi durumu	
Kadın	15630	Var	10516	Var	16461	0 çocuk	15908	200.000-350.000 TL	1012	İlköğretim/Ortaöğretim	16995	Evli	19642	Genç	3436	Kredi yok	6699
Erkek	9504	Yok	14618	Yok	8673	1 çocuk	6118	350.000-450.000 TL	8524	Lisans	8125	Bekar	5492	Yetişkin	20953	Ödendi	12319
						2 çocuk ve üzeri	3108	450.000-550.000 TL	15267	Akademik derece	14			Yaşlı	745	Vadesi 3 ay geçen borçlar	6050
								550.000 TL ve üzeri	331							Vadesi 6 ay geçen borçlar	7
																Vadesi geçmiş veya kötü borçlar	59
Toplam	25134		25134		25134		25134		25134		25134		25134		25134		25134

Örneğin, tüm veri seti içerisinde kadın müşteri sayısı 15.630 adet, erkek müşteri sayısı 9.504 adet, bekar kişilerin sayısı ise 5.492 adettir. Kredi kartı ödeme durumlarında ise tüm veri setinde vadesini geçirmiş kötü borca sahip kişilerin sayısı 59 olurken kredi kartını zamanında ödeyen kişilerin sayısı 12.319 adettir.

Veri kümesini analiz etmek için ilk aşamada Oracle-SQL veri tabanından yararlanılmıştır. Veritabanı dili olarak en sık kullanılan dil SQL'dir. SQL, veriler üzerinde farklı işlemler gerçekleştirerek veritabanlarını yönetir. Yapılandırılmış sorgu dili olarak adlandırılmaktadır. İlişkisel veritabanı yönetim sistemlerindeki verileri paylaşmak ve verileri yönetmek için kullanılmaktadır. Verileri yeniden düzenleyebilir, güncelleyebilir,

sorgulayabilir ve veritabanının şemasını oluşturarak verilerin erişimini sağlayabilmektedir.

Veritabanında yapılan analizde, SQL kodları örnek olarak şu şekilde açıklanmaktadır:

Veri tabanında TBLBCK.CUS_DATA adında bir tablo yaratılarak tablonun içerisine verikümesi aktarılmıştır. Veri kümesinin çağırılma işlemi;

*SELECT * FROM TBLBCK.CUS_DATA* şeklindedir.

Algoritma için belirlenen bazı sütun bilgileri kullanılmıştır. Kullanılan sütunlar veya nitelikler şunlardır:

SELECT

ID "MÜŞTERİ NUMARASI",

CODE_GENDER "CİNSİYET",

FLAG_OWN_CAR "ARABASI VAR MI?",

FLAG_OWN_REALTY "MÜLKÜ VAR MI?",

CNT_CHILDREN "ÇOCUK SAYISI",

AMT_INCOME_TOTAL "YILLIK GELİR TUTARI",

NAME_EDUCATION_TYPE "EĞİTİM DURUMU",

NAME_FAMILY_STATUS "MEDENİ DURUMU",

AGE "YAŞ DURUMU",

STATUS "KREDİ DURUMU"

FROM TBLBCK.CUS_DATA;

Veri kümesindeki her bir değer için string ifadeye çevrilerek algoritmada kullanılmıştır. Örneğin;

BEGİN

UPDATE TBLBCK.CUS_DATA SET FLAG_OWN_REALTY = 'MÜLKÜ VAR' WHERE FLAG_OWN_REALTY= '1';

UPDATE TBLBCK.CUS_DATA SET FLAG_OWN_REALTY = 'MÜLKÜ YOK' WHERE FLAG_OWN_REALTY = '0';

END;

Burada ‘Mülkü var mı?’ sütunundaki nitelikte kullanılan 1, mülkü var olarak 0 ise mülkü yok olarak veritabanında değiştirilmiştir. Bunun yanında veri kümesi algoritma için belli kategorilerde kategorize edilerek kullanılmıştır. Örneğin;

BEGİN

UPDATE TBLBCK.CUS_DATA SET AGE='GENÇ' WHERE AGE İN ('21','22','23','24','25','26', '27','28','29');

UPDATE TBLBCK.CUS_DATA SET AGE='YETİŞKİN' WHERE AGE İN ('30','31','32','33','34','35','36','37','38','39','40','41','42','43','44','45','46','47','48','49','50', '51','52','53','54','55','56','57','58','59');

UPDATE TBLBCK.CUS_DATA SET AGE='YAŞLI' WHERE AGE İN ('60','61','62','63','64','65','66','67','68');

END;

Burada müşterilerinin yaş bilgileri için yaşı 21-29 olanlar için “Genç”, yaşı 30-59 arasında olanlar için “Yetişkin”, yaşı 60 ve üzeri olanlar için ise “Yaşlı” olacak şekilde kategorize edilerek veri tabanında güncellenmiştir. Bunun yanında diğer bir örnek de yıllık gelir tutarları içinde algoritma için belli aralıklar belirlenerek kategorize edilmiştir.

Tüm veri seti içerisindeki kadın sayısını öğrenmek için;

SELECT COUNT(*) FROM TBLBCK.CUS_DATA WHERE CODE_GENDER='KADIN';

Farklı kredi kartı ödeme durumlarını öğrenmek için;

SELECT DİSTİNCT STATUS FROM TBLBCK.CUS_DATA;

Tüm veri setindeki kredi kartı ödeme durumu vadesi geçmiş ve kötü borca sahip kişilerin bilgilerini öğrenmek için;

SELECT ID "MÜŞTERİ NUMARASI",

CODE_GENDER "CİNSİYET",

FLAG_OWN_CAR "ARABASI VAR MI?",

FLAG_OWN_REALTY "MÜLKÜ VAR MI?",

CNT_CHILDREN "ÇOCUK SAYISI",

AMT_INCOME_TOTAL "YILLIK GELİR TUTARI",

NAME_EDUCATION_TYPE "EĞİTİM DURUMU",

NAME_FAMILY_STATUS "MEDENİ DURUMU",

AGE "YAŞ DURUMU",

STATUS "KREDİ DURUMU" FROM TBLBCK.CUS_DATA WHERE STATUS='VADESİ GEÇMİŞ VEYA KÖTÜ BORÇLAR';

Sonucu Şekil 7'de gösterilmiştir:

MÜŞTERİ NUMARASI	CİNSİYET	ARABASI VAR MI?	MÜLKÜ VAR MI?	ÇOCUK SAYISI	YILLIK GELİR TUTARI	EĞİTİM DURUMU	MEDENİ DURUMU	YAŞ DURUMU	KREDİ DURUMU
5021430	Kadın	Arabası yok	Mülkü var	0 çocuk	350.000-450.000 TL	Lisans	Bekar	Yetişkin	Vadesi geçmiş veya kötü borçlar
5033514	Kadın	Arabası var	Mülkü yok	1 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5038807	Erkek	Arabası var	Mülkü var	0 çocuk	450.000-550.000 TL	Lisans	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5069475	Kadın	Arabası yok	Mülkü var	2 çocuk ve üzeri	450.000-550.000 TL	Lisans	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5065518	Kadın	Arabası var	Mülkü yok	1 çocuk	450.000-550.000 TL	Lisans	Evli	Genç	Vadesi geçmiş veya kötü borçlar
5085886	Erkek	Arabası var	Mülkü var	0 çocuk	450.000-550.000 TL	Lisans	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5078801	Kadın	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5078805	Kadın	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi geçmiş veya kötü borçlar
5090737	Kadın	Arabası yok	Mülkü yok	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Bekar	Yetişkin	Vadesi geçmiş veya kötü borçlar

Şekil 7. Veri setindeki kredi kartı durumu vadesi geçmiş ve kötü borca sahip kişilerin bilgileri

Tüm müşteriler içinde kredi kartını zamanında ödeyen, cinsiyeti kadın olup, yıllık geliri 350.000-450.000 TL olan kişi sayısını öğrenmek için;

SELECT COUNT(*) SAYISI FROM TBLBCK.CUS_DATA WHERE STATUS='ÖDENDİ' AND CODE_GENDER ='KADIN' AND AMT_INCOME_TOTAL ='350.000-450.000 TL'

şeklinde ifade edilmektedir.

3.4 Verilerin Çözümlenmesi ve Yorumlanması

Analiz yapılabilmesi için minimum destek ve güven değerlerinin belirlenmesinden sonra bu değerlere eşit ya da büyük olacak şekilde algoritmaya devam edilir.

Algoritmayı uygulamadan önce destek ve güven değerlerini bir örnekle açıklamak istersek;

Tablo 3. Kredi durumu vadesi 6 ay geçen borca sahip müşterilerin bilgileri

Müşteri sayısı	Cinsiyet	Arabası var mı?	Mülkü var mı?	Çocuk sayısı	Yıllık gelir tutarı	Eğitim durumu	Medeni durumu	Yaş durumu	Kredi durumu
1	Erkek	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	Lisans	Bekar	Genç	Vadesi 6 ay geçen borçlar
2	Erkek	Arabası var	Mülkü yok	1 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Evli	Genç	Vadesi 6 ay geçen borçlar
3	Erkek	Arabası var	Mülkü yok	1 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Bekar	Yetişkin	Vadesi 6 ay geçen borçlar
4	Erkek	Arabası yok	Mülkü yok	0 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Bekar	Genç	Vadesi 6 ay geçen borçlar
5	Kadın	Arabası var	Mülkü yok	2 çocuk ve üzeri	200.000-350.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi 6 ay geçen borçlar
6	Erkek	Arabası yok	Mülkü var	0 çocuk	450.000-550.000 TL	İlköğretim/Ortaöğretim	Evli	Yetişkin	Vadesi 6 ay geçen borçlar
7	Erkek	Arabası var	Mülkü yok	1 çocuk	350.000-450.000 TL	İlköğretim/Ortaöğretim	Bekar	Yetişkin	Vadesi 6 ay geçen borçlar

Yukarıdaki Tablo 3'den yola çıkarak toplam müşteri sayısından hareketle vadesi 6 ay geçen borçlar için {Erkek, Mülkü yok} ile yıllık gelir tutarı (350.000-450.000 TL) olması arasındaki ilişki şu şekilde açıklanabilir:

Destek değeri= ({Erkek, Mülkü yok, Yıllık gelir tutarı (350.000-450.000 TL)})/(Toplam müşteri sayısı)=4/7=0.57

Güven değeri= ({ Erkek, Mülkü yok, Yıllık gelir tutarı (350.000-450.000 TL) })/({ Erkek, Mülkü yok })=4/4=1

Bu sonuçlardan yola çıkarak {Erkek, Mülkü yok}⇒ Yıllık gelir tutarı (350.000-450.000 TL) kuralı %57 destek, %100 güven eşik değerlerine sahiptir diyebiliriz.

Sonraki aşamada ise veri seti tarama işlemi gerçekleşmektedir ve her nesnenin frekans sayıları belirlenmektedir. Eğer destek değerinden küçük ise çözümden çıkartılmaktadır.

Örneğin, Tablo 3'den hareketle bazı hesaplamaları yapacak olursak;

Destek parametre değerinin %70'i ve Güven parametre değerinin %70 olarak belirlendiğinde;

Destek Değeri: 0,70 ve Güven Değeri: 0,70 ve N:7 (Toplam müşteri sayısı)

Destek Sayısı: 0,70 x 7 = 4,9'tür. Burada 4,9 ve üzeri değerler ile algoritmaya devam edilir.

Frekanslar hesaplandıktan sonra nitelik kümesi belirlenmektedir. Nitelik kümesini Tablo 3'den açıklamak istersek:

Nesne elemanları = {Erkek, Kadın, Arabası yok, Arabası var, Mülkü yok, Mülkü var, 0 çocuk, 1 çocuk, 2 çocuk ve üzeri, 200.000-350.000 TL, 350.000-450.000 TL, 450.000-550.000 TL, İlköğretim/Ortaöğretim, Lisans, Bekar, Evli, Genç, Yetişkin} dan yola çıkarak 2^18=262.144 adet olası aday nitelik kümesi bulunmaktadır. Devamında ise güven değeri üzerinden birliktelik kuralları oluşturulmaktadır. Bu işlemin ardından nesneler ikişerli olarak gruplanarak frekans değerleri belirlenir ve aynı şekilde üçlü, dörtlü vb. olacak şekilde işleme devam edilir. Destek sayısından küçük olanlar işlemden çıkartılır. Frekans sayıları Tablo 4'de gösterilmiştir.

Tablo 4. Müşteri bilgileri ve frekans sayıları

No	**Müşteri Bilgileri**	**Frekans**
1	Erkek	6
2	İlköğretim/Ortaöğretim	6
3	Mülkü yok	5
4	Arabası var	4
5	350.000-450.000 TL	4
6	Bekar	4
7	Yetişkin	4
8	Arabası yok	3
9	0 çocuk	3
10	1 çocuk	3

11	Evli	3
12	Genç	3
13	Mülkü var	2
14	450.000-550.000 TL	2
15	Kadın	1
16	2 çocuk ve üzeri	1
17	200.000-350.000 TL	1
18	Lisans	1

Tablo 4'de frekans değeri 4,9 ve üzeri olan çözümlemeler ile devam edilir. Çünkü minimum destek değerinin altında olduğu için listeden çıkarılmıştır. Sonrasında 2'li, 3'lü ve artarak diğer gruplandırmalar yapılır.

Tablo 5. Destek değerinin üzerinde olan müşteri bilgileri ve frekansları

No	**Müşteri Bilgileri**	**Frekans**
1	Erkek	6
2	İlköğretim/Ortaöğretim	6
3	Mülkü yok	5

Sonrasında 2'li gruplama yapılarak Tablo 6'da gösterilmiştir.

Tablo 6. İkili gruplandırma sonucunda müşteri bilgileri ve frekansları

No	**Müşteri Bilgileri**	**Frekans**
1	Erkek, İlköğretim/Ortaöğretim	5
2	İlköğretim/Ortaöğretim, Mülkü yok	5
3	Erkek, Mülkü yok	4

Tablo 7. Destek değerinin üzerinde olan müşteri bilgileri ve frekansları

No	**Müşteri Bilgileri**	**Frekans**
1	Erkek, İlköğretim/Ortaöğretim	5
2	İlköğretim/Ortaöğretim, Mülkü yok	5

Sonrasında 3'lü gruplama yapılarak Tablo 8'de gösterilmiştir.

Tablo 8. Üçlü gruplandırma sonucunda müşteri bilgileri ve frekansları

No	Müşteri Bilgileri	Frekans
1	Erkek, İlköğretim/Ortaöğretim, Mülkü yok	4

Tablo 8'de minimum destek değerinin altında olduğu için listeden çıkarılmıştır.

Tekli, ikili ve üçlü nitelik listesi Tablo 9'da gösterilmiştir.

Tablo 9. Tüm gruplandırma sonucunda müşteri bilgileri ve frekansları

No	Müşteri Bilgileri	Frekans
1	Erkek	6
2	İlköğretim/Ortaöğretim	6
3	Mülkü yok	5
4	Erkek, İlköğretim/Ortaöğretim	5
5	İlköğretim/Ortaöğretim, Mülkü yok	5

Tablo 9'da örneğin erkek ile ilköğretim/ortaöğretim arasındaki destek ve güven değeri hesaplarsak:

Destek (Erkek -> İlköğretim/Ortaöğretim): Frekans(Erkek, İlköğretim/Ortaöğretim) / N

Destek: 5/7: 0,714

Güven(Erkek -> İlköğretim/Ortaöğretim): Frekans(Erkek, İlköğretim/Ortaöğretim) / Frekans(Erkek)

Güven: 5/6 = 0,833 yani %83,3 olduğu görülmüş olup Minimum Güven değeri olan %70'ın üzerindedir.

Sonuç olarak; cinsiyeti erkek olup, eğitim seviyesi ilköğretim/ortaöğretim olması halinde tüm müşterilerin arasında %71,4'ında görülmektedir. Yüzde 83,33 doğruluk ile erkek müşterilerin eğitim durumu ilköğretim/ ortaöğretim den mezun olmuştur.

Apriori algoritması, bu çalışmada Python Jupyter Notebook platformu üzerinde gerçekleştirilmiştir.

Ptyhon, her alanda sınırsız sayıda modüle sahiptir yani çok zengin bir kütüphanesi vardır. Pyhton yorumlamaya dayalı bir dildir. Makine öğrenmesi uygulamaları için kullanılan şirketlerin en çok talep edilen bir dildir. Jupyter Notebook gibi web tabanlı notebooklar ile tarayıcı üzerinden programlama yapılabilmektedir. Jupyter Notebook, çeşitli programlama dilleri için etkileşimli bir ortam sağlayan açık kaynak kodlu bir programdır. Bu notebook ile artık veri bilimi çalışmaları için gerekli ortama sahiptir.

Algoritmada kullanılacak olan destek ve güven değerleri için; bu değerlerin yüksek olası az sayıda kuralın ortaya çıkmasını sağlarken, çok düşük tutulması da çok fazla kuralın oluşmasına neden olmaktadır. Bu değerler kural sayısını kontrol etmektedir. Deneme yanılma yoluyla destek ve güven değerleri belirlenmiştir.

Örneğin, Tablo 3'den hareketle asgari destek değeri %70 olarak belirlendiğinde sadece 5 adet kuralın oluşturulduğu, %50 olarak belirlendiğinde ise 23 tane kuralın oluştuğu gözlemlenmiştir. Asgari destek değeri %70 ve asgari güven değeri %70 olarak belirlendiğinde elde edilen kural sayısı bu örnek için uygun olarak değerlendirilmiştir.

Tablo 10'da kredi kartı borcunu zamanında ödeyen müşterilerin Python Jupyter Notebook uygulaması üzerinde çözülen algoritmanın kodları şu şekilde çözümlenmiştir:

Tablo 10. Kredi kartı borcunu zamanında ödeyen müşterilerin Python Jupyter Notebook uygulaması üzerinde algoritmanın çözümü

In[1]	#Veri kümesinin çağırılma işlemi:
	import pandas as pd #kütüphanenin çağırılması
	dataset = pd.read_csv('Ödendi.csv')
	dataset

Out[1]	Müşteri Bilgileri 0 Kadın,Arabası yok,Mülkü var,2 çocuk ve üzeri,3... 1 Kadın,Arabası yok,Mülkü var,2 çocuk ve üzeri,3... 2 Kadın,Arabası yok,Mülkü var,2 çocuk ve üzeri,3... 3 Kadın,Arabası yok,Mülkü var,2 çocuk ve üzeri,3... 4 Kadın,Arabası yok,Mülkü var,2 çocuk ve üzeri,3... ... ... 12314 Kadın,Arabası var,Mülkü var,1 çocuk,450.000-55... 12315 Kadın,Arabası yok,Mülkü var,0 çocuk,450.000-55... 12316 Kadın,Arabası yok,Mülkü var,1 çocuk,450.000-55... 12317 Kadın,Arabası yok,Mülkü var,0 çocuk,450.000-55... 12318 Kadın,Arabası yok,Mülkü var,0 çocuk,450.000-55... 12319 rows × 1 columns
In[2]	#Veri kümesi satır sayısı: dataset.shape
Out[2]	(12319, 1)

In[3]	#Her müşterinin bilgilerini ayırma işlemi:
	data = list(dataset["Müşteri Bilgileri"].apply(lambda x:x.split(',')))
	data

Out[3]

```
[['Kadın',
  'Arabası yok',
  'Mülkü var',
  '2 çocuk ve üzeri',
  '350.000-450.000 TL',
  'Lisans',
  'Evli',
  'Yetişkin'],
 ['Kadın',
  'Arabası yok',
  'Mülkü var',
  '2 çocuk ve üzeri',
  '350.000-450.000 TL',
  'Lisans',
  'Evli',
  'Yetişkin'],....
```

In[4]

```
#Veri Ön İşleme
#String ifadesindeki müşteri bilgilerinin mantıksal yapıya çevrilmesi:
#!pip install mlxtend
from mlxtend.preprocessing import TransactionEncoder

te = TransactionEncoder()
te_data = te.fit(data).transform(data)
df = pd.DataFrame(te_data,columns=te.columns_)
df
```

Out[4]

	0 çocuk	1 çocuk	2 çocuk ve üzeri	200.000-350.000 TL	350.000-450.000 TL	450.000-550.000 TL	550.000 TL ve üzeri	Akademik derece	Arabası var	Arabası yok	…
0	False	False	True	False	True	False	False	False	False	True	…
1	False	False	True	False	True	False	False	False	False	True	
…	…	…	…	…	…	…	…	…	…	…	…

In[5]

```
#Veri İlişkilendirme Kuralları
#Deneme yanılma yolu ile belirlenen minimum destek eşik değerinin uygulanması:
from mlxtend.frequent_patterns import apriori
ruless = apriori(df,min_support=0.5, use_colnames=True,verbose = 1)
ruless
```

Out[5]

	support	itemsets
0	0.639581	(0 çocuk)
1	0.601916	(450.000-550.000 TL)
2	0.580323	(Arabası yok)
3	0.784398	(Evli)
4	0.622534	(Kadın)
5	0.647049	(Mülkü var)
6	0.831967	(Yetişkin)
7	0.679195	(İlköğretim/Ortaöğretim)
8	0.517331	(Yetişkin, 0 çocuk)
9	0.507996	(450.000-550.000 TL, Yetişkin)
10	0.508077	(Evli, Mülkü var)
11	0.665151	(Evli, Yetişkin)
12	0.535352	(Evli, İlköğretim/Ortaöğretim)
13	0.537219	(Kadın, Yetişkin)
14	0.544768	(Mülkü var, Yetişkin)
15	0.580404	(İlköğretim/Ortaöğretim, Yetişkin)

In[6]

```
#Oluşan kural sayısını bulma işlemi:
ruless.shape
```

Out[6]

```
(16, 2)
```

In[7]

```
# Apriori Fonksiyonunu
# df = True-False array ile dönüştürülmüş dataframe
# min_support = Belirlenen minimum destek eşik değerinin üzerindeki bilgileri getirmek için
# use_colnames = True, sütun isimlerini gösterir.
# verbose = 1, Toplam kombinasyon sayısını verir.
df1 = apriori(df, min_support=0.5, use_colnames=True, verbose = 1)
df1
```

Out[7]

	support	itemsets
0	0.639581	(0 çocuk)
1	0.601916	(450.000-550.000 TL)
2	0.580323	(Arabası yok)
3	0.784398	(Evli)
4	0.622534	(Kadın)
5	0.647049	(Mülkü var)
6	0.831967	(Yetişkin)
7	0.679195	(İlköğretim/Ortaöğretim)
8	0.517331	(Yetişkin, 0 çocuk)
9	0.507996	(450.000-550.000 TL, Yetişkin)
10	0.508077	(Evli, Mülkü var)
11	0.665151	(Evli, Yetişkin)
12	0.535352	(Evli, İlköğretim/Ortaöğretim)
13	0.537219	(Kadın, Yetişkin)
14	0.544768	(Mülkü var, Yetişkin)
15	0.580404	(İlköğretim/Ortaöğretim, Yetişkin)

In[8]

```
#Çıkan ilişkileri destek sayısına göre büyükten küçüğe göre sıralama işlemi:
df1.sort_values(by="support", ascending=False)
```

Out[8]

	support	itemsets
6	0.831967	(Yetişkin)
3	0.784398	(Evli)
7	0.679195	(İlköğretim/Ortaöğretim)
11	0.665151	(Evli, Yetişkin)
5	0.647049	(Mülkü var)
0	0.639581	(0 çocuk)
4	0.622534	(Kadın)
1	0.601916	(450.000-550.000 TL)
15	0.580404	(İlköğretim/Ortaöğretim, Yetişkin)
2	0.580323	(Arabası yok)
14	0.544768	(Mülkü var, Yetişkin)
13	0.537219	(Kadın, Yetişkin)
12	0.535352	(Evli, İlköğretim/Ortaöğretim)
8	0.517331	(Yetişkin, 0 çocuk)
10	0.508077	(Evli, Mülkü var)
9	0.507996	(450.000-550.000 TL, Yetişkin)

In[9]

```
from mlxtend.frequent_patterns import association_rules
# Aprioride "support" hesabı yapabilmektedir, "confidence(güven) ve diğerleri" için association rules kullanılmaktadır
#Deneme yanılma yoluyla belirlenen minimum güven değeri bu çalışma için %70 dir. Yüksek tutulması önemlidir.
association_rules(df1, metric = "confidence", min_threshold = 0.7)
```

In[10]

```
#Sonuçları Seçme ve Filtreleme
#minimum destek eşik değeri %50, minimum güven eşik değeri %70 ile kurallar listenir.
rules = association_rules(df1, metric = "confidence", min_threshold = 0.7)
rules[ (rules['confidence'] >= 0.7) & (rules['support'] >= 0.5) ]
```

Out[10]

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(0 çocuk)	(Yetişkin)	0.639581	0.831967	0.517331	0.808859	0.972225	-0.014779	0.879106
1	(450.000-550.000 TL)	(Yetişkin)	0.601916	0.831967	0.507996	0.843965	1.014421	0.007222	1.076893
2	(Mülkü var)	(Evli)	0.647049	0.784398	0.508077	0.785221	1.001050	0.000533	1.003833
3	(Evli)	(Yetişkin)	0.784398	0.831967	0.665151	0.847977	1.019243	0.012558	1.105312
4	(Yetişkin)	(Evli)	0.831967	0.784398	0.665151	0.799493	1.019243	0.012558	1.075282
5	(İlköğretim/Ortaöğretim)	(Evli)	0.679195	0.784398	0.535352	0.788216	1.004867	0.002593	1.018026
6	(Kadın)	(Yetişkin)	0.622534	0.831967	0.537219	0.862955	1.037247	0.019291	1.226114
7	(Mülkü var)	(Yetişkin)	0.647049	0.831967	0.544768	0.841927	1.011972	0.006445	1.063010
8	(İlköğretim/Ortaöğretim)	(Yetişkin)	0.679195	0.831967	0.580404	0.854548	1.027141	0.015337	1.155245

Diğer kredi kartı ödeme vadeleri olan; kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip kişiler, kredi kartı borcunun vadesi 6 ay geçen kişiler, kredi kartı borcunun

vadesi 3 ay geçen kişiler ve kredi kartına sahip olmayan kişiler içinde Tablo 10'daki gibi Python Jupyter Notebook uygulaması üzerinde apriori algoritması ayrı ayrı çözülerek gerçekleştirilmiştir ve çıkan sonuçlar bulgular kısmında sunulmuştur. Bunun yanında veri seti Power BI platfromunda görselleştirilmiştir. Power BI, tüm verileri tek bir yerde görselleştirmek ve analiz etmek için kullanılan bir iş analizi hizmetidir yani Power BI verileri analiz eden ve yorumlayan bir iş analitiği paketidir.

4. BULGULAR VE TARTIŞMA

4.1 Bulgular

Bu çalışmada Apriori algoritması kullanılarak kredi kartı ödeme durumlarına etki eden müşteri bilgilerine ilişkin kurallar çıkartılmıştır. Elde edilen kurallar farklı müşterilerin hangi nitelikler etkisinde kredi kartı ödeme durumlarını etkilediği hususunda bilgi vermektedir.

Kredi durumları şu şekildedir:

1) *Kredi kartını zamanında ödeyen müşteriler*
2) *Kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip kişiler*
3) *Kredi kartı borcunun vadesi 6 ay geçen kişiler*
4) *Kredi kartı borcunun vadesi 3 ay geçen kişiler*
5) *Kredi kartına sahip olmayan kişiler*

1) Kredi kartını zamanında ödeyen müşteriler için:

Kredi kartı zamanında ödeyen 12319 adet müşteri için; %50 destek değeri, %70 güven değeri belirlenmiştir.

Tablo 11. Tüm sık geçen müşteri bilgileri

Kural Sayısı	Destek	Nesneler
1	0.831967	(Yetişkin)
2	0.784398	(Evli)
3	0.679195	(İlköğretim/Ortaöğretim)
4	0.665151	(Evli, Yetişkin)
5	0.647049	(Mülkü var)
6	0.639581	(0 çocuk)
7	0.622534	(Kadın)
8	0.601916	(450.000-550.000 TL)
9	0.580404	(İlköğretim/Ortaöğretim, Yetişkin)
10	0.580323	(Arabası yok)
11	0.544768	(Mülkü var, Yetişkin)
12	0.537219	(Kadın, Yetişkin)
13	0.535352	(Evli, İlköğretim/Ortaöğretim)

14	0.517331	(Yetişkin, 0 çocuk)
15	0.508077	(Evli, Mülkü var)
16	0.507996	(450.000-550.000 TL, Yetişkin)

Tablo 11'e göre, tek başına görülme olasılıkları en yüksek destek değerine sahip müşteri nitelikleri şu şekildedir:

0,831967 destek değeri yetişkin grubunda olması, 0,784398 destek değeri ile medeni durumu evli olması ve 0,679195 destek değeri ile eğitim durumu ilköğretim/ortaöğretim olan kişilerden oluşmaktadır.

Aşağıdaki Tablo 12'de çıkartılan birliktelik kuralları gösterilmiştir.

Tablo 12. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar

Kural Sayısı	Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence	Lift	Leverage	Conviction
0	(0 çocuk)	(Yetişkin)	0.639581	0.831967	0.517331	0.808859	0.972225	-0.014779	0.879106
1	(450.000-550.000 TL)	(Yetişkin)	0.601916	0.831967	0.507996	0.843965	1.014421	0.007222	1.076893
2	(Mülkü var)	(Evli)	0.647049	0.784398	0.508077	0.785221	1.001050	0.000533	1.003833
3	(Evli)	(Yetişkin)	0.784398	0.831967	0.665151	0.847977	1.019243	0.012558	1.105312
4	(Yetişkin)	(Evli)	0.831967	0.784398	0.665151	0.799493	1.019243	0.012558	1.075282
5	(İlköğretim/Ortaöğretim)	(Evli)	0.679195	0.784398	0.535352	0.788216	1.004867	0.002593	1.018026
6	(Kadın)	(Yetişkin)	0.622534	0.831967	0.537219	0.862955	1.037247	0.019291	1.226114
7	(Mülkü var)	(Yetişkin)	0.647049	0.831967	0.544768	0.841927	1.011972	0.006445	1.063010
8	(İlköğretim/Ortaöğretim)	(Yetişkin)	0.679195	0.831967	0.580404	0.854548	1.027141	0.015337	1.155245

Tablo 12'de kullanılan tanımlar şu şekilde açıklanmaktadır:

Antecedents= X nesnesi

Consequents=Y nesnesi

Antecedent support= X'in tek başına görülme olasılığı

Consequent support= Y'nin tek başına görülme olasılığı

Support= X ve Y'nin birlikte görülme olasılığı

Confidence= X nesnesini içeren işlemlerin B nesnesini de içermesi güven eşik değerini göstermektedir.

Lift= Kuralın farklılığı

Leverage= X ve Y'nin birlikte görülmesinin X ve Y'nin ayrı ayrı görülme olasılıklarından ne kadar fazla olduğunu gösterir.

Conviction= X'in Y olmadan görülme olasılığıdır.

Tablo 12'ye göre:

Kural 0: Hiç çocuğu olmayıp ⟹ yaş durumu= Yetişkin olan < sup:(0.517331) > conf:(0.808859) >lift:(0.972225) lev:(-0.014779) conv:(0.879106)

Kural 1: Gelir durumu= 450.000-550.000 TL olup ⟹ yaş durumu= Yetişkin olan < sup:(0.507996)>conf:(0.843965)>lift:(1.014421)lev:(0.007222) conv:(1.076893)

Kural 2: Mülk sahibi olan ⟹ medeni durumu= Evli < sup:(0.508077) > conf:(0.785221) > lift:(1.001050) lev:(0.000533) conv:(1.003833)

Kural 3: Medeni durumu= Evli ⟹ Yaş grubu= Yetişkin olan < sup:(0.665151) >conf:(0.847977) > lift:(1.019243) lev:(0.012558) conv:(1.105312)

Kural 4: Yaş grubu = Yetişkin olup ⟹ Medeni durumu= Evli olan < sup:(0.665151)> conf:(0.799493) > lift:(1.019243) lev:(0.012558) conv:(1.075282)

Kural 5: Eğitim durumu= İlköğretim/Ortaöğretim olup ⟹ Medeni durumu =Evli olması < sup:(0.535352) > conf:(0.788216) >lift:(1.004867) lev:(0.002593) conv:(1.018026)

Kural 6: Cinsiyet= Kadın olup ⟹ Yaş grubu= Yetişkin olması >sup:(0.537219) > conf:(0.862955) > lift:(1.037247) lev:(0.019291) conv:(1.226114)

Kural 7: Mülkü olup ⟹ Yaş Grubu= Yetişkin olması < sup:(0.544768) > conf:(0.841927) > lift:(1.011972) lev:(0.006445) conv:(1.063010)

Kural 8: Eğitim durumu= İlköğretim/Ortaöğretim olup ⟹ Yaş Grubu= Yetişkin olması < sup:(0.580404) > conf:(0.854548) > lift:(1.027141) lev:(0.015337) conv:(1.155245)

Kural 0'a göre; hiç çocuk sahibi olmayan müşterilerin tek başına görülme olasılığı %63,9 olurken, yetişkin grubunda olan müşterilerin tek başına görülme olasılığı %83,2'dir. Hiç çocuğu olman müşterilerin yetişkin olmasının birlikte görülme olasılığı ise %51,7'dir.Yüzde 81 doğruluk ile de yetişkin grubunda olan müşterilerin hiç çocuğu yoktur. Hiç çocuğu olmayan müşteriler ile yetişkin grubunda olan müşterilerin görülme olasılığı 0,972 kat artmaktadır.

Kural 1'e göre; yüzde 84,4 doğruluk ile yetişkin kişilerin gelir durumu 450.000-550.000 TL arasındadır.

Kural 2'ye göre; yüzde 78,6 doğruluk ile evli olan kişilerin mülkleri de vardır.

Kural 3'e göre; yüzde 84,8 doğruluk ile yetişkin kişiler evlidir.

Kural 5'e göre; yüzde 78,9 doğruluk ile evli olan kişilerin eğitim durumu ilköğretim/ortaöğretimdir.

Kural 6'ya göre yüzde 86,3 doğruluk ile yetişkin kişilerin cinsiyeti kadındır.

Kural 7'ye göre; yüzde 84,2 doğruluk ile yetişkin kişiler mülk sahibidir.

Kural 8'e göre; yüzde 85,5 doğruluk ile yetişkin kişilerin eğitim durumu ilköğretim/ortaöğretimdir.

Şekil 8'e göre kredi kartını zamanında ödeyen müşteriler için Power BI'da veri kümesinde en yüksek tekrar eden özellikler birlikte gösterilmiştir.



Şekil 8. Kredi kartını zamanında ödeyen müşterilerin Power BI'da görselleştirilmesi

2) Kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip kişiler için:

Kredi kartı ödeme vadesi geçmiş ve kötü borca sahip 59 adet müşteri için; %60 destek değeri, %70 güven değeri belirlenmiştir.

Tablo 13. Tüm sık geçen müşteri bilgileri

Kural sayısı	**Destek**	**Nesneler**
1	0.881356	(Yetişkin)
2	0.745763	(Mülkü var)
3	0.711864	(Evli)
4	0.694915	(Yetişkin, Mülkü var)
5	0.677966	(0 çocuk)
6	0.661017	(Arabası yok)
7	0.644068	(İlköğretim/Ortaöğretim)
8	0.644068	(0 çocuk, Yetişkin)
9	0.610169	(Arabası yok, Yetişkin)
10	0.610169	(İlköğretim/Ortaöğretim, Yetişkin)

Tablo 13'e göre, tek başına görülme olasılıkları en yüksek destek değerine sahip müşteri nitelikleri şu şekildedir:

0.881356 destek değeri ile yetişkin olması, 0.745763 destek değeri ile mülke sahip olması ve 0.711864 destek değeri ile evli olan kişilerden oluşmaktadır.

Aşağıdaki Tablo 14'de çıkartılan birliktelik kuralları gösterilmiştir.

Tablo 14. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar

Kural Sayısı	Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence	Lift	Leverage	Conviction
0	(0 çocuk)	(Yetişkin)	0.677966	0.881356	0.644068	0.950000	1.077885	0.046538	2.372881
1	(Yetişkin)	(0 çocuk)	0.881356	0.677966	0.644068	0.730769	1.077885	0.046538	1.196126
2	(Arabası yok)	(Yetişkin)	0.661017	0.881356	0.610169	0.923077	1.047337	0.027578	1.542373
3	(Yetişkin)	(Mülkü var)	0.881356	0.745763	0.694915	0.788462	1.057255	0.037633	1.201849
4	(Mülkü var)	(Yetişkin)	0.745763	0.881356	0.694915	0.931818	1.057255	0.037633	1.740113
5	(İlköğretim/Ortaöğretim)	(Yetişkin)	0.644068	0.881356	0.610169	0.947368	1.074899	0.042517	2.254237

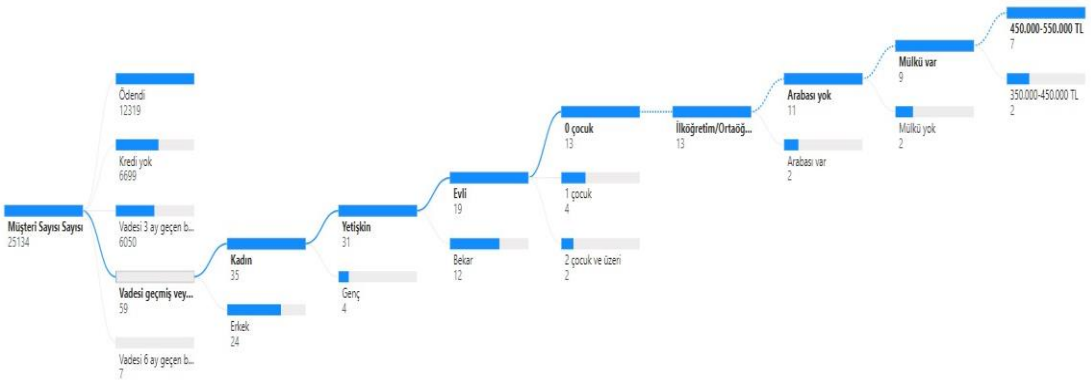
Kural 0'göre; yüzde 95 doğruluk ile yetişkin kişilerin hiç çocuğu yoktur.

Kural 2'ye göre; yüzde 92,3 doğruluk ile yetişkin kişilerin arabası yoktur.

Kural 4'e göre yüzde 93,2 doğruluk ile yetişkin kişilerin mülkü vardır.

Kural 5'e göre yüzde 94,7 doğruluk ile yetişkin kişilerin eğitim durumu ilköğretim/ortaöğretimdir.

Şekil 9'a göre kredi kartı borcunun vadesini geçirmiş veya kötü borca sahip müşteriler için Power BI'da veri kümesinde en yüksek tekrar eden özellikler birlikte gösterilmiştir.



Şekil 9. Kredi kartı ödeme vadesi geçmiş veya kötü borca sahip müşterilerin Power BI'da görselleştirilmesi

3) Kredi kartı borcunun vadesi 6 ay geçen kişiler için:

Kredi kartı ödeme vadesi 6 ay geçen borca sahip 7 adet müşteri için; %70 destek değeri, %80 güven değeri belirlenmiştir.

Tablo 15. Tüm sık geçen müşteri bilgileri

Kural Sayısı	**Destek**	**Nesneler**
1	0.857143	(Erkek)
2	0.857143	(İlköğretim/Ortaöğretim)
3	0.714286	(Mülkü yok)
4	0.714286	(Erkek, İlköğretim/Ortaöğretim)
5	0.714286	(İlköğretim/Ortaöğretim, Mülkü yok)

Tablo 15'e göre, tek başına görülme olasılıkları en yüksek destek değerine sahip müşteri nitelikleri şu şekildedir:

0.857143 destek değeri ile cinsiyeti erkek olması, 0.857143 destek değeri ile eğitim durumu ilköğretim/ortaöğretim olması ve 0.714286 destek değeri ile mülke sahip olmayan kişilerden oluşmaktadır.

Aşağıdaki Tablo 16'da çıkartılan birliktelik kuralları gösterilmiştir.

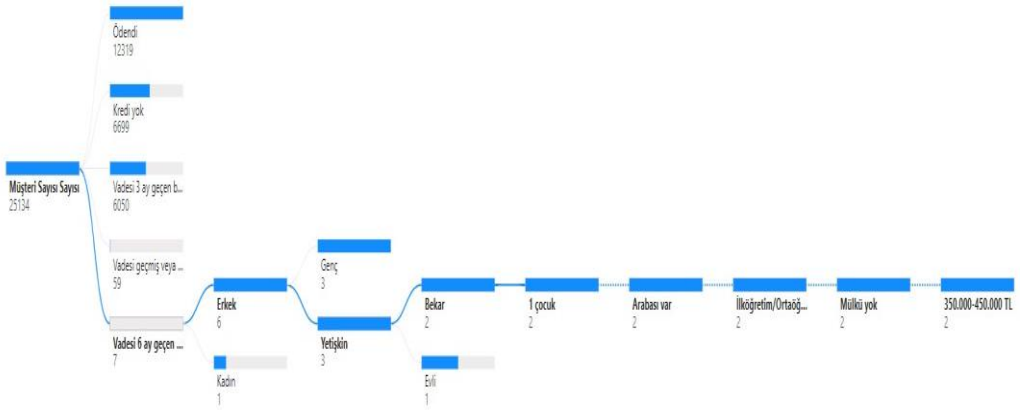
Tablo 16. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar

Kural Sayısı	Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence	Lift	Leverage	Conviction
0	(Erkek)	(İlköğretim/ Ortaöğretim)	0.857143	0.857143	0.714286	0.833333	0.972222	-0.020408	0.857143
1	(İlköğretim/ Ortaöğretim)	(Erkek)	0.857143	0.857143	0.714286	0.833333	0.972222	-0.020408	0.857143
2	(İlköğretim/ Ortaöğretim)	(Mülkü yok)	0.857143	0.714286	0.714286	0.833333	1.166667	0.102041	1.714286
3	(Mülkü yok)	(İlköğretim/ Ortaöğretim)	0.714286	0.857143	0.714286	1.000000	1.166667	0.102041	inf

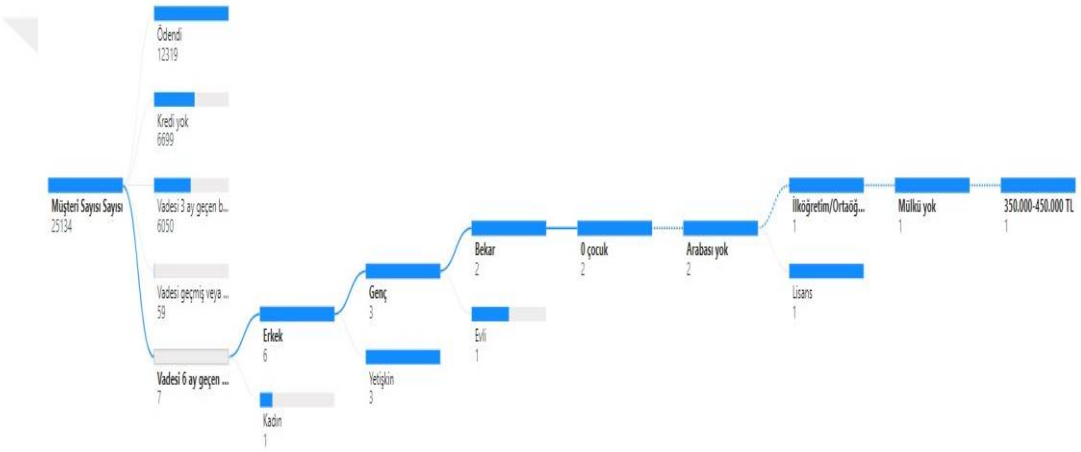
Kural 0'a göre; yüzde 83,3 doğruluk ile cinsiyeti erkek olan müşterilerin eğitim durumu ilköğretim/ortaöğretimden mezun olmuşlardır.

Kural 3'e göre yüzde 100 doğruluk ile mülkü olmayan müşterilerin eğitim durumu ilköğretim/ortaöğretimden mezun olmuştur.

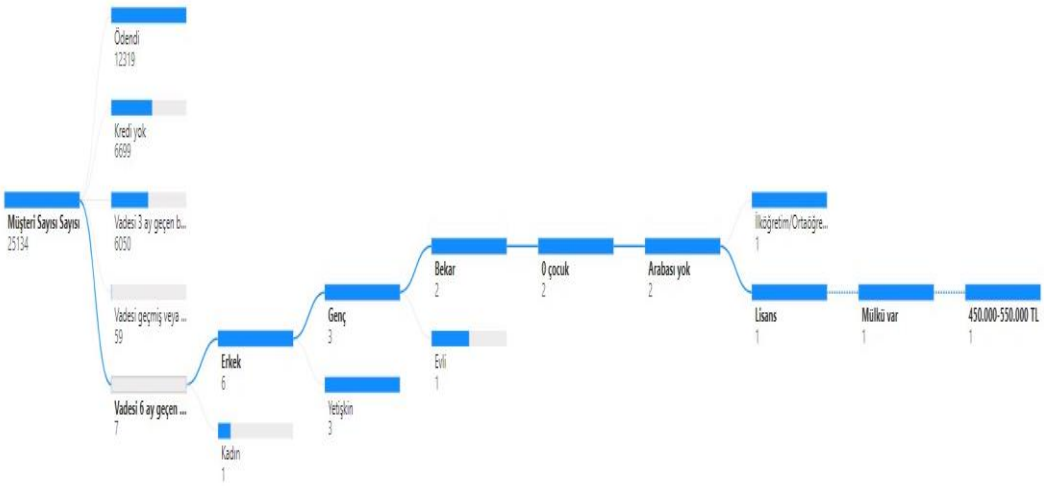
Şekil 10, Şekil 11 ve Şekil 12'ye göre kredi kartı borcunun vadesi 6 ay geçen müşterilerin için Power BI'da veri kümesinde en yüksek tekrar eden özelliklerin tüm kombinasyonları birlikte gösterilmiştir.



Şekil 10. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi



Şekil 11. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi



Şekil 12. Kredi kartı borcunun vadesi 6 ay geçen müşterilerin Power BI'da görselleştirilmesi

4) Kredi kartı borcunun vadesi 3 ay geçen kişiler için:

Kredi kartı ödeme vadesi 3 ay geçen borca sahip 6050 adet müşteri için; %50 destek değeri, %80 güven değeri belirlenmiştir.

Tablo 17. Tüm sık geçen müşteri bilgileri

Kural Sayısı	**Destek**	**Nesneler**
1	0.833554	(Yetişkin)
2	0.777025	(Evli)
3	0.687438	(İlköğretim/Ortaöğretim)
4	0.671074	(Mülkü var)
5	0.666281	(Evli, Yetişkin)
6	0.634215	(0 çocuk)
7	0.624132	(Kadın)
8	0.606942	(450.000-550.000 TL)
9	0.587934	(Yetişkin, İlköğretim/Ortaöğretim)
10	0.587603	(Arabası yok)
11	0.566446	(Mülkü var, Yetişkin)
12	0.537521	(Kadın, Yetişkin)
13	0.536033	(Evli, İlköğretim/Ortaöğretim)
14	0.524132	(Evli, Mülkü var)
15	0.520661	(0 çocuk, Yetişkin)
16	0.517851	(450.000-550.000 TL, Yetişkin)

Tablo 17'ye göre, tek başına görülme olasılıkları en yüksek destek değerine sahip müşteri nitelikleri şu şekildedir:

0.833554 destek değeri ile müşterilerin yetişkin grubunda olması, 0.777025 destek değeri ile medeni durumlarının evli olması ve 0.687438 destek değeri ile eğitim durumlarının ilköğretim/ortaöğretimden mezun olan kişilerden oluşmaktadır.

Aşağıdaki Tablo 18'de çıkartılan birliktelik kuralları gösterilmiştir.

Tablo 18. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar

Kural Sayısı	Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence	Lift	Leverage	Conviction
0	(0 çocuk)	(Yetişkin)	0.634215	0.833554	0.520661	0.820954	0.984884	-0.007991	0.929628
1	(450.000-550.000 TL)	(Yetişkin)	0.606942	0.833554	0.517851	0.853214	1.023586	0.011932	1.133935
2	(Evli)	(Yetişkin)	0.777025	0.833554	0.666281	0.857477	1.028701	0.018589	1.167857
3	(Kadın)	(Yetişkin)	0.624132	0.833554	0.537521	0.861229	1.033201	0.017273	1.199430
4	(Mülkü var)	(Yetişkin)	0.671074	0.833554	0.566446	0.844089	1.012639	0.007070	1.067570
5	(İlköğretim/Ortaöğretim)	(Yetişkin)	0.687438	0.833554	0.587934	0.855254	1.026033	0.014917	1.149917

Tablo 18'e göre birlikte görülme olasılığı en yüksek olan %66,6 ile evli ve yetişkin grubunda olan müşterilerdir.

Kural 0'a göre; yüzde 82,1 doğruluk ile hiç çocuk sahibi olmayan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 1'e göre; yüzde 85,3 doğruluk ile gelir durumu 450.000-550.000 TL olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

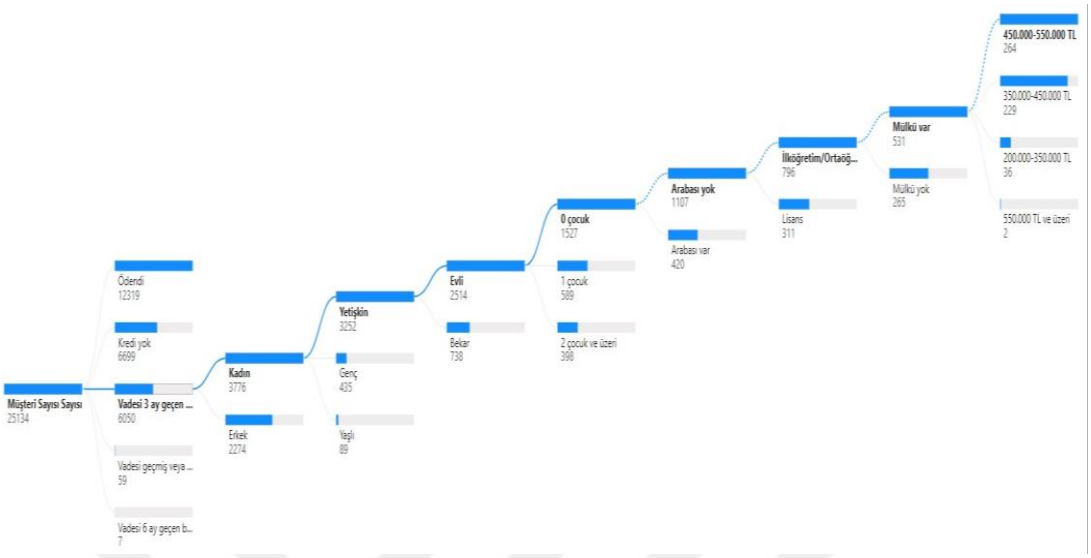
Kural 2'ye göre; yüzde 85,7 doğruluk ile evli olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 3'e göre; yüzde 86,1 doğruluk ile cinsiyeti kadın olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 4'e göre; yüzde 84,4 doğruluk ile mülk sahibi olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 5'e göre; yüzde 85,5 doğruluk ile eğitim durumu ilköğretim/ortaöğretimden mezun olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Şekil 13'e göre kredi kartı borcunun vadesi 3 ay geçen müşteriler için Power BI'da veri kümesinde en yüksek tekrar eden özellikler birlikte gösterilmiştir.



Şekil 13. Kredi kartı borcunun vadesi 3 ay geçen müşterilerin Power BI'da görselleştirilmesi

5) Kredi kartına sahip olmayan kişiler için:

Kredi kartına sahip olmayan 6699 adet müşteri için; %50 destek değeri, %70 güven değeri belirlenmiştir.

Tablo 19. Tüm sık geçen müşteri bilgileri

Kural Sayısı	Destek	Nesneler
1	0.836692	(Yetişkin)
2	0.781161	(Evli)
3	0.672041	(Yetişkin, Evli)
4	0.660546	(İlköğretim/Ortaöğretim)
5	0.654426	(Mülkü var)
6	0.619346	(0 çocuk)
7	0.619346	(Kadın)
8	0.618898	(450.000-550.000 TL)
9	0.577997	(Arabası yok)
10	0.568891	(İlköğretim/Ortaöğretim, Yetişkin)
11	0.553665	(Mülkü var, Yetişkin)
12	0.543365	(Kadın, Yetişkin)
13	0.525153	(Yetişkin, 450.000-550.000 TL)
14	0.519630	(İlköğretim/Ortaöğretim, Evli)
15	0.509778	(Mülkü var, Evli)
16	0.509330	(0 çocuk, Yetişkin)

Tablo 19'a göre, tek başına görülme olasılıkları en yüksek destek değerine sahip müşteri nitelikleri şu şekildedir:

0.836692 destek değeri ile müşterilerin yetişkin grubunda olması, 0.781161 destek değeri ile medeni durumlarının evli olması ve 0.672041 destek değeri ile yetiştin grubunda olup evli olan müşterilerden oluşmaktadır.

Aşağıdaki Tablo 20'de çıkartılan birliktelik kuralları gösterilmiştir.

Tablo 20. Apriori algoritmasının veri setine uygulanması ile ortaya çıkan kurallar

Kural Sayısı	Antecedents	Consequents	Antecedent support	Consequent support	Support	Confidence	Lift	Leverage	Conviction
0	(0 çocuk)	(Yetişkin)	0.619346	0.836692	0.509330	0.822367	0.982879	-0.008872	0.919355
1	(450.000-550.000 TL)	(Yetişkin)	0.618898	0.836692	0.525153	0.848529	1.014147	0.007326	1.078145
2	(Mülkü var)	(Evli)	0.654426	0.781161	0.509778	0.778969	0.997193	-0.001435	0.990081
3	(Yetişkin)	(Evli)	0.836692	0.781161	0.672041	0.803211	1.028227	0.018449	1.112049
4	(Evli)	(Yetişkin)	0.781161	0.836692	0.672041	0.860310	1.028227	0.018449	1.169071
5	(İlköğretim/Ortaöğretim)	(Evli)	0.660546	0.781161	0.519630	0.786667	1.007048	0.003637	1.025806
6	(Kadın)	(Yetişkin)	0.619346	0.836692	0.543365	0.877320	1.048558	0.025163	1.331168
7	(Mülkü var)	(Yetişkin)	0.654426	0.836692	0.553665	0.846031	1.011162	0.006112	1.060655
8	(İlköğretim/Ortaöğretim)	(Yetişkin)	0.660546	0.836692	0.568891	0.861243	1.029343	0.016217	1.176934

Tablo 20'ye göre birlikte görülme olasılığı en yüksek olan %67,2 ile evli ve yetişkin grubunda olan müşterilerdir.

Kural 0'a göre; yüzde 82,2 doğruluk ile hiç çocuğu olmayan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 1'e göre; yüzde 84,5 doğruluk ile gelir durumu 450.000-550.000 TL arasında olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 2'ye göre; yüzde 77,9 doğruluk ile mülkü olan müşterilerin medeni durumu evlidir.

Kural 4'e göre; yüzde 86 doğruluk ile medeni durumu evli olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

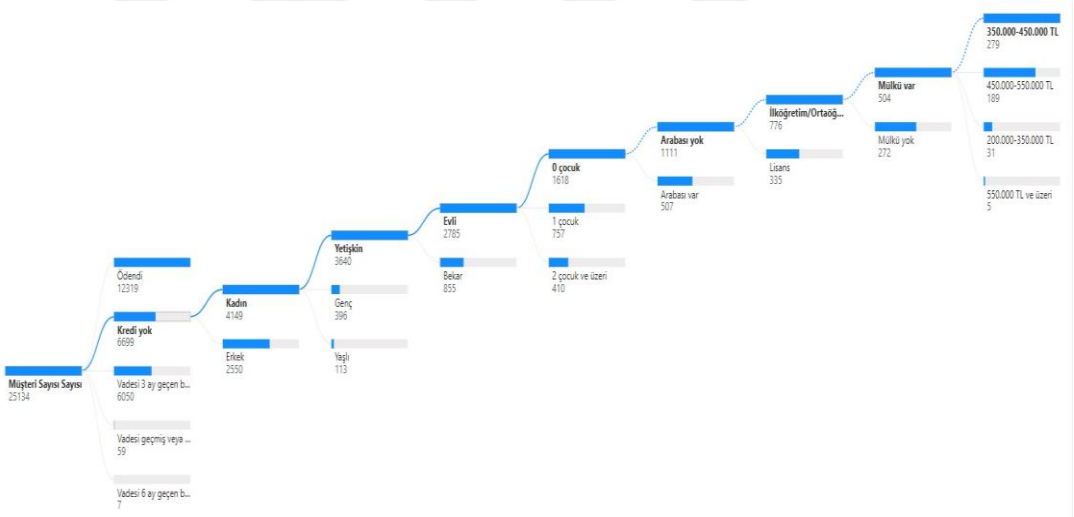
Kural 5'e göre; yüzde 78,7 doğruluk ile eğitim durumu ilköğretim/ortaöğretimden mezun olan kişilerin medeni durumu evlidir.

Kural 6'ya göre; yüzde 87,7 doğruluk ile cinsiyeti kadın olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 7'ye göre; yüzde 84,6 doğruluk ile mülkü olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

Kural 8'e göre; yüzde 86,1 doğruluk ile eğitim durumu ilköğretim/ortaöğretimden mezun olan müşterilerin yaş grubu yetişkin kişilerden oluşmaktadır.

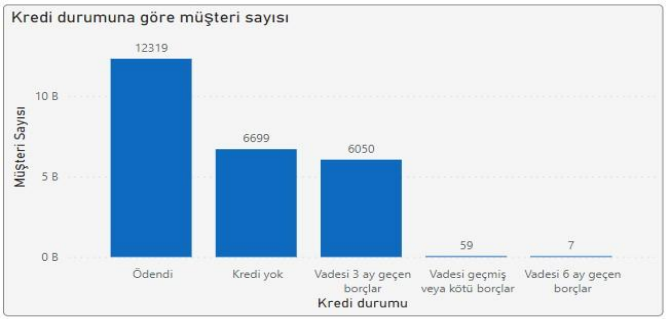
Şekil 14'e göre kredi kartına sahip olmayan müşterilerin için Power BI'da veri kümesinde en yüksek tekrar eden özellikler birlikte gösterilmiştir.



Şekil 14. Kredi kartına sahip olmayan müşterilerin Power BI'da görselleştirilmesi

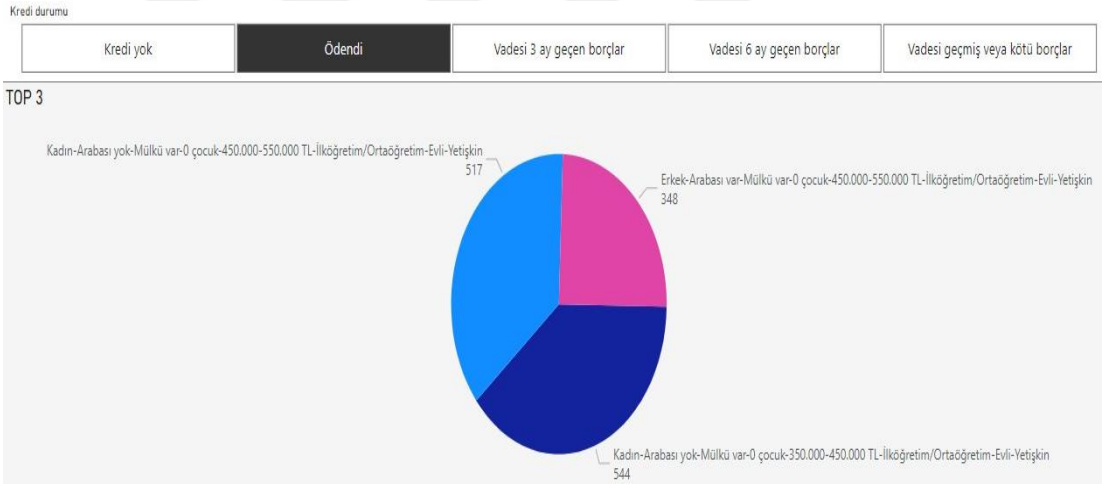
Genel olarak Power BI'da bazı örnekler için raporlamalar şu şekilde açıklanmaktadır:

Şekil 15'e göre kredi kartı ödeme durumlarına göre müşteri sayılarında kredi kartı olmayan müşteri sayısı 6699 kişi iken vadesi 6 ay geçen müşteri sayısı ise 7 kişiden oluşmaktadır.



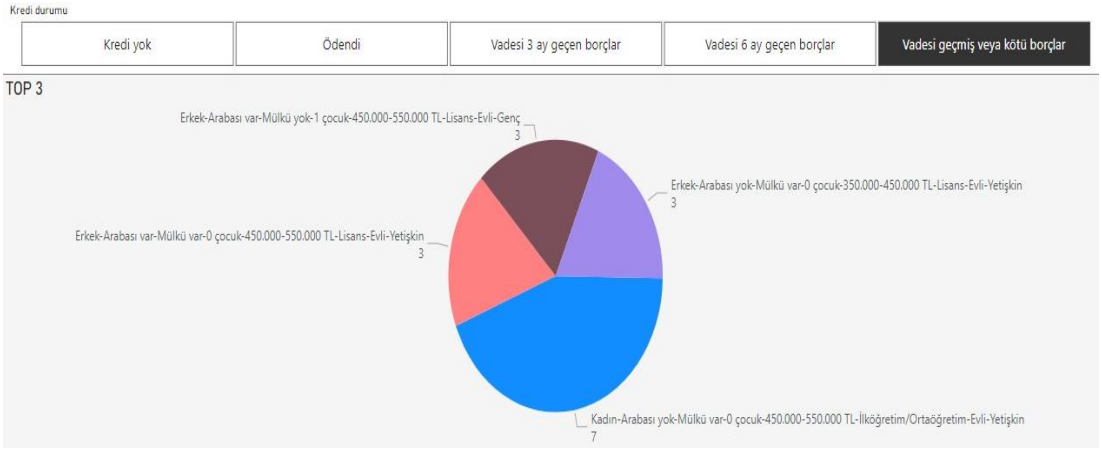
Şekil 15. Kredi durumlarına göre müşteri sayısı

Şekil 16'ya göre kredi kartını zamanında ödeyen müşteriler için en yüksek üç tekrar eden tüm özellikleri kapsayan müşteri sayıları belirtilmiştir. Buna göre en yüksek özelliğe sahip olan 544 adet müşteri için cinsiyeti kadın olup, arabası olmayan fakat mülkü olan, hiç çocuğu olmayan, yıllık gelir tutarı 350.000-450.000 TL arasında olan, eğitim durumu ilköğretim/ortaöğretim mezunu olan evli ve yetişkin kişilerden oluşmaktadır.



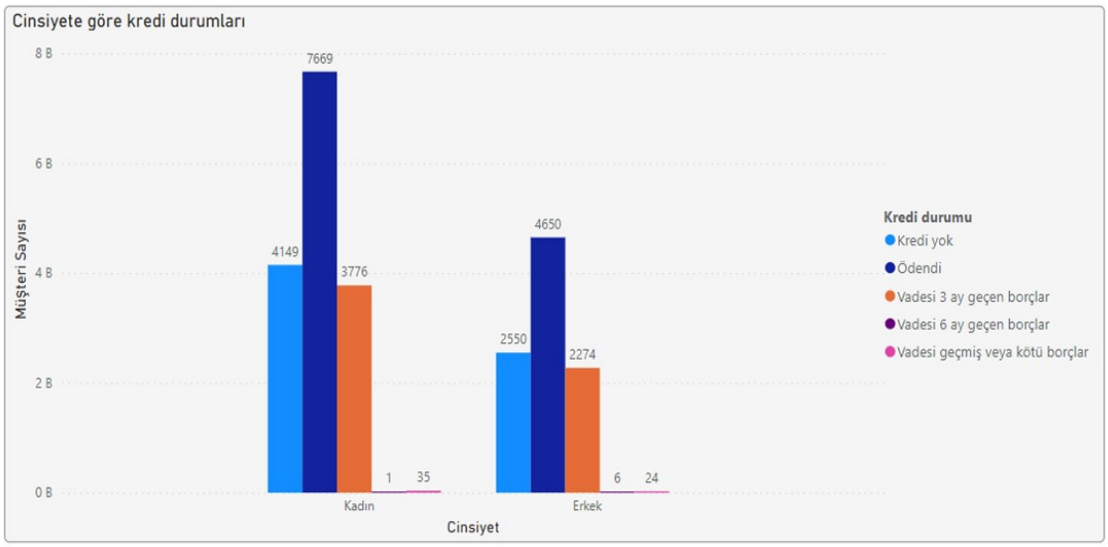
Şekil 16. Kredi kartını zamanında ödeyen müşterilerin tüm özelliklerini birlikte kapsayan durumların gösterimi

Şekil 17'ye göre ise kredi kartı ödeme vadesi geçmiş kötü borca sahip müşteriler için en yüksek üç tekrar eden tüm özellikleri kapsayan müşteri sayılarına göre en yüksek özelliğe sahip olan 7 adet müşteri için cinsiyeti kadın olup, arabası olmayan fakat mülkü olan, hiç çocuğu olmayan, yıllık gelir tutarı 450.000-550.000 TL arasında olan, eğitim durumu ilköğretim/ortaöğretim mezunu olan evli ve yetişkin kişilerden oluşmaktadır.



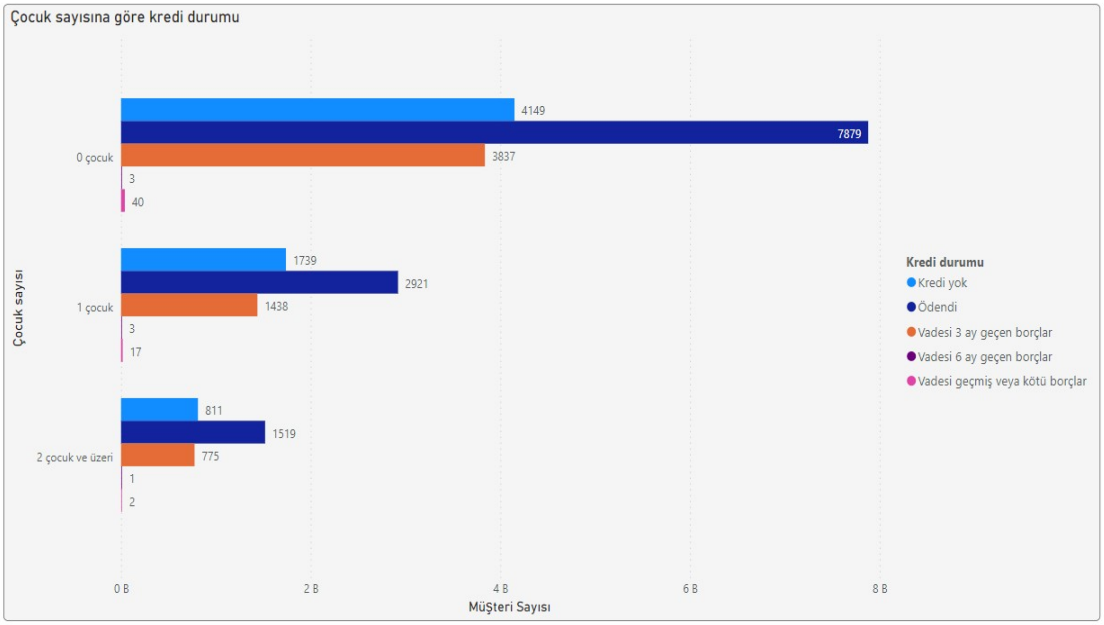
Şekil 17. Vadesi geçmiş veya kötü borca sahip müşterilerin tüm özelliklerini birlikte kapsayan durumların gösterimi

Şekil 18'e göre cinsiyete göre kredi kartı durumlarına baktığımızda kadın müşteriler en çok kredi kartını zamanında öderken, vadesini 6 ay geçiren müşterilerin çoğunluğu ise erkek müşterilerden oluşmaktadır.



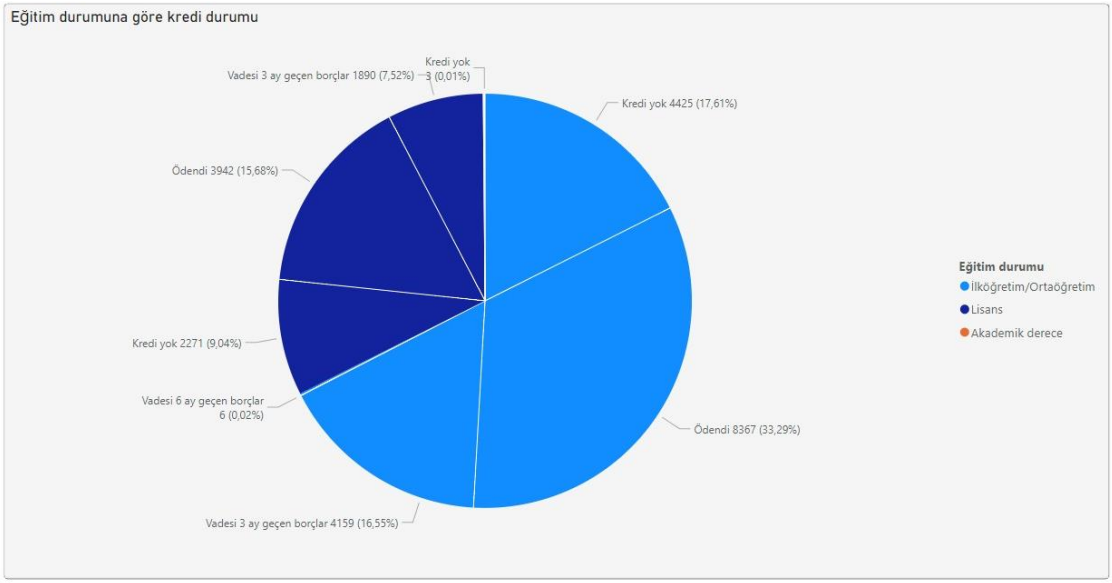
Şekil 18. Cinsiyete göre kredi durumları

Şekil 19'a göre çocuk sayısına göre kredi kartı ödeme durumunda vadesi geçmiş veya kötü borca sahip kişilerin hiç çocuğu yoktur.



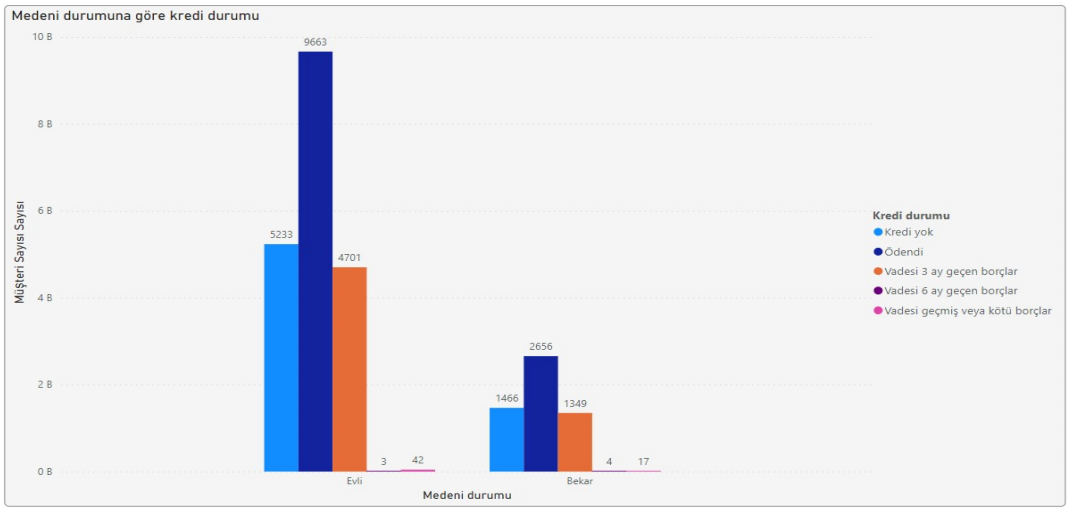
Şekil 19. Çocuk sayısına göre kredi durumları

Şekil 20'ye göre eğitim durumları ile kredi kartı ödeme durumları arasında yüzdelik ilişkide %33,29 ile kredisini zamanında ödeyen müşteriler ilköğretim/ortaöğretim mezunudur. Lisans sahibi kişiler arasında ise %15,68 ile kredi kartını zamanında ödemiştir.



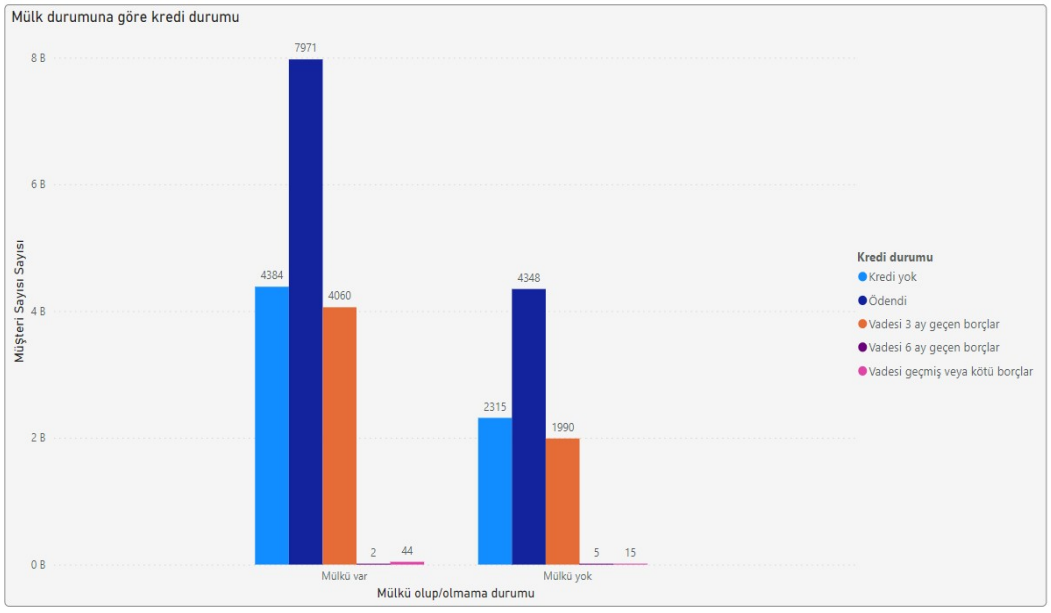
Şekil 20. Eğitim durumuna göre kredi kartı ödeme durumları

Şekil 21'e göre vadesi 6 ay geçen borca sahip müşterilerin çoğunluğunun medeni durumu bekardır. Kredi kartını zamanında ödeyen müşterilerin çoğunluğu ise evli kişilerden oluşmaktadır.



Şekil 21. Medeni durumuna göre kredi durumları

Şekil 22'ye göre vadesi geçmiş veya kötü borca sahip müşterilerin çoğunluğunun mülkü varken, vadesini 6 ay geçen borca sahip müşterin ise mülkü yoktur.



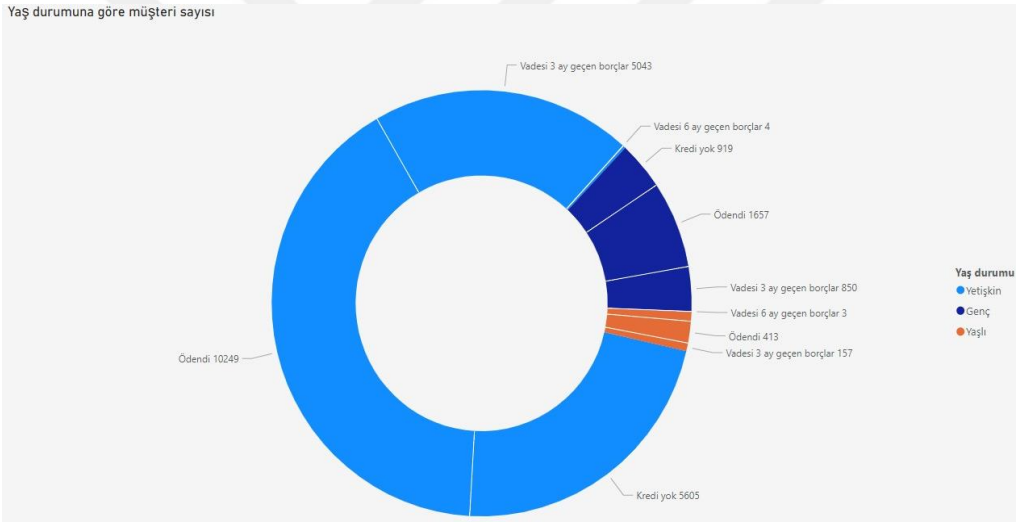
Şekil 22. Mülk durumuna göre kredi durumları

Şekil 23'e göre kredi kartı olmayan müşterilerin çoğunluğunun yıllık gelir durumu 450.000-550.000 TL arasındadır.

Yıllık gelir tutarı	Kredi yok	Ödendi	Vadesi 3 ay geçen borçlar	Vadesi 6 ay geçen borçlar	Vadesi geçmiş veya kötü borçlar	Toplam
200.000-350.000 TL	239	531	238	1	3	**1012**
350.000-450.000 TL	2203	4237	2058	4	22	**8524**
450.000-550.000 TL	4146	7415	3672	2	32	**15267**
550.000 TL ve üzeri	111	136	82		2	**331**
Toplam	**6699**	**12319**	**6050**	**7**	**59**	**25134**

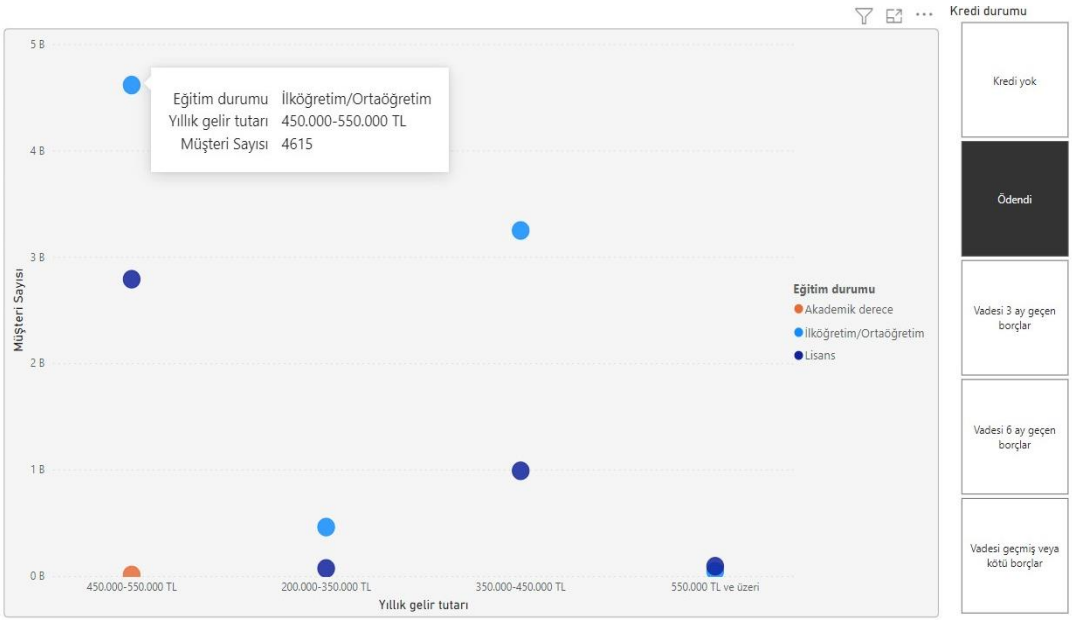
Şekil 23. Yıllık gelir durumuna göre kredi durumları

Şekil 24'e göre yaşlı grubundaki 413 müşteri kredi kartı borcunu zamanında öderken aynı zamanda genç grubundaki 1657 müşteri de kredi kartı borcunu zamanında ödemiştir.



Şekil 24. Yaş durumuna göre kredi kartı ödeme durumları

Şekil 25'e göre yıllık gelir tutarı ile eğitim durumları arasındaki ilişkide kredi kartını zamanıda ödeyen müşteriler için yıllık gelir tutarı 450.00-550.000 TL olan müşterilerin eğitim durumu ilköğretim/ortaöğretimdir ve en yüksek tekrar eden değerdir.



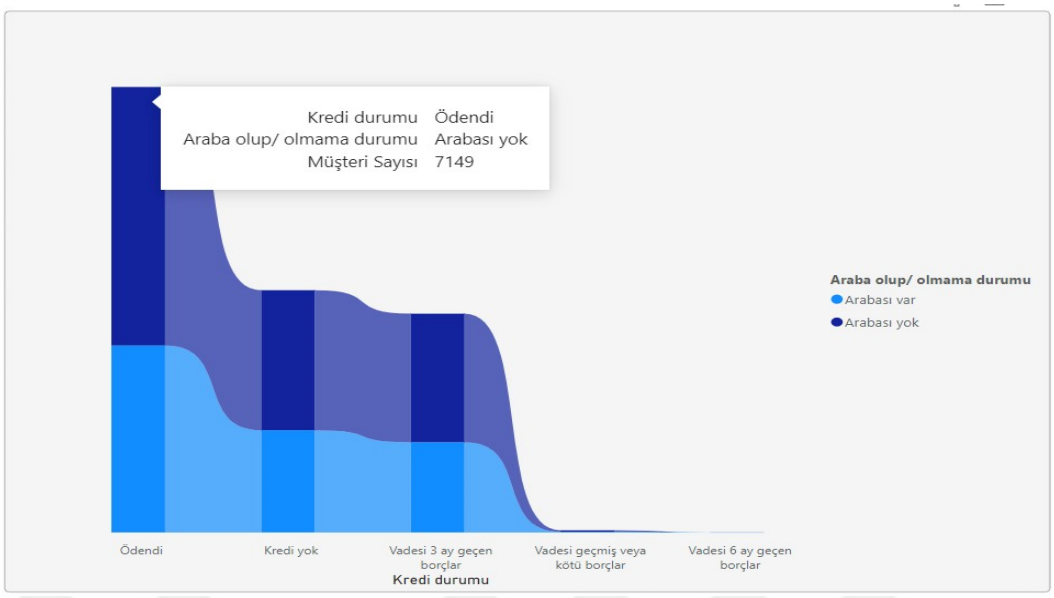
Şekil 25. Yıllık gelir durumu ve eğitim durumu arasındaki ilişkiyi kredi kartı ödeme durumlarına göre gösterilmesi

Şekil 26'göre mülk durumu ile çocuk sayısı arasındaki ilişkide kredi kartını zamanında ödeyen müşteriler için hiç çocuğa sahip olmayan müşteriler mülk sahibidir.



Şekil 26. Çocuk sayısı ile mülk durumu arasındaki ilişkiyi kredi kartı ödeme durumlarına göre gösterilmesi

Şekil 27'ye göre arabası olup/olmama durumuna göre kredi kartını zamanında ödeyen müşterilerin çoğunluğunun arabası yoktur.



Şekil 27. Arabası olup/olmama durumuna göre kredi kartı ödeme durumları

4.2 Tartışma

Elde edilen her bir bulgu deneme yanılma yoluyla belirlenen destek ve güven eşik değerlere göre kurallar oluşturulmuştur. Buna bağlı olarak kural sayıları farklılık göstermektedir. Güven ve destek değerlerindeki artış nesnelerin birlikte olma olasılığını da arttırmaktadır. Yani 0 ve 1 arasında belirlenen destek ve güven eşik değerleri 1'e yaklaştıkça aralarındaki güçlü birliktelikte o kadar artmaktadır. Düşük değerler ise kural sayısının artmasına neden olur. Buna göre algoritmada belirlenen destek ve güven değerleri en optimal sonuç verecek şekilde her durumda farklı değerler oluşturulmuştur.

Bu çalışmaya göre, çıkan sonuçlardaki kurallar ile bir müşterinin bilgilerine bakılarak kredi kartının borcunu ödeyebilme durumlarından söz edilebilir. Ayrıca hangi müşteri özelliğine sahip olan kişiler kredi kartını zamanında veya geç ödedikleri konusunda da fikir vermektedir. Bu sayede banka müşteri bazlı veya firma bazlı farklı stratejiler belirleyebilir.

Sonuçta çıkan kurallar farklı kredi kartı borcu ödeme vadeleriyle aynı sonucu verebilir. Örneğin, çıkan en yüksek doğruluk değerlerine göre değerlendirildiğinde kredi kartını zamanında ödeyen müşteriler yetişkin grubunda olup cinsiyeti kadın müşteriden oluşmaktadır. Aynı zamanda kredi kartı borcunun vadesini 3 ay geçiren müşteriler ile herhangi bir kredi kartına sahip olmaya müşterinde yaş grubu yetişkin olup cinsiyeti kadın

müşterilerden oluşmaktadır. Oluşan her üç durum için birbirleriyle kıyaslandığında bu sonucun yanında diğer en yüksek doğruluğa sahip kurallarında dikkate alınması gerektiğinde daha iyi sonuçlar verecektir. Ayrıca algoritma, müşteri bilgilerindeki çok az rastlanan durumları yok saydığından çıkan kurallarda her müşteri bilgisi gözlenememiştir. Örneğin, tüm veri setinde yetişkinlerin sayısı, genç ve yaşlı grubundaki müşteri sayısından çok fazla olduğundan çıkan kurallarda genç ve yaşlı müşteri grubuna rastlanılmamıştır.

5. SONUÇ ve ÖNERİLER

Çalışmanın amacı apriori algoritması kullanarak banka sektöründe kredi kartı ödeme vadeleri ile müşterilerin genel özellikleri arasındaki ilişkiyi belirlemektir. Böyle bir çalışma yapılmasının çeşitli avantajları vardır. İlk olarak banka sektörü veri madenciliği yöntemi ile incelenmektedir. Müşterilerin kredi kartı ödeme vadelerindeki durumlarından vadesini geçirmiş, kötü borca sahip müşteri bilgilerini saptayarak, bu müşteriler için yeniden kredi kartı yapılandırma politikaları oluşturularak kredi kartı borcunu ödeyebilmelerini sağlanabilir. Bunun yanında kötü borca sahip müşteri bilgilerine benzeyen müşteriler için ise ilk aşamada limit azaltılarak kredi kartı verilebilir. Ayrıca kredi kartını zamanında ödeyebilen müşteriler için de borçları bittikten sonra sürekliliği sağlamak adına farklı stratejiler geliştirilerek bankaya yeniden kazandırılabilir. Yani mevcut/potansiyel müşterilerin doğru şekilde kredi kartı satış sürecinde firmalara önemli katkılar sağlayabilir.

Çalışmanın amacını gerçekleştirebilmek için ilk olarak banka sektöründe kredi kartı ödeme vadeleri durumlarının müşteri bilgilerini kapsayan veri seti kullanılarak bir veritabanı oluşturulmuştur. Bu veritabanındaki her bir bilgi algoritma öncesinde veri ön işleme sürecinden geçerek algoritmaya hazır hale getirilmiştir. Veriler belli kategorilerde kategorize edilerek yeniden oluşturulmuştur ve veri kümesindeki bazı nitelikler çalışma için daha az önemli olduğu belirlenerek veri setinden çıkarılmıştır. Bundan sonra veritabanında yer alan verilere veri madenciliği yöntemlerinden olan apriori algoritması uygulanmıştır. Sonuçta banka sektöründe işlem yapacak kişi veya firmalar için müşteriyi daha iyi analiz etmeye olanak sağlayacak bilgiler ortaya çıkmıştır. Ayrıca veriler Power BI aracında görselleştirilerek sunulmuştur.

Birliktelik kuralı sonucuna göre, kredi kartına sahip olmayan 6699 adet müşteri için ve kredi kartını zamanında ödeyen 12319 adet müşteri için %50 destek, %70 güven eşik değerleri uygulandığında 9 adet kural oluşmuştur. Kredi kartı borcunun vadesi 3 ay geçen borca sahip 6050 adet müşteri için %50 destek ve %80 güven eşik değeri uygulandığında 6 adet kural oluşurken, kredi kartı borcunun vadesi 6 ay geçen borca sahip 7 adet müşteri

için %70 destek ve %80 güven eşik değeri sonucunda 4 adet kuralın çıktığı gözlemlenmiştir. Bunların yanında kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip 59 adet müşteri için ise %60 destek ve %70 güven değeri sonucunda 6 adet kural çıkmıştır.

Çalışmanın sonucunda her kategoride en yüksek doğruluğa sahip kural için sonuçlar değerlendirilmiştir. Bu sonuçlara göre; kredi kartı sahibi olmayan, kredi kartını zamanında ödeyen ve kredi kartı vadesinin borcunu 3 ay geçiren yetişkin müşterilerin cinsiyeti kadındır. Hiç çocuğu olmayan yetişkin müşteriler ise kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip kişilerdir. Ayrıca mülk sahibi olmayan ve eğitim durumu ilköğretim/ortaöğretimden mezun olan müşteriler ise kredi kartı borcunun 6 ay geçmesine rağmen ödemeyen müşterilerdir.

Bunların yanında diğer kurallarda ise kredi kartı vadesini 6 ay geçirmiş eğitim durumu ilköğretim/ortaöğretim mezunu olan müşteriler erkek müşterilerdir. Kredi kartı borcunun vadesini geçirmiş ve kötü borca sahip yetişkin kişilerin arabası olmayıp, mülkü olan ve eğitim durumu ilköğretim/ortaöğretimden mezun olan müşterilerden oluşmaktadır. Kredi kartını zamanında ödeyen ve kredi kartı olmayan yetişkin müşterilerin hiç çocuğu yoktur. Yıllık gelir durumları 450.000-550.000 TL arasındadır. Evli ve mülk sahibilerdir. Aynı zamanda ilköğretim/ortaöğretimden mezun olan müşterilerdir. Evli müşterilerin ise mülkü olup, ilköğretim/ortaöğretimden mezun olan müşterilerdir. Ek olarak kredi kartı vadesinin borcunu 3 ay geçiren yetişkin müşterilerin de hiç çocuğu yoktur. Gelir durumları 450.000-550.000 TL arasında olup evli ve mülk sahibilerdir. Aynı zamanda eğitim durumu ilköğretim/ortaöğretimden mezun olan müşterilerden oluşmaktadır.

Apriori algoritması ortaya çıktığından daha iyi geliştirilmiş algoritmalarda literatüre sunulmaktadır. İlerleyen çalışmalarda geliştirilmiş apriori algoritmaları kullanılabilir. Çalışmanın karşılaşılan problemlerden biri de sayısal değişkenleri kategorik değişkenlere çevrilmesi problemidir. Örneğin, müşterilerinin yaş bilgileri için yaşı 21-29 olanlar için "Genç", yaşı 30-59 arasında olanlar için "Yetişkin", yaşı 60 ve üzeri olanlar için ise "Yaşlı" olacak şekilde kategorize edilerek veri tabanında güncellenmiştir. İlerleyen çalışmalarda sayısal değişkenler kategorik değişkenlere çevrilerek başka yöntemlerde

kullanılabilir. Farklı güven ve destek eşik değerleri ile daha az veya daha fazla sayıda kurallar ile sektör daha detaylı bir şekilde incelenebilir. Çalışmaya ek olarak farklı müşteri özellikleri eklenip veya çıkartılarak yeniden çözümlenebilir. Örneğin, emekli olması, mesleği, kirada oturup oturmaması gibi özellikler eklenebilir. Bu çalışmada sadece 30 günlük veri alınmıştır. Daha detaylı analiz etmek için dönem bazında algoritma uygulanabilir. Görselleştirme araçları olarak Power BI yanında başka araçlarda kullanılabilir.

KAYNAKLAR

Alan, B. (2016). *Veri Madenciliği ve Market Veritabanında Birliktelik Kurallarının Belirlenmesi* (Yayın no. 436682) [Yüksek Lisans Tezi, Recep Tayyip Erdoğan Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Bayram, O. (2014). *Birliktelik Analizi ve Bir Uygulaması* (Yayın no. 374663) [Yüksek Lisans Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Bayram, S., ve Dündar, S. (2021). Türkiye'de Banka Şubesi Lokasyonunun Veri Madenciliği İle Analizi. *Uluslararası Bankacılık, Ekonomi Ve Yönetim Araştırmaları Dergisi*, 4(1), 34-52. 22 Temmuz 2023'de https://dergipark.org.tr/en/download/article-file/1765507 adresinden indirildi.

Bilgin, S. (2013). *Banka Müşterilerinin İnternet Bankacılığına İlişkin Yaklaşımlarının Veri Madenciliği Teknikleri İle İncelenmesi* (Yayın no. 351610) [Yüksek Lisans Tezi, İstanbul Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Çalış, A., ve Baynal, K. (2016). Kümeleme Analizi İle Bankacılık Sektöründe Satış Stratejilerinin Belirlenmesi, *Beykent Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, 1(9), 13 – 41. 16 Ağustos 2023'de https://dergipark.org.tr/tr/download/article-file/224545 adresinden indirildi.

Çelik, B. (2019). *Mekanik Arızaların Veri Madenciliği Apriori Algoritması İle Analiz Edilmesi* (Yayın no. 598685) [Yüksek Lisans Tezi, Kocaeli Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Demir, T. (2018). *Veri Madenciliği Yöntemleri İle Banka Çağrı Merkezi Müşterilerine Uygulanan Pazarlama Stratejilerinin Belirlenmesi* (Yayın no. 509804) [Yüksek Lisans Tezi, Bahçeşehir Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Gülce, A. C. (2010). *Veri Madenciliğinde Apriori Algoritması Ve Apriori Algoritmasının Farklı Veri Kümelerinde Uygulanması* (Yayın no. 269275) [Yüksek Lisans Tezi, Trakya Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Gündüz. (2015). *İş Zekası Uygulamaları Ve Pazar Sepeti Analizi* (Yayın no. 394507) [Yüksek Lisans Tezi, İstanbul Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Gündüz, S. (2017). *Veri Madenciliğinde Kullanılan Birliktelik Analizi Ve Market Sepet Analizi: Bir Uygulama* (Yayın no. 484324) [Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Gürsoy, T., ve Bilgin, S. (2016). Banka Müşterilerinin İnternet Bankacılığına İlişkin Yaklaşımlarının Veri Madenciliği Teknikleri İle İncelenmesi. *Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 7(14). 5 Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/545748 adresinden indirildi.

İldaş, T. (2021). Kredi Riski Ölçüm Modellerinin Değerlendirilmesi. *Finansal Araştırmalar Ve Çalışmalar Dergisi*, 13(25), 516-547. 22 Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/1901150 adresinden indirildi.

Karaatlı, M., Kocabıyık,T., Çal, D.Y., ve Çolak, M. (2021). Bıst-30 Endeksinde Yer Alan Payların Ortak Hareketlerinin Veri Madenciliği Kapsamında Birliktelik Kuralı İle İncelenmesi. *Finansal Araştırmalar ve Çalışmalar Dergisi*, 13(25), 548-571. 22 Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/1901149 adresinden indirildi.

Karaatlı, M., ve Yıldız, E. (2021). Mevduat Bankaların Finansal Yapılarının Kümeleme Analizi İle İncelenmesi. *Bmij Araştırma Makalesi*, 9(1), 1-17. 10 Temmuz 2023'de https://www.bmij.org/index.php/1/article/download/1594/1507/8491 adresinden alındı.

Karademir, M. (2021). *Veri Madenciliği Market Sepeti Analizi: Aprıorı ve Fp-Growth Algoritmaları Karşılaştırması* (Yayın no. 659482) [Yüksek Lisans Tezi, İstanbul

Medeniyet Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Kök, D., Aksu, G. (2013). Müşteri Kredi Değerliliğinin Belirlenmesinde Analitik Ağ Süreci Kullanımı: Bir Model Önerisi. *Muhasebe ve Finansman Dergisi*, (59), 167-186. 22 Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/427468 adresinden indirildi.

Özdamar, Y. (2020). *SQL Nedir? SQL Nerelerde Kullanılır?*. 22 Temmuz 2023'de https://teknoloji.org/sql-nedir-sql-nerelerde-kullanilir/ adresinden alındı.

Özdem, K., Akcayol, M. (2021). Müşteri Davranış Tahmini İçin Bir Model: Bankacılık Sektörü İçin Uygulama. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 10(1), 001- 008. 10 Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/1193415 adresinden alındı.

Öztemiz, F. (2017). *Aprıorı Algoritması İle Müşteri Bazlı Market Sepet Analizi ve Ürün Satış Tahmini* (Yayın no. 477653) [Yüksel Lisans Tezi, İnönü Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Rencüzoğulları, B. (2019). *Veri Madenciliğinde Birliktelik Kuralı ve Hepatit Hastalığı Üzerine Bir Uygulama* (Yayın no. 601842) [Yüksek Lisans Tezi, Eskişehir Teknik Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Tanyeli, E. (2021). *Çok Ölçütlü Karar Verme ve Birliktelik Analizi Yöntemleri İle Bir Tekstil İşletmesinin Aksesuar Deposu Yerleşimi Üzerine Bir Uygulama* (Yayın no. 695830) [Yüksek Lisans Tezi, Kırıkkale Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Tunç, Z. (2022). *Apriori Algoritmasıyla Market Sepet Analizi: Bir Perakende Firması Uygulaması* (Yayın no. 723398) [Yüksek Lisans Tezi, Bahçeşehir Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Yüksel, T., ve Zontul, M. (2019). Dağıtık Sistemlerde Birliktelik Kuralları ile Sepet Analizi. *Aurum Mühendislik Sistemleri Ve Mimarlık Dergisi*, 3(1), 65-77. 22

Temmuz 2023'de https://dergipark.org.tr/tr/download/article-file/773189 adresinden alındı.

ÖZGEÇMİŞ

Dilay TERLEMEZ

Eğitim

Derece	Yıl	Üniversite, Enstitü, Anabilim/Anasanat Dalı
Yüksek Lisans	2024	Maltepe Üniveristesi, Lisansüstü Eğitim Enstitüsü, Endüstri Mühendisliği Anabilimdalı
Lisans	2021	Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği

İş/İstihdam (Varsa)

Yıl	Görev
2022	Veri Analiz ve Raporlama Yetkilisi
2021	Satın Alma Uzmanı

Alınan Burs ve Ödüller

Yıl	Burs/Ödül
2020-2021	Endüstri Mühendisliği Bölümü İkincisi Ödülü
2019-2020	TEV-Eğitim Bursu ve Başarı Bursu

Yayınlar ve Diğer Bilimsel/Sanatsal Faaliyetler

Yıl	Ay	Yayın Yeri	Konusu
2022	Ekim	Uluslararası Kitap Bölümü(UAKB)	Mesleki Kas İskelet Sistemi Rahatsızlıkları