

**İSTANBUL TEKNİK ÜNİVERSİTESİ  
BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ**

**BANKACILIK SEKTÖRÜNDE MÜŞTERİ  
KAYBI ÖNGÖRÜSÜ**

**Bitirme Ödevi**

**Ecem SÖZERİ  
040090598**

**Bölüm: Bilgisayar Mühendisliği**

**Danışman: Prof. Dr. Eşref ADALI**

Ocak 2014

**İSTANBUL TEKNİK ÜNİVERSİTESİ  
BİLGİSAYAR VE BİLİŞİM FAKÜLTESİ**

**BANKACILIK SEKTÖRÜNDE MÜŞTERİ  
KAYBI ÖNGÖRÜSÜ**

**Bitirme Ödevi**

**Ecem SÖZERİ  
040090598**

**Bölüm: Bilgisayar Mühendisliği**

**Danışman: Prof. Dr. Eşref ADALI**

Ocak 2014

## **Özgünlük Bildirisi**

1. Bu çalışmada, başka kaynaklardan yapılan tüm alıntılar, ilgili kaynaklar referans gösterilerek açıkça belirtildiğini,
2. Alıntılar dışındaki bölümlerin, özellikle projenin ana konusunu oluşturan teorik çalışmaların ve yazılım/donanımın benim tarafımdan yapıldığını bildiririm.

İstanbul, Ocak 2014

Ecem SÖZERİ

## ÖNSÖZ

Bu çalışma süresince benden yardım ve desteğini hiçbir zaman esirgemeyen, tecrübeleri ile beni her zaman daha iyiye yönlendiren tez danışmanım Prof. Dr. Eşref ADALI'ya, çalışmanın başlangıç aşamasından bitiş aşamasına kadar bankacılık sektörüne ait bilgi birikimi ve tecrübelerini benimle paylaşan ve her zaman için destekleyici ve teşvik edici bir rol bürünen Sn. Uğur GÖKDERE'ye, çalışma sırasında karşılaştığım zorlukları çözüme ulaştırma sürecimde uzman bakış açısı ile yanımda bulunan Sn. Nihal Benlikaya ALPAN'a ve hayatım boyunca karşılıksız olarak beni seven ve her koşulda destekleyen aileme teşekkür ve şükranlarımı sunarım.

Ocak, 2014

Ecem SÖZERİ

# BANKACILIK SEKTÖRÜNDE MÜŞTERİ KAYBI ÖNGÖRÜSÜ

## ( ÖZET )

Günümüzde her sektör için müşteri kaybı istenmeyen bir durumdur. Müşteri kaybı en genel tanımıyla bir servis ya da hizmet sektöründeki kayıp müşterilerin ölçümü olarak açıklanabilir. Müşteri kaybı oranının en düşük seviyeye indirilmesi ile o sektöre ait organizasyon, pazardaki yerini iyileştirebilir ve kar payını arttırabilir. Bu nedenle müşteri kaybının öngörüsüne ve ölçülmesine yönelik olarak yapılan çalışmalar, Müşteri İlişkileri Yönetimi (MİY) sistemlerinin önemli çalışmalarından olmuştur.

Müşteri kaybının öngörüsünün yapılması ve ölçülmesi amacıyla oluşturulan çalışmalar bankacılık sektöründe de oldukça geniş yer tutmaktadır. Hizmet ve servislerini kullanan müşterilerini kaybetmek istemeyen bankalar bu amaçla yoğun analizler yapmakta ve yazılımlar kullanmaktadır.

Bu çalışmada bankacılık sektöründeki müşterilerin davranışları analiz edilerek, bu analizler sonucunda müşteri kaybının öngörülmesi amaçlanmıştır. Bu amaç doğrultusunda, ilk aşamada bankadan gerekli verilerin elde edilmesi ve bu verilerin yoğun analizler sonucunda çalışma için anlam ifade edecek şekle kavuşturulması sağlanmıştır. Ham veriden çalışma için daha anlamlı olacak veriye dönüşümün sağlanmasının ardından, elde edilen ve “eğitim verileri” ve “test verileri” olarak ayrılan bu verilerden eğitim verileri yardımıyla müşterilerin davranışlarını karakterize eden bir model oluşturulmuştur. Bu süreci izleyen aşamada ise müşterilere ait test verilerinin bu model ile test edilerek çalışmaya katılması sağlanmıştır. Böylelikle test verileri ile çalışmaya katılan müşterilerin o bankaya ait hizmet ve servisleri kullanmaya devam edeceklerine ya da etmeyeceklerine ilişkin öngörü yapılması sağlanmıştır.

Eğitim verilerinden elde edilen model ile test verilerine ilişkin öngörülerin yapılmasının sağlandığı bu çalışma sonucunda bankaya ait müşterilerin davranışları doğrultusunda sınıflandırılması gerçekleşmiştir. Bu gerçekleştirme sonuçları ile bankanın hizmet ve servislerini kullanmayı bırakma yöneliminde olan müşteriler ortaya çıkarılmıştır.

# '7<|FB'DF98 7H CB':CF'65B?' 7|GHCA9F

## ( SUMMARY )

Customer churn means the termination of membership agreement between the company and the customer. Companies never like emerged such a situation where customer churn is a condition encountered in each sector. It is very difficult to minimize customer churn and nowadays companies finance a lot of research in order to minimize this situation. These researches have a wide area in the Customer Relationship Management (CRM) strategies. These researches have also lead companies to have revolution from product-oriented to customer-focused. The researches, which try to minimize customer churn, are covered under Customer Churn Prediction.

In customer churn prediction researches main object is to identify the customers before customers get the decision to terminate the financial agreement, which leads maximization of profit, and protect market share. Therefore, firms now give a value on the data and use this data to identify these customers.

Particularly in banking sector where there is intense competition there is lots of data stored for each customer. Raw data is meaningless for the companies, unless analyzed. But if this data analyzed then there will be enhancing factors for the companies in terms of profit margin. Acquisition of new customers compared to the existing customer retention has very high cost, which is becoming another conscious point of banks as well. Because of this reality a lot of software have been developed in order to identify customer churn.

In customer churn researches, data mining methods are utilized. Data mining is an approach, which leads big data to be converted to be meaningful and useful data for companies. By this approach all hidden and radical data can be figured out.

Especially for rapidly growing companies, big data as a concept became quite familiar in the literature. However, unless the data remains raw, it does not make any sense for companies. In particular, modeling and the estimate approach are meaningful to this study. However, the data obtained on the data mining method to create a powerful and accurate predictive models in order for the big stage, passing through a set of data is required to be ready. First, the data set must be analyzed in a good way in order to prevent useless data to exist in this analyze process. This is because the data within the company database involves a lot of extra information. By determining which service or services to be implemented some reduction steps may be carried out in data. In parallel filtering method will be applied as well for the same purpose.

After analyzing the data with various methods work set will be ready. Usually using this model uses to model data and to make some predictions data mining methods. Also in analyze step decision trees will be applied. With decision trees work set will be formulized.

This project is also a customer churn prediction research. In this project bank customers' data has been analyzed. These data is based on customers' financial data and customers' demographic data. Financial data consist of time deposit and demand deposit balances and

transactions in three currencies, Turkish Liras, Euros and US Dollars. Demographic data consists of age and gender. Also call center calling counts of customers are taken into account.

For this research project all data are grouped as training set and test set individually. Customers are well known in training set in order to develop the algorithm, which gives best results. Each customer in this set is identified as active, inactive or churn individually. By knowing the customers, behaviors' of these customers are modeled. These models applied to test set customers according to.

In modeling phase data for each customer have been analyzed in depth in order to get properties which leads customer churn. These properties generate the decision mechanisms in the decision trees' steps.

In this research in phase of defining decision trees a lot of algorithm have been applied. This is because to compare the algorithms' effectiveness and identifying best algorithm. After applying these algorithms and comparing the results effectiveness of this research have been optimized.

In this research 250 customer data have been analyzed by keeping 50 as training and 200 for test set. Training set has been analyzed deeply in order to understand the signals on the way to terminate the membership agreement. Signals have been applied to decision tree algorithm in order to build decision tree for this research. Modeling the final Customer Churn Model, two algorithms have been applied, which are Simple Cart and LAD Tree. The reason applying these two algorithms is to find out the most successful algorithm for this research. In addition to this properties have been tested in terms of sensitivity. By doing this the properties can be compared in terms of efficiency, as well.

When the results of this research have been analyzed, it has been seen that LAD tree algorithm is most effective to use in such a research. This proofs that selection of algorithm and fine tuning of properties have a deep effect on the success rate of Customer Churn Prediction Researches.

Briefly, customer churn analysis has a wide area in customer relation management studies. It is so priceless to convert big data to meaningful data for companies to analyze and identify customers before the customers get the decision to terminate financial agreement. Conversion and analysis period of big data, decision trees and data mining methods have been applied in order to identify customers. This identification will be used to contact the customers in order to build strong relationship with the customers, which will also take companies to maximize profit and cover market share. This type of communication with the customers will be applied to general communication skills of companies and will be the base of communication strategy of the companies.

# İÇİNDEKİLER

1. GİRİŞ.....	1
<b>1.1 Genel Bilgi</b> .....	1
<b>1.2 Rapor Organizasyonu</b> .....	2
2. PROJENİN TANIMI VE PLANI.....	4
<b>2.1 Proje Tanımı ve Amacı</b> .....	4
<b>2.2 Proje Hedefi</b> .....	4
<b>2.3 Başarım Kriteri</b> .....	4
<b>2.4 Risk Yönetimi</b> .....	5
<b>2.5 Proje Planı</b> .....	6
2.5.1 Ekip Organizasyonu.....	7
<b>2.6 Proje Geliştirme Ortamı ve Proje Kaynakları</b> .....	7
3. KURAMSAL BİLGİLER.....	8
4. ANALİZ VE MODELLEME.....	12
<b>4.1 Analiz</b> .....	12
5. TASARIM, GERÇEKLEME VE SINAMA.....	13
6. DENEYSEL SONUÇLAR.....	18
7. SONUÇ ve ÖNERİLER.....	34
8. KAYNAKLAR.....	35
9. EKLER.....	37



# 1. GİRİŞ

## 1.1 Genel Bilgi

Günümüzde yeni hizmet ve servis oluşturulmasına yönelmiş her sektör için müşteri kaybı tahmini ve ölçümü önemli bir konu olmuştur. Müşteri sadakatinin analizi olarak da adlandırılan bu çalışmalar sektöre ait kurumların pazar payları üzerinde değişimlere sebebiyet vermektedir. Özellikle rekabetin üst seviyede olduğu ve yeni müşterilerin kazanımının var olan müşteriyi elde tutmaya çalışmaktan daha zor olduğu pazarlarda bu konu çok daha fazla önem teşkil etmektedir. Bankacılık da müşterilere ilişkin davranışların analiz edilmesinin ve bu davranışlar doğrultusunda müşteri kaybı tahmininin yapılmasının oldukça önemli bir çalışma konusu olduğu sektörlerdendir.

Müşteri kaybı tahmininin konu olduğu çalışmaların sonuçları, bankaların belirlediği stratejiler üzerinde önemli etkilere sahiptir. Çünkü bu çalışmalar yardımıyla ortaya çıkan sonuçlar ile bankalar müşterilere yönelik yeni stratejiler belirlemekte ya da var olan stratejilerini iyileştirmektedirler.

Müşteri kaybına ilişkin oranların önceden belirlenmesinin son derece önemli olduğu bankacılık sektöründe büyük veri kavramı özellikle son yıllarda üzerinde çalışılan konulardan olmuştur. Müşteri kaybına ilişkin ön kestirimlerin temeli de var olan ham veriden çalışma için kullanışlı olabilecek verilerin elde edilmesi ve bu verilerin modellenmesine dayanmaktadır. Bu aşama problemin çözümüne ilişkin en önemli aşama olarak kendine yer edinmiştir.

Müşteri kaybına yönelik olarak yapılan çalışmaların [1,2] kilit aşaması, çalışmada kullanılacak verilerin elde edilmesi ve elde edilen verilerden çalışma için yararlı verilerin belirlenmesi sürecidir. Literatürde yer edinen çalışmaların en büyük ortak noktası ise verilere ilişkin bu sürecin çalışmanın büyük bir kısmını oluşturduğu gerçeğidir. Bu sürecin projenin toplam süresinin %60-70'ini oluşturduğu literatür araştırmaları sırasında rastlanılan bir gerçek olmuştur [1]. Bu çalışmalarda verilere ilişkin modellemelerin ve ön kestirimlerin yapılması için veri madenciliği yöntemleri -çeşitli araçlar yardımıyla sınıflandırma algoritmaları- kullanılmıştır. Çalışma sonuçlarında ise müşterilerin verilerinden yola çıkılarak o müşterinin ileriki dönemlerde banka için Kayıp müşteri olup olmayacağına sonucuna varılmıştır [1].

Müşteri kaybına ait tahminlerin yapılmasının amaçlandığı bu çalışmada da literatürdeki çalışmalara paralel süreçlerden geçilmiştir. Banka vasıtasıyla elde edilen verilerin ham halden çalışma için anlamlı olacak hale dönüştürülmesi yani ham verinin birçok filtreleme işlemlerinden geçirilmesi aşaması çalışmada uzun zaman alan bir süreç olmuştur. Bu süreç sonucunda elde edilen "eğitim verileri" ile modelleme yapılması yoluna gidilmiş, bu amaçla karar ağacı yöntemleri ile oluşturulmuş bir model elde edilmiştir. Bu model temel alınarak ise müşterilere ilişkin "test verileri" test edilmiş ve müşterilerin Kayıp olup olmama durumları belirlenmiştir. Bu belirleme sürecinde karar ağaçlarına ilişkin algoritmalar uygulanmış buna paralel olarak da karar ağacını oluşturan ve müşterilere ait verilerin analiz edilmesi sonucu belirlenen özelliklerin değişmesi durumunda sonuçların bu

değişimden ne kadar etkileneceği de analiz edilmiştir. Böylelikle çalışmanın belirlenen özelliklere ne kadar duyarlı olduğu saptanarak bir sonuç oluşturulmuştur.

## 1.2 Rapor Organizasyonu

Projeye ilişkin olarak hazırlanan bu raporda, Giriş, Projenin Tanımı ve Planı, Kuramsal Bilgiler, Analiz ve Modelleme, Tasarım, Gerçekleme ve Sınama, Deneysel Sonuçlar, Sonuç ve Öneriler, Kaynaklar ve Ekler olmak üzere toplamda 9 bölüm bulunmaktadır. Organizasyonun bu şekilde olduğu raporda bölümlere ilişkin genel bilgiler ise şu şekildedir:

### 1. GİRİŞ

Proje raporuna ait bu bölüm içerisinde çalışmanın konusu hakkında genel bilgi verilmiş, projenin çalışma konusu üzerine yapılan literatür çalışmalarına değinilmiş ve proje raporunun bölümleri genel hatlarıyla tanıtılmıştır.

### 2. PROJENİN TANIMI VE PLANI

Raporun bu bölümü dahilinde proje tanımı yapılarak, projenin iş planı anlatılmıştır.

### 3. KURAMSAL BİLGİLER

Bu bölümde ise projenin oluşturulması süresince edinilen ve projenin daha iyi anlaşılabilir, anlatılmasını sağlayan kuramsal bilgilere yer verilmiştir.

### 4. ANALİZ VE MODELLEME

Bu bölüm içerisinde projenin analizine ilişkin açıklamalar yer almaktadır.

### 5. TASARIM, GERÇEKLEME VE SINAMA

Bu bölümde ise projenin gerçekleşmesine ilişkin aşamalara değinilmiş ve projedeki test verilerinin nasıl sınındığı açıklanmıştır.

### 6. DENEYSEL SONUÇLAR

Raporun bu alt başlığı içerisinde ise projede elde edilen sonuçlar ayrıntılı bir şekilde açıklanmıştır.

### 7. SONUÇ VE ÖNERİLER

Bu bölümde ise elde edilen deneysel sonuçlar değerlendirilmiş ve ilerleyen zamanlarda böyle bir çalışma içerisinde olacak ilgililere bir takım önerilerde bulunulmuştur.

### 8. KAYNAKLAR

Raporun bu kısmı ise rapor boyunca atıfta bulunulan referansların format dahilinde belirtildiği bölümdür.

## 9. EKLER

Raporun son kısmını oluřturan bu bölümde ise raporun akıřını engellememesi için bu başlık altında belirtilmesi uygun bulunan bilgilere yer verilmiřtir.

## 2. PROJENİN TANIMI VE PLANI

Projenin Tanımı ve Planı başlıklı bu bölümde proje tanımı, kapsamı ve proje hedefi belirtilmiştir. Bölümün ilerleyen aşamalarında ise proje planına yer verilmiştir.

### 2.1 Proje Tanımı ve Amacı

Bu proje, bankacılık sektörüne ait müşterilerin bankaya olan sadakatlerinin/ bağlılıklarının ölçülmesine yönelik olarak yapılan bir çalışmadır. Proje, bankanın sahip olduğu müşterilerin davranışlarından yola çıkarak müşterilerin bankadan ayrılma potansiyellerinin olup olmadığını ya da böyle bir potansiyelin olmaması durumunda ise bankayı ne kadar aktif olarak kullandıklarını belirlemeye yönelik olarak oluşturulmuştur.

Projede bankadan ham olarak elde edilen verilerden çalışma için anlam teşkil edecek verilerin elde edilmesi ve bunu takiben bu yararlı verilerle makine öğrenmesi yöntemleri kullanılarak anlamlı sonuçlar çıkarılması amaçlanmıştır. Bu anlamlı sonuçların ise müşteri kaybının tahmin edilmesine yönelik olması temel hedefdir. Bu şekilde amacı ortaya konulan proje ile birlikte ham olarak bankada yer edinen büyük miktardaki verilerden anlamlı bilgiler oluşturulması sağlanmıştır. Bu anlamlı verilerin banka içerisinde daha verimli olarak kullanılmasının sağlanması ise projenin en genel ve birincil hedefidir.

### 2.2 Proje Hedefi

Projeye ilişkin hedef ise yukarıda belirtilmiş olan amaç doğrultusunda bankadaki büyük verilerden bankanın strateji belirleme aşamasında faydalanabileceği ve müşteri ilişkilerine yönelik bir takım iyileştirmeler yapabileceği bilgi dönüşümünün sağlanmasıdır. Genel olarak ham veriden müşterilerin Kayıp olup olmayacaklarının ya da Kayıp değilse bankayı hangi aktiflikte kullandıklarına ilişkin bilgilerin elde edilmesi bu proje sonunda görülmesi beklenen sonuçtur.

### 2.3 Başarım Kriteri

Projenin gerçekleşmesi için bir takım başarı kriterleri belirlenmiştir. Bu bağlamda proje bütününün başarıya ulaşmasını sağlayan kriterler aşağıdaki gibi belirtilebilir:

- Projede Kullanılacak Yazılım Bileşenlerinin Belirlenerek Elde Edilmesi
- Belirlenen Banka Yardımıyla Analizi Yapılacak Verilerin Elde Edilmesi
- Elde Edilen Verilerin Kapsamlı Analizler Sonucunda Filtrelenerek Bir Elemeden Geçirilmesi
- Verilere İlişkin Modellemelerin Yapılması
- Elde Edilen Modeller Yardımıyla Verilerin Sınıflandırılması

## 2.4 Risk Yönetimi

Projenin gerçekleştirilmesi sırasında karşılaşılan riskler ve bu risklerin hangi yöntemler yardımı ile ortadan kaldırıldığına ilişkin açıklamalar aşağıda belirtilerek tablo haline getirilmiştir.

Projedeki en önemli risk faktörü projede kullanılacak olan verilerin bankadan elde edilememesidir. Verilerin elde edilmesi projeye ilişkin en temel adım olduğundan bu riske karşı önlemlerin alınması son derece önemlidir. Bu nedenle böyle bir risk ile karşılaşılmasını engellemek ya da karşılaşılması durumunda alınacak önlemleri belirlemek gerekmektedir. Böyle bir riskin varlığından ötürü verilerin elde edilmesi için birden fazla bankaya istekte bulunulmuştur. Verilerin elde edilmesi ve kullanılması aşamasında “veri gizliliği” ilkeleri benimsenerek, bankanın çalışmaya güven duyması sağlanmıştır.

Projede karşılaşılabilecek bir diğer risk ise projede kullanılacak algoritmanın seçimine ilişkindir. Ancak çalışma sırasında birden fazla algoritmanın kullanılması ve bu algoritmalarından elde edilen sonuçların karşılaştırılması bu riski minimize eden bir etken olmuştur.

Verilerin elde edilmesinin sağlanmasının ardından verilerle ilgili olarak karşılaşılabilecek bir diğer kuvvetli risk faktörü ise verilerin çalışma için gerekenden çok daha fazla bilgiye sahip olmasıdır. Bu durum çalışmanın sağlıklı bir şekilde yürütülmesini engelleyeceğinden bu riskin ortadan kaldırılması gerekmektedir. Bu ise elde edilen verilerin gerekli analizler ve uzman- tez danışmanı eşliğinde gerçekleştirilecek toplantılar sonucunda verilen kararlar doğrultusunda veri üzerinde filtrelemenin yapılması ile sağlanacaktır. Böylelikle verinin istenilen formata dönüşümü sağlanacak, çalışmaya bu şekilde devam edilecektir.

**Tablo 2.1:** Risk

Risk	Riske Karşı Alınan Önlem
Proje Kapsamında Kullanılacak Verilerin Elde Edilememesi	Projede yavaşlamaya sebep olan bu riski ortadan kaldırmaya yönelik olarak birden fazla bankadan veri isteme işlemi gerçekleştirilmiştir. Bu doğrultuda banka ile verilerin gizliliği hakkında anlaşmaya varılmıştır.
Projede Kullanılacak Algoritmalar Hakkında Karar Verilememesi	Bu riski ortadan kaldırmak için birden fazla algoritmanın kullanılması kararı alınmıştır.
Projede Elde Edilen Verilerin Çalışma Kapsamında Olmayan Bir Çok Ekstra Bilgiye Sahip Olması	Bu durumda elde edilen veriler üzerinde çeşitli filtreleme işlemleri uygulanmıştır.

Ön kestirimlerle belirlenen bu risk sınıflandırılmasına ilişkin etki çarpanları ise raporun devamındaki tabloda görülmektedir. Bu etki çarpanları yardımı ile hangi riskin daha fazla önem teşkil ettiği ve çalışmayı ne oranda yavaşlatabileceği kolaylıkla anlaşılmıştır. Bu ise riskleri minimize etme aşamasında yararlanılacak bir hesaplama olmuştur.

**Tablo 2.2:** Belirtilen Risklere İlişkin Etki Çarpanlarının Hesaplanması

Risk	Önem Faktörü (Düşük:1, Orta:2, Yüksek:3)	Olasılık [0.0,1.0]	Etki Çarpanı (Önem Faktörü x Olasılık)
Proje Kapsamında Kullanılacak Verilerin Elde Edilememesi	3	0.8	2.4
Projede Kullanılacak Algoritmalar Hakkında Karar Verilememesi	2	0.5	1.0
Projede Elde Edilen Verilerin Çalışma Kapsamında Olmayan Bir Çok Ekstra Bilgiye Sahip Olması	3	0.7	2.1

## 2.5 Proje Planı

Projenin tanım ve amacının belirlenmesinin ardından proje planının oluşturulması aşamasına geçilmiştir. Bu aşamada projeye ilişkin zaman kestirimlerinde bulunmuş ve proje süresince bu kestirimlere uyulmaya çalışılmıştır. Aynı zamanda oluşturulan bu proje planının proje boyunca bir rota olarak kullanılması söz konusu olmuştur. Proje aşamalarının göz önüne alınarak oluşturulduğu bu planda bir diğer dikkat edilen nokta ise projeye ilişkin zaman kestirimlerinin proje süresince yardımları alınan uzmanların etkisi nedeniyle esnek olması gerektiğidir. Uzman desteği ile çözüme kavuşturulan noktalarda

uzmanların zaman planlarına uyum söz konusu olmuştur. Bu uyumun sağlanmaya çalışılması ise planda bir takım değişikliklere sebep olmuştur. Ancak planda zaman aralıklarının yukarıda belirtilen riskler göz önüne alınarak %20 oranında arttırılması bu değişikliklerin projeyi aksatmasını engellemiştir.

Projeye ilişkin zaman kestirimlerinin yer aldığı diyagram “Ekler” bölümünde yer almaktadır.

## 2.5.1 Ekip Organizasyonu

Proje çalışmasında bir grup çalışması söz konusu değildir. Proje bireysel olarak oluşturulmuştur. Ancak bankacılık sektörüne ait uzmanlarla yoğun iş birliği gerektiren bu projede, uzmanlar ile gerçekleştirilen toplantılar belirli bir organizasyon dahilinde gerçekleşmiştir.

## 2.6 Proje Geliştirme Ortamı ve Proje Kaynakları

Çalışmada kaynak olarak önceki yıllarda konu ile ilgili oluşturulmuş projelere ilişkin makalelerden yararlanılmıştır. Rapor boyunca gerekli yerlerde bu kaynaklara atıfta bulunularak çalışmanın güvenilirliği arttırılmıştır.

Çalışmada kullanılan algoritmaların gerçekleşmesinde kullanılan yazılım bileşeni Weka olmuştur. Weka veri madenciliği uygulamaları için kullanılacak olan makine öğrenmesi algoritmaları topluluğudur [3].

Çalışmanın veri elde edilmesi/analizi ve filtrelenmesi aşamalarında ise banka veri tabanından elde edilen veriler Microsoft Excel 2013 ortamına taşınarak, verilere ilişkin çalışmaların burada yapılması sağlanmıştır.

Proje boyunca dokümantasyon için MS Word 2013, zaman diyagramlarının oluşturulması için ise GanttProject kullanılmıştır.

Proje boyunca kullanılan kaynakların sınıflandırılması ile birlikte gösterimi ise şu şekildedir:

Yazılım Bileşeni:

Weka Tool

Eclipse

Verilerin Saklandığı Ortam:

Microsoft Excel 2013

Dokümantasyon:

Microsoft Excel 2013

GanttProject

### 3. KURAMSAL BİLGİLER

Günümüzde müşteri kaybının tahmini olarak kendine yer edinmiş olan kavram, dünya literatüründe karşımıza “churn prediction” olarak çıkmaktadır. Müşteri kaybı (customer churn) bir hizmet/ servisi kullanmayı bırakan müşterilerin ölçülmesi olarak tanımlanabilir [1]. Bir başka tanıma göre ise müşteri kaybı, müşterinin o ana kadar bağlı olduğu kurum/firma ile üyelik anlaşmasını sonlandırması ve buna paralel olarak başka bir rakip firmanın müşterisi haline gelmesini ifade etmektedir[9]. Müşteri kaybı müşteri sadakati kavramı ile birebir olarak ilgilidir. Özellikle her bir bireyin belli bir servis/ hizmeti belirli bir kurumdan sağladığı (Banka, Telekomünikasyon) ve pazar rekabetinin en üst noktaya ulaştığı sektörlerde kurumlar Müşteri İlişkileri Yönetimine (CRM) daha çok odaklanmaya başlamıştır [4]. Bu alanın müşteri kaybını önlemeye yönelik çalışmaları özellikle kurumlar açısından büyük önem taşımaktadır. Çünkü araştırmalar gösteriyor ki, bir kurum için yeni müşteri kazanmak, var olan müşteriyi elde tutmaya yönelik çalışmalar yapmaktan çok daha fazla maliyetlidir [5]. Yeni müşteri kazanma çalışmalarının gitme potansiyeli olan müşteriye yönelik olarak yapılan çalışmalardan beş kat daha maliyetli olduğu gözler önüne serilmiştir [6,7]. Bu nedenle özellikle son yıllarda iş zekası uygulamalarının hedef noktası müşterilere ait büyük verilerden kurumlar için yarar teşkil edebilecek bilgilere dönüşümün sağlanması ve bu bilgiler ışığında kuruma ait müşteri kayıp oranının tahmin edilmesidir. Bu hedef noktasının seçiminin altında ayrıca kaybedilen müşterinin aynı zamanda yeni müşteri kazanımına negatif etkide bulunabileceği gerçeği de yatmaktadır. Kısacası, rekabetin had safhaya ulaştığı sektörler için müşteri kaybı tahminlerinin gerçekleştirilmesine yönelik olarak yapılan çalışmalar kurumlar için oldukça kritik çalışmalar olarak görülebilir. Ayrıca bu çalışma sonuçları ile bankaların gelecek dönemlere ait stratejilerinde bir takım değişimler ya da iyileştirmeler yapılabilir.

Bir bankaya ait müşteri aşağıda belirtilen üç durumdan birine ait olabilir. Bu durumlar Aktif olma durumu, İnaktif olma durumu ve Kayıp olma durumudur. Aktif- İnaktif- Kayıp müşteri sınıflandırması hemen hemen her sektör için geçerli olmakla beraber çalışmanın odak noktasının bankacılık sektörü olmasından dolayı tanımlar bu sektör doğrultusunda verilmiştir.

**Aktif Müşteri:** Bu sınıfa mensup müşteriler banka için en yararlı müşteriler olarak tanımlanabilir. Çünkü müşterinin banka ile olan ilişkisi son derece kuvvetlidir ve müşteri finansal işlemleri için bankaya sürekli olarak para yatırımı yapmaktadır.

**İnaktif Müşteri:** Bu müşteri gruplarının banka ile olan ilişkileri daha azdır. Ancak yine de bankanın müşterisi olmaya devam etmektedirler. Müşterilerin finansal işlemleri aktif müşteriler kadar olmamakla beraber, kayıp potansiyeli olan müşteriler kadar da az değildir.

**Kayıp Müşteri:** Bu müşteri grubu banka ile olan tüm ilişkisini kesmiştir ya da bankayı terk etme eğilimine yönelik davranışlarda bulunmaktadır. Müşteri kaybının tahmini çalışmalarının da odak noktası bu müşteri grubudur. Bu nedenle bu müşteri grubuna ait davranışların analiz edilerek belirlenmesi, bankalar açısından hayati önem taşımaktadır.

Bankacılık sektörünün ilgi alanı içerisinde önemli bir payı olan müşteri kaybı tahminin yapılması banka müşterilerine ait büyük verilerle analizlerin gerçekleştirilmesi demektir. Bu aşama ise oldukça kapsamlı çalışmaları gerektirmektedir. Bu amaçla öncelikle müşteri kaybının çeşitleri bilinmeli ve banka için gerçekten negatif etki yaratacak kayıp tipinin detaylı bir şekilde belirlenmesi gerekmektedir.



Bir bankaya ait müşteriler birçok sebepten bankadan servis/ hizmet almayı kesebilirler. Bu açıdan bakıldığında genel olarak müşteri kayıpları 2 sınıfa ait olabilir. Bu 2 sınıf gönüllü kayıplar ve gönülsüz kayıplardır.

**Gönüllü Kayıp:** Bu kayıp türünde müşteriler bankadan servis ya da hizmet almayı isteyerek bırakmıştır. Müşteri başka bir servis, ürün ya da hizmet sağlayıcısına geçme kararı almış olabilir. Bir başka olasılık olarak da o sektöre ait servis, ürün ya da hizmet kullanımını tamamen bırakmış olabilir.

**Gönülsüz Kayıp:** Bu kayıp tipinde ise müşteriler istemeyerek ya da zorunda olarak servis/ hizmet sağlayıcısını kullanmayı bırakmıştır. Bu eylemin altında yatan sebepler, müşteri ölümü, müşterinin başka bir lokasyona yerleşmek durumunda olması ve müşterinin bankanın “kara listesine” alınması gibi zorunda kalmayı gerektiren durumlar olabilir [2,8].

Müşteri kaybının tahminine ilişkin çalışmalarda genellikle gönülsüz kayıp sınıfına ait olan kayıp müşteriler çalışmanın dışarısında tutulmaktadır[8]. Çünkü bu sınıfa ait kayıp müşteriler davranış analizleri için sağlıklı sonuçları temsil etmemektedir.

Müşteri kaybının tahmini için yapılan çalışmalarda en önemli nokta büyük verilerin yönetiminin sağlanmasıdır. Firmaların veri tabanlarında bulunan ham veriler işlenmedikçe çoğu zaman firma için bir anlam teşkil etmemektedir. Bununla birlikte her geçen gün bu ham veri miktarı artmaktadır. Bu durum ise firmalar için büyük verilerin işlenmesine yönelik çalışmaları zorunlu hale getirmiştir. Bu nedenle başta bankacılık sektörü olmak üzere birçok kurum için veri madenciliği oldukça önemli bir konumda yer almaktadır. Çünkü veri madenciliği metotları yardımı ile büyük veriler üzerinde değişik bakış açıları ile analiz işlemleri gerçekleştirilebilmekte ve büyük veriler bankalar için anlamlı bilgilere dönüştürülebilmektedir [1,10]. Veri madenciliği bu açıdan bakıldığında “büyük ve anlaşılması zor veri gruplarından saklı bilgileri gün ışığına çıkararak, kullanılabilir bir formata sokmaktır” denilebilir. Bu nedenle veri madenciliği uygulamaları ile müşteri kaybının tahminine yönelik çalışmalar oluşturulabilmektedir.

Müşteri kaybı çalışmalarında veri madenciliği yöntemleri yardımı ile kayıp müşterileri modelleyerek kayıp oranının tahmin edilmeye çalışılmasından önce aşılması gereken çok önemli bir köşe taşı vardır. Bu köşe taşı ise büyük miktardaki verilerin filtrelenmesi gerektiğidir. Veri madenciliği çalışmalarında kullanılacak olan verinin filtrelenerek hazırlanması hayati bir adım olarak görülmektedir. Bu işlemin örnek çalışmaların zaman planının %60-70’ini kapsaması bu hayati önemi ispatlar niteliktedir [1,2]. Filtreleme işlemlerinin gerçekleştirilmesi için verinin bir takım aşamalardan geçirilmesi gerekmektedir. Bu aşamalar verinin temizlenme sürecine yönelik çalışmalardan oluşmaktadır. Verilerin temizlenmesinde eksik değerler ve gürültülü değerler ele alınabilir. Eksik veriler üzerinde uygulanabilecek metotlar ise şu şekilde sıralanabilir:

- Eksik verinin analiz işleminden çıkarılması
- Eksik olan değerlerin el ile doldurulması
- Eksik değerler için önceden belirlenmiş olan değerlerin atanması
- Eksik değerlerin ortalama değer ile değiştirilmesidir.

Gürültülü değerler ise çalışmayı hataya sürükleyebileceğinden çeşitli yöntemlerle yok edilmesi ya da değiştirilmesi gereken verilerdir. Bu amaçla kutulama, demetleme ve regresyon yöntemleri kullanılabilir. Bu üç yöntem de gürültülü değerlerin değiştirilmesine ya da elenmesine yöneliktir [11].

Eksik ve gürültülü değerlere yönelik yöntemin uygulanmasının ardından verinin hazırlanması ve filtrelenmesi aşamasına geçilir. Bu aşama verinin niteliğine göre değişiklik göstermektedir. Her veri türü için veri küçültülmesine ve filtrelenmesine yönelik olarak göz önüne alınacak özellikler farklı olmakla beraber bu bölümün bu kısmında projenin çalışma alanının banka sektörü olması nedeniyle banka müşterilerine yönelik verilerin küçültülmesi ve filtrelenmesi aşamasına ilişkin kuramsal bilgiler açıklanacaktır.

Bir veri setindeki veriler üç türden birine ait özelliklere ilişkin olabilir [2]:

**Müşteri Davranışı:** Müşteri davranışlarını temsil eden veriler müşterilerin bir servis ya da hizmetin hangi bölümlerini kullandığına ya da bu servis/hizmeti ne sıklıkta kullandığına yönelik bilgiler içeren verilerdir.

**Müşteri Algısı:** Müşteri algısını temsil eden veriler ise müşterilerin servis/hizmete yönelik düşüncelerine ilişkin verilerdir.

**Demografik Ögeler:** Müşteriye ait demografik ögelerin hepsi bu veri tipinde yer alır. Bu demografik ögeler müşterinin yaşı, cinsiyeti, eğitim seviyesi ve sosyal statüsü gibi özelliklerine ait olabilir.

Bankacılık sektöründe müşteri kaybı çalışmaları yapılırken bu üç türe ait veriler de kullanılabilir. Elde edilen bu verilerden belirli özellikleri taşımayanların elenmesi yöntemiyle veri madenciliği çalışmaları daha yüksek başarı oranları ile sonuçlandırılabilir. Banka ürünlerinden bir kısmına yönelik verilerin elde edilmesi literatürdeki çalışmalarda gözlemlenen veri küçültülmesi işlemlerinden biridir [1, 2, 12 ve 13]. Bununla birlikte çalışma, müşterilerden sadece bir kısmına yönelik olarak da gerçekleştirilebilir. Literatürde yer edinen çalışmaların büyük bir kısmı bireysel müşterileri temel alarak oluşturulmuş çalışmalardır [1].

Müşteri kaybı tahmini için filtrelenen ve küçültülen veriler, veri madenciliği uygulamaları yardımı ile yararlı bilgiler üretecek hale dönüştürülebilir. Bu dönüştürme veri madenciliği metotları yardımı ile sağlanmaktadır. Veri madenciliği, veri tabanlarında bulunan veriler arasında ilişkiler, örüntüler, yapılar gibi bilgilerin keşfedilerek ortaya çıkarılması işlemidir [11]. Bankacılıkta da müşteri kayıp oranının önceden belirlenmesi adına yapılan çalışmalarda veri madenciliğinin modelleme ve tahmin etmeye yönelik olan metotları kullanılmaktadır. Bu amaca yönelik çalışmalarda modelleme amacıyla kullanılan temel teknik karar ağaçlarıdır (decision trees).

Karar ağaçları veriyi sınıflandırma ve veri üzerinde tahminlerde bulunma çalışmalarında sıkça kullanılan algoritmalarındandır. Az sayıda özellik yardımıyla bile hızlı bir öğrenme mekanizmasına sahip olan karar ağaçları her öğrenme gerektiren algoritmalarda kullanılabilir. [14]. Karar ağaçları ağaç yapısının oluşturularak makine öğrenmesinin sağlandığı yapılardır. Bu yapılarda her ağaçta olduğu gibi düğümler ve yapraklar bulunmaktadır. Ağaç üzerindeki yapraklarda hedef nitelikler yer almaktadır. Bu hedef nitelik ait olunan sınıf ya da tahmin sonucunda verinin alacağı değer olabilir. Ağaç

üzerinde yer alan diğer düğümler ise karar mekanizmalarını oluşturmaktadır. Bu düğümler yardımı ile ise veri test edilerek, verinin bu düğümden yer alan özelliğe ilişkin sahip olduğu değer sonucu ile bir sonraki düğümün hangisi olacağına karar verilir.

Karar ağacının bu yapısı yardımı ile müşteri kaybına ilişkin çalışmalar yürütülebilir. Müşteri kaybının tahmin edilmesi için veriler “eğitim seti” ve “test seti” olarak ikiye ayrılmaktadır. Eğitim seti içerisindeki veriler müşterilere ilişkin özelliklere yönelik olmakla beraber bu verilerin ait olduğu sınıf bilinmektedir. Bu veriler ve her verinin ait olduğu sınıf yardımıyla karar ağaçları oluşturulur. Daha sonraki aşamada bu karar ağaçları test seti için model oluşturacaktır. Elde edilen bu model yardımı ile ait olduğu sınıfa ilişkin bilginin bilinmediği test setine ait verilerin sınıflarının belirlenmesi sağlanır.

Karar ağacı algoritmaları çok çeşitlidir. Bu çalışmada dahilinde kullanılan karar ağacı algoritmaları ise CART (Classification And Regression Tree) ve LAD Tree algoritmalarıdır. Bu algoritmaların uygulanmasına ilişkin bilgiler raporun ilerleyen kısımlarında kolaylıkla görülebilmektedir.

## 4. ANALİZ VE MODELLEME

### 4.1 Analiz

Projenin konu aldığı problem, bankacılık sektöründeki müşteri kaybı oranının tahmin edilmesidir. Buna paralel olarak probleme ilişkin çözüme kavuşturulması gereken bir diğer nokta ise Kayıp olma potansiyeli olmayan müşterilerin bankayı aktif olarak kullanıp kullanmadıklarıdır.

Problem, banka müşterilerinin dahil olduğu bir senaryo içerisinde yer almaktadır. Bu senaryoda banka müşterileri banka ile etkileşimlerini gösteren hesap hareketleri ve demografik özellikleri doğrultusunda ele alınacak veri kümesini oluşturmaktadır. Bu veri kümesi modellemenin yapılacağı “eğitim seti” ve müşteri kaybına ilişkin tahminlerin yapılacağı “test seti” olarak ayrılmaktadır. Bu açıdan bakıldığında veri setleri problem domenine ilişkin “eğitim seti” ve “test seti” olarak iki ayrı yapıyı temsil etmektedir.

Probleme ilişkin bileşenlerden olan veri setinin analiz edilmesi aşaması problemin temelini teşkil etmektedir. Büyük veriler ile analiz işleminin gerçekleştirilmesi oldukça dikkatli bir şekilde yürütülmesi gereken bir süreçtir. Bu süreçte verilerin çalışma için anlamlı hale getirilmesi veri küçültme ve filtreleme çalışmaları ile sağlanmıştır.

Verilerin anlamlı hale getirilmesi ile problem domenine ait olan veri seti bileşeninin küçülmesi ve bu verilerin daha etkili özelliklere sahip olması sağlanmıştır.

Oluşturulan anlamlı verilerden yola çıkılarak problem domenine ilişkin bileşenlerin arasındaki bağlantılardan anlamlı sonuçların çıkarılması problemin çözüme kavuşturulduğu aşamayı temsil etmektedir.

## 5. TASARIM, GERÇEKLEME VE SINAMA

Bankacılık sektörü için müşteri kaybı tahmininin yapılmasını amaçlayan bu projede ilk süreç proje süresince analiz edilecek olan verilerin elde edilmesidir. Projenin tasarımı ve gerçekleştirilmesi bu süreçle başlar. Literatürde yer alan başarılı çalışmada da belirtildiği gibi verinin modellenmesi ve modele ilişkin tahminlerin yapılmasına dayalı çalışmalarda verinin elde edilmesi ve hazırlanması süreci çalışmanın büyük bir kısmını oluşturmaktadır [1]. Bu projede de verilerin elde edilmesi ve hazırlanması sürecinin oldukça zorlayıcı ve zaman alıcı bir aşama olduğunu belirtmek doğru bir açıklama olacaktır.

Verilerin elde edilmesi için birçok banka ile iletişime geçilmiştir. Bir kısmı “müşteri bilgilerinin gizliliği” gerekçesi ile böyle bir paylaşımda bulunmayı kabul etmemiştir. İletişime geçilen bankalardan bir kısmı ise istenilen formatta veri paylaşımında bulunamamıştır. Çalışmada kullanılan veriler ise veri alışverişinde verici taraf olmayı kabul eden büyük bir bankadan elde edilmiştir. Ancak bu çalışma süresince banka adının gizli tutulması konusunda anlaşmaya varılmıştır.

Banka yardımıyla alınan verilerin hangi ürün, hizmet ya da servise ait olacağı ya da hangi müşteri grubunu içereceği proje kapsamı doğrultusunda belirlenmiştir. Bankalar birçok ürün ile müşterilerine hizmet sunmaktadır. Bu proje kapsamında ise üzerinde çalışılacak ürünler müşterilere ait vadeli, vadesiz mevduat hesapları ile bireylere ait kredi hesaplarıdır. Bu nedenle veri içerisinde diğer ürünlere ait olan bilgiler analizin dışında tutulmuştur. Buna paralel olarak değerlerinde büyük oranda bilinmeyen taşıyan veriler de analize dahil edilmemiştir. Çalışma bireysel banka müşterilerine yönelik olduğundan verilerin yalnızca bireysel müşterilere ait hesaplardan oluşmasına dikkat edilmiştir. Ayrıca veriler içerisindeki bir takım eksik değerlerin çalışmanın verimliliği açısından doldurulması işlemi gerçekleştirilmiştir. Bu aşamada raporun “Kuramsal Bilgiler” bölümünde değinilen yöntemlerden faydalanılmıştır. Yukarıda da belirtildiği gibi verinin büyük bir kısmının eksik değerlerden oluştuğu veriler analizden çıkarılmıştır. Bu “Kuramsal Bilgiler” bölümünde değinilen “Eksik Değeri Analizden Çıkarmak” seçeneği doğrultusunda uygulanan bir yöntemdir. Veri içerisindeki eksik değerlerin verinin boyutu yanında oldukça küçük kaldığı durumlarda ise veri elle doldurulmuş ya da eksik değerler içeren niteliğin ortalama değeri ile veri tamamlanmıştır. Örneğin, bir aya ya da haftaya ait vadeli TL mevduat hesabında eksik değer bulunduran veri, bir önceki ve bir sonraki ay ya da haftaya bakılarak doldurulmuştur. Aynı zamanda bu doldurma işlemi sırasında bankaya ait o ayki faiz oranları da göz önünde bulundurulmuştur. Ancak verilerin önemli bir kısmının çok fazla eksik değer içermemesi bu sürecin daha hızlı tamamlanmasını sağlamıştır.

Müşterilere ait veriler vadeli TL, Dolar ve Euro hesapları ile vadesiz TL, Dolar ve Euro hesaplarından oluşmaktadır. Bunun yanı sıra veriler içerisinde müşterilerin ev ve araba kredilerine yönelik bilgiler de yer almaktadır. Bu özellikteki veriler, müşterilerin davranışlarına yönelik verilerdir ve “Kuramsal Bilgiler” bölümünde açıklanan veri türlerinden müşteri davranışına ait değerler içeren veri türü sınıfıyla bire bir örtüşmektedir. Müşterilerin hesap hareketlerini göz önüne seren bu veri tipleri için üç senelik bir zaman aralığı belirlenmiştir. Üç senelik zaman dilimi için net bir tarih aralığı belirlenmemiştir.

Elde edilen veriler arasında müşterilere ait demografik değerler içeren özellikler de bulunmaktadır. Müşterilerin yaş ve cinsiyet bilgileri bu tür veriler arasında gösterilebilir. Bu türdeki veriler yine “Kuramsal Bilgiler” başlığı altında belirtilen veri türleri arasından demografik öğeler içeren veri türüne dahildir.

Yukarıda adı geçen özelliklere ek olarak verilerin ayrıca müşteri algısını yansıtan bir takım özellikler barındırmasına da önem verilmiştir. “Kuramsal Bilgiler” bölümünde açıklaması yapılan ve müşterinin kuruma yönelik olarak düşüncelerini yansıtan veriler özellikle müşteri kaybının tahminine ilişkin çalışmalarda önem arz etmektedir. Bu çalışmaya da bu amaçla müşterinin son zamanlarda yaptığı çağrı merkezi görüşmelerinin sayısı dahil edilmiştir. Özelliğin adı anlaşılabilirliği arttırmak için “Çağrı Merkezi Görüşme Sayısı” olarak belirtilmiş olsa da gerçekte bu özellik müşterinin banka şubesini arama sayısına aittir. Bu özelliğin kullanılması ise banka uzmanları tarafından önerilen bir konu olmuştur. Uzmanlardan alınan bilgiler doğrultusunda bankayı arayan müşterilerin çoğunlukla mutsuz ve verilen hizmetten memnuniyetsiz müşteri grubuna ait müşteriler olduğu sonucuna ulaşılmıştır.

Veri hazırlanmasına ilişkin aşamanın tamamlanması sonucunda 250 kişilik müşteri grubuna ait veriler üzerinde çalışılmaya karar verilmiştir. Daha sonrasında bu 250 kişiye ait veri gurubu “eğitim seti” ve “test seti” olarak ikiye ayrılmıştır. Eğitim setini oluşturan veri grubuna ait verilerin hangi sınıfa ait olduğu önceden bilinmemektedir. Yani bu gruba ait müşterilerin Aktif, İnaktif ya da Kayıp olup olmadıkları belirtilmiştir. Çalışma, müşterilerin kayıp olup olmaması durumu yanında müşterilerin aktif ya da İnaktif olup olmama durumlarını da inceleyeceğinden verilerin 3 sınıfa ait olması gerekmiştir. Eğitim seti 50 kişilik müşteri grubundan oluşmaktadır. Bu müşteri grubu, durumu bilinmeyen müşterilerin oluşturduğu “test seti” için bir model teşkil etmiştir. Bu model yardımıyla “test veri seti” üzerinden tahminlerde bulunulmuştur.

“Eğitim seti” doğrultusunda model oluşturma ise karar ağaçları algoritmaları ile gerçekleştirilmiştir. Karar ağaçları oluşturulurken birden fazla algoritma kullanılmıştır. Böylelikle çalışmanın hangi algoritma ile daha sağlıklı sonuçlar ürettiğine ilişkin sonuçlar çıkarılmıştır. Proje dahilinde kullanılan algoritmalar CART ve LAD Tree algoritmalarıdır. Bu algoritmalar yardımıyla “eğitim setine” ait modeller oluşturulmuştur. Bu algoritmaların kullanılması ise Weka aracı sayesinde gerçekleşmiştir.

Karar ağaçlarının oluşturulması için bir takım karar mekanizmalarının belirlenmesi gerekmektedir. Bu çalışma için karar verme mekanizmaları, verilerden yola çıkılarak belirlenen müşterilere ait özelliklerdir. Bu özelliklerin adları ve açıklamaları aşağıda sıralanmıştır:

**Toplam Para Ortalaması:** Bu özellik müşteriye ait tüm vadeli – vadesiz hesaplardaki para toplamalarının belirlenen süre içerisindeki ortalamasıdır. Bu özellik sayısal bir değerdir.

**Toplam Paranın Azalmayan Olması Durumu:** Bu özellik ise müşterinin o anki parasının belirlenen zaman aralığının en başındaki miktara göre durumunu belirtir. Buna göre müşteri o an elinde zaman aralığının başındaki miktardan daha az ya da daha fazla miktarda para bulundurabilir. Bu özellik her bir müşteri için “Evet” ya da “Hayır” değerlerinden birini almaktadır.

**Tehlike Artışı:** Bu ise müşterinin tehlike potansiyeline işaret eden bir özelliktir. Tehlike artışının olduğu durumlarda “Var” , tehlike artışının olmadığı durumlarda ise “Yok” değerini almaktadır. Müşteriye ait tehlike potansiyeli ise bayrak ataması yöntemi ile hesaplanmaktadır. Bu bayrak ataması yönteminin nasıl uygulandığı ise bir örnek ile betimlenebilir. Bu amaçla müşterinin öncelikle son 6 aydaki toplam para ortalaması hesaplanır. Bu para ortalamasının müşterinin 3 senelik toplam para ortalamasına bölümü

ile bir değere ulaşılır. Aynı işlem müşterinin 1 senelik toplam para ortalamasının 3 senelik toplam para ortalamasına bölümü ile tekrarlanır. Bu iki işlem sonucunda elde edilen değerler ise birbiri ile karşılaştırılarak tehlike artışının olup olmadığı sonucuna ulaşılır. Bu iki değer karşılaştırılması ise değerlere içinde buldukları aralık doğrultusunda bayrak atamasının yapılması ile gerçekleştirilir. Buna göre değer aralıkları ve bu değerlere ilişkin bayrak değerleri şöyledir:

**Tablo 5.1:** Müşteriye Ait Tehlike Artışı Hesabında Kullanılan Bayrak Değerleri

Değer Aralığı	Bayrak Değeri
%1- %30	Kırmızı
%31- % 60	Turuncu
%61- %90	Sarı
%91- %120	Mavi
%121- %150	Koyu Yeşil
%151- %180	Yeşil
%181 ve üzeri	Açık Yeşil

Bu tablo yardımıyla bayrak değerinde negatif bir değişim olan müşteriler için bu özellik “Var ” değerini alırken, bayrak değerinde değişiklik olmayan ya da bayrak değeri iyileşen müşteriler için bu özellik “Yok” değerini almaktadır.

**Dalgalanma:** Bu özellik ise müşteriye ilişkin hesap hareketlerindeki dalgalanmayı izah etmektedir. Ancak burada belirtmeye çalışılan dalgalanma negatif yönlü bir dalgalanmadır. Müşterinin son aylara ilişkin toplam para miktarının önceki aylara ilişkin toplam para miktarından az olması durumunda müşteriye ilişkin negatif bir dalgalanmanın söz konusu olduğu söylenebilir. Bu özelliğe ilişkin değer, müşterinin son aylara ilişkin para ortalamasının daha önceki aylara ilişkin para ortalamasından %20 veya daha fazlası kadar az olması durumunda “Var” değerini alırken, diğer durumlarda “Yok” değerini almaktadır. Bu %20’lik oran çalışmanın ilerleyen bölümlerinde değiştirilerek çalışma özelliklere olan duyarlılık açısından da ele alınmıştır.

**Kapanmış Hesap:** Bu özellik ise müşterinin son yılında toplam hesap sayısında azalma olup olmaması durumunu temsil etmektedir. Son bir yılda eğer müşteri toplam hesap sayısının bir bölü dördü ya da daha fazlası kadar hesabını kapatmışsa bu özellik o müşteri için “Var” değerini almaktadır. Diğer durumlarda ise bu özellik o müşteri için “Yok” değerini almaktadır.

**Ev veya Araba Kredisi:** Bu özellik ise müşteriye ait kredi hesaplarının olması halinde “Var” değerini alırken, kredi hesabı olmayan müşteriler için “Yok” değerini almaktadır.

**Çağrı Merkezi Görüşme Sayısı:** Bu özellik ise müşterinin son dönemlerde kaç defa banka şubasını aradığını gösterir. Değeri sayısaldır.

**Pozitif İvme:** Pozitif ivme özelliği ise müşterinin son 6 aya ilişkin toplam para ortalamasında pozitif bir ivmenin olup olmadığını gösterir. Son 6 aya ilişkin toplam para ortalamasının bir önceki 6 aya ilişkin toplam para ortalamasının %20 veya daha fazlası

olması durumunda bu özellik o müşteri için “Var” değerini alırken, diğer durumlarda “Yok” değerini almaktadır. 6 ay olarak belirtilen zaman aralığı ve %20 olarak belirtilen oran çalışmanın ilerleyen kısımlarında daha farklı sayılarla değiştirilerek de denenmiştir. Bu sayede proje sonucunun özelliklere ilişkin sayısal belirteçlere ne oranda bağlı olduğu sonucuna ulaşılmıştır.

Hesap Sayısında Artma: Bu özellik ise müşterinin son aylarında hesap sayısında artma olması durumunda “Var” değerini alırken, diğer durumlarda “Yok” değerini almaktadır.

Yaş: Bu özellik o müşterinin yaşını belirten bir özelliktir.

Cinsiyet: Müşterinin cinsiyetine ilişkin bilgi içeren özelliktir.

Durum: Çalışmanın sonucunda gözlemlenmek istenen sonuç bu özellik değerinde saklıdır. Eğitim seti verileri için bu özellik Aktif, İnaktif ya da Kayıp değerlerinden birini alırken, test seti için bu özellik çalışma sonucunda belirlenecektir.

Bu özellikler doğrultusunda karar ağaçları oluşturulmuştur. Karar ağaçları oluşturulurken CART ve LAD Tree algoritmaları kullanılmıştır. Bu algoritmalar Weka aracı vasıtası ile uygulanmıştır. Weka makine öğrenmesi çalışmalarının yapıldığı bir araçtır ve literatürde yer alan çalışmalarda da sıklıkla kullanılmıştır [1]. Weka, ARFF formatındaki veriler üzerinde çalışmaktadır. ARFF, veriler ve bu verilere ilişkin özellik ile bu özelliklere ait değerlerin belirtildiği bir dosya türüdür. Proje dahilinde de belirlenen özelliklere ilişkin ARFF dosyaları oluşturulmuştur.

Projede verilere ilişkin özelliklerin belirlenmesinin ardından, her bir müşteriye ait veri için bu özelliklerin değerleri doldurulmuştur. Ancak çalışmada belirlenen özelliklerin barındırdığı sayısal kısımlar değiştirilerek çalışma birçok kez tekrarlanmıştır. Böylelikle projenin belirlenen özelliklere ne kadar duyarlı olduğu ve bu özelliklerin değişiminin çalışma sonucunu ne oranda etkilediği gözlemlenmiştir. Daha açık bir şekilde anlatmak gerekirse çalışma boyunca kullanılan özellikler şu değişikliklere maruz kalmıştır:

Öncelikle hesap hareketlerine ilişkin zaman aralıkları sırasıyla 3 sene, 2 sene ve 1 sene olarak değiştirilmiştir. Daha sonrasında ise negatif dalgalanmanın ve pozitif ivmenin olup olmadığı kararı verilirken dikkat edilen oran sırasıyla %10 - %20 - %30 olarak değiştirilmiştir. Bunu takiben tehlike artışı özelliğinin değerinin belirlenmesinde kullanılan bayrak aralıkları sırasıyla %30 ve %20 olarak değiştirilmiştir. Bunlara ek olarak ise yukarıdaki özellikleri açıklarken kullanılan “müşterinin son aylarına ilişkin” şeklindeki ibarelerde belirtilen son ay zaman aralığı sırasıyla 6 ay ve 3 ay olarak değiştirilmiştir. Bu şekilde özellikler değişik kombinasyonlarla değiştirilmiştir. Bu kombinasyonlarla birlikte eğitim setindeki verilere ait olmak üzere toplamda 30 adet ARFF dosyası hazırlanmıştır. Bu 30 ARFF dosyası model oluşumu sırasında kullanılacaktır. Diğer taraftan yine aynı şekilde test setine ait veriler için de 30 adet ARFF dosyası hazırlanmıştır. Bu 30 adet dosya ise oluşturulan model yardımıyla durumları tahmin edilecek olan müşterilere ait verileri içermektedir.

Kullanılan algoritmalar ve oluşturulan modeller ile bu modeller baz alınarak yapılan tahminlere ilişkin sonuçlar bir sonraki bölümde görsel veriler ile açıklanmıştır. Ayrıca çalışmanın doğruluğunun değişik özellik kombinasyonları ve farklı algoritma kullanımları



ile test edilme aşaması sonucunda elde edilen sonuçlara ilişkin başarı oranları da ilerleyen bölümde belirtilen bir konu olmuştur.

## 6. DENEYSEL SONUÇLAR

Çalışma sırasında karar ağaçları, “Tasarım Gerçekleme Sınama” bölümünde açıklanan özelliklerin değişik kombinasyonları ile oluşturulmuştur. Her bir kombinasyon için ayrı bir karar ağacı elde edilmiştir. Bunun yanı sıra her bir kombinasyon için CART ve LAD Tree algoritmaları için kullanılmıştır. Bu sayede her bir kombinasyona ilişkin 2 ayrı ağaç yapısı elde edilmiştir. Toplamda ise 30 adet özellik kombinasyonuna karşılık gelen 60 tane ağaç yapısı elde edilmiştir.

Yukarıda bahsedilen özellik kombinasyonları ise şu şekilde listelenebilir:

- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon

- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon

- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon
- Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon

Bu aşamada bu kombinasyonlardan birini örnek olarak açıklamak yerinde olacaktır. Örneğin “müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyonu” şu anlama gelmektedir:

Müşterinin 3 senelik hesap hareketleri baz alınarak bu hesap hareketleri 6 aylık periyodlara bölünecektir. Her bir 6 ayın kendi içerisindeki karşılaştırılmasında sınır %30 olarak belirlenecektir. Yani negatif dalgalanmanın ya da pozitif ivmenin varlığı birbirini izleyen 6 aylık periyodların oranlarının bu sınır dahilinde olup olmaması durumuna göre belirlenecektir. Ayrıca müşteriye ait özelliklerden biri olan bayrak ataması yöntemi ile tehlike artışının tespitinde ise bayrak aralıkları 30 olacaktır.

Yukarıda açıklandığı üzere her bir kombinasyon bu şekilde CART ve LAD Tree algoritmaları ile modellenmiştir. Daha sonra ise bu modeller yardımı ile test verilerine

ilişkin sonuçlar çıkarılmıştır. Her bir kombinasyonun farklı algoritmalar ile modellenmesi sonucu ortaya çıkan karar ağaçları bölümün aşağıdaki kısımlarında görülmektedir. Ayrıca bu model ile test verilerinin test edilmesi sağlanarak müşterilerin “aktif”, “İnaktif” ve “kayıp” sonuçlarından hangisine ait oldukları saptanmıştır. Yine bu sonuçlara ilişkin görseller ve açıklamalar ileriki kısımda yer almaktadır.

Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyona ilişkin veri setinin “LAD Tree” karar ağacı algoritması ile modellenmesi sonucu oluşan karar ağacı EK-B bölümünde görülmektedir. Bu modelin baz alınarak test verilerinin test edilmesi sonucunda ise elde edilen bir tablo olarak EK-C bölümünde görülmektedir.

Bu şekilde tüm kombinasyonlara ait 2 farklı algoritmanın kullanılması sonucu elde edilen modeller ve bu modellerin kullanılması ile test veri seti için üretilen sonuçlar EK-B ve EK-C bölümlerinde görülebilir. Ancak verilerin oldukça fazla yer kaplamasından dolayı tüm sonuçlara ilişkin tabloların rapora eklenmesi söz konusu olamamaktadır. Bu amaçla EK bölümüne çalışmanın örnek bir yansıması eklenerek, çalışmanın bu yöntem doğrultusunda yürütüldüğü açıklanmıştır. Aynı zamanda Ekler bölümünde çalışmanın en başarılı sonuç verdiği müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyonun LAD Tree algoritması ile modellenmesi sonucu elde edilen sonuçlar görülebilir.

Tüm kombinasyonların 2 farklı algoritma ile modellenmesi ve bu modeller doğrultusunda test verilerine ilişkin sonuçların üretilmesiyle birlikte elde edilen sonuçlar aşağıdaki tabloda belirtilmiştir:

**Tablo 6.1:** Deneysel Sonuçlara Ait Tablo

Kombinasyon	Algoritma	Başarı Oranı (Kayıp)	Başarı Oranı (Aktif)	Başarı Oranı (İnaktif)
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,882352941	0,757575758	0,240601504
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır	LAD TREE	0,681818182	0,612903226	0,223404255

oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon				
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,875	0,757575758	0,236220472
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,85	0,737704918	0,222222222
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,882352941	0,79245283	0,265486726
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,928571429	0,75862069	0,245614035
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,875	0,79245283	0,261682243

Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,939393939	0,75862069	0,247706422
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,882352941	0,561643836	0,161290323
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,8	0,679245283	0,205607477
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,875	0,591549296	0,168539326
Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,822222222	0,68627451	0,211538462
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış	Simple CART	0,88	0	0,211428571

ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon				
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,851851852	0,587301587	0,2
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,88	0	0,211428571
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,884615385	0,587301587	0,198198198
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,88	0	0,211428571
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının	LAD TREE	0,851851852	0,587301587	0,2



30 olduđu kombinasyon				
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,88	0	0,211428571
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	LAD TREE	0,896551724	0,606557377	0,2
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	Simple CART	0,693877551	0,793103448	0,193548387
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	LAD TREE	0,711538462	0,811320755	0,210526316
Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,701754386	0,867924528	0,2

Müşterinin 3 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,705882353	0,814814815	0,189473684
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,842105263	0,757575758	0,217054264
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,878787879	0,709677419	0,219047619
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,853658537	0,757575758	0,222222222
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD Tree	0,909090909	0,709677419	0,228571429
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış	Simple CART	0,842105263	0,79245283	0,23853211

ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon				
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,80952381	0,75862069	0,23
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,80952381	0,75862069	0,23
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,818181818	0,745762712	0,226804124
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,842105263	0,761363636	0,22972973
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının	LAD TREE	0,857142857	0,742857143	0,221052632

30 olduđu kombinasyon				
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,846153846	0,761363636	0,232876712
Müşterinin 2 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	LAD TREE	0,861111111	0,732394366	0,215053763
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	Simple CART	0,814814815	0,554216867	0,144444444
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	LAD TREE	0,756756757	0,5	0,165048544
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,814814815	0,554216867	0,144444444

Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,8	0,558823529	0,184782609
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,814814815	0,545454545	0,186046512
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,756756757	0,5	0,165048544
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,814814815	0,545454545	0,186046512
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,8	0,558823529	0,184782609
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış	Simple CART	0,745098039	0,803571429	0,247311828

ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon				
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,794117647	0,793650794	0,242718447
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,8	0,803571429	0,235294118
Müşterinin 2 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,742857143	0,793650794	0,235294118
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,652173913	0,666666667	0,102272727
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının	LAD TREE	0,75	0,590909091	0,160714286

30 olduđu kombinasyon				
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,652173913	0,666666667	0,102272727
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	LAD TREE	0,764705882	0,634146341	0,166666667
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	Simple CART	0,652173913	0,947368421	0,2
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyon	LAD TREE	0,75	0,590909091	0,160714286
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 20 olduđu kombinasyon	Simple CART	0,652173913	0,947368421	0,2

Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,764705882	0,634146341	0,166666667
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	Simple CART	0,652173913	0,884615385	0,215686275
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyon	LAD TREE	0,763157895	0,727272727	0,208333333
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	Simple CART	0,652173913	0,884615385	0,215686275
Müşterinin 1 senelik hesap hareketlerine ilişkin 3 aylık periyodları inceleyen artış ve azalışların %10 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyon	LAD TREE	0,763157895	0,727272727	0,208333333

Tablo incelendiğinde her bir kombinasyona ait 2 farklı algoritmanın verdiği sonuçlar doğrultusunda oluşan başarı oranları görülecektir. Bu başarı oranları algoritma ve kombinasyonların kayıp, inaktif ve aktif müşterileri saptama başarılarına göre belirlenmiştir. Algoritmaların test verileri için ürettiği sonuçlar test verisi olarak kullanılan



müşterilere ait gerçek değerler ile karşılaştırılmıştır. Bu karşılaştırma sonucunda ise elde edilen oranlar başarı oranlarını oluşturmuştur.

Tablo detaylı bir şekilde incelendiğinde projede en iyi sonucun 3 senelik hesap hareketlerinin 6 aylık periyodlara bölüdüğü ve özelliklere ilişkin oran sınırlarının %20 olarak belirlendiği aynı zamanda da tehlike artışını simgeleyen bayrak aralıklarının 20 olduğu kombinasyonun LAD Tree algoritması ile modellenmesi ile elde edildiği görülmektedir. Bu algoritma ile kayıp ve aktif müşterilerin tespiti oldukça başarılıdır. İnaktif müşterilerin saptanmasına ilişkin sonuçlar ise bu başarı oranlarına nazara daha düşük kalmaktadır.

Tabloda dikkat çekici olan bir diğer nokta ise 1 senelik hesap hareketlerine ilişkin kombinasyonların bir kısmı dışında tüm kombinasyonların kayıp müşterileri saptamak konusunda oldukça başarılı olduğudur.

Aktif müşterileri saptama konusunda LAD Tree algoritması ile modellenen veriler Simple Tree algoritması ile modellenen verilere oranla çok daha iyi sonuçlar vermektedir. Özellikle Simple Tree algoritması 3 aylık periyodlar söz konusu olduğunda aktif müşteri saptama konusunda başarısız kalmaktadır. LAD Tree modeli ise özellikle 3 senelik hesap hareketleri söz konusu olduğunda aktif müşterileri yakalama konusunda oldukça yüksek başarı oranları yakalamıştır.

Zaman aralığı 1 seneye indirildiğinde ise başarı oranları düşüş göstermektedir. Bu nedenle bu çalışma için verilerin 1 senelik zaman dilimi içerisinde değerlendirilmesi çok iyi sonuçlar vermeyecektir. Verilerin 3 senelik zaman dilimi dahilinde değerlendirilmesi çok daha sağlıklı sonuçlar üretecektir. Zaman periyodları söz konusu olduğunda ise tablodan hareketle 6 aylık periyodların daha iyi başarı oranları getirdiği açıktır.

Kısacası çalışma LAD Tree algoritması kullanıldığında daha yüksek başarı oranları ile sonuçlanmıştır. Zaman diliminin 3 sene olduğu zaman periyodlarının ise 6 ay ile sınırlandırıldığı çalışmalarda sonuçlar oldukça tatmin edicidir. Özellikle oranın %20 olduğu bayrak ataması yönteminde de aralıkların 20 olarak belirlendiği kombinasyon için çalışma en yüksek başarı oranını yakalamıştır. Bu da çalışmadaki özellik kombinasyonu ve algoritma seçiminin ne kadar önemli olduğunu ispatlar niteliktedir.

## 7. SONUÇ ve ÖNERİLER

Bu proje ile birlikte banka müşterilerine ait verilerden yola çıkılarak müşteri kaybı tahmininin yapılması sağlanmıştır. Bu sonucun yanı sıra Kayıp olma potansiyeli olmayan müşterilerin bankayı hangi aktiflikte kullandıklarına ilişkin sonuçlar da üretilmiştir. Bu sonuçların belirlenen özelliklere hangi oranlarda bağlı oldukları ise özelliklere ilişkin sayısal niceliklerin değiştirilmesi ile belirlenmiştir. Bunun yanı sıra proje kapsamında birden fazla karar ağacı kullanılarak, çalışmanın hangi karar ağacı için daha iyi sonuçlar verdiği analiz edilmiştir.

Proje performans açısından oldukça hızlı ve etkili sonuçlar vermektedir. Ancak projede verilerin elde edilme ve bu verilerin proje için anlamlı hale getirilmesi aşaması oldukça uzun süren süreçler olmaktadır. Verileri temsil edecek özelliklerin belirlenmesinin ardından model oluşturma ve test verilerine ait tahminlerin yapılması süreci oldukça hızlı gelişmiştir.

Proje maliyet açısından ise oldukça olumlu sonuçlar içermektedir. Proje dahilinde kullanılan tüm yazılım bileşenleri ücretsizdir. Ayrıca bankadan verilerin elde edilmesinin ve bu verilerin anlamlı hale dönüştürülmesi sürecinin de hiçbir maliyeti olmamıştır.

Çalışma literatürde kendine yer edinmiş çalışmalardaki başarılı sonuçlara yakın sonuçlar vermektedir. Bu açıdan bakıldığında projenin başarı oranının yüksek olduğu söylenebilir [1].

Projenin olumlu yönleri büyük veriler içerisindeki gizli bilgileri çıkarmakken olumsuz yönü verinin elde edilmesi sürecinin oldukça zorlayıcı olmasıdır. Bu nedenle verinin elde edilmesi süresi boyunca araştırmacılar sabırlı olmalı ve mümkün olan en yararlı veri seçimini yapabilmelidirler.

Projenin veri elde edilmesi aşaması oldukça zorlu ve uzun bir süreç olarak proje zaman planında yer almıştır. Bu konu üzerine çalışacak kişilerin verileri daha hızlı ve etkili bir biçimde elde etmeleri faydalarına olacaktır. Bununla beraber elde edilen verilerin sayısı da arttırılarak müşteri kaybına ilişkin kestirimler yapılabilir. Tüm bunlara ek olarak aynı zamanda daha farklı karar ağacı algoritmaları kullanılarak da çalışma geliştirilebilir.

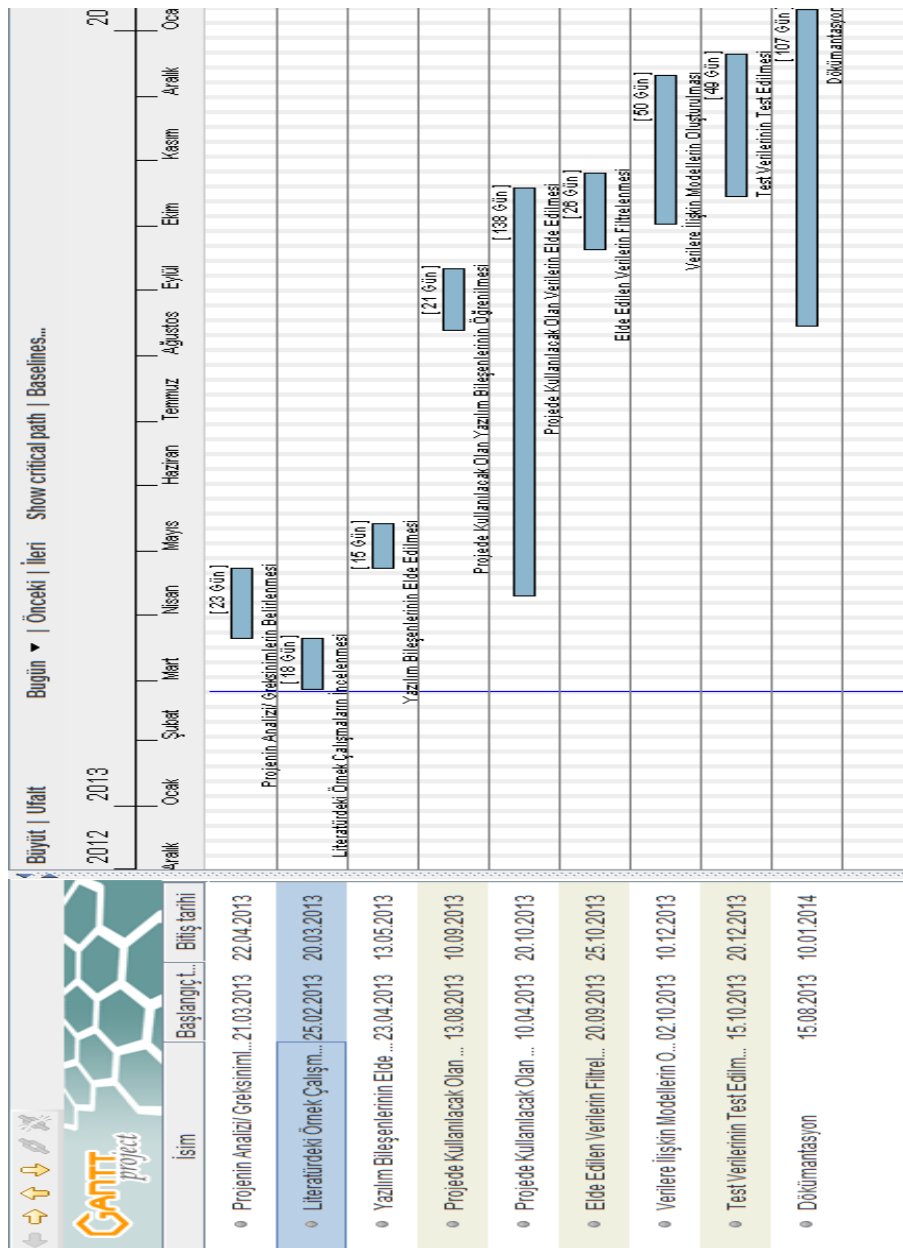
## 8. KAYNAKLAR

- [1] Devi Prasad, S. Madhavi, "Prediction of Churn Behavior of Bank Customers Using Data Mining Tools", *Business Intelligence Journal*, vol. 5, No. 1, January 2012.
- [2] Vladiaslav Lazarov, Marius Capota, "Churn Prediction", 2008, <http://home.in.tum.de/~lazarov/files/research/papers/churn-prediction.pdf>
- [3] David Scuse, Peter Reutemann, *WEKA Experimenter Tutorial for Version 3-5-5*, University of Waikato, January 26, 2007.
- [4] Luo Bin; Peiji, Shao; Liu Juan, "Customer Churn Prediction Based on the Decision Tree in Personal Handyphone System Service," *Service Systems and Service Management*, 2007 International Conference on , vol., no., pp.1,5, 9-11 June 2007
- [5] Jae-Hyeon Ahna., Sang-Pil Hana, Yung-Seop Leeb, "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry", *Telecommunications Policy*, 2006, pp. 561-562.
- [6] Dixon Mary, "39 Experts Predict the Future", *America's Community Banker*, vol. 8, No. 7, 1999, pp. 20-31.
- [7] Slater, Stanley F. and Narver John C, "Intelligence Generation and Superior Customer Value", *Journal of the Academy of Marketing Science*, vol. 28, No. 1, 2000, pp.120-127.
- [8] Gordon S. Linoff, "Survival Data Mining for Customer Insight ", *Data Miners*, June 2004, pp.7
- [9] M. Richeldi, A. Petrucci, "Churn Analysis Case Study", *Telecom Italia Lab.*, Torino, 2002, p. 3-6.
- [10] Feng Bao, Xu He, Fengzhi Zhao, "Applying data mining to the geosciences data", *Computer, Mechatronics, Control and Electronic Engineering (CMCE)*, 2010 International Conference on , vol.5, no., pp.290,293, 24-26 Aug. 2010.
- [11] Tuğba Tosun, "Veri Madenciliği Teknikleriyle, Kredi Kartlarında Müşteri Kaybetme Analizi", 2004.
- [12] Ning Wang; Dong-xiao Niu, "Credit card customer churn prediction based on the RST and LS-SVM," *Service Systems and Service Management*, 2009. ICSSSM '09. 6th International Conference on , vol., no., pp.275,279, 8-10 June 2009
- [13] Yin Wu; Jiayin Qi; Wang, C., "The study on feature selection in customer churn prediction modeling," *Systems, Man and Cybernetics*, 2009. SMC 2009. IEEE International Conference on , vol., no., pp.3205,3210, 11-14 Oct. 2009
- [14] Iqbal, M.R.A.; Rahman, S.; Nabil, S.I.; Chowdhury, I.U.A., "Knowledge based decision tree construction with feature importance domain knowledge," *Electrical & Computer Engineering (ICECE)*, 2012 7th International Conference on , vol., no.,

pp.659,662, 20-22 Dec. 2012.

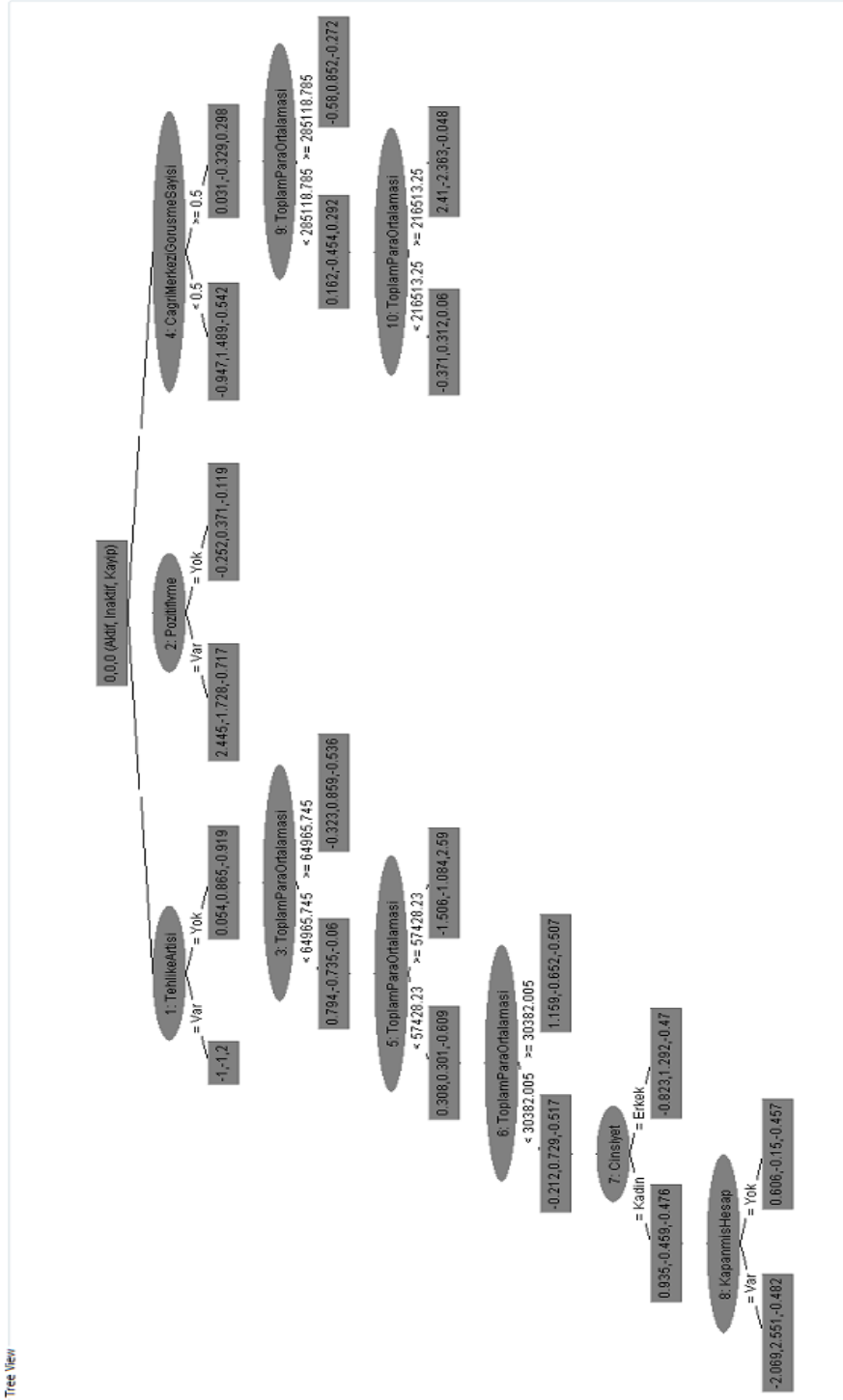
## 9. EKLER

### EK-A: Zamanlama Diyagramı

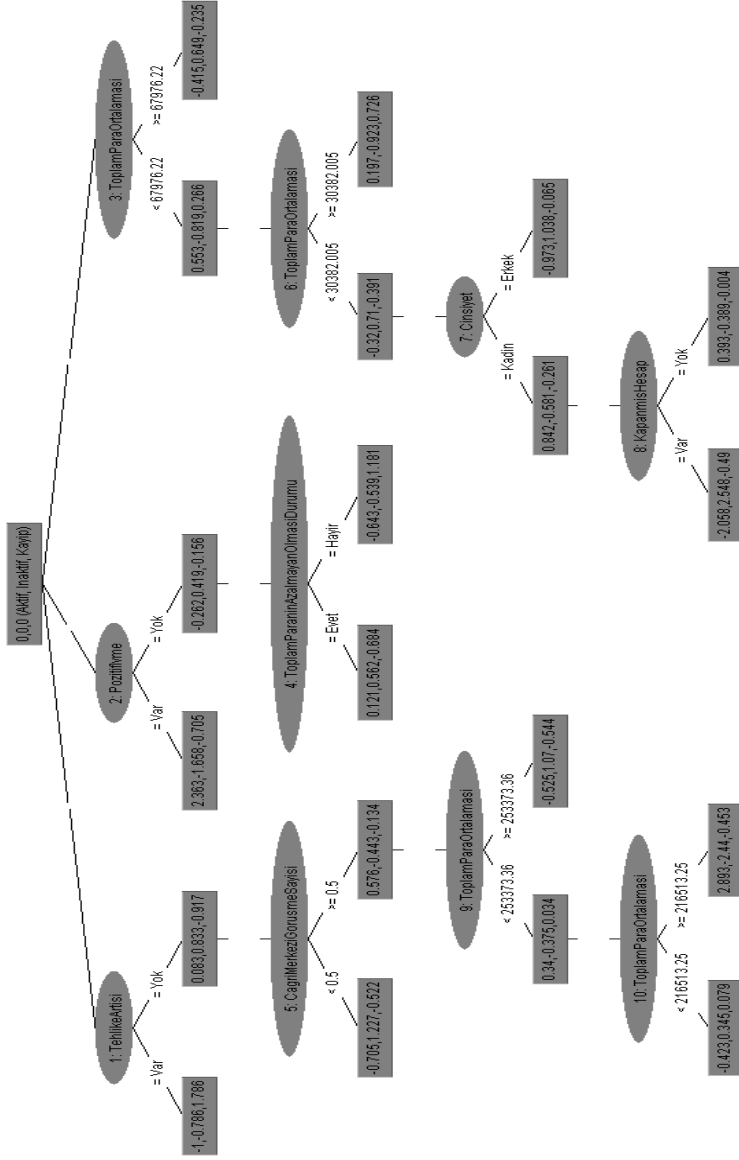


Şekil 1: Projeye İlişkin Zaman Planı

## EK-B: Karar Ağaçları



**Şekil 2:** Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyona ilişkin veri setinin “LAD Tree” karar ağacı algoritması ile modellenmesi



**Őekil 3:** MüŐterinin 3 senelik hesap hareketlerine iliŐkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralıđının 20 olduđu kombinasyona iliŐkin veri setinin “LAD Tree” karar ađacı algoritması ile modellenmesi

## EK-C: Sonular

**Tablo 9.1:** Mşterinin 3 senelik hesap hareketlerine iliřkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiđi tehlike artışına ait bayrak aralığının 30 olduđu kombinasyona iliřkin veri setinin "Simple CART" karar ađacı algoritması ile modellenmesinin ardından bu model ile test verisinin test edilmesi sonucunda elde edilen sonular.

Veri	Sonu
A2	İnaktif
A3	İnaktif
A5	İnaktif
A6	İnaktif
A7	İnaktif
A8	İnaktif
A9	İnaktif
A10	Kayıp
A11	İnaktif
A12	Kayıp
B2	İnaktif
B3	İnaktif
B4	İnaktif
B5	Aktif
B6	İnaktif
B9	İnaktif
B10	Aktif
B11	İnaktif
B12	İnaktif
C5	İnaktif
C6	Aktif
C7	Aktif
C8	Kayıp
C9	İnaktif
C10	Kayıp
C11	Aktif
D2	İnaktif
D3	İnaktif
D4	İnaktif
D5	İnaktif
D6	İnaktif
D7	İnaktif
D8	İnaktif
D9	İnaktif
D10	İnaktif
D11	İnaktif



D12	Kayıp
E2	Aktif
E3	İnaktif
E4	Aktif
E5	İnaktif
E6	Kayıp
E7	İnaktif
E8	İnaktif
E9	İnaktif
E10	İnaktif
E11	Kayıp
E12	İnaktif
F4	İnaktif
F5	İnaktif
F6	İnaktif
F7	Kayıp
F9	Aktif
F10	İnaktif
F12	Kayıp
G2	Aktif
G3	İnaktif
G4	Kayıp
G6	İnaktif
G7	İnaktif
G9	İnaktif
G10	İnaktif
G11	İnaktif
G12	Kayıp
H	Aktif
H2	İnaktif
H3	Kayıp
H5	İnaktif
H6	İnaktif
H7	Kayıp
H9	İnaktif
H11	Aktif
H12	Kayıp
J3	İnaktif
J4	İnaktif
J7	İnaktif
J8	İnaktif
J9	İnaktif
J10	İnaktif
J11	İnaktif
J12	Kayıp

K2	İnaktif
K3	İnaktif
K4	İnaktif
K5	İnaktif
K8	Aktif
K9	Aktif
K11	İnaktif
K12	Kayıp
L1	İnaktif
L2	Kayıp
L3	İnaktif
L4	İnaktif
L5	İnaktif
L6	İnaktif
L7	İnaktif
L8	İnaktif
L9	İnaktif
L10	Kayıp
L11	İnaktif
M	İnaktif
M1	İnaktif
M2	Aktif
M3	İnaktif
M4	İnaktif
M5	İnaktif
M6	Aktif
M7	İnaktif
M8	İnaktif
M9	İnaktif
M10	İnaktif
M11	İnaktif
N1	İnaktif
N2	İnaktif
N3	Aktif
N4	Kayıp
N5	İnaktif
N6	İnaktif
N7	Aktif
N8	İnaktif
N9	İnaktif
N10	İnaktif
N11	İnaktif
O	İnaktif
O1	Aktif
O2	Aktif

O3	Kayıp
O4	Kayıp
O5	İnaktif
O6	İnaktif
O7	İnaktif
O8	Aktif
O9	Aktif
O10	İnaktif
O11	Kayıp
P2	Kayıp
P3	İnaktif
P4	İnaktif
P5	İnaktif
P6	İnaktif
P7	İnaktif
P8	Aktif
P9	İnaktif
P10	İnaktif
P11	Aktif
R	Kayıp
R1	Aktif
R3	İnaktif
R4	Aktif
R5	İnaktif
R6	İnaktif
R7	İnaktif
R8	İnaktif
R9	İnaktif
R10	Kayıp
R11	İnaktif
S1	İnaktif
S2	Kayıp
S3	İnaktif
S4	İnaktif
S5	İnaktif
S6	İnaktif
S7	İnaktif
S8	İnaktif
S9	Kayıp
S10	Kayıp
S11	İnaktif
T2	Kayıp
T3	İnaktif
T4	Aktif
T5	İnaktif

T6	İnaktif
T7	İnaktif
T8	İnaktif
T9	Aktif
T10	İnaktif
T11	Kayıp
V	İnaktif
V2	İnaktif
V3	İnaktif
V4	Aktif
V5	İnaktif
V6	İnaktif
V7	İnaktif
V8	İnaktif
V9	Aktif
V10	Kayıp
V11	Kayıp
Y	Aktif
Y1	Aktif
Y2	İnaktif
Y3	Aktif
Y4	İnaktif
Y5	İnaktif
Y6	İnaktif
Y7	İnaktif
Y8	Aktif
Y9	Kayıp
Y10	İnaktif
Y11	Kayıp

**Tablo 9.2:** Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %30 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 30 olduğu kombinasyona ilişkin veri setinin “LAD TREE” karar ağacı algoritması ile modellenmesinin ardından bu model ile test verisinin test edilmesi sonucunda elde edilen sonuçlar.

Veri	Sonuç
A2	Kayıp
A3	İnaktif
A5	İnaktif
A6	Aktif
A7	İnaktif
A8	İnaktif
A9	Aktif
A10	Kayıp
A11	İnaktif
A12	Kayıp
B2	İnaktif
B3	İnaktif
B4	İnaktif
B5	Aktif
B6	İnaktif
B9	İnaktif
B10	Aktif
B11	Aktif
B12	Aktif
C5	Aktif
C6	Aktif
C7	Aktif
C8	Kayıp
C9	Aktif
C10	Kayıp
C11	Aktif
D2	İnaktif
D3	Aktif
D4	İnaktif
D5	İnaktif
D6	Aktif
D7	Aktif
D8	İnaktif
D9	İnaktif
D10	İnaktif
D11	İnaktif
D12	Kayıp
E2	Aktif

E3	İnaktif
E4	Aktif
E5	İnaktif
E6	Kayıp
E7	İnaktif
E8	Kayıp
E9	Aktif
E10	Aktif
E11	Kayıp
E12	İnaktif
F4	İnaktif
F5	İnaktif
F6	İnaktif
F7	Kayıp
F9	Aktif
F10	İnaktif
F12	Kayıp
G2	Aktif
G3	İnaktif
G4	Kayıp
G6	İnaktif
G7	İnaktif
G9	İnaktif
G10	Aktif
G11	İnaktif
G12	Kayıp
H	Aktif
H2	Aktif
H3	Kayıp
H5	Aktif
H6	İnaktif
H7	Kayıp
H9	İnaktif
H11	Aktif
H12	Kayıp
J3	Aktif
J4	Aktif
J7	İnaktif
J8	Aktif
J9	Aktif
J10	İnaktif
J11	İnaktif
J12	Kayıp
K2	İnaktif
K3	İnaktif

K4	İnaktif
K5	Aktif
K8	İnaktif
K9	İnaktif
K11	İnaktif
K12	Kayıp
L1	Aktif
L2	Kayıp
L3	İnaktif
L4	İnaktif
L5	İnaktif
L6	İnaktif
L7	İnaktif
L8	İnaktif
L9	İnaktif
L10	Kayıp
L11	İnaktif
M	İnaktif
M1	Aktif
M2	Aktif
M3	İnaktif
M4	İnaktif
M5	İnaktif
M6	Aktif
M7	İnaktif
M8	Aktif
M9	İnaktif
M10	İnaktif
M11	İnaktif
N1	İnaktif
N2	Aktif
N3	Aktif
N4	Kayıp
N5	İnaktif
N6	İnaktif
N7	Aktif
N8	Kayıp
N9	İnaktif
N10	İnaktif
N11	İnaktif
O	İnaktif
O1	Aktif
O2	Aktif
O3	Kayıp
O4	Kayıp

O5	İnaktif
O6	Kayıp
O7	Aktif
O8	Aktif
O9	İnaktif
O10	İnaktif
O11	Kayıp
P2	Kayıp
P3	Aktif
P4	Aktif
P5	Aktif
P6	Aktif
P7	İnaktif
P8	Aktif
P9	İnaktif
P10	İnaktif
P11	Aktif
R	Kayıp
R1	Aktif
R3	İnaktif
R4	Aktif
R5	Aktif
R6	İnaktif
R7	İnaktif
R8	İnaktif
R9	Kayıp
R10	Kayıp
R11	İnaktif
S1	İnaktif
S2	Kayıp
S3	İnaktif
S4	Kayıp
S5	Kayıp
S6	İnaktif
S7	İnaktif
S8	Kayıp
S9	Kayıp
S10	Kayıp
S11	Kayıp
T2	Kayıp
T3	İnaktif
T4	İnaktif
T5	İnaktif
T6	İnaktif
T7	İnaktif



T8	İnaktif
T9	Aktif
T10	Aktif
T11	Kayıp
V	İnaktif
V2	Kayıp
V3	İnaktif
V4	İnaktif
V5	Aktif
V6	Aktif
V7	İnaktif
V8	İnaktif
V9	Aktif
V10	Kayıp
V11	Kayıp
Y	İnaktif
Y1	Aktif
Y2	İnaktif
Y3	Aktif
Y4	İnaktif
Y5	Aktif
Y6	Aktif
Y7	İnaktif
Y8	Aktif
Y9	Kayıp
Y10	Aktif
Y11	Kayıp

**Şekil 9.3:** Müşterinin 3 senelik hesap hareketlerine ilişkin 6 aylık periyodları inceleyen artış ve azalışların %20 sınır oranı ile belirtildiği tehlike artışına ait bayrak aralığının 20 olduğu kombinasyona ilişkin veri setinin “LAD TREE” karar ağacı algoritması ile modellenmesinin ardından bu model ile test verisinin test edilmesi sonucunda elde edilen sonuçlar.

Veri	Sonuç
A2	İnaktif
A3	İnaktif
A5	İnaktif
A6	İnaktif
A7	Aktif
A8	Aktif
A9	İnaktif
A10	Kayıp
A11	İnaktif
A12	İnaktif
B2	İnaktif
B3	İnaktif
B4	Aktif
B5	Aktif
B6	İnaktif
B9	İnaktif
B10	Aktif
B11	İnaktif
B12	İnaktif
C5	İnaktif
C6	Aktif
C7	Aktif
C8	Kayıp
C9	Kayıp
C10	Kayıp
C11	Aktif
D2	İnaktif
D3	İnaktif
D4	İnaktif
D5	Aktif
D6	İnaktif
D7	Aktif
D8	İnaktif
D9	İnaktif
D10	İnaktif
D11	İnaktif
D12	İnaktif
E2	Aktif

E3	Aktif
E4	Aktif
E5	Aktif
E6	Kayıp
E7	Kayıp
E8	İnaktif
E9	İnaktif
E10	İnaktif
E11	Kayıp
E12	İnaktif
F4	Kayıp
F5	İnaktif
F6	İnaktif
F7	Kayıp
F9	Aktif
F10	İnaktif
F12	Kayıp
G2	Aktif
G3	Aktif
G4	Kayıp
G6	İnaktif
G7	İnaktif
G9	İnaktif
G10	Aktif
G11	İnaktif
G12	Kayıp
H	Aktif
H2	Aktif
H3	İnaktif
H5	İnaktif
H6	İnaktif
H7	Kayıp
H9	İnaktif
H11	Aktif
H12	İnaktif
J3	İnaktif
J4	İnaktif
J7	Kayıp
J8	Aktif
J9	İnaktif
J10	İnaktif
J11	İnaktif
J12	Kayıp
K2	İnaktif
K3	İnaktif

K4	İnaktif
K5	İnaktif
K8	Aktif
K9	Aktif
K11	İnaktif
K12	Kayıp
L1	İnaktif
L2	Kayıp
L3	İnaktif
L4	İnaktif
L5	Aktif
L6	İnaktif
L7	İnaktif
L8	Aktif
L9	Aktif
L10	Kayıp
L11	İnaktif
M	İnaktif
M1	Aktif
M2	Aktif
M3	İnaktif
M4	Aktif
M5	Aktif
M6	Aktif
M7	İnaktif
M8	İnaktif
M9	İnaktif
M10	İnaktif
M11	İnaktif
N1	İnaktif
N2	İnaktif
N3	Aktif
N4	Kayıp
N5	İnaktif
N6	İnaktif
N7	Aktif
N8	Aktif
N9	Kayıp
N10	İnaktif
N11	İnaktif
O	İnaktif
O1	Aktif
O2	Aktif
O3	Kayıp
O4	Kayıp

O5	İnaktif
O6	İnaktif
O7	Kayıp
O8	Aktif
O9	Aktif
O10	İnaktif
O11	Kayıp
P2	Kayıp
P3	İnaktif
P4	İnaktif
P5	İnaktif
P6	İnaktif
P7	İnaktif
P8	Aktif
P9	Aktif
P10	İnaktif
P11	Aktif
R	Kayıp
R1	Aktif
R3	İnaktif
R4	Aktif
R5	İnaktif
R6	Aktif
R7	İnaktif
R8	İnaktif
R9	İnaktif
R10	Kayıp
R11	İnaktif
S1	Aktif
S2	Kayıp
S3	İnaktif
S4	İnaktif
S5	İnaktif
S6	İnaktif
S7	Kayıp
S8	İnaktif
S9	Kayıp
S10	İnaktif
S11	Aktif
T2	Kayıp
T3	İnaktif
T4	Aktif
T5	İnaktif
T6	İnaktif
T7	İnaktif

T8	İnaktif
T9	Aktif
T10	İnaktif
T11	İnaktif
V	Aktif
V2	İnaktif
V3	Aktif
V4	Aktif
V5	İnaktif
V6	İnaktif
V7	İnaktif
V8	İnaktif
V9	Aktif
V10	İnaktif
V11	Kayıp
Y	Aktif
Y1	Aktif
Y2	İnaktif
Y3	Aktif
Y4	İnaktif
Y5	Aktif
Y6	İnaktif
Y7	İnaktif
Y8	Aktif
Y9	Kayıp
Y10	İnaktif
Y11	İnaktif