

**T.C.  
YILDIZ TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**YAZILIM PROJELERİNDE RİSK FAKTÖRLERİ VE FARKLI KESTİRİM  
YAKLAŞIMLARI İLE RİSK DEĞERLENDİRME MODELLERİ ÇIKARIMI**

**AHMET UNUDULMAZ**

**YÜKSEK LİSANS TEZİ  
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI  
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN  
PROF. DR. OYA KALIPSIZ**

**İSTANBUL, 2013**

**T.C.**  
**YILDIZ TEKNİK ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**YAZILIM PROJELERİNDE RİSK FAKTÖRLERİ VE FARKLI KESTİRİM  
YAKLAŞIMLARI İLE RİSK DEĞERLENDİRME MODELLERİ ÇIKARIMI**

Ahmet UNUDULMAZ tarafından hazırlanan tez çalışması 20.06.2013 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

**Tez Danışmanı**

Prof. Dr. Oya Kalıpsız  
Yıldız Teknik Üniversitesi

**Jüri Üyeleri**

Prof Dr. Oya KALIPSIZ  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Doç Dr. Alev TAŞKIN GÜMÜŞ  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Yrd. Doç Dr. Gökhan BİLGİN  
Yıldız Teknik Üniversitesi

\_\_\_\_\_

Risk yönetimi alanında yapılmış bu tez çalışmasında risk faktörleri, sınıflandırıcı yöntemlerine ve model performanslarına ulaşılabilinmektedir. Yazılım mühendisliğinde risk yönetimine geniş bir şekilde yer verilmiş ve risk yönetimi adımları anlatılmıştır. Dergi ve makale çalışmalarıyla desteklenen tez çalışmasında veri setleri içindeki risk faktörlerinin etki değerleri, risk faktörlerinin birbirleriyle kıyaslanmaları ve risk faktörü eleme yöntemlerinden bahsedilmiş olup, bu faktörler yardımıyla veriler çok kritik, kritik, koşullu hata, orta seviyeli ve ihmal edilebilir olarak sınıflandırılmıştır. Bu sınıflandırma işleminde ise Destek Vektör Makineleri, Karar Ağaçları ve Naive Bayes sınıflandırıcılarından faydalanılmıştır.

Tez çalışması kapsamında bilgi birikimi ve anlayışıyla sürekli yanımda olan ve bana yol gösteren hocam Prof. Dr. Oya Kalıpsız'a, bu süreçte yanımda olan meslektaşım ve arkadaşım M. Özgür Cingiz'e, bu süreçte moral motivasyonumu sürekli üst düzeyde tutmamı sağlayan ve desteklerini bir an olsun bırakmayan aileme çok teşekkür ederim.

Mayıs, 2013

Ahmet UNUDULMAZ

## İÇİNDEKİLER

|                                                                                      | Sayfa |
|--------------------------------------------------------------------------------------|-------|
| KISALTMA LİSTESİ.....                                                                | vi    |
| ŞEKİL LİSTESİ.....                                                                   | vii   |
| ÇİZELGE LİSTESİ .....                                                                | viii  |
| ÖZET .....                                                                           | xi    |
| ABSTRACT.....                                                                        | xii   |
| BÖLÜM 1                                                                              |       |
| GİRİŞ.....                                                                           | 1     |
| 1.1    Literatür Özeti .....                                                         | 1     |
| 1.2    Tezin Amacı .....                                                             | 2     |
| 1.3    Tezin Yapısı.....                                                             | 2     |
| BÖLÜM 2                                                                              |       |
| YAZILIM PROJELERİNDE RİSK YÖNETİMİ .....                                             | 4     |
| 2.1    Risk Yönetim Paradigmaları .....                                              | 6     |
| 2.2    Risk Öznitelikleri.....                                                       | 13    |
| 2.3    Yazılım Risk Taksonomisi.....                                                 | 14    |
| BÖLÜM 3                                                                              |       |
| İLGİLİ ÇALIŞMALAR.....                                                               | 16    |
| 3.1    Risk Yönetimi ve Analizi ile İlgili Yapılan Çalışmalar.....                   | 17    |
| 3.2    Risk Faktörleri ve Risk Modellerinin Kullanılması ile İlgili Çalışmalar ..... | 19    |
| BÖLÜM 4                                                                              |       |
| UYGULAMA TASARIMI ve VERİ HAZIRLAMA.....                                             | 23    |
| 4.1    Veri Eldesi.....                                                              | 25    |
| 4.2    Veri Ön İşleme.....                                                           | 25    |
| 4.3    Özellik Seçimi ve Faktör Kararı.....                                          | 27    |
| 4.4    Kümeleme İşlemleri .....                                                      | 31    |

|                                                |                                                              |    |
|------------------------------------------------|--------------------------------------------------------------|----|
| 4.5                                            | Sınıflandırma İşlemleri .....                                | 32 |
| 4.5.1                                          | Destek Vektör Makineleri .....                               | 32 |
| 4.5.2                                          | Karar Ağaçları .....                                         | 33 |
| 4.5.3                                          | Naive Bayes Yöntemi .....                                    | 34 |
| BÖLÜM 5                                        |                                                              |    |
| BULGULAR .....                                 |                                                              | 36 |
| 5.1                                            | Destek Vektör Makinesi Sınıflandırma Yöntemi Sonuçları ..... | 37 |
| 5.2                                            | Karar Ağaçları Sınıflandırma Yöntemi Sonuçları .....         | 41 |
| 5.3                                            | Naive Bayes Sınıflandırma Yöntemi Sonuçları .....            | 45 |
| 5.4                                            | Test Sonuçlarının Karşılaştırılması .....                    | 48 |
| BÖLÜM 6                                        |                                                              |    |
| SONUÇ VE ÖNERİLER .....                        |                                                              | 51 |
| KAYNAKLAR .....                                |                                                              | 53 |
| EK-A .....                                     |                                                              | 55 |
| VERİ KÜMESİNDEN ÇIKARILAN GENEL SONUÇLAR ..... |                                                              | 55 |
| ÖZGEÇMİŞ .....                                 |                                                              | 79 |

## KISALTMA LİSTESİ

---

|       |                                        |
|-------|----------------------------------------|
| BK    | Bilgi Kazanımı                         |
| ÇKYSA | Çok Katmanlı Yapay Sinir Ağları        |
| DVM   | Destek Vektör Makinesi                 |
| KA    | Karar Ağaçları                         |
| kNN   | k-En Yakın Komşuluk                    |
| NB    | Naive Bayes                            |
| RHİY  | Risk Hafifletme İzleme ve Yönetme      |
| UYMS  | Ulusal Yazılım Mühendisliği Sempozyumu |
| UYMK  | Ulusal Yazılım Mimarisi Konferansı     |
| YSA   | Yapay Sinir Ağları                     |

## ŞEKİL LİSTESİ

---

|                                            | Sayfa |
|--------------------------------------------|-------|
| Şekil 1.1 Risk Yönetim Paradigmaları ..... | 6     |
| Şekil 4.1 Sistem Mimarisi .....            | 24    |
| Şekil 4.2 Destek Vektör Makineleri .....   | 33    |

## ÇİZELGE LİSTESİ

|                                                                                          | Sayfa |
|------------------------------------------------------------------------------------------|-------|
| Çizelge 2.1 Risk Yönetim Süreçleri .....                                                 | 7     |
| Çizelge 2.2 Risk Tipleri .....                                                           | 8     |
| Çizelge 2.3 Etki Analizi .....                                                           | 9     |
| Çizelge 2.4 Örnek Risk Çizelgesi .....                                                   | 10    |
| Çizelge 2.5 Olasılık/Etki Seviyesi.....                                                  | 14    |
| Çizelge 2.6 Risk Taksonomi Çizelgesi.....                                                | 14    |
| Çizelge 4.1 Arff Format Örneği .....                                                     | 26    |
| Çizelge 4.2 Projelerde Yer Alan 32 Risk Faktörü .....                                    | 27    |
| Çizelge 4.3 Hata Risk Faktörleri ve Değerleri .....                                      | 30    |
| Çizelge 4.4 Bilgi Kazanım Özellik Seçimine Göre Etki Değer Sıralaması ve Etki Değeri     | 31    |
| Çizelge 5.1 DVM Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....                        | 38    |
| Çizelge 5.2 DVM Tüm Projeler için Karışıklık Matrisi .....                               | 38    |
| Çizelge 5.3 DVM İlk Ürün Gurubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....   | 39    |
| Çizelge 5.4 DVM İkinci Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri ...   | 39    |
| Çizelge 5.5 DVM Birinci Ürün Grubu Karışıklık Matrisi.....                               | 39    |
| Çizelge 5.6 DVM İkinci Ürün Grubu Karışıklık Matrisi .....                               | 40    |
| Çizelge 5.7 DVM Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....            | 40    |
| Çizelge 5.8 KA Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....                         | 42    |
| Çizelge 5.9 KA Tüm Projeler için Karışıklık Matrisi .....                                | 42    |
| Çizelge 5.10 KA İlk Ürün Gurubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....   | 43    |
| Çizelge 5.11 KA İkinci Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri.....  | 43    |
| Çizelge 5.12 KA Birinci Ürün Grubu Karışıklık Matrisi .....                              | 44    |
| Çizelge 5.13 KA İkinci Ürün Grubu Karışıklık Matrisi .....                               | 44    |
| Çizelge 5.14 KA Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....            | 45    |
| Çizelge 5.15 NB Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....                        | 46    |
| Çizelge 5.16 NB Tüm Projeler için Karışıklık Matrisi.....                                | 46    |
| Çizelge 5.17 NB İlk Ürün Gurubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri .....   | 46    |
| Çizelge 5.18 NB İkinci Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri ..... | 47    |
| Çizelge 5.19 NB Birinci Ürün Grubu Karışıklık Matrisi .....                              | 47    |
| Çizelge 5.20 NB İkinci Ürün Grubu Karışıklık Matrisi .....                               | 48    |
| Çizelge 5.21 NB Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri.....             | 48    |
| Çizelge 5.22 Çalışmanın Başarı Oranları ve Yüzdeleri .....                               | 49    |
| Çizelge A1.1 Karar Ağacı Tüm Projeler Özet Tablosu .....                                 | 57    |
| Çizelge A1.2 Karar Ağacı Tüm Projeler Karışıklık Matrisi .....                           | 57    |
| Çizelge A1.3 Karar Ağacı 1.Ürün Grubu Özet Tablosu .....                                 | 57    |
| Çizelge A1.4 Karar Ağacı 1.Ürün Grubu Karışıklık Matrisi .....                           | 58    |

|               |                                                    |    |
|---------------|----------------------------------------------------|----|
| Çizelge A1.5  | Karar Ağacı 2. Ürün Grubu Özet Tablosu.....        | 58 |
| Çizelge A1.6  | Karar Ağacı 2. Ürün Grubu Karışıklık Matrisi ..... | 58 |
| Çizelge A1.7  | Karar Ağacı 1. Proje Özet Tablosu.....             | 59 |
| Çizelge A1.8  | Karar Ağacı 1. Proje Karışıklık Matrisi .....      | 59 |
| Çizelge A1.9  | Karar Ağacı 2. Proje Özet Tablosu.....             | 59 |
| Çizelge A1.10 | Karar Ağacı 2. Proje Karışıklık Matrisi .....      | 60 |
| Çizelge A1.11 | Karar Ağacı 3. Proje Özet Tablosu .....            | 60 |
| Çizelge A1.12 | Karar Ağacı 3. Proje Karışıklık Matrisi.....       | 60 |
| Çizelge A1.13 | Karar Ağacı 4. Proje Özet Tablosu .....            | 61 |
| Çizelge A1.14 | Karar Ağacı 4. Proje Karışıklık Matrisi.....       | 61 |
| Çizelge A1.15 | Karar Ağacı 5. Proje Özet Tablosu .....            | 61 |
| Çizelge A1.16 | Karar Ağacı 5. Proje Karışıklık Matrisi.....       | 62 |
| Çizelge A1.17 | Karar Ağacı 6. Proje Özet Tablosu .....            | 62 |
| Çizelge A1.18 | Karar Ağacı 6. Proje Karışıklık Matrisi.....       | 62 |
| Çizelge A1.19 | Karar Ağacı 7. Proje Özet Tablosu .....            | 63 |
| Çizelge A1.20 | Karar Ağacı 7. Proje Karışıklık Matrisi.....       | 63 |
| Çizelge A1.21 | Karar Ağacı 8. Proje Özet Tablosu .....            | 63 |
| Çizelge A1.22 | Karar Ağacı 8. Proje Karışıklık Matrisi.....       | 64 |
| Çizelge A2.1  | Naive Bayes Tüm Projeler Özet Tablosu .....        | 64 |
| Çizelge A2.2  | Naive Bayes Tüm Projeler Karışıklık Matrisi .....  | 64 |
| Çizelge A2.3  | Naive Bayes 1.Ürün Grubu Özet Tablosu .....        | 65 |
| Çizelge A2.4  | Naive Bayes 1.Ürün Grubu Karışıklık Matrisi.....   | 65 |
| Çizelge A2.5  | Naive Bayes 2. Ürün Grubu Özet Tablosu .....       | 65 |
| Çizelge A2.6  | Naive Bayes 2. Ürün Grubu Karışıklık Matrisi.....  | 66 |
| Çizelge A2.7  | Naive Bayes 1. Proje Özet Tablosu .....            | 66 |
| Çizelge A2.8  | Naive Bayes 1. Proje Karışıklık Matrisi.....       | 66 |
| Çizelge A2.9  | Naive Bayes 2. Proje Özet Tablosu .....            | 67 |
| Çizelge A2.10 | Naive Bayes 2. Proje Karışıklık Matrisi.....       | 67 |
| Çizelge A2.11 | Naive Bayes 3. Proje Özet Tablosu .....            | 67 |
| Çizelge A2.12 | Naive Bayes 3. Proje Karışıklık Matrisi.....       | 68 |
| Çizelge A2.13 | Naive Bayes 4. Proje Özet Tablosu .....            | 68 |
| Çizelge A2.14 | Naive Bayes 4. Proje Karışıklık Matrisi.....       | 68 |
| Çizelge A2.15 | Naive Bayes 5. Proje Özet Tablosu .....            | 69 |
| Çizelge A2.16 | Naive Bayes 5. Proje Karışıklık Matrisi.....       | 69 |
| Çizelge A2.17 | Naive Bayes 6. Proje Özet Tablosu .....            | 69 |
| Çizelge A2.18 | Naive Bayes 6. Proje Karışıklık Matrisi.....       | 70 |
| Çizelge A2.19 | Naive Bayes 7. Proje Özet Tablosu .....            | 70 |
| Çizelge A2.20 | Naive Bayes 7. Proje Karışıklık Matrisi.....       | 70 |
| Çizelge A2.21 | Naive Bayes 8. Proje Özet Tablosu .....            | 71 |
| Çizelge A2.22 | Naive Bayes 8. Proje Karışıklık Matrisi .....      | 71 |
| Çizelge A3.1  | DVM Tüm Projeler Özet Tablosu .....                | 71 |
| Çizelge A3.2  | DVM Tüm Projeler Karışıklık Matrisi.....           | 72 |
| Çizelge A3.3  | DVM 1.Ürün Grubu Özet Tablosu .....                | 72 |
| Çizelge A3.4  | DVM 1.Ürün Grubu Karışıklık Matrisi .....          | 72 |
| Çizelge A3.5  | DVM 2. Ürün Grubu Özet Tablosu .....               | 73 |
| Çizelge A3.6  | DVM 2. Ürün Grubu Karışıklık Matrisi .....         | 73 |
| Çizelge A3.7  | DVM 1. Proje Özet Tablosu.....                     | 73 |

|               |                                       |    |
|---------------|---------------------------------------|----|
| Çizelge A3.8  | DVM 1. Proje Karışıklık Matrisi ..... | 74 |
| Çizelge A3.9  | DVM 2. Proje Özet Tablosu .....       | 74 |
| Çizelge A3.10 | DVM 2. Proje Karışıklık Matrisi ..... | 74 |
| Çizelge A3.11 | DVM 3. Proje Özet Tablosu .....       | 75 |
| Çizelge A3.12 | DVM 3. Proje Karışıklık Matrisi ..... | 75 |
| Çizelge A3.13 | DVM 4. Proje Özet Tablosu .....       | 75 |
| Çizelge A3.14 | DVM 4. Proje Karışıklık Matrisi ..... | 76 |
| Çizelge A3.15 | DVM 5. Proje Özet Tablosu .....       | 76 |
| Çizelge A3.16 | DVM 5. Proje Karışıklık Matrisi ..... | 76 |
| Çizelge A3.17 | DVM 6. Proje Özet Tablosu .....       | 77 |
| Çizelge A3.18 | DVM 6. Proje Karışıklık Matrisi ..... | 77 |
| Çizelge A3.19 | DVM 7. Proje Özet Tablosu .....       | 77 |
| Çizelge A3.20 | DVM 7. Proje Karışıklık Matrisi ..... | 78 |
| Çizelge A3.21 | DVM 8. Proje Özet Tablosu .....       | 78 |
| Çizelge A3.22 | DVM 8. Proje Karışıklık Matrisi ..... | 78 |

**YAZILIM PROJELERİNDE RİSK FAKTÖRLERİ VE FARKLI KESTİRİM  
YAKLAŞIMLARI İLE RİSK DEĞERLENDİRME MODELLERİ ÇIKARIMI**

Ahmet UNUDULMAZ

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Prof. Dr. Oya KALIPSIZ

Risk yönetiminin önemi son yıllarda daha çok öne çıkmıştır. Başarısız giden projeler ve süreçlere bakıldığında risk yönetiminin iyi bir şekilde ele alınmadığı gözlemlenmektedir. Bu nedenle risk yönetimi, risk analizi ve risk önleme alanları şirketlerde önemli bir yere gelmiştir.

Bu çalışmada da hata veri setleri üzerinde faktör kararı verilerek gereksiz risk faktörlerinin projelerden çıkarılması sağlanmıştır. Faktörlerde kendi içinde sıralanarak hangi alt alanın en çok risk taşıdığı ortaya konulmuştur. Faktör kararından sonra bilgisayar mühendisliği kestirim yaklaşımları kullanılarak sınıflandırmalar yapılmış ve modeller oluşturulmuştur. Bu modellerin birbirleriyle karşılaştırılmasına yer verilmiştir.

**Anahtar Kelimeler:** Yazılım Mühendisliğinde Risk Yönetimi, Risk Faktörleri, Risk Modelleri, Risk Tahmini

**RISK FACTORS AND RISK ASSESMENT MODELS USING DIFFERENT  
APPROACH IN SOFTWARE PROJECT**

Ahmet UNUDULMAZ

Department of ComputerEngineering

MSc. Thesis

Adviser: Prof. Dr. Oya KALIPSIZ

Risk management isn't used properly when investigated unsuccessful project and unsuccessful process. Therefore risk management, risk analysis and risk avoiding areas are more important than other phases in companies.

In this work bug data sets are examined. Factor decision is made in those data set and eliminate some of them. All risk factors are arranged in itself. In this manner more critical factors are identified. After that computer engineering approach are used for modelling and classification. These models are compared in itself.

**Keywords :** Risk Management in Software Engineering, Risk Factors, Risk Models, Risk Assesment

### GİRİŞ

Yazılım mühendisliğinde risk yönetiminin önemi son yıllarda daha çok ortaya çıkmıştır. Başarısız giden birçok proje, hatalı yönetilen süreçler, ürünlerin ve projelerin zamanında teslim edilememesi, maliyetlerde oluşan aşırı artışlar risk yönetiminin önemini günden güne arttırmaktadır. Özellikle risklerin düzgün bir şekilde ele alınamaması, projeler arası risk türlerinin ve gruplarının farklılaşması, şirket/kurum içi yapılan risk yönetimi ve risk analizi çalışmalarının ölçülememesi de bu durumu karmaşıktırmaktadır. Tez çalışması ise yukarıda belirtilen problemler üzerine yapılmıştır. Risklerin her projede standartlaştırılması, risk faktörlerinin düzgün bir şekilde analiz edilmesi ve uygulanan risk önleme ve risk analiz çalışmalarının başarı oranlarının ölçülebilmesi hedeflenmiştir.

#### 1.1 Literatür Özeti

Yazılım mühendisliğinde risk yönetimi ile ilgili çalışmalara bakıldığında ise son yıllarda artış gözlemlenmektedir. Bu alanda yapılan çalışmalar geçmiş yıllara göre çeşitlilik göstermekte ve önem kazanmaktadır. Genel risk yönetimi, risk faktörleri, risk önlemek için modellerin kullanılması ile ilgili çalışmaların yayınlandığı, bu konular üzerinde daha detaylı durulduğu gözlemlenmiştir. Bu kısımda yazılım mühendisliğindeki risk yönetimi konusunda yapılan çalışmalar iki alt başlık altında toplanmıştır. İlk başlıkta risk yönetimi ve analizi ile ilgili dergi ve makale çalışmaları incelenmiş ve derlenmiştir. İkinci başlıkta ise risk faktörlerine yer verilmiş, faktörlerin birbirleriyle kıyaslanması, eleminize edilmesi ve sıralanması ile ilgili yapılan çalışmalardan bahsedilmiştir. Ayrıca yazılım

mühendisliği yaklaşımları kullanılarak oluşturulan modeller ortaya konulmuş ve modellerinde birbirleriyle kıyaslanmasına yer verilmiştir.

## **1.2 Tezin Amacı**

Yazılım risk yönetimi alanında yapılan bu çalışmada temel olarak risk faktörleri ve risk değerlendirme modellerinin birbirleriyle karşılaştırılmasından bahsedilecektir. İncelenen çalışmalar göz önünde bulundurulduğunda bu iki alanda şirketlerde yaşanan problemler ve zaman kayıpları dikkate alınmıştır. Risk faktörlerinin analiz edilmemesi veya hatalı analiz edilerek gerkesiz bir faktör için önlem alınması zaman kaybına ve maddi kayıplara yol açmaktadır. Bu nedenle faktörlerin doğru seçilmesi, kritik faktörlerin belirlenmesi ve etki değeri düşük faktörlerin elenmesi gerekmektedir. Tez çalışması kapsamında risk faktörlerinin detaylı analizi amaçlanmıştır.

Diğer bir amaç ise yazılım mühendisliği yaklaşımlarını kullanarak indirgenmiş risk faktörleri ve veri setlerini doğru sınıflandırabilecek modeller oluşturmaktır. Kullanılan yöntemlerin başarı oranları ve başarı yüzdeleride çalışma süresince incelenecektir.

## **1.3 Tezin Yapısı**

Tez çalışması kapsamında ilk olarak yazılım projelerindeki risk yönetimi konusuna değinilecektir. Risk ve risk yönetimi tanımlamaları ve risk yönetimi yapmanın önemi vurgulanıp şirketlerin kayıpları ile ilgili örneklerde gösterilecektir. Risk yönetim paradigmalarından da detaylıca bahsedilecektir. Risk tanımlama, risk analizi, risk planlama, risk izleme, risk kontrol etme adımlarından bahsedilip örnek tablolar ve iş listeleriyle detaylandırılacaktır. Ayrıca örnek risk tablolarına da yer verilecektir. Risk öznitelikleri ve risk taksonomisinden de bahsedilecektir. Gereksinim, tasarım, kod ve birim testler ile iletişim konularından da bahsedilip yazılım risk yönetimi tanım kısmı tamamlanacaktır.

Risk yönetimi konusundan sonra ilgili çalışmalar konusu gelecektir. Bu kısımda dergilerde çıkmış veya konferanslarda yayınlanmış olan makale ve bildiriler yer alacaktır. İlgili çalışmalar bölümü de kendi içinde ikiye ayrılmaktadır. İlk kısımda genel risk yönetimi ve günümüzde risk yönetimi ile ilgili yapılan çalışmalar anlatılacak olup

ikinci kısımda da yazılım mühendisliği yaklaşımları kullanarak oluşturulan modellerin risk veri setlerine uygulanması ile ilgili makaleler anlatılacaktır. Destek vektör makineleri, NaiveBayes, yapay sinir ağları, k en yakın komşuluk ve bulanık mantık yaklaşımları ile oluşturulan çalışmalardan bahsedilecek ve bu sınıflandırıcıların başarı oranları detaylandırılacaktır.

İlgili çalışmalar anlatıldıktan sonra uygulama tasarımı ve veri hazırlama konularına değinilecektir. Sistem mimarisinin nasıl olacağı, çalışmanın hangi evrelerden geçtiği bu bölümde anlatılacaktır. Veri eldesi ve veri seti üzerinde detaylara yer verilecek olup, veri ön işlemede hangi aşamalardan geçtiğinden bahsedilecektir. Bu aşamadan sonra tez çalışması için önemli sonuçlardan biri olan faktör kararına yer verilecektir. Veri setlerindeki faktörler incelenecek ve sonuçlara ne derecede etki ettiklerinin bilgisi bu kısımda yer alacaktır. Faktör kararı verildikten sonra kümeleme ve sınıflandırma işlemlerinden bahsedilecektir. Kullanılan sınıflandırıcılar (Destek Vektör Makineleri, Karar Ağaçları ve NaiveBayes) hakkında kısa bir bilgi verilip bulgular ve sonuçlar kısmı anlatılacaktır.

Bulgular ve sonuçlar kısmında kullanılacak proje sayısı ve her bir projenin kaç veriden oluştuğu bilgisine erişilebilecektir. Yapılan çalışmada kullanılan 8 projenin ortak alındığı durum, aynı özelliklere sahip ürün gruplarının ortak olarak ele alındığı durumlar (4 proje içerir) ve her bir projenin kendi içinde ele alındığı durumlar detaylandırılarak anlatılacaktır. Destek vektör makineleri, karar ağaçları ve NaiveBayes yöntemleri ile aynı veriler sınıflandırılacak ve elde edilen modeller ve sonuçlar birbirleriyle kıyaslanacaktır. Kıyaslama aşamasında ilgili çalışmalarda göz önünde bulundurularak genel bir yoruma varılacak ve sınıflandırma yöntemleri hakkında çıkarımlarda bulunulacaktır.

Ekler kısmında ise kullanılan üç yöntem için detaylı tablolara yer verilecektir. Her bir yöntem için 8 projenin ortak olarak ele alınma durumu, aynı ürün gruplarının ortak olarak ele alınma durumu ve 8 projenin ayrı ayrı olarak ele alındığı durumlar için özet tabloları, sınıf bazında doğruluk tabloları ve karışıklık matrislerinin yer aldığı ve hataların sınıflandırıldığı tablolara yer verilerek çalışma tamamlanacaktır.

## BÖLÜM 2

---

### YAZILIM PROJELERİNDE RİSK YÖNETİMİ

Risk kaybın olasılıksal gösterilmiş halidir. Bir bakıma kumardır da denebilir. Risk alındığı zaman olasılıklara bağlı olarak kazanma ve kaybetme şansı da göze alınmış olmaktadır. Yazılım projelerindeki riskte bu olasılıklara bağlıdır. Olasılıkların yanında riski tanımlayacak başka bir faktörde kaybın önemidir. Bu öneme göre risk değişkenlik gösterecektir. Bu olasılıklar kontrol altındaki faktörlerden kaynaklanırsa bu risk olarak değerlendirilemez. Basitleştirilmiş şekilde faktörün risk olarak adlandırılabilmesi için kontrol altındaki faktörlerden bağımsız olmalıdır[1].

Riskin basitleştirilmiş tanımı ise projenin başarısızlığına neden olacak kayıplar olarak söylenebilir. Bu kayıplar henüz yaşanmamıştır. Bu potansiyel problemler maliyete, zaman planına, projenin teknik başarısına, yazılım ürün kalitesine veya proje takımının moraline etki edebilir. Risk yönetimi tanımlayıcı, adres gösterici ve bu potansiyel problemleri eleyici bir süreçtir. Yani risk normal olmayan olayların veya başarısızlıkların ve bu olayın sonucu olarak sistemin, kullanıcıların ve ortamın etkilenmesi durumudur. Riskler çok kritikten(tüm sistemin kaybı), etki değeri olmayana(sistem zararı olmayan) doğru derecelendirilebilir[2].

Riskler muhtemel kayıp olasılıkları olarakda nitelendirilmektedir. Buradaki kayıp projeye eksi yönde etki eden faktörlerdir(Son ürün kalitesini etkileyecek, maliyetleri arttıracak, zamanda beklemeye neden olacak problemler). Riskler olma ve olama ihtimalleri olan potansiyel problemlerdir. Risklerin belirlenmesi, etki değerlerinin ortaya konulması, olasılıkların çıkarılması ve buna uygun planların yapılması iyi bir

fikirdir. Ayrıca risk yönetiminin yapılması kesin olmayan verilerin yazılım ekiplerince daha iyi anlaşılmasına da olanak vermektedir[3][4].

Proje yönetimi açısından bakıldığında da yazılım risk yönetimi risk faktörlerinin sistematik bir şekilde tanımlandığı, değerlendirildiği ve hafifletildiği biçimsel bir süreçtir. İç ve dış sebepler göz önünde bulundurulduğunda projedeki risklere karar verilmesi temel proje yönetim sürecini oluşturmaktadır. Projede tüm gereksinimler aynı değerleri almadan sıralanır ve tüm çıktılar olasılıklar halinde gösterilir[5]. PMBOK'da riskler açık bir şekilde listelenmiştir. PMBOK'a göre risk yönetiminin etkilediği alanlar şu şekildedir. Kapsam, İletişim, PM entegrasyonu, Kalite, İnsan Kaynakları, Zaman, Maliyet ve Kontratlar(Servis, Materyal performansı) olarak listelenmiştir[6].

Yazılım geliştirme projelerinde riskler kavram keşfinden başlayarak yazılım projelerinin kabulüne kadar devam etmektedir. Bu nedenle iyi bir proje yöneticisi aynı zamanda iyi bir risk yöneticisi de olmak zorundadır[7].

Standish Grup 8380 yazılım projesini baz alarak anket yapmıştır. Bu ankete göre aşağıdaki sonuçlar çıkmıştır :

- %53 proje tamamlanmıştır. Fakat maliyet ve zamanlama açısından oluşan sıkıntılar yüzünden beklenen özelliklerin hepsini sağlayamamıştır.
- %31 proje iptal edilmiştir. Yazılım geliştirme sürecinde projelerin %31'i iptal edilmiştir.
- İptal edilen ve tamamlanan projelerin toplamında ortalama maliyet başta tahmin edilen maliyetin %190 fazlasına karşılık gelmiştir.
- Tamamlanmış ve iptal edilmiş projelerin başlangıçta tahmin edilen zamandan %222 fazla zaman harcadığı gözlemlenmiştir
- Ortalama olarak %31 proje istenilen özellikte ve tüm fonksiyonlarını yerine getirecek şekilde tamamlanmıştır.

1998 senesinde müşteri kredileriyle ilgili 22 saatlik bir yazılım problemi nedeniyle eBay 4 milyon dolar zarara girmiştir. Yatırımcı güvenini geri kazanabilmek için ise 5.7 milyar dolarlık bir maliyet harcamışlardır. 1998 senesinde Hershey şirketi %12 zarar ederek 150 milyon dolarlık bir gelirden olmuştur. Üst düzey yazılımlarından birinin aksaklığı

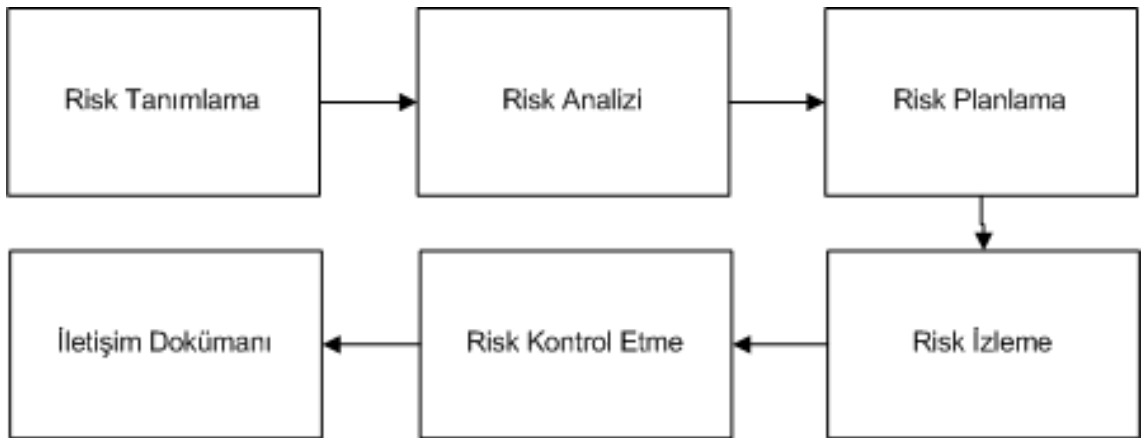
sonucu kayıtlarda cadılar bayramı için şeker getirdiği bilgisi yazmıştır. Bunun sonucu olarak şeker üreticilerin %19 net karı düşmüştür[1].

Yukarıda da görüleceği üzere yazılım hataları pahalı ve organizasyon yapısını ciddi ölçüde zarara uğratabilecek şekilde olabilmektedir. Önlem almak her zaman hatayı düzeltmekten daha iyi sonuçlar vermektedir. Bu noktada da risk yönetimi ortaya çıkmaktadır. Risk yönetimi aktiviteleri yazılım projelerini daha az karmaşık yapmaktadır[1]. Ayrıca :

- Risklerin tanımlaması ve önceliklendirilmesi proje yöneticilerine ve proje personeline büyük faydalar sağlayacak ve projeye daha iyi odaklanmalarına yardımcı olacaktır.
- Risk hafifletme toplam proje riskini azaltacağı gibi proje tamamlanma hızını da arttıracaktır.
- Yazılım risk yönetimi kullanan projelerin zaman planları daha kolay kestirilebilecek, daha az sürprizle karşılaşılacaktır

## 2.1 Risk Yönetim Paradigmaları

Risk paradigmasında tüm risk süreçleri analiz edilir ve tartışılır. Şekil 2.1'den de görüleceği üzere 6 süreç içermektedir. Bunlar tanımlama, analiz etme, planlama, izleme, kontrol etme ve iletişim dokümanıdır.



Şekil 2.1 Risk Yönetim Paradigmaları

Risk yönetimi şekil 2.1 'de ki adımlarla incelenmektedir. Risk yönetimi kendi içinde de alt dallara ayrılmaktadır. Risk değerlendirme ve risk kontrolü şeklinde belirtilen süreçleri ikiye ayırmak mümkündür[8]. Çizelge 2.1'den de bu ayrım detaylı bir şekilde incelenebilmektedir.

Çizelge 2.1 Risk Yönetim Süreçleri

| Risk Değerlendirme                |                                                            | Risk Kontrol                                      |                                 |                                                         |
|-----------------------------------|------------------------------------------------------------|---------------------------------------------------|---------------------------------|---------------------------------------------------------|
| Tanımlama                         | Analiz                                                     | Planlama                                          | Kontrol                         | İzleme                                                  |
| Listeleme<br>Karar Sürücü Analizi | Performans Modelleri<br>Maliyet Modelleri<br>Karar Analizi | Riskten kaçınma<br>Risk transferi<br>Risk azaltma | Prototip<br>Analiz<br>Benchmark | Kilometre taşı takibi<br>İzleme<br>Düzeltilici Eylemler |

Risk yönetiminde ilk adım risk tanımlamadır. Risk tanımlamanın amacı ise riskler daha problem haline gelmeden önce bu bilginin yazılım proje yönetim sürecine dahil edilmesidir. Proje içindeki herhangi biri riskleri projeye tanımlayabilir. Her bir birey projenin belirli kısımlarında bilgi sahibidir. Tanımlama aşamasında belirli olmayan problemler ve riskler somut hale getirilir, tanımlanır ve ölçülebilir. Aynı zamanda proje planına etki edecek tüm risklerin tanımlanması anlamında gelmektedir. Risk tanımlama için bir çok yöntem bulunmaktadır. Olası risklerin listelenmesi, anketler, toplantılar, beyin fırtınaları, planların gözden geçirilmesi yöntemleri sayılabilir[9][10].

Risk tanımlamada diğerlerinden öne çıkan yöntem risk elemanlarının listelenmesi şeklindedir. SEI de bununla ilgili bir liste yayınlamıştır[11]. Bu liste risk tanımlamasında kullanılacağı gibi önceden bilinen veya kestirilebilen risk kategorileri de belirlenebilmektedir. Çizelge 2.2'ye bakıldığında bu listenin özetine erişilebilir.

Risk gruplarına çizelge 2.2'den bakıldığında şu alt alanlara da eklemeler yapabilir. Personel boyutu ve deneyim risklerine ek olarak takımların değiştirilmesini ve personel kayıpları gösterilebilir. Başarılı olan bir projedeki elemanların başka bir ekibe aktarılması sonucunda yeni projede görülecek risklerde göz ardı edilmemelidir. İş etki risklerine ek olarak ise belirli bir projenin veya görevin arttırılması veya takım elemanlarına başka işlerin verilmesidir[12][13].

Çizelge 2.2 Risk Tipleri [1][3]

| Risk Tipleri                        | Tanım                                                                                                                                                                 | Örnek Riskler                                                                                                                                                                                    |
|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ürün boyutu riskleri                | Proje riskleri ile doğrudan orantılıdır. Geçmiş tecrübelerden çıkarılan bilgiler kullanılmaktadır                                                                     | LOC, FP<br>Program sayısı<br>Kullanıcı sayısı<br>Veritabanı boyutu                                                                                                                               |
| İş Etki Riskleri                    | Genel risklerle ilintilidir. Proje bitiş zamanları ile ilgilidir.                                                                                                     | Şirket gelir etkisi<br>Makul teslim tarihi<br>Gecikmenin ücrete etkisi<br>Ürün hatalarına etkisi                                                                                                 |
| Müşteri Bazlı Riskler               | Müşterilerin farklı karakterlere sahip olmalarından ve isteklerinin farklı olmasından kaynaklı risk grubudur. Müşterinin süreci bilmemesi de bu riski arttırmaktadır. | Müşteri ile daha önce çalışıldı mı?<br>Müşteri ne istediğini biliyor mu?<br>Proje kapsamını belirleyebilmek için müşterinin zamanı var mı?<br>Müşteri yazılım süreçlerinden anlıyor mu?          |
| Ürün Riskleri                       | Ürün geliştirme süreçlerindeki etmenlerin hatalı tanımlanmasından kaynaklı ortaya çıkan risk grubudur. Kalitenin göz ardı edilmesinden kaynaklanmaktadır.             | Yazılım geliştirme sürecinin bir standardı var mı?<br>Yazılım süreçleri başka projelere de uygulanıyor mu?<br>Dokümanlardan elde edilen geri bildirimler yeni projede kullanılıyor mu?           |
| Teknoloji Riskleri                  | Teknolojik gelişmelere ayak uydurmak ve teknolojinin sürekli kendini yenilemesi sonucunda ortaya çıkan risk grubudur.                                                 | Kullanılacak teknoloji organizasyonda yeni mi?<br>Yeni algoritma tasarında müşteri gereksinimleri entegre edildi mi?<br>Gereksinimler yeni analiz dizayn ve test methodları gerektiriyor mu?     |
| Geliştirme Çevre Riskleri           | Yazılım ürününün geliştirildiği ortam riskleri bu grupta yer almaktadır.                                                                                              | Yazılım proje yönetim araçları mevcut mu?<br>Yazılım süreç yönetim araçları mevcut mu?<br>Kodlama ve test araçları ürün için uygun mu?                                                           |
| Personel Boyutu ve Deneyim Riskleri | Proje personelinin boyutu ve deneyimleri ile ilgili riskler bu grupta yer almaktadır                                                                                  | Proje için en iyi insanlar seçildi mi?<br>Yeteri sayıda insan var mı?<br>Personel bu iş için yeterli eğitimi adlı mı                                                                             |
| Risk Bileşenleri ve Sürücüleri      | Performans, maliyet, destek ve zamanlama riskleri risk bileşenleri olarak adlandırılır ve bu grupta yer alırlar                                                       | Performans (Gereksinimleri ne ölçüde karşılayacağı ile ilgilenir)<br>Maliyet (Proje bütçesi ile ilgilenir)<br>Destek (Yazılımın kolay adaptasyonu)<br>Zamanlama (Ürün zamanında üretililecek mi) |

Risk analizi yapmanın temel amacı risk karar tahmin bilgisinin çıkarılabilmesini sağlamaktır. Bu analizler proje yöneticilerine birçok riskin ortaya çıkmadan gözlemlenmesine olanak sağlamaktadır. Analiz aşamasında risk alanlarına göz atılır ve olası riskler gözden geçirilir. Her bir risk için risk indeksi yaratılır. Olası risk sonuçları ortaya konulur, kritik değişkenlerin tanımı yapılır, olası sonuçlarında tanımları bu aşamada yapılmaktadır[14].(Çizelge 2.3)

Çizelge 2.3 Etki Analizi

| Kategori\ Bileşen | Performans                                                                                              | Maliyet/Zaman                                                                                                   |
|-------------------|---------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| Felaket           | Gereksinimdeki başarısızlık sonucunda başarısız olarak etkiler                                          | Başarısızlık, maliyeti ve zamanda gecikmeyi artırır                                                             |
|                   | Perfonmansı direkt olarak etkiler ve yazılımların tepkisiz kalmasına yol açar                           | Finansal kesintilerle birlikte zamanda telafi edilemeyecek süreçler doğmaktadır                                 |
| Kritik            | Gereksinimlerdeki başarısızlık sistem performansını düşürmektedir.                                      | Maliyet ve zamanda gecikme artar fakat felaket kategorisine göre daha az zarar bulunur. Kaliteden ödün verilir. |
| Marjinal          | Gereksinimlerdeki başarısızlık daha az etki eder.                                                       | Maliyet ve zamandaki sarkmalara bağlı olarak kalite tolere edilebilir.                                          |
| İhmal Edilebilir  | Gereksinimlerdeki başarısızlık ihmal edilebilir. Etkisi yoktur. Performansda düşüş söz konusu değildir. | Zaman ve maliyet açısından en az kayıp bu gruptadır. Kalitede sıkıntı yoktur                                    |

Planlama aşamasında ise hangi risklerin olacağının kararını vermek için kullanılan bir fonksiyondur. Bu fonksiyonda temel proje risk bilgilerinden yola çıkılarak karar verme ve hafifletme stratejileri geliştirilir. Ürün tasarımı veya geliştirme süreçlerindeki değişikliklerden oluşacak risklerden uzak durulur. Riskler kabul edilir ve ona göre stratejiler ve önlem planları uygulanır. Risk planlamasındaki anahtar alan gelecek sonuçların tahminin şimdiden yapılabilmesidir[1].

Çizelge 2.4 Örnek Risk Çizelgesi[3][16]

| Riskler                                                  | Kategori               | Olasılık (%) | Etki     |
|----------------------------------------------------------|------------------------|--------------|----------|
| Proje müşterinin gereksinimlerinin iyi analiz edilmemesi | Müşteri Bazlı          | 90           | Kritik   |
| Kötü tanımlanmış yazılım gereksinimleri                  | İş Etki                | 70           | Marjinal |
| Daha az yeniden kullanılabilirlik                        | Ürün Boyutu            | 70           | Kritik   |
| Son kullanıcı engeli                                     | İş Etki                | 40           | Marjinal |
| Zaman aşımı olabilir                                     | İş Etki                | 50           | Kritik   |
| Finansman azalabilir                                     | Müşteri Bazlı          | 40           | Felaket  |
| Müşteri gereksinimlerini değiştirebilir                  | Ürün Boyutu            | 80           | Kritik   |
| Teknoloji istenileni karşılamayabilir                    | Teknoloji Bazlı        | 30           | Felaket  |
| Kullanılan araçlardaki problemler                        | Geliştirme Çevre Bazlı | 80           | Marjinal |
| Personelin deneyimsiz olması                             | Personel Bazlı         | 30           | Kritik   |
| Personelin bilgisiz olması                               | Personel Bazlı         | 70           | Marjinal |

İzleme adımında bu aşamada kazanılmış, derlenmiş ve raporlanmış risk durum bilgisi bulunmaktadır. İzleme aşamasında doğruluk, zaman ve ilişkili risk bilgisi tüm gruplara ve insanlara en şeffaf halde aktarılır. İzleme durum bilgisi kontrol fonksiyonu için kritik bir aşamadır. Bu aşamanın statik olmaması, periyodik olarak kendini yenilemesi de gerekmektedir[9].

İzleme aşamasında risk tablolarından faydalanılır. Örnek risk tablosuna çizelge 2.4' den ulaşılabilir. Tabloya bakıldığında riskler belirli kategorilere göre sınıflandırılmış, olasılık değerleri ve etki değerleri hazırlanmıştır. Bu sonuçlara göre ise RHİY (Risk hafifletme, izleme ve yönetme) planı belirlenmelidir[3].

Kontrol adımında ise amaç zamanında alınan ve etkili kararların hafifletme planında kullanımını sağlamaktır. İzleme durum bilgisini kullanır. Bu bilgiden yola çıkarak

raporlanacak verileri belirler. Kontrol fonksiyonu analiz durum raporlarını, nasıl işletileceğini ve bu kararların ne şekilde entegre edileceğini de belirtmektedir. RHİY planına göre risklerden kaçınılır, riskler izlenir ve kontrol altına alınarak yönetilir[1][3].

İletişim adımında riskler azaltılabilmektedir. Bu yüzden bu süreç modelde ortada yer almaktadır. Doğru bir analiz ve iyi yönetim için yüksek seviyede iletişim olmalıdır.

Risk yönetim paradigmalarına aynı zamanda belirli kurumlar tarafından çıkartıldığını ve standartlaştığı gözlemlenmektedir. Proje Yönetimi Enstitüsü(PMI) ,NewtownSquare ve PA risk yönetimi için eğitici bir rehber çıkarmışlardır. Tüm projelere uygulanabilmektedir. PMI'a göre işlem adımlar şu şekildedir:

- Risk yönetim planlaması
- Risk tanımı
- Nitel risk analizi
- Nicel risk analizi
- Risk tepki planlaması
- Risk izleme ve takip

Her bir işlem adımı için PMI girdiler, araçlar, teknikler ve çıktılar üretmektedir[15].

Risk Yönetim Enstitüsü (IRM), Risk Yöneticileri ve Garanti Birliği(AIRMIC) ve Genel Sektördeki Risk Yönetim Ulusal Forumu(ALARM) risk yönetimi üzerinde genel bir standart geliştirmişlerdir. Bu standart aşağıdaki temel adımları kapsamaktadır:

- Risk Tanımı
- Risk Yönetimi
- Risk Değerlendirme
- Risk Analizi
- Risk Evrimi
- Risk raporlama ve iletişim
- Risk davranışı
- İzleme ve risk yönetim işlemi

Süre gelen risk yönetimi modeli de geliştirilmiştir. Bu paradigma evrensel risk yönetim elementlerini içermektedir[1] :

- Tanımlama
- Analiz
- Plan
- Hafifletme
- İzleme
- İletişim

Güney Kaliforniya Üniversitesi Yazılım Mühendisliği departmanı yöneticisi Barry Boehm risk yönetim modeli ortaya koymuştur. Bu model iki fazdan oluşmaktadır: Risk değerlendirme ve risk kontrolü. Her faz üç element içermektedir[1].

- Risk değerlendirme
  - Risk tanımlama
  - Risk analizi
  - Risk önceliklendirme
- Risk kontrolü
  - Risk yönetim planı
  - Risk çözümü
  - Risk izleme

CMMI(Yetenek Olgunluk Model Entegrasyonu) ise risk yönetimi için aşağıdaki adımların yapılmasını önermektedir[1] :

- Organizasyonel önlem yayınlama
- Tanımlı işlemlerin yayınlanması
- Kaynakların doğrulanması
- Sorumlulukların atanması
- İnsanların eğitilmesi
- Konfigürasyon eğitimi
- Paydaşların tanımı
- Gözleme ve kontrol süreçleri

## 2.2 Risk Öznitelikleri

Risk öznitelikleri istenen aksiyon stratejisi veya çözülmek istenen probleme göre seçilmelidir. Bu amaçla risk öznitelikleri gruplanabilir. Risk öznitelikleri ve sınıflandırmalar karıştırılmamalıdır. Örnek vermek gerekirse etkilenecek anahtar sonuç alanları öznitelikken, etkilenecek olan kalite, performans gibi alt alanlar sınıfları oluşturmaktadır. Aynı şekilde alan bir öznitelikken proje, işlem ve ürün sınıf olarak nitelendirilebilir. Riskleri anlamak için seçilen öznitelik sayısı olabildiğince düşük tutulmalıdır. Odaklanılmak istenen projeye veya stratejilerde görülen problemlere göre öznitelik seçilmelidir[1][14].

Risk özniteliklerine baktığımızda iç ve dış riskleri görmekteyiz. İç riskler organizasyon yapısında bulunan risk faktörlerinden ortaya çıkmaktadır. İç riskler için alınan önlem planı daha sağlam temellere basmaktadır ve bu planlar sayesinde iç riskler ortadan önemli ölçüde kaldırılabilir. Aynı zamanda iç riskler kontrol edilebilir, ölçülebilir ve kolaylıkla izlenebilir. Dış riskler ise iç risklere oranla gözlemlenmesi, kontrolü ve metrik olarak ölçümü daha zordur. Alınan önlem planları daha farazidir[1].

Risk özniteliklerinden biride riskleri gözlemler. Risk gözlemler sınıfları ise risk tehlikeleri, risk sınırlama, nominal riskler ve önemsiz risklerdir. Risk tehlikelerine baktığımızda bu sınıfa giren riskler projelere etki değeri en yüksek olan risk grubundadırlar. Potansiyel olarak en üst seviyede zarar verecek olan risklerdir. Risk tehlikeleri yönetiminde Murphy kanunlarından dersler çıkarılabilir. Murphy kanunlarının temeli şu söze dayanmaktadır : “Eğer bir işi halletmek için birden fazla olasılık varsa ve bu olasılıklardan biri istenmeyen sonuçlar ve felaket doğuracaksa kesinlikle bu olasılık gerçekleşecektir”. Aslen Finagle Kanunu olarak geçen ve daha yaygın olarak bilinen bir söz ise şöyledir : “Ters gidebilecek her şey, ters gidecektir”.

Eğer risk tehlikeleri göze alınacaksa bunun geri dönüşünün çok büyük olması ve büyük faydalar sağlaması gerekmektedir. Göze alındığında da devamlı olan risk izlemesi yapılmalıdır. Önceliklendirme yapılarak risk ortaya çıkma numarası (REN) verilmesi de gerekmektedir. Bu riskler risk veritabanına girilmeli ve rutin analizler her risk için yapılmalıdır. Bu tarz risklerde her risk kendi içinde devam eden kontrol mekanizması da içermelidir[1].

Nominal risklerde ise derecelendirme yapılabilir. Pareto yasası kullanılarak bu risklerde önceliklendirme yapılabilir. Olasılık etki uzayında da bu riskler sınıflandırılabilir. Çizelge 2.5’ den de bu durum görülebilir.

Çizelge 2.5 Olasılık/Etki Seviyesi

| Sınıf | Olasılık | Etki Seviyesi |
|-------|----------|---------------|
| Q1    | Yüksek   | Düşük         |
| Q2    | Yüksek   | Yüksek        |
| Q3    | Düşük    | Yüksek        |
| Q4    | Düşük    | Düşük         |

Önemsiz riskler ise bu grupta etki değeri en az olan risk grubudur. İhmal edilebilirler.

### 2.3 Yazılım Risk Taksonomisi

SEI’nın yaptığı taksonomi bazlı araştırma sonucunda üç temel sınıf ortaya çıkmıştır. Bunlar ürün mühendisliği, geliştirme çevresi ve program kısıtlamalarıdır. Bu kategorilerin her biri temel olarak gereksinim, tasarım, kod-birim testleri ve iletişim alt alanlarını etkilediği gözlemlenmiştir. Risk taksonomi tablosunda bu adımlar detaylıca açıklanmıştır[1][16].

Çizelge 2.6 Risk Taksonomi Çizelgesi

| Adım       | Alt Alanlar                                                                                                                |                                                                                                                                                                                                                                                                                                                               |
|------------|----------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Gereksinim | Değişkenlikleri<br>Anlaşılabilir olması<br>İmplementasyon<br>İzleme                                                        | Gereksinimlerin değişmesi, eklemesi veya kararsız olması durumları<br>Anahtar alt alanlar anlaşılabilir şekilde yazılmış mı?<br>İmplementasyon yapmak kolay mı?<br>Tasarım,kodlama ve test fazlarında kolay bir şekilde izleme yapılabilir mi?                                                                                |
| Tasarım    | Fonksiyonallite<br>Zorluk<br>Arayüzler<br>Performans ve kalite<br>Test<br>Donanımsal kısıtlar<br>Yeniden kullanılabilirlik | Gereksinimleri karşılayacak özel algoritmalar var mı?<br>Gerçekleştirilmesi zor gereksinim veya fonksiyon var mı?<br>İç ve dış arayüzler iyi tanımlanmış mı?<br>Beklenen performans, kalite veya dizaynda problemler var mı?<br>Yazılım kolay bir şekilde test edilebiliyor mu?<br>Yazılım yeniden kullanılabilirliği var mı? |

Çizelge 2.6 (Devam)

|                   |                                                                              |                                                                                                                                                                                                                                                                                                     |
|-------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kod ve birim test | İmplementasyon<br>Test edilme<br>Kodlama<br>Çevre<br>Ürün<br>Sistem<br>Bakım | Seçilen algoritma ve tasarım kolay bir şekilde uygulanabiliyor mu?<br>Birim test için çok zaman gereklimi<br>Kodlama için uygulanan dil yeterli mi?<br>Takımın kodlama zamanı yeterli mi?<br>Donanımsal veya yazılımsal bir engel var mı?<br>Tüm gereksinimler için kabul kriterleri onaylanmış mı? |
| İletişim          | Takım içi uyum<br>Motivasyon                                                 | Takım arasında iyi bir uyum var mı?<br>İletişim takım içinde yeterli düzeyde mi?<br>Görevler düzgün bir şekilde dağıtılabiliyor mu?<br>İyi bir ürün çıkartabilmek için takım yeteri kadar motive mi?                                                                                                |

### İLGİLİ ÇALIŞMALAR

Günümüzde yazılım ve teknoloji alanındaki gelişmeler insanların hayatlarını kolaylaştırmakta, değer katmakta, kalite açısından da üst sıralara getirmektedir. Bununla paralel olarak teknoloji ve yazılımlarda görülen risklerde artmaktadır. Risklerin artması sonucu sistemler ve süreçler üzerindeki riskler şirketler ve kuruluşlar tarafından daha fazla fark edilmekte ve önemli olarak görülmektedir. Bilgi sistemleri, süreç, ürün üzerindeki hatalar ve güvenlik açıkları ciddi iş krizlerine ve itibar kayıplarına yol açmaktadır. Bu durum yazılım mühendisliğinde risk yönetiminin önemini arttırmaktadır.

Risklerin arttığı bu ortamda yazılım risk yönetiminde önemini arttırmıştır. Yazılım risk yönetiminde risk faktörlerinin tanımlandığı, kıyaslandığı, gereksiz faktörlerin çıkarıldığı ve etki değeri yüksek faktörlerin belirlendiği ve gerekirse hafifletildiği bir süreçtir. Riskler tanımlanır, analizi yapılır, planlanır, kritiklik seviyesi gerekli riskler izlenir ve önlemler alınır. Kurum ve kuruluşlar tüm süreçlerini gözden geçirmekte ve riskleri önleme, indirgeme ve önlem alma adına çalışmalarda bulunmaktadır.

Yazılım mühendisliğinde risk yönetimi ile ilgili çalışmalara bakıldığında ise son yıllarda artış gözlemlenmektedir. Bu alanda yapılan çalışmalar geçmiş yıllara göre çeşitlilik göstermekte ve önem kazanmaktadır. Genel risk yönetimi, risk faktörleri, risk önlemek için modellerin kullanılması ile ilgili çalışmaların yayınlandığı, bu konular üzerinde daha detaylı durulduğu gözlemlenmiştir. Bu kısımda yazılım mühendisliğindeki risk yönetimi konusunda yapılan çalışmalar iki alt başlık altında toplanmıştır. İlk başlıkta risk yönetimi ve analizi ile ilgili dergi ve makale çalışmaları incelenmiş ve derlenmiştir. İkinci başlıkta

ise risk faktörlerine yer verilmiş, faktörlerin birbirleriyle kıyaslanması, eleminize edilmesi ve sıralanması ile ilgili yapılan çalışmalardan bahsedilmiştir. Ayrıca yazılım mühendisliği yaklaşımları kullanılarak oluşturulan modeller ortaya konulmuş ve modellerinde birbirleriyle kıyaslanmasına yer verilmiştir.

### **3.1 Risk Yönetimi ve Analizi ile İlgili Yapılan Çalışmalar**

Risk yönetimi ve analizi ile ilgili yapılan genel çalışmalar bu başlık altında toplanmıştır. Tahminler, belirsizlikler ve riskler ele alınmış ve bu belirsizliklerin/risklerin sebepleri incelenmiştir. Ayrıca risk yönetim metodolojilerinden ve risk adımlarından bahsedilmiş ve risklerin bağımlılıkları da bu kısımda değinilmiştir. Risk analizi alanında yapılan çalışmalarda bahsedilerek risk yönetimi ve analizi ile ilgili yapılan çalışmalar tamamlanmıştır.

Kitchenham ve Linkman bu çalışmasında tahminler, belirsizlik ve risk konularını ele almışlardır. Yöneticilerin tahmin etmedeki başarısızlıklarına değinmişler, yazılım organizasyonundaki problemlerden bahsetmişlerdir. Belirsizliklerin ve riskin efektif bir şekilde yönetilemediğine ve faktörlerin organizasyonel yapıda tam anlaşılamadığına değinmişlerdir. Projelere göre kişiler atanmakta, yanlış atamalar sonucu riskler ortaya çıkmaktadır. Ölçüm hataları ve kullanılan model hataları kestirimi gerçeklikten uzaklaştırmakta ve hatalara sebebiyet vermektedir. Aynı şekilde yanlış tanımlanan kapsamda risklere sebebiyet verecektir. Belirsizliği ve riskleri ortadan kaldıracı için ise organizasyona odaklanması gerektiği vurgulanmıştır. Tüm organizasyon yapısı açık bir şekilde ortaya konulması gerektiğinden bahsedilmiştir. Aynı şirket içinde yapılan büyük çapta bir projenin, küçük birkaç projeden daha kritik olduğunu ortaya koymuş ve tahminlerin elde edilen projelerdeki hataların ortalamalarına göre yapılması gerektiğini belirtmiştir. Tahmin etme modellerini ise bu faktörleri göz önünde bulundurarak oluşturulması gerektiğine vurgu yapılmış ve belirli olmayan risk faktörleri için varsayımlarda bulunması gerektiğini belirtmiştir. Hatalı varsayımların etki değerleri göz önünde bulundurularak değerlendirme modellerinin çıkarılabileceğinden bahsedilmiştir[17].

Kwan ve Leung çalışmasında ise risk yönetim metodolojisini ve proje risk bağımlılıklarını ele almıştır. Günümüzdeki proje risk tanımını ve risk yönetiminden

bahsetmişlerdir. Risk tanımında iki etmenin olduğunu, bunların ise risk olasılık değeri ve risk etki değeri olduğunu belirtmişlerdir. Genel risk yönetim uygulamalarından ve bu uygulamaların alt başlıkları olan proje riskleri tanımlama, riskleri önceliklendirme, risk tepki planı oluşturma, durum izleme ve aksiyon alma ve risk kontrol etmeden bahsetmişlerdir. Burada anlatılanlarda bölüm ikide anlatılanlara paralellik göstermektedir. Risk bağımlılığı alanında farklı metotlar kullanılarak çalışmalarını gerçekleştirmişlerdir. Konservatif metod, optimistik metod ve ağırlık metodları kullanılarak risk bağımlılığı grafi oluşturulmuştur. Riskler birbirlerinden ayrılarak grafların oluşturulması birbirini etkileyen riskleri ortaya çıkarmak adına yapılmıştır. Yani bir riskin oluşması başka bir riskin oluşma ve oluşmama olasılığını arttırabilmekte veya azaltabilmektedir. Risk grafları oluşturulduktan sonra ise risk yönetim paradigmaları risk bağımlılığı yönünden ele alınmıştır. Buna göre 3 farklı durum için kritiklik seviyeleri belirlenmiştir. Bunlardan ilki iç ciddiyeet seviyesidir. Risk yönetim paradigmalarında direkt olarak kullanılmaktadır. İkinci ise diğer riskin oluşma durumuna için kullanılan ciddiyeet seviyesidir. Üçüncü olarak ise tepki ciddiyeet seviyesi kullanılmaktadır. Bu seviyelere göre farklı aksiyon planları alınmaktadır[18].

Benaroch ve Goldstein ise risk yönetimini ekonomik optimizasyon yaklaşımı ile incelemişlerdir. Bu yaklaşımın göre nasıl bir hafifletme senaryosu uygulayacaklarını, risklerin nasıl kontrol altına alınacağını ve proje değerini en üst seviyeye nasıl taşıyacaklarını anlatmışlardır. Çalışma kapsamında da projede görülebilecek risklerin maliyet tablolarını çıkartarak hafifletme senarosunun hangi faktörler üzerinde yapılacağına karar vermişlerdir. Sistem risk yönetim problemleri ekonomik bir bakış açısıyla çözmeye uğraşmışlardır. Risk faktörleri ve risk önleme planları tanımlanmış ve etki değerlerine göre bütçelendirme yapılmıştır[19].

Ray ve Mohapatra ise risk analizi ile ilgili bir çalışma yayınlamışlardır. Yazılım geliştirme yaşam döngüsü içinde analiz ve dizayn adımları için durum bazlı risk önleme metodolojisi üzerine çalışmalar yapmışlardır. Risk tahmin metodu geliştirerek komponentler arası karmaşıklık indirgeyip herbir bileşeni graf yöntemleriyle incelemişlerdir. Bunun yanında kritiklik analizide yaparak sistem seviyelerindeki

başarısızlık senaryolarınında analizini yapmışlardır. Buna göre hata seviyelendirmeyi felaket, major, marjinal ve minör olarak 4 seviyede toplamışlardır[20].

### 3.2 Risk Faktörleri ve Risk Modellerinin Kullanılması ile İlgili Çalışmalar

Risk faktörlerinin değerlendirilmesi, sınıflandırılması ve eleminize edilmesi ile ilgili çalışmalarla risk modellerinin oluşturulması ve oluşturulan bu modellerin birbiriyle kıyaslanması ile ilgili çalışmalar bu başlık altında toplanmıştır. İncelenen çalışmalarda yapay sinir ağları, destek vektör makineleri, Naive Bayes, kNN yöntemi, bulanık mantık, çok katmanlı yapay sinir ağları kullanılmıştır. Bu sınıflandırma yöntemleri sayesinde risk verileri ve risk faktörleri sınıflandırılabilmiş, bilgisayar mühendisliği yaklaşımları kullanıldıktan sonra ise risk modelleri oluşturulmuştur. Yapılan çalışmalar aşağıda listelenmiştir.

Sun Yat Sen Üniversitesinden yayınlanan makalede yapay sinir ağları ve destek vektör makineleri yöntemleri kullanılarak yazılım projelerinin risk yönetiminin modellenmesi sağlanmıştır. 300 farklı firmaya gönderilen 194 anket sorusuyla birlikte kategoriler daha da netleştirilmiştir. Anket sonuçlarında 6 kategori bulunmuştur. Bunlar ortam riskleri(müşteri tabanlı riskler), proje gereksinimi riskleri, işbirliği riskleri, takım riskleri, proje yönetimi riskleri , mühendislik riskleridir(yazılım müh adımlarının uygulanması ve yönetilebilirliği) . Anket çalışmasına ek olarak risk faktörlerinin belirlenmesinde 30 uzman( 4 tanesi bu alanda profesör ve 26 tanesi uzman) çalışmıştır. Belirlenen 64 farklı risk yukarıda yer alan 6 ana risk grubu içerisinde tanımlanmıştır. 300 anketten geri dönüş olarak 133 anketin geri dönüşü olmuş ve bu 133 tane anket sonucundan 120 tanesi belirleyici anket olarak ele alınmıştır. Anketlerin %71'i proje yöneticilerine, proje liderlerine ve müşteri müdürlerine sorularak elde edilmiştir.%25'i ise takım çalışanlarına sorularak elde edilmiştir. Standish grubun başarılı proje oranı %27-29 oranındayken anketlerden elde edilen sonuçta bu veriyi doğrular nitelikte %26 olarak saptanmıştır. Rastgele veriseti ikiye bölünerek 100 örnek(anket sonucu) eğitim verisi olarak kullanılırken 20 örnek test amaçlı kullanılmıştır. Orijinal proje sonuçları üçe ayrılarak başarılı proje, üzerine uğraşılan(Yarı Başarısız) ve tamamen başarısız olarak sınıflandırılmıştır. Riskler standart yapay sinir ağlarıyla, genetik algoritma ile iyileştirilmiş yapay sinir ağlarıyla ve SVM yaklaşımlarıyla tahminleri yapılmıştır. NN elde

edilen sonuçlar projelerin gerçek sonuçlarıyla kıyaslandığında %70 oranında bir doğruluk oranı elde edilmektedir. SVM kullanarak elde edilen başarı oranı ise %80'nin üzerinde yer almaktadır[21].

Sun Yat Sen Üniversitesinin yayınladığı bir diğer makalede yazılım geliştirme sırasında oluşabilecek riskleri bir faktöre veya proje sonuçlarına bağlı değil projenin tümünü ele alarak tahmin ve kontrol eden akıllı bir modelin oluşturulması amaçlanmıştır. Risklerin belirlenmesi için yazılım firmalarından elde edilen veriler kullanılarak risk tahmin modeli oluşturulmuştur. Modelin performansını değerlendirmek için iki makine öğrenmesi yaklaşımı olan ANN ve SVM yöntemleri kullanılmış ve sonuçları kıyaslanmıştır. 64 risk faktörü 30 uzman ve birçok akademisyene sorularak 41'e indirgenmiştir. Bu faktörlerden oluşan anketten 300 tane hazırlanarak Çin'de yer alan firmalara yollanmıştır. Bu firmaların 133 tanesinden geri dönüş alınmış ve 120 anket sonucu projede değerlendirilmeye alınmıştır. Anket yollanan firmalarda projelerin başarı oranı beklenildiği gibi %22 çıkmıştır. Buda anket sonuçlarının doğruluğunu işaret etmektedir. 120 proje örneği rastsal olarak 2 parçaya bölünerek 100 proje verisi eğitim 20 proje verisi ise anket için kullanılmıştır. 64 risk faktörü girdi olarak ve 8 faktör çıktı olarak belirlenmiştir. Çıktı faktörleri doğrusal dönüşümle 0-1 arasına normalize edilip birleştirilerek 3 tip proje çıktısı belirlenmiştir. Bunlarda başarılı, yarı başarılı veya üstünde uğraşılmaya devam edilen, tamamen başarısız olarak gruplandırılmıştır. SVM ile oluşturulan modelin başarı oranı 20 test projesi için %85, ANN için ise %75 oranında başarı sağlamaktadır. Bu çalışma SVM tabanlı oluşturulan risk tahmin modelinin ANN ile oluşturulan modelden daha iyi çalıştığını gözlemlenmiştir. Ayrıca risk tahmini hataları 2 kategoriye ayrılmıştır. Bunlarda riski yüksek beklemek yani projede gerçekleşen riskin maliyetine oranla daha fazla önlem için zaman ve bütçe harcamak ve riski düşük almak yani gerçekleşen riskin için yeterli önlem almamaktır buda projenin başarısız olmasını sağlayabilmektir[22].

Windsor üniversitesinde yapılan çalışmada ise NaiveBayes sınıflandırıcısı ile K-en yakın komşuluk sınıflandırıcısının performansları incelenmiştir. Veri setlerine uygulanan bu iki yöntemin performansları belirtilmiştir. 671 veri 9 sınıfa ayrılmak istenmiş. NaiveBayes sınıflandırıcısı %12,43 hatayla verileri doğru sınıflandırabilmiştir. K en yakın komşuluk metodunda ise k değerine göre yüzdeler değişkenlik göstermiştir. K değerinin

5 seçildiği durumlarda hata oranı en düşük çıkmıştır. %9,45 hata oranıyla sınıflandırma yapabilen K en yakın komşuluk yöntemi NaiveBayes yöntemine göre daha başarılı olmuştur[23].

Bulanık mantık alanında da risk değerlendirme alanında çalışmalar yapılmıştır. Hunan üniversitesinde yapılan bir çalışmada risklerde oluşan kayıplar üzerine yoğunlaşmıştır. Riskleri değerlendirmek, risklerin olasılıkları bulmak ve risklerin etkisini hesaplamak için bulanık mantık dili kullanılmıştır. Risklerin hesaplanması ve hesaplanan değerlerin bulanıklığı giderilmesinden sonra sıralanması gerekmektedir. Bunun içinde risk faktörlerini belirlemişler ve risklerin önem sırasına göre risk analiz ağı oluşturmuşlar(Semantik indirgemeler kullanılarak). Sonuç olarak yaptıkları çalışmada yazılım projelerindeki risk faktörlerin kritiklik seviyesini analiz etmişlerdir. Bulanık mantık alandaki bir diğer çalışmada ise yazılım projelerindeki tasarım risk analizi incelenmiştir. Yazılım projelerindeki risklerin tanımı ve kritiklik seviyelerinin farklılık göstermesinden dolayı böyle bir çalışma yayınlamışlardır. Proje riskleri ve ayrı bileşenler arasında bulanık mantık yöntemi kullanılmıştır. Bunun sonucunda da düşük, orta ve yüksek seviyeli riskleri tespit edebilmişlerdir[24][25].

UYMS 2012 ve UYMK 2012 de yayınlanan Yıldız Teknik Üniversitesinde M.ÖzgürCingiz ve Oya Kalıpsız ile yaptığım makale çalışması çerçevesince risk faktörleri belirlenmiş, risk verileri toplanmış ve model oluşturulmuştur. TurkcellICT’den elde edilen, etki analizi için kullanılan verilere göre uygulama gerçekleştirilmiştir. 6 risk faktörü ve 7 problem etki özelliğinden oluşan 384 veri çalışmada kullanılmıştır. Destek vektör makineleri(DVM), karar ağacı(J48) ve naivebayes(NB) yaklaşımlarıyla projede yer alan problemlerin risk faktörleri üzerindeki etkileri ve problemlerin öneminden oluşan veri kümesi eğitilerek 10 kere çapraz doğrulama yöntemiyle test edilmiştir. DVM, J48 ve NB sınıflandırıcılarından elde edilen test sonucuna dair F değerleri(kesinlik ve geri çağırma ölçüm değerlerinin harmonik ortalamasından elde edilen ölçüm değeridir) sırasıyla 0.96, 0.904 ve 0.885’dir. F değerlerinden yola çıkarak sınıflandırıcıların başarıları ölçüldüğünde başarı oranlarına göre DVM,J48 ve NB olduğu gözlemlenmiştir. DVM 384 problem etki analiz verisinin 371 adedini doğru olarak belirleyebilmiştir. 13 tanesini ise farklı etiketlemiştir. Modelin doğruluk oranı ise %96,61’dir. J48, 384 problem etki analiz verisinin 350 adedini doğru olarak belirleyebilmiştir. 34 tanesini ise farklı etiketlemiştir.

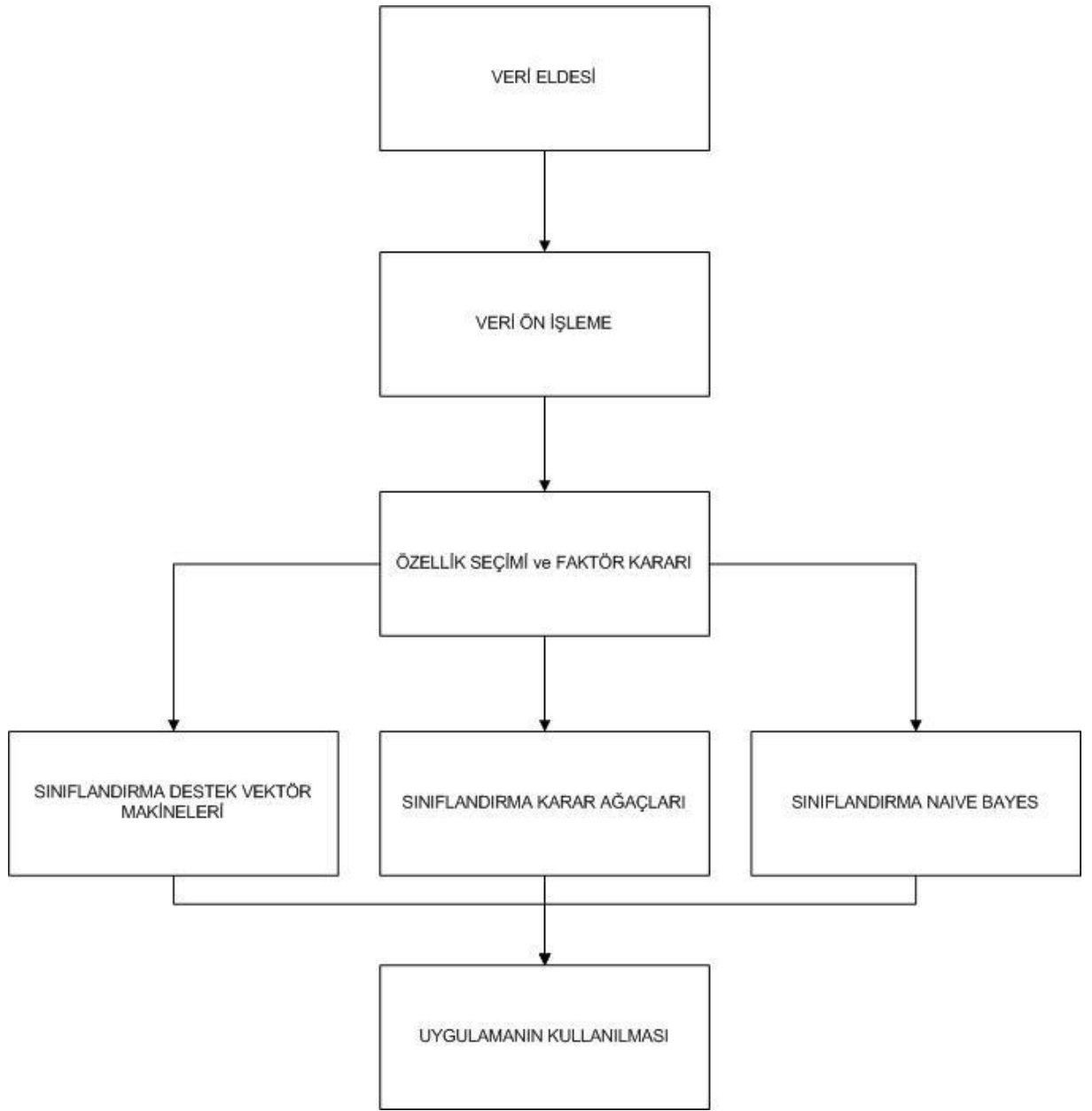
Modelin doğruluk oranı ise %91,14'dür. NB ise 384 problem etki analiz verisinin 352 adedini doğru olarak modelleyebilmiştir. 32 tanesini ise farklı etiketlemiştir. Modelin doğruluk oranı ise %91,66 dır. DVM sınıflandırıcı diğer sınıflandırıcılardan daha başarılı sonuç verdiği bu makalede de gözlemlenmektedir. Problemlerin proje üzerindeki etkisi %96'lara varan sınıflandırma başarı ölçütü değeriyle kestirilmektedir. Ayrıca karar ağaçları ile yapılan sınıflandırmanın NaiveBayes ile yapılan sınıflandırmadan da iyi sonuçlar verdiği gözlemlenmiştir. Çalışma çerçevesinde tek proje ve tek projede yer alan problemlerin projeye etkisi üzerine çalışılmış böylece risk faktörlerinin her bir probleme göre aldığı değerlerin projeye etkisi de gözlemlenmiştir. UYMK 2012'de bunlara ek olarak bazı sınıflandırıcılar eklenmiştir. Sınıflandırıcı olarak Destek Vektör Makineleri (DVM), k-En Yakın Komşuluk (kNN), Karar Ağacı (J48), NaiveBayes (NB) ve Çok Katmanlı Yapay Sinir Ağları (ÇKYSA) kullanılmıştır. Sınıflandırma performansını incelemek için Kesinlik, Anma, F-ölçeği ve Kappa istatistiği çalışmamızda kullanılmıştır. En iyi başarıyı Çok Katmanlı Yapay Sinir Ağları sınıflandırıcısından elde edilmiştir. Çok Katmanlı Yapay SinirAğları'nı yüzde 96 ile Destek Vektör Makineleri takip etmiştir. Başarısı en düşük olan sınıflandırıcılar ise NaiveBayes ve k-En Yakın Komşuluk sınıflandırıcılarıdır[26][27].

### UYGULAMA TASARIMI ve VERİ HAZIRLAMA

Uygulamayı tasarlayabilmek için öncelikli olarak ilgili çalışmalar incelenmiştir. İlgili çalışmalar kısmına bakıldığında yapılan çalışmalar birbirleri ile benzerlik göstermektedir. Öncelikli olarak firmalardan risk verileri toplanmıştır. Elde edilen risk verileri benzer projeler üzerinedir. İlgili firma sektöre yönelik projeler üretmekte, üretim sektörü içinde gömülü sistem tasarımlarında yapmaktadırlar. Ayrıca elde edilen tüm veri setleri de birbirine benzer projeleri kapsamaktadır.

Bu aşamadan sonra uzman görüşleri yardımıyla, yapılan anket sonuçlarına göre veya özelliklerin proje/probleme etki derecesine göre özellik sıralamaları yapılarak faktör kararına varılmıştır. Projeye/probleme en çok etki edecek faktörler belirlenip önemsiz faktörler elenmiştir. Risk faktörleri belirlendikten sonrada ilgili sınıflandırıcılar kullanılıp risk modelleri ortaya konulmuştur.

Tez çalışması kapsamında yapılan çalışma da ilgili çalışmalara paralellik göstermektedir. Bir yazılım firmasından projelerindeki hata verileri toplanmıştır. Projeler seçilirken uzman görüşleri alınmıştır. Bunun yanında projelerdeki hataların özelliklerine bakılmış ve ortak parametreleri ve benzer risk faktörleri olanlar tercih edilmiştir. İlgili projeler ve risk faktörleri belirlendikten sonra her projede tekil olan (unique) değerlerde elenecektir. Daha sonra kalan faktörler için hata seviyelendirmesi üzerindeki etkilerine bakılır. Bu işlem için bilgi kazanımı yöntemi kullanılacaktır. Bilgi kazanımı özellik seçiminde sınıf bilgisini belirleyen en değerli özelliklerin çıkarımını sağlamaktadır. Değerli özellikler sınıf bilgisi için ayırt edici özelliklerdir. Bilgi Kazanımı farklı sınıflar içerisinde ayırt edici özelliği fazla olan özellikleri belirlemeye çalışır.



Şekil 4.1 Sistem Mimarisi

Faktörler sıralandıktan sonra sınıflandırıcı model seçilir ve veriler sınıflandırılır. Sınıflandırma işleminden önce kümeleme işlemleri de yapılacaktır. Kümeleme işlemlerinde uzman görüşü alınmıştır. Birbirine benzer dört proje kendi aralarında ikiyeşerli olarak kümelenecek ve sonuçları ortaya konacaktır.

Sınıflandırıcı olarak ise Destek Vektör Makineleri (DVM), Karar Ağaçları (KA) ve NaiveBayes(NB) sınıflandırıcıları kullanılacaktır. Bu sınıflandırma özelliklerinin seçiminde ilgili çalışmalar baz alınmıştır. DVM sınıflandırıcısının KA ve NB'den daha iyi

sonular vereceėi beklenmektedir. KA'nında NB ye gre daha iyi sonular vereceėi hedeflenmiřtir. Sınıflandırma iřlemlerinden sonra kullanılan sınıflandırıcılar detaylı řekilde listelenecektir. Sistem mimarisi akıř diyagramına řekil 4.1'den eriřilebilmektedir.

řekil 4.1'e gre ilk olarak veri eldesinden bahsedilecek olup, daha sonra verilere ne řekilde bir n iřleme yapılacaėından bahsedilecektir. zelliklerin ve faktrlerin elenmesi ve kullanılacak faktrlerin kararından sonra sınıflandırma yntemleri anlatılacak olup son olarak da Weka zerinde uygulamanın nasıl kullanıldıėına deėinilecektir.

#### **4.1 Veri Eldesi**

Veri elde iřlemi iin bir yazılım firmasına bařvurulmuřtur. Firmanın yaptıėı projeler incelenmiřtir. Toplamda 13 proje iin hata veri setleri elde edilmiřtir. Bazı projelerin verilerinin diėerlerine gre daha farklı olduėu gzlemlenmiřtir. 8 projenin birbiriyle benzerlik gstermesi zerine uzman grřleri de alınarak proje sayısı 8'e indirgenmiřtir. Veriler Excel formatında proje bazlı řekilde temin edilmiřtir. Her proje kendi iinde en az 206 en fazla 3007 hata sayısına sahiptir. Her hata birok etkenden ve faktrden oluřmuřtur. Bu faktrlerinde tm projeler iin ortak olmasına uėrařılacaktır. Veri n iřleme ařamasında detaylı bir řekilde bu sreten bahsedilecektir.

#### **4.2 Veri n iřleme**

Veri n iřleme iin elde edilen 13 proje ilk olarak uzman grřleri yardımıyla 8 projeye indirgenmiřtir. Excel olarak verilen veri setlerinin Wekaarayzyle uyumlu hale getirilmesi gerekmektedir. Wekaarayz iin verilerin "arff" uzantılı zel bir formata getirilmesi gerekmektedir.

Arff formatı iin ilk olarak proje ismi girilmelidir. Proje ismi girildikten sonra ilgili zellikler yani faktrlerin girilmesi gerekmektedir. Faktrler girilirken her bir faktrn alabileceėi deėerlerde yan tarafında belirtilmelidir. Faktr seimi tamamlandıktan sonra tm hata verileri aralara virgl konularak tek tek girilmelidir.

Çizelge 4.1 Arff Format Örneği

| Adım | Bilgi                                                    |
|------|----------------------------------------------------------|
| 1    | @Relation Proje1                                         |
| 2    | @attributeTurBilgisi {Fonksiyonel,Arayuz,Baglanti,.....} |
| 3    | @attributeSirketPayda {Paydas,SirketBazli,Outsource}     |
| 4    | ...                                                      |
| 5    | ...                                                      |
| 6    | @data                                                    |
| 7    | Fonksiyonel,SirketBazli...                               |
| 8    | Arayuz,Outsource,...                                     |
| 9    | ...                                                      |

Çizelge 4.1’ de de belirtildiği üzere “Relation” kısmına proje ismi girilmelidir. 8 proje tek tek ele alındığında Proje1’den Proje8’e kadar hepsi “@” işaretiyle belirtilmelidir. Tüm projeler ortak olarak ele alındığı durum içinde ProjeGenel olarak belirtilmelidir. Ayrıca kümelenen veriler içinde ProjeGrup1 ve ProjeGrup2 içinde projeler belirtilmelidir. Toplamda 11 proje için farklı arff dosyaları oluşturulmuştur.

Oluşturulan 11 dosya için @attribute yani özellikler Türkçe karakter içermeyen girilmelidir. Format olarak da tablo x.x’de belirtilen yapı kullanılmıştır. Belirlenen 13 faktör için(Bölüm 4.3’de anlatılmıştır) “attribute” bilgisi gene çizelge 4.1’de belirtilen formata göre girilmiştir.

Veri kısmına bakıldığında ise “@data” ifadesinden sonra projelerin hata verileri formatlı bir şekilde yazılmalıdır. Veriler Türkçe karakter,boşluk ve noktalama işaretleri içermemelidir. Veri hücreleri içindeki tüm Türkçe karakterler, boşluk ve noktalama işaretleri veri setinden kaldırılmış ve uygun formata getirilmiştir. 8135 hata verisi için bu işlemler gerçekleştirilmiştir.

Format tamamlanıp veriler hazırlandıktan sonra \*.arff uzantısıyla veriler dosyaya kaydedilmelidir. Bu şekilde Weka arayüzü üzerinden bu dosyalar hatasız bir şekilde açılabilme ve sınıflandırmaya uygun hale gelmektedir.

#### 4.3 Özellik Seçimi ve Faktör Kararı

Yazılım alanında faaliyet gösteren firmadan 12 proje için hataların listelendiği veri setlerine bakıldığında tüm projelerin eş faktörleri içermediği gözlemlendi. Bu nedenle 12 projeden 8 tanesi tez çalışması kapsamında kullanılmıştır. Uzman görüşleri de aynı yöndedir. 8 projenin hata verileri incelendiğinde 32 faktörün hatanın kritiklik seviyesine direkt olarak etki etmediği gözlemlenmiştir. Bazı risk faktörleri tekil değerler aldığı için kritiklik seviyesine hiçbir etkisi bulunmamaktadır. Hata bazlı risk faktörlerine baktığımızda aşağıdaki gibi sıralandığı gözlemlendi. 32 faktörden oluşan bu set aşağıda listelenmiştir.

Çizelge 4.2 Projelerde Yer Alan 32 Risk Faktörü

| RISK ID | RISK FAKTÖRLERİ                                |
|---------|------------------------------------------------|
| 1       | Hatanın Numarası                               |
| 2       | Hatanın Global Numarası                        |
| 3       | Hatanın Özeti                                  |
| 4       | Hatanın Tanımı                                 |
| 5       | Hatanın Global Tanımı                          |
| 6       | Kişi-Grup-Şirket bazında atanmış hata          |
| 7       | Kişi-Grup-Şirket bazında çözülen hata          |
| 8       | Kişi-Grup-Şirket bazında değiştirilen hata     |
| 9       | Kişi-Grup-Şirket bazında bulunan hata          |
| 10      | Hatanın Tekrarlandığı ve Değiştirildiği Tarih  |
| 11      | Hatanın Çözüldüğü Tarih                        |
| 12      | Hatanın Bulunduğu Tarih                        |
| 13      | Hatanın Paydaşlar Tarafından Yaratıldığı Tarih |
| 14      | Hatanın Paydaşlar Tarafından Çözümleme Tarihi  |
| 15      | Hatanın Paydaşlar Tarafından Kapatılma Tarihi  |
| 16      | İç Müşteriye Etki Durumu                       |

Çizelge 4.2(Devam)

|    |                                               |
|----|-----------------------------------------------|
| 17 | Dış Müşteriye Etki Durumu                     |
| 18 | Paydaşlara Etki Durumu                        |
| 19 | Proje Seviyesine Etki Durumu                  |
| 20 | Müşteriye ve Marka Bilinirliğine Etki Durumu  |
| 21 | Performansa Etki Durumu                       |
| 22 | Yönetim İlgisi Seviyesi                       |
| 23 | Zamana Etki Durumu                            |
| 24 | Risk Sahiplik Seviyesi                        |
| 25 | Hatanın Proje İçindeki Seviyesi               |
| 26 | Hatanın Tür Bilgisi                           |
| 27 | Hatanın Görüldüğü İşletim Sistemi             |
| 28 | Hatanın Çözülmesinde Öncelik Gösterilmeli Mi? |
| 29 | Hatanın Hangi Grupta Gösterileceği Bilgisi    |
| 30 | Hatanın Öneri Olup Olmama Durumu              |
| 31 | Hatanın yazılımsal-tasarımsal olup olmadığı   |
| 32 | Hatanın Şirket Bazlı Olup Olmama Durumu       |

Yapılan çalışmada risk faktörleri belirlenmiş, risk verileri toplanmış ve model oluşturulmuştur. Yazılım sektöründe çalışan bir firmadan elde edilen, hataları seviyelendirmek için kullanılan alanlara göre faktörler belirlenecektir. Risk kestirimi sektörden sektöre farklılaşabileceği gibi, aynı işi yapan firmalar arasında bile riskler farklı olarak belirlenebilmektedir. Her yeni projeye göre önceki projeler için tanımlanan riskler farklılaşabilmektedir. Bu farklılık projelerde oluşan problemlerde dahi görülmektedir. Bu da risk yönetiminde standart oluşturmayı zorlaştırmaktadır. Bu nedenle bir projeye başlamadan önce ele alınması gereken ilk kısım risk faktörlerinin analizi olmalıdır. Belirlenen riskler ve etkileri tüm projeyi etkileyebileceği gibi proje içinde oluşacak problemler içinde önemli birer faktördür. Farklı risk faktörlerinin projeye veya probleme hangi oranda etki edeceği bulunmalı, gereksiz riskler projeye veya probleme dahil edilmemeli, önemli olan risklerde gözden geçirilmelidir.

Çizelge 4.3 Hata Risk Faktörleri ve Değerleri

| FAKTÖRLER                                   | ALT ALAN    | DEĞERLER                                                                                                                            |
|---------------------------------------------|-------------|-------------------------------------------------------------------------------------------------------------------------------------|
| Hatanın öneri olup olmama durumu            | 2 alt alan  | "Hata öneridir", "Hata öneri değildir"                                                                                              |
| Hatanın proje içindeki durumu               | 10 alt alan | "Yeni hata", "Çözülen Hata", "Kapatılan Hata", "Paydaşlara gönderilen hata", "Tekrarlanamayan hata", "Otomasyonda görülen hata" vs. |
| Hatanın görülme seviyesi                    | 6 alt alan  | "Her zaman", "Bazen", "Nadiren", "Sık sık", "Hiçbir zaman", "Otomotize edildiğinde"                                                 |
| Kişi-grup-şirket bazında bulunan            | Değişken    | Kişi-grup-şirket sayısına göre değişkenlik gösteriyor                                                                               |
| Kişi-grup-şirket bazında çözümleyen         | Değişken    | Kişi-grup-şirket sayısına göre değişkenlik gösteriyor                                                                               |
| Kişi-grup-şirket bazında değiştiren         | Değişken    | Kişi-grup-şirket sayısına göre değişkenlik gösteriyor                                                                               |
| Hatanın tür bilgisi                         | 30 alt alan | "Fonksiyonel", "Arayüz", "Bağlantı", "Kullanılabilirlik", "Otomasyon", "Faturalama", "Güvenlik" vs.                                 |
| Hatanın önceliğinin olma durumu             | 4 alt alan  | "Öncelikli", "Öncelikli değil", "Koşullu öncelikli", "Koşullu öncelikli değil"                                                      |
| Hatanın görüldüğü işletim sistemi           | 2 alt alan  | "Windows", "Linux"                                                                                                                  |
| Hatanın şirket bazlı olup olmama durumu     | 3 alt alan  | "Şirket bazlı", "Paydaşlar", "Outsource"                                                                                            |
| Hatanın hangi grupta gösterileceği          | 10 alt alan | Tür bilgisi ile paralellik gösteriyor                                                                                               |
| Hatanın yazılımsal-tasarımsal olup olmaması | 2 alt alan  | "Yazılımsal", "Donanımsal"                                                                                                          |
| Kişi grup şirket bazında atanmış hata       | Değişken    | Kişi-grup-şirket sayısına göre değişkenlik gösteriyor                                                                               |

Yapılacak projede temel olarak 32 risk faktörü etki değeri belirlenmiştir. Faktörlerden 19 tanesi tekil ve diğer projelerde ortak olmadığı için elenmiştir. Kalan faktörler ve alt alanları detaylı şekilde çizelge 4.3'de listelenmiştir.

Çizelge 4.4 Bilgi Kazanım Özellik Seçimine Göre Etki Değer Sıralaması ve Etki Değerleri

| RISK FAKTÖRLERİ                             | ETKİ DEĞERLERİ |
|---------------------------------------------|----------------|
| Kişi grup şirket bazında atanmış hata       | 0.538717       |
| Kişi grup şirket bazında çözümleyen         | 0.433654       |
| Hatanın proje içindeki durum bilgisi        | 0.419192       |
| Kişi-grup-şirket bazında değiştiren         | 0.323085       |
| Kişi-grup-şirket bazında yaratan            | 0.190352       |
| Hatanın tür bilgisi                         | 0.126425       |
| Hatanın görüldüğü işletim sistemi           | 0.049086       |
| Hatanın önceliğinin seviyesi                | 0.04682        |
| Hatanın hangi grupta gösterileceği          | 0.040754       |
| Hatanın görülme sıklığı                     | 0.006596       |
| Hatanın yorum olup olmadığı                 | 0.006101       |
| Hatanın yazılımsal-tasarımsal olup olmadığı | 0.004704       |
| Hatanın şirket bazlı olup olmadığı          | 0.000965       |

Listelenen bu 13 faktörün hata seviyelendirme üzerindeki etkilerine bakmak için bilgi kazanımı yöntemi kullanılacaktır. Bilgi kazanımı özellik seçiminde sınıf bilgisini belirleyen en değerli özelliklerin çıkarımını sağlamaktadır. Değerli özellikler sınıf bilgisi için ayırt edici özelliklerdir. Yukarıda belirtilen faktörler birden fazla sınıfta bulunabilmektedir. Bilgi Kazanımı farklı sınıflar içerisinde ayırt edici özelliği fazla olan özellikleri belirlemeye çalışır.

Bilgi Kazanımı bir özelliğin sınıf bilgisine ait belirsizliğinin, özellik bilgisi kullanılarak azaltılması olarak da tanımlanabilir. Belirsizlik ölçütü olarak rastsal değişkenlerde kullanılan entropi, bilgi kazanımı hesabında da kullanılmaktadır. Kısaca, bilgi kazanımı özelliğin bir veya birden fazla sınıf için var olup olmamasına göre entropideki azalımı ölçmektedir. Bilgi kazanımı yüksek olan özellikler sınıflandırma için değerli özellikler olarak alınır ve belirli bir eşik değerinin altında kalan özellikler elenerek özellik seçimi gerçekleştirilmiş olur.

Weka arayüzünde bilgi kazanımı kullanılarak 13 faktör için yapılan çalışma sonrasında çizelge 4.4 elde edilmiştir.

Yazılım risk paradigmlarında da bu konudan bahsedilmiştir. Tanımlama aşamasında belirli olmayan problemler ve riskler somut hale getirilir, tanımlanır ve ölçülebilir. Bu ölçüm sayesinde de risk analizi yapılabilir. Risk analizi yapmanın temel amacı risk karar tahmin bilgisinin çıkarılabilmesini sağlamaktır. Bu analizler proje yöneticilerine bir çok riskin ortaya çıkmadan gözlemlenmesine olanak sağlamaktadır. Ayrıca yapılan bu ön işleme aşaması yazılım risk yönetiminin şu alt adımlarına da paralellik göstermektedir[1].

- Risk yönetiminde risk kültürü belirlenmeli ve nelerin risk olarak kabul edilebileceği, yüksek risk alınıp yüksek yararların mı yoksa düşük risk alınıp düşük yararların mı elde edileceğinin kararı verilmeli.
- Risk seviyelendirmesinin yapılması ve zararın ölçülüp karşılığında elde edilecek yararın organizasyon ve proje için yeterli olup olmayacağına karar

Oluşturulan 8135hatanın belirlenen risk faktörlerine göre gruplanmış sonuçları ve etki değerleri listelenmiştir. Model oluşturmak için bu veri setine başvurulmuştur. Çalışma kapsamında destek vektör makineleri, karar ağaçları ve naıvebayes sınıflandırıcıları kullanılarak modeller ayrı ayrı oluşturulmuş ve elde edilen test sonuçları birbirleriyle kıyaslanmıştır.

#### 4.4 Kümeleme İşlemleri

Kümeleme işlemi risk faktörleri belirlendikten sonra yapılmıştır. Uzman görüşleri alınarak 4 proje için kümeleme işlemleri gerçekleştirilmiştir. Kümeleme işlemi için bilgisayar mühendisliği yaklaşımları kullanılmayarak uzman görüşleri baz alınmıştır. Ele alınan 4 projede kendi içinde ikiye ayrılarak kümelenebilir.

Kümeleme işleminin yapılması için birbirine benzer veya eşdeğer projeler kullanılmıştır. Kullanılan projeler benzer ürün gruplarına aittir ve benzer özelliklere sahiptirler.

#### 4.5 Sınıflandırma İşlemleri

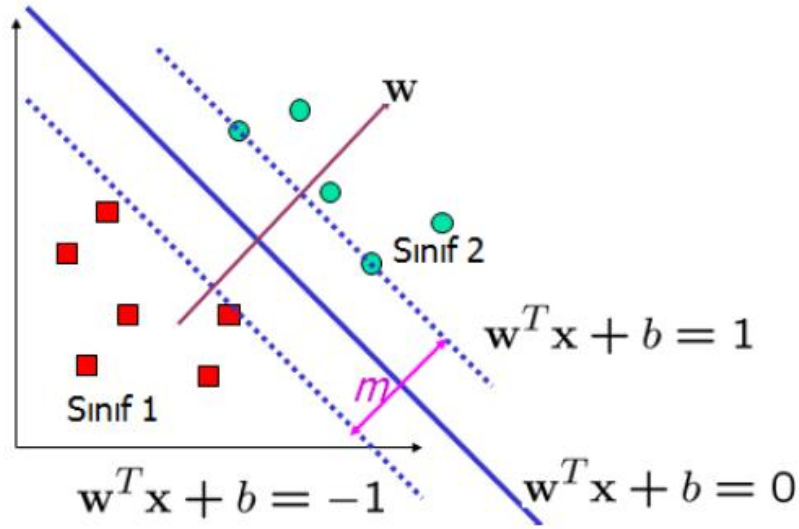
Yazılım risk yönetiminde risk tahmin modelleri oluştururken veya risk faktör kararı verilirken farklı makine öğrenmesi yaklaşımları kullanılarak pek çok çalışma gerçekleştirilmiştir. Yapay sinir ağları, bayes sınıflandırıcısı, karar ağaçları, k-en yakın komşuluk, destek vektör makineleri gibi farklı sınıflandırma yaklaşımları yazılım risk yönetiminde kullanılmıştır. Bu sınıflandırma yaklaşımlarını kullanan çalışmalarda sınıflandırma algoritmalarını kendi aralarında karşılaştırmıştır.

Yapılan çalışmalara bakıldığında genel olarak en iyi performans destek vektör makinelerinin verdiği gözlemlenmiştir. NaiveBayes sınıflandırıcısının performansının ise diğer sınıflara nazaran daha düşük olduğu gözlemlenmiştir. Yukarıdaki çalışmalar göz önünde bulundurularak tez kapsamında Destek Vektör Makineleri, Karar ağaçları ve NaiveBayes sınıflandırıcılarının kullanılmasına karar verilmiştir. Destek Vektör Makine yaklaşımıyla yapılan sınıflandırmaların diğer iki yöntemle göre daha iyi sonuçlar vermesi amaçlanmıştır[21][22][23][26][27].

##### 4.5.1 Destek Vektör Makineleri

Risk analizi için en çok kullanılan sınıflandırma yöntemlerinden biri destek vektör makinesidir. Destek vektör makineleri metin sınıflandırma, yüz tanıma, el yazısı tanıma ve buna benzer makine öğrenmesi algoritmalarının kullanıldığı alanlarda başarılı sonuçlar vermesinden dolayı çalışmada tercih edilen sınıflandırma yöntemlerden biridir.

Destek Vektör Makineleri karar düzlemlerinin ideal şekilde belirlenmesi prensibine dayanır. İdeal karar sınırı ayırdığı sınıflara ait verilere mümkün olduğu kadar karar düzleminde uzak olmalıdır.



Şekil 4.2 Destek Vektör Makineleri

Şekil 4.2'ye göre kesikli doğrular her sınıfa ait sınırları göstermektedir. Sınırlar üzerindeki örneklere destek vektör denilmekle birlikte karar sınırları sadece bu vektörler tarafından belirlenmektedir. İdeal bir karar düzlemi belirlemek için şekil 4.2'de gösterilen  $m$  değerinin maksimize edilmesi gerekmektedir.

$$m = \frac{2}{||w||} \quad (4.1)$$

Bir başka deyişle  $||w||$  değerinin minimize edilerek karar düzleminin ayırdığı sınıflara ait sınır değerlerine uzaklık maksimum yapılmaktadır. Sınıf etiketleri  $y_i \in \{1, -1\}$  olmak üzere tüm veri setindeki değerlerin  $\{x_1, x_2, x_3, \dots, x_n\}$  denklem 4.2'de belirlenmiş karar sınırı tarafından sınıflandırılması gerekir.

$$y_i(w^T x + b) \geq 1 \quad (4.2)$$

#### 4.5.2 Karar Ağaçları

Karar ağaçları örnek verileri bir ağaç yapısına benzetmektedir. Ağacın kökünden ağacın alt dallarına kadar sıralayarak verilerin sınıflandırılmasını sağlar. Ağaçtaki düğümler özellik, dallar özellik değer bilgisini ve yaprak düğümler ise sınıf etiketini belirtmektedir.

Karar ağaçlarının yaygın kullanımının nedeni ağaç yapılarının kurallarla ve sade bir şekilde tanımlanabilmesidir. Bu şekilde öğrenilen kurallar kolay bir şekilde aktarılmış

olur. Karar ağaçları kullandığı algoritma farklılıklarına göre en çok kullanılanlar ID3, ASSISTANT ve J48'dir.

Karar ağacı sınıflandırıcısıyla veriler aç gözlü yaklaşım (greedy) olan belirli kriterleri en iyi özelliğe dayalı bölünürler. Düğümler bölünme şekline göre çoklu veya ikili bölünmektedir. En iyi düğümler bölümlemesi için homojen sınıf dağılımına sahip düğümler tercih edilir. Homojenliği belirlemek için düğümün saf olmama değeri (impurity) hesaplanır. Saf olmama değerini ölçmek için gini indeksi, entropi gibi değerler kullanılır. Tüm veriler karar ağacında doğru şekilde sınıflandırılana kadar karar ağacı dallandırılır.

#### 4.5.3 Naive Bayes Yöntemi

Klasik Naive Bayes (KNB) uzun yıllardır farklı sınıflandırma problemlerinde yaygın olarak kullanılan makine öğrenmesi yaklaşımıdır. KNB bağımsızlık varsayımına dayanmaktadır ve özellikler sınıftan bağımsızdır.

Bir sınıflandırıcıya ait olasılık değeri,  $F_1$ 'den  $F_N$ 'e kadar koşullu özellik değişkenleri ve farklı  $C$  sınıf etiketleri kullanan bağımlı  $C$  değişkeni üzerine tanımlanmış koşullu modeldir.

$$p(C|F_1, F_2, \dots, F_N) \quad (4.3)$$

NB'de temel fikir özelliklerin ve sınıf bilgilerinin birlikte gerçekleşme olasılık değerlerinin kullanarak sınıf etiketi atamaktır. Bayes teoremiyle elde edilen verinin  $C$  sınıfına ait olma olasılığı denklem x.x'deki şekildedir.

$$p(C|F_1, F_2, \dots, F_N) = \frac{p(C) p(F_1, \dots, F_N)}{p(F_1, \dots, F_N)} \quad (4.4)$$

Etiket değeri bilinen bir sınıfta bir özelliğin ilgili sınıfta koşullu bulunma olasılığı diğer özelliklerin aynı sınıfta koşullu bulunma olasılığından bağımsızdır. Özelliklerin koşullu olasılık değerleri kullanılarak verinin ilgili sınıfa ait olma olasılığı hesaplanmıştır.

$$p(C, F_1, \dots, F_N) \propto p(C) \prod_{i=1}^N p(F_i|C) \quad (4.5)$$

Denklem 4.6’da gösterildiği üzere NB sınıflandırıcısı test verisine sınıf etiketi atarken her sınıf değeri için elde edilen verinin ilgili sınıf etiketine ait olma olasılıklarını kullanır ve bu değerlerden maksimum olanı sınıf etiketi olarak belirler.

$$\operatorname{argmax}_c p(C = c) \prod_{i=1}^n p(F_i = f_i | C = c) \quad (4.6)$$

## BÖLÜM 5

---

### BULGULAR

Çalışma çerçevesince dördüncü bölümde de anlatıldığı üzere ilk olarak sistem tasarımı gerçekleştirilmiştir. Kullanılacak uygulama için öncelikli olarak veriler elde edilmiştir. Verilerin elde edilebilmesi için yazılım firmasından 8 farklı proje için hata verileri toplanmıştır. Toplanan bu verilerin uygulamaya uygun hale getirilmesi için ön işleme aşaması yapılmıştır. Bu ön işleme aşamasında tüm projedeler e bakılmış. Birbirine benzer özellikteki projeler seçilerek faktörlere karar verilmiştir. Faktör kararından sonra deneyim bilgisinden faydalanarak kümeleme işlemleri de yapılmıştır. Kümelenen veriler aynı zamanda diğer projelerle de karşılaştırılmıştır.

Kullanılacak olan 8 proje 3 farklı şekilde ele alınmıştır. Bunlar da aşağıda belirtilmiştir.

- Tüm projeler ortak olarak ele alınmıştır. Bunun için 8 proje verileri ortak faktörler kullanılarak tek proje verisiymiş gibi işlenmiştir. 8135 hata verisi kullanılmıştır.
- İkinci olarak kümelenen veriler ele alınmıştır. Kümeleme işlemi için uzman görüşüne başvurulmuş ve 4 farklı proje ikiye bölünmüş ve kümelenebilir. Kümelenen bu projelerin sonuçlarına bakılmış ve genel olarak proje oranlarıyla da karşılaştırılmıştır. Kümelenen ilk iki projenin hata sayısı 1027 iken, diğer ikili proje grubunun toplam hata sayısı ise 3951'dir.
- Üçüncü olarak 8 farklı proje ayrı ayrı ele alınmıştır. İlk proje 511, ikinci proje 122, üçüncü proje 905, dördüncü proje 1948, beşinci proje 492, altıncı proje 3007, yedinci proje 944, sekizinci proje ise 206 hata içermektedir. Bu sonuçlarda kümelenmiş hata verileri ve tüm projelerin ortak olarak ele alındığı veri setiyle kıyaslanmıştır.

Verileri işleme aşamasında ise DVM,KA ve NB sınıflandırıcıları kullanılarak modeller oluşturulmuştur. Bu yöntemler sonucunda oluşacak modeller ayrı ayrı anlatılacak ve elde edilen bulgular listelenecektir.

### 5.1 Destek Vektör Makinesi Sınıflandırma Yöntemi Sonuçları

Risk analizi için en çok kullanılan sınıflandırma yöntemlerinden biri destek vektör makinesidir. Destek vektör makineleri metin sınıflandırma, yüz tanıma, el yazısı tanıma ve buna benzer makine öğrenmesi algoritmalarının kullanıldığı alanlarda başarılı sonuçlar vermesinden dolayı çalışmada tercih edilen sınıflandırma yöntemlerinden biridir.

Destek Vektör Makineleri karar düzlemlerinin ideal şekilde belirlenmesi prensibine dayanır. İdeal karar sınırı ayırdığı sınıflara ait verilere mümkün olduğu kadar karar düzleminden uzak olmalıdır.

Weka arayüzünde sınıflandırma fonksiyonlarından destek vektör makineleri kullanılmıştır. Bu işlem şu gruplara uygulanmıştır :

- Tüm projelerin ortak olarak ele alındığı durumlar için
- Benzer projelerin ortak olarak ele alındığı durumlar için(4 proje 2şerli olarak gruplanmıştır)
- Tüm projelerin ayrı olarak ele alındığı durumlar için( 8 proje ayrı olarak ele alınmıştır)

Tüm projelerin ortak olarak ele alındığı durumlarda 8135 hatanın 6382 adedi doğru olarak belirlenebilmiştir. 1753 tanesi de farklı etiketler almıştır. Modelin doğruluk oranı ise %78,45'dir. Test sonuçlarını doğrulama yöntemi 10 kere çapraz doğrulama olarak uygulanmıştır. Çizelge 5.1'den de incelenebilir.

Tüm projelerin karışıklık matrisi ise çizelge 5.2'de gösterilmiştir. Karışıklık matrisinde veriler 6 farklı sınıfta toparlanmıştır. Bu verilerin gerçek değerleri ve modelden elde edilen veriler karşılaştırıldığında çizelge 5.2 deki gibi doğru ve hatalı sınıflandırılan veriler görülebilmektedir.

Çizelge 5.1 DVM Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 6382         | %78,45           |
| Yanlış Sınıflandırılan Örnek | 1753         | %21,54           |
| Toplam Örnek Sayısı          | 8135         |                  |

Tabloya bakıldığında çok kritik olarak değerlendirilen hataların sadece 354 tanesi doğru sınıflandırıldığı, 690 adedinin çok kritik olduğu halde kritik olarak sınıflandırıldığı görülmektedir. Buradaki problem hataların ya kritik olduğu halde çok kritik olarak girilmesi veya kullanılan faktörlerin hatalı veya bilinçsizce girildiği görülmektedir. Bu durumda firma çok kritik olarak girdiği hataları gözden geçirmeli, kritiklik ve çok kritiklik arasındaki farkları iyi analiz etmelidir. Aynı şekilde ihmal edilebilir hatalara baktığımızda da yüzde olarak kötü sonuçlar görülmektedir. İhmal edilen hataların da üzerinden geçilmelidir. Bunun dışında kritik ve orta derece olarak girilen hataların oluşturulan modelle yüksek oranda uyduğu gözlemlenmektedir. Koşullu ve hata değil olarak girilen hatalar ise yüksek oranda tahmin edilmiş ve modelle uyumluluğu gözlemlenmiştir.

Çizelge 5.2 DVM Tüm Projeler için Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f  |
|----------------------|-----|------|-----|-----|-----|----|
| a =Çok kritik        | 354 | 690  | 16  | 1   | 3   | 0  |
| b = Kritik           | 193 | 5058 | 50  | 18  | 27  | 1  |
| c = Koşullu Hata     | 9   | 138  | 364 | 17  | 48  | 5  |
| d = Hata değil       | 1   | 41   | 33  | 162 | 98  | 22 |
| e = Orta derece      | 8   | 77   | 80  | 75  | 410 | 25 |
| f = İhmal edilebilir | 1   | 0    | 7   | 27  | 40  | 34 |

Benzer projelerin ortak olarak ele alındığı durumlarda ise çizelge 5.3 ve çizelge 5.4'den yararlanılabilir.

Çizelge 5.3 DVM İlk Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 906          | %88.21           |
| Yanlış Sınıflandırılan Örnek | 121          | %11.78           |
| Toplam Örnek Sayısı          | 1027         |                  |

Çizelge 5.4 DVM İkinci Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 3098         | %78.41           |
| Yanlış Sınıflandırılan Örnek | 853          | %21.58           |
| Toplam Örnek Sayısı          | 3951         |                  |

Çizelge 5.3'e bakıldığında ortak olan iki proje için toplam 1027 hata girilmiş ve başarı oranı %88 olarak görülmüştür. Doğru sınıflandırılan örnek sayısı ise 906'dır. Çizelge 5.4'e baktığımızda ise 3951 hata sayısına sahip projelerin 3098 verisi doğru olarak sınıflandırılmıştır. Başarı oranı ise %78,41'dir.

Çizelge 5.5 DVM Birinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b   | c | d  | e  | f |
|----------------------|----|-----|---|----|----|---|
| a =Çok kritik        | 9  | 66  | 0 | 1  | 0  | 0 |
| b = Kritik           | 18 | 863 | 0 | 6  | 2  | 0 |
| c = Koşullu Hata     | 0  | 1   | 0 | 0  | 0  | 1 |
| d = Hata değil       | 1  | 4   | 0 | 22 | 4  | 0 |
| e = Orta derece      | 0  | 4   | 0 | 3  | 10 | 4 |
| f = İhmal edilebilir | 0  | 0   | 0 | 0  | 6  | 2 |

Çizelge 5.5 Birinci Ürün Grubu Karışıklık Matrisi ve Çizelge 5.6 İkinci Ürün grubu karışıklık matrisine bakıldığında ise projelerin genel olarak ele alındığında görülen

durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir hataların daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

Çizelge 5.6 DVM İkinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b    | c   | d  | e   | f |
|----------------------|----|------|-----|----|-----|---|
| a =Çok kritik        | 73 | 397  | 3   | 0  | 0   | 0 |
| b = Kritik           | 71 | 2620 | 18  | 9  | 14  | 1 |
| c = Koşullu Hata     | 6  | 61   | 244 | 5  | 33  | 3 |
| d = Hata değil       | 3  | 21   | 28  | 53 | 36  | 4 |
| e = Orta derece      | 0  | 26   | 69  | 31 | 108 | 2 |
| f = İhmal edilebilir | 0  | 0    | 4   | 2  | 6   | 0 |

Çizelge 5.7 DVM Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| Projeler | Doğru Sınıflandırılan | Toplam Hata Sayısı | Başarı Oranı(%) |
|----------|-----------------------|--------------------|-----------------|
| 1.Proje  | 358                   | 511                | %70.05          |
| 2.Proje  | 109                   | 122                | %89.34          |
| 3.Proje  | 796                   | 905                | %87.95          |
| 4.Proje  | 1448                  | 1948               | %74.33          |
| 5.Proje  | 397                   | 492                | %80.69          |
| 6.Proje  | 2412                  | 3007               | %80,21          |
| 7.Proje  | 691                   | 944                | %73.19          |
| 8.Proje  | 171                   | 206                | %83.0           |

8 proje tek başına ele alındığında ise başarı oranları çizelge 5.7’de başarı oranlarındaki gibi olmaktadır. Başarı oranları %89.34 ile %73.19 arasında değişkenlik göstermektedir. Ayrıntılar için Ek A’ya bakılabilir.

8 proje için karışıklık matrislerine ise Ek-A’dan bakılabilir. Projelerin genel olarak ele alındığında görülen durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir

hataların daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

## 5.2 Karar Ağaçları Sınıflandırma Yöntemi Sonuçları

Karar ağaçları örnek verileri bir ağaç yapısına benzemektedir. Ağacın kökünden ağacın alt dallarına kadar sıralayarak verilerin sınıflandırılmasını sağlar. Ağaçtaki düğümler özellik, dallar özellik değer bilgisini ve yaprak düğümler ise sınıf etiketini belirtmektedir. Karar ağaçlarının yaygın kullanımının nedeni ağaç yapılarının kurallarla ve sade bir şekilde tanımlanabilmesidir. Bu şekilde öğrenilen kurallar kolay bir şekilde aktarılmış olur. Karar ağaçları kullandığı algoritma farklılıklarına göre en çok kullanılanları ID3, ASSISTANT ve J48'dir.

Karar ağacı sınıflandırıcısıyla veriler aç gözlü yaklaşım (greedy) olan belirli kriterleri en iyileyen özelliğe dayalı bölünürler. Düğümler bölünme şekline göre çoklu veya ikili bölünmektedir. En iyi düğümler bölümlenmesi için homojen sınıf dağılımına sahip düğümler tercih edilir. Homojenliği belirlemek için düğümün saf olmama değeri (impurity) hesaplanır. Saf olmama değerini ölçmek için gini indeksi, entropi gibi değerler kullanılır. Tüm veriler karar ağacında doğru şekilde sınıflandırılana kadar karar ağacı dallandırılır.

Weka arayüzünde sınıflandırma fonksiyonlarından Karar ağacı yaklaşımı olarak J48 karar ağacı sınıflandırıcı kullanılmıştır. 8135 hatanın 6231 adedini doğru olarak belirleyebilmiştir. 1904 tanesini ise farklı etiketlemiştir. Modelin doğruluk oranı ise %76,59'dur. Test sonuçlarının doğrulama yöntemi 10 kere çapraz doğrulama olarak uygulanmıştır. Çizelge 5.8'de de bu durum gözlemlenmektedir.

Çizelge 5.8 KA Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 6231         | %76,595          |
| Yanlış Sınıflandırılan Örnek | 1904         | %23,405          |
| Toplam Örnek Sayısı          | 8135         | -                |

Tüm projelerin karışıklık matrisi ise tablo 5.9’da gösterilmiştir. Karışıklık matrisinde veriler 6 farklı sınıfta toparlanmıştır. Bu verilerin gerçek değerleri ve modelden elde edilen veriler karşılaştırıldığında tablo x.x deki gibi doğru ve hatalı sınıflandırılan veriler görülebilmektedir. Burada da DVM’dekine paralel sonuçlar gözlemlenmektedir. Yani hataların ya kritik olduğu halde çok kritik olarak girilmesi veya kullanılan faktörlerin hatalı veya bilinçsizce girildiği görülmektedir. Bu durumda firma çok kritik olarak girdiği hataları gözden geçirmeli, kritiklik ve çok kritiklik arasındaki farkları iyi analiz etmelidir. Aynı şekilde ihmal edilebilir hatalara baktığımızda da yüzde olarak kötü sonuçlar görülmektedir. İhmal edilen hataların da üzerinden geçilmelidir. Bunun dışında kritik ve orta derece olarak girilen hataların oluşturulan modelle yüksek oranda uyduğu gözlemlenmektedir.

Çizelge 5.9 KA Tüm Projeler için Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f  |
|----------------------|-----|------|-----|-----|-----|----|
| a =Çok kritik        | 246 | 881  | 5   | 0   | 2   | 0  |
| b = Kritik           | 70  | 5184 | 33  | 17  | 40  | 3  |
| c = Koşullu Hata     | 3   | 230  | 302 | 16  | 27  | 3  |
| d = Hata değil       | 9   | 75   | 39  | 137 | 85  | 12 |
| e = Orta derece      | 6   | 120  | 127 | 70  | 334 | 18 |
| f = İhmal edilebilir | 0   | 12   | 9   | 26  | 34  | 28 |

Benzer projelerin ortak olarak ele alındığı durumlarda ise çizelge 5.10 ve çizelge 5.11’den yararlanılabilir.

Çizelge 5.10 KA Tüm Projeler için Karışıklık Matrisi

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 889          | %86.5628         |
| Yanlış Sınıflandırılan Örnek | 138          | %13.4372         |
| Toplam Örnek Sayısı          | 1027         |                  |

Çizelge 5.11 KA İlk Ürün Gurubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 3038         | %76.8919         |
| Yanlış Sınıflandırılan Örnek | 913          | %23.1081         |
| Toplam Örnek Sayısı          | 3951         |                  |

Çizelge 5.10'a bakıldığında ortak olan iki proje için toplam 1027 hata girilmiş ve başarı oranı %86 olarak görülmüştür. Doğru sınıflandırılan örnek sayısı ise 889'dur. Çizelge 5.11'e baktığımızda ise 3951 hata sayısına sahip projelerin 3038 verisi doğru olarak sınıflandırılmıştır. Başarı oranı ise %76,89'dır.

Çizelge 5.12 ve çizelge 5.13'e bakıldığında ise projelerin genel olarak ele alındığında görülen durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir hataların daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

Çizelge 5.12 KA Birinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 76  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 889 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 2   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 31  | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 8   | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

Çizelge 5.13 KA İkinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a | b    | c   | d  | e  | f |
|----------------------|---|------|-----|----|----|---|
| a =Çok kritik        | 0 | 471  | 2   | 0  | 0  | 0 |
| b = Kritik           | 0 | 2677 | 20  | 9  | 27 | 0 |
| c = Koşullu Hata     | 0 | 75   | 265 | 2  | 10 | 0 |
| d = Hata değil       | 0 | 49   | 37  | 39 | 20 | 0 |
| e = Orta derece      | 0 | 57   | 101 | 21 | 57 | 0 |
| f = İhmal edilebilir | 0 | 3    | 8   | 0  | 1  | 0 |

8 proje tek başına ele alındığında ise başarı oranları çizelge 5.14 başarı oranlarındaki gibi olmaktadır. Başarı oranları %86.29 ile %68.68 arasında değişkenlik göstermektedir. Ayrıntılar için Ek A'ya bakılabilir.

8 proje için karışıklık matrislerine ise Ek-A'dan bakılabilir. Projelerin genel olarak ele alındığında görülen durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir hataların daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

Çizelge 5.14 KA Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| Projeler | Doğru Sınıflandırılan | Toplam Hata Sayısı | Başarı Oranı(%) |
|----------|-----------------------|--------------------|-----------------|
| 1.Proje  | 351                   | 511                | %68.68          |
| 2.Proje  | 109                   | 122                | %89             |
| 3.Proje  | 781                   | 905                | %86.29          |
| 4.Proje  | 1391                  | 1948               | %71.40          |
| 5.Proje  | 396                   | 492                | %80.48          |
| 6.Proje  | 2383                  | 3007               | %79.24          |
| 7.Proje  | 672                   | 944                | %71.18          |
| 8.Proje  | 172                   | 206                | %83.49          |

### 5.3 Naive Bayes Sınıflandırma Yöntemi Sonuçları

Klasik NaiveBayes(KNB) uzun yıllardır farklı sınıflandırma problemlerinde yaygın olarak kullanılan makine öğrenmesi yaklaşımıdır. KNB bağımsızlık varsayımına dayanmaktadır ve özellikler sınıftan bağımsızdır. Bir sınıflandırıcıya ait olasılık modeli, F1'denFN'e kadar koşullu özellik değişkenleri ve farklı C sınıf etiketleri kullanan bağımlı C değişkeni üzerine tanımlanmış koşullu modeldir.

NaiveBayes Yöntemi kullanıp Weka arayüzünde sınıflandırma fonksiyonlarından yararlanılarak risk modeli oluşturulmuştur. 8135 problem etki analiz verisinin 5914 adedini doğru olarak modelleyebilmiştir. 2221 tanesini ise farklı etiketlemiştir. Modelin doğruluk oranı ise %72,69dur. Test sonuçlarının doğrulama yöntemi 10 kere çapraz doğrulama olarak uygulanmıştır. Çizelge 5.15'den sonuçlar görülebilir.

Tüm projelerin karışıklık matrisi ise Çizelge 5.16'da gösterilmiştir. Karışıklık matrisinde veriler 6 farklı sınıfta toparlanmıştır. Bu verilerin gerçek değerleri ve modelden elde edilen veriler karşılaştırıldığında tablo x.x deki gibi doğru ve hatalı sınıflandırılan veriler görülebilmektedir. Burada da DVM'dekine ve KA'daki sonuçlara paralel sonuçlar gözlemlenmektedir. Yani hataların ya kritik olduğu halde çok kritik olarak girilmesi veya kullanılan faktörlerin hatalı veya bilinçsizce girildiği görülmektedir. Bu durumda firma çok kritik olarak girdiği hataları gözden geçirmeli, kritiklik ve çok kritiklik arasındaki farkları iyi analiz etmelidir. Aynı şekilde ihmal edilebilir hatalara baktığımızda da yüzde olarak kötü sonuçlar görülmektedir. İhmal edilen hataların da üzerinden geçilmelidir. Bunun dışında kritik ve orta derece olarak girilen hataların oluşturulan modelle yüksek oranda uyduğu gözlemlenmektedir.

Çizelge 5.15 NB Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 5914         | %72.69           |
| Yanlış Sınıflandırılan Örnek | 2221         | %27.30           |
| Toplam Örnek Sayısı          | 8135         |                  |

Çizelge 5.16 NB Tüm Projeler için Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f |
|----------------------|-----|------|-----|-----|-----|---|
| a =Çok kritik        | 353 | 685  | 13  | 6   | 5   | 2 |
| b = Kritik           | 308 | 4740 | 94  | 98  | 99  | 8 |
| c = Koşullu Hata     | 22  | 147  | 266 | 29  | 117 | 0 |
| d = Hata değil       | 12  | 23   | 53  | 135 | 132 | 2 |
| e = Orta derece      | 23  | 51   | 106 | 73  | 420 | 2 |
| f = İhmal edilebilir | 2   | 0    | 11  | 28  | 67  | 0 |

Çizelge 5.17 NB İlk Ürün Gurubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 885          | %86,17           |
| Yanlış Sınıflandırılan Örnek | 142          | %13,82           |
| Toplam Örnek Sayısı          | 1027         |                  |

Çizelge 5.18 NB İkinci Ürün Grubu için Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 2918         | %73.85           |
| Yanlış Sınıflandırılan Örnek | 1033         | %26.14           |
| Toplam Örnek Sayısı          | 3951         |                  |

Benzer projelerin ortak olarak ele alındığı durumlarda ise çizelge 5.17 ve çizelge 5.18'den yararlanılabilir.

Çizelge 5.17'ye bakıldığında ortak olan iki proje için toplam 1027 hata girilmiş ve başarı oranı %86 olarak görülmüştür. Doğru sınıflandırılan örnek sayısı ise 885'dir. Çizelge 5.18'e baktığımızda ise 3951 hata sayısına sahip projelerin 2918 verisi doğru olarak sınıflandırılmıştır. Başarı oranı ise %73,85'dir.

Çizelge 5.19 ve çizelge 5.20'ye bakıldığında ise projelerin genel olarak ele alındığında görülen durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir hataların

daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

Çizelge 5.19 NB Birinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b   | c | d | e | f |
|----------------------|----|-----|---|---|---|---|
| a =Çok kritik        | 5  | 71  | 0 | 0 | 0 | 0 |
| b = Kritik           | 11 | 877 | 0 | 1 | 0 | 0 |
| c = Koşullu Hata     | 0  | 2   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0  | 28  | 0 | 3 | 0 | 0 |
| e = Orta derece      | 0  | 21  | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0  | 8   | 0 | 0 | 0 | 0 |

Çizelge 5.20 NB İkinci Ürün Grubu Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d  | e   | f |
|----------------------|-----|------|-----|----|-----|---|
| a =Çok kritik        | 157 | 310  | 4   | 1  | 1   | 0 |
| b = Kritik           | 202 | 2360 | 69  | 37 | 65  | 0 |
| c = Koşullu Hata     | 16  | 33   | 248 | 1  | 54  | 0 |
| d = Hata değil       | 3   | 14   | 42  | 33 | 53  | 0 |
| e = Orta derece      | 0   | 10   | 84  | 22 | 120 | 0 |
| f = İhmal edilebilir | 0   | 0    | 4   | 0  | 8   | 0 |

8 proje tek başına ele alındığında ise başarı oranları çizelge 5.21'deki gibi olmaktadır. Başarı oranları %89.34 ile %68.49 arasında değişkenlik göstermektedir. Ayrıntılar için Ek A'ya bakılabilir.

8 proje için karışıklık matrislerine ise Ek-A'dan bakılabilir. Projelerin genel olarak ele alındığında görülen durum gözlemlenmektedir. Burada da çok kritik ve ihmal edilebilir hataların daha kötü sınıflandırıldığı, diğer sınıflandırmaların ise daha başarılı olduğu gözlemlenmektedir.

Çizelge 5.21 NB Proje Bazlı Sınıflandırılan Örnek Sayıları ve Yüzdeleri

| Projeler | Doğru Sınıflandırılan | Toplam Hata Sayısı | Başarı Oranı(%) |
|----------|-----------------------|--------------------|-----------------|
| 1.Proje  | 350                   | 511                | %68.49          |
| 2.Proje  | 109                   | 122                | %89.34          |
| 3.Proje  | 784                   | 905                | %86.62          |
| 4.Proje  | 1394                  | 1948               | %71.56          |
| 5.Proje  | 390                   | 492                | %79.26          |
| 6.Proje  | 2245                  | 3007               | %74.65          |
| 7.Proje  | 698                   | 944                | %73.94          |
| 8.Proje  | 172                   | 206                | %83.49          |

#### 5.4 Test Sonuçlarının Karşılaştırılması

Sonuçları daha net bir şekilde anlayabilmek ve gerekli karşılaştırmaları yapabilmek için çizelge 5.22'den yararlanılabilir. Belirtilen tabloda toplam hata sayıları, NB sınıflandırma yüzdesi, KA sınıflandırma yüzdesi, DVM sınıflandırma yüzdeleri listelenmiştir. Listelenen bu değerler tüm projelerin ortak alınma durumunu, benzer projelerin ortak alınma durumunu ve projelerin tek başlarına olan durumlarını göstermektedir.

Tablo x.x'den de görüleceği üzere DVM sınıflandırma yüzdesi sekizinci proje dışında tüm projelerde NB ve KA'ya göre daha yüksek çıkmıştır. KA'da tüm projeler, ürün grubu I, ürün grubu II, ilk proje, beşinci proje, altıncı projede NB'ye göre daha başarılı sonuçlar ortaya koymuştur. NB sadece birkaç projede KA'ya göre üstünlük sağlasada genel olarak KA daha başarılı sonuçlar ortaya koymuştur.

Çizelge 5.22 Çalışmanın Başarı Oranları ve Yüzdeleri

| Projeler        | Toplam Hata Sayısı | NB Doğru Sınıf. (%) | KA Doğru Sınıf. (%) | DVM Doğru Sınıf. (%) |
|-----------------|--------------------|---------------------|---------------------|----------------------|
| Tüm Projeler    | 8135               | %72,69              | %76,59              | %78,45               |
| Ürün Grubu I    | 1027               | %86,17              | %86,56              | %88,21               |
| Ürün Grubu II   | 3951               | %73,85              | %76,89              | %78,41               |
| İlk Proje       | 511                | %68,49              | %68,68              | %70,05               |
| İkinci Proje    | 122                | %89,34              | %89                 | %89,34               |
| Üçüncü Proje    | 905                | %86,62              | %86,29              | %87,95               |
| Dördüncü Proje  | 1948               | %71,56              | %71,40              | %74,33               |
| Beşinci Proje   | 492                | %79,26              | %80,48              | %80,69               |
| Altıncı Proje   | 3007               | %74,65              | %79,24              | %80,21               |
| Yedinci Proje   | 944                | %73,94              | %71,18              | %73,19               |
| Sekizinci Proje | 206                | %83,49              | %83,49              | %83,0                |

Bulunan sonuçlarda görüldüğü üzere SVM ile sınıflandırılan verilerde doğru sınıflandırma oranı karar ağaçlarından ve NaiveBayes sınıflandırıcılarından daha iyi sonuçlar verdiği gözlemlenmiştir. Tüm projelere bakıldığında DVM %78,45 oranında başarı gösterirken, KA %76,59, NB ise %72,69 oranında başarılı sınıflandırma yapmıştır.

İlgili çalışmalar kısmına bakıldığında da belirtilen çalışmalarla tez kapsamında bulunan sonuçlar paralellik göstermektedir. Sun Yat Sen Üniversitesinden yayınlanan makalede yazılım projelerinin risk yönetiminin modellenmesi sağlanmıştı. Rastgele veri seti ikiye bölünerek 100 örnek(anket sonucu) eğitim verisi olarak kullanılırken 20 örnek test amaçlı kullanılıp bu örnekler NN ve SVM sınıflandırıcılarıyla modellenmişti. NN elde edilen sonuçlar projelerin gerçek sonuçlarıyla kıyaslandığında %70 oranında bir doğruluk oranı elde edilmektedir. DVM kullanarak elde edilen başarı oranı ise %80'nin üzerindedir. Sun Yat Sen Üniversitesinin yaptığı diğer çalışmada da sonuçlar DVM ile oluşturulan modelin başarı oranı 20 test projesi için %85, ANN için ise %75 oranında başarı sağlamıştı[21][22].

UYMS 2012’de yayınlanan Yıldız Teknik Üniversitesinde M.ÖzgürCingiz ve Oya Kalıpsız ile yaptığım makale çalışma çerçevesince TurkcellICT’den elde edilen, etki analizi için kullanılan verilere göre DVM modelin doğruluk oranı ise %96,61’dir. J48 modelin doğruluk oranı ise %91,14’dür.NBmodelinin doğruluk oranı ise %91,66 dır. UYMK 2012’de yayınlanan Yıldız Teknik Üniversitesinde M.ÖzgürCingiz ve Oya Kalıpsız ile yaptığım diğer bir makale çalışmada iseTurkcell Teknolojide gerçekleştirilen projelerdeki problem etki veri setleri kullanılmıştır. Çok Katmanlı Yapay SinirAğları’nı % 96 ile Destek Vektör Makineleri takip etmiştir. F-ölçeğine göre başarısı en düşük olan sınıflandırıcılar NaïveBayes ve k-En Yakın Komşuluk sınıflandırıcılarıdır. Elde edilen veriler ilgili çalışmalarla paralellik göstermektedir. DVM büyük veri setlerinde bile diğer yöntemlere göre daha iyi sonuçlar verdiği gözlemlenmiştir[26][27].

Kümelemenmiş verilere baktığımızda iseürün grubu I ikinci ve üçüncü projelerden, ürün grubu II ise altıncı ve yedinci projelerden oluşmuştur. Ürün grubu I’in başarı oranına baktığımızda DVM için %88,21 iken genel proje başarı oranına baktığımızda %78,45 olduğu görülmektedir. Ürün grubu II’ye baktığımızda ise DVM için %78,41 oranında başarılı sınıflandırma yaptığı gözlemlenmiştir. Tüm projeler için DVM sınıflandırıcısının oranına yakın değerlerde olduğu gözlemlenmiştir. Bu durumda benzer projelerin kendi içinde değerlendirilmesi başarı oranlarını yükseltecektir.

Projelerin DVM, KA ve NB sınıflandırıcılar için karışıklık matrislerine bakıldığında ise yönteme göre çok farklı olmadığı gözlemlenmiştir. Çok kritik olarak ve ihmal edilebilecek olan hataların diğer sınıflara göre kötü sonuçlar ortaya koyduğu görülmüştür. Buradaki problem hataların ya kritik olduğu halde çok kritik olarak girilmesi veya kullanılan faktörlerin hatalı veya bilinçsizce girildiği görülmektedir. Bu durumda firma çok kritik olarak girdiği hataları gözden geçirmeli, kritiklik ve çok kritiklik arasındaki farkları iyi analiz etmelidir. Bunun dışında kritik ve orta derece olarak girilen hataların oluşturulan modelle yüksek oranda uyduğu gözlemlenmektedir. Koşullu ve hata değil olarak girilen hatalar ise yüksek oranda tahmin edilmiş ve modelle uyumluluğu gözlemlenmiştir.

### SONUÇ VE ÖNERİLER

Tez çalışması kapsamında temel olarak üç sonuca ulaşılmıştır. Bunlardan ilki sınıflandırıcı ve model performansları ve başarı oranlarıdır. Yapılan çalışmaya bakıldığında SVM yöntemi KA ve NB sınıflandırma yöntemlerinden daha başarılı olmuştur. Proje bazlı baktığımızda NB bazı projelerde KA'dan daha iyi olsa da, KA genel olarak daha yüksek başarı oranları sergilemektedir.

İkinci çıkarım ise kümeleme işlemleri sonucu elde edilen verilerin başarı oranlarının daha yüksek olduğudur. Benzer ürün gruplarının sınıflandırılması sonucu elde edilen veriler gerçeği daha iyi yansıtmakta ve başarı oranlarını daha da yükseltmektedir.

Üçüncü sonuç ise sınıflandırma işleminin yapıldığı hata türlerinden “çok kritik” hataların %50'sinden çoğunun hatalı değerlendirildiği gözlemlenmiştir. Yanlış değerlendirilen hatalar genellikle “kritik” grubunda çıktığı gözlemlenmiştir. Buradaki problem hataların ya kritik olduğu halde çok kritik olarak girilmesi veya kullanılan faktörlerin hatalı veya bilinçsizce girildiği görülmektedir. Bu durumda firma çok kritik olarak girdiği hataları gözden geçirmeli, kritiklik ve çok kritiklik arasındaki farkları iyi analiz etmelidir.

Yayınladığım makalelere bakıldığında veri setinin az olması ve verinin homojen dağılmamasından kaynaklı sıkıntılar bu tez çalışması kapsamında ortadan kaldırılmıştır. Yeterli sayıda veri setleri kullanılmış ve kullanılan sınıflandırıcı sonuçları da önceki makale çalışmalarını ve ilgili çalışmalar kısmında bahsedilen çalışmaları doğrulamaktadır.

Yapılan çalışmayı geliřtirmek adına faktör kararları, kümeleme işlemleri ve sınıflandırma yöntemleri ile ilgili ek çalışmaların yapılması düşünülmektedir. Sınıflandırma yöntemleri arttırılarak faktör kararlarının ne yönde deęiřeceęi, projeye etki eden faktörlerin sıralamasında deęiřiklik olup olmayacağına ilerleyen zamanlarda yapılacak makale çalışmalarında yer verilecektir.

Kümeleme işlemleri tecrübeye dayalı olarak uzman görüşleri alınarak yapılmıştır. Uzman görüşleri ile veri setlerinin birbiriyle kıyaslanması sonucu ortaya çıkacak sonuçlar karşılaştırılabilir ve uzman görüşlerinin doğruluęu verilerle de doğrulanması ile ilgili çalışmalarında yapılması düşünülmektedir. Ayrıca yapay sinir aęlarının da başarı oranları göz önüne alındığında destek vektör makineleri ile detaylı bir karşılaştırılmasının yapılması da düşünülmektedir.

## KAYNAKLAR

---

- [1] Pandian, C.R.,(2007), Applied Software Risk Management A Guide for Software Project Managers,Auerbach Publications,ABD
- [2] Higuera,R.P., Gluch,D.,Dorofee,A.J. ve Murphy,R.L.,(1994), "An introduction to Team Risk Management",Software Engineering Institute, Pennslyvania
- [3] Pressman, R.S.,(2001),Software Engineering A Practitioner's Approach,5.baskı, McGrawHill,New York
- [4] Software Engineering Institute, [www.sei.cmu.edu/programs/sepm/risk/](http://www.sei.cmu.edu/programs/sepm/risk/), 20 Ekim 2012
- [5] DeMarco,T. ve Lister T.,(2003), "Risk Management During Requirements",Software IEEE,20(05):99-101
- [6] Duncan,W.R.,(1996), "A Guide to the Project Management Body of Knowledge. NewtownSquare",Project Management Institue , 8
- [7] Futrell,R.T.,Shafer,D.F. ve Safer,L.I.,(2002),Quality Software Project Management,Prentice Hall PTR,ABD
- [8] Boehm,B.W.,(1991),"Software RiskManagement:Principles and Practices",Software IEEE,8(1):32-41
- [9] Jalote,P.,(2002),Software Project management in Practice,Addison Wesley,Boston
- [10] Hall,E.M.,(1998), Managing Risk: Methods for Software Systems Development,Addison-Wesley,ABD
- [11] Software Engineering Institute(SEI), <http://www.sei.cmu.edu/reports/93tr006.pdf>, 27 Ocak 2012
- [12] Wysocki,R.K.,(2006),Effective\_Software\_Project\_Management,Wiley Publishing,ABD
- [13] Futrell,R.T.,Shafer,D.F. ve Safer,L.I.,(2002),Quality Software Project Management,Prentice Hall PTR,ABD
- [14] Galorath,D.D. ve Evans,M.W.,(2006),Software Sizing Estimation and Risk Management,Auerbach Publications,ABD
- [15] Chapman,C. ve Ward,S.,(2009),Project Risk Management Process,Techniques and Insights,ikinci basım, John Wiley&Sons Ltd, İngiltere

- [16] Mcmanus,J.,(2004),Risk Management in Software Development Projects,Elsevier Butterworth-Heinemann,Oxford
- [17] Kitchenham,B. ve Linkman,S.,(1997),"Estimates,Uncertainty,and Risk",Software IEEE,14(3):69-74
- [18] Kwan,T.W. ve Leung,H.K.N.,(2011),"A Risk Management Methodologyfor Project Risk Dependencies",IEEE TRANSACTIONS ON SOFTWARE ENGINEERING,14(5)
- [19] Benaroch,M. ve Goldstein,J.,(2009),"An Integrative Economic Optimization Approachto Systems Development Risk Management",IEEE TRANSACTIONS ON SOFTWARE ENGINEERING,35(5)
- [20] Ray,M. ve Mohapatra,D.P.,(2013),"Risk analysis: a guiding force in the improvement of testing",Software IET,7(1):29-46
- [21] Hu,Y.,Huang,J.,Chen,J. ve Liu,M.,(2007),"Software Project Risk Management Modeling with Neural Network and Support Vector Machine Approaches",Natural Computation, 2007.ICNC,Haikou
- [22] Hu,Y.,Zhang,X.,Sun,X. ve Liu,M.(2009),"An Intelligent Model for Software Project Risk Prediction",Information Management, Innovation Management and Industrial Engineering, 2009 International Conference,Xi'an
- [23] Islam M.J.,Wu,Q.M.J. ve M.Ahmadi,(2007),"Investigating the Performance of Naive- Bayes Classifiers and K- Nearest Neighbor Classifiers",Convergence Information Technology, 2007. International Conference,Gyeongju
- [24] Tang,A. ve Wang,R.,(2010),"Software project risk assessment model based on fuzzy theory",Computer and Communication Technologies in Agriculture Engineering (CCTAE), 2010 International Conference,Chengdu
- [25] Bragina,T. ve Tabunshcyk,G.,(2011),"Fuzzy model for the software projects design risk analysis",CAD Systems in Microelectronics (CADSM), 2011 11th International Conference The Experience of Designing and Application,Polyana-Svalyava
- [26] Unudulmaz,A.,Cingiz,M.Ö. ve Kalipsiz,O.,(2012),"Yazılım Projelerindeki Problem Etkilerinin Yazılım Mimarisi ile İlişkilendirilmesi",Ulusal Yazılım Mimarisi Konferansı, İzmir
- [27] Cingiz,M.Ö.,Unudulmaz,A. ve Kalipsiz,O.,(2012),"Farklı Kestirim Yaklaşımları Kullanılarak Yazılım Projelerindeki Problem Etkilerinin Risk Yönetiminin Gerçeklenmesi",Ulusal Yazılım Mühendisliği Sempozyumu, Ankara

---

## VERİ KÜMESİNDEN ÇIKARILAN GENEL SONUÇLAR

Bu kısım altında projede işlenmiş olan tüm veri setlerine ait tablolar listelenecektir. 3 başlık altından tüm sonuçlara ulaşılabilinmektedir. Bu ana başlıklar altta listelenmiştir.

- Karar Ağaçları
- Destek Vektör Makineleri
- NaiveBayes

Her grupta kendi içinde 3'e ayrılmaktadır. Bu alt başlıklarda aşağıda listelenmiştir.

- Tüm projelerin sonuçları
- Aynı özelliklere sahip projelerin sonuçları (2 proje içerir)
- Proje bazlı oranlar (8 proje içerir)

Her bir madde 3 tablodan oluşmaktadır. Bu tablolarla aşağıda belirtilmiştir.

- Özet Tablosu
- Sınıf bazında doğruluk oranları
- Karışıklık Matrisi

Özet tablosunda o grupta yer alan doğru ve yanlış sınıflandırılan örnekler ve yüzdeleri yer almaktadır. Tez çalışmasında da bu oranlara yer verilmiştir.

Sınıf bazında doğruluk tablosunda ise TP oranı,FPoranı,Hassasiyet,Geri Çağırma, F-Ölçümü ve ROC alanı değerlerinin sınıflar bazında ayrıntılı listesine yer verilmiştir.

Karışıklık matrisinde ise doğru sınıflandırılan veriler gösterilmiştir. Hatalı sınıflandırılan veriler ise hangi alt sınıfta yer aldıysa burada da yer almaktadır.

## A1 KARAR AĞAÇLARI

### 1.Tüm projeler

Çizelge A1.1 Karar Ağacı Tüm Projeler Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 6231         | %76,595          |
| Yanlış Sınıflandırılan Örnek | 1904         | %23,405          |
| Toplam Örnek Sayısı          | 8135         | -                |

Çizelge A1.2 Karar Ağacı Tüm Projeler Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f  |
|----------------------|-----|------|-----|-----|-----|----|
| a =Çok kritik        | 246 | 881  | 5   | 0   | 2   | 0  |
| b = Kritik           | 70  | 5184 | 33  | 17  | 40  | 3  |
| c = Koşullu Hata     | 3   | 230  | 302 | 16  | 27  | 3  |
| d = Hata değil       | 9   | 75   | 39  | 137 | 85  | 12 |
| e = Orta derece      | 6   | 120  | 127 | 70  | 334 | 18 |
| f = İhmal edilebilir | 0   | 12   | 9   | 26  | 34  | 28 |

### 2.Aynı Özelliklere Sahip Proje Oranları

#### 2.1.Ürün Grubu – 1 (iki benzer proje içerir –Daha az hata sayısı)

Çizelge A1.3 Karar Ağacı 1.Ürün Grubu Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 889          | %86.5628         |
| Yanlış Sınıflandırılan Örnek | 138          | %13.4372         |
| Toplam Örnek Sayısı          | 1027         |                  |

Çizelge A1.4 Karar Ağacı 1.Ürün Grubu Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 76  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 889 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 2   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 31  | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 8   | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

## 2.2.Ürün Grubu – 2 (iki benzer proje içerir –Daha fazla hata sayısı)

Çizelge A1.5 Karar Ağacı 2. Ürün Grubu Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 3038         | %76.8919         |
| Yanlış Sınıflandırılan Örnek | 913          | %23.1081         |
| Toplam Örnek Sayısı          | 3951         |                  |

Çizelge A1.6 Karar Ağacı 2. Ürün Grubu Karışıklık Matrisi

| Sınıf                | a | b    | c   | d  | e  | f |
|----------------------|---|------|-----|----|----|---|
| a =Çok kritik        | 0 | 471  | 2   | 0  | 0  | 0 |
| b = Kritik           | 0 | 2677 | 20  | 9  | 27 | 0 |
| c = Koşullu Hata     | 0 | 75   | 265 | 2  | 10 | 0 |
| d = Hata değil       | 0 | 49   | 37  | 39 | 20 | 0 |
| e = Orta derece      | 0 | 57   | 101 | 21 | 57 | 0 |
| f = İhmal edilebilir | 0 | 3    | 8   | 0  | 1  | 0 |

### 3.Proje Bazlı Oranlar

#### 3.1. İlk Proje

Çizelge A1.7 Karar Ağacı 1. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 351          | %68.68           |
| Yanlış Sınıflandırılan Örnek | 160          | %31.31           |
| Toplam Örnek Sayısı          | 511          |                  |

Çizelge A1.8 Karar Ağacı 1. Proje Karışıklık Matrisi

| Sınıf                | a | b   | c  | d  | e  | f |
|----------------------|---|-----|----|----|----|---|
| a =Çok kritik        | 2 | 27  | 0  | 0  | 1  | 0 |
| b = Kritik           | 0 | 213 | 7  | 0  | 2  | 0 |
| c = Koşullu Hata     | 0 | 11  | 26 | 0  | 3  | 0 |
| d = Hata değil       | 0 | 2   | 3  | 15 | 34 | 0 |
| e = Orta derece      | 0 | 9   | 21 | 17 | 95 | 1 |
| f = İhmal edilebilir | 0 | 0   | 3  | 4  | 15 | 0 |

#### 3.2. İkinci Proje

Çizelge A1.9 Karar Ağacı 2. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 109          | %89              |
| Yanlış Sınıflandırılan Örnek | 13           | %10.65           |
| Toplam Örnek Sayısı          | 122          |                  |

Çizelge A1.10 Karar Ağacı 2. Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 11  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 109 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 1   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 1   | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

### 3.3. Üçüncü Proje

Çizelge A1.11 Karar Ağacı 3. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 781          | %86.29           |
| Yanlış Sınıflandırılan Örnek | 124          | %13,70           |
| Toplam Örnek Sayısı          | 905          |                  |

Çizelge A1.12 Karar Ağacı 3. Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 1 | 63  | 0 | 1 | 0 | 0 |
| b = Kritik           | 1 | 776 | 0 | 3 | 0 | 0 |
| c = Koşullu Hata     | 0 | 2   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 25  | 0 | 3 | 2 | 0 |
| e = Orta derece      | 0 | 18  | 0 | 1 | 1 | 0 |
| f = İhmal edilebilir | 0 | 7   | 0 | 0 | 1 | 0 |

### 3.4. Dördüncü Proje

Çizelge A1.13 Karar Ağacı 4. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 1391         | %71.4066         |
| Yanlış Sınıflandırılan Örnek | 557          | %28.5934         |
| Toplam Örnek Sayısı          | 1948         |                  |

Çizelge A1.14 Karar Ağacı 4. Proje Karışıklık Matrisi

| Sınıf                | a | b    | c  | d  | e   | f  |
|----------------------|---|------|----|----|-----|----|
| a =Çok kritik        | 2 | 202  | 0  | 0  | 0   | 0  |
| b = Kritik           | 4 | 1132 | 2  | 3  | 1   | 3  |
| c = Koşullu Hata     | 0 | 145  | 9  | 10 | 16  | 4  |
| d = Hata değil       | 0 | 4    | 2  | 75 | 23  | 13 |
| e = Orta derece      | 1 | 22   | 10 | 36 | 146 | 18 |
| f = İhmal edilebilir | 0 | 5    | 1  | 22 | 10  | 27 |

### 3.5. Beşinci Proje

Çizelge A1.15 Karar Ağacı 5. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 396          | %80.48           |
| Yanlış Sınıflandırılan Örnek | 96           | %19.51           |
| Toplam Örnek Sayısı          | 492          |                  |

Çizelge A1.16 Karar Ağacı 5. Proje Karışıklık Matrisi

| Sınıf                | a   | b   | c | d | e  | f |
|----------------------|-----|-----|---|---|----|---|
| a =Çok kritik        | 243 | 18  | 0 | 0 | 0  | 0 |
| b = Kritik           | 62  | 124 | 0 | 0 | 0  | 0 |
| c = Koşullu Hata     | 1   | 0   | 1 | 0 | 1  | 0 |
| d = Hata değil       | 0   | 0   | 0 | 7 | 1  | 0 |
| e = Orta derece      | 4   | 4   | 1 | 0 | 21 | 0 |
| f = İhmal edilebilir | 1   | 0   | 0 | 0 | 1  | 0 |

### 3.6. Altıncı Proje

Çizelge A1.17 Karar Ağacı 6. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 2383         | %79.2484         |
| Yanlış Sınıflandırılan Örnek | 624          | %20.75           |
| Toplam Örnek Sayısı          | 3007         |                  |

Çizelge A1.18 Karar Ağacı 6. Proje Karışıklık Matrisi

| Sınıf                | a | b    | c   | d  | e  | f |
|----------------------|---|------|-----|----|----|---|
| a =Çok kritik        | 0 | 263  | 0   | 0  | 1  | 0 |
| b = Kritik           | 0 | 2083 | 14  | 8  | 25 | 0 |
| c = Koşullu Hata     | 0 | 49   | 230 | 0  | 4  | 0 |
| d = Hata değil       | 0 | 48   | 32  | 31 | 15 | 0 |
| e = Orta derece      | 0 | 52   | 89  | 13 | 39 | 0 |
| f = İhmal edilebilir | 0 | 3    | 8   | 0  | 0  | 0 |

### 3.7.Yedinci Proje

Çizelge A1.19 Karar Ağacı 7. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 672          | %71.18           |
| Yanlış Sınıflandırılan Örnek | 272          | %28.81           |
| Toplam Örnek Sayısı          | 944          |                  |

Çizelge A1.20 Karar Ağacı 7. Proje Karışıklık Matrisi

| Sınıf                | a | b   | c  | d | e  | f |
|----------------------|---|-----|----|---|----|---|
| a =Çok kritik        | 0 | 207 | 2  | 0 | 0  | 0 |
| b = Kritik           | 2 | 599 | 2  | 0 | 0  | 0 |
| c = Koşullu Hata     | 0 | 27  | 38 | 0 | 4  | 0 |
| d = Hata değil       | 0 | 3   | 1  | 6 | 9  | 0 |
| e = Orta derece      | 0 | 2   | 4  | 8 | 29 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0  | 1 | 0  | 0 |

### 3.8. Sekizinci Proje

Çizelge A1.21 Karar Ağacı 8. Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 172          | %83.49           |
| Yanlış Sınıflandırılan Örnek | 34           | %16.50           |
| Toplam Örnek Sayısı          | 206          |                  |

Çizelge A1.22 Karar Ağacı 8. Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 20  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 172 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 2   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 12  | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

## A-2 NAIVE BAYES

### 1.Tüm projeler

Çizelge A2.1 Naive Bayes Tüm Projeler Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 5914         | %72.69           |
| Yanlış Sınıflandırılan Örnek | 2221         | %27.30           |
| Toplam Örnek Sayısı          | 8135         |                  |

Çizelge A2.2 Naive Bayes Tüm Projeler Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f |
|----------------------|-----|------|-----|-----|-----|---|
| a =Çok kritik        | 353 | 685  | 13  | 6   | 5   | 2 |
| b = Kritik           | 308 | 4740 | 94  | 98  | 99  | 8 |
| c = Koşullu Hata     | 22  | 147  | 266 | 29  | 117 | 0 |
| d = Hata değil       | 12  | 23   | 53  | 135 | 132 | 2 |
| e = Orta derece      | 23  | 51   | 106 | 73  | 420 | 2 |
| f = İhmal edilebilir | 2   | 0    | 11  | 28  | 67  | 0 |

## 2. Aynı Özelliklere Sahip Proje Oranları

### 2.1.Ürün Grubu – 1 (iki benzer proje içerir –Daha az hata sayısı)

Çizelge A2.3 Naive Bayes 1.Ürün Grubu Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 885          | %86,17           |
| Yanlış Sınıflandırılan Örnek | 142          | %13,82           |
| Toplam Örnek Sayısı          | 1027         |                  |

Çizelge A2.4 Naive Bayes 1.Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b   | c | d | e | f |
|----------------------|----|-----|---|---|---|---|
| a =Çok kritik        | 5  | 71  | 0 | 0 | 0 | 0 |
| b = Kritik           | 11 | 877 | 0 | 1 | 0 | 0 |
| c = Koşullu Hata     | 0  | 2   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0  | 28  | 0 | 3 | 0 | 0 |
| e = Orta derece      | 0  | 21  | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0  | 8   | 0 | 0 | 0 | 0 |

### 2.2.Ürün Grubu – 2 (iki benzer proje içerir –Daha fazla hata sayısı)

Çizelge A2.5 Naive Bayes 2.Ürün Grubu Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 2918         | %73.85           |
| Yanlış Sınıflandırılan Örnek | 1033         | %26.14           |
| Toplam Örnek Sayısı          | 3951         |                  |

Çizelge A2.6 Naive Bayes 2.Ürün Grubu Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d  | e   | f |
|----------------------|-----|------|-----|----|-----|---|
| a =Çok kritik        | 157 | 310  | 4   | 1  | 1   | 0 |
| b = Kritik           | 202 | 2360 | 69  | 37 | 65  | 0 |
| c = Koşullu Hata     | 16  | 33   | 248 | 1  | 54  | 0 |
| d = Hata değil       | 3   | 14   | 42  | 33 | 53  | 0 |
| e = Orta derece      | 0   | 10   | 84  | 22 | 120 | 0 |
| f = İhmal edilebilir | 0   | 0    | 4   | 0  | 8   | 0 |

### 3.Proje Bazlı Oranlar

#### 3.1. İlk Proje

Çizelge A2.7 Naive Bayes 1.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 350          | %68.49           |
| Yanlış Sınıflandırılan Örnek | 161          | %31.50           |
| Toplam Örnek Sayısı          | 511          |                  |

Çizelge A2.8 Naive Bayes 1.Proje Karışıklık Matrisi

| Sınıf                | a   | b | c | d | e   | f |
|----------------------|-----|---|---|---|-----|---|
| a =Çok kritik        | 26  | 0 | 0 | 0 | 3   | 0 |
| b = Kritik           | 209 | 1 | 1 | 0 | 11  | 0 |
| c = Koşullu Hata     | 9   | 6 | 6 | 0 | 25  | 0 |
| d = Hata değil       | 2   | 0 | 0 | 5 | 47  | 0 |
| e = Orta derece      | 9   | 4 | 4 | 1 | 129 | 0 |
| f = İhmal edilebilir | 0   | 0 | 0 | 0 | 22  | 0 |

### 3.2. İkinci Proje

Çizelge A2.9 Naive Bayes 2.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 109          | %89.34           |
| Yanlış Sınıflandırılan Örnek | 13           | %10.65           |
| Toplam Örnek Sayısı          | 122          |                  |

Çizelge A2.10 Naive Bayes 2.Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 11  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 109 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 1   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 1   | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

### 3.3. Üçüncü Proje

Çizelge A2.11 Naive Bayes 3.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 784          | %86.62           |
| Yanlış Sınıflandırılan Örnek | 121          | %13.37           |
| Toplam Örnek Sayısı          | 905          |                  |

Çizelge A2.12 Naive Bayes 3.Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d  | e | f |
|----------------------|---|-----|---|----|---|---|
| a =Çok kritik        | 4 | 60  | 0 | 1  | 0 | 0 |
| b = Kritik           | 8 | 768 | 0 | 4  | 0 | 0 |
| c = Koşullu Hata     | 0 | 2   | 0 | 0  | 0 | 0 |
| d = Hata değil       | 0 | 18  | 0 | 12 | 0 | 0 |
| e = Orta derece      | 0 | 18  | 0 | 2  | 0 | 0 |
| f = İhmal edilebilir | 0 | 8   | 0 | 0  | 0 | 0 |

### 3.4. Dördüncü Proje

Çizelge A2.13 Naive Bayes 4.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 1394         | %71.56           |
| Yanlış Sınıflandırılan Örnek | 554          | %28.43           |
| Toplam Örnek Sayısı          | 1948         |                  |

Çizelge A2.14 Naive Bayes 4.Proje Karışıklık Matrisi

| Sınıf                | a  | b    | c  | d  | e   | f |
|----------------------|----|------|----|----|-----|---|
| a =Çok kritik        | 15 | 185  | 3  | 0  | 1   | 0 |
| b = Kritik           | 26 | 1095 | 17 | 3  | 4   | 0 |
| c = Koşullu Hata     | 7  | 96   | 44 | 11 | 26  | 0 |
| d = Hata değil       | 0  | 0    | 3  | 44 | 70  | 0 |
| e = Orta derece      | 2  | 18   | 2  | 15 | 196 | 0 |
| f = İhmal edilebilir | 0  | 0    | 1  | 8  | 56  | 0 |

### 3.5.Beşinci Proje

Çizelge A2.15 Naive Bayes 5.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 390          | %79.26           |
| Yanlış Sınıflandırılan Örnek | 102          | %20.73           |
| Toplam Örnek Sayısı          | 492          |                  |

Çizelge A2.16 Naive Bayes 5.Proje Karışıklık Matrisi

| Sınıf                | a   | b   | c | d | e  | f |
|----------------------|-----|-----|---|---|----|---|
| a =Çok kritik        | 244 | 17  | 0 | 0 | 0  | 0 |
| b = Kritik           | 61  | 125 | 0 | 0 | 0  | 0 |
| c = Koşullu Hata     | 0   | 0   | 0 | 0 | 3  | 0 |
| d = Hata değil       | 7   | 0   | 0 | 0 | 1  | 0 |
| e = Orta derece      | 5   | 4   | 0 | 0 | 21 | 0 |
| f = İhmal edilebilir | 0   | 1   | 0 | 0 | 1  | 0 |

### 3.6. Altıncı Proje

Çizelge A2.17 Naive Bayes 6.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 2245         | %74.65           |
| Yanlış Sınıflandırılan Örnek | 762          | %25.34           |
| Toplam Örnek Sayısı          | 3007         |                  |

Çizelge A2.18 Naive Bayes 6.Proje Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d  | e  | f |
|----------------------|-----|------|-----|----|----|---|
| a =Çok kritik        | 51  | 209  | 1   | 1  | 2  | 0 |
| b = Kritik           | 127 | 1852 | 59  | 27 | 65 | 0 |
| c = Koşullu Hata     | 5   | 26   | 215 | 1  | 36 | 0 |
| d = Hata değil       | 0   | 20   | 37  | 30 | 39 | 0 |
| e = Orta derece      | 0   | 11   | 65  | 20 | 97 | 0 |
| f = İhmal edilebilir | 0   | 3    | 3   | 0  | 5  | 0 |

### 3.7. Yedinci Proje

Çizelge A2.19 Naive Bayes 7.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 698          | %73.94           |
| Yanlış Sınıflandırılan Örnek | 246          | %26.05           |
| Toplam Örnek Sayısı          | 944          |                  |

Çizelge A2.20 Naive Bayes 7.Proje Karışıklık Matrisi

| Sınıf                | a  | b   | c  | d | e  | f |
|----------------------|----|-----|----|---|----|---|
| a =Çok kritik        | 87 | 118 | 4  | 0 | 0  | 0 |
| b = Kritik           | 61 | 533 | 8  | 0 | 1  | 0 |
| c = Koşullu Hata     | 6  | 12  | 46 | 0 | 5  | 0 |
| d = Hata değil       | 3  | 0   | 1  | 0 | 15 | 0 |
| e = Orta derece      | 0  | 2   | 8  | 1 | 32 | 0 |
| f = İhmal edilebilir | 0  | 0   | 0  | 0 | 1  | 0 |

### 3.8. Sekizinci Proje

Çizelge A2.21 Naive Bayes 8.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 172          | %83.49           |
| Yanlış Sınıflandırılan Örnek | 34           | %16.50           |
| Toplam Örnek Sayısı          | 206          |                  |

Çizelge A2.22 Naive Bayes 8.Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 20  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 172 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 2   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 12  | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

### A-3 DESTEK VEKTÖR MAKİNELERİ

#### 1.Tüm projeler

Çizelge A3.1 DVM Tüm Projeler Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 6382         | %78,45           |
| Yanlış Sınıflandırılan Örnek | 1753         | %21,54           |
| Toplam Örnek Sayısı          | 8135         |                  |

Çizelge A3.2 DVM Tüm Projeler Karışıklık Matrisi

| Sınıf                | a   | b    | c   | d   | e   | f  |
|----------------------|-----|------|-----|-----|-----|----|
| a =Çok kritik        | 354 | 690  | 16  | 1   | 3   | 0  |
| b = Kritik           | 193 | 5058 | 50  | 18  | 27  | 1  |
| c = Koşullu Hata     | 9   | 138  | 364 | 17  | 48  | 5  |
| d = Hata değil       | 1   | 41   | 33  | 162 | 98  | 22 |
| e = Orta derece      | 8   | 77   | 80  | 75  | 410 | 25 |
| f = İhmal edilebilir | 1   | 0    | 7   | 27  | 40  | 34 |

## 2. Aynı Özelliklere Sahip Proje Oranları

### 2.1.Ürün Grubu – 1 (iki benzer proje içerir –Daha az hata sayısı)

Çizelge A3.3 DVM 1.Ürün Grubu Özet Tablosu

|                              |      |        |
|------------------------------|------|--------|
| Doğru Sınıflandırılan Örnek  | 906  | %88.21 |
| Yanlış Sınıflandırılan Örnek | 121  | %11.78 |
| Toplam Örnek Sayısı          | 1027 |        |

Çizelge A3.4 DVM 1.Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b   | c | d  | e  | f |
|----------------------|----|-----|---|----|----|---|
| a =Çok kritik        | 9  | 66  | 0 | 1  | 0  | 0 |
| b = Kritik           | 18 | 863 | 0 | 6  | 2  | 0 |
| c = Koşullu Hata     | 0  | 1   | 0 | 0  | 0  | 1 |
| d = Hata değil       | 1  | 4   | 0 | 22 | 4  | 0 |
| e = Orta derece      | 0  | 4   | 0 | 3  | 10 | 4 |
| f = İhmal edilebilir | 0  | 0   | 0 | 0  | 6  | 2 |

## 2.2.Ürün Grubu – 2 (iki benzer proje içerir –Daha fazla hata sayısı)

Çizelge A3.5 DVM 2.Ürün Grubu Özet Tablosu

|                              |      |        |
|------------------------------|------|--------|
| Doğru Sınıflandırılan Örnek  | 3098 | %78.41 |
| Yanlış Sınıflandırılan Örnek | 853  | %21.58 |
| Toplam Örnek Sayısı          | 3951 |        |

Çizelge A3.6 DVM 2.Ürün Grubu Karışıklık Matrisi

| Sınıf                | a  | b    | c   | d  | e   | f |
|----------------------|----|------|-----|----|-----|---|
| a =Çok kritik        | 73 | 397  | 3   | 0  | 0   | 0 |
| b = Kritik           | 71 | 2620 | 18  | 9  | 14  | 1 |
| c = Koşullu Hata     | 6  | 61   | 244 | 5  | 33  | 3 |
| d = Hata değil       | 3  | 21   | 28  | 53 | 36  | 4 |
| e = Orta derece      | 0  | 26   | 69  | 31 | 108 | 2 |
| f = İhmal edilebilir | 0  | 0    | 4   | 2  | 6   | 0 |

## 3.Proje Bazlı Oranlar

### 3.1. İlk Proje

Çizelge A3.7 DVM 1.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 358          | %70.05           |
| Yanlış Sınıflandırılan Örnek | 153          | %29.94           |
| Toplam Örnek Sayısı          | 511          |                  |

Çizelge A3.8 DVM 1.Proje Karışıklık Matrisi

| Sınıf                | a | b   | c  | d  | e   | f |
|----------------------|---|-----|----|----|-----|---|
| a =Çok kritik        | 2 | 27  | 0  | 0  | 1   | 0 |
| b = Kritik           | 6 | 208 | 4  | 0  | 4   | 0 |
| c = Koşullu Hata     | 0 | 9   | 27 | 0  | 4   | 0 |
| d = Hata değil       | 0 | 3   | 0  | 18 | 33  | 0 |
| e = Orta derece      | 0 | 9   | 11 | 18 | 103 | 2 |
| f = İhmal edilebilir | 0 | 0   | 1  | 4  | 17  | 0 |

### 3.2.İkinci Proje

Çizelge A3.9 DVM 2.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 109          | %89.34           |
| Yanlış Sınıflandırılan Örnek | 13           | %10.65           |
| Toplam Örnek Sayısı          | 122          |                  |

Çizelge A3.10 DVM 2.Proje Karışıklık Matrisi

| Sınıf                | a | b   | c | d | e | f |
|----------------------|---|-----|---|---|---|---|
| a =Çok kritik        | 0 | 11  | 0 | 0 | 0 | 0 |
| b = Kritik           | 0 | 109 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 1   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 1   | 0 | 0 | 0 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

### 3.3. Üçüncü Proje

Çizelge A3.11 DVM 3.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 796          | %87.95           |
| Yanlış Sınıflandırılan Örnek | 109          | %12.04           |
| Toplam Örnek Sayısı          | 905          |                  |

Çizelge A3.12 DVM 3.Proje Karışıklık Matrisi

| Sınıf                | a  | b   | c | d  | e | f |
|----------------------|----|-----|---|----|---|---|
| a =Çok kritik        | 13 | 51  | 0 | 1  | 0 | 0 |
| b = Kritik           | 17 | 754 | 0 | 6  | 3 | 0 |
| c = Koşullu Hata     | 0  | 1   | 0 | 0  | 0 | 1 |
| d = Hata değil       | 1  | 7   | 0 | 19 | 3 | 0 |
| e = Orta derece      | 0  | 4   | 0 | 4  | 8 | 4 |
| f = İhmal edilebilir | 0  | 0   | 0 | 0  | 6 | 2 |

### 3.4. Dördüncü Proje

Çizelge A3.13 DVM 4.Proje Özet Tablosu

| ÖRNEK                        | ÖRNEK SAYISI | BAŞARI YÜZDELERİ |
|------------------------------|--------------|------------------|
| Doğru Sınıflandırılan Örnek  | 1448         | %74.33           |
| Yanlış Sınıflandırılan Örnek | 500          | %25.66           |
| Toplam Örnek Sayısı          | 1948         |                  |

Çizelge A3.14 DVM 4.Proje Karışıklık Matrisi

| Sınıf                | a  | b    | c  | d  | e   | f  |
|----------------------|----|------|----|----|-----|----|
| a =Çok kritik        | 4  | 186  | 14 | 0  | 0   | 0  |
| b = Kritik           | 13 | 1098 | 29 | 1  | 4   | 0  |
| c = Koşullu Hata     | 3  | 74   | 86 | 10 | 8   | 3  |
| d = Hata değil       | 0  | 3    | 7  | 68 | 21  | 18 |
| e = Orta derece      | 0  | 21   | 5  | 30 | 160 | 17 |
| f = İhmal edilebilir | 0  | 0    | 2  | 21 | 10  | 32 |

### 3.5.Beşinci Proje

Çizelge A3.15 DVM 5.Proje Özet Tablosu

| ÖZET                         |     |        |
|------------------------------|-----|--------|
| Doğru Sınıflandırılan Örnek  | 397 | %80.69 |
| Yanlış Sınıflandırılan Örnek | 95  | %19.30 |
| Toplam Örnek Sayısı          | 492 |        |

Çizelge A3.16 DVM 5.Proje Karışıklık Matrisi

| Karışıklık Matrisi   |     |     |   |   |    |   |
|----------------------|-----|-----|---|---|----|---|
| Sınıf                | a   | b   | c | d | e  | f |
| a =Çok kritik        | 245 | 16  | 0 | 0 | 0  | 0 |
| b = Kritik           | 63  | 123 | 0 | 0 | 0  | 0 |
| c = Koşullu Hata     | 1   | 0   | 0 | 0 | 2  | 0 |
| d = Hata değil       | 0   | 0   | 0 | 8 | 0  | 0 |
| e = Orta derece      | 5   | 3   | 1 | 0 | 21 | 0 |
| f = İhmal edilebilir | 1   | 0   | 0 | 0 | 1  | 0 |

### 3.6. Altıncı Proje

Çizelge A3.17 DVM 6.Proje Özet Tablosu

| ÖZET                         |      |        |
|------------------------------|------|--------|
| Doğru Sınıflandırılan Örnek  | 2412 | %80.21 |
| Yanlış Sınıflandırılan Örnek | 595  | %19.78 |
| Toplam Örnek Sayısı          | 3007 |        |

Çizelge A3.18 DVM 6.Proje Karışıklık Matrisi

| Karışıklık Matrisi   |    |      |     |    |    |   |
|----------------------|----|------|-----|----|----|---|
| Sınıf                | a  | b    | c   | d  | e  | f |
| a =Çok kritik        | 5  | 257  | 0   | 0  | 2  | 0 |
| b = Kritik           | 14 | 2079 | 10  | 11 | 15 | 1 |
| c = Koşullu Hata     | 1  | 41   | 204 | 7  | 29 | 1 |
| d = Hata değil       | 0  | 25   | 25  | 39 | 33 | 4 |
| e = Orta derece      | 1  | 30   | 56  | 19 | 85 | 2 |
| f = İhmal edilebilir | 0  | 0    | 4   | 2  | 5  | 0 |

### 3.7. Yedinci Proje

Çizelge A3.19 DVM 7.Proje Özet Tablosu

| ÖZET                         |     |         |
|------------------------------|-----|---------|
| Doğru Sınıflandırılan Örnek  | 691 | %73.199 |
| Yanlış Sınıflandırılan Örnek | 253 | %26.80  |
| Toplam Örnek Sayısı          | 944 |         |

Çizelge A3.20 DVM 7.Proje Karışıklık Matrisi

| Karışıklık Matrisi   |    |     |    |    |    |   |
|----------------------|----|-----|----|----|----|---|
| Sınıf                | a  | b   | c  | d  | e  | f |
| a =Çok kritik        | 83 | 124 | 2  | 0  | 0  | 0 |
| b = Kritik           | 73 | 522 | 7  | 0  | 1  | 0 |
| c = Koşullu Hata     | 4  | 13  | 47 | 0  | 5  | 0 |
| d = Hata değil       | 0  | 0   | 1  | 10 | 8  | 0 |
| e = Orta derece      | 2  | 1   | 7  | 4  | 29 | 0 |
| f = İhmal edilebilir | 0  | 0   | 0  | 0  | 1  | 0 |

### 3.8. Sekizinci Proje

Çizelge A3.21 DVM 8.Proje Özet Tablosu

| ÖZET                         |     |        |
|------------------------------|-----|--------|
| Doğru Sınıflandırılan Örnek  | 171 | %83.0  |
| Yanlış Sınıflandırılan Örnek | 35  | %16.99 |
| Toplam Örnek Sayısı          | 206 |        |

Çizelge A3.22 DVM 8.Proje Karışıklık Matrisi

| Karışıklık Matrisi   |   |     |   |   |   |   |
|----------------------|---|-----|---|---|---|---|
| Sınıf                | a | b   | c | d | e | f |
| a =Çok kritik        | 2 | 18  | 0 | 0 | 0 | 0 |
| b = Kritik           | 7 | 165 | 0 | 0 | 0 | 0 |
| c = Koşullu Hata     | 0 | 0   | 0 | 0 | 0 | 0 |
| d = Hata değil       | 0 | 2   | 0 | 0 | 0 | 0 |
| e = Orta derece      | 0 | 8   | 0 | 0 | 4 | 0 |
| f = İhmal edilebilir | 0 | 0   | 0 | 0 | 0 | 0 |

## ÖZGEÇMİŞ

---

### KİŞİSEL BİLGİLER

**Adı Soyadı** : Ahmet UNUDULMAZ  
**Doğum Tarihi ve Yeri** : 24.01.1988 , İstanbul  
**Yabancı Dili** : İngilizce  
**E-posta** : ahmetunudulmaz@gmail.com

### ÖĞRENİM DURUMU

| Derece | Alan                       | Okul/Üniversite               | Mezuniyet Yılı |
|--------|----------------------------|-------------------------------|----------------|
| Lisans | Bilgisayar<br>Mühendisliği | Kocaeli Üniversitesi          | 2010           |
| Lise   | Fen Bilimleri              | Köy Hizmetleri Anadolu Lisesi | 2006           |

### İŞ TECRÜBESİ

| Yıl        | Firma/Kurum          | Görevi         |
|------------|----------------------|----------------|
| Mart 2011- | Elektronik İşletmesi | ArGe Mühendisi |

## **YAYINLARI**

### **Bildiri**

1.Unudulmaz, A, Cingiz, M. Ö,,Kalıpsız, O., “*Yazılım Projelerindeki Problem Etkilerinin Yazılım Mimarisi ile İlişkilendirilmesi*”, 4.Ulusal Yazılım Mimarisi Konferansı,UYMK 2012, İzmir, (7-8 Aralık), 2012

2.Cingiz, M. Ö, Unudulmaz, A., Kalıpsız, O., “*FarklıKestirim Yaklaşımları Kullanılarak Yazılım Projelerindeki Problem Etkilerinin Risk Yönetiminin Gerçeklenmesi*”, 6. Ulusal Yazılım Mühendisliği Sempozyumu, UYMS 2012, Ankara, (30 Mayıs-1 Haziran), 2012

3.Cingiz, M. Ö, Unudulmaz, A., Kalıpsız, O., “*Software Project Problem Effects on Risk Analysis*”,12th International Conference on Intelligent Software Methodologies, Tools and Techniques, Budapeşte (22-24 Eylül), 2013