T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Özdevimli Öğrenme Yaklaşımı ile Bilgi Güvenliği Risklerinin Nitel Değerlendirilmesine Yönelik Bir Model

Mete EMİNAĞAOĞLU

Doktora Tezi

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Şaban EREN

2011

EDİRNE

T.C.

TRAKYA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

ÖZDEVİMLİ ÖĞRENME YAKLAŞIMI İLE BİLGİ GÜVENLİĞİ RİSKLERİNİN NİTEL DEĞERLENDİRİLMESİNE YÖNELİK BİR MODEL

Mete EMİNAĞAOĞLU

DOKTORA TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

Bu tez ….. / ….. / 2011 tarihinde aşağıdaki jüri tarafından kabul edilmiştir.

Prof. Dr. Şaban EREN

(Danışman)

Yrd. Doç. Dr. Erdem Uçar

(2. Danışman)

Doç. Dr. Yılmaz Kılıçaslan

(Üye)

Doç. Dr. Tahir Altınbalık

(Üye)

Yrd. Doç. Dr. Aydın Carus

(Üye)

Yrd. Doç. Dr. Yılmaz Gökşen

(Üye)

Doktora Tezi

Trakya Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

ÖZET

Örgütlerde ve kurumsal yapılarda bilgi güvenliği yönetiminin en önemli aşaması bilgi güvenliği risklerinin belirlenmesi ve bu risklerin çeşitli nitel veya nicel yöntemlerle hesaplanıp değerlendirilmesi sürecidir. Bilgi güvenliği risklerini değerlendirmede doğrusal olmayan modeller üretilmesi; bilinen yöntemlere alternatif olabilecek gelişime açık güncel bir araştırma konusu olarak kabul edilmektedir. Bu çalışmada, bilgi güvenliği risklerinin nitel değerlendirmesine yönelik yeni bir model ortaya konulmaktadır.

Yapılan çalışma iki temel amaca yöneliktir. Birinci amaç, bir kuruma özel bir bilgi güvenliği risk anketi oluşturulması, kurumda anketin uygulanması ve elde edilen sonuçların değerlendirilmesidir. Çalışmanın ikinci amacı da anket verilerinden yola çıkarak özdevimli öğrenme sınıflandırıcılarına uygun özgün bir nitel risk değerlendirme modeli geliştirmek ve hangi sınıflandırıcıların tasarlanan model için en başarılı sonuçları verdiğini belirlemektir. Sonuçlar değerlendirildiğinde geliştirilen modelin başarılı ve iyileştirmeye açık özgün bir model olduğu görülmüştür. Çalışma süresince bu iki amaca ek olarak diğer bazı önemli ve özgün bulgular ve sonuçlar elde edilmiştir. Özdevimli öğrenme yaklaşımının bilgi güvenliği risk değerlendirme sürecinde denetim mekanizması olarak katkı sağlayacağı görülmüş ve çalışmadaki modelin iyileştirilmesinde bu sonuçlardan yararlanılmıştır. Bir başka bulgu, değişik anket uygulamalarında özdevimli öğrenme yaklaşımının hata denetim işlevi olarak kullanılabileceğinin ortaya çıkarılmış olmasıdır. Geliştirilen örnek modelin uygulanabileceği yeni çalışma alanları ve modelin iyileştirilmesine yönelik öneriler çalışma sonucunda ortaya konulmuştur.

2011, 164 sayfa

Anahtar Kelimeler: Bilgi Güvenliği Risk Değerlendirmesi, Nitel Risk Analizi, Bilgi Güvenliği Risk Anketi, Özdevimli Öğrenme, İkili Sınıflandırıcılar.

Doctorate Thesis

Trakya University

Institute of Natural Sciences

Department of Computer Engineering

ABSTRACT

The definition, analysis and assessment of information security risks by the aid of quantitative or qualitative methods is the most crucial process in information security management amongst the institutions and corporations. Instead of the common methodologies and models; derivation and usage of non-linear models for information security risk assessment has become an alternative hot topic. In this study, a new model has been proposed for assessing the qualitative information security risks.

This dissertation's basic aim is twofold. The first aim is to design, derive and implement a unique information security risk analysis survey for a specific institution. The second aim of this study is to implement an original machine learning classification model that deduces and prioritizes the risks with the data set that is derived from the survey results. The model is refined by observing and comparing the performance values of binary classifier algorithms' train and test results. The results show that the model can be accepted as a successful prototype. In addition, it is shown that some machine learning classifiers could be used as a cross-check control mechanism in information security risk evaluations and assessments. It is also shown that machine learning algorithms can be adapted and used as a supporting control mechanism for discovering biased answers in generic purpose surveys. Some recommendations for further improvements and new research areas have also been included in the study.

2011, 164 pages

Keywords: Information Security Risk Assessment, Qualitative Risk Analysis, Information Security Risk Survey, Machine Learning, Binary Classifiers.

ÖNSÖZ

Tez çalışmamın en başından sonuna kadar tüm aşamalarında birçok konuda engin bilgi birikimi ve deneyiminden yararlandığım, büyük ilgi ve desteğini gördüğüm tez danışmanım Prof. Dr. Şaban EREN başta olmak üzere; değerli hocalarım Yrd. Doç. Dr. Erdem UÇAR ve Doç. Dr. Yılmaz KILIÇASLAN'a, çalışmanın anket uygulaması sürecinde önemli yardımlarda bulunan Yrd. Doç. Dr. Yılmaz GÖKŞEN'e, anket uygulamasının gerçekleşmesine olanak sağlayan, ankete destek veren Sn. Yılmaz KADIZ başta olmak üzere tüm Karşıyaka Devlet Hastanesi yetkililerine ve çalışanlarına, sevgi ve sabrıyla bana güç ve moral veren sevgili eşim Neslihan EMİNAĞAOĞLU'na ve çalışmanın çeşitli aşamalarında katkıda bulunan diğer tüm değerli meslektaşlarıma sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

SİMGELER VE KISALTMALAR

BGYS	:	Bilgi Güvenliği Yönetim Sistemi
COBIT	:	Bilgi ve İlgili Teknolojiler için Kontrol Hedefleri (Control Objectives for Information and Related Technologies)
FN	:	Yanlış Negatif (False Negative)
FP	:	Yanlış Pozitif (False Positive)
ISO	:	Uluslararası Standartlar Örgütü (International Organization for Standardization)
ISO/IEC 27001	:	Uluslararası Bilgi Güvenliği Yönetim Sistemleri Standardı
ISO/IEC 27002	:	Uluslararası Bilgi Güvenliği Yönetimi için Uygulama Kuralları Standardı
ISO/IEC 27005	:	Uluslararası Bilgi Güvenliği Risk Yönetimi Standardı
ITGI	:	Bilgi Teknolojileri Yönetişim Enstitüsü (Information Technology Governance Institute)
K.D.H.	:	Karşıyaka Devlet Hastanesi
PCI DSS	:	Kartlı Ödeme için Veri Güvenliği Endüstri Standardı (Payment Card Industry Data Security Standard)
PUKÖ	:	Planla-Uygula-Kontrol Et-Önlem al
TCK	:	Türk Ceza Kanunu
TN	:	Doğru Negatif (True Negative)
TP	:	Doğru Pozitif (True Positive)
TSE	:	Türk Standartları Enstitüsü
TS ISO/IEC 27001	:	TSE Bilgi Güvenliği Yönetim Sistemleri Standardı
vd	:	varsayılan değer (default value)

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

21. yüzyılda şirketler, devletler, kurumlar, bireyler ve toplumların tamamının ortak bileşkesi bilgi çağında yaşamaları ve bilgi çağının gerekliliklerine ayak uydurmak zorunda olmalarıdır. Üretim, hizmet veya tüketim sürecinde, bilgi en değerli rekabet ve başarı unsuru haline gelmiştir. Bu kadar vazgeçilmez bir unsur olan bilginin güvenliği de, artık yadsınamaz bir kavram olarak karşımıza çıkmaktadır. İşin niteliği veya sürecin yapısı ne olursa olsun, teknoloji bağlantılı olmayan süreçlerin yönetiminde bile, bilgi güvenliğinin de etkin, sürekli ve başarılı bir şekilde sağlanarak yönetilmesi çok önemli bir gereksinim olarak kabul edilmektedir. Bilgi güvenliği yönetiminin başarısı da ancak etkin, güvenilir, başarılı bir bilgi güvenliği risk ölçüm yöntemi, sistemi ve araçlarıyla sağlanmaktadır.

Günümüzde çok çeşitli bilgi güvenliği risk ölçüm yöntemleri, modelleri ve bunlara ilişkin yazılımlar var olmasına karşın, bunların hiçbiri bütün kurumlar için genelleştirilecek bir şekle dönüşememekte veya bir kurumdaki her bir farklı risk için başarılı ve doğru sonuçlar vermemektedir. Bu nedenle de, kurumlar ya risklerini sağlıklı biçimde öngörememekte ve risklerini kontrol edememekte, ya da sürekli olarak konunun uzmanlarına danışarak insana bağlı yorum ve müdahalelerle risk ölçümü ve risk yönetimi yapmak zorunda kalmaktadır.

Bu tez çalışmasının amacı, günümüzde her türlü sektörde faaliyet gösteren kurumların bilgi güvenliği risklerini ölçümleme ve değerlendirmede kullandıkları yöntem, araç ve yazılımların yetersiz kaldığı alanlar için, yeni bir yaklaşımla alternatif yöntemler kullanarak ilk örnek olabilecek bir model ortaya koymaya çalışmaktır. Bu yeni yaklaşımda; özdevimli öğrenme (makine öğrenimi) algoritma ve modellerinin kullanılması, böylelikle özdevimli öğrenme yaklaşımı ile bilgi güvenliği risklerinin ölçümlenmesinde hangi kısımlarda mevcut yöntemlerden daha başarılı veya daha başarısız olduğunun bilimsel ve nesnel bir şekilde ortaya konması amaçlanmaktadır. Çalışma, bu alandaki eksiklik ve yetersizlikleri belli ölçüde giderebilecek şekilde bilgi güvenliği risklerinin belli alanları için, güvenilir, esnek, kullanışlı yeni bir yaklaşım

ortaya koyması ve bu sayede bilgi güvenliği yönetiminde kurumlara yardımcı olacak örnek bir model oluşturması açısından önem ve yarar arz etmektedir.

Çalışmanın ilk aşamasına tez konusu ile ilgili ulusal ve uluslararası kaynakların taranmasıyla başlanılmış ve bir referans eserler kütüphanesi oluşturulmuştur. Toplanan tüm eserler belirli kategoriler altında uygun biçimde isimler verilerek konumlandırılmıştır. Sonraki aşamada her bir eserin ön incelemesi yapılarak hangi eserden ne ölçüde yararlanılıp yararlanılmayacağına yönelik bir karşılaştırma yapılarak önemli eserler belirlenmiş ve detaylı bir şekilde incelenmiştir. Bilgi güvenliği, bilgi güvenliğinde risk analizi ve ölçümlemesi, veri madenciliği ve özdevimli öğrenmeye ilişkin kuramsal temeller, önceki araştırmalar, ilgili kavramlar, standartlar, yöntemler ve örnek uygulamalar çalışmanın ikinci ve üçüncü bölümlerinde detaylı olarak verilmiştir.

Çalışmanın dördüncü bölümünde Özdevimli Öğrenme ile Bilgi Güvenliğinde Nitel Risk Değerlendirmesi Modeli başlığı ile çalışmada ortaya konan, geliştirilen ve uygulamalarla sonuç alınan model anlatılmıştır. İlgili modelde kullanılan gerçek örnek veri kümesinin edinilmesindeki anket uygulaması, modelin tasarımı, modele uyarlanarak veri kümesi üzerinden yapılan deneyler, seçilen yöntemler, gözlemler ve modelin yapılandırılmasına ilişkin kapsamlı bilgilendirme yapılmıştır. Çalışmadaki bilgi güvenliği risk alanlarının belirlenmesi, ilgili anketin oluşturulması ve anketin uygulanması sürecinde Türkiye'deki sağlık sektöründeki bir kurumun belli bir birimi ve o birimin çalışanları kapsama alınmıştır. Bilgi güvenliği risk anketinin içeriğindeki risklere ilişkin sorular özgün bir biçimde oluşturulmuş, ayrıca basitleştirilmiş şekilde psikometrik ölçüm amaçlı değerlendirme soruları da ankette kapsanmıştır. Çalışmanın risk değerlendirmesi aşamasında nitel risk değerlendirme yöntemi kullanılmıştır. Kurumdaki öncelikli riskleri belirleme ve tahmin etme sürecinde ikili sınıflandırıcı türündeki özdevimli öğrenme algoritmaları çalışmanın kapsamına alınmıştır. Toplam 68 değişik özdevimli öğrenim algoritması ve bu algoritmaların çeşitli alt işlevleri, değişik parametre seçeneklerinin denenmesi ile birlikte 3200 civarında gözlem ve test yapılmıştır. Elde edilen istatistiksel değerler ve ölçüm sonuçları incelenerek model için en uygun algoritmalar saptanmış ve bu algoritmaların bileşkeleri sonucunda öncelikli risklerin hangileri olduğu belirlenerek modele son şekli verilmiştir. Ortaya konan model sonucunda elde edilen bulgular ve sonuçlar incelenmiş, bilgi güvenliği risk değerlendirmesinin belli alt alanlarında özdevimli öğrenmedeki ikili sınıflandırıcıların

başarım düzeyi belirlenmiş ve bunlara ilişkin yeni gereksinimler, yeterlilikler veya geliştirme alanları ortaya çıkarılmıştır.

Çalışmada elde edilen sonuçlar, ilgili çıktılar ve elde edilen bulguların değerlendirilmesi sonuçlar bölümünde yer almıştır. Bu çalışma sonucunda ortaya konan çeşitli yeni araştırma konuları ve geliştirme alanları ise çalışmanın son bölümü olan tartışma kısmında irdelenmiştir.

2. BİLGİ GÜVENLİĞİ YÖNETİMİ VE BİLGİ GÜVENLİĞİNDE RİSK DEĞERLENDİRMESİ

2.1. BİLGİ GÜVENLİĞİNDE TEMEL KAVRAMLAR

Veriler (data), bir kavram, varlık veya konunun herhangi bir simge öbeği ile ifade edilmesidir. Bilgi (information) ise herhangi bir konu veya varlık ile ilgili verilerin bir araya gelmesi ile ortaya çıkan açıklayıcı bir bütündür. Veri, gerçeklik üzerinde yapılan gözlemlerin sonucu ve bu anlamda bilginin üretildiği hammaddedir. Veri, kullanıcılar için herhangi bir anlam ifade etmeyen olgulardır ve sayısal değerlerin yanı sıra sayısal olmayan değerleri de kapsayabilir (Gökçen, 2007). Bilgi, kişiler veya kurumlar için anlamlı olması yanı sıra belli maddi veya manevi değerleri de içermektedir. Belli süreçler ya da işlemler sonunda bu bilgiler kazanılmış bilgiye (knowledge) dönüşmektedir. Günümüzün rekabetçi ekonomileri ve teknolojileri düşünüldüğünde; verileri bilgi ve kazanılmış bilgiye dönüştürebilen kurumlar, çok daha başarılı, hızlı ve etkin stratejik kararlar alabilmektedir (Vercellis, 2009).

Küreselleşmenin alabildiğince yaygınlaştığı günümüz dünyasında, bilgi ve iletişim teknolojilerinin yaygın kullanımı nedeni ile ülkelerin ve kuruluşların en önemli değerlerinden olan bilginin güvenliğinin sağlanması, ülkelerin ve kuruluşların güvenliğinin sağlanması ile eşdeğer tutulmakta ve önem kazanmaktadır (TBD 4. Çalışma Grubu, 2006).

Bilgi, günümüzdeki diğer önemli tüm ekonomik etmenlere benzer şekilde; kurum, şirket, ülke veya birey için değeri olan ve bu nedenle gereken düzeyde ve istenildiği sürece sürekli korunması gereken bir varlıktır. Bilgiler sadece Internet, bilgisayarlar, vb. elektronik ortamda bulunmayabilir. Bilgi birçok değişik biçimlerde üretilebilir, işlenebilir, kaydedilebilir veya iletilebilir. Kâğıt belgeler, manyetik ortamlarda tutulan ses, görüntü kayıtları, vb. diğer bilgi biçimleri olabildiği gibi, kişiler arası sözlü iletişimle de ifade edilebilir. Bilgi güvenliği, hangi biçimde veya yapıda olursa olsun, bilginin gereken düzeyde ve şekilde korunmasını sağlayan tüm süreçler ve çözümleri kapsamaktadır (Calder ve Watkins, 2008).

Bilgi güvenliğinin, birçok ulusal ve uluslararası kaynakçaya göre üç temel öğeden oluştuğu ve bu üç temel öğenin korunması ile bilgi güvenliğinin sağlanabileceği belirtilmektedir (TSE, 2006). Bu üç öğe gizlilik, bütünlük ve kullanılabilirlik olarak tanımlanmaktadır (Pfleeger, 2007).

Gizlilik (confidentiality), bilginin yetkisiz kişiler, varlıklar ya da süreçlerce kullanılmasının önlenmesi ya da bu gibi yetkisiz unsurlara ifşa edilmemesidir. Bütünlük (integrity), bilgi varlıklarının doğruluğunu ve tamlığını koruma özelliğidir. Kullanılabilirlik (availability), bir bilginin veya bilgi sisteminin sadece yetkililer tarafından ve o yetkililerin gereksinim duyduğu anlarda erişilebilir ve kullanılabilir olma özelliğidir (ISO/IEC, 2005).

Bilgi güvenliğinin gizlilik, bütünlük ve kullanılabilirlik öğeleri üç ayaklı bir taburenin ayakları olarak düşünülebilir. Bilgi güvenliğinin etkin, verimli ve başarılı bir şekilde sağlanması ve sürdürülmesi de; bu üç ayaktan oluşan taburenin ayaklarının birbirinin dengesini bozmayacak bir biçimde yapılanması ve bunun sonucunda taburenin sürekli dengede ve ayakta kalması şeklinde düşünülebilir.

Günümüzdeki bazı ulusal ve uluslararası standartlar ile bazı ilgili kaynaklarda bilgi güvenliğinin bu üç temel öğesine ek olarak bazı yeni ikincil özellikler veya öğeler de katılmaktadır. Bu ek özellikler, doğruluk (authenticity), açıklanabilirlik (accountability), inkâr edememe (non-repudiation) ve güvenilirlik (reliability) olarak tanımlanmaktadır (TSE, 2006, ISO/IEC, 2005).

Bilgi güvenliğinin başarısının temelinde yatan bir başka önemli etmen, bilgi güvenliğine ilişkin riskleri belirlemek, analiz etmek ve bu riskleri kontrol edip yönetebilmektir. Bilgi güvenliği, risk odaklı bir olgudur. Güvenlik risklerinin neler olduğu bilinmediği ve irdelenmediği sürece; hangi düzeyde güvenliğin, ne kadar süreyle, kime, nasıl bir çözümle sağlanacağı da doğru bir şekilde tanımlanamayacaktır. Çalışmanın ilerleyen bölümlerinde bilgi güvenliği riskleri konusuna ayrıntılı değinildiği için bu bölümde bu konudan detaylı olarak bahsedilmemiştir.

Son yıllarda bilgisayar ve internet kullanımının hızla yaygınlaşarak artması sonucu, kişiler, kurumlar ve kuruluşlar işlerini artık çok büyük oranda elektronik ortamlarda gerçekleştirmektedirler. Bunun sonucu olarak e-ticaret, e-kurum, e-devlet, e-imza, e-posta gibi kavramlar hızla klasik çalışma biçimlerinin yerini almaktadır. Bu değişim,

kurumların ticari faaliyetlerinde ve bireylerin günlük yaşantısında neredeyse her alanda görülebilmektedir. Ancak, günümüzde bilişimde yaşanan müthiş gelişim hızı ile beraber kişiler ve kurumlar; yazılımlar ve bilgisayarların kullanılmasıyla yapılan sahtekârlıklar, bilgi hırsızlığı, bilgisayar korsanları, elektronik saldırılar, bilgi sızdırma ve ilgili kuruluşların kendi çalışanlarınca oluşturulabilecek potansiyel iç saldırılar gibi çok çeşitli tehditlerle karşı karşıyadır (Rost ve Glass, 2011). Özellikle yazılımlardaki güvenlik açıklarından yararlanılarak yapılan saldırılar, bilgisayar virüsleri, bilgisayarları ağ üzerinden ele geçirerek bilgisayarlara zarar veren kişilerin kullandığı yöntemler, kişisel ve kurumsal bilgilerin izinsiz olarak elde edilmesi veya değiştirilmesi konusundaki tehditler artarak sürmekte olup, kişiler ve kurumlar bu tehlikeler karşısında giderek daha riskli bir duruma gelmektedir. Hizmetlerin internet ortamında sunulma eğiliminin artması, açık ve özel ağlar arasındaki geçişler, bilgilerin halka açık sistemlerle paylaşılması gibi uygulamaların artması sonucu bilgilere erişimin yetkilendirilmesi ve denetlenmesi güçleşmektedir (Calder ve Watkins, 2008).

Bilgi güvenliğinde yaşanan sorunlar ve bunların sonucunda oluşan maddi kayıplara ilişkin ulusal veya uluslar arası çeşitli sayısal gözlemler ve istatistiksel araştırmalar da yapılmaktadır. CSI ve FBI kurumlarının 2008 yılında yaptıkları bir çalışmanın sonucuna göre; ABD'deki 522 kurumun (devlet veya özel sektör) % 49'unda virüs, truva atı, solucan, vb zararlı kod saldırısı yaşanmış, % 42'sinde dizüstü bilgisayar, cep bilgisayarı, vb mobil cihazlar çalınmış, % 44'ünde şirket çalışanları İnternet ve diğer yetkilerini, erişimlerini suiistimal etmiş ve bir yıl içerisinde bu 522 kurumun toplam maddi kaybı 156 Milyon ABD Doları olmuştur (Richardson, 2008). Bilişim suçlularının değerli bilgi içeren büyük firmalara saldırı olasılığı son yıllarda çok daha fazla olmaktadır ve yapılan araştırmalar sonucu elde edilen istatistiksel veriler de bu savı desteklemektedir. Dünya genelinde yapılan bir başka araştırma raporunda 2008 yılı boyunca yeni tehditlerin yayılması ve amacına ulaşmasında İnternet ortamının ve web sitelerinin ana kaynak olarak kullanıldığı özellikle vurgulanmıştır (Symantec-1, 2009). Aynı çalışmada, saldırganların bu tehditleri geliştirirken ve kullanıcılara yöneltirken eskisine oranla çok daha fazla "kişiye özel" zararlı kod eylemleri düzenlediklerinin de altı çizilmektedir. İlgili çalışmanın sonuç raporuna göre, belirlenen tüm saldırıların neredeyse % 90'ının, kullanıcıya ait kritik bilgilerin çalınması amacını taşıdığı vurgulanmaktadır. Klavyedeki tuş basımlarının gizlice kaydedilmesi yolu (key logger)

ile çevrim içi banka hesap bilgileri gibi kritik bilgilerin çalınmasına yönelik eylemler, tüm saldırıların % 76'sını oluşturmaktadır. Bu sonuç, 2007 yılında % 72 olarak belirlenen oranla karşılaştırıldığında, yalnızca bir yıl içerisinde yaşanan artışı ortaya koymaktadır (Symantec-1, 2009). Aynı araştırmada ülkemizle ilgili dikkat çekici istatistiksel değerler ve bulgular da mevcuttur. Türkiye'deki kurumların % 50'si, yaşadıkları herhangi bir bilişim saldırısının sistemlerinin durmasına neden olduğunu belirtmiştir. Sistemi duran kurumların %50'sinde 8 saati aşan bir kesinti yaşandığı da bulgulanmıştır (Symantec-1, 2009). Herhangi bir saldırıya uğrayan kurumların % 35'i, bu saldırının belli düzeyde bir bilgi kaybına neden olduğunu belirtirken, % 10'u ise sistemlerinin yavaşladığını belirtmişlerdir (Symantec-1, 2009). Aynı araştırmanın sonuçlarına göre Türkiye'de 2008 yılında ülke genelindeki zararlı kod saldırıları; bir önceki yıla göre iki kata yakın bir düzeyde artmış, dünya genelindeki toplam zararlı kod eylemlerinin % 6'sını oluşturarak genel sıralamada Türkiye dokuzuncu sıraya yükselmiştir. Symantec firmasının bir başka araştırmasına göre, çöp (spam) e-posta eylemleri de Türkiye'de bir önceki yıla göre 12 kat artarak dünya genelinde üçüncü, Avrupa-Orta Doğu (EMEA) bölgesinde de ikinci sıraya yükselmiştir (Symantec-2, 2009). Aynı araştırmanın ülkemizle ilgili ortaya koyduğu çarpıcı sonuçlardan birisi de, 2008 yılında virüs tipindeki zararlı kodların üretildiği ve yayılma kaynağı olarak çıktığı ülkeler arasında Türkiye'nin, Avrupa-Orta Doğu bölgesi genelinde ikinci sırada yer almasıdır (Symantec-2, 2009).

Deloitte firmasının TMT (Teknoloji, Medya, Telekomünikasyon) Küresel Güvenlik Araştırması raporuna göre dünya genelindeki kurumların % 41'inin son bir yıl içerisinde kurum içinden kaynaklanan en az bir tehditle uğraşmak zorunda kaldıkları açıklanmıştır (Deloitte, 2009). Aynı araştırmadan elde edilen bir başka sonuca göre; bu kurumların % 70'inden fazlasının zararlı kod saldırısı, gene aynı orana yakın düzeyde kazalar sonucunda bilgi kaybı ve hizmet kesintisi (Denial of Service) saldırıları yaşadıkları ve zarar gördükleri (kurumların yaklaşık % 5'inin, 1 ile 5 Milyon ABD Doları arasında kayıp yaşadığı) kaydedilmiştir. Türkiye'nin de kapsama alındığı diğer bir başka uluslararası araştırmaya göre, genel amaçlı bilgi sistemlerinin kurulumunda bilgi güvenliği birimleri süreçlere büyük oranda katılırken veya kapsatılırken, insan kaynakları sistemlerinin kurulumunda bu destek ve katılımın yarı yarıya azaldığı görülmektedir (Ernst & Young, 2008). Aynı rapora göre, Türkiye'deki kurumların %

31'inin iş sürekliliğine yönelik planları olduğu ve bilgi sistemlerinin krizlere, felaketlere hazırlıklı olduğu, dünya genelinde de bu oranın % 40 civarında olduğu bulgulanmıştır (Ernst & Young, 2008).

Bilgi güvenliğini tehdit eden etmenler sadece elektronik ortamda yapılan saldırılarla sınırlı kalmamaktadır. Sel, deprem, hortum, v.b. doğal afetler veya kullanıcı hataları sonucunda da bilgiler ve bilgi sistemleri tamamen ya da kısmen zarar görebilmektedir. Özellikle kurum ve kuruluşların kuruldukları günden bu yana tüm faaliyetlerini içeren bilgilerin yok olması, tüm kurumsal belleğin bir anda silinmesi anlamına gelecektir (TBD 4. Çalışma Grubu, 2006).

Bilgi güvenliğini sağlayacak çözümleri sadece teknik sistem güvenliği veya bilişim teknolojileri ile ilişkilendirmemek gerekmektedir. Bilgi güvenliği, kurumlardaki iş süreçlerinde bilginin devreye girdiği her noktada ve her aşamada yer almaktadır. Bilgi güvenliği ve bilgi güvenliği yönetimi, insan, süreç ve teknoloji üçlüsünün birlikte uyumlu ve birbirini destekler biçimde hareket etmesiyle başarılmaktadır (Mitnick, 2005).

Son yıllarda yapılan çeşitli araştırmaların ve kurumlarda yaşanan deneyimlerin ortaya koyduğu önemli bir gerçek de, bilgi güvenliğinin başarısındaki veya başarısızlığındaki en önemli etmenin, insan faktörü olduğudur (Ashenden, 2008, Schneier, 2003, Sasse, vd., 2001). Günümüzde, kurumların bilgi güvenliğine ilişkin çözümleri etkin ve başarılı bir biçimde uygulayabilmesi için kurum çalışanlarının bilgi güvenliği konusunda bilinçli ve destekleyici olması zorunludur (Williams, 2008). Kurumlarda bilgi güvenliğinin bir kurum kültürü haline gelmesi; bilgi güvenliği yönetiminin nihai hedefidir ve bu hedefe kurumlardaki bireylerin bilgi güvenliğini sağlıklı bir biçimde algılaması ve sahiplenmesi ile ulaşılabilmektedir (Veiga ve Eloff, 2010, Zakaria, 2006). Öte yandan, bilgi güvenliğine ilişkin saldırıların başarısı ve gerçekleşen risk sonucu oluşan zararın büyüklüğü de çoğunlukla insan faktörlerine dayalı olmaktadır. Sosyal mühendislik gibi, hem saldırganın hem de kurbanın psikolojik, sosyolojik, insani etmenlerinden (saflık, aldanma eğilimi, dikkatsizlik, telaş, umursamazlık, vb.) yararlanılarak yapılan bilgi güvenliği saldırıları buna örnek olarak gösterilebilir (Mitnick, 2005).

İnsanlar; hem bilgi güvenliğini hem de buna ilişkin riskleri anlamakta ve algılamakta zorlanmakta veya yanlış yorumlarda bulunma eğilimine girebilmektedir (Schneier, 2000). Bilgi güvenliğini algılama ve değerlendirme konusundaki benzer sorun ve olgu, bireylerin günlük yaşamlarındaki her türlü güvenlik, tehdit ve risklere ilişkin durumlarda da yaşanmaktadır. İnsanın doğasından kaynaklanan bu durum, bireylerin farklı psikolojik, sosyolojik ve kültürel yapılardan gelmelerine ve buna bağlı olarak güvenliği farklı şekillerde algılayarak değerlendirmeler yapmalarına veya yanılgıya düşmelerine yol açmaktadır (Zakaria, 2006). Ayrıca, insanların güvenlikle ilgili bir olay veya konu hakkında yetersiz ön bilgi veya deneyiminin olması da sorunlara yol açmaktadır. (Schneier, 2000).

Bilgi güvenliğinde ve bilgi güvenliği risk değerlendirmesinde insan faktörünün etkili olduğu durumlar için evrensel bir model veya genel kabul görmüş bir standart henüz ortaya konmamıştır. Bunun en temel nedeni, insan temelli risklerin ve tehditlerin sayısal olarak ölçülmesinin zorluğu olarak görülmektedir (Tipton ve Krause, 2007, Stewart, 2009). Öte yandan, bu konuya ilişkin çalışmalarda; öncelikle bireylerin psikolojik özelliklerinin veya güvenlik ve riski algılamada kişilik özelliklerinin ne ölçüde ve ne şekilde etki ettiğinin belirlenmesinin gerekli olduğu son yıllarda tartışılmaya başlanan bir olgudur (Stewart, 2009, Schneier, 2008). Bilgi güvenliği risk analizi ve değerlendirmesinde bireylerin kişilik özelliklerini ve psikolojik etmenleri de katacak yeni modeller geliştirilmesi yönünde az sayıda da olsa çalışmalar yapılmaya başlanmıştır. Henüz bu çalışmalar, psikoloji ve psikiyatride günümüzde yaygın şekilde kullanılan MMPI (Minnesota Multiphasic Personality Inventory) gibi detaylı ve kapsamlı kişilik analizi testlerini kapsayacak veya bilgi güvenliği risk ölçümlerine uyarlayacak düzeye gelmemiştir. Bu konuyla ilgili literatür taramasında bulunabilen tek çalışmada, günümüzde psikologların pek tercih etmediği eski bir yöntem olan "16 kişilik faktörü" (Cattel's 16 PF Questionnaire) testinin kullanıldığı görülmektedir (Mazareanu, 2009). Tezin; "4.2. Bilgi Güvenliği Nitel Risk Anketi Uygulaması" başlıklı bölümünde daha detaylı olarak irdelenecek olan anket çalışmasında ise, Cloninger' in mizaç ve karakter boyutları tanımlamalarından yola çıkarak bu çalışma için türetilmiş basit bir kişilik analiz testi uygulanmıştır. Bu testin çok daha kapsamlı hale getirilerek gerçek anlamda psikometrik ölçüm uygulanabileceği çözümlere ilişkin yorumlar da çalışmanın "6. Öneriler" başlıklı kısımda değinilmektedir.

2.2. BİLGİ GÜVENLİĞİ YÖNETİM SİSTEMLERİ

Bilgi güvenliği yönetimi, "şirket için kritik önem taşıyan bilgilerin kontrol altında tutulması" anlamına gelir. Bilgi güvenliği yönetim sisteminin kurulum aşamasında en önemli faktörlerden biri de her yönetim sisteminde olduğu gibi şirket üst yönetiminin beyanı, somut desteği, uygun kaynakların ayrılması ve şirket içinde bilginin korunmasına yönelik başarılı uygulamaların yürütülmesini teşvik etmesidir (Tipton ve Krause, 2007).

Etkin bir bilgi güvenliği yönetim sistemi, kurumda bir güvenlik ve güvenlik yönetimi anlayışının benimsenmesini, kurumun karşı karşıya bulunduğu risklerin kabul edilebilir seviyeye indirgenmesi için politika ve yönergeler oluşturulmasını, bilgi varlıklarına yönelik meydana gelebilecek tehditlerin önceden tanımlanarak gerekli önlemlerin alınmasını sağlar. Etkili ve etkin bir bilgi güvenliği yönetim sistemi kuruluşun güvenlik ihtiyaçlarının adreslenmesini ve risk tolerans seviyesini belirlemesini sağlar (ITGI, 2008). Günümüzdeki kurumların, şirketlerin, örgütlerin; bilgi güvenliğini yeterli şekilde sağlatması, iş süreçlerini uluslararası standartlara ve bu amaçlı sertifikalara uyumlu hale getirmesi, Türkiye'deki ve yurtdışındaki ilgili ticari, sınai ve diğer yasalara, tüzük ve yönetmeliklere uyulması çok önemli bir olgu olmuştur. T.C. 5809 no.lu Elektronik Haberleşme Kanunu ve ilgili Elektronik Haberleşme Güvenliği Yönetmeliği (T.C., 2008), TCK sınai mülkiyet ve bilişim suçlarıyla ilgili kanunlar ve maddeler (TCK, 2005) ülkemizdeki hemen her şirketi veya her kurumu uygun bir Bilgi Güvenliği Yönetim Sistemi (BGYS) oluşturması veya BGYS' ye uygun hareket etmesi, işlerini yürütürken BGYS' ye uygun çalışması, destek vermesi ve farkındalık sahibi olmasını sorumlu kılmaktadır.

Daha önceki bölümde de belirtildiği gibi, bilgi güvenliği sadece teknolojik konuları içermemektedir. Bilginin durduğu, saklandığı, paylaşıldığı, işlendiği, iletildiği her türlü kaynak, ortam ve sistemi içermektedir. Aynı zamanda yazılı kâğıt belge ve sözlü bilginin de güvenliğinin sağlanması gerekmektedir. Bunu sağlarken de, sadece teknolojik yaklaşım değil, insan ve iş süreçlerini de katacak şekilde etkin bir bilgi güvenliği yönetimi sağlatılmalıdır (Schneier, 2000).

2.2.1. Uluslararası Bilgi Güvenliği Yönetim Sistemleri Standardı

Bilginin güvenliği son yıllara kadar işletmelerin Bilgi Teknolojileri Departmanlarına verilmiş bir görev olarak algılanmaktaydı. Ancak, günümüzde bilginin korunması bir işletmenin tüm bölümlerinin ve tüm çalışanlarının görevidir. Bunun gerçekleşmesi ise bilgi güvenliği yönetim sisteminin uygulanması ve sürdürülmesi ile mümkün olmaktadır (Schneier, 2008).

Bilgi güvenliği yönetimi alanında, günümüzde birçok farklı standart, ölçümler ve kurallar bulunmaktadır. Örnek vermek gerekirse, Sarbanes-Oxley Kanunu, HIPAA kanunu, COBIT' in bilgi teknolojilerinin güvenliği bölümleri, E-ticarette PCI standardı, NIST, FIPS, vb sayılabilir. Bunların içinde en yaygın olarak kabul gören ve uygulananlarından birisi, uluslararası ISO örgütünce tanımlanmış olan ISO 27001 Bilgi Güvenliği Yönetim Sistemleri standardıdır. İlk adıyla BS 7799 ve daha sonraları ISO 17799 adını almış, son olarak ISO/IEC 27001 olarak güncellenmiştir (ISO/IEC, 2005).

ISO 27001'de 11 temel bilgi güvenliği alanı, bir başka deyişle 11 temel bilgi güvenliği kontrol amacı bulunmaktadır (TSE, 2006);

- Güvenlik Politikası
- Bilgi Güvenliği Organizasyonu
- Varlık Yönetimi
- İnsan Kaynakları Güvenliği
- Fiziksel ve Çevresel Güvenlik
- Haberleşme ve İşletim Güvenliği
- Erişim Kontrolü
- Bilgi Sistemleri Edinim, Geliştirme ve Bakımı
- Bilgi Güvenliği İhlal Olayı Yönetimi
- İş Sürekliliği Yönetimi
- Yasal Uyum

Bu 11 ana alana bağlı olarak 39 kontrol amacı ve toplam 133 kontrol maddesinden oluşmaktadır. ISO 27001 BGYS' nin uygulanmasında süreç yönetimi yaklaşımının benimsenmesi ile özellikle aşağıdaki konularda yarar sağlanmaktadır:

- Süreçlerin tanımlanması
- Süreçlerin belirlenmesi yoluyla firma varlıklarının ve dolayısıyla korunması hedeflenen bilginin belirlenmesi
- Bilgi Güvenliği Yönetim Sistemi kapsamının belirlenmesi
- Bilgi Güvenliği Yönetim Sistemi politikasının ve hedeflerinin belirlenmesi
- Varlık sahiplerinin ve sorumluluklarının belirlenmesi
- Risklerin yönetimi için uygulanacak kontrollerin belirlenmesi
- Bilgi Güvenliği Yönetim Sistemi'nin etkinliğinin ölçülmesi
- İzleme ve ölçüm sonuçlarına bağlı olarak sürekli iyileştirmenin yapılması

"Planla, Uygula, Kontrol Et, Önlem Al" (PUKÖ) modeli, Bilgi Güvenliği Yönetim Sistemleri için temel alınmıştır ve Şekil 2.1.' de görüleceği gibi bir yaşam döngüsü şeklinde betimlenmektedir (TSE, 2006). ISO 27001 açısından bu döngü incelendiğinde standart gereklilikleri ve kapsamı aşağıdaki gibi tanımlanmaktadır.

"Planla" aşamasının kapsamı:

- BGYS kapsamının belirlenmesi
- BGYS politikasının belirlenmesi
- Hedeflerin belirlenmesi
- Risklerin belirlenmesi
- Risklerin değerlendirilmesi
- Kontrol hedeflerinin ve kontrollerin seçilmesi
- Uygulanabilirlik Şartnamesinin hazırlanması

"Uygula" aşamasının kapsamı:

- Risk İşleme/Sağaltım (Treatment) Planının hazırlanması.

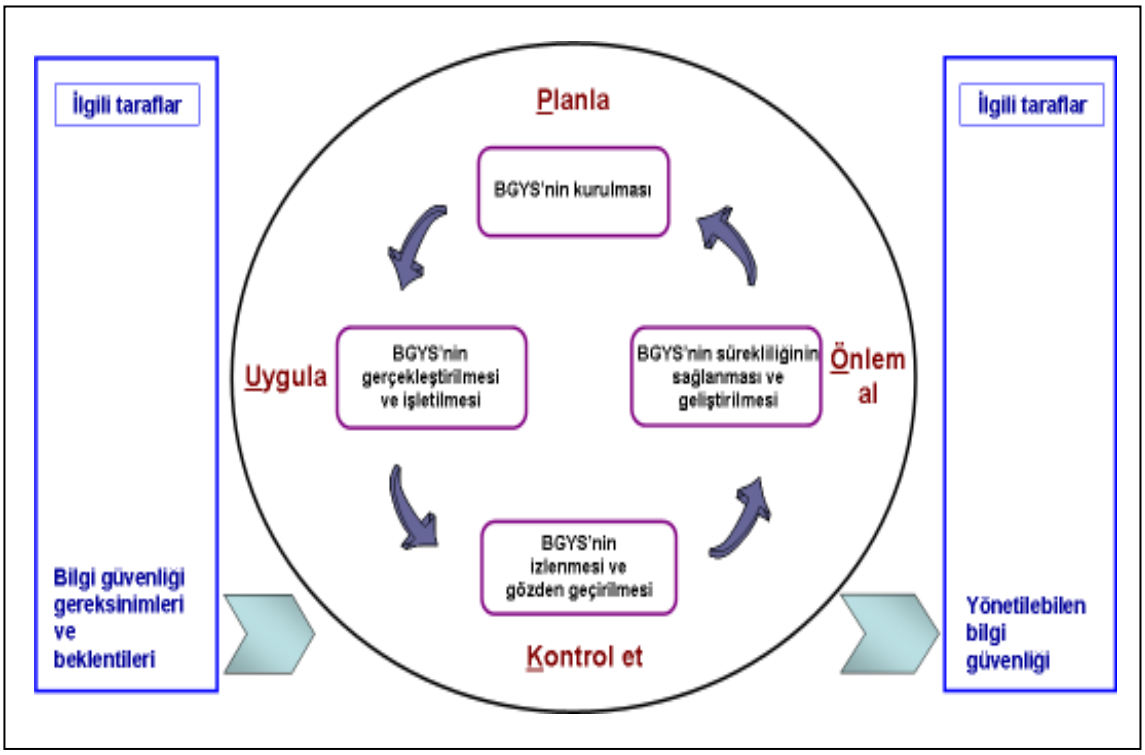
- Risk İşleme Planının uygulanması.
- Kontrol hedeflerine göre seçilmiş kontrollerin uygulanması.

"Kontrol Et" aşamasının kapsamı:

- Gözetim süreçlerinin uygulanması
- Planlanmış aralıklarla BGYS iç denetimlerinin gerçekleştirilmesi
- BGYS' nin etkinliğinin belli aralıklarla ölçülmesi ve değerlendirilmesi
- Kalan risk seviyesinin ve kabul edilebilir risk seviyesinin gözden geçirilmesi

"Önlem Alarak İyileştir" aşamasının kapsamı:

- Belirlenmiş iyileştirmelerin uygulanması
- Gerekli düzeltici ve önleyici faaliyetlerin uygulanması
- Sonuçların ve yapılan işlemlerin iletişiminin sağlanması
- İyileştirmelerin düşünülmüş hedeflere ulaşılıp ulaşılmadığının değerlendirilmesi



Şekil 2.1. BGYS için PUKÖ Modeli

Bilginin yaygınlaşması, paylaşılması ve bilgi temelli hizmetlerin artması karşısında hız kazanan güvenlik problemlerine uygulanabilir bir standart ve yöntem arayışına karşı geliştirilmiş olan uluslararası ISO/IEC 27001:2005 Bilgi Güvenlik Yönetimi Standardı, sektördeki bağımsız her alana uygulanabilmektedir. Bu standart ile bu standarda bağlı olarak alınacak belge (sertifika) ile şirket ve kurumlar; bilgi varlıklarını en üst düzeyde korumayı, doğru ve gerekli bilgi güvenlik yatırımları yapmayı, tehlike oluşmadan önlemlerini almayı ve kurumsal bilgi yönetimini devreye alarak bilgi güvenliğinde sürekliliği sağlamayı hedeflemektedirler. Bu nedenle, kurumların, işletmelerin ve şirketlerin ISO 27001 belgelendirmesine olan talebi gün geçtikçe artmaktadır. Türkiye'de de, dünyada da birçok farklı devlet kurumu ve özel sektör kuruluşu bu sertifikayı almakta ve bu sertifikanın değeri ve bilinilirliği gün geçtikçe ülkemizde ve dünyada artmaktadır (Richardson, 2008).

ISO 27001 gibi diğer bilgi güvenliği yönetimi sistemleri standartları da risk odaklı bir yaklaşıma sahiptir. Bilgi güvenliğini sağlarken, önce o kurumdaki bilgi değeri olan kaynaklar, varlıklar ve içerdikleri bilginin değeri belirlenmeli, sonra da bu kaynaklara özgü zayıflıklar ve bu zayıflıklardan yararlanabilecek tüm olası tehditler analiz edilerek belirlenmelidir. Bu tehditlerin gerçekleşme olasılığı da ilgili riskin değerinin hesaplanmasını sağlayacaktır. Daha sonra da, bu riskle baş etmeye / risk işlemesine (risk treatment) yönelik alternatif çözümler belirlenir ve bu çözümlerden hangisinin en etkin ve en az maliyetli olacağı üzerinde çalışılır, ardından da karar verilen en uygun çözümler uygulanarak riskin ne kadar azaldığı tekrardan uygulanacak denetimlerle ölçülür.

ISO 27001 ve diğer bilgi güvenliği yönetimi sistemleri ve ilgili yöntemlerinden hangisi kullanılırsa kullanılsın, kurumlarda bilgi güvenliği yönetimi sistemi ve yapısını kurmanın çok büyük yararları olduğu bilinmektedir. Bu yararlar aşağıda maddeler halinde verilmektedir (Calder ve Watkins, 2008):

- İş sürekliliğini sağlamak, meydana gelebilecek zararı en aza indirebilmek, kazancı ve iş fırsatlarını arttırabilmek amacıyla bilginin birçok tehlikeye karşı korunmasını sağlar.

• Kuruluşun kurumsal değerlerini, yatırımlarını ve hedeflerini sürdürüp koruyabilmesi için ortaya konması gereken kontrollerin firma içinde yerleştirilmesi ve uygulanması yolunda temel teşkil eder.

• Olumlu izlenim sağlar ve rakip şirketlerin önüne geçmede katkı sağlar.

• Kurum müşterileri, bilgilerinin güvende tutulacağı konusunda kurumun beyanından dolayı kendilerini güvende hisseder.

• Tek bir bilgi güvenlik ihlalinin bile çıkaracağı zarar ve masraf çok büyük olabilir. Belgelendirme işlemi, maruz kalınabilecek bu tür masrafları azaltır ve bu da ilgili iş alanındaki yatırımcılar ve müşteriler için olumlu bir izlenim sağlar.

• Bilgi güvenliği ile ilgili sigorta primlerinde düşüş sağlanır.

• Kurumun, ilgili geçerli tüm kanun ve tüzüklere uygunluğunun yetkili makamlara kanıtlanmasına yardımcı olur.

• Organizasyonun tüm aşamalarında ve birimlerinde güvenliğe bağlılığın sağlanması ve kanıtlanmasında yardımcı olur.

• Kurum içinde, bilgi sistemleri ve iş süreçlerindeki bilgi zayıflıklarının nasıl korunacağı konusundaki farkındalık artar.

• Bilgiye ve sistemlere daha güvenilir erişim sağlanır.

• Çalışanların kuruluş içerisindeki sorumlulukları ve bilgi güvenliği konularındaki bilinçlerinin artmasını sağlar.

• Düzenli olarak gerçekleştirilen denetimler, sistemlerin güvenliğini ve etkinliğini izlemeye ve iyileştirmeye yardımcı olur.

• Gizlilik sağlanır.

• Bilginin sadece yetkili kişiler tarafından erişilebilir olması sağlanır.

• Bütünlük sağlanır.

• Bilginin ve işlem metotlarının doğruluğunun ve bütünlüğünün korunması ve içeriğinin değişmemesi sağlanır.

• Bilgiye ve sistemlere gerektiği her anda yetkililerin ulaşabilmesi sağlanır.

- Başarılı bir e-ticaret için gerekli olan işlevsellik, güvenlik, güvenilirlik ve veri korumasına ilişkin gereklilikler sağlanır.
- Bilgi güvenliğinin, o şirket veya örgütte kurum kültürünün bir parçası haline gelmesi ve kalıcı olması sağlanır.
- İş süreçlerinde risk azalımı yanı sıra, iş süreçlerinde kalite de artar; kalite güvencesine, toplam kalite yönetimine de katkı sağlar.
- Yönetimsel yapılarda iş etkinliğini ve kaliteyi arttıracak değişimlere de katkı sağlar.
- Personelin iş yapış şekillerini iyileştirerek performanslarını arttırır.
- Personelin bireysel beceri, iş ve kariyer gelişimlerine katkıda bulunur.

Daha önce de vurgulandığı şekilde, bilgi güvenliği yönetiminin uluslararası en geçerli standardı ve belgelendirmesi ISO 27001 ve ona bağlı uygulama standartları olan ISO 27002, ISO 27003, ISO 27004 ve ISO 27005'tir. ISO standartları dışında da bilgi güvenliği ve bilgi güvenliği yönetimi sistemi yapılanması için birçok değişik yöntem, mimari tanımlamaları, ölçütler ve standartlar bulunmaktadır. Öte yandan, güvenliğe teknolojik değil, süreçsel olarak bakan ve sadece bilgi teknolojilerinin değil, tüm iş süreçlerinde bilginin güvenliğine odaklanan ve bunun etkin, sürekli, yönetilebilir olması için, tam anlamıyla tepeden yönetim bakış açısı ve yetkinliğine sahip olan günümüzdeki en geçerli standart, ISO 27001 olarak kabul edilmektedir.

2.2.2. Bilgi Güvenliği Yönetim Sistemlerindeki Diğer Modeller

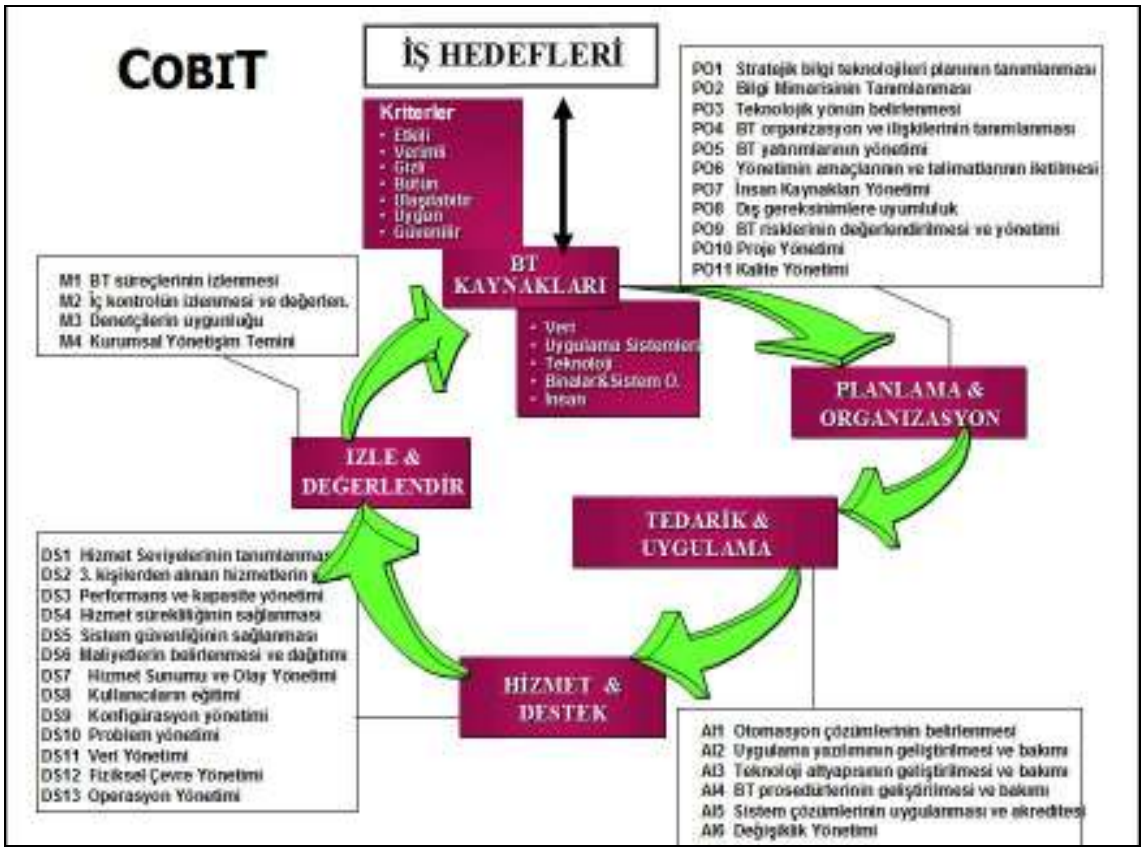
Günümüzde ülkemizde ve dünya genelinde kullanılan belli başlı diğer bilgi güvenliği yönetimi standart veya modelleri arasında COBIT 4.1'in güvenlikle ilgili kısımları ve PCI v1.1 (E-ticaret ve kredi kartı güvenliği standardının en güncel sürümü) örnek verilebilir. "Bilgi ve İlgili Teknolojiler için Kontrol Hedefleri" olarak tanımlanan COBIT (The Control Objectives for Information and related Technology) tüm iş süreçlerini değil, sadece bilgi teknolojileri ve ilgili bilişim süreçlerini kapsamına alır.

COBIT bir standarttan daha çok, ilkeler, kontrol hedefleri, ölçütleri, denetim kuralları ve bunların yapıtaşlarını tanımlayan uluslararası bir BT ilkeleri mimarisi ve yardımcı el kitabıdır. Aynı zamanda COBIT, bilgi teknolojileri kontrol hedefleri ve denetim

rehberinin yanı sıra uygulanabilir bir model olabilmek ve yönetime yön verebilmek için gerekli araçları da barındıran bir yöntem bilimi olarak da tanımlanmaktadır. (ITGI, 2007). ISO 27001'den bir başka temel farkı ise, bilgi teknolojilerinde sadece güvenliği değil, başka ana ölçüt ve hedefleri de içermesidir. Toplamda yedi tane olan bu bilişim kontrol kıstasları "Gizlilik", "Bütünlük", "Kullanılabilirlik", "Etkinlik", "Verimlilik", "Güvenilirlik" ve "Uyum"' dur (ITGI, 2007). ISO 27001 'de ise, sadece üçü yani, "Gizlilik", "Bütünlük", "Kullanılabilirlik" ilkeleri mevcuttur. COBIT; insan, süreç, altyapı uygulamaları, kurumlarda bilgi teknolojilerinde iş süreçlerinin olgunluğu ve bilişim süreçlerine hükmetme hedefleri doğrultusunda neyi hangi düzeyde ve hangi ölçütlerle, hangi birimin, kimin, ne sıklıkta, neye bakarak yapması gerektiğini anlatır. 34 üst düzey iş süreci (ana kontrol) ve bunlara bağlı 210 adet kontrol hedefinden oluşur. Bu 34 üst düzey süreç 4 ana gruba dağıtılmış şekilde ayrılmaktadır (ITGI, 2007);

- Planlama ve Organizasyon (Plan & Organize; PO)
- Tedarik ve Uygulama (Acquire & Implement; AI)
- Hizmet Sunumu ve Destek (Deliver & Support; DS)
- İzle ve Değerlendir (Monitor & Evaluate; ME)

COBIT, diğer iç denetim ve bilgi teknolojileri risk ve kontrol yöntemlerinden bağımsız olarak hazırlanmamış, aksine oluşumunda mevcut yöntem bilimler dikkate alınmıştır. Mevcut ve genel kabul görmüş iç denetim yöntemlerinin bilgi teknolojileri kontrol ihtiyaçlarını karşılayan yönleri COBIT' e eklenmiştir. COBIT özgün içeriğinin yanı sıra yüksek kalite güvencesinin sağlanabilmesi için çok geniş bir katılımla oluşturulmuş ve çeşitli gözden geçirme aşamalarına tabi tutulmuştur. Şekil 2.2.'de COBIT mimarisinin, kurumlarda bir yaşam döngüsü şeklinde yönetimsel yapısı ve ilgili modeli gösterilmektedir (ITGI, 2007).



Şekil 2.2. COBIT Yönetim Modeli

COBIT, bilgi teknolojileri süreçlerini iş süreçlerinin destekçisi olarak görmekte ve aralarındaki ilişkileri bire bir ortaya koymaktadır. Böylece bilgi teknolojileri kaynaklarının iş stratejileri doğrultusunda etkin biçimde kullanılmasına olanak sağlamaktadır. Kurumların bilgi teknolojileri politikasına temel olabilme özelliğinin yanında yönetim odaklı olması nedeniyle kurum içi kullanım için son derece uygundur. COBIT' in bakış açısı ise bilgi teknolojileri yönetişimi (IT governance) için evrensel bir model olmak ve ilgili mimariyi ve yöntemi tanımlamaktır (ITGI, 2007).

COBIT, "Ne?" ve "Hangi kontroller?" sorularına ve 0 ile 5 arası skaladaki bir ölçümlemeyle "Kurumda nasıl uygulanıyor? Örnek var mı?", "Ne kadar başarılı?" "Ne kadar uygun?" sorularına, tüm bilişim süreçleri ve bilgi teknolojileri konuları açısından bakmaktadır. ISO 27001 ise "Ne?" ve "Hangi kontroller?" sorularına tüm iş süreçlerinin bilgi güvenliği açısından yanıt vermektedir. ISO 27002 ise bu konulara öneri tabanlı uygulama örnekleri ve kısmen "Nasıl yapılabilir?" sorusuna yanıt verecek şekilde

yaklaşmaktadır. ISO/IEC 27002, ilgili literatürde Uluslararası Bilgi Güvenliği Yönetimi için Uygulama Kuralları Standardı olarak adlandırılmaktadır (ISO/IEC, 2005).

Kartlı Ödeme için Veri Güvenliği Endüstri Standardı (PCI DSS - Payment Card Industry Data Security Standard) ise, uluslararası düzeyde en geçerli elektronik ticaret güvenliği standartlarından birisidir. PCI DSS; e-ticaret güvenliği ve ilgili sayısal bilgilerin ve bireylerin kredi kartları bilgilerinin korunması için temel gerekler, e-ticaret yapan ticari kurumların sorumlulukları, uyması gereken teknik ve süreçsel standartlar, kurallar ve cezai yaptırımları da içeren bir standarttır. En son sürümü, PCI v.1.1 adıyla 2006 yılında yürürlüğe girmiştir (PCI Security Standards Council, 2006). Elektronik kart ödeme ve üretimindeki dünyanın en büyük iki kuruluşunun destek verdiği ve bu kuruluşlardan temsilcilerin oluşturduğu bir konsey tarafından ortak hazırlanıp belirlenen bir standart olup, dünyadaki tüm e-ticaret yapan veya kredi kartı bilgilerini Internet ve diğer elektronik ortamlarda saklayan, ileten, taşıyan tüm kurumları sorumlu kılmaktadır. PCI DSS, ISO ve COBIT' ten farklı olarak bazı güvenlik maddelerinde, teknolojik detaylara da iner ve teknik standart kitabı mantığına yakındır. Ayrıca cezai yaptırımların parasal bedellerini, hangi hacimde ticaret yapan kurumun hangi güvenlik teknik düzeylerden sorumlu olduğunu da sayısal ölçütlerle tanımlar. PCI DSS' in diğer bir farkı da, sadece e-ticaret ve kredi kartı, vb ödeme kartlarının güvenliğine odaklanmasıdır; e-ticaret dışı konular ve e-ticaret yapmayan kurumlar PCI DSS' in kapsamı dışındadır (PCI Security Standards Council, 2006).

Bu standart ve ölçüt kitaplarına ek olarak, başka çeşitli bilgi güvenliği standartları ve yönetmelikleri de mevcuttur. Bunlar içinde; ITIL (sadece güvenlik süreçleriyle ilgili olan kısmı), FIPS, GASSP, NIST, Common Criteria en çok bilinen ve kullanılanları olarak kabul edilmektedir. Ayrıca, ülkemizde de özellikle bankacılık sektörüne yönelik bilgi sistemleri ve bilgi güvenliğine ilişkin bazı yönetmelikler bulunmaktadır. T.C. Bankacılık Düzenleme ve Denetleme Kurumu'nun (BDDK) "Bankalarda Bilgi Sistemleri Yönetiminde Esas Alınacak İlkelere İlişkin Tebliğ"'i ve diğer ilgili bilgi sistemleri yönetmelikleri bunlara örnek olarak gösterilmektedir (BDDK, 2007).

Bilgi güvenliği yönetiminin ana konularından birisi olan iş sürekliliği ise, son yıllarda artık sadece bilginin kesintisizliği ve bilginin sürekliliğinin güvenceye alınması olarak görülmemekte, her türlü işletimsel etkinliğin de kapsandığı başlı başına ayrı bir yönetim

konusu ve alanına dönüşmektedir (British Standards Institute, 2006, Business Continuity Institute, 2007).

2.3. BİLGİ GÜVENLİĞİ RİSKLERİ

Kurumlarda bilgi güvenliği yönetimi sisteminin etkin ve başarılı bir biçimde oluşturulması ve yönetilmesinde, risk yönetimi zorunlu ve hayati yapı taşlarından birisi olmaktadır. ISO 27001 başta olmak üzere, ilgili tüm bilgi güvenliği standartları ve yönetmeliklerinde de, bilgi güvenliği risk analizi, ölçümü ve değerlendirmesi en öncelikli ve önemli aşamalardan birisi olarak kabul edilmektedir (Dhillon, 2007). Bu bölümde, önce genel olarak risk kavramına değinilmiş, ardından da bilgi güvenliği risklerine ilişkin çeşitli model ve yöntemler aktarılmıştır.

2.3.1. Risk Kavramı ve Riskin Temel Bileşenleri

Risk, Türk Dil Kurumu Sözlüğü' nün 2005 yılına ait baskısında zarara uğrama tehlikesi olarak tanımlanmıştır (TDK, 2005). Bir başka sözcük tanımında ise risk; bir zarara, kayba, tehlikeye yol açabilecek bir olayın ortaya çıkma olasılığı olarak ifade edilmektedir (Meydan Larousse, 1998). Risk sözcüğünün kökeninin Arapça rızık ya da Latince risicum sözcüklerinden geldiği düşünülmektedir (Tevfik, 1997). Risk için çok çeşitli ve daha farklı tanımlara da rastlamak mümkündür. Bu farklılık, kullanılan risk yönetim metodunun, riski algılama biçiminden kaynaklanmaktadır. Ancak hangi tanım olursa olsun, risk kavramı ile iki temel özellik anlatılmaktadır. Bu özellikler; kayıp ve olasılıktır. Kayıp ifadesiyle; zaman, para, itibar, kalite gibi konularda karşılaşılacak zararlar ve tehlikeler; olasılık ifadesi ile de bu zararların oluşma potansiyeli anlatılmaktadır.

Risk; belirli bir zaman aralığında hedeflenen bir sonuca ulaşmada, kayba ya da zarara uğrama olasılığıdır, ya da bir olayın beklenenden farklı olarak gerçekleşebilme olanağıdır. Olabilecek sonuçların sayısının artması ile risk meydana gelir. Riskin var olması demek bir olayın sonucunun tam olarak tahmin edilemeyeceği demektir. Risk, gelecekte oluşabilecek potansiyel sorunlara, tehdit ve tehlikelere işaret eder. Risk genellikle tam ve net olarak bilinemez ya da öngörülemez, belirsizdir. Belirsizliğin

sonuç üzerinde olumsuz etkileri vardır. Riskin öznel ve nesnel yanları olduğunu savunanlar ve bu yargıya karşı çıkanlar olmasına rağmen, genel olarak riskin nesnel ve ölçülebilen bir faktör olduğu söylenebilir. Risk, nesnel ve ölçülebilen bir faktör olduğu sürece yönetilebilir bir olgudur. Riskin temel bileşenleri, oluşma olasılığı ve oluşması durumunda sonucu ne ölçüde etkileyeceğidir. Riskler birbiriyle etkileşim içerisinde olan 3 temel alanda ele alınır (Fıkırkoca, 2003):

- Teknik risk, hedeflenen (tahmin edilen ve planlanan) başarım değerine ulaşamamanın bir ölçüsüdür. Teknik riskler, maliyet ve çizelge risklerinin temel nedenidir.
- Maliyet riski, tahmin edilen ve planlanan maliyet değerinin aşılması durumudur. Örneğin ekonomik koşullardaki belirsizlikler önemli maliyet risk kaynaklarından biridir.
- Çizelge riski ise bir işin tahmin edilen ve planlanan sürede gerçekleştirilememesinin bir ölçüsüdür.

Risk yönetimi, risklerin oluşma olasılıklarını veya oluşan risklerin etkilerini azaltacak ya da riskleri tamamen ortadan kaldıracak her türlü eylem olarak tanımlanabilir. Bir başka deyişle, risk yönetimi karar vericilerin riski azaltmak veya ortadan kaldırmak üzere yararlandıkları yol veya yöntemlerdir (Cadoğlu, 2000). Bu tanım, risk yönetiminin hedeflerini de ortaya koymaktadır. Risk yönetimi iki önemli aşamadan oluşur. Mevcut durumun incelenmesi, bilgi varlıklarının gizlilik, bütünlük ve kullanılabilirlik bileşenlerine karşı zarar verebilecek tehditlerin belirlenmesi ve bu tehditlerin kullanabilecekleri zayıf noktaların ortaya listelenmesinden oluşan risk analizi birinci aşamadır. İkinci aşamada risk alma isteğine bağlı olarak risk işlemesi süreci yer alır. Risk işlemesi içinde dört farklı seçenek yer almaktadır. Riskten kaçınma, riske neden olan tehdidin ortadan kaldırılarak riskin tümüyle yok edilmesidir. Riski azaltma, ilgili tehdidin ortaya çıkma olasılığını ve/veya ortaya çıkması durumunda kuruma olan olumsuz etkisinin azaltılmasıdır. Riski transfer etme, dış kaynak kullanımı ya da sigortalama gibi yöntemlerle ilgili riskin sorumluluğunu bir başkasına devretme anlamına gelmektedir (ISO / IEC, 2005).

Gerçek risk, ya da teknik kullanımı ile “kalan risk”, varlığın ve ortamın kendisinden kaynaklanan yapısal risk üzerine risk tedavi seçeneklerinin uygulanmasından geride

kalan anlamına gelmektedir. Yapısal risk ise, söz konusu riskin gerçekleşme olasılığı ve olasılıktan bağımsız her bir ortaya çıkışında kuruma vereceği zarar ile hesaplanabilir. Her kurum için risk analizi ve riski ele alma, tedavi etme yöntemleri farklılıklar gösterebilir. Önemli olan kurum kültürüne, boyutuna, çalışanlarına, sektörüne ve önem verdiği değerlere uygun bir yapının kurulmasıdır. Doğru şekilde yaratılacak risk yönetim yöntem bilimi kurumun ihtiyaçlarına çabuk ve doğru cevaplar verirken, bu amaçla harcanacak zaman ve kaynakların verimli kullanımını sağlayacaktır (Finne, 2008).

Risk yönetimi uygulama yöntemleri; nicel ve nitel yaklaşımın her ikisinde de iki temel alt sürece dayanmaktadır. Bu iki alt süreç; risklerin değerlendirilmesi ve çözümlenmesidir. Risklerin değerlendirilmesi, hangi risklerin hangi olasılıkla oluşacağını ve hangi risklerin hangi öncelikle çözümlenmeye değer olduğunun belirlenmesini içermektedir. Risklerin çözümlenmesi süreci ise, çözümleyici uygulamaların planlanması ve icrasını içermektedir. Bu kapsamda; her iki alt süreç de farklı alt süreçlere ayrılarak, risk yönetimine ilişkin metotları belirlemektedir. Konuyla ilgili olarak alt süreçlere göre farklılık gösteren bazı farklı yöntemlerden en yaygın olanlarından bir tanesi kısaca bu bölümde özetlenmektedir.

Proje Yönetimi Bilgi Dağarcığı'nda (PMBOK – Project Management Body of Knowledge) tanımlanan yöntem bilimine göre risk yönetim süreci yedi adımı içermektedir (PMI, 2008). Bu adımlar, sırasıyla aşağıdaki biçimde sıralanabilir:

− Risk planlaması (risk planning),

− Risk tanımlaması (risk identification),

− Risk değerlendirmesi (risk assessment),

− Riskler için strateji geliştirilmesi (risk strategies),

− Risklerin gözlenmesi ve denetimi (risk monitoring and control),

− Risk eylemi (risk response),

− Risklerin tekrar değerlendirilmesi (risk evaluation).

Proje Yönetimi Bilgi Dağarcığı'ndaki (PMBOK) yöntem bilimi, bilgi teknolojilerinde risk yönetimi veya bilişim projelerinde risk yönetiminde de en çok kullanılan yapı olarak benimsenmektedir (Marchewka, 2009).

Risk yönetimindeki önemli bir konu, risklerin ölçülmesi ve değerlendirilmesi aşamasıdır. Risk hesaplaması veya değerlendirmesinde iki temel yaklaşım bulunmaktadır. Bunlardan birincisi nicel (quantitative) yöntemdir. Bu yöntemde sayısal veriler, ölçümlenmiş, gözlemlenmiş gerçek finansal maliyetler ve olasılık değerlerinden yola çıkarak değerlendirme ve hesaplamalar yapılır. Diğer temel yaklaşım veya yöntem ise nitel (qualitative) olarak tanımlanmaktadır. Nitel risk ölçüm ve değerlendirmesinde gerçek sayısal değerler, finansal maliyetler ve daha önceden hesaplanmış istatistiksel gözlemlere bağlı olasılık ölçümleri yerine, öznel değerlendirmeye bağlı olan, göreceli, sayısal olmayan ölçeksel ağırlıklar veya puanlamalar kullanılır. Bazı durumlarda da bu iki yöntemin belli özellikleri bir araya getirilip melez yöntemler uygulanmaktadır (Peltier, 2001). Bilgi güvenliği risk değerlendirmesinde de benzer yaklaşımlar uygulanmakta olup, ilgili nitel ve nicel yöntemler Bölüm 2.3.4 ve 2.3.5' de daha detaylı olarak irdelenecektir.

Bilgi güvenliği risk yönetimine ilişkin bazı kaynaklarda ise, risk değerlendirmesini de içeren ilgili aşamalar risk analizi başlığı altında toplanmaktadır. Risk analizi, stratejik kararlarda riskin kapsamlı olarak anlaşılmasını sağlayan yöntemlerin bütünü olarak tanımlanmaktadır (Tevfik, 1997). Risk analizinde önemli olan, riske ilişkin değişkenin kestiriminin olasılık dağılımı biçiminde ortaya konmasıdır. Bu olasılılık dağılımını elde etmede iki temel yöntem ve model vardır (Tevfik, 1997):

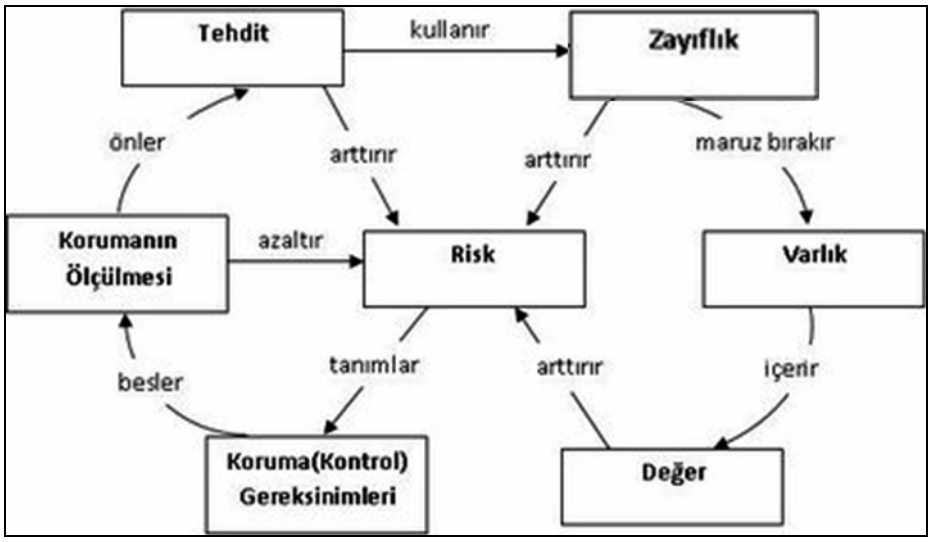
- Analitik yöntemler: Bireysel kestirimler matematiksel yöntemlerle birleştirilir.
- Simülasyon yöntemi: Monte-Carlo, vb simülasyon yöntemi ve yapısal modeller kullanılır.

Günümüzde, bu yöntemlere alternatif olabilecek risk analiz yöntem bilimleri geliştirilmeye çalışılmaktadır.

2.3.2. Bilgi Güvenliği Risklerinde Temel Kavramlar

Bilgi güvenliğine ilişkin risklerin analizi, ölçümü ve değerlendirmesinde 1980'li yılların ortasından itibaren çeşitli yöntemler, yaklaşımlar ve modeller ortaya konmuştur ve günümüzde de yaygın olarak kullanılmaktadır. Richard Baskerville, bilgi sistemlerinin güvenlik amaçlı risk analizinde kontrol listeleri (checklists) araçlarından yararlanarak bir yöntem geliştirmiştir (Baskerville, 1993). Yapılan bazı çalışmalarda risk analizinin, bilgi sistemlerinin güvenli bir şekilde tasarımı için kullanımına yönelik ilk modeller ortaya konmuştur (Fisher, 1984). Ayrıca, benzer bir başka çalışmada Donn B. Parker, nitel risk hesaplaması için etki ve olabilirlik çarpımını ilk olarak literatürde tanımlamıştır (Parker, 1981). Bu çalışmaların tümünde de, kontrol listelerinin çok yoğun bir şekilde, fazla sayıda kullanılması ve risk analizinin bu listelere dayalı olması dikkati çeken ortak bir özelliktir. Bu da, sürekli güncelleme, tutarlılık, verimli kullanım ve yönetimsel zorlukları beraberinde getirmektedir. Diğer bazı araştırmacılar, buna alternatif olarak bilgi güvenliği risk analizi için mantıksal bir model oluşturmaya çalışmışlardır (Backhouse ve Dhillon, 1996). Anderson, Longley ve Kwok; ilk defa ilgili çevreye ve koşullara ait tehditlerin belirlenip tanımlanmasından yola çıkan güvenlik risk analizi modelini ortaya koymuştur (Anderson, vd., 1994).

Bilgi güvenliğinde riskin ortak tanımı şu şekilde ifade edilmektedir. Bir bilgi varlığına (asset), o varlığın zayıflıklarından (vulnerabilities) yararlanan tehditlerin (threats) ortaya çıkardığı veya gerçekleştirdiği kayıp (loss) veya etki (impact) bileşkesidir (Peltier, 2001). Ayrıca bu riskin gerçekleşmesi, ilgili riskin ne olasılıkla gerçekleşeceği, bir başka deyişle, ilgili tehdidin o varlığın zayıflığından yararlanması durumunun hangi olasılıkla (likelihood, probability) gerçekleşebileceğine bağlıdır (Layton, 2007). Bilgi güvenliğinde riske ilişkin kavramlar ve etkileşimleri Şekil 2.3.'te gösterilmektedir (Farn, vd., 2004).



Şekil 2.3. Bilgi güvenliğinde temel risk kavramları ve etkileşimleri

Bilgi güvenliği ve bilgi güvenliği yönetimi, risk odaklı bir yaklaşım içermektedir. Bilgi güvenliğinin en öncelikli aşaması; kurumların bilgi envanterindeki varlıkların gizliliğini, bütünlüğünü, kullanılabilirliğini tehdit eden risklerin tanımlanıp, bu konuda risk yönetimi gereklerinin yapılmasıdır (ISO/IEC, 2005). Bilgilerin karşı karşıya kalabileceği tehditler bilgilerin önemine, tehlikenin olabilirliğine ve gerçekleştiğindeki etkisine göre çeşitli risk işleme (treatment) seçeneklerinden birisi üzerinden gidilerek riskler yönetilir ve kontrol edilir. Dört temel risk işleme seçeneği bulunmaktadır (Layton, 2007):

- Riskin azaltılması,
- Riskin kabul edilmesi,
- Tamamen üçüncü bir tarafa devredilmesi (sigorta etmek gibi)
- Risk kaynağının yok edilmesi

Bilgi güvenliği risklerine ilişkin olarak, ilgili bilgi varlıklarının belirlenmesi gerekmektedir. Bilgi varlıkları çok çeşitli olabilmektedir, ama genelde temel bazı gruplara ayrılarak sınıflandırılmaktadır. Bunlar; insan, elektronik veri, fiziksel altyapı, kağıt belgeler, iş süreçleri, yazılımlar, bilişim teknolojileri olarak tanımlanmaktadır (Landoll, 2006). Bilgi güvenliğinde tehditler; fiziksel, elektronik, sayısal nitelikli veya doğal afetler de olmak üzere çok çeşitli olabilmektedir. Çizelge 2.1.'de örnek olması amacıyla belli başlı bilgi güvenliği tehditleri listelenmektedir (Layton, 2007).

Çizelge 2.1. Bilgi güvenliği tehditlerine örnekler

Gizli Kanallardan Bilgi Sızdırılması
Kaza veya arızalardan oluşabilecek hasar
Malzemelerin tahrip edilmesi
Deprem
Su baskını
Gizli bilginin ortaya çıkması
Zararlı kodlar (bilgisayar virüsleri, vb.)
Tozlanma, kirlenme
Yangın
Bilgisayara yetkisiz erişim
Ekipmana yetkisiz erişim
Kullanıcı hataları
Elektrik kesintisi
Bilişim suçluları (hackers, cyber criminals)
Aygıtların kaybolması
Art niyetli personel

Bilgi güvenliği risklerine ilişkin zayıflıklar da, ilgili bilgi varlığının zayıflığıyla ilintilidir ve bunlar da tehditler kadar çok olmasa da çeşitlilik göstermektedir. Çizelge 2.2.'de örnek olması amacıyla belli başlı bilgi güvenliği zayıflıkları listelenmektedir (ISO/IEC, 2008).

Çizelge 2.2. Bilgi güvenliği zayıflıklarına örnekler

Kritik görevdeki personelin yedeğinin olmaması
Personelde güvenlik bilincinin eksikliği
Bilgisayarlardaki koruma yamalarının eksik güncellenmesi
Sistemlerde zararlı kodlara karşı korunma eksikliği
Kilitsiz bırakılan kapılar
Açıktan giden elektrik kabloları
Masada unutulan gizli bilgiler (kullanıcının parolası, vb)
Avuç içi bilgisayarların kolay kaybedilecek, çaldırılabilecek hafiflikte ve boyutlarda olması

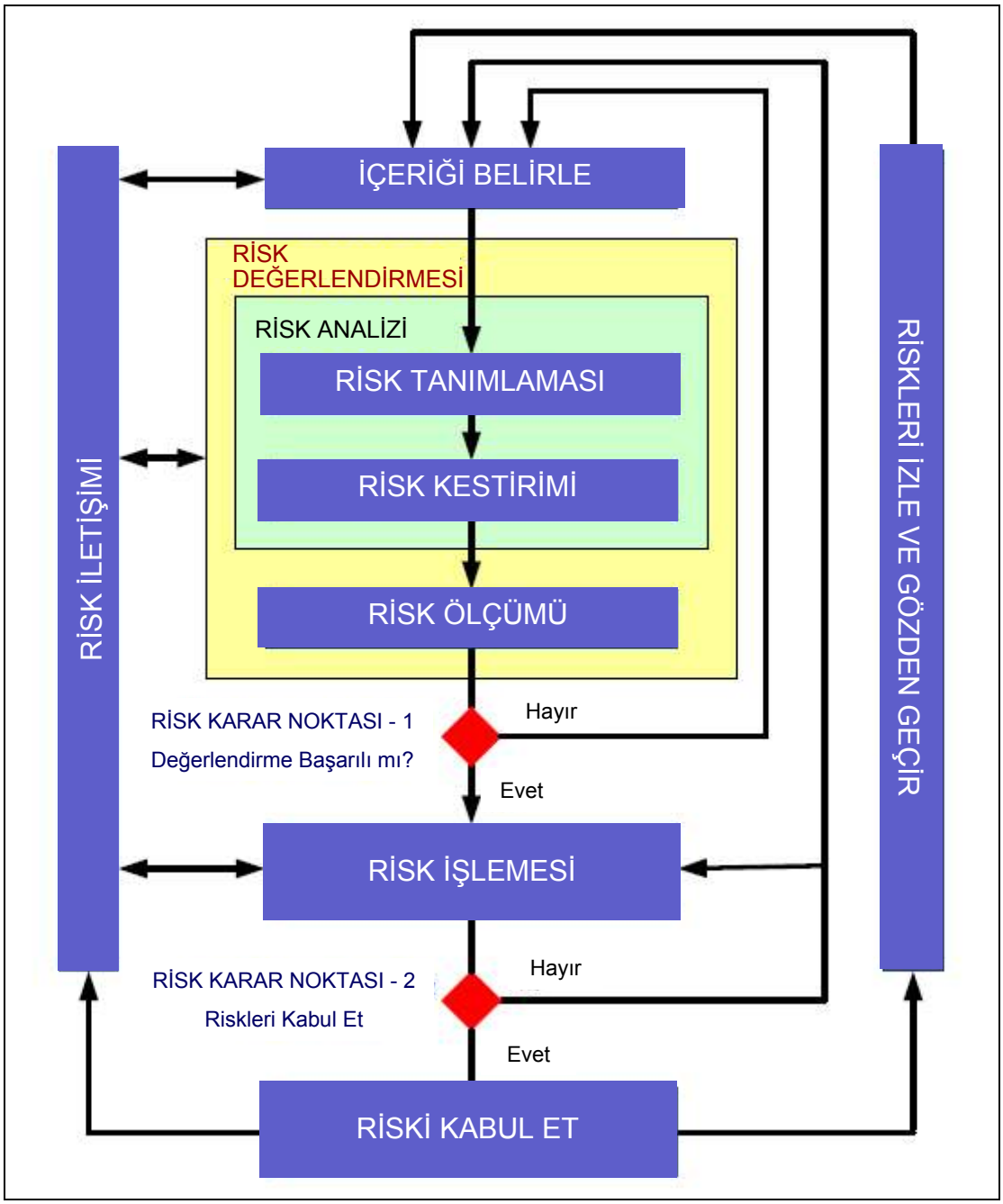
Bilgi güvenliği yönetimi süreçlerinde, güvenlik planlaması risk analizi ile başlamaktadır. Risk analizi, bir sistemdeki zayıflıkların bilinmesi ve bunlardan dolayı uğranabilecek kayıpların hesaplanması işlemidir. Risk değerlendirmesi ise, varlıkların karşısındaki olası tehditler, zayıflıklar, bunların etkileri, olasılıklar ve sonuçta olası risklerin derecesinin veya miktarının belirlenmesidir. Riske değer biçme (evaluation) veya risk değerlendirmesi, belirli bir riskin diğer risklere göreceli derecesinin ve önceliğinin saptanmasıdır. Risk değerlendirmesi, analiz ve değer biçmenin tamamından oluşan bir süreç olarak tanımlanmaktadır (Peltier, 2001).

2.3.3. Uluslararası Bilgi Güvenliği Risk Yönetimi Standardı

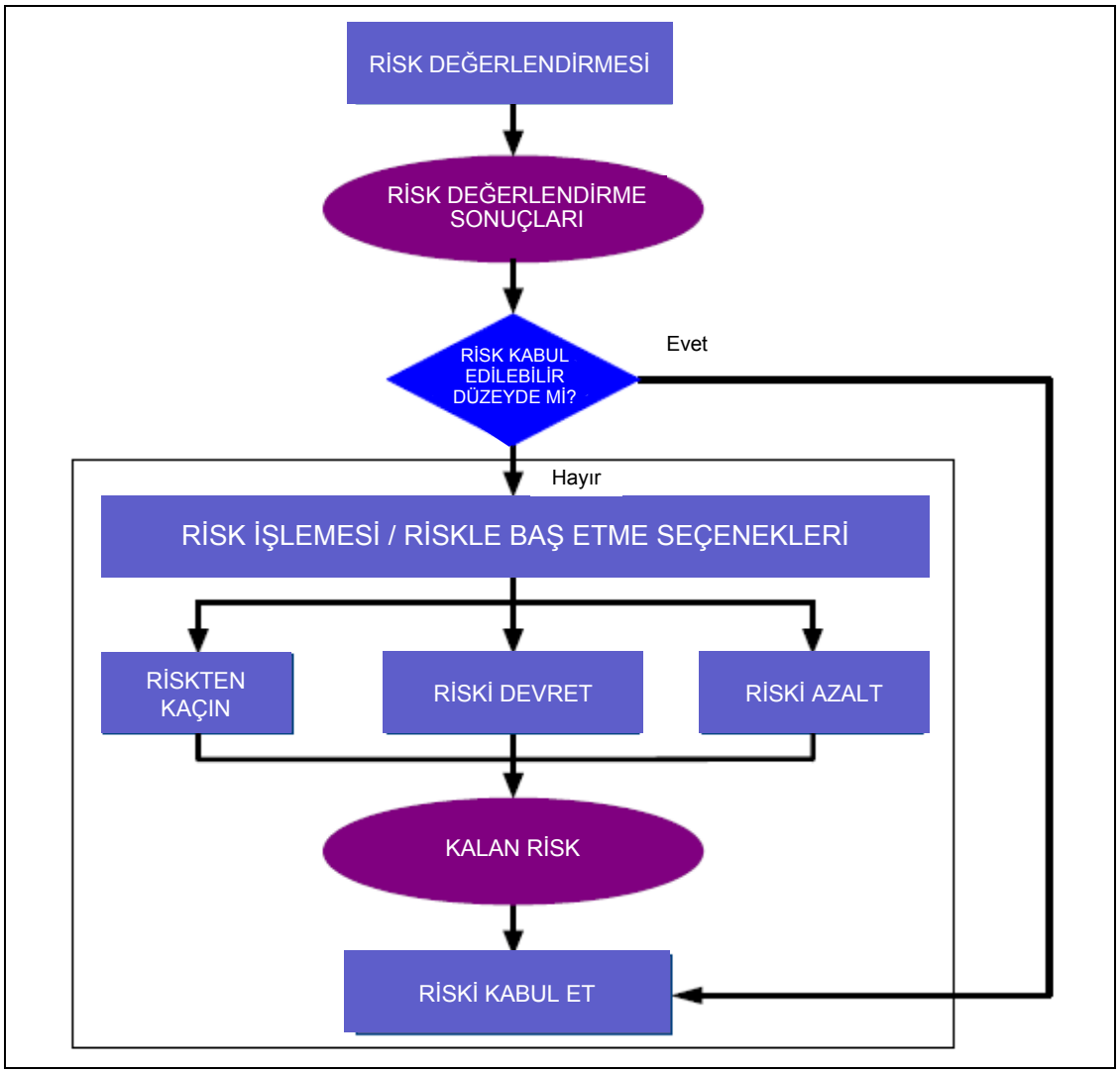
Son yıllarda bilgi güvenliği yönetiminde risk yönetimi kavramının önemli bir noktaya gelmesi, ISO‘ nun bilgi güvenliği risk yönetimine ilişkin yeni bir uluslararası standart geliştirmesine yol açmıştır. ISO/IEC 27005 adıyla 2008 yılında uluslararası bilgi güvenliği risk yönetimi standardı tanımlanmış ve yürürlüğe girmiştir (ISO/IEC, 2008). ISO/IEC 27005, kurumların bilgi güvenliği risklerini değerlendirmek amacı ile genel

kapsamda bir çerçeve sunmaktadır. Özel, adlandırılmış ya da zorunlulukları belirlenmiş bir yaklaşıma sahip değildir. Kurumlar ve örgütler ISO 27001 BGYS' e göre bilgi güvenliği süreçlerini yapılandırırken ilgili risk değerlendirmelerinde ISO 27005 standardından yararlanmalarına olanak sağlanmıştır.

ISO/IEC 27005 standardı, bilgi güvenliği risklerinin tanımlarını, bilgi güvenliği yönetim sürecinde risklerin hangi aşamada, ne şekilde ve nasıl bir yöntemle kapsanacağını belirleyen bir belgedir. Bu standart, kurumların ihtiyaçları doğrultusunda bir risk yönetim yöntem bilimi geliştirmesi gerektiğini ifade eder ve risk yönetim yöntem bilimlerine ilişkin örnekler verir. Kurumların kendi ihtiyaçlarına özgü bir risk yöntem bilimi geliştirmesi veya kabul görmüş bir yöntemi benimseyerek bunu risk yönetim araçları ile otomatik hale getirmesi gerektiğini vurgular. ISO/IEC 27005 standardı, bir kurumda bilgi güvenliği risklerinin tanımlanması, adlandırılması, gruplanması, analizi, değerlendirilmesi, hesaplanması, ölçümü ve sağaltımına (işleme) ilişkin adımları belirlemektedir. Bu standart, bilgi güvenliği risklerinin sürekli ve güncel şekilde izlenmesi, denetimi ve yönetimine ilişkin kuralları ve yöntemleri tanımlar. Şekil 2.4.'te ISO/IEC 27005' in bilgi güvenliği risk yönetim süreci gösterilmektedir (ISO/IEC, 2008). Şekil 2.5.'te de ISO/IEC 27005' te tanımlanmış olan risk işleme (treatment) ile ilgili süreçteki çeşitli seçeneklerin akış şeması gösterilmektedir (ISO/IEC, 2008).



Şekil 2.4. ISO/IEC 27005 Bilgi Güvenliği Risk Yönetimi Süreci



Şekil 2.5. ISO/IEC 27005' e göre bilgi güvenliği risk işleme süreci

ISO/IEC 27001 BGYS'de olduğu gibi, ISO 27005'te de, PUKÖ yaşam döngüsü olgusu bulunmaktadır. ISO 27005, bilgi güvenliği risk yönetimi sürecini PUKÖ modeli çerçevesinde yorumlamaktadır ve çizelge 2.3.'te bu model kısaca açıklanmaktadır (ISO/IEC, 2008).

Çizelge 2.3. ISO/IEC 27005'te tanımlanan PUKÖ modeli

Planla	• İçeriğin belirlenmesi ve tanımlanması • Risk değerlendirmesi • Risk işleme planının oluşturulması • Risk kabulü
Uygula	• Risk işleme planının uygulanması
Kontrol Et	• Risklerin sürekli izlenmesi ve risklerin yeniden gözden geçirilmesi
Önlem Al	• Gerekli değişiklikler ve iyileştirmeleri yaparak bilgi güvenliği risk yönetimi sürecinin etkin şekilde yönetilmesi

ISO/IEC 27005 standardında; bilgi güvenliğine ilişkin varlık tipleri, tehdit tipleri ve örnekleri, zayıflık tipleri ve örnekleri, risk türleri, risk hesaplamasında kullanabilecek yöntemler, riski yönetme için alternatif stratejiler detaylı olarak verilmiştir. Ayrıca standardın ekler bölümünde, açıklamalı çeşitli örnekler gösterilmiştir. Eklerin "E" bölümü ve ilgili alt maddelerinde yer alan çeşitli risk değerlendirmesi ve ölçümü örneklerinin tamamında nitel risk analizi yöntemi esas alınmıştır. Bu örneklerden bazıları çizelgeler 2.4. ve 2.5.'te verilmektedir (ISO/IEC, 2008).

Çizelge 2.4.'te verilmiş olan örnekte, tehdit düzeyi ve zayıflık düzeyi üçlü nitel ölçekte tanımlanmıştır. Bu tanıma göre, "Düşük" değerlere karşılık 0, "Orta" değerlere karşılık 1, "Yüksek" değerlere karşılık olarak da 2 sayısı atanmaktadır. Bilgi varlığı değeri de, beşli ölçekte tanımlanmış olup 0 ile 4 arası değerler almaktadır. Buna göre, her bir riskin hesaplanması da bilgi varlığı değeri, tehdit düzeyi ve zayıflık düzeyinin sayısal toplamı şeklinde bulunmaktadır. Örneğin, çizelgede bilgi varlığı değeri = 4 için, o varlığın "Orta" düzeydeki bir tehdide karşı "Düşük" düzeyde zayıflık durumu söz konusu ise, ilgili risk değeri 4 + 0 + 1 = 5 olarak bulunacaktır (tehdit değeri = 1, zayıflık değeri = 0). Aynı çizelgeden bir başka örnek vermek gerekirse, bilgi varlığı değeri = 0 için (bir kurum için en düşük değerde bilgi içeren varlık veya sistem şeklinde düşünülebilir), o varlığın "Yüksek" düzeydeki bir tehdide karşı "Yüksek" düzeyde zayıflık durumu söz konusu ise, ilgili risk değeri 0 + 2 + 2 = 4 olarak bulunacaktır.

(tehdit değeri = 2, zayıflık değeri = 2). İlgili bütün riskler için benzer şekilde toplanarak hesaplamalar yapıldıktan sonra, risk değerlendirmesinin bir sonraki aşamasına geçilir ve kurum için belirlenen risk kabul değerinden büyük olan riskler belirlenerek risk işleme sürecinde kapsatılır. Bu örnekte, eğer kurumun risk kabul eşik düzeyi 5 ise, Çizelge 2.4.' teki 5 değerinden büyük olan, bir başka deyişle 6, 7 veya 8 değerini alan riskler (çizelgede kırmızı renkle tanımlı olan sayısal değerler) kurum için ele alınması gereken öncelikli riskler olacak ve risk işleme sürecine konulmaları gerekecektir.

Çizelge 2.4. ISO/IEC 27005'e göre örnek bir nitel risk değerlendirmesi

	Tehdit Düzeyi	Düşük			Orta			Yüksek		
	Zayıflık Düzeyi	Düşük	Orta	Yüksek	Düşük	Orta	Yüksek	Düşük	Orta	Yüksek
Bilgi Varlığının Değeri	0	0	1	2	1	2	3	2	3	4
	1	1	2	3	2	3	4	3	4	5
	2	2	3	4	3	4	5	4	5	**6**
	3	3	4	5	4	5	**6**	5	**6**	**7**
	4	4	5	**6**	5	**6**	**7**	**6**	**7**	**8**

Çizelge 2.4.'te gösterilen ve ISO 27005'te de tanımlı olan bu basit bilgi güvenliği nitel risk hesaplama yöntemi, sıkça olmasa da günümüzde kullanılmaya devam etmektedir. Nitel ölçekte karşılık gelen göreceli sayısal değerlerin toplanması yerine, bu değerlerin çarpılarak risklerin öncelik sırasına konulduğu alternatif bir nitel risk hesaplama ve değerlendirme yöntemi de bulunmaktadır ve bu ikinci nitel yöntem günümüzde daha çok tercih edilmektedir. Bu yönteme ilişkin bir örnek Çizelge 2.5'te gösterilmektedir. Çizelge 2.5'te verilen örnekte, her bir bilgi varlığı için 1 ile 5 arası değerler alacak şekilde beşli ölçek kullanılmıştır. Bilgi varlığına çeşitli zayıflıklar sonucunda herhangi bir tehdidin etki etmesi ve bu tehdidin gerçekleşmesinin olabilirlik değeri de benzer şekilde 1 ile 5 arası değerler almaktadır. Bu iki değerin çarpımı sonucunda elde edilen

puan da risk ölçüm değeri olarak bulunmaktadır (Çizelge 2.5' teki "d" değerleri). Her bir bilgi varlığı ve etki eden tehdit için bu değerler hesaplandıktan sonra, Çizelge 2.5' te "e" olarak tanımlanan risk öncelik sırasına göre, ilgili riskler en büyükten en küçüğe doğru azalarak gidecek şekilde sıralanmaktadır. İlgili çizelgedeki örnekte, "d = 15", yani risk değeri en büyük olan, birinci sırada gelmekte ve "e = 1" olmaktadır. Risk değeri 10 bulunan risk de en büyük ikinci değeri aldığı için, öncelik olarak ikinci sırada gelmekte ve "e = 2" olarak tanımlanmaktadır. Çizelge 2.5'te verilmiş olan ve ISO 27005'ten alıntılanmış olan bu nitel risk ölçümünün de basitleştirilmiş bir örnek olduğunu, zayıflık değerleri veya başka etmenlerin de çarpım hesabına katılarak farklı yöntemler geliştirildiğini ve bu değişik yöntemlerin de günümüzde kullanıldığını belirtmekte yarar vardır.

Çizelge 2.5. ISO/IEC 27005'te nitel riskler için tehditlerin derecelendirmesi örneği

Tehdit Adı	Etkilenen varlığın değeri b	Tehdidin gerçekleşmesinin olabilirlik değeri c	Risk Ölçümü d d = b x c	Risk Öncelik Sırası e (d' nin en büyük değeri için e = 1 olacak şekilde; azalan d değerlerine göre, e sıralanmaktadır.)
A	5	2	10	**2**
B	2	4	8	**3**
C	3	5	15	**1**
D	1	3	3	**5**
E	4	1	4	**4**
F	2	4	8	**3**

2.3.4. Bilgi Güvenliği Risk Analizinde Nicel Yöntemler

Bilgi güvenliği risk analizindeki iki temel yaklaşımdan birisi nicel (gerçek sayısal değerlere bağlı) yöntemler ve yaklaşımlardır. Bilgi varlıklarının gerçek veya gerçeğe yakın ölçülmüş parasal değerleri, tehdit ve zayıflıkların etki etme olasılığının (probability) istatistiksel olarak ölçüldüğü değerler ve riskin gerçekleşmesi durumunda oluşacak kaybın da somut bir maliyet ölçeği (parasal değeri) olarak tanımlanabildiği yöntemler nicel yöntemler olarak kabul edilmektedir (Landoll, 2006). Farklı modeller, uyarlamalar ve yaklaşımlar olsa da, temelde tüm nicel yöntemlerin dayandığı ortak yöntem bir sonraki paragrafta özetlenmektedir.

Nicel bilgi güvenliği risk analizinde öncelikle gerekenler şunlardır:

- Bilgi varlığının parasal değeri
- Bilgi varlığına ilgili zayıflıklardan ötürü etki edebilecek tehdidin gerçekleşme olasılığı (p, probability). p, 0 ile 1 arası sayısal bir değer alır.
- Zararın parasal değeri

Nicel risk analizinde ilgili riskin hesaplanması aşağıdaki gibi yapılır (Landoll, 2006):

YBK : YıllıkBeklenenKayıp(Annual Loss Expectancy (ALE))
TBK: TekilBeklenenKayıp(Single Loss Expectancy (SLE))
YTS: Yıl içindeki Tekrar Sayısı(Annual Rate of Occurrence (ARO))
VD : VarlıkDeğeri(Asset Value)
EF:Etki Faktörü(Exposure Factor) (2.1)

*TBK=VD*EF*
*(SLE=Asset Value * EF)* (2.2)

$$YBK = TBK * YTS$$
$$(ALE = SLE * ARO) \quad (2.3)$$

Yukarıdaki denklem (2.2)'de gösterilmiş olan "Tekil Beklenen Kayıp", ilgili bir riskin herhangi bir anda bir seferlik gerçekleşmesi durumunda, bilgi varlığına yaşatacağı zarar veya kaybın parasal değeridir. İlgili bilgi varlığının maddi değerine yaşatacağı zarar,

“Etki Faktörü“ denilen sayısal katsayı ile bilgi varlığının kendi değerinin çarpımıdır ve bu “Etki Faktörü”, 0 ile 1 arası çeşitli gerçek sayı (real number) değerlerini alabilir. “Yıllık Beklenen Kayıp” ise, “Tekil Beklenen Kayıp” değerinin, bir yıl içerisinde aynı riskin gerçekleşme sıklığına bağlı olarak hesaplanmaktadır ve denklem (2.3)’de gösterilmektedir. “Yıllık Beklenen Kayıp” değerini hesaplayabilmek için, ilgili riskin “Yıl içindeki Tekrar Sayısı”, başka bir deyişle, bir yıl içerisinde bu riskin gerçekleşme sıklığının tahmini değerinin bulunması gerekmektedir. Örnek vermek gerekirse, eğer bir bina yangını ve buna ilişkin sistem kesintisi riski son 5 yılda bir kez yaşandı ise; “Yıl içindeki Tekrar Sayısı” $=\frac{1}{5}=0.2$ olacaktır. “Yıl içindeki Tekrar Sayısı” değeri, 0’dan büyük herhangi bir gerçek sayı değeri alabilir. “Tekil Beklenen Kayıp” değerinin, “Yıl içindeki Tekrar Sayısı” ile çarpılmasıyla “Yıllık Beklenen Kayıp” parasal değeri bulunacaktır.

Yıllık Beklenen Kayıp hesaplandıktan sonra, ayrıca, bu riski azaltacak güvenlik önlemleri veya çözümlerinin (safeguards, controls) maliyeti hesaplanır ve Yıllık Beklenen Kayıp’la karşılaştırılarak güvenlik çözümünün ve risk işlemesinin ne kadar etkin ve verimli olduğu sayısal olarak ortaya konur (Landoll, 2006). Bir riski azaltacak çok farklı güvenlik çözümleri, ürünleri, işlevleri olduğu düşünülürse, üst yönetim ve karar verici yetkililerinin güvenlik yatırımlarına karar verme aşamasında bu hesabın yapılması da oldukça önemli ve etkili bir unsurdur. Bu olanak, nitel risk analizinde bulunmamaktadır.

GED : GüvenlikEtkinlikDeğeri(SafeguardValue)
YGM: Yıllık(GüvenlikMaliyeti (Annual Safeguard Cost) (2.4)
GED=YBK (ÇözümÖncesi)-YBK(ÇözümSonrası)-YGM

Yukarıdaki denklem (2.4)’te görüleceği gibi, güvenlik çözümleri / yatırımlarının uygulanmadan önceki mevcut riske ilişkin YBK yanı sıra, riskin azaltılması sonrası kalan risk (residual risk) için yeni YBK değeri hesaplanır. Yeni YBK değeri ve bunu azaltacak güvenlik çözümü maliyeti (YGM) toplamının eski YBK değerinden çıkartılmasıyla kalan değer, ilgili risk azaltıcı güvenlik çözümünün ne kadar etkin ve verimli olduğunu gösterecektir. GED parasal değerinin eksi çıkması, ilgili risk işleme veya risk azaltıcı çözümünün uygun olmadığı ve başka bir çözüme gidilmesi gerektiğini kesin olarak gösterecektir. GED değerinin 0’dan büyük pozitif değer olması da her

zaman çözümün ideale yakın, uygun veya başarılı olduğunu göstermeyebilir. Örneğin, YGM değerlerinin milyon TL veya daha büyük olduğu durumlarda, GED' in pozitif ama bir iki basamaklı göreceli olarak YGM' den çok daha düşük çıktığı durumlarda; güvenlik çözümünün verimliliği ve uygunluğu gözden geçirilmelidir.

2.3.5. Bilgi Güvenliği Risk Analizinde Nitel Yöntemler

Bilgi güvenliğinin nitel risk ölçüm ve değerlendirmesinde gerçek sayısal değerler, finansal maliyetler ve daha önceden hesaplanmış istatistiksel gözlemlere bağlı olasılık ölçümleri yerine, öznel değerlendirmeye bağlı olan, göreceli, sayısal olmayan ölçeksel ağırlıklar veya puanlamalar kullanılmaktadır (Landoll, 2006). Nitel risk analizi yöntemlerinde, riskin gerçekleşme olasılığı veya başka bir deyişle, bir bilgi varlığına, zayıflıklarından yararlanarak bir tehdidin etki etme ve zarar verme olasılığı genelde olabilirlik (likelihood) olarak tanımlanır (Layton, 2007). Bunun nedeni, riskin gerçekleşme sıklığı veya olasılığına dair istatistiksel veriler, gözlem sonuçları, gerçek sayılar, nicel ölçümler olmaması ve bu nedenle istatistikteki genel olarak bilinen olasılık (probability) kavramını tam olarak karşılamamasıdır. Riskin yaşanma sıklığı veya tehdidin etki etme olasılığı "Az, Çok, vb" gibi nitel ifadelerle belirlendiği için, çoğu ilgili kaynakta olasılık yerine olabilirlik olarak tanımlanmaktadır (Dhillon, 2007).

Nitel risk analizinde değişik yöntemler, modeller ve hesaplama araçları geliştirilmiştir. Bunların çoğu temelde aynı mantıktan yola çıkmaktadır. Bu temel mantık ve yöntem aşağıda özetlenmiştir (Peltier, 2001):

- Bilgi varlıklarının değeri "Düşük", "Orta", "Yüksek", vb nitel tanımlamalarla sınıflandırılır. Bazı yöntemlerde, bu değer gizlilik, bütünlük ve kullanılabilirlik şeklinde üç alt kümeye ayrılır ve bu üçünün bileşkesi hesaplanır.
- Bilgi varlığına, belli bir zayıflığından yararlanarak etki edebilecek tehdidin gerçekleşme olabilirlik değeri "Düşük", "Orta", "Yüksek", vb nitel tanımlamalarla sınıflandırılır.
- Riskin gerçekleşmesi durumunda oluşacak zarar veya etki, "Düşük", "Orta", "Yüksek", vb nitel tanımlamalarla sınıflandırılır. Bazı yöntemlerde, etki değeri

gizlilik, bütünlük ve kullanılabilirlik şeklinde üç alt kümeye ayrılır ve bu üçünün bileşkesi hesaplanır.

- Yukarıda belirtilen tüm nitel sınıflandırmalar aynı zamanda göreceli bir ölçekte tanımlanır. Genelde beşli ölçek (scale) kullanılır. (1 en düşük, 5 en yüksek, vb) Literatürdeki diğer bazı yöntemlerde, üçlü, sekizli veya onlu ölçekler de kullanılmaktadır.
- Ölçekte karşılık gelen olabilirlik değeri ile zararın kartezyen çarpımı alınır. Bazı yöntemlerde, olabilirlik değeri ve zarar değerinin düz toplamı alınarak da hesaplanabilir. Olabilirlik - Zarar matrisi oluşturulur. Böylece risk değerleri belli bir ölçek içerisinde hesaplanır ve sıralanır.
- Belirlenen kabul edilebilir risk eşik düzeyinden yüksek olan riskler, risk işleme sürecine katılarak bu riskler için uygun güvenlik önlemleri ve çözümleri belirlenir.

Çizelge 2.6. ve 2.7.'de 1–9 arası ölçeklendirilmiş bir nitel risk analiz modeli örneği gösterilmektedir. Çizelge 2.6.'da, Olabilirlik - Zarar matrisi tanımı yer almaktadır. Çizelge 2.7.'de ise, Olabilirlik – Zarar matrisine göre risk sıralaması gösterilmektedir (Layton, 2007).

Çizelge 2.6. Nitel risk analizi için Olabilirlik - Zarar matrisi örneği

Etki				
	Yüksek	6	8	9
	Orta	3	5	7
	Düşük	1	2	4
		Düşük	Orta	Yüksek

Olabilirlik

Çizelge 2.7. Nitel risk analizi için risk sıralaması örneği

Etki	Olabilirlik	Nitel risk sıralama değeri	Risk Matris puanı
Yüksek	Yüksek	Yüksek	9
Yüksek	Orta	Yüksek	8
Orta	Yüksek	Yüksek	7
Yüksek	Düşük	Orta	6
Orta	Orta	Orta	5
Düşük	Yüksek	Orta	4
Orta	Düşük	Düşük	3
Düşük	Orta	Düşük	2
Düşük	Düşük	Düşük	1

Nitel risk analizi ve değerlendirmesinde, ilgili literatürde çok çeşitli modeller ve yöntemler var olmasına rağmen, genelde temel mantığı bu bölümde açıklanan yönteme dayanmaktadır. Günümüzde en yaygın olarak bilinen ve ticari amaçlı yazılım olarak da üretilip kullanılan nitel risk analizi modellerinden birisi CRAMM' dir. CRAMM (CCTA Risk Analysis and Management Method), ilk olarak 1980'lerde askeri amaçlı geliştirilmiş, daha sonra yazılım platformu güncellenip kamu veya özel tüm sektörlerin ticari amaçlı kullanımına sunulmuş nitel bilgi güvenliği risk analizi yazılımıdır (Landoll, 2006). Benzer şekilde, OCTAVE (The Operationally Critical Threat, Asset, and Vulnerability Evaluation) da yaygın olarak bilinen ve kullanılan bir başka nitel risk değerlendirmesi modeli ve aracıdır (Landoll, 2006).

Bilgi güvenliği risklerinin analizi ve değerlendirilmesinde kullanılan nitel ve nicel yöntemlerin birbirlerine göreceli olarak üstünlük sağladıkları veya zayıf kaldıkları bazı temel özellikleri olduğu bilinmektedir. Çizelge 2.8.'de nitel ve nicel yöntemlerin avantaj ve dezavantajlarının karşılaştırılması özetlenerek gösterilmektedir (Peltier, 2001, Tittel, vd., 2003).

Çizelge 2.8. Nicel ve nitel risk analizi yöntemlerinin karşılaştırılması

Nicel Risk Analizi / Değerlendirmesi	**Nitel Risk Analizi / Değerlendirmesi**
Avantajları	**Avantajları**
Ölçütler, hesaplamalar ve elde edilen sonuçlar; nesnel, somut, bağımsız gözlemlerden elde edilen istatistiksel verilere dayanır.	Hesaplamalar basittir.
Bilgi varlıklarının parasal değerleri ve risk yaklaşımı kavramları çok ayrıntılı ve belirgindir.	Bilgi varlıklarının ve risklerin parasal değerlerini bilmeye veya ölçmeye gerek yoktur.
Risk işleme sürecindeki güvenlik çözümlerinin kar-zarar analizini yapma, bu şekilde de üst yönetimin; bilgi güvenliğindeki yatırımların verimliliğini irdeleme olanağı vardır.	Tehditlerin yoğunluğunu sayısal olarak ölçerek belirlemeye gerek yoktur.
Üst yönetimi tatmin edecek şekilde tüm bilgiler ve tüm sonuçlar parasal ve gerçek sayılarla ortaya konabilir.	Teknik bilgi becerisi yetersiz olan veya güvenlik konusunda bilgisi olmayan kişiler de risk değerlendirme sürecine katkı yapabilir.
Otomatik bir süreç biçimine getirilebilir.	Raporlama, sonuçları değerlendirme ve iş süreçleri ile bütünleştirmede esnektir.
	Çok fazla miktarda ön bilgiye veya daha önceden yapılmış istatistiksel verilere gerek duymaz. Parasal değerin ölçülemediği veya ön bilginin hiç olmadığı koşullarda da kullanılabilir.
Dezavantajları	**Dezavantajları**
Hesaplamalar karmaşıktır.	Öznellik ve taraflı yargıda bulunma sorunu vardır.
Genelde, ilgili bir bilgi bankasıyla ve uygun araçlarla birlikte kullanılırsa başarılı sonuç verebilir.	Bilgi varlıklarının gerçek parasal değerlerini sürece katma olanağı kısıtlıdır.
Çok fazla sayıda ön çalışmaya ve ön bilgiye gerek duyulur.	Risk işleme sürecindeki güvenlik çözümlerinin kar-zarar analizini yapma olanağı yoktur.
Üst yönetime veya uzman olmayan kişilere sunulacak kadar yalınlaştırılması çok zordur.	Tahmine dayalı olduğu için başarılı kestirimler yapılması zorunludur.
Risk değerlendirme sürecine katılan kişilerin yönlendirilmesi zor ve karmaşık bir işlemdir.	Otomatik bir biçime gelmesi çok zordur, otomasyona elverişliliği düşüktür.
Risk değerlendirme süreci sırasında karar değiştirmek çok zor ve zahmetlidir.	
Kapsam dışı bırakılan riskleri sonradan hesaba katmak genelde çok zordur.	

Her iki yöntemin güçlü olduğu yönleri kullanarak kısmen nitel ve kısmen nicel özellikler gösteren melez (hibrid) risk analiz ve değerlendirme yöntemleri de son yıllarda ortaya konmakta ve uygulanmaktadır. Melez yöntemlerin en yaygın olarak kullanılanlarından bir tanesi, Türk araştırmacıların geliştirdiği BİGRA (Bilgi Güvenliği Risk Analizi Yöntemi) veya uluslararası literatürde bilinen adıyla ISRAM (Information Security Risk Analysis Method) adlı yöntem ve bu yöntemi kullanan yazılımdır. ISRAM, nitel yöntemlerle oluşturulan karar tablolarından, nicel yöntemlerle elde edilen gerçek sayısal değerlerden, bulanık mantık ve hata ağaçları analizi yöntemlerinden yararlanılarak geliştirilmiş bir modeldir (Karabacak ve Soğukpınar, 2005). Bu model, daha çok teknoloji bağlantılı risklerin değerlendirilmesinde veya ölçülmüş sayısal verilerin hazırda bulunduğu durumlarda daha etkin olarak çalışmaktadır. Nitel değerlendirmelerde kullanıcıya bağlı durumlarda, diğer tüm nitel yöntemlerdeki gibi öznellik sorunu içermektedir.

Nite ve nicel yöntemleri birlikte kullanan bir başka alternatif de, COBRA (Cost Estimation, Benchmarking and Risk Assessment) adıyla bilinen modeldir. COBRA modelinde; uzman deneyimi ile gelen nitel değerler, yaklaşık sayısal değerlere çevrilmektedir. Sonraki aşamada güvenlikle ilgili nicel sayıları kullanarak gerçek maliyete dönüştürülmektedir. Bu model, ilk başta kurumların herhangi bir projedeki proje yatırım maliyetlerini tahmin etme amacıyla geliştirilmiştir. Bilgi güvenliği risk değerlendirmesi amaçlı kullanıma da sonraki süreçlerde uyarlanmıştır (Dhillon, 2007).

3. BİLGİ GÜVENLİĞİNDE ÖZDEVİMLİ ÖĞRENME YÖNTEMLERİ

3.1 ÖZDEVİMLİ ÖĞRENMEDE GENEL KAVRAMLAR

Özdevimli öğrenme (machine learning), bilgisayarların algılayıcı verileri ya da veritabanları gibi veri türlerine dayalı öğrenimini olanaklı kılan algoritmaların tasarım ve geliştirme süreçlerini konu edinen bir bilim dalıdır. Türkçe' de makine öğrenimi, makine öğrenmesi, yapay zeka, yapay akıl, otomatik öğrenim, yapay öğrenim, vb. terimler de özdevimli öğrenme ile eş anlamlı kullanılmaktadır. Bu çalışmada, İngilizce' de "machine learning" olarak ifade edilen kavramın güncel T.D.K. bilim ve sanat terimleri ana sözlüğünde ve diğer güncel kaynaklarda "özdevimli öğrenme" olarak karşılık bulmuş olması nedeniyle özdevimli öğrenme terimi tercih edilmiştir (TDK, 2011).

Özdevimli öğrenme araştırmalarının odaklandığı konu, bilgisayarlara karmaşık örüntüleri algılama ve veriye dayalı akılcı kararlar verebilme becerisi kazandırmayı içermektedir. Bu; özdevimli öğrenmenin istatistik, olasılık kuramı, veri madenciliği, örüntü tanıma, yapay zeka, uyarlamalı denetim ve kuramsal bilgisayar bilimi gibi alanlarla yakından ilintili olduğunu göstermektedir. Bazı güncel kaynaklarda, istatistiksel öğrenme (statistical learning) olarak da adlandırılmaktadır (Hastie, vd., 2009). Özdevimli öğrenmenin kullandığı algoritma ve dayandığı modellerin, istatistik ve ileri istatistikle doğrudan bağlantılı olması nedeniyle kayda değer bir yaklaşım olarak görülebilir. Bazı istatistiksel özdevimli öğrenme araştırmacıları Bayes istatistiği çerçevesi kapsamında kullanılabilen yöntemler geliştirmektedir. Özdevimli öğrenme kuramı ve istatistik temelde farklı kavramlar olsalar da birbiriyle yakından ilintilidir.

Birçok kaynakta özdevimli öğrenme ile veri madenciliği (data mining) arasında yakın ilişki kurulmaktadır. Bundan ötürü; özdevimli öğrenme, veri madenciliğinin bilgisayar destekli otomatik hale getirilmiş ve geliştirilmiş sistemi olarak tanımlanmaktadır. Çeşitli kaynaklarda özdevimli öğrenme, veri madenciliğinin sınıflandırıcı amaçlı alt kümesi olarak da tanımlanmaktadır. Veri madenciliği, büyük ölçekli veriler arasından bilgiye ulaşma ve bilgiyi anlamlandırma işidir (Alpaydın, 2010). Ya da bir anlamda büyük veri

yığınları içerisinden gelecekle ilgili tahminde bulunabilmemizi sağlayabilecek bağıntıların bilgisayar programı kullanarak aranmasıdır. Veri madenciliğinin de ulusal ve uluslar arası terminolojide eş değer anlamlı tanımları ve adlandırmaları bulunmaktadır. Bu tanımlar içindeki en yaygın kullanım Veritabanlarında Bilgi Keşfi (Knowledge Discovery From Databases - KDD)' dir. Alternatif olarak veri madenciliği aslında bilgi keşfi sürecinin bir parçası şeklinde kabul görmektedir. Veri madenciliği adımı, kullanıcı ve bilgi tabanıyla etkileşim halindedir. İlginç örüntüler kullanıcıya gösterilir, bunun ötesinde istenirse bilgi tabanına da kaydedilebilir. Buna göre, veri madenciliği işlemi, gizli kalmış örüntüler bulunana kadar devam eder.

Veri madenciliği, eldeki verilerden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılmasıdır. Bu da; kümeleme, veri özetleme, değişikliklerin analizi, sapmaların tespiti gibi belirli sayıda teknik yaklaşımları içerir (Gökşen, vd., 2011).

Başka bir deyişle, veri madenciliği, verilerin içerisindeki desenlerin, ilişkilerin, değişimlerin, düzensizliklerin, kuralların ve istatistiksel olarak önemli olan yapıların yarı otomatik olarak keşfedilmesidir. Özdevimli öğrenme de, bu yarı otomatik yapının tamamen otomatik hale getirilmesi ve veri madenciliğinde keşfedilen örüntü veya desenlerin, ileride ortaya çıkacak yeni bilgi / veri yığınlarında kendi başına en doğru yorumlama, kestirim ve çıkarımları yapabilmesidir.

Veri madenciliğini istatistiksel bir yöntemler serisi olarak görmek mümkün olabilir. Ancak veri madenciliği, geleneksel istatistikten birkaç yönde farklılık gösterir. Veri madenciliğinde amaç, kolaylıkla mantıksal kurallara ya da görsel sunumlara çevrilebilecek nitel modellerin çıkarılmasıdır. Bu bağlamda, veri madenciliği insan merkezlidir ve bazen insan – bilgisayar ara yüzü birleştirilir.

Özdevimli öğrenme algoritmaları hedeflenen sonuca göre birçok öbeğe ayrılabilmektedir. Sıkça kullanılan algoritma türleri şunlardır:

- Gözetimli öğrenme: Girdileri hedef çıktılara eşleyen bir işlev üretir.
- Gözetimsiz öğrenme: Bir girdi kümesi modeller.
- Yarı gözetimli öğrenme: Uygun işlev ya da sınıflandırıcılar oluşturmak için etiketli ve etiketsiz örnekleri birlikte ele alır.

- Pekiştirici öğrenme: Dünya algısına dayalı bir öğrenme biçimi. Her eylem ortamda bir etki oluşturmakta ve ortam, öğrenme algoritmasına yol gösteren ödüller biçiminde dönütler vermektedir.
- Uyum: Deney girdileri ve çıktılarına dayanarak çıktıları öngörmeye çalışır.
- Öğrenmeyi öğrenme: Önceki deneyimlerden yararlanır.

Özdevimli öğrenme veya makine öğrenmesi, aslında insanın öğrenmesine benzer bir yapıdadır. İnsanlar çocukluk döneminde temel kavram tanımlarını şekillendirmede tümevarım yöntemini kullanırlar. Birey, hayvanlar, bitkiler, eşyalar ve bunun gibi kavramları ifade eden örnekleri görür. Bu görme ve algılama süreci sırasında birey, örneklere verilen adları işitir ve kavram özelliklerini tanımladığına inandığı sözcükleri seçerek sınıflama modelini oluşturur. Daha sonra bu modelleri benzer yapıdaki nesneleri tanımlamada kullanır. Bu tür öğrenme "tümevarıma dayalı denetimli kavram öğrenme" veya kısaca "denetimli öğrenme" (supervised learning) olarak tanımlanır. Denetimli öğrenmeyle, girdi verilerinin değerleri kullanılarak çıktı değerleri tahmin edilmeye veya öğrenilmeye çalışılır. Bu süreçte öncelikle sonuçları bilinen veriler üzerinde bir sınıflama yapılır ve sonuçları bilinmeyen veri kümesi için sonuçlar tahmin edilmeye çalışılır. "Denetimsiz öğrenmede" (unsupervised learning) önceden tanımlanmış bir sınıfa ait olmayan verilerden model oluşturulur. Veri örnekleri, kümeleme sistemleri tarafından tanımlanan bir benzerlik taslağına göre gruplandırılır. Elde edilen kümelerin anlamı, bir veya birden çok değerlendirme tekniğinin yardımıyla kullanıcı tarafından belirlenir. Denetimsiz öğrenme modellerinde bir çıktı alanı söz konusu değildir (Witten, vd., 2011).

Özdevimli öğrenme sistemlerinin bir bölümü insan sezgisine olan gereksinimi tümüyle ortadan kaldırmaya çalışırken, bazıları insan ve makine arasında işbirliğine dayalı bir yaklaşım benimsemektedir. Öte yandan, sistemi tasarlayan kişinin verinin kodlanma biçimi üzerinde tümüyle egemen oluşu nedeniyle insan sezgisinin tümüyle ortadan kaldırılması olanaksızlaşmaktadır. Özdevimli öğrenme, deneysel yöntemin otomatikleştirilmesi çabası olarak görülmektedir.

Özdevimli öğrenme algoritmalarında çoğunlukla ikili sınıflandırıcı (binary classifier) yapılar ve modeller kullanılmakta ve uygulamada temel amaç, kestirim veya tahmini yapılacak sonuç değerlerinin iki değer veya iki sınıftan birisine atanması şeklinde olmaktadır. Bu iki değer, 0 ve 1 gibi sayısal (numeric) olabileceği gibi, herhangi bir şekilde betimlenmiş sayısal olmayan / isimli (nominal) değerler de alabilir (Witten, vd., 2011).

Özdevimli öğrenme algoritmalarında diğer bir kullanılan yapı da, çoklu sınıflandırıcı (multi-classifier) olarak bilinmektedir. İkili sınıflandırıcılardan tek temel farkı, sonuç kestirim sınıflarının ikiden daha fazla değer içermesidir. İlgili sınıflandırıcı algoritma, sonuçları tahminlerken, üç veya daha fazla sınıftan hangisine girmesi gerektiğini olabildiğince doğru veya az hatayla bilmeye çalışır. Çoklu sınıflandırıcılar, daha karmaşık bir yapıya sahip olmaları, daha yavaş çalışmaları ve başarım değerlerinin genelde daha düşük olması nedeniyle ikili sınıflandırıcılar kadar yaygın kullanılmamaktadır.

Özdevimli öğrenmede hangi sınıflandırıcı ve buna bağlı olarak hangi algoritma ve ilgili model kullanılırsa kullanılsın, oluşturulan modelin başarım (performance) ölçümü ve ölçütleri bilinen bazı standart yöntemlerle ele alınmaktadır. Temelde, seçilen sınıflandırıcı algoritması ve bunun ne ölçüde başarılı olduğu irdelenirken iki ana süreç kullanılmaktadır. Bunlardan birisi, modelin eğitim veya daha yaygın bir deyişle öğrenme (train) aşamasında verdiği tahmin veya kestirim (prediction) sınıflandırma başarım sonuçlarıdır. İkinci aşama ise aynı modelin yeni veri kümelerinde kestirim ve başarımının sınanması, yani test aşamasıdır. Özdevimli öğrenme için geliştirilen bir modelin sadece öğrenme sürecinde elde ettiği başarım sonuçlarına göre karar vermek genelde yanıltıcı ve eksik olacaktır. Çünkü bazı sınıflandırıcılar; bazı veri kümeleri veya bazı modellerde öğrenme yerine ezberleme veya taklit etme eğilimine girebilir, bunu da anlayabilmenin tek yolu, modelin test sürecine sokulmasıdır. Eğer öğrenme sürecindeki başarım değerleri, sınama sürecindeki başarım değerleri için göreceli olarak belirlenen kıstaslara göre kayda değer düzeyde düşerse, modelde aşırı uyum (over-fitting) sorunu olabilir (Kılıçaslan, vd., 2009).

İkili sınıflandırıcı modellerde, seçilen sınıflandırıcı algoritmanın, eldeki örneklem veri kümesi üzerinden sonuç değerlerini iki ayrı sınıfta ayrıştırması ile sonuçlar elde edilir.

Elde edilen sonuçlarda, orijinal veri kümesindeki sonuç değerinin girdiği sınıf ile özdevimli öğrenme algoritmasının tahmin ettiği sonucun aynı veya doğru olup olmadığı karşılaştırılır ve başarım değerleri bulunur. Buna göre de, dört ayrı durum için başarım değerleri ve veri ayrıştırma sonuçları elde edilmiş olur. Bu dört durum aşağıdaki şekilde tanımlanmaktadır:

FN: Yanlış Negatif (False Negative)

FP: Yanlış Pozitif (False Positive)

TN: Doğru Negatif (True Negative)

TP: Doğru Pozitif (True Positive)

Bu dört duruma göre, karışıklık matrisi (confusion matrix) değerleri elde edilir. TP (True Positive) ve TN (True Negative) ile gösterilen sayısal değerler, sınıfları doğru tahmin edilen (Doğru Pozitif ve Doğru Negatif) örneklerin sayılarını, FP (False Positive) ve FN (False Negative) ile gösterilen değerler ise sınıfı yanlış tahmin edilen örneklerin sayılarını (Yanlış Pozitif ve Yanlış Negatif) göstermektedir. Bu dört durumun toplamını oluşturan sayı (TP+TN+FP+FN) ise, tüm veri kümesi adedini verecektir (Sammut ve Webb, 2011). Çizelge 3.1.'de karışıklık matrisi yapısı gösterilmektedir.

Çizelge 3.1. Karışıklık matrisi

		Gerçek sınıf	
		Pozitif	**Negatif**
Tahmin edilen sınıf	**Pozitif**	TP	FP
	Negatif	FN	TN

Karışıklık matrisinden elde edilen bu sayısal değerlerle çeşitli ölçütler hesaplanır. Bunlardan doğruluk (accuracy) ölçütü, tüm veri içinde doğru tahmin edilenlerin oranını ölçmek için kullanılır. Bütün hata tiplerini dikkate alarak, pozitif ve negatif örnekleri

aynı derecede önemsemeyi sağlar. Sınıflandırıcının toplam başarımını değerlendirmeye yardımcı olur. Öte yandan, eğer veri kümesinde dengesiz dağılım var ise, doğruluk ölçütü yeterli olmamaktadır. Bu durumda kullanılan geri çağırım (recall) ve duyarlılık (precision) ölçütleri, sırasıyla, pozitif örneklerin negatif olarak sınıflandırılmasından oluşan yanlışlar ile negatif örneklerin pozitif olarak sınıflandırılmasından oluşan yanlışları belirtir. Geri çağırım ve duyarlılık hesabında, TP adedi sistem tarafından belirlenen / bulunan art gönderimsel ilişki sayısı (total number of documents retrieved) olarak da tanımlanmaktadır (Witten, vd., 2011). F-ölçütü (F-measure) geri çağırım ve duyarlılık ölçütlerini, her ikisinin harmonik ortalamasını (harmonic mean) alarak birleştirir (Sammut ve Webb, 2011). Örnek bir basit harmonik ortalama hesabı aşağıda gösterilmektedir (İkiz, vd., 2006).

Örnek:

1, 2 ve 4 sayılarının basit harmonik ortalaması şu şekilde bulunur:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n.\prod_{j=1}^{n} x_j}{\sum_{i=1}^{n} \frac{\prod_{j=1}^{n} x_j}{x_i}} \tag{3.1}$$

$$H = \frac{1}{\frac{1}{3}\left(\frac{1}{1} + \frac{1}{2} + \frac{1}{4}\right)} = \frac{12}{7}$$

Bu bölümde anlatılan ölçütler, ikili sınıflandırıcılarda kullanılmaktadır. Benzer yöntemler çoklu sınıflandırıcıların başarım ölçütleri için kullanılmakta olup, bazı farklılıklar vardır (Alpaydın, 2010). Tez çalışması sadece ikili sınıflandırıcıları kapsadığından, burada ikili sınıflandırıcıların başarım ölçüt formülleri ve hesaplamaları verilmektedir. Özdevimli öğrenme sınıflandırıcılarının öğrenme ve test süreçlerinde başarım değerlerini ölçmede kullanılan Doğruluk, Duyarlılık, Geri çağırım, F-ölçütü hesaplamaları ilgili denklemleriyle birlikte aşağıda özetlenmektedir (Sammut ve Webb, 2011).

$$Doğruluk = \frac{TP + TN}{TP + TN + FN + FP} = \frac{sistem\ tarafından\ yapılan\ doğru\ sınıflandırma\ sayısı}{sınamada\ kullanılan\ toplam\ örnek\ sayısı} \tag{3.2}$$

$$Duyarlılık = \frac{TP}{TP + FP} = \frac{sistem\ tarafından\ doğru\ olarak\ belirlenen\ artgönderimsel\ ilişki\ sayısı}{sistem\ tarafından\ bulunan\ toplam\ artgönderimsel\ ilişki\ sayısı} \quad (3.3)$$

$$Geriçağırım = \frac{TP}{TP+FN} = \frac{sistem\ tarafından\ doğru\ olarak\ bulunan\ artgönderimsel\ ilişki\ sayısı}{verideki\ toplam\ artgönderimsel\ ilişki\ sayısı} \quad (3.4)$$

$$F\text{-}ölçütü = \frac{2 \cdot duyarlılık \cdot geriçağırım}{duyarlılık + geriçağırım} \quad (3.5)$$

Sınıflandırıcıların başarım ölçütü olarak kullanılabilecek bir başka seçenek de Kappa istatistiği veya Kappa ölçütüdür (Kappa statistics). Rastgele tahminde bulunan dayanak algoritmalarından elde edilen sonuçlarla, özdevimli öğrenme algoritmalarından elde edilen sonuçların istatistiksel yollarla da karşılaştırılmasının gerektiği durumlarda Kappa istatistiği uygun bir istatistiksel ölçüt olarak görülebilir. Kappa istatistiği, sınıflandırıcıların değerlendirilmesi için, doğruluk ölçütüne sunulmuş bir alternatif seçenektir (Kılıçaslan, vd., 2009). Literatürde, iki gözlemci arasındaki uyumun derecesini ölçmek için psikoloji ve tıpta çeşitli alanlarda yapılan deneyler ve gözlemlerde kullanım amaçlı bir ölçüt olarak ortaya atılmıştır ve özellikle sayısal olmayan (nominal) değerler ve ölçekler için geliştirilmiştir (Cohen, 1960). Günümüzde özdevimli öğrenme sınıflandırıcılarının başarımının ölçülmesi de dahil olmak üzere, çeşitli alanlarda kullanılmaktadır. Özdevimli öğrenme alanında bir sınıflandırıcının doğruluğunun, rastsal tahminde bulunan herhangi bir sınıflandırıcının doğruluğuyla karşılaştırılması amacıyla Kappa ölçütü kullanılmaktadır (Kılıçaslan, vd., 2009). Denklem (3.6)'da Kappa istatistiği denklemi tanımlanmaktadır (Cohen, 1960):

$$\kappa = \frac{P_o - P_c}{1 - P_c} \quad (3.6)$$

P_o ilgili sınıflandırıcının doğruluğu, P_c ise aynı veri kümesi üzerinde rastgele tahminle elde edilen doğruluktur veya rastgele bir şekilde sınıflandırıcıların tümünün anlaşma (overall agreement) oranıdır. P_c değeri hesaplanırken, sınıflandırıcıların doğru ve yanlışlardaki ortak anlaşma / aynı kestirim değerlerinin toplamı ile bulunur. Bu hesaplama, bir sonraki paragrafta verilmiş olan örnekte açıklanmaktadır. Kappa istatistiği -1 ile 1 aralığındaki çeşitli eksi veya artı gerçek sayı değerlerini alabilir. -1 tam uyumsuzluğu (tamamen yanlış sınıflandırma), 1 ise tam uyumu gösterir (tamamen doğru sınıflandırma). Kappa, herhangi bir sınıflandırıcının, 0 Kappa değeri veren rastgele sınıflandırıcılardan ne kadar iyi olduğunun değerlendirilmesinde kullanılır. 0.4 ve üzerindeki Kappa ölçüm değerinin, rastlantının ötesinde kabul edilebilir bir uyumu tanımladığı kabul edilmektedir. (Landis ve Koch, 1977). Özdevimli öğrenme sınıflandırıcılarında Kappa değerinin 1 veya 1'e çok yakın değerler (0.996, 0.997, vb.) vermesi, sınıflandırıcının sınama sürecinde çok başarılı olduğunu gösterebileceği gibi, modelin ve uygulamanın gerçekte herhangi bir öğrenme yapmadığını veya öğrenmeye / tahmin etmeye gereksinim olmadığını da gösteriyor olabilir.

İkili sınıflandırıcılarda, Cohen' in Kappa istatistiği yöntemi aşağıda açıklanan örnekteki gibi uyarlanmaktadır:

Örnek 3.1. Toplam 1750 adetlik örnek veri kümesinde, A ve U şeklinde adlandırılmış ikili sınıfa ait aşağıdaki gibi gözlem sonuçları elde edildiğini varsayalım. A sınıfı, olumlu sonuçlar yani pozitif değerleri, U sınıfı da negatif değerleri betimliyor olsun.

		Gerçek sınıflar	
		A	U
Tahmin edilen sınıflar	A	739	37
	U	71	903

Bu sonuçlara göre,

Doğru Pozitif (True Positive, TP) = 739

Doğru Negatif (True Negative, TN) = 903

Yanlış Pozitif (False Positive, FP) = 37

Yanlış Negatif (False Negative, FN) = 71

…değerlerini alacaktır.

İlgili sınıflandırıcının doğruluğu P_o şu şekilde hesaplanır:

$$P_o = \frac{TP + TN}{TP + TN + FP + FN} = \frac{739 + 903}{739 + 903 + 37 + 71} = \frac{1642}{1750} = 0.938 \tag{3.7}$$

Tümünün anlaşma (overall agreement) olasılığı olan P_c değerini bulmak için, öncelikle gerçek sınıf ve tahmin edilen sınıfın rastgele A ve U sınıflarında tanımlama oranları hesaplanmalı ve sonra da bu ikisinin toplamı bulunmalıdır.

$$P_{A1} = \frac{TP + FP}{TP + TN + FP + FN} = \frac{776}{1750} = 0.443$$

$$P_{A2} = \frac{TP + FN}{TP + TN + FP + FN} = \frac{810}{1750} = 0.4628 \tag{3.8}$$

$$P_{U1} = \frac{TN + FP}{TP + TN + FP + FN} = \frac{940}{1750} = 0.537$$

$$P_{U2} = \frac{TN + FN}{TP + TN + FP + FN} = \frac{974}{1750} = 0.5565$$

Gerçek sınıf ve tahmin edilen sınıfın; ortak olarak rastgele herhangi bir örneği A sınıfında tanımlama olasılığı P_A şu şekilde elde edilir:

$$P_A = P_{A1} * P_{A2} = 0.443 * 0.4628 = 0.205 \tag{3.9}$$

Benzer şekilde, gerçek sınıf ve tahmin edilen sınıfın; ortak olarak rastgele herhangi bir örneği U sınıfında tanımlama olasılığı P_U aşağıdaki gibi elde edilir:

$$P_U = P_{U1} * P_{U2} = 0.537 * 0.5565 = 0.2988 \tag{3.10}$$

P_c , tümünün anlaşma olasılığı da aşağıdaki gibi hesaplanır:

$$P_c = P_A + P_U = 0.205 + 0.2988 = 0.50386 \quad (3.11)$$

İlgili denklemde, yukarıda elde edilen değerler yerlerine konarak Kappa istatistiği değeri de aşağıdaki gibi 0.875 olarak bulunur:

$$K = \frac{P_o - P_c}{1 - P_c} = \frac{(0.938 - 0.50386)}{(1 - 0.50386)} = 0.875 \quad (3.12)$$

Burada verilmiş olan örnekteki Kappa değerinin 0.875 olması; örnekteki sınıflandırıcının rastlantının ötesinde kabul edilebilir bir uyuma sahip olduğunu ve yüksek düzeyde başarılı kestirim yapabildiğini göstermektedir.

Şu ana kadar anlatılan özdevimli öğrenme sınıflandırıcıları ve ilgili algoritmaların başarım ölçütleri ve bu ölçütlere ilişkin kullanılan yöntemler daha çok sınıflandırıcının başarımına ve sayısal olmayan verilere ilişkin sınıflandırıcıların başarımına odaklanmaktadır. Buna alternatif bir başka yöntem de, sınıflandırıcıların sayısal olarak kestirim / tahmin etme başarımlarının ölçülmesi ve buna göre değerlendirilmesidir. Özdevimli öğrenme algoritmalarının sayısal kestirimlerine ilişkin başarım ölçütlerinde yaygın olarak bilinen bazı istatistiksel hesaplamalar kullanılır (Witten, vd., 2011). Ortalama mutlak hata (Mean absolute error, MAE), Hata kareler ortalaması (Mean squared error, MSE), Hata kareler ortalamasının karekökü (Root mean squared error, RMSE), Göreceli mutlak hata (Relative absolute error, RAE), Göreceli hata kareler (Relative squared error, RSE), Göreceli hata karelerin karekökü (Root relative squared error, RRSE) ve Korelasyon katsayısı (Correlation coefficient), kestirimlerde sayısal başarım ölçümünde en çok kullanılan istatistiklerdir. Bu hesaplamalara ilişkin özet bir bilgi Çizelge 3.2.'de verilmektedir (Witten, vd., 2011).

Çizelge 3.2.'deki p_1 , p_2 ... p_n özdevimli öğrenme sınıflandırıcısının sınama / test aşamasında ölçülen sayısal tahmin değerleri, a_1 , a_2 ... a_n ise gerçek sayısal değerlerdir. p_i değeri, testteki *i*' nci örnek veya kayda (instance) ait tahminin sayısal değerine karşılık gelmektedir. Tezdeki ilgili model ve çalışmada, sayısal veri kümesi

olmadığı ve sayısal tahmin yapılmadığı için, bu başarım ölçüt değerleri göz önüne alınmamıştır.

Çizelge 3.2. Özdevimli öğrenmede sayısal kestirimcilerin başarım ölçütleri

Hata kareler ortalaması	$\frac{(p_1 - a_1)^2 + ... + (p_n - a_n)^2}{n}$
Hata kareler ortalamasının karekökü	$\sqrt{\frac{(p_1 - a_1)^2 + ... + (p_n - a_n)^2}{n}}$
Ortalama mutlak hata	$\frac{\lvert p_1 - a_1 \rvert + ... + \lvert p_n - a_n \rvert}{n}$
Göreceli hata kareler *	$\frac{(p_1 - a_1)^2 + ... + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + ... + (a_n - \bar{a})^2}$
Göreceli hata karelerin karekökü *	$\sqrt{\frac{(p_1 - a_1)^2 + ... + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + ... + (a_n - \bar{a})^2}}$
Göreceli mutlak hata *	$\frac{\lvert p_1 - a_1 \rvert + ... + \lvert p_n - a_n \rvert}{\lvert a_1 - \bar{a} \rvert + ... + \lvert a_n - \bar{a} \rvert}$
Korelasyon katsayısı **	$\frac{S_{PA}}{\sqrt{S_P S_A}}$ $S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n-1}$ $S_P = \frac{\sum_i (p_i - \bar{p})^2}{n-1}, S_A = \frac{\sum_i (a_i - \bar{a})^2}{n-1}$

* bu denklemlerdeki $\bar{a}$, öğrenme veri kümesinin aritmetik ortalamasıdır.

** bu denklemdeki $\bar{a}$, test veri kümesinin aritmetik ortalamasıdır.

3.2 ÖZDEVİMLİ ÖĞRENMEDE SINIFLANDIRICI ALGORİTMALAR

Özdevimli öğrenmede ikili veya çoklu sınıflandırıcıların oluşturulması ve verilerin sınıflandırma işlemine sokulup sonuçların elde edilmesinde kullanılan çok çeşitli algoritmalar ve yöntemler bulunmaktadır. Bu sınıflandırıcıların pek çoğu istatistikteki yöntemler ve ilgili modellere dayanmaktadır. AODE, AODEsr, BayesNet, NaiveBayes, Lojistik Regresyon, Doğrusal Regresyon gibi istatistikteki çeşitli yöntemlerin doğrudan kullanıldığı sınıflandırıcılar yanı sıra, diğer pek çok sınıflandırıcının uyarlanmasının belli aşamalarında istatistiğin temel kuramları ve ilkelerinden yararlanılmaktadır (Alpaydın, 2010). Karar Ağaçları (Decision Trees), Destek Vektör Makinesi (SVM – Support Vector Machine), Yapay Sinir Ağları (ANN – Artificial Neural Networks), k-En Yakın Komşu (kNN), Voting Feature Intervals, Karar Tabloları diğer belli başlı sınıflandırıcı algoritmalar olarak tanımlanabilir. Bu algoritmaların, çok çeşitli alt türevleri, farklı işlev ve seçenekli alternatif sürümleri de kullanılmakta ve geliştirilmektedir (Witten, vd., 2011). Ayrıca, birden çok sınıflandırıcının birlikte kullanıldığı birleşik (ensemble) algoritmalar da bulunmaktadır.

Bir sonraki bölüm ve onu takip eden bölümlerde, bu çalışmada en başarılı sonuçları veren sınıflandırıcı algoritmalara kısaca değinilecektir. Çalışmanın 5. bölümünde açıklanacak olan modeldeki en başarılı on sınıflayıcı algoritma; Karar Ağaçları, k-En Yakın Komşu ve Destek Vektör Makinesi sınıflandırıcılarının değişik alt türevleridir.

3.2.1. Karar Ağaçları

Karar ağaçları, her iç düğümün (internal node) bir nitelik üzerindeki testini ve her dalın bu testin çıktısını gösterdiği, her yaprak düğümünün (leaf node) ise sınıfları ya da sınıf dağılımlarını temsil ettiği ağaç yapılı akış şemasıdır. En üstteki düğüm ise kök düğüm (root node) olarak adlandırılır (Alpaydın, 2010). Tahmin edici ve tanımlayıcı özelliklere sahip olan karar ağaçları, özdevimli öğrenmede yaygın olarak kullanılmaktadır.

Karar ağacı, ağacın kökünden başlayıp yaprak düğümlere doğru hareket ederek örnek üzerinden sınıflamayı sağlamada kullanılan tümevarımsal bir yaklaşımdır. Güçlendirme (boosting) ile çeşitli hipotezler birleştirilir. Budama (pruning) ile genelleme gücü arttırılır (Kılıçaslan, vd., 2009).

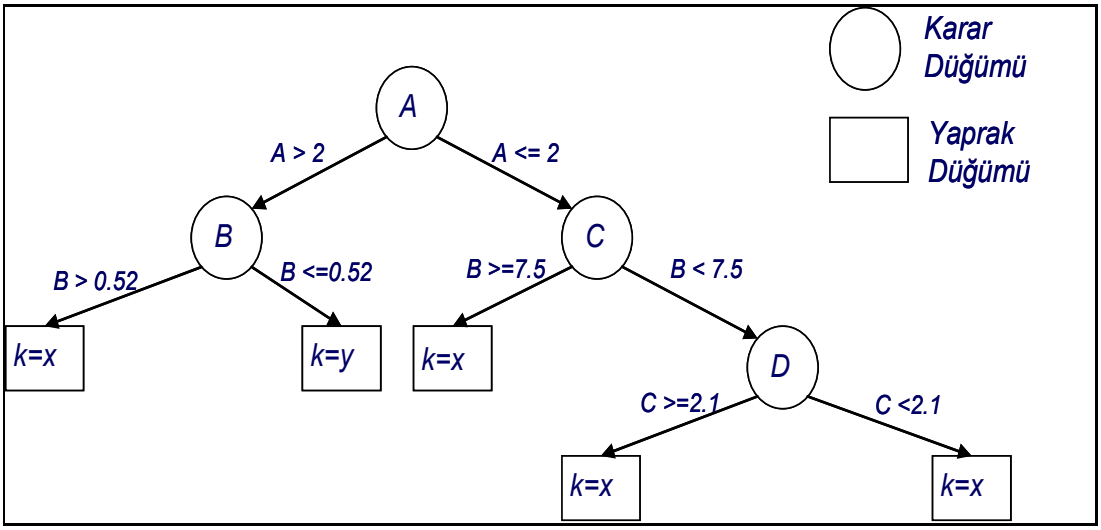
Karar ağacı öğrenmesinde, ağaçlar karar düğümleri ve sonlandırıcı yapraklardan oluşmaktadır. Sınıflandırılacak yeni bir örnek verildiğinde, bir yaprak düğüme ulaşıp da örneğe bir değer ataması yapılana kadar, test fonksiyonları düğümlerde özyinelemeli bir şekilde uygulanmaktadır. Dallanma için, her düğümde örneğin bir özelliği test edilmektedir. Ağacı oluşturmak için, seçilen özellikten meydana gelecek bilgi kazancı hesaplanmalı ve ağacın derinliğini azaltmak için önceden belirlenmiş sayıda, en çok bilgi sağlayan özellikler seçilmelidir. Bir karar ağacı oluşturulurken A özelliğinin kullanılması durumunda elde edilen bilgi kazancını veren denklemler aşağıda gösterilmektedir:

$$I\left(P(v_1),.....,P(v_n)\right)=\sum_{i=1}^{n}-P(v_i)\log_2 P(v_i) \tag{3.13}$$

$$Kazanç(A)=I\left(\frac{p}{p+n},\frac{n}{p+n}\right)-\sum_{i=1}^{v}\frac{p_i+n_i}{p+n}I\left(\frac{p_i}{p_i+n_i},\frac{n_i}{p_i+n_i}\right) \tag{3.14}$$

Denklem 3.13 ve 3.14'te I bilgi içeriği, v_i özelliğin alabileceği sonlu sayıdaki olası değer sayısı, p ve n de sırasıyla eğitim kümesindeki pozitif ve negatif örnek sayılarıdır. Öğrenme kümesinden birden fazla hipotez çıkarılan durumlarda, birleşik (ensemble) öğrenme yöntemleri hipotez uzayı içindeki bir grup hipotezi seçip birleştirerek sınıflandırıcının etkinliğini yükseltirler (Alpaydın, 2010). Bu hipotezler, bileşenlerini derecelendirerek tahmin yürüten tek bir sınıflandırıcı altında toplanır.

Şekil 3.1.'de, karar ağacı ile basit bir sınıflandırma örneği gösterilmektedir. Örnekte, kök karar düğümündeki A parametresinin 2 değerinden küçük veya büyük eşit olma durumuna göre bir alt karar düğümlerine dallanma yapılmakta ve karar ağacı yapısı oluşturulmaktadır. Benzer şekilde, B, C ve D için de çeşitli sayısal değerlerle karşılaştırılmasına göre hangi düğüme geçileceğine karar verilmektedir. Son aşamadaki yaprak düğümlerinde de, k 'nın; x veya y sınıflarından hangisine gireceği belirlenmektedir.



Şekil 3.1. Karar Ağacı ile sınıflandırma örneği

3.2.2. k-En Yakın Komşu Algoritması

k-en yakın komşu (k-Nearest Neigbour - kNN) algoritması ve diğer örnek-tabanlı yöntemler, benzerlik yoluyla öğrenme teknikleridir. k-en yakın komşu sınıflandırması, hevesli (eager) olarak da adlandırılan diğer sınıflandırma yöntemlerinden, bütün öğrenim (eğitim) örneklerini *n*-boyutlu bir uzayda noktalar halinde tutması ve etiketsiz bir örnek sınıflandırılmak isteninceye kadar bir sınıflandırıcı oluşturmaması yönleriyle ayrılır. Bu yöntem, sınama öncesinde bir kural veya fonksiyonlar kümesi oluşturmadığı için eğitim zamanı açısından istekli yöntemlerden daha etkin olmaktadır. Ancak sınama (test) aşamasında her örnek için yeniden hesaplama gerektirdiğinden daha yavaş olmaktadır. Bu yöntemin bir diğer dezavantajı da her özelliğe eşit ağırlık verdiği için ilgisiz özellikleri de sınıflandırmada kullanmak zorunda kalmasıdır. Yeni bir örnek sınanırken, *k* adet en yakın komsusu arasında en sık geçen sınıf etiketi bu örneğe atanır (Wang, vd., 2006).

n-boyutlu bir uzayda $X = (x_1,...,x_n)$ ve $Y = (y_1,...,y_n)$ gibi iki veri noktası kNN algoritmasına verildiğinde, bunlar arasındaki Öklid uzaklığı su şekilde hesaplanır:

$$\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+...+(x_n-y_n)^2}=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \qquad (3.15)$$

Bir başka deyişle, öğrenme kümesine ait örnekler yerleştirildikleri özellik uzayında birer nokta ile temsil edilirler. Sınıfı bulunacak olan örnek, bu uzayda kendine en yakın ve sayıca belirli bir örneklemin sınıf değerini alır. Örneğin diğerlerine uzaklığı Öklid uzaklığı ile hesaplanır. kNN algoritması, *k* değeri iyi belirlendiği takdirde olumlu sonuçlar verebilmektedir (Kılıçaslan, vd., 2009).

Weka yazılımındaki kNN algoritmasına IBk adı verilmektedir (Weka, 2011). Çalışmada en başarılı sınıflandırıcılardan birisi de, kNN' in bir türevi olan Lazy.Ibk algoritması olarak gözlemlenmiştir. Öte yandan, çalışmada Lazy.Ibk varsayılan değerler ve işlevlerle değil, farklı seçeneklerle bu başarı düzeyini yakalamıştır. Bu durum, çalışmanın 5. bölümünde daha detaylı olarak irdelenmiş ve açıklanmıştır.

3.2.3. Destek Vektör Makinesi

Destek vektör makinesi (SVM – Support Vector Machine) sınıflandırmasının temel mantığı, öğrenme (eğitim) verisini bir esleme fonksiyonuyla daha yüksek boyutlu bir uzaya taşımak ve burada maksimum genişlikte ayırıcı çoklu-düzlem (hyperplane) oluşturmaktır. Destek vektör makinesi sınıfları ayıran sonsuz sayıdaki çoklu-düzlem arasında her sınıftan en uzak noktada bulunan düzlemi arar. Sınır çoklu-düzlemleri üzerinde bulunan vektörler destek vektörleri olarak adlandırılmaktadır. Destek vektörleri, çoklu-düzlemleri yeniden kurmak için gerekli tüm bilgiyi taşır. İlk ortaya konan özgün ve ideal çoklu-düzlem algoritması doğrusal bir sınıflandırıcıyken, daha sonra destek vektör makinelerinin, her nokta çarpımının doğrusal olmayan bir $K(x_i, x_j)$ çekirdek fonksiyonuyla yer değiştirilmesi sonucu doğrusal olmayan sınıflandırma problemlerini de çözebilecekleri gösterilmiştir (Hamel, 2009).

Destek vektör makinesi, düşük boyutta doğrusal olarak ayrılamayacak bir veri kümesini, daha yüksek boyuta taşıyarak bir düzlem yardımıyla ayırmayı sağlar. Sınıflandırmada çeşitli çekirdek fonksiyonları kullanılabilmektedir (Kılıçaslan, vd., 2009).

3.3 ÖZDEVİMLİ ÖĞRENME UYGULAMALARI

Özdevimli öğrenmenin başlıca uygulamaları; makine algılaması, bilgisayarlı görme, doğal dil işleme, söz dizimsel örüntü tanıma, arama motorları, tıbbi tanı, bio-informatik, beyin-makine ara yüzleri, kredi kartı dolandırıcılığı denetimi, kurumların ticari risklerini tahmin etme, müşteri memnuniyeti ve pazarlama amaçla müşteri profillerini belirleme, borsa çözümlemesi, DNA dizilerinin sınıflandırılması, konuşma ve el yazısı tanıma, bilgisayarlı görmede nesne tanıma, oyun oynama, yazılım mühendisliği, uyarlamalı web siteleri ve robotların karmaşık hareketleri olarak sıralanabilir (Giudici ve Figini, 2009, Berry ve Kogan, 2010). Tüm bu uygulama alanlarında, özdevimli öğrenme modelleri, yöntemleri ve ilgili yazılımlar sürekli artan bir şekilde başarıyla geliştirilmekte ve kullanılmaktadır.

Özdevimli öğrenmenin en ilgi çekici ve güncel başarılı uygulamalarından birisi de IBM şirketinin geliştirdiği Watson adlı teknolojik ürünüdür. Watson; veri madenciliği, doğal dil işleme, özdevimli öğrenme, yapay zeka, paralel işleme, stratejik karar verme gibi birçok işlevi bünyesinde toplayan yazılım ve bu yazılımın çalıştığı özel bir platformdur (IBM, 2011). İngiltere'deki Jeopardy adlı bir genel kültür yarışmasında dünya şampiyonluğu unvanı olan rakiplerini yenmiştir (The New York Times, 2011). Konunun uzmanları, Watson sisteminin çeşitli evrensel düzeyde önemli projeler için uyarlanıp geliştirileceğini, bilişim dünyası ve yapay zeka alanında çığır açacağını vurgulamaktadır (IBM, 2011). Yarışmanın eğitim süresince sürekli artan başarım değerleri, öğrenme sürecinin karmaşıklığı, çoklu sınıflandırıcı gerektiren karmaşık kestirim koşullarının olması ve zamana karşı yarıştığı rakiplerinin insan olmaları nedeniyle; Watson günümüzdeki en başarılı özdevimli öğrenme uygulamalarından birisi olarak kabul edilmektedir.

Son yıllarda, anti-virüs teknolojileri, saldırı belirleme / önleme sistemleri, güvenlik duvarları, e-posta içerik denetimi sistemleri gibi bilgi güvenliği teknolojilerinde de özdevimli öğrenme uygulamaları yaygın olarak kullanılmaya başlanmıştır.

3.4 BİLGİ GÜVENLİĞİNDE ÖZDEVİMLİ ÖĞRENME

Bir önceki bölümde de kısaca bahsedildiği gibi, son yıllarda sürekli artan bir düzeyde, özdevimli öğrenme modelleri ve algoritmaları, bilgi güvenliği uygulamalarında kullanılmaya başlanmıştır. Bilgi güvenliğinde özdevimli öğrenme, ağırlıklı olarak teknoloji temelli güvenlik uygulamalarında kullanılmaktadır. Ayrıca, bilgi güvenliğinde nicel risk analizlerinde ve nicel risk kestiriminde de özdevimli öğrenme modelleri hem kuramsal olarak hem de teknolojik çözümlere yönelik olarak geliştirilmeye başlanmıştır. Bu modeller ve ilgili uygulamalardan bazılarına burada kısaca değinilmiştir.

Bilgi güvenliği yönetiminde tehditleri tanımlama, gruplandırma, ilgili ölçüm birimini belirleme ve güvenliği ölçerek analiz etme amaçlı Destek Vektör Makinesi (SVM - Support Vector Machine) tipinde özdevimli öğrenme sınıflandırıcılarının kullanıldığı bir model geliştirilmiştir (Qu ve Zhang, 2007).

Bir başka çalışmada; bilgisayar ağlarındaki saldırıları belirleme amaçlı sistemler için çeşitli özdevimli öğrenme algoritmaları kullanılarak bir model oluşturulmuştur (Giacinto, vd., 2003). Ağ güvenliğine yönelik benzer diğer çalışmalarda da çeşitli sınıflandırıcı algoritmalar uyarlanarak modeller ve teknolojik çözümlerin geliştirildiği bilinmektedir (Pietraszek ve Tanner, 2005, Abraham, vd., 2005). Bilgi güvenliğinde özdevimli öğrenme model ve yöntemlerinin en sık olarak uyarlandığı ve uygulandığı alanlar genelde ağ saldırı belirleme / önleme sistemleri ve anti-virüs, vb zararlı kod koruma sistemleri olarak göze çarpmaktadır. Bir başka dikkat çekici olgu da; bilgi güvenliğinde özdevimli öğrenme modellerinin çoğunlukla sayısal verilerin, istatistiksel ölçümlerin ve nicel yöntemlerin kullanılabildiği alanlara uygulanmış olmasıdır.

Bu çalışmanın kuramsal araştırması ve literatür taraması sırasında, kayda değer herhangi bir özdevimli öğrenme yaklaşımlı nitel risk değerlendirme modeli veya çalışması bulunamamıştır.

4. ÖZDEVİMLİ ÖĞRENME İLE BİLGİ GÜVENLİĞİNDE NİTEL RİSK DEĞERLENDİRMESİ MODELİ

4.1. ÇALIŞMANIN AMACI VE KAPSAMI

Bu çalışmanın amacı, günümüzde her türlü sektörde faaliyet gösteren kurumların bilgi güvenliği risklerini ölçümleme ve değerlendirmede kullandıkları yöntem, araç ve yazılımların yetersiz kaldığı alanlar için, özdevimli öğrenme algoritmalarından yararlanarak yeni bir alternatif model geliştirmeye çalışmaktır. Bu amaç çerçevesinde, özdevimli öğrenme yaklaşımı ile bilgi güvenliği risklerinin değerlendirilmesinde hangi kısımlarda mevcut yöntemlerden daha başarılı veya başarısız olduğunun ortaya konulmasına çalışılmıştır. Özdevimli öğrenme modelinin tasarlanması, geliştirilmesi ve uygulanması için bu çalışmada çeşitli aşamalarda farklı çalışmalar ve uygulamalar gerçekleştirilmiştir.

Günümüzde çok çeşitli bilgi güvenliği risk ölçüm yöntemleri, formülleri, modelleri ve bunlara ilişkin yazılımlar mevcut olmasına rağmen, hiçbir yöntem veya uygulama, tüm kurumlar için veya bir kurumdaki her risk için sağlıklı ve başarılı sonuç vermemektedir. Bu nedenle de, kurumlar ya risklerini sağlıklı öngörememekte ve risklerini kontrol edememekte, ya da sürekli konunun uzmanlarına danışarak insana bağlı yorum ve müdahalelerle risk ölçümü ve risk yönetimi yapmak zorunda kalmaktadır. Bu tez çalışması; bu alandaki eksiklik ve yetersizlikleri belli ölçüde giderebilecek şekilde bilgi güvenliği risklerinin belli alanları için güvenilir, esnek, kullanışlı yeni bir yaklaşım ortaya koyması ve bu sayede bilgi güvenliği yönetiminde kurumlara yardımcı olacak örnek bir model oluşturması açısından önem ve fayda arz etmektedir.

Çalışmada amaçlanan model için öncelikle, seçilecek bir kuruma ait nitel bilgi güvenliği risk analizi ve ilgili verilerin elde edilmesi gerekmektedir. Bu nedenle, çalışma kapsamına ve amacına uygun bilgi güvenliği nitel risk analizi için Karşıyaka Devlet Hastanesi'nde anket çalışması yapılmıştır. Bir sonraki bölümde anket çalışması detaylı olarak verilmektedir.

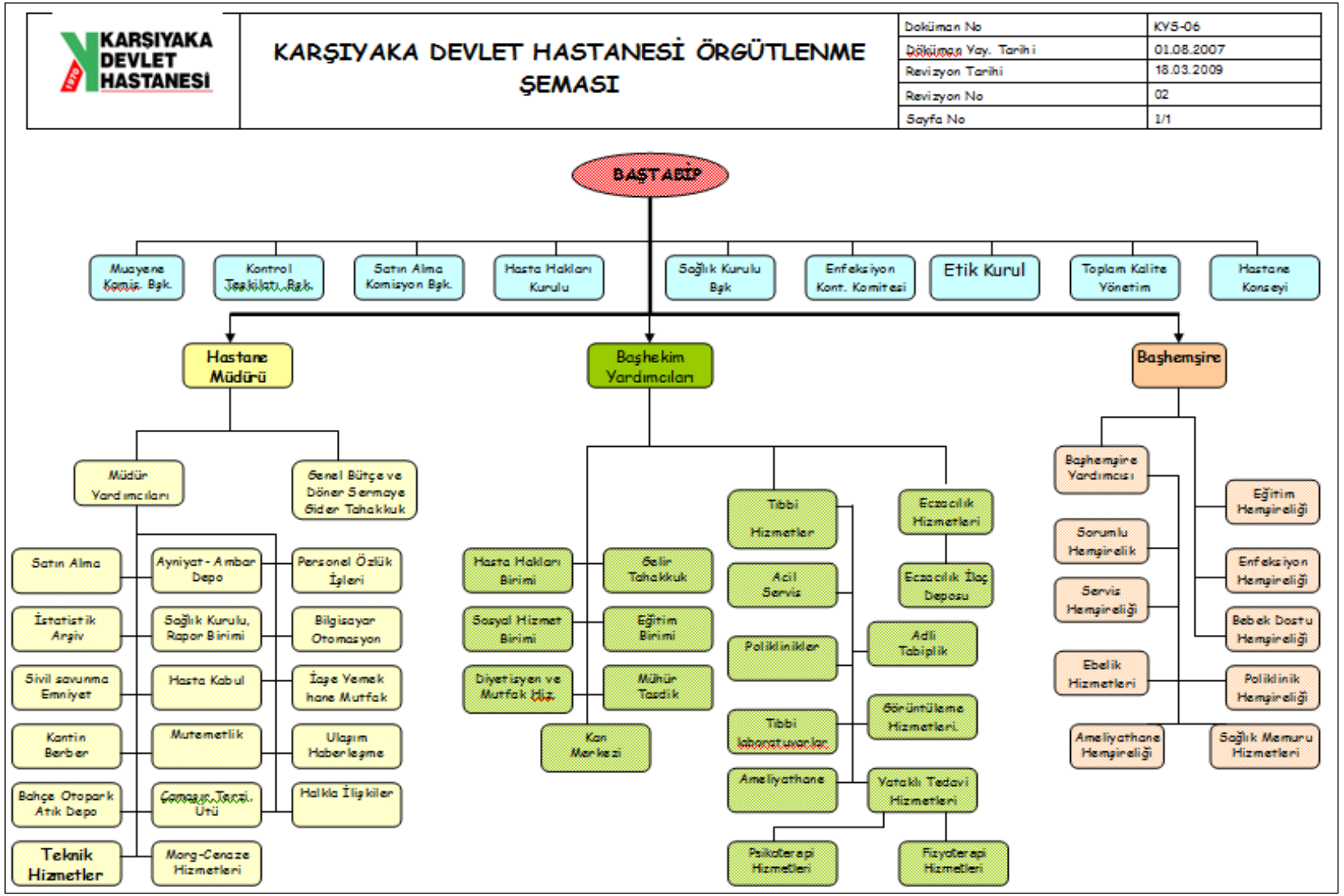
4.2. BİLGİ GÜVENLİĞİ NİTEL RİSK ANKETİ UYGULAMASI

Bilgi güvenliği nitel risk anketi uygulaması için İzmir' in en büyük devlet hastanelerinden birisi olan Karşıyaka Devlet Hastanesi seçilmiştir. Karşıyaka Devlet Hastanesi (K.D.H.), İzmir'in kuzey bölgesine hitap eden tek devlet hastanesidir. Menemen, Foça, Aliağa, Bergama, Dikili, Kınık, Ayvalık, Edremit, Burhaniye ilçelerine de hizmet veren bir bölge hastanesi konumundadır. 1970 yılında 50 yatak kapasiteyle hizmete açılmıştır. A ve B bloklarından oluşmaktadır. Günümüzde, hastanenin 300 yatak kapasitesine sahip olduğu bildirilmektedir. Spiral Bilgisayarlı Tomografi cihazının faaliyete başlaması, kan merkezi, yeni acil servis ve yeni ameliyathanenin de açılması ile hastanenin hizmet yelpazesi daha da genişlemiştir. Karşıyaka Devlet Hastanesinde Ocak 2010 itibariyle 154'ü doktor olmak üzere 580 sağlık personeli, toplam 937 personel görev yapmaktadır. K.D.H.'ye bağlı olarak açılan Egekent Semt Polikliniği ayrılarak, Egekent Devlet Hastanesi olarak hizmet vermeye başlamıştır. 01.01.2005 tarihinde K.D.H. bünyesinde Semt 1 Polikliniği açılmıştır. 21.02.2005 tarihinde SSK ' ya ait sağlık kurumlarının Bakanlığa devrinden sonra Karşıyaka SSK Dispanseri Karşıyaka Semt 2 Polikliniği adı altında K.D.H.'ye bağlanmıştır. Karşıyaka Devlet Hastanesi 2008 yılında ISO 9001 Kalite Belgesini almıştır (KDH-1, 2011).

K.D.H. Ana İş Süreçleri ve Ana Hizmet Birimleri aşağıdaki şekilde tanımlıdır:

- Hasta Kabul ve Acil Servis
- L.B.S.
- Radyoloji
- Ameliyathane
- Satın Alma
- Muhasebe
- Gelir / Gider Gerçekleştirme
- Özlük İşleri

Şekil 4.1.'de Karşıyaka Devlet Hastanesi (K.D.H.) örgüt şeması gösterilmektedir.



Şekil 4.1. K.D.H. örgütlenme şeması

K.D.H. ilgili tüm bilgi sistemleri, iş süreçleri ve birimlerine yönelik yazılım ve iş uygulamaları; toplam 21 ayrı modülden oluşan HBYS (Hastane Bilgi Yönetim Sistemi) içerisinde bütünleşik bir şekilde çalışmaktadır. İlgili yazılım, altyapı, donanım, güvenlik, vb hizmetler ve sistemler dışarıdan ayrı bir BT şirketince sağlanmaktadır. Bilgi sistemleri bakım, yönetim, destek ve veri giriş işlemlerine hizmet veren bu şirket çerçevesinde çalışan toplam 110 personel mevcuttur ve K.D.H.' nin yazılım bakımı da kapsanacak şekilde, tüm bilişim hizmetlerine 7 x 24 kesintisiz destek vermektedir. Ana veri tabanı sistemi Oracle olup, veri / bilgi işlemeye yönelik tüm işlemler bu veri tabanı üzerinden gerçekleşmektedir. Raporlama, evrak ve diğer bazı işletim amaçlı günlük işler için MS Office ve Open Office yazılımları da kullanılmaktadır. Hastanede iş amaçlı e-posta kullanımı kurum politikası gereğince yoktur. Elektronik haberleşme tamamen hastane içi portal ve HBYS ile bütünleşmiş bir yapı içerisinden sağlanmaktadır. İş sürekliliği ve yedekleme konusunda politikalar, yönergeler (prosedürler) ve bitirilmiş bir proje mevcuttur ve tam yedekli bir altyapı kesintisiz hizmet vermekte, ayrıca, verilerden geri dönüş testleri de dahil olmak üzere, belli aralıklarda Kontrol Teşkilatı'nca denetimler ve testler yapılmaktadır. Kontrol Teşkilatı, diğer tüm bilgi sistemleri güvenliği izleme ve denetimlerini de sürekli yapmaktadır. K.D.H., güvenli şekilde e-randevu, laboratuar sonuç, vb web tabanlı Internet'e açık uygulamalar ve Kiosk üzerinden otomasyon projelerine de başlamıştır. Bazı modüller faal olarak hizmete geçmiştir (KDH-2, 2011).

Karşıyaka Devlet Hastanesi; Genel Sağlık Sigortası sağlık bakım hizmetleri ödemeleri için altyapı geliştirilmesi amacıyla başlatılan ve kısaca DRG (Tanı İlişkili Gruplar) olarak adlandırılan "sağlık hizmetleri finansman yapısının güçlendirilmesi ve yeniden yapılandırılması için altyapı geliştirme projesi" için Türkiye çapında belirlenen ilk 50 pilot hastaneden birisidir. 2010 yılı itibarı ile Türkiye çapında 262 hastanede mevcut olan "Sağlık Bakanlığı DRG Uygulamaları" kapsamında hastane, DRG'ye dayalı maliyet uygulamaları yapmakta ve elde edilen verileri Sağlık Bakanlığı'na göndermektedir. Daha önce Hacettepe Üniversitesi bünyesinde Haziran 2007'den beri DRG Projesi şeklinde yürütülen bu hizmet, Kasım 2009'dan itibaren Sağlık Bakanlığı Tedavi Hizmetleri Genel Müdürlüğü Performans Yönetimi ve Kalite Geliştirme Daire Başkanlığı bünyesinde kurulan Tanı İlişkili Gruplar Şubesi tarafından yürütülmektedir (KDH-3, 2011).

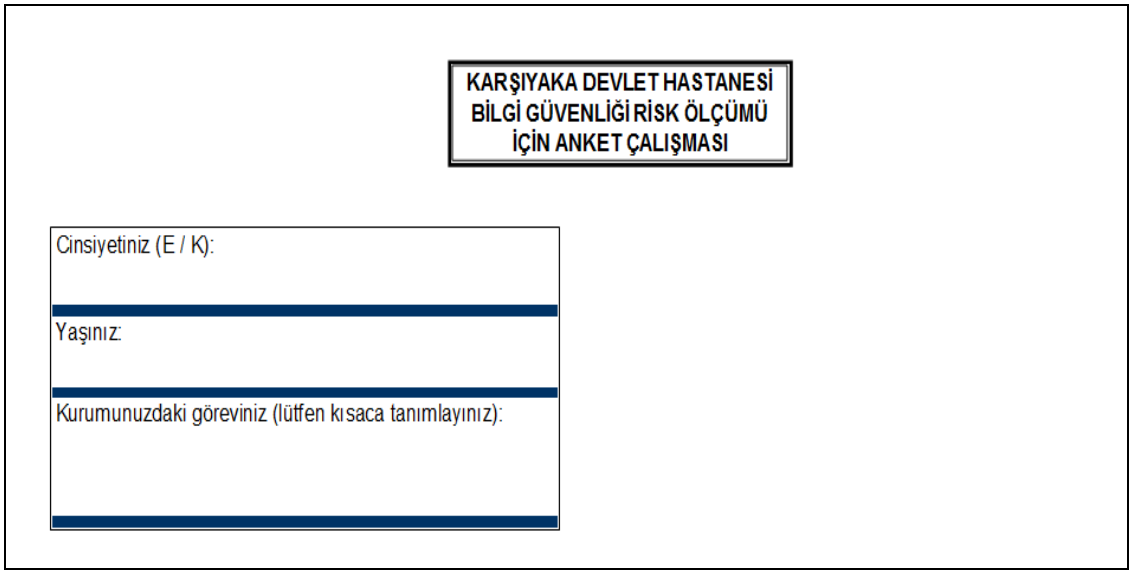
Anket çalışması için, öncelikle Karşıyaka Devlet Hastanesi Müdür Yardımcısı, Başhekimi ve bilgi işlem yetkilisinden randevu alınıp 12 Temmuz ve 16 Temmuz 2010 tarihinde toplantılar yapılmıştır. Çalışmadaki anket uygulamasının gerçekleştirilebilmesi için, Karşıyaka Devlet Hastanesi'nden gerekli resmi izinler 26 Temmuz 2010 tarihinde alınmıştır. Anket kapsamı, ilgili sorular ve anketin uygulama yöntemi, hastane yetkililerinin de görüşleri alınarak belirlenmiş ve buna uygun şekilde hem Microsoft Excel yazılımı ile elektronik ortamda, hem de kağıt çıktı olarak hazırlanmıştır.

Anketler, hastanedeki iki ayrı denek grubuna, yetkililerin gözetimi ve yönetiminde toplam 64 kişiye 3 Ağustos - 4 Ağustos 2010 tarihlerinde uygulanmıştır. Anket dört bölümden oluşmakta, ilk üç bölümde toplam 6 adet soru ile kişiye ait kısa bilgiler elde edilmekte, dördüncü bölümde de 270 farklı seçeneği puanlayarak veya işaretleyerek doldurmaları istenen risk analiz soruları bulunmaktadır. Anket katılımcılarının anketi bitirme süresinin, 36 dakika ile 63 dakika arasında değiştiği gözlemlenmiştir. Anket sonuçları hem kağıt belgede 5 yıl saklanmak üzere gizli olarak arşivlenmiş, hem de bilgisayar ortamına (Microsoft Excel yazılımı) aktarılmıştır. Anket ve risk ölçümü için geliştirilen modelleme ve yöntemdeki bazı yaklaşımların, ülkemizde ve yurtdışında bir ilk örnek olduğu söylenebilir. Ayrıca, anketin uygulanmasının da Türkiye'de bir hastane sektöründe ilk örnek olduğu belirlenmiştir.

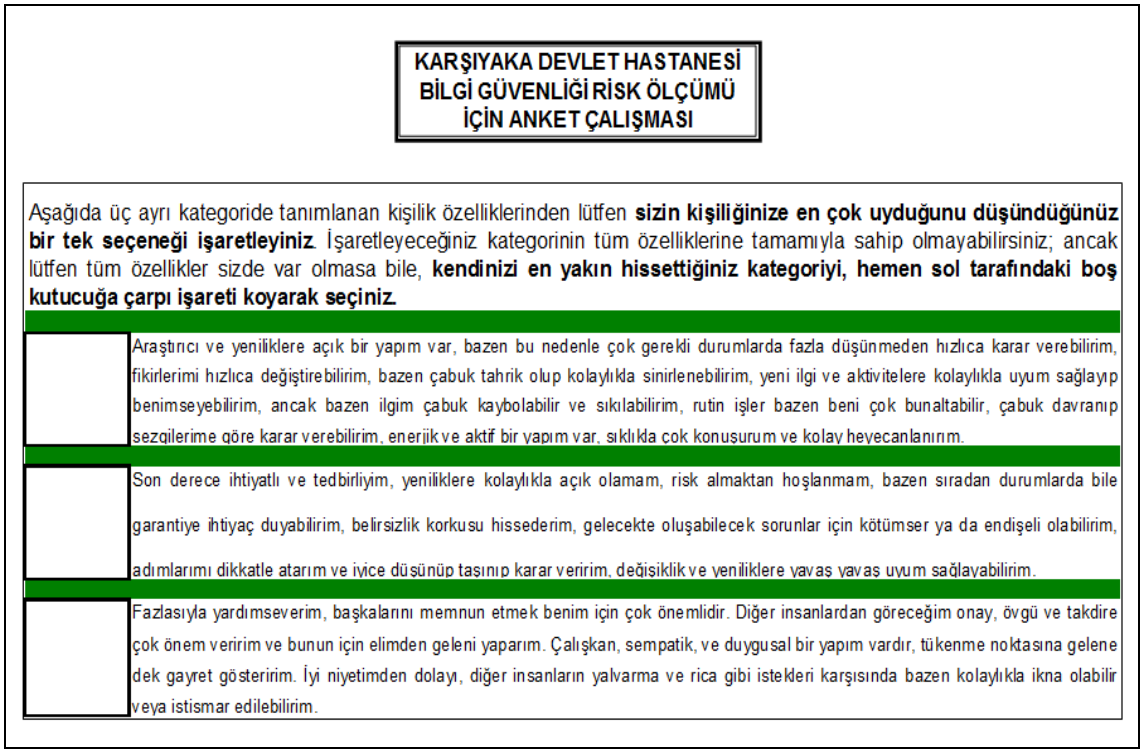
Anket, dört bölümden oluşmaktadır. Birinci bölümde ankete katılan kişinin cinsiyeti, yaşı, kurumdaki görevine ilişkin basit demografik sorular yer almaktadır. İkinci bölümde ise, ankete katılan kişinin risk algılaması ve iş yaşamında tehlikeler, riskli durumlara ne düzeyde hassasiyet gösterdiğini belirlemeye yönelik basitleştirilmiş psikometrik ölçüm amaçlı sorular yer almaktadır. Üçüncü bölümde ise, ikinci bölümdeki soru ve içeriğin aynısı yer almaktadır. Bu bölümün ikinci bölümden farkı, anket katılımcısının kendisi yerine çalışma arkadaşlarını değerlendirdiği yanıt kısmı bulunmaktadır. Şekil 4.2., 4.3. ve 4.4.'de anketin ilk üç bölümüne ait sayfaların çıktı örnekleri yer almaktadır.

Çok basitçe de olsa, kişilik analizi yaparak çalışanların risk olgusunu algılama farklılığını ortaya çıkarmak amacıyla geliştirilen bu sorular, C. Robert Cloninger' in ilgili literatürdeki çalışmalarından yola çıkılarak türetilmiştir. Cloninger; bireylerde zarardan kaçınma, yenilik arama, ödül bağımlılığı ve sebat etme (ısrarcılık) şeklinde

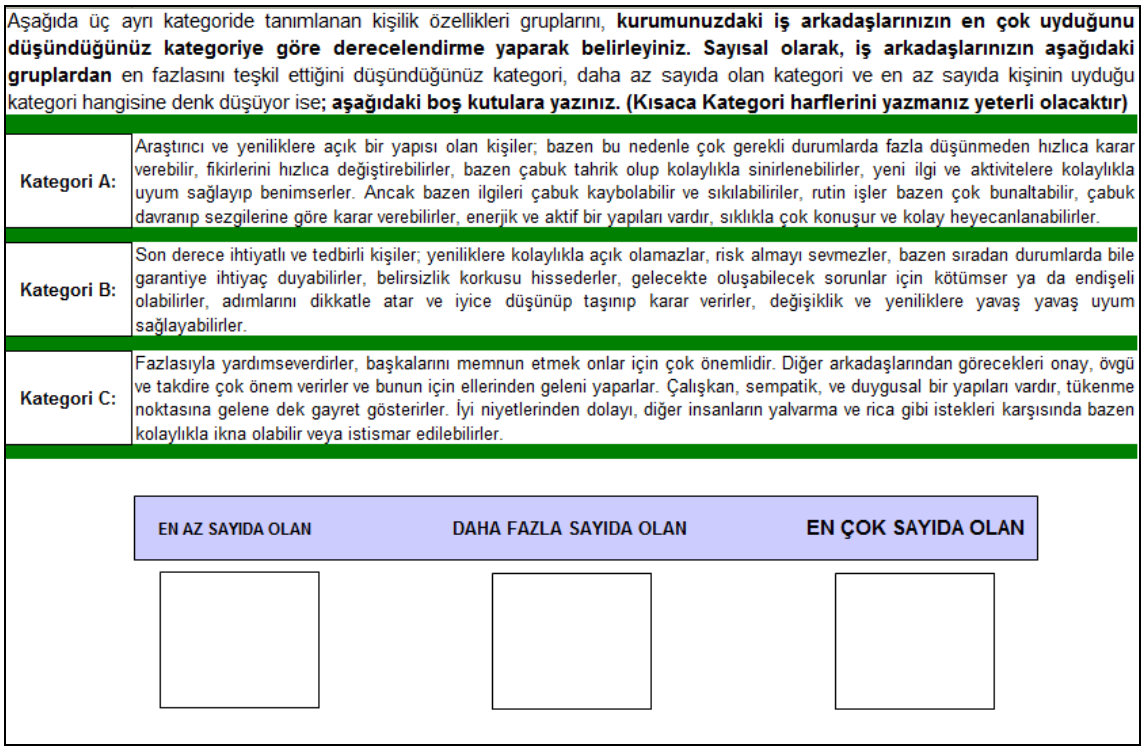
dört farklı mizaç boyutu olduğunu ortaya koymuştur. Yine aynı çalışmalarda; bireylerde kendini yönetme, işbirliği yapma ve kendini aşma şeklinde üç farklı karakter boyutu bulunduğu ortaya konulmuştur (Arkar, 2004).



Şekil 4.2. Anket uygulamasının birinci bölümüne ilişkin örnek çıktı

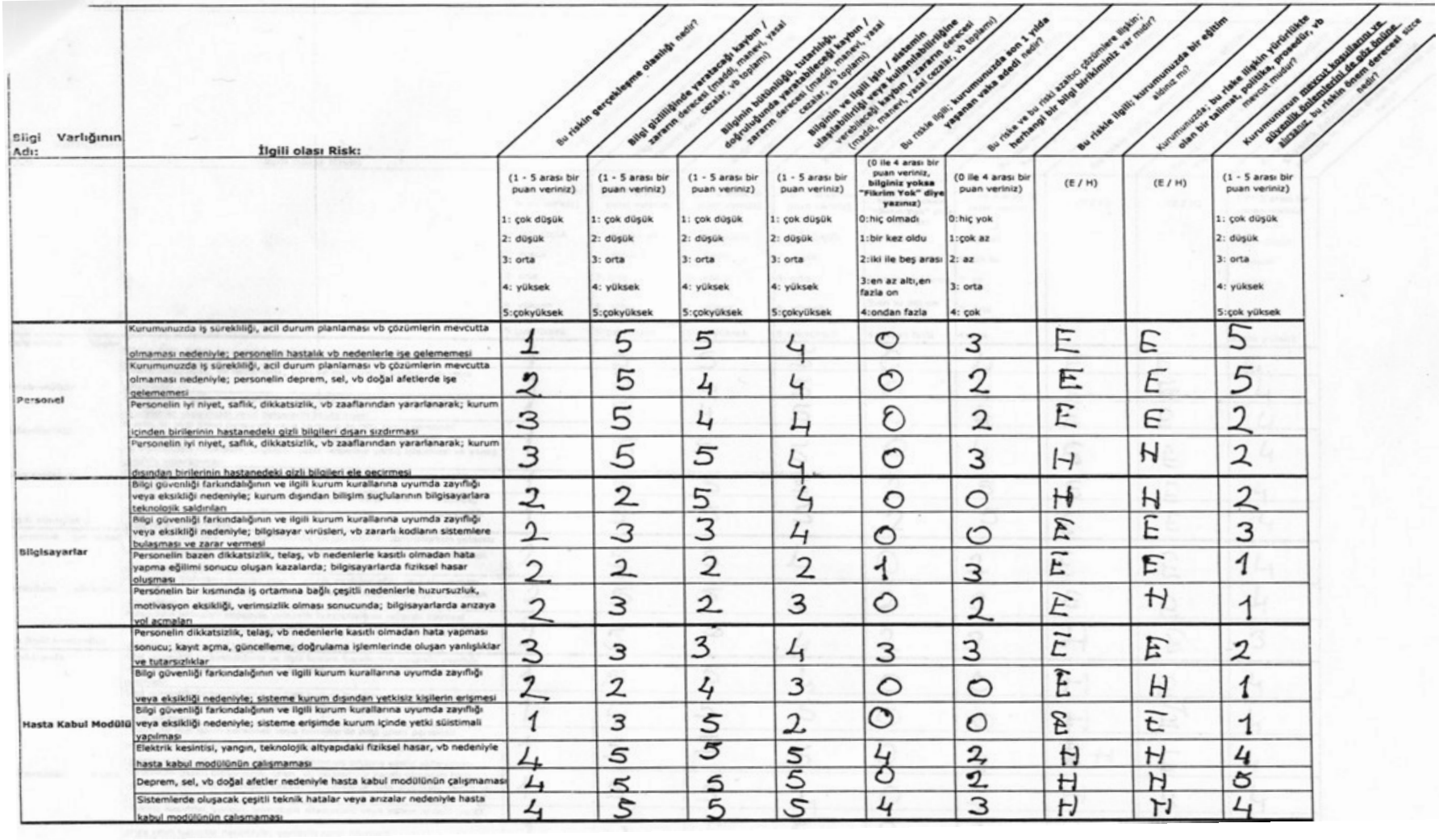


Şekil 4.3. Anket uygulamasının ikinci bölümüne ilişkin örnek çıktı



Şekil 4.4. Anket uygulamasının üçüncü bölümüne ilişkin örnek çıktı

Anketin dördüncü ve son bölümü ise, bilgi güvenliği risklerine ilişkin olarak katılımcılara soruların yöneltildiği ve farklı ölçeklerde nitel puanlama yaparak yanıtlamalarının istendiği sorulardan oluşmaktadır. Toplam 30 farklı risk konusu için sorular bulunmaktadır. Anketin bilgi güvenliği risk kapsamına alınması uygun görülen 6 farklı bilgi varlığı belirlenmiştir. Bunlar; "Personel", "Bilgisayarlar", "Hasta Kabul Modülü", "Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler", "Kurumun personel istihdam süreçleri ve ilgili kontroller" ve "Elektronik ortamdaki hasta kayıt verileri" dir. 6 bilgi varlığının bir tanesi donanım, bir tanesi yazılım, bir tanesi elektronik veri, bir tanesi insan, iki tanesi de iş süreci tipindedir. 6 varlığa ilişkin çeşitli risk konuları belirlenmiş ve bunlara uygun risk soruları oluşturulmuştur. Bu 30 sorunun her birisi için, farklı konularda risk analizi için veri toplama amaçlı ölçüm soruları sorulmaktadır. Bunlar, toplam 9 adettir. Böylece, her bir anket katılımcısının 270 adet risk konusuna ilişkin puan vererek değerlendirme yapması sağlanmıştır. Şekil 4.5.' de anketin kurumdaki risklerin değerlendirildiği dördüncü bölümüne ilişkin örnek bir ekran çıktısı gösterilmektedir.



Bilgi Varlığının Adı:	İlgili olası Risk:	Bu riskin gerçekleşme olasılığı nedir?	Bilgi gizliliğinde yaratacağı kaybın / zararın derecesi (maddi, manevi, yasal cezalar, vb toplamı)	Bilginin bütünlüğü, tutarlılığı, doğruluğunda yaratabileceği kaybın / zararın derecesi (maddi, manevi, yasal cezalar, vb toplamı)	Bilginin ve ilgili işin / sistemin ulaşılabilirliğine verebileceği kaybın / zararın derecesi (maddi, manevi, yasal cezalar, vb toplamı)	Bu riskle ilgili; kurumunuzda son 1 yılda yaşanan vaka adedi nedir?	Bu riske ve bu riski azaltıcı çözümlere ilişkin; herhangi bir bilgi birikiminiz var mıdır?	Bu riskle ilgili; kurumunuzda bir eğitim aldınız mı?	Kurumunuzda; bu riske ilişkin yürürlükte olan bir talimat, politika, prosedür, vb mevcut mudur?	Kurumunuzun mevcut kontrollerini ve güvenlik önlemlerini de göz önüne alırsanız, bu riskin önem derecesi sizce nedir?
		(1 - 5 arası bir puan veriniz) 1: çok düşük 2: düşük 3: orta 4: yüksek 5:çokyüksek	(1 - 5 arası bir puan veriniz) 1: çok düşük 2: düşük 3: orta 4: yüksek 5:çokyüksek	(1 - 5 arası bir puan veriniz) 1: çok düşük 2: düşük 3: orta 4: yüksek 5:çokyüksek	(1 - 5 arası bir puan veriniz) 1: çok düşük 2: düşük 3: orta 4: yüksek 5:çokyüksek	(0 ile 4 arası bir puan veriniz, bilginiz yoksa "Fikrim Yok" diye yazınız) 0:hiç olmadı 1:bir kez oldu 2:iki ile beş arası 3:en az altı,en fazla on 4:ondan fazla	(0 ile 4 arası bir puan veriniz) 0:hiç yok 1:çok az 2: az 3: orta 4: çok	(E / H) 	(E / H) 	(1 - 5 arası bir puan veriniz) 1: çok düşük 2: düşük 3: orta 4: yüksek 5:çok yüksek
Personel	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin hastalık vb nedenlerle işe gelememesi	1	5	5	4	0	3	E	E	5
	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin deprem, sel, vb doğal afetlerde işe gelememesi	2	5	4	4	0	2	E	E	5
	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum içinden birilerinin hastanedeki gizli bilgileri dışarı sızdırması	3	5	4	4	0	2	E	E	2
	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum dışından birilerinin hastanedeki gizli bilgileri ele geçirmesi	3	5	5	4	0	3	H	H	2
Bilgisayarlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının bilgisayarlara teknolojik saldırıları	2	2	5	4	0	0	H	H	2
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların sistemlere bulaşması ve zarar vermesi	2	3	3	4	0	0	E	E	3
	Personelin bazen dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapma eğilimi sonucu oluşan kazalarda; bilgisayarlarda fiziksel hasar oluşması	2	2	2	2	1	3	E	E	1
	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik olması sonucunda; bilgisayarlarda arızaya yol açmaları	2	3	2	3	0	2	E	H	1
Hasta Kabul Modülü	Personelin dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapması sonucu; kayıt açma, güncelleme, doğrulama işlemlerinde oluşan yanlışlıklar ve tutarsızlıklar	3	3	3	4	3	3	E	E	2
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme kurum dışından yetkisiz kişilerin erişmesi	2	2	4	3	0	0	E	H	1
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme erişimde kurum içinde yetki süstimali yapılması	1	3	5	2	0	0	E	E	1
	Elektrik kesintisi, yangın, teknolojik altyapıdaki fiziksel hasar, vb nedeniyle hasta kabul modülünün çalışmaması	4	5	5	5	4	2	H	H	4
	Deprem, sel, vb doğal afetler nedeniyle hasta kabul modülünün çalışmaması	4	5	5	5	0	2	H	H	5
	Sistemlerde oluşacak çeşitli teknik hatalar veya arızalar nedeniyle hasta kabul modülünün çalışmaması	4	5	5	5	4	3	H	H	4

Şekil 4.5. Anket uygulamasının dördüncü bölümüne ilişkin örnek çıktı

Çizelge 4.1.' de anketteki hastane üst yönetimiyle ortak belirlenen öncelikli risk konularına ilişkin (anketin uygulanacağı bilgi işlem ve veri girişi personeli ve ilgili iş süreçlerine yönelik riskler) 30 adet risk sorusu yer almaktadır. Çizelge 4.2.'de de, her bir risk konusu için puanlanması / yanıtlanması istenen 9 adet risk ölçüm sorusu gösterilmektedir.

Çizelge 4.1. Anket uygulamasında kapsanan bilgi varlıkları ve ilgili riskler

Bilgi Varlığının Adı:	**İlgili olası Risk:**
Personel	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin hastalık vb nedenlerle işe gelememesi
	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin deprem, sel, vb doğal afetlerde işe gelememesi
	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum içinden birilerinin hastanedeki gizli bilgileri dışarı sızdırması
	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum dışından birilerinin hastanedeki gizli bilgileri ele geçirmesi
Bilgisayarlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının bilgisayarlara teknolojik saldırıları
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların sistemlere bulaşması ve zarar vermesi
	Personelin bazen dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapma eğilimi sonucu oluşan kazalarda; bilgisayarlarda fiziksel hasar oluşması
	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik olması sonucunda; bilgisayarlarda arızaya yol açmaları
Hasta Kabul Modülü	Personelin dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapması sonucu; kayıt açma, güncelleme, doğrulama işlemlerinde oluşan yanlışlıklar ve tutarsızlıklar
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme kurum dışından yetkisiz kişilerin erişmesi
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme erişimde kurum içinde yetki süistimali yapılması
	Elektrik kesintisi, yangın, teknolojik altyapıdaki fiziksel hasar, vb nedeniyle hasta kabul modülünün çalışmaması
	Deprem, sel, vb doğal afetler nedeniyle hasta kabul modülünün çalışmaması
	Sistemlerde oluşacak çeşitli teknik hatalar veya arızalar nedeniyle hasta kabul modülünün çalışmaması
Belgelerin kaşe-	Dikkatsizlik, telaş, vb nedenlerle; belgelerin kaybolması veya çalınması

imza-mühür işlemlerinin tamamlanması ve ilgili süreçler	Dikkatsizlik, telaş, vb nedenlerle; kaşe ve mühürlerin kaybolması veya çalınması
	Belgelere ait güncel ve eksiksiz bir envanterin bulunmaması nedeniyle; aksaklıklar yaşanması veya belgelerin kaybolması
	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin yazılı belgelere yanlış işlenmesi ve yanlış şekilde onaylanması
	Dikkatsizlik, telaş, vb nedenlerle; hatalı işlemler yapılması veya çeşitli aksamalar yaşanması
	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, vb nedenlerle; hatalı işlemler yapılması veya çeşitli aksamalar yaşanması
Kurumun personel istihdam süreçleri ve ilgili kontroller	Personelin istihdam sırasında yetkinlik ve güvenilirlik kontrollerinin yetersiz kalması nedeniyle; kurum içinde bilgi güvenliğine kasıtlı zarar verebilecek potansiyelde kişilerin görev alması
	Personelin istihdamı sonrası yeterli teknik ve diğer eğitimleri alamaması nedeniyle; kurumdaki bilgi sistemlerinde arıza, hata, aksaklığa yol açacak durumların oluşması
	Personelin istihdam sırasında yetkinlik kontrollerinin yetersiz kalması nedeniyle; kurumdaki bilgi sistemlerinde arıza, hata, aksaklığa yol açacak yetersiz / deneyimsiz kişilerin görev alması
Elektronik ortamdaki hasta kayıt verileri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının verilere erişmesi
	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların verilere zarar vermesi
	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin elektronik ortama yanlış aktarılması
	Personelin bazen dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapma eğilimi sonucu oluşan kazalarda; verilerde hata oluşması, verilerin zarar görmesi
	Elektrik kesintisi, yangın, teknolojik altyapıdaki veya sistemlerdeki çeşitli arıza veya hasarlar nedeniyle; verilerin zarar görmesi
	Deprem, sel, vb doğal afetler nedeniyle verilerin zarar görmesi
	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği sonucunda; verilere zarar verilmesi, verilerde uygun olmayan işlem yapılması

Çizelge 4.2. Anket uygulamasındaki risk ölçme ve değerlendirme soruları

Bu **riskin gerçekleşme olasılığı** nedir?	**Bilgi gizliliğinde yaratacağı kaybın / zararın** derecesi (maddi, manevi, yasal cezalar, vb toplamı)	**Bilginin bütünlüğü, tutarlılığı, doğruluğunda yaratabileceği kaybın / zararın** derecesi (maddi, manevi, yasal cezalar, vb toplamı)	**Bilginin ve ilgili işin / sistemin ulaşılabilirliği veya kullanılabilirliğine** verebileceği **kaybın / zararın** derecesi (maddi, manevi, yasal cezalar, vb toplamı)	Bu riskle ilgili; **kurumunuzda son 1 yılda yaşanan vaka adedi** nedir?	Bu riske ve bu riski azaltıcı çözümlere ilişkin; **herhangi bir bilgi birikiminiz** var mıdır?	**Bu riskle ilgili; kurumunuzda bir eğitim** aldınız mı?	Kurumunuzda; **bu riske ilişkin yürürlükte olan bir talimat, politika, prosedür, vb** mevcut mudur?	**Kurumunuzun mevcut koşullarını ve güvenlik önlemlerini de göz önüne** alırsanız, bu **riskin önem derecesi** sizce nedir?
(1 - 5 arası bir puan veriniz)	(1 - 5 arası bir puan veriniz)	(1 - 5 arası bir puan veriniz)	(1 - 5 arası bir puan veriniz)	(0 ile 4 arası bir puan veriniz, **bilginiz yoksa "Fikrim Yok" diye yazınız**)	(0 ile 4 arası bir puan veriniz)	(E / H)	(E / H)	(1 - 5 arası bir puan veriniz)
1: çok düşük	1: çok düşük	1: çok düşük	1: çok düşük	0:hiç olmadı	0:hiç yok			1: çok düşük
2: düşük	2: düşük	2: düşük	2: düşük	1:bir kez oldu	1:çok az			2: düşük
3: orta	3: orta	3: orta	3: orta	2:iki ile beş arası	2: az			3: orta
4: yüksek	4: yüksek	4: yüksek	4: yüksek	3:en az altı, en fazla on	3: orta			4: yüksek
5:çok yüksek	5:çok yüksek	5:çok yüksek	5:çok yüksek	4:ondan fazla	4: çok			5:çok yüksek

Ankette, önceden belirlenmiş olan 30 farklı riske ilişkin tehdit ve zayıflık unsurları ve bunların etkileşimi şeklinde tehdit ve zayıflık maddeleri yer almamaktadır. Sadece, bu tehdit ve zayıflıklara ilişkin oluşacak her bir risk ankete eklenmiş ve sorularda riskin kendisi anketi yanıtlayanlara soru olarak yöneltilmiştir. Bu yöntem; anketi yanıtlayanlara kolaylık sağlaması ve yanıtların daha sağlıklı, etkin bir şekilde verilmesi amacıyla uygulanmıştır ve özgün bir yaklaşımdır. Bilinen birçok bilgi güvenliği risk değerlendirme anketlerinde ve ilgili risk ölçümü / risk analizi yazılımlarında; tehditler ve zayıflıkların ayrı ayrı gruplanarak bunların etkileşimi sonucu oluşacak riskleri kullanıcıların anlaması ve puanlaması beklenmektedir. Bu çalışmadaki ankette ise, çalışanlara sadece cümleler şeklinde riskin kendisi sorulmuştur. Yanıtların değerlendirilmesi ve özdevimli öğrenmeye uyarlanması gibi çalışanların katılmadığı sonraki aşamalarda ilgili tehdit ve zayıflık ayrıştırması yapılmıştır.

Şekil 4.2.' de gösterilmiş olan anket uygulamasının birinci bölümündeki verilerin anket çalışmasının tamamlanması sonrası incelenmesi sonucunda, aşağıdaki basit tanımlayıcı istatistikler ve demografik sonuçlar elde edilmiştir:

```
Anketi yanıtlayan Erkek sayısı:     7
Yaş ortalaması (Erkek):             31.29
Görev Tanımına Göre Dağılımlar (Erkek):
            Bilgisayar Operatörü:               6
            Müdür Yardımcısı:                   1

Anketi yanıtlayan Kadın sayısı:     57
Yaş ortalaması (Kadın):             29.44
Görev Tanımına Göre Dağılımlar (Kadın):
            Bilgisayar Operatörü:               40
            Müdür Yardımcısı:                   0
            Veri Kayıt / Hazırlama Elemanı:     6
            Bilgisayar Elemanı:                 6
            Bilgisayar Elemanı (Danışma):       1
            Hastane Bilgi İşlem Sorumlusu:      1
            Tıbbi Sekreter:                     3

Toplam katılımcı sayısı:            64
Genel yaş ortalaması:               29.64
```

Görev Tanımına Göre Dağılımlar (Genel):

Bilgisayar Operatörü:	46
Müdür Yardımcısı:	1
Veri Kayıt / Hazırlama Elemanı:	6
Bilgisayar Elemanı:	6
Bilgisayar Elemanı (Danışma):	1
Hastane Bilgi İşlem Sorumlusu:	1
Tıbbi Sekreter:	3

Yukarıdaki sonuçlarda, genel olarak yedi farklı görev tanımı ve unvanı bulunmaktadır. Öte yandan, bu farklı görevdeki personelin hepsi aynı birime (bilgi işlem – veri kayıt) bağlı olarak görev yapmakta ve bu birim de hastane Müdür Yardımcısı'na bağlı olarak çalışmaktadır. Bu nedenle, anket çalışmasının tek bir bölüm / departman kapsamında yapıldığı kabul edilmektedir.

Benzer biçimde, Şekil 4.3 ve 4.4.' de gösterilmiş olan anket uygulamasının ikinci ve üçüncü bölümündeki verilerin anket çalışmasının tamamlanması sonrası incelenmesi sonucunda, aşağıdaki sonuçlar elde edilmiştir:

Anket Uygulamasının 2. Bölümü – Sonuçlar:

İlk seçeneği işaretleyen kişi adedi:	17
İkinci seçeneği işaretleyen kişi adedi:	27
Üçüncü seçeneği işaretleyen kişi adedi:	20
Toplam:	64

Anket Uygulamasının 3. Bölümü – Sonuçlar:

EN AZ SAYIDA OLAN

Kategori A'yı işaretleyen	Kategori B'yi işaretleyen	Kategori C' yi işaretleyen
23	18	**23**

DAHA FAZLA SAYIDA OLAN

Kategori A'yı işaretleyen	Kategori B'yi işaretleyen	Kategori C' yi işaretleyen
14	**23**	**27**

	EN ÇOK SAYIDA OLAN	
Kategori A'yı	Kategori B'yi	Kategori C' yi
işaretleyen	işaretleyen	işaretleyen
27	**23**	14

Anketin ikinci bölümünde elde edilen ve yukarıda özetlenen sonuçlara göre, ankete katılanların kendilerini tanımladıkları kişilik gruplarından en fazla sayıda olanı (27 kişi), ikinci seçenektir ("Son derece ihtiyatlı ve tedbirliyim, yeniliklere kolaylıkla açık olamam, risk almaktan hoşlanmam, bazen sıradan durumlarda bile garantiye ihtiyaç duyabilirim, belirsizlik korkusu hissederim, gelecekte oluşabilecek sorunlar için kötümser ya da endişeli olabilirim, adımlarımı dikkatle atarım ve iyice düşünüp taşınıp karar veririm, değişiklik ve yeniliklere yavaş yavaş uyum sağlayabilirim."). İkinci sırada toplam 20 kişiyle anketteki üçüncü seçenek olarak verilmiş olan kişilik tanımlamasıdır ("Fazlasıyla yardımseverim, başkalarını memnun etmek benim için çok önemlidir. Diğer insanlardan göreceğim onay, övgü ve takdire çok önem veririm ve bunun için elimden geleni yaparım. Çalışkan, sempatik, ve duygusal bir yapım vardır, tükenme noktasına gelene dek gayret gösteririm. İyi niyetimden dolayı, diğer insanların yalvarma ve rica gibi istekleri karşısında bazen kolaylıkla ikna olabilir veya istismar edilebilirim."). Anketin ikinci bölümünde az tercih edilen grup, birinci seçenek olan ("Araştırıcı ve yeniliklere açık bir yapım var, bazen bu nedenle çok gerekli durumlarda fazla düşünmeden hızlıca karar verebilirim, fikirlerimi hızlıca değiştirebilirim, bazen çabuk tahrik olup kolaylıkla sinirlenebilirim, yeni ilgi ve aktivitelere kolaylıkla uyum sağlayıp benimseyebilirim, ancak bazen ilgim çabuk kaybolabilir ve sıkılabilirim, rutin işler bazen beni çok bunaltabilir, çabuk davranıp sezgilerime göre karar verebilirim, enerjik ve aktif bir yapım var, sıklıkla çok konuşurum ve kolay heyecanlanırım.") tanımıdır ve bunu toplamda 17 kişi tercih etmiştir. Öte yandan, bir amacı da anketin ikinci bölümünün sağlaması ve kontrol soruları olan üçüncü bölüm ile ikinci bölümdeki yanıtlar karşılaştırıldığında 2. ve 3. bölümler arası tutarsızlık olduğu görülmüştür. Ankete katılanların yanıtlarının; kişilik gruplarını ayırt edemediği ve buna göre risk değerlendirmesinde ağırlıklı katsayı olarak kullanabilecek anlamlı bir sayısal değer sağlayamadığı bulgulanmıştır. Ayrıca, anketin üçüncü bölümündeki yanıtların; ikinci bölümden bağımsız olarak kendi içinde dağılımları çapraz analiz ve karşılaştırmalı olarak incelendiğinde de ayırt edici ve anlamlı dağılımlar elde edilemediği

görülmektedir. Bu nedenle, anketin ikinci ve üçüncü bölümündeki yanıtlar; çalışmadaki risk değerlendirme modelinin oluşturulması ve özdevimli öğrenmeye uyarlanması süreçlerine herhangi bir etmen olarak katılmamış, diğer risk etmenlerinin puanlarını değiştirecek biçimde ağırlık katsayısı olarak kullanılmamıştır.

4.3. ANKET SONUÇLARININ ÖZDEVİMLİ ÖĞRENMEYE UYARLANMASI

Toplam 64 kişiye uygulanan anket sonucunda, 1920 farklı örneklem kaydı (instance) elde edilmiştir. Çalışmadaki özdevimli öğrenme modeli ikili sınıflandırıcı algoritmalar kullanılarak geliştirileceği için, anketteki son soru olan ve 1-5 arası puanlanmış olan "Kurumunuzun mevcut koşullarını ve güvenlik önlemlerini de göz önüne alırsanız, bu riskin önem derecesi sizce nedir?" sonuçları, ikili değere dönüştürülmüştür. Hastane yetkilileriyle yapılan ortak çalışma sonucu, bu soruya (özdevimli öğrenmede sınıf ayırımı ve risk kestirimi için kullanılacak sonuç parametresi) 1, 2 veya 3 puanı verilenler, alt düzey risk (kurumun risk eşik düzeyi altında, göz ardı edilebilir düzeyde riskler) "A" olarak tanımlanmış, 4 veya 5 puanı verilenler de risk eşik düzeyinin üstü "U" olarak tanımlanmıştır. Böylece, 1920 adet risk örneğinin 1099 tanesi kabul / göz ardı edilebilir risk düzeyinin üstü "U", geri kalan 821 tanesi de "A" olarak sınıflandırılmıştır. Ayrıca, ankette risk sorusu olarak sorulmuş olan 30 risk konusunun her birisi, ilgili zayıflık ve tehdit ilişkisine göre tekrar tanımlanmıştır. Çizelge 4.3.'de anketteki risklerin ilgili tehdit ve zayıflıklara ayrıştırılmış listesi verilmektedir. Böylelikle, risk değerlendirme modelinde 6 adet bilgi varlığı, 10 farklı tehdit, 9 farklı zayıflık tanımlanmış ve bunların her birisi için bir endeks kod değeri belirlenmiştir. Çizelge 4.4., 4.5. ve 4.6.'da buna ilişkin bilgiler yer almaktadır. Şekil 4.6.'da ise anketteki toplam 12 parametre için verilen yanıtların dağılımlarının adetleri gösterilmektedir.

Çizelge 4.3. Anketteki risklerin ilgili tehdit ve zayıflıklara ayrıştırılmış listesi

Tehdit	**Zayıflık**	**İlgili olası Risk**
Hastalık	Kurumda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin hastalık vb nedenlerle işe gelememesi
Deprem, sel, vb doğal afetler	Kurumda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması	Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin deprem, sel, vb doğal afetlerde işe gelememesi
Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum içinden birilerinin hastanedeki gizli bilgileri dışarı sızdırması
Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum dışından birilerinin hastanedeki gizli bilgileri ele geçirmesi
Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının bilgisayarlara teknolojik saldırıları
bilgisayar virüsleri, vb zararlı kodlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların sistemlere bulaşması ve zarar vermesi
Fiziksel hasarlar, yangınlar, kazalar, vb	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin bazen dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapma eğilimi sonucu oluşan kazalarda; bilgisayarlarda fiziksel hasar oluşması
Yazılım vb de teknik hatalar, arızalar	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik vb durumlar	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik olması sonucunda; bilgisayarlarda arızaya yol açmaları
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapması sonucu; kayıt açma, güncelleme, doğrulama işlemlerinde oluşan yanlışlıklar ve tutarsızlıklar
Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme kurum dışından yetkisiz kişilerin erişmesi
Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; sisteme erişimde kurum içinde yetki süistimali yapılması

Fiziksel hasarlar, yangınlar, kazalar, vb	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Elektrik kesintisi, yangın, teknolojik altyapıdaki fiziksel hasar, vb nedeniyle hasta kabul modülünün çalışmaması
Deprem, sel, vb doğal afetler	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Deprem, sel, vb doğal afetler nedeniyle hasta kabul modülünün çalışmaması
Yazılım vb de teknik hatalar, arızalar	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Sistemlerde oluşacak çeşitli teknik hatalar veya arızalar nedeniyle hasta kabul modülünün çalışmaması
Bilgi, belge, veri kaybı	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Dikkatsizlik, telaş, vb nedenlerle; belgelerin kaybolması veya çalınması
Eksik, kayıp Malzeme	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Dikkatsizlik, telaş, vb nedenlerle; kaşe ve mühürlerin kaybolması veya çalınması
Bilgi, belge, veri kaybı	Güncel ve tutarlı , güvenilir envanter olmaması	Belgelere ait güncel ve eksiksiz bir envanterin bulunmaması nedeniyle; aksaklıklar yaşanması veya belgelerin kaybolması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Karmaşık ve özel dil, terminoloji kullanımının getirdiği zorluklar ve karışıklıklar	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin yazılı belgelere yanlış işlenmesi ve yanlış şekilde onaylanması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Dikkatsizlik, telaş, vb nedenlerle; hatalı işlemler yapılması veya çeşitli aksamalar yaşanması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik vb durumlar	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, vb nedenlerle; hatalı işlemler yapılması veya çeşitli aksamalar yaşanması
Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri	İstihdam öncesi ilgili süreçlerde ve kontrollerdeki eksiklikler	Personelin istihdam sırasında yetkinlik ve güvenilirlik kontrollerinin yetersiz kalması nedeniyle; kurum içinde bilgi güvenliğine kasıtlı zarar verebilecek potansiyelde kişilerin görev alması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	İstihdam süresi boyunca gerekli eğitimlerin verilmemesi	Personelin istihdamı sonrası yeterli teknik ve diğer eğitimleri alamaması nedeniyle; kurumdaki bilgi sistemlerinde arıza, hata, aksaklığa yol açacak durumların oluşması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	İstihdam öncesi ilgili süreçlerde ve kontrollerdeki eksiklikler	Personelin istihdam sırasında yetkinlik kontrollerinin yetersiz kalması nedeniyle; kurumdaki bilgi sistemlerinde arıza, hata, aksaklığa yol açacak yetersiz / deneyimsiz kişilerin görev alması
Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının verilere erişmesi

bilgisayar virüsleri, vb zararlı kodlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların verilere zarar vermesi
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Karmaşık ve özel dil, terminoloji kullanımının getirdiği zorluklar ve karışıklıklar	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin elektronik ortama yanlış aktarılması
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin bazen dikkatsizlik, telaş, vb nedenlerle kasıtlı olmadan hata yapma eğilimi sonucu oluşan kazalarda; verilerde hata oluşması, verilerin zarar görmesi
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Elektrik kesintisi, yangın, teknolojik altyapıdaki veya sistemlerdeki çeşitli arıza veya hasarlar nedeniyle; verilerin zarar görmesi
Deprem, sel, vb doğal afetler	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Deprem, sel, vb doğal afetler nedeniyle verilerin zarar görmesi
Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik vb durumlar	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği sonucunda; verilere zarar verilmesi, verilerde uygun olmayan işlem yapılması

Çizelge 4.4. Anketteki bilgi varlıklarının özdevimli öğrenme için yapılandırılması

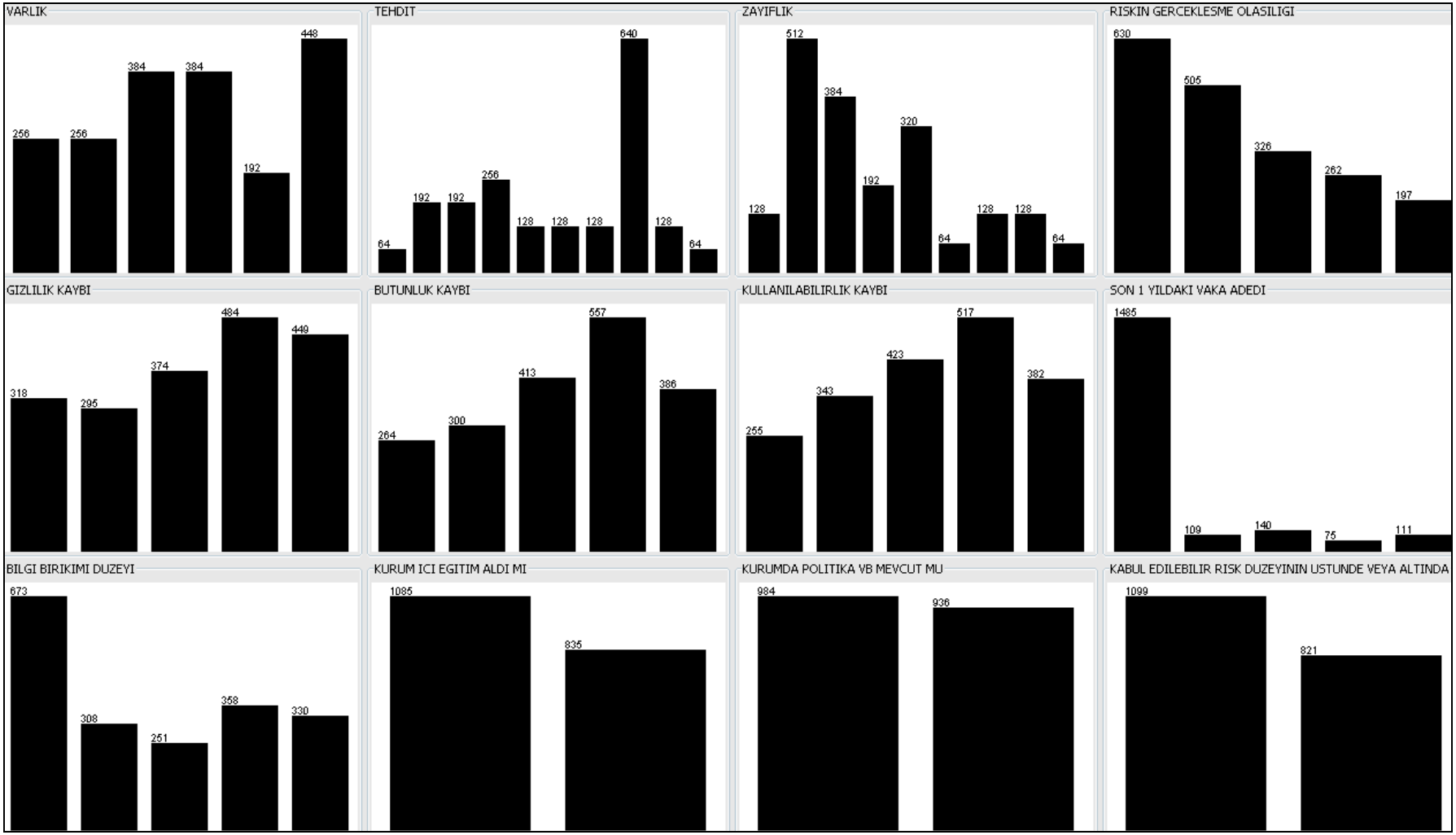
Bilgi Varlığı Kodu:	**Bilgi Varlığının Adı:**
V1	Personel
V2	Bilgisayarlar
V3	Hasta Kabul Modülü
V4	Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler
V5	Kurumun personel istihdam süreçleri ve ilgili kontroller
V6	Elektronik ortamdaki hasta kayıt verileri

Çizelge 4.5. Tehditlerin özdevimli öğrenme için yapılandırılması

Tehdit Kodu	**Tehdit**
T1	Hastalık
T2	Deprem, sel, vb doğal afetler
T3	Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri
T4	Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri
T5	Bilgisayar virüsleri, vb zararlı kodlar
T6	Fiziksel hasarlar, yangınlar, kazalar, vb
T7	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem
T8	Yazılım vb de teknik hatalar, arızalar
T9	Bilgi, belge, veri kaybı
T10	Eksik, kayıp Malzeme

Çizelge 4.6. Zayıflıkların özdevimli öğrenme için yapılandırılması

Zayıflık Kodu	**Zayıflık**
Z1	Kurumda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması
Z2	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı
Z3	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği
Z4	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik vb durumlar
Z5	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı
Z6	Güncel ve tutarlı , güvenilir envanter olmaması
Z7	Karmaşık ve özel dil, terminoloji kullanımının getirdiği zorluklar ve karışıklıklar
Z8	İstihdam öncesi ilgili süreçlerde ve kontrollerdeki eksiklikler
Z9	İstihdam süresi boyunca gerekli eğitimlerin verilmemesi



Şekil 4.6. Anketteki 1920 adetlik veri kümesinin yanıtlara göre sayısal dağılımları

Anket sonuçlarının özdevimli öğrenme yaklaşımına uyarlanması ve uygulanması aşaması için, çeşitli özdevimli öğrenme yazılımı ve araçları önceden incelenmiş ve Weka yazılımının kullanılmasına karar verilmiştir. Weka yazılımı, bu çalışmanın yapıldığı tarihlerde uluslararası düzeyde en çok kabul gören ve akademik araştırma projelerinde sıkça kullanılan açık kaynak kodlu, parasız şekilde serbest kullanıma açık veri madenciliği ve özdevimli öğrenme yazılımlarından birisidir (Khattak, vd., 2010). Weka yazılımı, Yeni Zelanda'daki Waikato Üniversitesi (University of Waikato) akademisyenleri tarafından üretilmiş olan ve sürekli geliştirilerek desteklenen bir yazılımdır. Bu yazılım, veri madenciliği ve özdevimli öğrenme amaçlı birçok algoritma, seçenek ve araç içermekte olup Java programlama dili ile geliştirilmiştir (Witten, vd., 2011). Çalışmanın yapıldığı dönemde en güncel ve güvenilir sürümü (version) olarak kabul gören Weka 3.6 sürümü ve bunun alt sürümü (subversion) olan 3.6.0. bu çalışmada kullanılmıştır. (Weka, 2011).

Çizelgelerde gösterilen kod tanımlamalarının yapılması, Weka yazılımına veri girdisi olarak tanımlanırken sistemin ve modelin etkinliğini ve hızını arttırmak ve standart bir yapıya kavuşturabilmek için gerekli bir işlemdir. Anketlere verilen yanıtlar yeniden yapılandırıldıktan sonra, tüm MS Excel'deki kayıtlar, virgül ayraçlı (comma seperated file, .csv) biçimine dönüştürülmüş ve bu şekilde Weka yazılımına girdi verisi olarak yüklenmiştir. Ham verilerde, yanıtlar ve puanların bir kısmı Weka sisteminde otomatik olarak sayısal (numeric) tanımlanmıştır. Bu veriler ve tüm örneklem kümesi, ikili sınıflandırıcıları hatasız ve verimli kullanabilmek için, sayısal olmayan (nominal) biçime dönüştürülmesi Weka yazılımının ilgili ara yüzünden sağlatılmış ve sonra da tüm veri kümesi (1920 adet kayıt) Weka yazılımının kendi veri biçimi olan ".arff" dosyasına dönüştürülmüştür. Şekil 4.7.' de Weka girdi (input) veri dosyası (.arff) biçimi ve içeriğine ilişkin örnek gösterilmektedir.

```
@relation data-forWeka-V1-nounknowns-
weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last

@attribute VARLIK {V1,V2,V3,V4,V5,V6}
@attribute TEHDIT {T1,T2,T3,T4,T5,T6,T8,T7,T9,T10}
@attribute ZAYIFLIK {Z1,Z2,Z3,Z4,Z5,Z6,Z7,Z8,Z9}
@attribute 'RISKIN GERCEKLESME OLASILIGI' {1,2,3,4,5}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'SON 1 YILDAKI VAKA ADEDI' {0,1,2,3,4}
@attribute 'BILGI BIRIKIMI DUZEYI' {0,1,2,3,4}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'KURUMDA POLITIKA VB MEVCUT MU' {H,E}
@attribute 'KABUL EDILEBILIR RISK DUZEYININ USTUNDE VEYA ALTINDA' {U,A}

@data
V1,T1,Z1,5,4,3,2,4,2,H,H,U
V1,T2,Z1,3,2,2,3,1,1,H,H,U
V1,T3,Z2,5,5,5,5,3,0,E,E,U
V1,T4,Z2,5,5,5,5,4,0,H,H,U
V2,T4,Z3,3,3,3,2,3,1,H,H,A
V2,T5,Z3,5,5,5,4,3,2,E,E,U
V2,T6,Z2,3,3,3,3,4,2,H,H,A
V2,T8,Z4,4,5,4,5,2,2,H,H,A
V3,T7,Z2,5,5,4,5,4,4,H,H,U
V3,T4,Z3,4,4,4,3,3,2,E,E,A
V3,T3,Z3,5,4,2,3,2,1,H,H,A
V3,T6,Z5,5,5,5,4,3,2,H,H,A
V3,T2,Z5,5,5,4,5,3,4,E,E,U
V3,T8,Z5,5,5,4,5,3,4,E,E,U
V4,T9,Z2,4,5,5,5,0,0,H,H,U
V4,T10,Z2,4,5,5,5,0,0,H,H,U
V4,T9,Z6,4,5,5,5,0,0,H,H,U
```

Şekil 4.7. Weka girdi veri dosyası (.arff) içeriğine ilişkin örnek

4.4. MODELİN GELİŞTİRİLMESİ VE UYGULANMASI

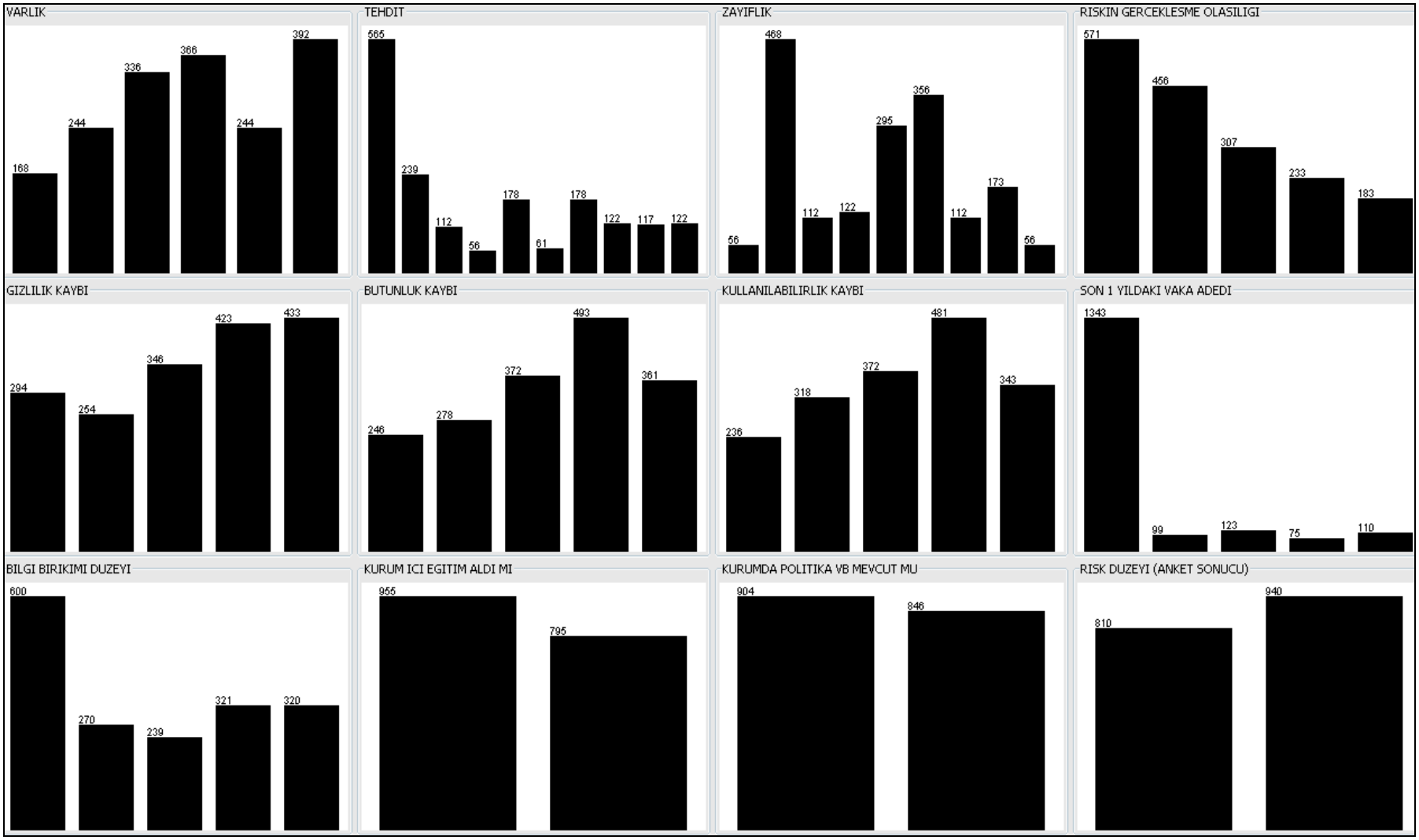
Weka için uygun şekle getirilen anket veri kümesi üzerinde, ilgili ikili sınıflandırıcı algoritmaların tek tek seçilip hem eğitim / öğrenim (train) hem de sınama (test) amaçlı denenmesi ve başarım değerlerinin gözlemlenmesi yapılmıştır. İlk aşamada orijinal veri kümesi 1920 adet kayıttan oluşurken, yapılan gözlemlerde elde edilen sayısal sonuçlar ve sonrasında anket yanıtlarının anket kâğıt formları ele alınarak incelenmesi ile bazı kişilerin ankete gelişigüzel anlamsız yanıtlar ve puanlamalar verdiği (biased answers) bulgulanmıştır. Bu işlemde öncelikle rastgele seçilen Weka yazılımında tanımlanmış 13 adet farklı sınıflandırıcının (AODE, AODEsr, BayesNet, BayesHNB, NaiveBayes, WAODE, DecisionTree, Functional Tree, J48, LogisticRegression, NBTree, NestedDicthotomies-ClassBalanced, SimpleLogisticRegression) risk sınıflandırmaları

ile anketteki risk yanıtları MS Excel tablosunda karşılaştırılmıştır. Şekil 4.8.'de bu karşılaştırmanın bir örneği gösterilmektedir.

Sınıflandırıcıların tamamında ortak şekilde bulgulanan ve aynı zamanda anketteki risk yanıtlarından farklı olanlar gruplanmış ve sonra da bu yanıtları içeren anket formları tek tek incelenmiştir. Bunun sonucunda, sınıflandırıcıların risk tahminlerinden farklılık gösteren yanıtları içeren ilgili anket formlarında kullanıcıların gelişigüzel anlamsız yanıtlar ve puanlamalar verdiği bulgulanmıştır. Toplam 170 adet yanıt veri kümesinden çıkarılmış ve böylece toplam 1750 kayıttan oluşan yeni bir veri kümesi elde edilmiştir. 1750 adet risk örneğinin 940 tanesi kabul / göz ardı edilebilir risk düzeyinin üstü "U", geri kalan 810 tanesi de "A" olarak sınıflandırılmıştır. Şekil 4.9.'da 1750 adetlik yeni veri kümesindeki 12 parametre için verilen yanıtların dağılımlarının adetleri gösterilmektedir. Bu 1750 örnek üzerinden 68 değişik ikili sınıflandırıcı özdevimli öğrenim algoritması ve bunların alt seçenekleri, değişik parametre kombinasyonlarıyla beraber 3200 civarında değişik gözlem ve test yapılmıştır. Elde edilen başarım değerlerine göre, en başarılı sınıflandırıcıların kestirimde bulunduğu üst düzey riskler (U kodlananlar) gruplanmış, kesişim kümeleri alınarak her bir risk için toplam adetler belirlenmiş ve bu adetlere göre riskler küçükten büyüğe sıralanmıştır. Sıralanan 30 farklı risk, Kartil yöntemiyle dört kesite ayrılarak gruplanmış, dağılımda % 50 ve üzeri grubunda kalan riskler belirlenmiş ve böylece Karşıyaka Devlet Hastanesi için hangi risklerin daha öncelikli ve daha yüksek düzeyde riskli oldukları ortaya çıkarılmıştır. Şekil 4.10.'da Kartil yöntemiyle risklerin sıralanması ve üst düzey risk grubuna girenlerin belirlenmesi işlemi gösterilmektedir. Şekil 4.10.'da risklerin adları yerine; her bir riske ait bilgi varlığı, tehdit ve zayıflık sınıflarının kısaltılmış endeks kodlarıyla (Çizelge 4.4, 4.5 ve 4.6.'da bu kodlar ve karşılık gelen kavramlar tanımlıdır) betimlenerek gösterimi yapılmaktadır.

RISK DUZEYI (ANKET SONUCU)	RISK TAHMINI: AODE	RISK TAHMINI: AODEsr	RISK TAHMINI: BayesNet	RISK TAHMINI: BayesHNB	RISK TAHMINI: NaiveBayes	RISK TAHMINI: WAODE	RISK TAHMINI: DecisionTree	RISK TAHMINI: FT (FUNCTIONAL TREE)	RISK TAHMINI: J48	RISK TAHMINI: Logistic Regres.	RISK TAHMINI: NBTree	RISK TAHMINI: Nesteddict.class balanced	RISK TAHMINI: SimpleLogisticReg.
A	A	A	A	A	A	A	U	A	U	U	A	U	U
U	U	U	A	U	A	U	A	U	U	U	A	U	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
A	U	U	U	U	U	U	U	A	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	U	A	A	A	U	A	A	A	A	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	A	A	A	A	A	A	A	A	U	A	A	U	A
A	A	A	A	U	A	A	U	A	A	A	A	A	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	A	U	U	U	U	U
U	U	U	U	U	U	U	U	A	U	U	U	U	U
U	U	U	A	U	A	U	A	U	U	U	A	U	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
A	A	A	U	A	U	A	A	A	A	U	U	A	U
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	U	U	U	U	U	U	U	A	U	U	U	U	U
U	U	U	U	U	U	U	U	A	U	U	U	U	U
A	A	A	A	A	A	A	A	A	A	U	A	A	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A	U	A	A	A	A	A
A	U	U	U	A	U	U	U	A	U	U	U	U	U
A	A	A	A	A	A	A	A	A	A	A	A	A	A
A	A	U	U	U	U	A	U	A	U	U	U	U	U
U	A	A	A	A	A	A	A	U	A	A	A	A	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	A	U	U	A
U	U	U	U	U	U	U	U	U	U	U	U	U	U
U	U	U	U	U	U	U	U	U	U	U	U	U	U

Şekil 4.8. Anketteki yanıtlar ile sınıflandırıcıların kestirimlerinin karşılaştırılması



Şekil 4.9. Anketteki 1750 adetlik veri kümesinin yanıtlara göre sayısal dağılımları

Varlık Kodu:	Tehdit Kodu:	Zayıflık Kodu:	risk=U(üst) olanların adedi
V2	T8	Z4	1
V2	T6	Z2	3
V1	T1	Z1	5
V3	T7	Z2	8
V1	T2	Z1	11
V3	T3	Z3	11
V3	T4	Z3	12
V3	T6	Z5	13
V4	T7	Z2	17
V5	T7	Z9	18
V6	T4	Z3	18
V3	T8	Z5	19
V4	T7	Z4	19
V6	T7	Z2	19
V2	T4	Z3	20
V4	T9	Z6	20
V3	T2	Z5	22
V5	T3	Z8	22
V6	T2	Z5	22
V4	T10	Z2	23
V6	T7	Z4	23
V2	T5	Z3	25
V6	T5	Z3	25
V6	T7	Z5	25
V6	T7	Z7	25
V1	T3	Z2	26
V4	T7	Z7	26
V5	T7	Z8	26
V4	T9	Z2	28
V1	T4	Z2	36
TOPLAM:			568
Aritmetik ortalama			18,933
Ortanca			20
Standart Sapma			7,961

Açıklama: 25 farklı sınıflandırıcının tamamında ortak sonuç olarak Risk = U bulunanlar için Kartil yöntemi ile analiz

Kartiller (Quartiles)

Q1 %25	Q2 %50	Q3 %75	Q4 %100

1 14 20 25 3

Şekil 4.10. Kartil yöntemiyle risklerin sıralanması ve gruplanması

Çalışma sonucunda en başarılı doğruluk değerlerini sağlayan sınıflandırıcılar için, bir başka çalışma daha yapılarak bu sınıflandırıcıların öğrenme doygunluğuna ulaşıp ulaşmadıkları da belirlenmiştir. Bunun için, değişik, artan sayıda ve belli boyutlarda rastsal örneklem kümeleri elde edilmiş ve her bir sınıflandırıcı için öğrenme doğruluk başarım değerleri ölçülmüş, bu değerlere göre de her bir sınıflandırıcı için öğrenme eğrileri (learning curve) oluşturulmuştur (Weiss, 2003). En başarılı kabul edilen sınıflandırıcıların öğrenme eğrilerinin incelenmesi sonucunda tümünün belli düzeyde

öğrenme doygunluğuna (happy graph) eriştiği bulgulanmıştır (Russell, 2009). Öğrenme eğrilerine ilişkin sonuçlar, 5. bölümde ilgili şekillerde özetlenmektedir.

Çalışmada ortaya konan model ve ilgili yöntemin aşamaları aşağıda maddeler halinde özetlenmektedir:

1. Bilgi güvenliği nitel risk analizi değerlendirmesi için gerekli ölçüm veya örneklem verileri oluşturulur. Bu aşamada, her bir riske ilişkin etki eden etmenler (faktörler) tanımlanır:
    a. Riskin gerçekleşme olasılığı (olabilirlik nitel değeri),
    b. Bilgi gizliliğinde yaratacağı kaybın / zararın derecesi,
    c. Bilginin bütünlüğünde yaratabileceği kaybın / zararın derecesi,
    d. Bilginin ve ilgili işin / sistemin ulaşılabilirliği veya kullanılabilirliğine verebileceği kaybın / zararın derecesi,
    e. Riskle ilgili kurumda son 1 yılda yaşanan vaka / olay adedi,
    f. Riske ilişkin bilgi birikimi düzeyi,
    g. Riskle ilgili eğitim alınıp alınmadığı,
    h. Riske ilişkin bir politika veya yönergenin yürürlükte olup olmadığı
2. Nitel risk değerlendirmesi ön çalışması yapılır; Bu aşamada, 1-5 arası ölçekte ilgili her bir riskin öngörülen değeri alınır, kurumun kabul edilebilir risk eşik değerinin 1-5 arası hangi değer olduğu belirlenir ve buna göre risk değerleri iki sınıfa ayrılır. Risk eşik değerinin üstünde kalan riskler "üst" sınıfta, geri kalanlar "alt" sınıfta gruplanacak şekilde tanımlanır.
3. Veri kümesinin, modelde seçilen özdevimli öğrenme yazılımına uygun biçime (format) dönüştürülmesi sağlanır ve ilgili yazılıma girdi (input) olarak aktarılır.
4. Özdevimli öğrenmede seçilen ikili sınıflandırıcı farklı algoritmalarla öğrenme (train) analizi yapılır. Alt ve üst riskte tahmin edilen risklere ilişkin sonuçlar, modelin 2. aşamasındaki ön analizde tanımlanmış olan risk sonuçları ile karşılaştırılır. Tüm sınıflandırıcı algoritmaların öğrenme sonuçları ile ön analizdeki sonuçların tutarsızlık gösterdiği örnekler bulgulanırsa, bunlar incelemeye alınır. Modelin 1. aşamasında anket, vb yöntemlerle elde edilen veri

kümesinde hatalı, yanıltıcı (biased) örneklerin olup olmadığını anlamak için bu inceleme yapılır. Eğer bu tarz yanıltıcı, tutarsız örnekler var ise, bunlar veri kümesinden ayıklanır veya olanak var ise değiştirilir.

5. Düzeltilen veri örneklem kümesi için, modeldeki 3. aşama tekrar işletilir.
6. Özdevimli öğrenmede seçilen ikili sınıflandırıcı farklı algoritmalarla öğrenme (train) ve sınama (test) analizleri yapılır. Elde edilen sonuçlara göre, algoritmaların doğru pozitif ve doğru negatif değerleri için doğruluk (accuracy) en yüksek başarım (performance) değerleri vermesi için ek deneyler yapılır.
7. Deneyler sonunda en yüksek öğrenme ve test başarım değerlerini veren algoritmalar belirlenir. Test başarım değerlerini elde etmek için iki alternatiften biri kullanılır:

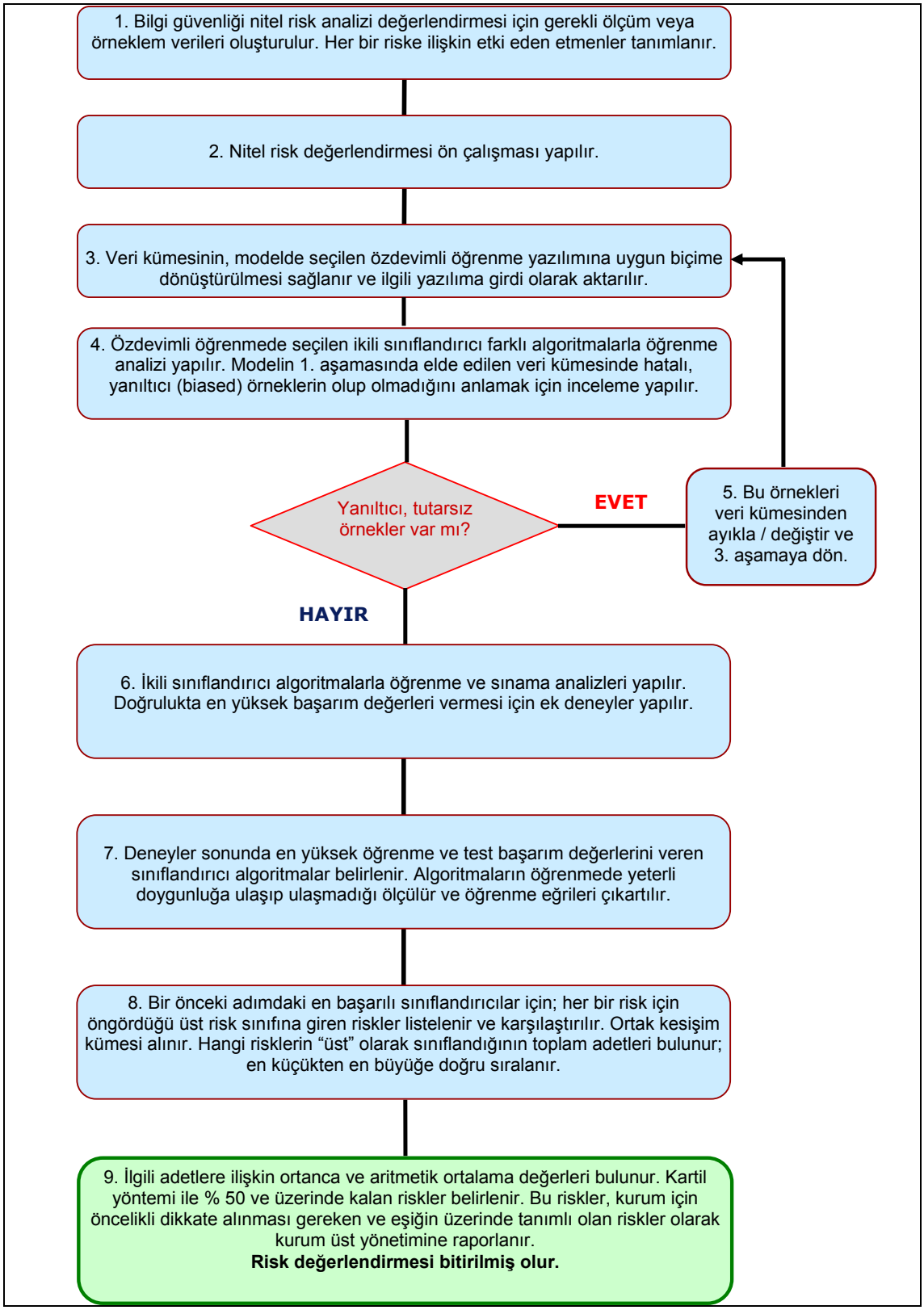
   Seçenek a: Modelde kullanılan örnek veri kümesine test amaçlı yeni veriler eklenmesi olanaklı değilse, onlu çapraz doğrulama (10-fold cross-validation) yöntemi ile testler gerçekleştirilir (Sammut ve Webb, 2011).

   Seçenek b: Modelde kullanılan örnek veri kümesine test amaçlı yeni veriler eklenmesi olanaklı ve bu veri kümesinin boyutu, öğrenme aşamasındaki veri kümesinin en az % 33'ü civarı bir büyüklükte ise, bu yeni veri kümeleri test amaçlı kullanılır.

   Belirlenen algoritmaların öğrenmede yeterli doygunluğa ulaşıp ulaşmadığı (happy graph) tüm veri kümesinden rastsal olarak farklı boyutlardaki alt örneklem kümeleri ile ölçülür ve öğrenme eğrileri (learning curves) çıkartılır.
8. Modelin 7. aşamasındaki sonuçlara göre, model için en uygun algoritmalar belirlenir. Bu algoritmaların her bir risk için öngördüğü üst risk sınıfına giren riskler listelenir ve karşılaştırılır. Ortak kesişim kümesi alınır. Bu kesişim kümesindeki tüm örnek uzayına ait risklerin hangilerinin "üst" sınıfta tanımlandığı ve bunların toplam adetleri bulunarak en küçükten en büyüğe sıralanır.
9. İlgili adetlere ilişkin ortanca (median) ve aritmetik ortalama (mean) değerleri bulunur. Kartil (quartile) yöntemi ile % 50 ve üzerinde kalan riskler belirlenir ve bu riskler, kurum için öncelikli dikkate alınması gereken ve eşiğin üzerinde

tanımlı olan riskler olarak kurum üst yönetimine raporlanır. Risk değerlendirmesi aşaması bitirilmiş olur.

Ortaya konan modelin 7. aşamasında tanımlanan alternatif test seçenekleri ve bunlardan hangisinin seçileceğine ilişkin yöntembilimi, ilgili kaynaklardaki temel kuramsal bilgilere dayandırılmaktadır (Witten, vd., 2011, Refaeilzadeh, vd., 2008, Perlich, vd., 2003). Çalışmada ortaya konan modelin süreç şeması biçiminde gösterimi, Şekil 4.11.'de verilmiştir.



Şekil 4.11. Özdevimli öğrenme yaklaşımı ile bilgi güvenliği risk değerlendirmesi modeli

5. SONUÇLAR VE TARTIŞMA

Çalışma sonucunda, özdevimli öğrenimdeki ikili sınıflandırıcı algoritmalardan en başarılı öğrenim ve test değerlerinin elde edildiği algoritmalar Çizelge 5.1.'de gösterilmektedir. Her sınıflandırıcı algoritmanın öğrenme ve test deneyinin ne kadar sürede gerçekleştiği bilgisi, en iyi başarım değerini vermesini sağlayan alt yordam ve parametre, işlev değişkenlerinde yapılan değişiklikler ve varsayılan değerler (vd olarak kısaltma ile betimlenmiştir) de Çizelge 5.1.' de belirtilmektedir. Bu on adet sınıflandırıcı kadar kayda değer başarım değerleri sağlamayan, fakat yapılan tüm gözlemlerde diğer sınıflandırıcılardan öne çıkan 15 sınıflandırıcı algoritma daha bulgulanmıştır. Toplam 25 algoritmanın hepsinin de risk değerini üst (U) olarak bulguladığı kesişim kümesi sonuçlarına göre, anketin uyguladığı hastane için en öncelikli ve risk işleme sürecine girmesi gereken toplam 16 adet risk ortaya çıkarılmıştır. Bu 16 adet üst düzey riskin listesi de Çizelge 5.2.' de verilmektedir.

Çizelge 5.1.'de Weka yazılımında tanımlandığı şekliyle kısaltmalarla ifade edilmiş sınıflandırıcı algoritmalarla ilgili kısa bilgiler ve açıklamalar aşağıda maddeler halinde verilmektedir. Bu algoritmalar, işlevleri, kullandıkları denklemler, parametreler hakkında detaylı bilgiler ilgili kaynaklardan incelenebilir (Witten, vd., 2011, Weka, 2011).

Meta-RandomSubspace: Karmaşık karar ağaçları mekanizmasına sahiptir. Çeşitli sınıflandırıcı algoritmalarla birlikte kullanılabilir.

RepTree: Hızlı öğrenebilen bir karar ağacı algoritmasıdır. Karar / regresyon ağaçlarını bilgi kazançlarına göre kurar.

J48: Weka yazılımına özgü bir karar ağacı algoritması olup C4.5 karar ağacı algoritmasının geliştirilmiş bir türevidir. Budamalı veya budamasız çalışacak şekilde farklı seçeneklerle kullanılabilir.

SMO (Sequential Minimized Optimization): Destek Vektör Makinesi sınıflandırıcısının başarım düzeyini arttırmak için geliştirilmiş bir algoritmadır.

Lazy-Ibk: k-En Yakın Komşu sınıflandırıcısının Weka için geliştirilmiş bir modelidir.

FT (Functional Tree): Karar ağacı sınıflandırıcısı grubunda bir algoritmadır. İçteki düğümler veya yapraklarda lojistik regresyon işlevleri tanımlanır ve işletilir.

RandomForest: Çeşitli rastsal karar ağaçlarının oluşturduğu bir ağaç kümesi şeklinde çalışır.

Çizelge 5.1. En iyi başarım değerlerinin gözlemlendiği sınıflandırıcılar

Algoritma Adı ve Türü:	**Parametreler / alt işlevler**	**Öğrenme Başarımı**	**Test Başarımı**	**Ek Bilgiler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: RepTree	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). RepTree parametreleri içinde; No Pruning: True (vd : false) Diğer tüm RepTree parametreleri aynı (vd).	0.957	0.782	Hız sorunu gözlenmemiştir. (2 ile 4 dk) Overfitting riski vardır.
Meta-RandomSubspace Alt sınıflayıcı algoritma: J48	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). J48 confidence factor: 0.25 (No pruning olduğundan bu parametre etkisizdir) Unpruned: True (vd: False) Diğer tüm parametreler aynı (vd).	0.938	0.786	Hız sorunu gözlenmemiştir. (1.5 ile 2 dk)
Meta-RandomSubspace Alt sınıflayıcı algoritma: J48	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). J48 confidence factor: 0.75 (vd: 0.25) Binary splits: True (vd: False) Diğer tüm parametreler aynı (vd).	0.929	0.796	Öğrenme kısmı 5 dakika, test kısmı 58 dakika sürdü. Yavaş çalıştı.
SMO (Sequential Minimized Optimization)	kernel: Normalized PolyKernel (vd: PolyKernel) Exponent: 3.5 (vd: 2.0) FilterType: Normalize Traning Data. Diğer tüm parametreler aynı (vd).	0.902	0.778	Hız sorunu gözlenmemiştir. (45 sn ile 2 dk)
Lazy-Ibk classifier	KNN = 12 (vd: 1) DistanceWeighing = Weight by 1-distance (vd: No distance weighting) NearestNeighborSearchAlg = LinearNNSEarch distance function Manhattan yapıldı (vd: Euclidean) Diğer tüm parametreler aynı (vd).	0.887	0.809	Hız sorunu gözlenmemiştir. (1.5 ile 3 dk)
Meta-RandomSubspace Alt sınıflayıcı algoritma: FT (Functional Tree)	Number of iterations:90 (vd: 10) Diğer tüm parametreler aynı (vd). FT parametreleri içinde; Binary Split: True (vd: false) Diğer tüm FT parametreleri aynı (vd).	0.882	0.794	Öğrenme kısmı 6 dakika, test kısmı 71 dakika sürdü. Yavaş çalıştı.
SMO (Sequential Minimized Optimization)	kernel: RBFKernel (vd: PolyKernel) FilterType: Normalize Training Data Gamma = 0.15 (vd 0.01) Diğer tüm parametreler aynı (vd).	0.869	0.777	Hız sorunu gözlenmemiştir. (38 sn ile 2 dk)
Meta-RandomSubspace Alt sınıflayıcı algoritma: FT (Functional Tree)	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd).	0.854	0.796	Öğrenme 6 dakika, test kısmı 68 dakika sürdü. Yavaş çalıştı.
RandomForest	NumofTrees = 92 MaxDepth = 3. Diğer tüm parametreler aynı (vd).	0.834	0.779	Hız sorunu gözlenmemiştir. (20 sn ile 82 sn.)
SMO (Sequential Minimized Optimization)	kernel: Puk (vd: PolyKernel) FilterType: Normalize Training Data Omega: 0.9 (vd 1.0) Sigma: 6.4 (vd 1.0) Diğer tüm parametreler aynı (vd).	0.818	0.781	Hız sorunu gözlenmemiştir. (38 sn ile 2 dk)

Çizelge 5.2. Özdevimli öğrenme sınıflandırıcılarına göre üst sınıfa giren riskler

Bilgi Varlığının Adı:	**Tehdit**	**Zayıflık**	**İlgili Risk (anketteki soru biçimi ile)**
Personel	Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum dışından birilerinin hastanedeki gizli bilgileri ele geçirmesi
Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler	Bilgi, belge, veri kaybı	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Dikkatsizlik, telaş, vb nedenlerle; belgelerin kaybolması veya çalınması
Personel	Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Personelin iyi niyet, saflık, dikkatsizlik, vb zaaflarından yararlanarak; kurum içinden birilerinin hastanedeki gizli bilgileri dışarı sızdırması
Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Karmaşık ve özel dil, terminoloji kullanımının getirdiği zorluklar ve karışıklıklar	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin yazılı belgelere yanlış işlenmesi ve yanlış şekilde onaylanması
Kurumun personel istihdam süreçleri ve ilgili kontroller	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	İstihdam öncesi ilgili süreçlerde ve kontrollerdeki eksiklikler	Personelin istihdam sırasında yetkinlik kontrollerinin yetersiz kalması nedeniyle; kurumdaki bilgi sistemlerinde arıza, hata, aksaklığa yol açacak yetersiz / deneyimsiz kişilerin görev alması
Bilgisayarlar	bilgisayar virüsleri, vb zararlı kodlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların sistemlere bulaşması ve zarar vermesi
Elektronik ortamdaki hasta kayıt verileri	bilgisayar virüsleri, vb zararlı kodlar	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; bilgisayar virüsleri, vb zararlı kodların verilere zarar vermesi
Elektronik ortamdaki hasta kayıt verileri	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Elektrik kesintisi, yangın, teknolojik altyapıdaki veya sistemlerdeki çeşitli arıza veya hasarlar nedeniyle; verilerin zarar görmesi
Elektronik ortamdaki hasta kayıt verileri	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Karmaşık ve özel dil, terminoloji kullanımının getirdiği zorluklar ve karışıklıklar	Hekim ile bilgi işlem personeli veya hemşire ile bilgi işlem personeli arasındaki sözlü iletişimde kullanılan karmaşık ve özel tıp dili ve terminolojileri nedeniyle; bilgilerin elektronik ortama yanlış aktarılması

Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler	Eksik, kayıp Malzeme	Personelin iyi niyet, saflık, dikkatsizlik, telaş, vb yapısı	Dikkatsizlik, telaş, vb nedenlerle; kaşe ve mühürlerin kaybolması veya çalınması
Elektronik ortamdaki hasta kayıt verileri	Yanlış, tutarsız veya bozulmuş hasarlı veri veya işlem	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği, verimsizlik vb durumlar	Personelin bir kısmında iş ortamına bağlı çeşitli nedenlerle huzursuzluk, motivasyon eksikliği sonucunda; verilere zarar verilmesi, verilerde uygun olmayan işlem yapılması
Hasta Kabul Modülü	Deprem, sel, vb doğal afetler	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Deprem, sel, vb doğal afetler nedeniyle hasta kabul modülünün çalışmaması
Kurumun personel istihdam süreçleri ve ilgili kontroller	Kurum içinden kötü niyetli birilerinin yetkisiz erişim eylemleri	İstihdam öncesi ilgili süreçlerde ve kontrollerdeki eksiklikler	Personelin istihdam sırasında yetkinlik ve güvenilirlik kontrollerinin yetersiz kalması nedeniyle; kurum içinde bilgi güvenliğine kasıtlı zarar verebilecek potansiyelde kişilerin görev alması
Elektronik ortamdaki hasta kayıt verileri	Deprem, sel, vb doğal afetler	Elektrik, işletim sistemleri, vb altyapının dayanıksızlığı	Deprem, sel, vb doğal afetler nedeniyle elektronik hasat kayıt verilerinin zarar görmesi
Bilgisayarlar	Kurum dışından kötü niyetli birilerinin yetkisiz erişim eylemleri	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği	Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının bilgisayarlara teknolojik saldırıları
Belgelerin kaşe-imza-mühür işlemlerinin tamamlanması ve ilgili süreçler	Bilgi, belge, veri kaybı	Güncel ve tutarlı , güvenilir envanter olmaması	Belgelere ait güncel ve eksiksiz bir envanterin bulunmaması nedeniyle; aksaklıklar yaşanması veya belgelerin kaybolması

Doğruluk başarım değerleri öğrenme ve test sürecinin bileşkesinde en başarılı olarak gözlemlenen sınıflandırıcı algoritmalar için ayrıca öğrenme eğrilerine yönelik ek gözlemler yapılmıştır. Anketteki sonuçlardan elde edilen toplam 1750 adet olan tüm örneklem kümesinden rastsal olarak seçilen sırasıyla 87, 175, 350, 525, 700, 875, 1050, 1225, 1400, 1575, 1662, 1715 adetlik alt örneklem kümeleri oluşturulmuştur. Rastsal alt veri kümelerinin oluşturulmasında, Weka yazılımındaki ilgili işlev kullanılmıştır. 12 farklı alt örneklem kümesi ve tüm veri kümesinin (1750 adet veri) de kapsanacağı şekilde, en başarılı sonuçları vermiş olan on ayrı sınıflandırıcının her birisi için öğrenme

(train) doğruluk (accuracy) başarım değerleri ölçülmüştür ve bu sonuçlar Çizelge 5.3.' de özetlenmektedir.

Çizelge 5.3. Sınıflandırıcıların doğruluk tabanlı öğrenme başarım değerleri

Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: RepTree	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). RepTree parametreleri içinde; No Pruning: True (vd : false) Diğer tüm RepTree parametreleri aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,977
175	0,983
350	0,974
525	0,975
700	0,977
875	0,971
1050	0,97
1225	0,973
1400	0,972
1575	0,964
1662	0,955
1715	0,962
1750	0,957
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: J48	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). J48 confidence factor: 0.25 (No pruning olduğundan bu parametre etkisizdir) Unpruned: True (vd: False) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,966
175	0,971
350	0,971
525	0,962

700	0,961
875	0,95
1050	0,946
1225	0,952
1400	0,951
1575	0,948
1662	0,943
1715	0,94
1750	0,938
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: J48	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd). J48 confidence factor: 0.75 (vd: 0.25) Binary splits: True (vd: False) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,966
175	0,971
350	0,957
525	0,949
700	0,95
875	0,936
1050	0,932
1225	0,936
1400	0,934
1575	0,929
1662	0,924
1715	0,928
1750	0,929
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: FT (Functional Tree)	Number of iterations:90 (vd: 10) Diğer tüm parametreler aynı (vd). FT parametreleri içinde; Binary Split: True (vd: false) Diğer tüm FT parametreleri aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,977

175	0,943
350	0,94
525	0,935
700	0,916
875	0,894
1050	0,883
1225	0,894
1400	0,883
1575	0,883
1662	0,881
1715	0,882
1750	0,882
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Meta-RandomSubspace Alt sınıflayıcı algoritma: FT (Functional Tree)	Number of iterations:180 (vd: 10) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,977
175	0,949
350	0,94
525	0,92
700	0,894
875	0,882
1050	0,882
1225	0,865
1400	0,862
1575	0,859
1662	0,86
1715	0,857
1750	0,854
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
SMO (Sequential Minimized Optimization)	kernel: Normalized PolyKernel (vd: PolyKernel) Exponent: 3.5 (vd: 2.0). FilterType: Normalize Traning Data. Diğer tüm parametreler aynı (vd).

Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	1
175	0,977
350	0,96
525	0,95
700	0,933
875	0,917
1050	0,905
1225	0,913
1400	0,913
1575	0,901
1662	0,897
1715	0,897
1750	0,902
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
SMO (Sequential Minimized Optimization)	kernel: RBFKernel (vd: PolyKernel) FilterType: Normalize Training Data Gamma = 0.15 (vd 0.01) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,977
175	0,954
350	0,929
525	0,918
700	0,889
875	0,88
1050	0,876
1225	0,879
1400	0,873
1575	0,867
1662	0,868
1715	0,868
1750	0,869
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
SMO (Sequential Minimized	kernel: Puk (vd: PolyKernel) FilterType: Normalize Training

Optimization)	Data Omega: 0.9 (vd 1.0) Sigma: 6.4 (vd 1.0) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	0,92
175	0,92
350	0,866
525	0,865
700	0,847
875	0,834
1050	0,834
1225	0,835
1400	0,83
1575	0,82
1662	0,82
1715	0,819
1750	0,818
Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
RandomForest	NumofTrees = 92 MaxDepth = 3. Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
87	1
175	0,994
350	0,949
525	0,922
700	0,887
875	0,875
1050	0,858
1225	0,86
1400	0,853
1575	0,837
1662	0,835
1715	0,832
1750	0,834

Sınıflandırıcı Algoritmanın Adı ve Türü:	**Parametreler / alt işlevler:**
Lazy-Ibk classifier	KNN = 12 (vd: 1) DistanceWeighing = Weight by 1-distance (vd: No distance weighting) NearestNeighborSearchAlg = LinearNNSEarch distance function Manhattan yapıldı (vd: Euclidean) Diğer tüm parametreler aynı (vd).
Örneklem Adedi:	**Doğruluk (Accuracy) Başarım Değeri (% 100 Doğrulukta Başarım = 1)**
9	1
32	0,909
142	0,82
280	0,944
405	0,932
558	0,91
713	0,915
894	0,902
1054	0,892
1215	0,892
1305	0,885
1358	0,888
1401	0,887

Çizelge 5.3.'de de görüleceği gibi, Lazy.Ibk adlı sınıflandırıcının örneklem adetleri diğerlerinden farklıdır. Bunun nedeni, bu sınıflandırıcının ilgili alt işlev ve parametrelerinin yapısıyla ilgili olup, en yakınındaki komşuyu arama (nearest neighbor search) algoritmasında kullanılan uzaklık hesaplama işlevinin Öklid (Euclidean) yerine Manhattan işlevi şeklinde tanımlanmış olmasından kaynaklanmaktadır. Manhattan uzaklık hesaplama işlevi, örneklem sayısına göre kendisi daha az sayıda bir alt küme oluşturmaktadır (Williams ve Li, 2010). Bu nedenle, Lazy.Ibk sınıflandırıcısının öğrenme başarım değerlerini ölçmede kullanılan alt küme boyutları sırasıyla 9, 32, 142, 280, 405, 558, 713, 894, 1054, 1215, 1305, 1358, 1401 değerlerini almıştır. Yapılan deneylerde, Lazy.Ibk sınıflandırıcısı için en yüksek öğrenme ve test başarım değerlerinin Manhattan uzaklık hesaplama işlevi kullanıldığı durumlarda oluştuğu bulgulanmıştır. Çalışmadaki risk değerlendirme parametrelerinin çok sayıda olması ve geliştirilen modelin doğrusal olmayan, karmaşıklığı yüksek bir yapı olması nedeniyle bu

bulgu, literatürdeki ilgili çalışmalarla tutarlı bir şekilde örtüşmektedir (Aggarwal, vd., 2001). Manhattan ve Öklid uzaklık hesaplama işlevleri ve ikili sınıflandırıcılarda kullanımı aşağıdaki tanımlamada ve denklemlerde özetlenmektedir (Wang, vd., 2006).

İkili sınıflandırıcı ile sınıflara ayrılacak verilerin, *d*-boyutlu bir Öklid uzayındaki vektörler şeklinde temsil edildiği varsayılsın.

Bu ikili sınıflandırıcı için öğrenme örnekleri $\left\{\left(\vec{X}_1, Y_1\right),.....,\left(\vec{X}_n, Y_n\right)\right\}$ gibi kümelerle tanımlanır ve sınıflandırıcı için sonuç deseni veya değeri $\vec{X}$ şeklinde betimlenir ise;

k-en yakın komşu (k-NN) kuralı $\vec{X}$ 'in k en yakın komşu değerlerini $\vec{X}_{(1)},....., \vec{X}_{(k)}$ bulur ve $\vec{X}$ değerini, $Y_{(1)},....., Y_{(k)}$ sınıfları içerisinden en fazla olanına atar. Her bir $Y_{(i)}$, karşılık gelen $\vec{X}_{(i)}$ ' nin sınıf etiketini temsil eder.

Bu tanımlara göre, Öklid uzaklığı (Euclidean distance) şu şekilde bulunur:

$$d\left(\vec{X}, \vec{X}_i\right) = \sqrt{\left(\sum_{j=1}^{d} \left|X^j - X_i^j\right|^2\right)} \tag{5.1}$$

Benzer şekilde, Manhattan uzaklığı (Manhattan distance) aşağıdaki gibi tanımlanır:

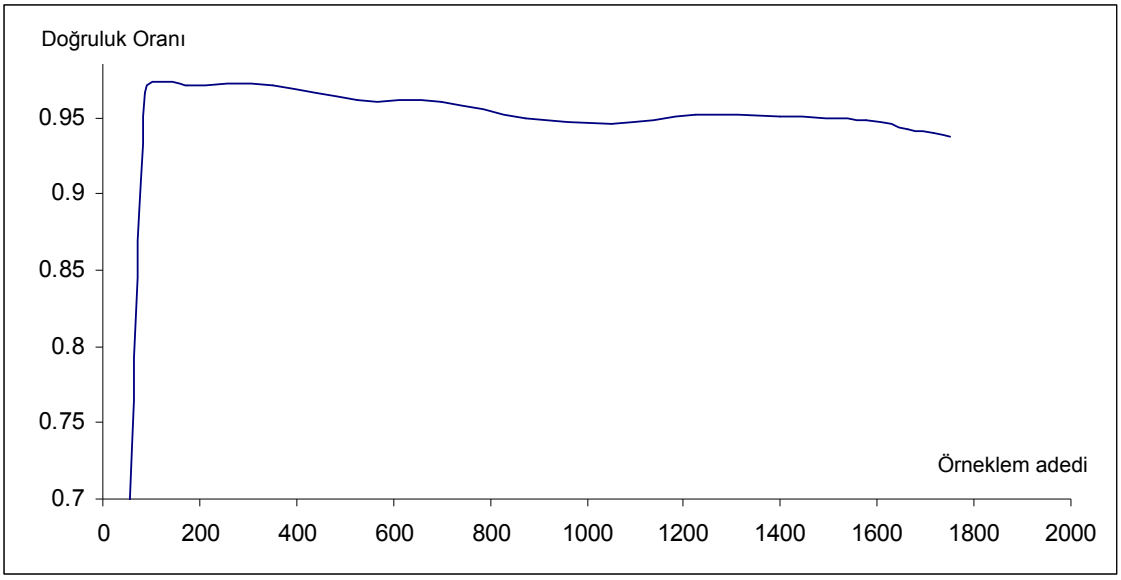
$$d\left(\vec{X}, \vec{X}_i\right) = \sum_{j=1}^{d} \left|X^j - X_i^j\right| \tag{5.2}$$

Denklem (5.1) ve (5.2) sonucuna göre, ikili sınıflandırıcı algoritmalar için (*Y* -1 veya 1 değeri alacak şekilde), k-en yakın komşu kuralı aşağıdaki gibi uygulanacaktır:

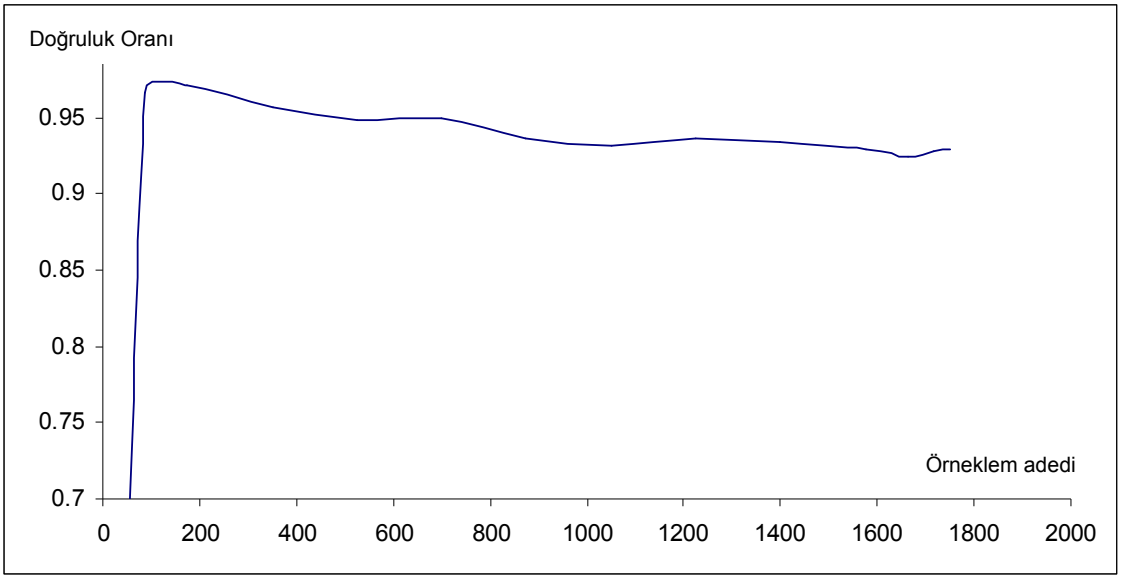
$$f\left(\vec{X}\right) = \text{sgn}\left(\sum_{i=1}^{k} Y_{(i)}\right) \tag{5.3}$$

İlgili her bir sınıflandırıcının öğrenmedeki doğruluk başarım değerleri kullanılarak oluşturulan öğrenme eğrileri de Şekil 5.1. - 5.10.' daki şekillerde gösterilmektedir. Elde

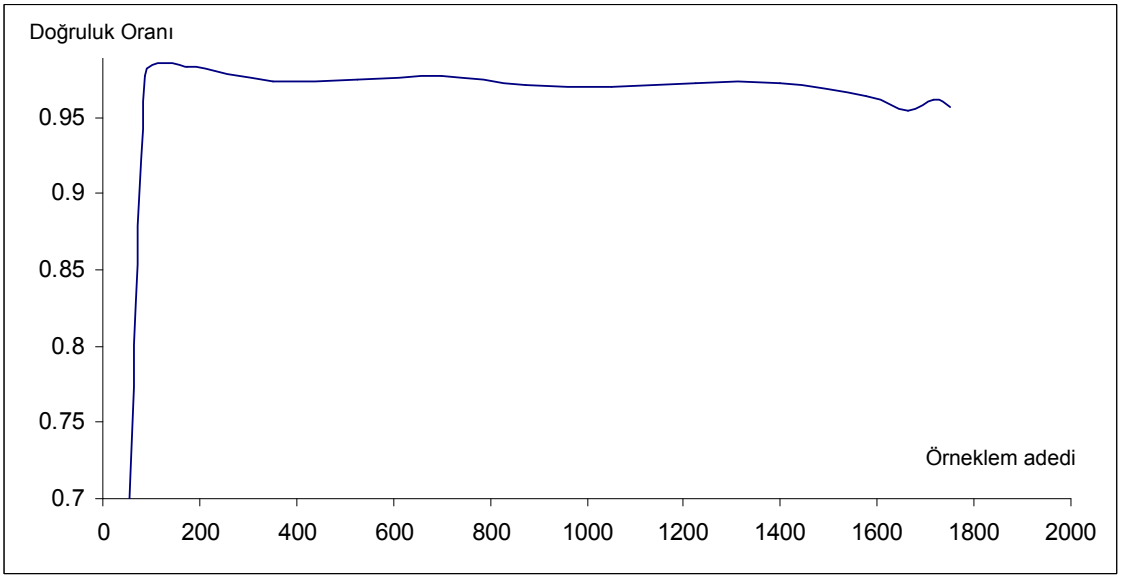
edilen sonuçlara göre öğrenme eğrileri; *y* ekseninde (düşey eksende) doğruluk değerleri, *x* ekseninde (yatay eksende) örneklem adedi olacak şekilde oluşturulmuştur. Öğrenme eğrilerine ilişkin şekiller incelendiğinde, on sınıflandırıcının tamamının öğrenme doygunluğuna (happy graph) eriştiği bulgulanmaktadır. Öğrenme eğrilerinin irdelenmesinde ve öğrenme doygunluğuna erişip erişmediğini belirlemede; ilgili kaynaklarda tanımlanmış olduğu şekilde, örneklem adedi arttıkça, modelin doğruluk veya hassasiyet başarım değerlerinin ne şekilde değiştiği ve artıp artmadığı incelenmektedir (Russell, 2009). Eğer öğrenme eğrisi belli bir sayıda örnek adedinden sonra yatayda düzleşen veya aşağı doğru eğim gösteren bir şekilde ise, bir başka deyişle, daha fazla veri kullanılsa da öğrenmedeki doğruluk düzeyinin daha fazla artamayacağı görülüyor ise, ulaşabilecek en yüksek öğrenme başarım değerlerine ulaştığı anlaşılmakta ve öğrenme doygunluğuna erişildiği yorumu yapılmaktadır (Weiss, 2003).



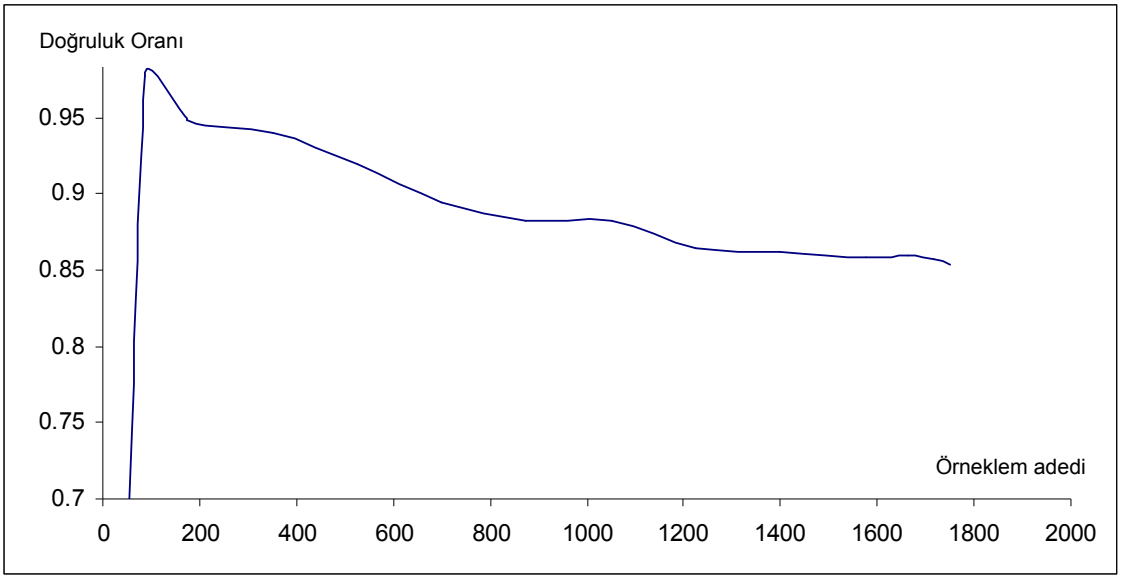
Şekil 5.1. Randomsubspace-J48 (unpruned) öğrenme eğrisi



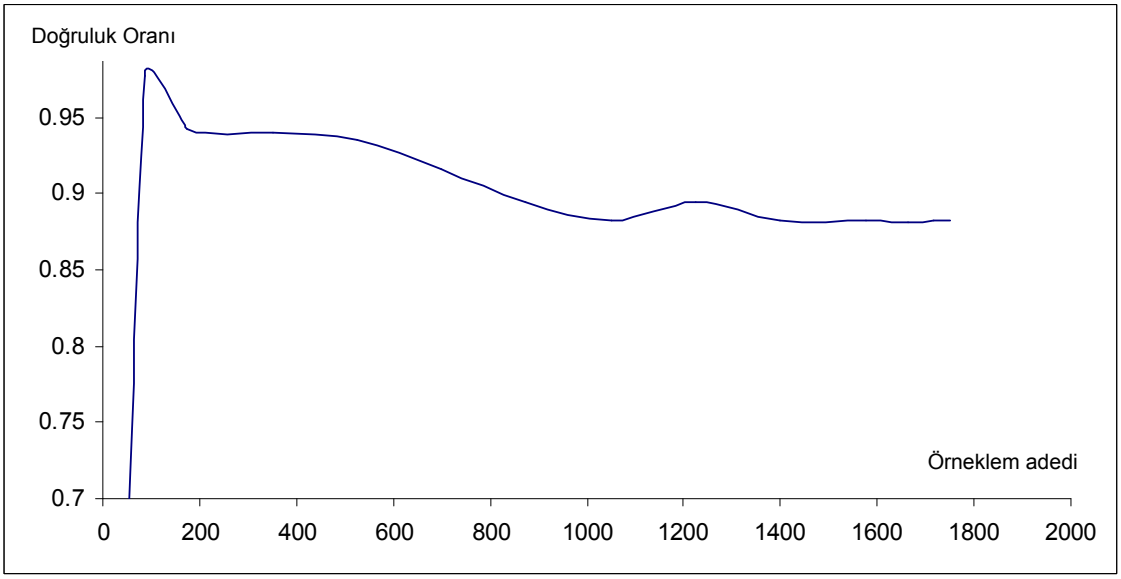
Şekil 5.2. Randomsubspace-J48 (binarysplits) öğrenme eğrisi



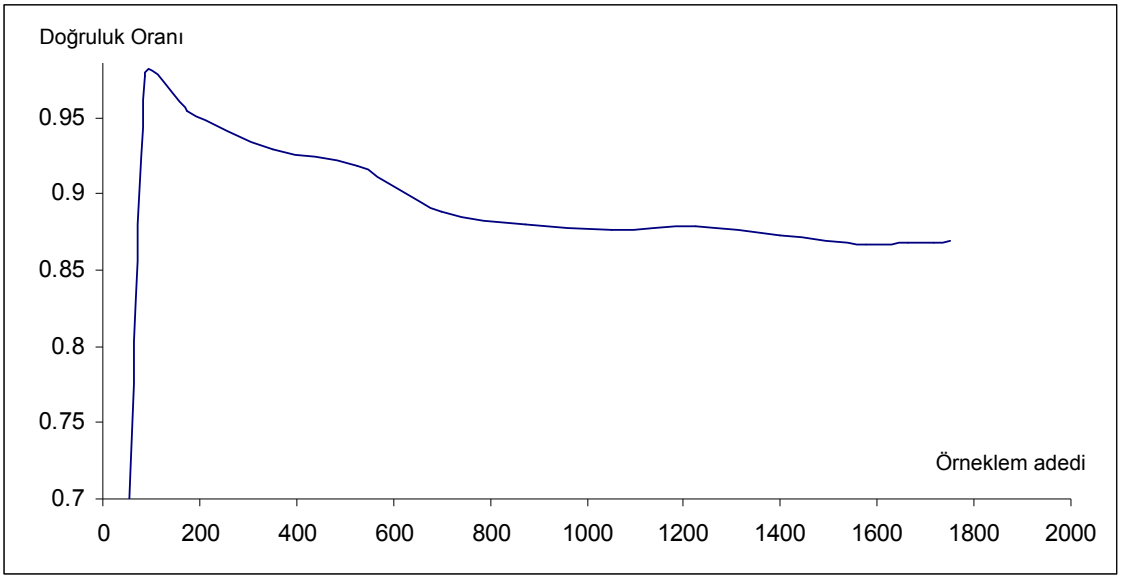
Şekil 5.3. Randomsubspace-REPTree öğrenme eğrisi



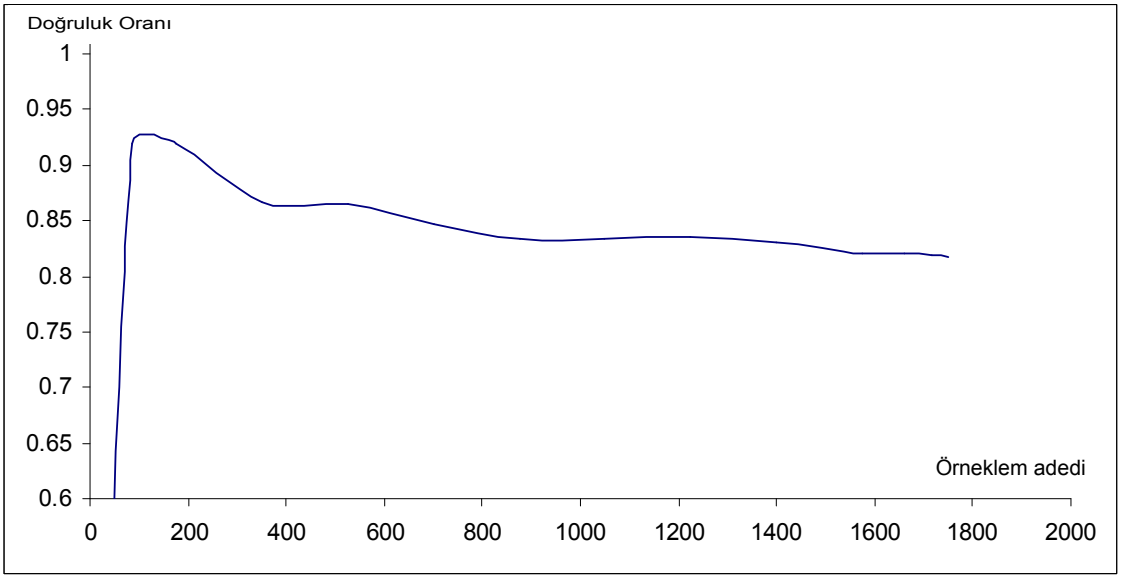
Şekil 5.4. Randomsubspace-FunctionalTree öğrenme eğrisi



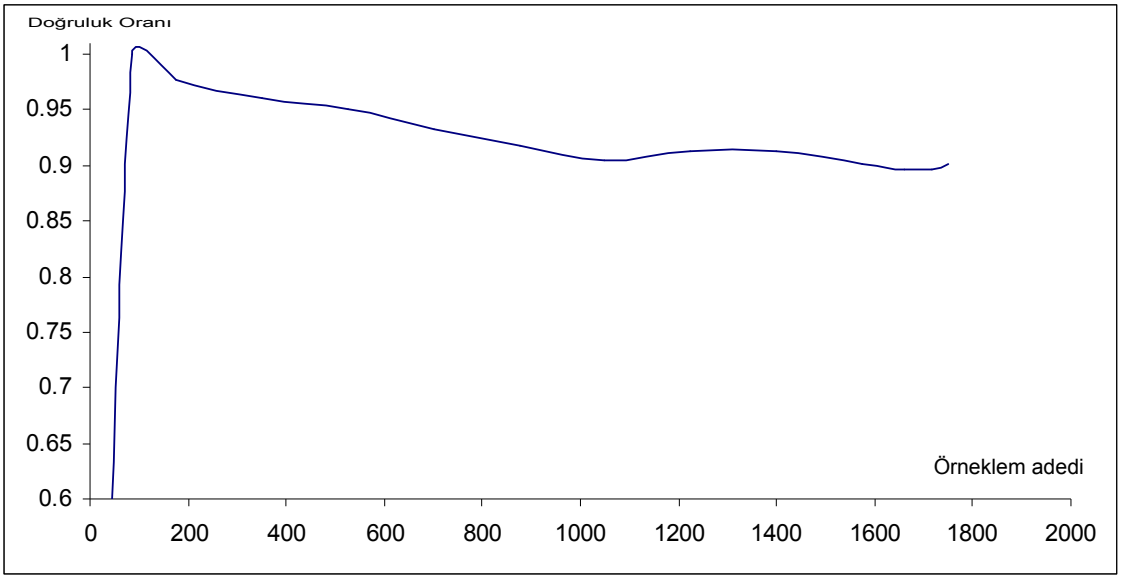
Şekil 5.5. Randomsubspace-FunctionalTree (binarysplits) öğrenme eğrisi



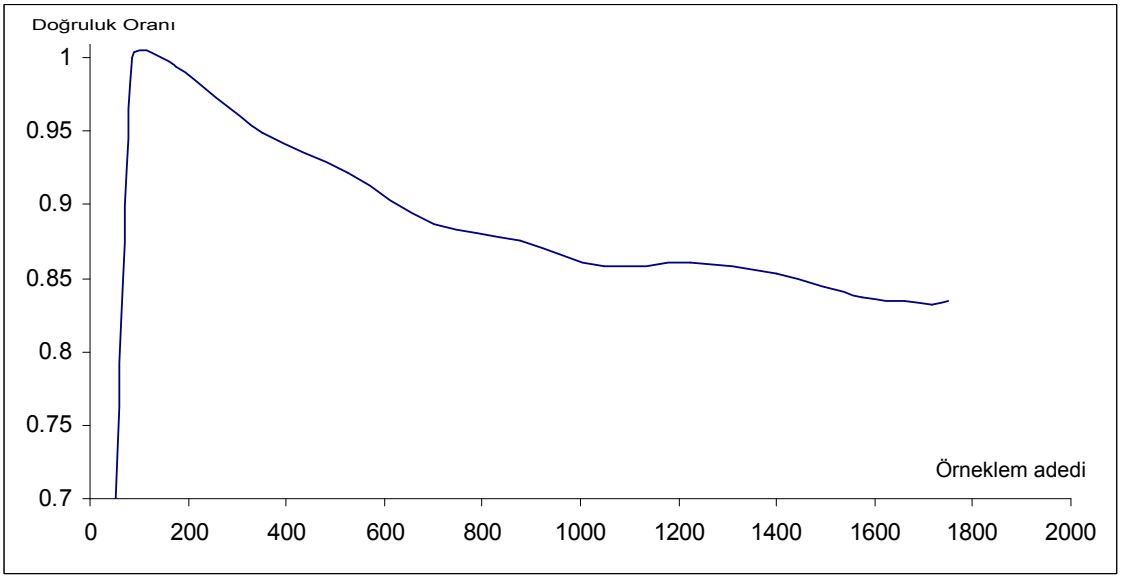
Şekil 5.6. Sequential Minimized Optimization (RBFKernel) öğrenme eğrisi



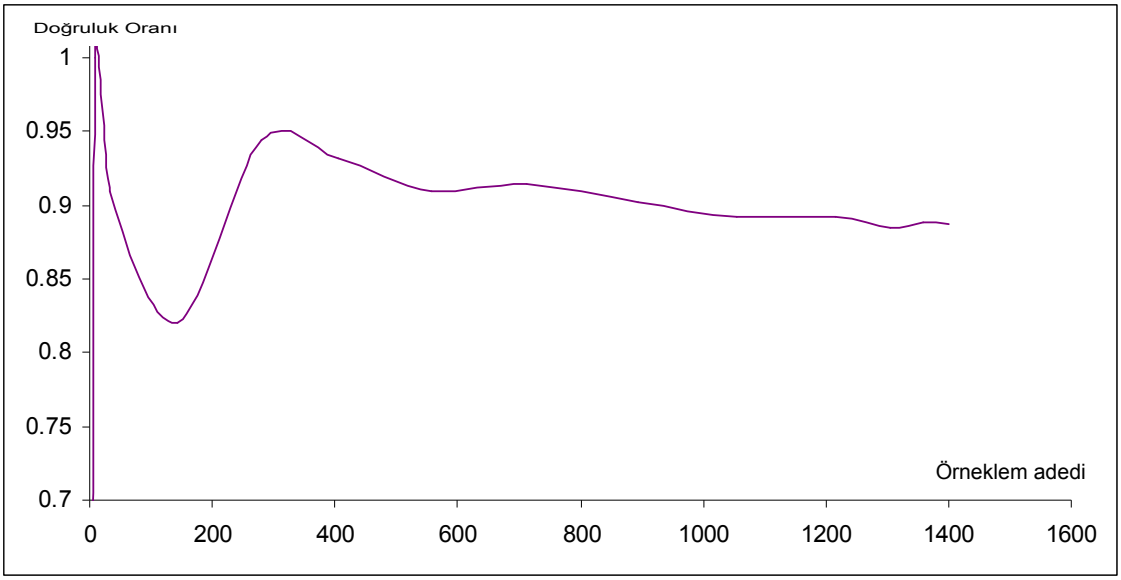
Şekil 5.7. Sequential Minimized Optimization (PUK kernel) öğrenme eğrisi



Şekil 5.8. Sequential Minimized Optimization (NormalizedPolyKernel) öğrenme eğrisi



Şekil 5.9. RandomForest öğrenme eğrisi



Şekil 5.10. Lazy.Ibk (KNN=12, Manhattan distance function) öğrenme eğrisi

Şekil 5.1. - 5.9.' daki öğrenme eğrileri incelendiğinde, belli sayıda veri kümesi ile modelde öğrenme uygulatıldıktan sonra, doğruluk değerlerinin olabileceği en yüksek başarım düzeyine geldiği ve bu noktadan sonra örneklem sayısının artmasının doğruluk başarımını etkilemediği, hatta düşürdüğü gözlemlenmiştir. Bu da, özdevimli öğrenme kuramı ve ilgili kaynaklara göre olması gereken normal bir durumdur; modelin öğrenme doygunluğuna ulaştığını ve deneylerde yeterli sayıda ve çeşitlilikte örnek veri kümesi kullanıldığını ortaya koymaktadır (Russell, 2009). Bir başka deyişle, ilgili algoritma ve bu algoritma kullanılarak oluşturulan öğrenme süreci, ulaşabileceği en yüksek öğrenme başarım değerlerine ulaştığını ve daha fazla veri kullanılsa da öğrenmedeki doğruluk düzeyinin daha fazla artamayacağını göstermektedir. Çalışmada ve deneyler sonucunda elde edilen bulgularda, sadece Şekil 5.10.'daki Lazy.Ibk adlı özdevimli öğrenme algoritmasının öğrenme eğrisinin farklılık gösterdiği görülmektedir. 87 adetlik ilk örneklem kümesinde doğruluk ve öğrenme başarım düzeyinin aşırı hızlı biçimde yükseldiği ve maksimuma ulaştığı (doğruluk oranı 0.909), ama veri örneklem kümesinin boyutu biraz daha artınca ciddi bir başarım kaybı olduğu (doğruluk oranı 0.82), örnek sayısı arttıkça öğrenmedeki doğruluk başarım değerinin tekrar arttığı ve örnek sayısı 1715'e geldiğinde öğrenme başarım değerinin doygunluğa ulaştığı görülmektedir. Bu sonuç da aslında kavramsal olarak tutarlı bir bulgudur, çünkü

Lazy.Ibk tasarımı ve yapısı gereği, çok hızlı ve az sayıda örnekle öğrenen ama çoğunlukla bu öğrenme süreci kopyalama ve ezberlemeye yönelik olabilen bir algoritmadır (Witten, vd., 2011).

Tez çalışması sonucunda, alanında ilk örnek olabilecek özgün bir nitel bilgi güvenliği risk değerlendirme modeli başarıyla ortaya konmuştur. Çalışmada elde edilen sayısal ve istatistiksel veriler ve bulgular ışığında varılan sonuçlar ve tartışma konuları aşağıda maddeler halinde verilmektedir:

- Çalışmada geliştirilen nitel bilgi güvenliği risk anketinin tasarımına ait bazı özelliklerinin alanındaki özgün örneklerden biri olduğu bulgulanmıştır. Ayrıca, anketin uygulamasının da Türkiye'de sağlık sektöründe bir ilk örnek olduğu belirlenmiştir.
- Elde edilen sonuçlara göre, hastanede hangi risklerin daha yüksek düzeyde / daha öncelikli oldukları belirlenmiş ve listelenmiştir.
- Tezdeki amaca uygun ve tamamen özgün bir model başarıyla oluşturulmuştur. Çalışmadaki nitel risk değerlendirmesi için özdevimli öğrenim yaklaşımı yöntemi, geliştirilen model ve elde edilen sonuçların, alanındaki özgün ilk örneklerden birisi olduğu belirlenmiştir.
- Anket sonucunda edinilen bilgi güvenliği risk değerleri, verileri ve parametreleri tezdeki özdevimli öğrenim modeline uygun hale getirilmiş, seçilen algoritmalarla farklı parametreler kullanılarak gözlemler ve rastsal örneklem kümeleri üzerinden testler yapılmıştır. Anketten elde edilen tüm veri kümesi (1920 örnek) ve ayıklanmış veri kümesi (1750 örnek) üzerinden, 68 değişik ikili sınıflandırıcı özdevimli öğrenim algoritması ve bunların alt seçenekleri, değişik parametre kombinasyonlarıyla beraber 3200 civarında değişik gözlem ve test yapılmıştır. Yapılan deneyler, karşılaştırmalı analizler ve gözlemlerin adet ve çeşitlilik açısından alanında özgün bir çalışma olduğu bulgulanmıştır.
- Çalışmadaki anket risklerini değerlendirme ve risk kestirimi için öğrenme (train) ve sınama (test) başarım (performance) değerleri birlikte göz önüne alındığında toplam 10 adet ikili sınıflandırıcı özdevimli öğrenim algoritmasının kabul edilebilir düzeyde güvenilir / başarılı sonuç verdiği bulgulanmıştır. Başarı kıstası

ve ölçütü olarak, öğrenme ve sınama gözlemlerindeki TP ve TN oranları ve bunlara ilişkin doğruluk (accuracy) değerleri göz önüne alınmıştır. Bunun iki temel nedeni vardır. Birincisi, modelde hem eşik değer üstü (Risk = U) hem de eşik değer altı (Risk = A) kestirim (prediction) başarım değerlerinin dikkate alınmış olmasıdır. Bir başka deyişle, kestirimde doğruluk kıstası riskin üst ("U") olduğu durumlar için ayrı, riskin alt ("A") olduğu durumlar için ayrı olarak ölçülmüş, bu ikisinin ağırlıklı ortalaması (weighted average) alınmış ve bu ağırlıklı ortalamadaki doğruluk başarım değerlerine göre sınıflandırıcıların başarısı irdelenmiştir. İkinci gerekçe ise, sınıflandırıcıların sınama başarım değerleri için onlu çapraz doğrulama (10-fold cross-validation) yönteminin kullanılmış olmasıdır ve bu sınama yöntemi, sınıflandırıcılarda doğruluğun öncelikli başarım ölçütü olarak değerlendirildiği durumlarda kullanılmaktadır (Hastie, vd., 2009, Refaeilzadeh, vd., 2008). Bölüm 4.4'te açıklanan modelin bir ilk model (prototype) olduğu göz önüne alınırsa, ileride benzer amaçlar için güncellenip geliştirilecek modellerde, F-ölçümü, Kappa istatistiği, ROC eğrileri de başarım kıstası olarak modele eklenebilir. Burada vurgulanması gereken bir diğer konu, çalışmada geliştirilen modelde ve ilgili özdevimli öğrenme sonuçlarında doğruluk (accuracy) değerlerinin başarım değerleri göz önüne alındıktan sonra, risk eşik değeri üst olan ("U") riskler için risk değerlendirmesinin sonuçlandırılmış olmasıdır. Eğer, ileride geliştirilecek bilgi güvenliği risk değerlendirme modellerinde, kurumun risk kabul eşik değeri üstünde kalanlar yanı sıra, altında kalanlar da ayrıca irdelenmesi istenecek ise, hem üst hem de alt değerler için çalışma ve başarım ölçümü yapılması gerekecektir. Bu durumda, özdevimli öğrenme sınıflandırıcılarının hassasiyet ve geri çağırım değerleri göz önüne alınmalı ve bu değerler üst ve alt sınıflar için ayrıştırılarak gözlemlenmeli ve ölçülmelidir.

- Benzer şekilde, bölüm 4.4.'te tanımlanan modelin 9. aşamasında kullanılan basit kartil yöntemi yerine, başka istatistiksel yöntemlerle de, hangi risklerin üst sınıfında kabul edileceği belirlenebilir.

- En başarılı bulunan on değişik ikili sınıflandırıcının doğruluk değerleri yanı sıra ROC alanı, F-ölçümü, Kappa istatistiği değerleri ve diğer tüm gözlem değerleri de çalışmanın Ekler Ek-A bölümünde yer almaktadır. Örnek olması açısından,

göreceli olarak başarım değerleri daha düşük olan çeşitli sınıflandırıcılara ait detaylı başarım değerleri de Ekler Ek-A'da verilmiştir.

- Çalışmada en başarılı sonuçların gözlemlendiği on adet sınıflandırıcı için öğrenme eğrilerinin irdelenmesinde ve sınıflandırıcıların öğrenme doygunluğuna erişip erişmediğini belirlemede; doğruluk başarım değerleri tabanlı ölçümler yapılmıştır. Benzer çalışma ve gözlemlerin, hassasiyet başarım değerleri için de ayrıca yapılması ve sınıflandırıcıların öğrenme sürecinde doğruluk başarım değerleri ile hassasiyet başarım değerlerinin karşılaştırılması da düşünülebilir.
- Özdevimli öğrenme amaçlı çalışan bazı ikili sınıflandırıcı algoritmaların, güvenlik risklerini nitel değerlendirmede, belli güvenilirlik ölçütlerinde, risk düzeyini öngörme / belirlemede kullanılabileceği ama bazı iyileşme / gelişme noktalarının da gerektiği kanıtlanmıştır. Bu algoritmaların, genelde, belirli grup / yapıda (ensemble / meta, SMO, RandomForest) olduğu ve bazı ortak alt bileşenlerde ve istatistiksel / matematiksel modellerde (rastsal dönüşümler, karar ağaçları) daha iyi sonuç verdiği saptanmıştır. Modelde kullanılan ankete ait veri kümesi ve parametrelerin fazla sayıda olması, doğrusal olmayan (non-linear) bir yapıda olması ve karmaşıklığının (complexity) oldukça yüksek olması çalışmada önemli bir kıstastır. Modelin, özellikle karar ağaçları tabanlı sınıflandırıcılarda daha başarılı olması bu açıdan da kıstasları ve gereklilikleri destekleyen tutarlı ve anlamlı bir bulgu ve sonuç olmuştur.
- İkili sınıflandırıcı kullanılması nedeniyle, bazı risklerin algoritma analizi öncesi sınıf tanımlamasının hatalı yapıldığı ve algoritmaların öğrenme ve test başarımlarını % 8 - % 14 düzeyinde düşürdüğü bulgulanmıştır.

Çalışmada geliştirilen model ve yöntemin; amaçlanan başarımı göstermesi yanı sıra, çeşitli uygulama ve alanlarla da katkı verecek yeni çözümler sağlayabileceği de ortaya konmuştur:

- Birçok değişik amaca yönelik uygulanan anketlerdeki tutarsız, gelişigüzel veya yanıltıcı, vb. (biased) yanıtların bulgulanması ve ayıklanması / düzeltilmesi için özdevimli öğrenme algoritmalarının denetim (control) amaçlı kullanılabileceği görülmüştür. Tezdeki anket veri kümesinde, rastgele seçilmiş 13 farklı sınıflandırıcının % 96 - %100 arasında oranlardaki başarıyla hataları yakaladığı

bulgulanmış ve bu yöntem, geliştirilen modele eklenmiştir. Bu şekilde, 1920 adet veri kümesinin 170 tanesi atılmış, 95 tanesi de düzeltilmiştir.

- Risk değerlendirmesi ve analizinin son aşaması olan hangi risklerin öncelikli olduğunu belirlemede; çeşitli özdevimli öğrenim algoritma sonuçlarının benzer sonuçlar için kesişim kümeleri kullanılarak, yönetimsel karar verme süreçlerine önemli katkı sağladığı gösterilmiştir. Ayrıca, anketlere verilen yanıtlarda daha çok sayıda kullanıcı tarafından daha öncelikli görülen bazı risklerin, özdevimli öğrenim algoritmaları ile sağlamasının yapılabildiği ve daha da önemlisi, daha farklı risklerin daha öne çıkabildiği bulgulanmıştır. Hastane anket sonuçlarına göre yapılan risk analizinde 16 "üst" düzey riskin 5 tanesi bu şekilde belirlenmiştir. Modelin 8. ve 9. aşamalarında sınıflandırıcıların tahmin ettiği riskler için kullanılan kartil yöntemi, benzer şekilde ankette kullanıcılardan gelen yanıtlar için uygulanmıştır. Anket yanıtlarına göre kartil yöntemiyle % 50 ve üzeri grubunda kalan risklerden 5 tanesi, özdevimli öğrenme modelinde % 50' lik grubun altında kalmaktadır. Bu riskler; "Kurumunuzda iş sürekliliği, acil durum planlaması vb çözümlerin mevcutta olmaması nedeniyle; personelin deprem, sel, vb doğal afetlerde işe gelememesi", "Elektrik kesintisi, yangın, teknolojik altyapıdaki veya sistemlerdeki çeşitli arıza veya hasarlar nedeniyle; verilerin zarar görmesi", "Elektrik kesintisi, yangın, teknolojik altyapıdaki fiziksel hasar, vb nedeniyle hasta kabul modülünün çalışmaması", "Sistemlerde oluşacak çeşitli teknik hatalar veya arızalar nedeniyle hasta kabul modülünün çalışmaması", "Bilgi güvenliği farkındalığının ve ilgili kurum kurallarına uyumda zayıflığı veya eksikliği nedeniyle; kurum dışından bilişim suçlularının bilgisayarlara teknolojik saldırıları" şeklindedir. Kullanıcı yanıtları ile özdevimli öğrenme sınıflandırıcılarının arasındaki bu farklılık dikkate değer bir durumdur. İleride yapılacak benzer çalışmalar sonucunda bu farklılığın nedeni konusunda kesin bir yargıya varılabilecektir. Öte yandan, elde edilen bu sonuç; kullanıcıların genelde yüksek düzeyde gördükleri risklerin, sınıflandırıcıların bulgulayacağı yüksek düzeyli risklerle karşılaştırılarak kurumda bilgi güvenliği risk yönetimi süreçlerindeki karar verme aşamalarında kontrol amaçlı kullanılabileceğini göstermektedir.

6. ÖNERİLER

Tez çalışması süresince ve çalışma sonunda elde edilen bulgular ışığında, bu çalışma sonrasında yapılabilecek yeni çalışma alanları ve araştırma konuları bir öneri paketi şeklinde belirlenmiştir. Bu öneriler aşağıda maddeler halinde verilmektedir:

- Nitel risk anketlerinde öznellik etkilerini azaltacak (taraflı, önyargılı, vb değerlendirme sorunları) şekilde, daha kapsamlı psikometrik ölçüm yöntemlerinin ve kişilik özelliklerinin belirlenip buna uygun ağırlık değerlerinin verildiği anket modeli geliştirilmesi, kurumlarda bu anketleri kullanarak uygulamalar yapılması ve bunun sonucunda tezde ortaya konan modelin iyileştirilmesi düşünülebilir. Böylece, tezin 5. bölümünde açıklanan modelin birinci aşamasında tanımlanmış olan toplam 8 adet etmene yeni etmen / parametreler katılması ve modelin geliştirilmesi sağlanabilir.
- Tezde ortaya konan modelin, ilgili tüm aşamalarını otomatik ve bütünleşik şekilde çalıştıracak bir yönetimsel yazılım / araç haline getirilmesi sağlanabilir.
- Tez çalışması sırasında yapılan deneylerde bazı ikili sınıflandırıcı algoritmalarda gözlemlenen yavaşlık sorununu azaltacak çalışmalar yapılabilir.
- Tez aşamasında kapsama alınmayan bazı melez / birleşik (hybrid / ensemble) sınıflandırıcıların, anket verileri kullanılarak yeni gözlemler üzerinden başarım testleri ve ölçümlerinin yapılması planlanabilir.
- Sayısal olmayan (nominal) veri kümesinin sayısal hale dönüştürülüp sadece sayısal veri kümelerinde çalışan algoritmalarla da gözlemlenmesi yönünde çalışmalar yapılabilir.
- İkili sınıflandırıcı kullanılması nedeniyle yaşanan kısıtlamaları ve sorunları aşmak amacıyla; risk değerinin ikili yerine çoklu ölçekte (“1-5” arası) bir yapıya dönüştürülüp çoklu-sınıflandırıcı (multi-classifier) algoritmalarla test ve gözlemler yapılması planlanmaktadır.
- Çeşitli nitel anket uygulamalarında (kalite yönetimi, İnsan Kaynakları performans ölçümü, vb çok geniş bir yelpazede) kişilerin anlamsız, tutarsız,

hatalı, vb yanıtlarını özdevimli öğrenme algoritmalarını kullanarak denetleme ve bulgulamaya yönelik yeni çalışmalar yapılması düşünülebilir.

- Tezde ortaya konan modelin son aşaması olan farklı sınıflandırıcı algoritmaların özdeş sonuçları ve kesişim kümelerine göre risklerin derecelendirilmesi ve önceliklendirilmesinde kartil (quartile) ve yalın sıralama (ranking) yöntemleri kullanılmıştır. Bu yöntem yerine, bulanık mantık (fuzzy logic) tabanlı bir yöntem geliştirilmesine ve bulanık mantık işlemleriyle değer atanmasına yönelik yeni bir çalışma yapılabilir.

KAYNAKLAR

1. Abraham, A., Grosan, C., Chen, Y., 2005. "Cyber Security and the Evolution in Intrusion Detection Systems". Journal of Engineering and Technology, 1(1), s. 74-81.
2. Aggarwal, C. C., Hinneburg, A., Keim, D. A., 2001. "On the Surprising Behavior of Distance Metrics in High Dimensional Space". In Proceedings of the ICDT '01 8th International Conference on Database Theory, Springer-Verlag, s. 420-434.
3. Alpaydın, E., 2010. *Introduction to Machine Learning, 2nd edition.* The MIT Press, USA.
4. Anderson, A., Longley, D., Kwok, L., 1994. "Security Modelling for Organisations". In Proceedings of the 2nd ACM Conference on Computer and Communications Security, s. 241-250.
5. Arkar, H., 2004. "Cloninger'in Psikobiyolojik Kişilik Kuramının Türk Örnekleminde Sınanması". Doktora Tezi, Ege Üniversitesi Sosyal Bilimler Enstitüsü, Psikoloji Ana Bilim Dalı, İzmir, Türkiye.
6. Ashenden, D., 2008. "Information Security management: A human challenge?". Elsevier Information Security Technical Report, 13, s. 195-201.
7. Backhouse, J., Dhillon, G., 1996. "Structures of Responsibility and Security of Information Systems". European Journal of Information Systems, 5(1), s. 2–9.
8. Barreno, M. B., 2008. "Evaluating the Security of Machine Learning Algorithms". PhD Dissertation, Electrical Engineering and Computer Sciences, University of California, Berkeley, USA.
9. Baskerville, R., 1993. "Information Systems Security Design Methods: Implications for Information Systems Development". Computing Surveys, 25(4), s. 375-414.

10. BDDK, 2007. *Bankalarda Bilgi Sistemleri Yönetiminde Esas Alınacak İlkelere İlişkin Tebliğ*. Bankacılık Düzenleme ve Denetleme Kurumu, T.C. Resmi Gazete, Sayı 26643, Türkiye.

11. Berry, M. W., Kogan, J., 2010. *Text Mining Applications and Theory*. John Wiley & Sons Ltd., UK.

12. Bouckaert, R. R., 2008. *Bayesian Network Classifiers in Weka for Version 3.5.7.* University of Waikato, New Zealand.

13. British Standards Institute, 2006. *BS 25999-1 Business Continuity Management Part I*, BSI, UK.

14. Business Continuity Institute, 2007. *A Management Guide to Implementing Global Good Practice in Business Continuity Management*. BCI.

15. Cadoğlu, K., 2000. *Risk Yönetimi ve TSK'daki Uygulamalar*. Harp Akademileri Basım Evi, İstanbul, Türkiye.

16. Calder, A., Watkins, S., 2008. *IT Governance A Manager's Guide to Data Security and ISO27001/ISO 27002, 4th edition.* Kogan Page Ltd., UK.

17. Cohen, J., 1960. "A Coefficient of Agreement for Nominal Scales". Educational and Psychological Measurement, 20(1), s. 37–46.

18. Delen, D., Sharda, R., Kumar, P., 2007. "Movie forecast Guru: A Web-based DSS for Hollywood managers". Decision Support Systems, 43(4), s. 1151-1170.

19. Deloitte, 2009. *Losing Ground; 2009 TMT Global Security Survey Full Report.* Deloitte Touche Tohmatsu, Netherlands.

20. Dhillon, G., 2007. *Principles of Information Systems Security.* John Wiley & Sons Inc., USA.

21. Ernst & Young, 2008. *Global Information Security Survey 2008.* EGYM Ltd., USA.

22. Farn, K-J., Lin, S-K., Fung, A. R-W., 2004. "A study on information security management system evaluation—assets, threat and vulnerability". Computer Standards and Interfaces, 26, s. 501-513.

23. Fıkırkoca, M., 2003. *Bütünsel Risk Yönetimi.* Kalder Yayınları, İstanbul, Türkiye.

24. Finne, T., 2008. "Managing Information Security and Business Risks in R&D Collaboration". In ISACA EuroCACS Conference Proceedings vol.1.

25. Fisher, R. P., 1984. *Information Systems Security*. Prentice-Hall, NJ, USA.

26. Giacinto, G., Roli, F., Didaci, L., 2003. "Fusion of multiple classifiers for intrusion detection in computer networks". Pattern Recognition Letters, 24(12), s. 1795-1803.

27. Giudici, P., Figini, S., 2009. *Applied Data Mining for Business and Industry 2nd Edition*. John Wiley & Sons Ltd., UK.

28. Gökçen, H., 2007. *Yönetim Bilgi Sistemleri.* Palme Yayıncılık, İstanbul, Türkiye.

29. Gökşen, Y., Eminağaoğlu, M., Doğan, O., 2011. "Data Mining in Medical Records for The Enhancement of Strategic Decisions: A Case Study". In EBEEC 2011 The 3rd International Conference The Economies of Balkan and Eastern Europe Countries in The Changed World, 5-8 May, Pitesti, Romania.

30. Hamel, L., 2009. *Knowledge Discovery with Support Vector Machines*. John Wiley& Sons Inc., NJ, USA.

31. Hand, C., Mannila, H., Smyth, P., 2001. *Principles of Data Mining*. The MIT Press, London, UK.

32. Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning, Data Mining, Inference and Prediction, 2nd edition*. Springer, New York, USA.

33. IBM, 2011. http://www.watson.ibm.com/index.shtml , IBM Watson Research Center, Internet sitesi, (Erişim: 22 Mayıs 2011).

34. ISO/IEC, 2005. *ISO/IEC 27001:2005*. International Organization for Standardization, Geneva, Switzerland.

35. ISO/IEC, 2005. *ISO/IEC 27002:2005*. British Standards Institute, UK.

36. ISO/IEC, 2008. *ISO/IEC 27005:2008*. International Organization for Standardization, Geneva, Switzerland.

37. ITGI Inst., 2007. *COBIT 4.1; Framework, Control Objectives, Management Guidelines and Maturity Models.* ISACA publishing.

38. ITGI Inst., 2008. *Information Security Governance: Guidance for Information Security Managers,* ISACA publishing.

39. İkiz, F., Püskülcü, H., Eren, Ş., 2006. *İstatistiğe Giriş.* Fakülteler Barış Kitabevi, İzmir, Türkiye.

40. Karabacak, B., Soğukpınar, İ., 2005. "ISRAM: Information Security Risk Analysis Method". Computers & Security, 24(2), s. 147-159.

41. KDH-1, 2011, http://www.karsiyakadh.gov.tr/HASTANEMIZ/tarihce.htm, Karşıyaka Devlet Hastanesi Internet sitesi, (Erişim: 20 Haziran 2011).

42. KDH-2, 2011, http://www.karsiyakadh.gov.tr/INDEX/index_ic.htm, Karşıyaka Devlet Hastanesi Internet sitesi, (Erişim: 20 Haziran 2011).

43. KDH-3, 2011, http://www.karsiyakadh.gov.tr/KALITE_YON/drg_projesi.htm, Karşıyaka Devlet Hastanesi Internet sitesi, (Erişim: 5 Temmuz 2011).

44. Khattak, A. M., Khan, A. M., Lee, S., Lee, Y-K., 2010. "Analyzing Association Rule Mining and Clustering on Sales Day Data with XLMiner and Weka". International Journal of Database Theory and Application, 3(1), s. 13-22.

45. Kılıçaslan, Y., Güner, E. S., Yıldırım, S., 2009. "Learning-based pronoun resolution for Turkish with a comparative evaluation". Computer Speech and Language, 23, s. 311–331.

46. Landis, J. R., Koch, G. G., 1977. "The Measurement of Observer Agreement for Categorical Data". Biometrics, vol. 33, s. 159-174.

47. Landoll, D. J., 2006. *The Security Risk Assessment Handbook - A Complete Guide for Performing Security Risk Assessments.* Auerbach Publications, USA.

48. Layton, T. P., 2007. *Information Security; Design, Implementation, Measurement and Compliance.* Auerbach Publications, USA.

49. Long, X., Yong, Q., Qianmu, L., 2008. "Information security risk assessment based on analytic hierarchy process and fuzzy comprehensive". In Proceedings

of the 2008 International Conference on Risk Management & Engineering Management. Beijing, China, 4-6 November. IEEE Computer Society.

50. Lv, J. J., Qiu, W. H., Wang, Y. Z., Zou, N., 2006. "A study on information security optimization based on MFDSM". In Proceedings of the Fifth International Conference on Machine Learning and Cybernetics. Dalian, China, 13-16 August. IEEE pub.

51. Marchewka, J. T., 2009. *Information Technology Project Management 3rd Edition*. John Wiley & Sons Inc., USA.

52. Mazareanu, V., 2009. "Risk assessment – a human psychology approach". MPRA, Munich, Germany.

53. Meydan Larousse, 1998. *Meydan Larousse Ansiklopedisi*. Nurdan Yayınları, İstanbul, Türkiye.

54. Mitnick, K. D., 2005. *Aldatma Sanatı.* ODTÜ Yayıncılık, Ankara, Türkiye.

55. Parker, D. B., 1981. *Computer Security Management.* Prentice Hall, USA.

56. PCI Security Standards Council, 2006. *Payment Card Industry (PCI) Data Security Standard v1.1*. PCI.

57. Peltier, T. R., 2001. *Information Security Risk Analysis.* Auerbach Publications, USA.

58. Perlich, C., Provost, F., Simonoff, J. S., 2003. "Tree Induction vs. Logistic Regression: A Learning-Curve Analysis". Journal of Machine Learning Research, 4, s. 211-255.

59. Pfleeger, C. P., 2007. *Security in Computing, 4th edition*. Prentice Hall, USA.

60. Pietraszek, T., Tanner, A., 2005. "Data mining and machine learning - towards reducing false positives in intrusion detection". Information Security Technical Report, 10(3), s. 169-183.

61. PMI, 2008. *PMBOK Project Management Body of Knowledge 4th edition*. Project Management Institute.

62. Qu, W., Zhang, D. Z., 2007. "Security metrics, models and application with SVM in information security management". In Proceedings of the Sixth

International Conference on Machine Learning and Cybernetics. Hong Kong, China, 19-22 August. IEEE pub.

63. Refaeilzadeh, P., Tang, L., Liu, H., 2008. "Cross-Validation". Arizona State University, s.1-6.

64. Richardson, R., 2008. *2008 CSI/FBI Computer Crime & Security Survey*. CSI, USA.

65. Rieck, K., 2009. "Machine Learning for Application-Layer Intrusion Detection". PhD Dissertation, Fraunhofer Institute FIRST & Berlin Institute of Technology, Berlin, Germany.

66. Rokach, L., Maimon, O., 2008. *Data Mining with Decision Trees - Theory and Applications.* World Scientific Publishing Co. Pte. Ltd., Singapore.

67. Rost, J., Glass, R. L., 2011. *The Dark Side of Software Engineering, Evil on Computing Projects*. IEEE Computer Society, John Wiley & Sons, Inc., Hoboken, New Jersey, USA.

68. Russell, S. J., Norvig, P., 2009. *Artificial Intelligence: A Modern Approach 3rd Edition.* Prentice Hall, USA.

69. Sammut, C., Webb, G. I., 2011. *Encyclopedia of Machine Learning*. Springer, USA.

70. Sasse, M. A., Brostoff, S., Weirich, D., 2001. "Transforming the 'weakest link' a human/computer interaction approach to usable and effective security". BT Technology Journal, 19(3), s.122-131.

71. Schneier, B., 2000. *Secrets & Lies; Digital Security in a Networked World*. John Wiley & Sons Inc., USA.

72. Schneier, B., 2003. *Beyond Fear - Thinking Sensibly About Security in an Uncertain World*. Copernicus Books, NY, USA.

73. Schneier, B., 2008. *Schneier on Security*. John Wiley & Sons Inc., USA.

74. Stewart, G., 2009. "Maximising the Effectiveness of Information Security Awareness Using Marketing and Psychology Principles". Technical Report, Royal Holloway University of London, UK.

75. Stokes J. W., Platt, J. C., Kravis, J., Shilman, M., 2008. "ALADIN: Active Learning of Anomalies to Detect Intrusion". Technical Report, MSR-TR-2008-24, USA.

76. Sun, L., Srivastava, R. P., Mock, T. J., 2006. "An information systems security risk assessment model under Dempster-Shafer Theory of belief functions". Journal of Management Information Systems, 22(4), s. 109-142.

77. Symantec-1, 2009. *Global Internet Security Threat Report 2008, vol. XIV*. Symantec Corporation, USA.

78. Symantec-2, 2009. *EMEA Internet Security Threat Report 2008, vol. XIV*. Symantec Corporation, USA.

79. Tevfik, A. T., 1997. *Risk Analizine Giriş*. Alfa Basım Yayım Dağıtım, İstanbul, Türkiye.

80. Tipton, H. F., Krause, M., 2007. *Information Security Management Handbook*. Auerbach Publications, USA.

81. Tittel, E., Stewart, J. M., Chapple, M., 2003. *Certified Information Systems Security Professional Study Guide 2nd Edition*. Sybex Inc., USA.

82. TBD 4. Çalışma Grubu, 2006. "E-devlet Uygulamalarında Güvenlik ve Güvenilirlik Yaklaşımları". Türkiye Bilişim Derneği Kamu - Bilişim Platformu VIII, Sonuç Raporu.

83. T.C., 2008. "Elektronik Haberleşme Kanunu, no:5809". Resmi Gazete, Sayı: 27050, Türkiye.

84. T.C., 2008. "Elektronik Haberleşme Güvenliği Yönetmeliği". Resmi Gazete, Sayı: 26942, Türkiye.

85. TCK, 2005. *Bilişim Alanında Suçlar, madde 243 - 246.* Türk Ceza Kanunu, Sayı: 5237, Bölüm 10, Türkiye.

86. TDK, 2005. *Büyük Türkçe Sözlük.* Türk Dil Kurumu, Ankara, Türkiye.

87. TDK, 2011. *Bilim ve Sanat Terimleri Ana Sözlüğü.* http://tdkterim.gov.tr/?kelime=%F6zdevimli+%F6%F0renme&kategori=terim&hng=md, Türk Dil Kurumu Internet sitesi, (Erişim: 8 Temmuz 2011).

88. The New York Times, 2011. http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html?pagewanted=all , The New York Times Company, Internet sitesi, (Erişim: 19 Şubat 2011).

89. TSE, 2006. *TS ISO/IEC 27001.* Türk Standartları Enstitüsü, Ankara, Türkiye.

90. Vercellis, C., 2009. *Business Intelligence; Data Mining and Optimization for Decision Making.* John Wiley & Sons, Ltd., UK.

91. Veiga, A. D., Eloff, J. H. P., 2010. "A framework and assessment instrument for information security culture". Computers and Security, 29, s. 196–207.

92. Vuk, M., Curk, T., 2006. "ROC curve, lift chart and calibration plot". Metodoloski Zvezki - Advances in Methodology and Statistics, 3(1), s. 89-108.

93. Wang, J., Neskovic, P., Cooper, L. N., 2007. "Improving nearest neighbor rule with a simple adaptive distance measure". Pattern Recognition Letters, 28, s. 207–213.

94. Weiss, G. M., Provost, F., 2003. "Learning when training data are costly: the effect of class distribution on tree induction". Journal of Artificial Intelligence Research, 19, s. 315-354.

95. Weka, 2011. *Weka 3: Data Mining Software in Java.* http://www.cs.waikato.ac.nz/ml/weka/index.html, University of Waikato, Internet sitesi, (Erişim: 15 Temmuz 2011).

96. Williams, J. W., Li, Y., 2010. "Comparative Study of Distance Functions for Nearest Neighbors". Advanced Techniques in Computing Sciences and Software Engineering, Springer, s. 79-84.

97. Williams, P., 2008. "In a 'trusting' environment, everyone is responsible for information security". Information Security Technical Report, Elsevier, 13, s. 207-215.

98. Witten, I. H., Frank, E., Hall, M. A., 2011. *Data Mining: Practical Machine Learning Tools and Techniques 3rd edition*. Elsevier Inc., Burlington, MA.

99. Zakaria, O., 2006. “Employee Security Perception in Cultivating Information Security Culture”. Security Management, Integrity, and Internal Control in Information Systems, IFIP International Federation for Information Processing, 193, s. 83-92.

100. Zhang, D., Jeffrey, J. P. T., 2007. Advances in Machine Learning Applications in Software Engineering. Idea Group Inc., UK.

101. Zhao, D-M., Wang, J-H., Wu, J., Ma, J-F, 2005. “Using fuzzy logic and entropy theory to risk assessment of the information security”. In Proceedings of the Fourth International Conference on Machine Learning and Cybernetics. Guangzhou, China, 18-21 August. IEEE pub.

ÖZGEÇMİŞ

Mete Eminağaoğlu 1973 yılında İstanbul'da doğdu. İlk ve orta öğrenimini İzmir'de tamamladı. Ege Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümünü 1996 yılında bitirdi. İzmir Yüksek Teknoloji Enstitüsü Bilgisayar Mühendisliği Bölümünde 1996 ile 1999 yılları arasında araştırma görevlisi olarak çalıştı ve aynı üniversitenin Bilgisayar Mühendisliği Bölümünde 1999 yılında Yüksek Lisans eğitimini tamamladı. Doktora eğitimine Trakya Üniversitesi Mühendislik Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümünde devam etti. Özel sektörde bilgi güvenliğinin çeşitli alanlarında baş denetçi, eğitmen, danışman ve üst düzey yönetici olarak toplam 14 yıl görev yaptı. Mete Eminağaoğlu, 2009 yılından itibaren Yaşar Üniversitesi Meslek Yüksek Okulu Bilgisayar Programcılığı'nda öğretim görevlisi olarak mesleki kariyerini sürdürmektedir. "CISSP" ve "CISA" gibi uluslararası bilgi güvenliği uzmanlığı ve bilgi sistemleri denetçisi sertifikalarına sahiptir. Bilgi güvenliği, risk analizi, veri madenciliği ve özdevimli öğrenim konuları başta olmak üzere uluslararası alan endeksli dergilerde yayınlanmış çeşitli makaleleri bulunmakta olup yurt içi ve yurt dışındaki akademik seminerlerde bildiriler sunmuştur.

EKLER

EK-A
İKİLİ SINIFLANDIRICILARIN GÖZLEM DEĞERLERİ

Bu bölümde, tez çalışmasındaki ilgili model için çeşitli ikili sınıflandırıcı tipinde özdevimli öğrenme algoritmalarının öğrenme ve test başarımlarına yönelik yapılmış olan deneyler, gözlemler ve elde edilen istatistiksel ve matematiksel sonuçlar, ilgili sayısal veriler yer almaktadır. Ekteki tüm veriler, Weka yazılımının ilgili her bir deney ve gözleme ait elde edilmiş olan sonuç çıktı raporlarından derlenmiştir.

Deneylerde kullanılan sınıflandırıcının adı, tipi, parametreleri, elde edilen sonuçlar, Doğru Negatif (TN rate), Doğru Pozitif (TP rate), Yanlış Negatif (FN rate), Yanlış Pozitif (FP rate), adetleri ve ilgili oranları, Kappa istatistiği (Kappa statistic) değerleri, Duyarlılık (Precision), Geri çağırım (Recall) değerleri, Hata Matrisi (Confusion Matrix) değerleri ve diğer istatistiksel hesaplar verilmektedir. Ayrıca, bu değerler, doğru kestirim kıstasının riskin üst (“U”) olduğu durumlar için ayrı, riskin alt (“A”) olduğu durumlar için ayrı olarak verilmekte ve bu ikisinin ağırlıklı ortalaması (Weighted Avg.) olarak ayrıca gösterilmektedir.

```
=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-M -V -1 -W 1 -K "weka.classifiers.functions.supportVector.RBFKernel -C 250007
-G 0.15"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===
```

```
SMO

Kernel used:
  RBF kernel: K(x,y) = e^-(0.15* <x-y,x-y>^2)

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1136

Number of kernel evaluations: 7757765 (52.265% cached)

Logistic Regression with ridge parameter of 1.0E-8
Coefficients...
               Class
Variable           A
====================
pred         -2.8281
Intercept     0.1293


Odds Ratios...
               Class
Variable           A
====================
pred          0.0591



Time taken to build model: 20.29 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1520               86.8571 %
Incorrectly Classified Instances       230               13.1429 %
Kappa statistic                          0.7353
Mean absolute error                      0.1698
Root mean squared error                  0.2977
Relative absolute error                 34.1417 %
Root relative squared error             59.7092 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.848     0.114      0.865     0.848     0.857      0.959
A
                 0.886     0.152      0.871     0.886     0.879      0.959
U
Weighted Avg.    0.869     0.134      0.869     0.869     0.868      0.959

=== Confusion Matrix ===

   a   b   <-- classified as
 687 123 |   a = A
 107 833 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-M -V -1 -W 1 -K "weka.classifiers.functions.supportVector.RBFKernel -C 250007
-G 0.15"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

SMO

Kernel used:
  RBF kernel: K(x,y) = e^-(0.15* <x-y,x-y>^2)

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1136

Number of kernel evaluations: 7757765 (52.265% cached)

Logistic Regression with ridge parameter of 1.0E-8
Coefficients...
               Class
Variable           A
====================
pred         -2.8281
Intercept     0.1293


Odds Ratios...
               Class
Variable           A
====================
pred          0.0591

Time taken to build model: 20.06 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1360               77.7143 %
Incorrectly Classified Instances       390               22.2857 %
Kappa statistic                          0.5513
Mean absolute error                      0.2604
Root mean squared error                  0.4085
Relative absolute error                 52.3777 %
Root relative squared error             81.9273 %
```

```
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.751     0.2        0.764     0.751     0.757      0.856
A
                 0.8       0.249      0.788     0.8       0.794      0.856
U
Weighted Avg.    0.777     0.227      0.777     0.777     0.777      0.856

=== Confusion Matrix ===

   a   b   <-- classified as
 608 202 |   a = A
 188 752 |   b = U


=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.REPTree -- -M 2 -V 0.0010 -N 3 -S 1 -L -1 -P
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

All the base classifiers:

FilteredClassifier using weka.classifiers.trees.REPTree -M 2 -V 0.0010 -N 3 -S
1 -L -1 -P on data filtered through weka.filters.unsupervised.attribute.Remove
-V -R 2,7,6,1,5,10,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data


Classifier Model
```

```
REPTree
============

BUTUNLUK KAYBI = 1
|   KULLANILABILIRLIK KAYBI = 1
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (5/0) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (2/1) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (4/1) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (8/1) [0/0]
|   |   |   |   VARLIK = V3 : A (4/1) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (7/2) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (9/0) [0/0]
|   |   |   TEHDIT = T4
|   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (17/0) [0/0]
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5 : A (0/0) [0/0]
|   |   |   |   |   VARLIK = V1 : U (1/0) [0/0]
|   |   |   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   |   |   VARLIK = V3 : A (2/0) [0/0]
|   |   |   |   |   VARLIK = V2 : A (2/1) [0/0]
|   |   |   |   |   VARLIK = V6 : A (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (12/0) [0/0]
|   |   |   TEHDIT = T10 : A (6/0) [0/0]
|   |   |   TEHDIT = T2 : A (11/0) [0/0]
|   |   |   TEHDIT = T1 : A (4/0) [0/0]
|   |   |   TEHDIT = T3
|   |   |   |   VARLIK = V5 : A (5/0) [0/0]
|   |   |   |   VARLIK = V1 : A (4/1) [0/0]
|   |   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   |   VARLIK = V3 : A (6/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (0/0) [0/0]
|   |   |   TEHDIT = T8 : A (15/0) [0/0]
|   |   |   TEHDIT = T5 : A (4/0) [0/0]
|   |   |   TEHDIT = T6 : A (6/0) [0/0]
|   |   GIZLILIK KAYBI = 2
|   |   |   VARLIK = V5 : A (2/0) [0/0]
|   |   |   VARLIK = V1 : A (2/1) [0/0]
|   |   |   VARLIK = V4 : A (1/0) [0/0]
|   |   |   VARLIK = V3 : A (2/0) [0/0]
|   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   VARLIK = V6 : A (2/1) [0/0]
|   |   GIZLILIK KAYBI = 3
|   |   |   TEHDIT = T7 : A (0/0) [0/0]
|   |   |   TEHDIT = T4 : U (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : A (2/0) [0/0]
|   |   |   TEHDIT = T1 : U (3/1) [0/0]
|   |   |   TEHDIT = T3 : A (0/0) [0/0]
|   |   |   TEHDIT = T8 : U (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (0/0) [0/0]
|   |   |   TEHDIT = T6 : A (1/0) [0/0]
|   |   GIZLILIK KAYBI = 4 : A (0/0) [0/0]
|   |   GIZLILIK KAYBI = 5 : A (2/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1
```

```
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (3/0) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (2/1) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4 : A (3/0) [0/0]
|   |   |   |   VARLIK = V3 : U (1/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (1/0) [0/0]
|   |   |   TEHDIT = T4 : A (2/0) [0/0]
|   |   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : A (4/0) [0/0]
|   |   |   TEHDIT = T1 : A (1/0) [0/0]
|   |   |   TEHDIT = T3 : A (7/0) [0/0]
|   |   |   TEHDIT = T8 : A (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (2/0) [0/0]
|   |   |   TEHDIT = T6 : A (0/0) [0/0]
|   |   GIZLILIK KAYBI = 2
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5 : A (1/0) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4 : A (2/0) [0/0]
|   |   |   |   VARLIK = V3 : A (0/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (2/1) [0/0]
|   |   |   TEHDIT = T4 : A (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (0/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : U (1/0) [0/0]
|   |   |   TEHDIT = T1 : A (1/0) [0/0]
|   |   |   TEHDIT = T3 : A (4/0) [0/0]
|   |   |   TEHDIT = T8 : A (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (2/1) [0/0]
|   |   |   TEHDIT = T6 : A (3/1) [0/0]
|   |   GIZLILIK KAYBI = 3 : U (3/1) [0/0]
|   |   GIZLILIK KAYBI = 4 : A (2/0) [0/0]
|   |   GIZLILIK KAYBI = 5 : U (2/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 3
|   |   TEHDIT = T7 : A (11/0) [0/0]
|   |   TEHDIT = T4 : U (1/0) [0/0]
|   |   TEHDIT = T9 : A (2/1) [0/0]
|   |   TEHDIT = T10 : A (2/0) [0/0]
|   |   TEHDIT = T2
|   |   |   VARLIK = V5 : A (0/0) [0/0]
|   |   |   VARLIK = V1 : A (2/1) [0/0]
|   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   VARLIK = V3 : U (1/0) [0/0]
|   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   VARLIK = V6 : A (2/0) [0/0]
|   |   TEHDIT = T1 : A (0/0) [0/0]
|   |   TEHDIT = T3 : A (0/0) [0/0]
|   |   TEHDIT = T8 : A (0/0) [0/0]
|   |   TEHDIT = T5 : A (0/0) [0/0]
|   |   TEHDIT = T6 : A (2/1) [0/0]
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7 : U (1/0) [0/0]
|   |   TEHDIT = T4 : A (1/0) [0/0]
|   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   TEHDIT = T10 : U (0/0) [0/0]
|   |   TEHDIT = T2 : U (4/0) [0/0]
|   |   TEHDIT = T1 : U (0/0) [0/0]
|   |   TEHDIT = T3 : U (1/0) [0/0]
|   |   TEHDIT = T8 : U (1/0) [0/0]
|   |   TEHDIT = T5 : U (1/0) [0/0]
```

```
|   |   TEHDIT = T6 : A (3/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 5 : U (2/0) [0/0]


Size of the tree : 1122

Time taken to build model: 3.48 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1675               95.7143 %
Incorrectly Classified Instances        75                4.2857 %
Kappa statistic                          0.9136
Mean absolute error                      0.1996
Root mean squared error                  0.2445
Relative absolute error                 40.1439 %
Root relative squared error             49.0365 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.937     0.026      0.969     0.937     0.953      0.995
A
                 0.974     0.063      0.947     0.974     0.961      0.995
U
Weighted Avg.    0.957     0.046      0.957     0.957     0.957      0.995

=== Confusion Matrix ===

   a   b   <-- classified as
 759  51 |   a = A
  24 916 |   b = U


=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.REPTree -- -M 2 -V 0.0010 -N 3 -S 1 -L -1 -P
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

All the base classifiers:
```

FilteredClassifier using weka.classifiers.trees.REPTree -M 2 -V 0.0010 -N 3 -S 1 -L -1 -P on data filtered through weka.filters.unsupervised.attribute.Remove -V -R 2,7,6,1,5,10,12

```
Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data


Classifier Model

REPTree
============

BUTUNLUK KAYBI = 1
|   KULLANILABILIRLIK KAYBI = 1
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (5/0) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (2/1) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (4/1) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (8/1) [0/0]
|   |   |   |   VARLIK = V3 : A (4/1) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (7/2) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (9/0) [0/0]
|   |   |   TEHDIT = T4
|   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (17/0) [0/0]
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5 : A (0/0) [0/0]
|   |   |   |   |   VARLIK = V1 : U (1/0) [0/0]
|   |   |   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   |   |   VARLIK = V3 : A (2/0) [0/0]
|   |   |   |   |   VARLIK = V2 : A (2/1) [0/0]
|   |   |   |   |   VARLIK = V6 : A (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (12/0) [0/0]
|   |   |   TEHDIT = T10 : A (6/0) [0/0]
|   |   |   TEHDIT = T2 : A (11/0) [0/0]
|   |   |   TEHDIT = T1 : A (4/0) [0/0]
|   |   |   TEHDIT = T3
|   |   |   |   VARLIK = V5 : A (5/0) [0/0]
|   |   |   |   VARLIK = V1 : A (4/1) [0/0]
|   |   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   |   VARLIK = V3 : A (6/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (0/0) [0/0]
|   |   |   TEHDIT = T8 : A (15/0) [0/0]
|   |   |   TEHDIT = T5 : A (4/0) [0/0]
|   |   |   TEHDIT = T6 : A (6/0) [0/0]
|   |   GIZLILIK KAYBI = 2
|   |   |   VARLIK = V5 : A (2/0) [0/0]
```

```
|   |   |   VARLIK = V1 : A (2/1) [0/0]
|   |   |   VARLIK = V4 : A (1/0) [0/0]
|   |   |   VARLIK = V3 : A (2/0) [0/0]
|   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   VARLIK = V6 : A (2/1) [0/0]
|   |   GIZLILIK KAYBI = 3
|   |   |   TEHDIT = T7 : A (0/0) [0/0]
|   |   |   TEHDIT = T4 : U (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : A (2/0) [0/0]
|   |   |   TEHDIT = T1 : U (3/1) [0/0]
|   |   |   TEHDIT = T3 : A (0/0) [0/0]
|   |   |   TEHDIT = T8 : U (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (0/0) [0/0]
|   |   |   TEHDIT = T6 : A (1/0) [0/0]
|   |   GIZLILIK KAYBI = 4 : A (0/0) [0/0]
|   |   GIZLILIK KAYBI = 5 : A (2/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = H : A (3/0) [0/0]
|   |   |   |   |   KURUM ICI EGITIM ALDI MI = E : A (2/1) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4 : A (3/0) [0/0]
|   |   |   |   VARLIK = V3 : U (1/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (1/0) [0/0]
|   |   |   TEHDIT = T4 : A (2/0) [0/0]
|   |   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : A (4/0) [0/0]
|   |   |   TEHDIT = T1 : A (1/0) [0/0]
|   |   |   TEHDIT = T3 : A (7/0) [0/0]
|   |   |   TEHDIT = T8 : A (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (2/0) [0/0]
|   |   |   TEHDIT = T6 : A (0/0) [0/0]
|   |   GIZLILIK KAYBI = 2
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5 : A (1/0) [0/0]
|   |   |   |   VARLIK = V1 : A (0/0) [0/0]
|   |   |   |   VARLIK = V4 : A (2/0) [0/0]
|   |   |   |   VARLIK = V3 : A (0/0) [0/0]
|   |   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   |   VARLIK = V6 : A (2/1) [0/0]
|   |   |   TEHDIT = T4 : A (1/0) [0/0]
|   |   |   TEHDIT = T9 : A (0/0) [0/0]
|   |   |   TEHDIT = T10 : A (0/0) [0/0]
|   |   |   TEHDIT = T2 : U (1/0) [0/0]
|   |   |   TEHDIT = T1 : A (1/0) [0/0]
|   |   |   TEHDIT = T3 : A (4/0) [0/0]
|   |   |   TEHDIT = T8 : A (1/0) [0/0]
|   |   |   TEHDIT = T5 : A (2/1) [0/0]
|   |   |   TEHDIT = T6 : A (3/1) [0/0]
|   |   GIZLILIK KAYBI = 3 : U (3/1) [0/0]
|   |   GIZLILIK KAYBI = 4 : A (2/0) [0/0]
|   |   GIZLILIK KAYBI = 5 : U (2/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 3
|   |   TEHDIT = T7 : A (11/0) [0/0]
|   |   TEHDIT = T4 : U (1/0) [0/0]
|   |   TEHDIT = T9 : A (2/1) [0/0]
|   |   TEHDIT = T10 : A (2/0) [0/0]
|   |   TEHDIT = T2
|   |   |   VARLIK = V5 : A (0/0) [0/0]
```

```
|   |   |   VARLIK = V1 : A (2/1) [0/0]
|   |   |   VARLIK = V4 : A (0/0) [0/0]
|   |   |   VARLIK = V3 : U (1/0) [0/0]
|   |   |   VARLIK = V2 : A (0/0) [0/0]
|   |   |   VARLIK = V6 : A (2/0) [0/0]
|   |   TEHDIT = T1 : A (0/0) [0/0]
|   |   TEHDIT = T3 : A (0/0) [0/0]
|   |   TEHDIT = T8 : A (0/0) [0/0]
|   |   TEHDIT = T5 : A (0/0) [0/0]
|   |   TEHDIT = T6 : A (2/1) [0/0]
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7 : U (1/0) [0/0]
|   |   TEHDIT = T4 : A (1/0) [0/0]
|   |   TEHDIT = T9 : A (1/0) [0/0]
|   |   TEHDIT = T10 : U (0/0) [0/0]
|   |   TEHDIT = T2 : U (4/0) [0/0]
|   |   TEHDIT = T1 : U (0/0) [0/0]
|   |   TEHDIT = T3 : U (1/0) [0/0]
|   |   TEHDIT = T8 : U (1/0) [0/0]
|   |   TEHDIT = T5 : U (1/0) [0/0]
|   |   TEHDIT = T6 : A (3/0) [0/0]
|   KULLANILABILIRLIK KAYBI = 5 : U (2/0) [0/0]

Size of the tree : 1122


Time taken to build model: 3.15 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1368               78.1714 %
Incorrectly Classified Instances       382               21.8286 %
Kappa statistic                          0.5597
Mean absolute error                      0.3208
Root mean squared error                  0.3904
Relative absolute error                 64.5141 %
Root relative squared error             78.2957 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.743     0.185      0.776     0.743     0.759      0.86
A
                 0.815     0.257      0.786     0.815     0.8        0.86
U
Weighted Avg.    0.782     0.224      0.782     0.782     0.781      0.86

=== Confusion Matrix ===

   a   b   <-- classified as
 602 208 |   a = A
 174 766 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -I 92 -K 0 -S 1 -depth 3
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

Random forest of 92 trees, each constructed while considering 4 random
features.
Out of bag error: 0.2314
Max. depth of trees: 3


Time taken to build model: 0.73 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1459               83.3714 %
Incorrectly Classified Instances       291               16.6286 %
Kappa statistic                          0.6648
Mean absolute error                      0.2819
Root mean squared error                  0.3444
Relative absolute error                 56.6905 %
Root relative squared error             69.0665 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.802     0.139      0.832     0.802     0.817      0.924
A
                 0.861     0.198      0.835     0.861     0.848      0.924
U
Weighted Avg.    0.834     0.171      0.834     0.834     0.833      0.924

=== Confusion Matrix ===

   a   b   <-- classified as
 650 160 |   a = A
 131 809 |   b = U
```

=== Run information ===

Scheme: **weka.classifiers.trees.RandomForest -I 92 -K 0 -S 1 -depth 3**
Relation: ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last
Instances: 1750
Attributes: 12
VARLIK
TEHDIT
ZAYIFLIK
RISKIN GERCEKLESME OLASILIGI
GIZLILIK KAYBI
BUTUNLUK KAYBI
KULLANILABILIRLIK KAYBI
SON 1 YILDAKI VAKA ADEDI
BILGI BIRIKIMI DUZEYI
KURUM ICI EGITIM ALDI MI
KURUMDA POLITIKA VB MEVCUT MU
RISK DUZEYI (ANKET SONUCU)
Test mode: **10-fold cross-validation**

=== Classifier model (full training set) ===

Random forest of 92 trees, each constructed while considering 4 random features.
Out of bag error: 0.2314
Max. depth of trees: 3

Time taken to build model: 0.81 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	1363	77.8857 %
Incorrectly Classified Instances	387	22.1143 %
Kappa statistic	0.5548	
Mean absolute error	0.3231	
Root mean squared error	0.3933	
Relative absolute error	64.9728 %	
Root relative squared error	78.8838 %	
Total Number of Instances	1750	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.753	0.199	0.765	0.753	0.759	0.856	A
	0.801	0.247	0.79	0.801	0.796	0.856	U
Weighted Avg.	0.779	0.225	0.779	0.779	0.779	0.856	

=== Confusion Matrix ===

a	b	<-- classified as
610	200	a = A
187	753	b = U

```
=== Run information ===

Scheme:       weka.classifiers.lazy.IBk -K 12 -W 0 -F -A
"weka.core.neighboursearch.LinearNNSearch -A \"weka.core.ManhattanDistance -R
first-last\""
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

IB1 instance-based classifier
using 12 similarity-weighted nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1099               78.444  %
Incorrectly Classified Instances       140                9.9929 %
Kappa statistic                          0.7717
Mean absolute error                      0.2866
Root mean squared error                  0.4913
Relative absolute error                 65.2045 %
Root relative squared error            104.8236 %
UnClassified Instances                 162               11.5632 %
Total Number of Instances             1401



=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.853     0.084      0.897     0.853     0.874      0.868
A
                 0.916     0.147      0.879     0.916     0.898      0.852
U
Weighted Avg.    0.887     0.118      0.887     0.887     0.887      0.86

=== Confusion Matrix ===

   a   b   <-- classified as
 486  84 |   a = A
  56 613 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.lazy.IBk -K 12 -W 0 -F -A
"weka.core.neighboursearch.LinearNNSearch -A \"weka.core.ManhattanDistance -R
first-last\""
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

IB1 instance-based classifier
using 12 similarity-weighted nearest neighbour(s) for classification


Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         815               71.9965 %
Incorrectly Classified Instances       193               17.0495 %
Kappa statistic                          0.6137
Mean absolute error                      0.2811
Root mean squared error                  0.3979
Relative absolute error                 63.4119 %
Root relative squared error             84.4636 %
UnClassified Instances                 124               10.9541 %
Total Number of Instances             1132

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.751     0.14       0.826     0.751     0.787      0.801
A
                 0.86      0.249      0.795     0.86      0.826      0.806
U
Weighted Avg.    0.809     0.198      0.81      0.809     0.808      0.804

=== Confusion Matrix ===

   a   b   <-- classified as
 356 118 |   a = A
  75 459 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-V -1 -W 1 -K "weka.classifiers.functions.supportVector.NormalizedPolyKernel -
C 250007 -E 3.5"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

SMO

Kernel used:
  Normalized Poly Kernel: K(x,y) = <x,y>^3.5/(<x,x>^3.5*<y,y>^3.5)^(1/2)

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1180

Number of kernel evaluations: 5950836 (86.731% cached)


Time taken to build model: 17.51 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1578               90.1714 %
Incorrectly Classified Instances       172                9.8286 %
Kappa statistic                          0.8017
Mean absolute error                      0.0983
Root mean squared error                  0.3135
Relative absolute error                 19.7661 %
Root relative squared error             62.8748 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.87      0.071      0.913     0.87      0.891      0.9
A
                 0.929     0.13       0.893     0.929     0.91       0.9
U
Weighted Avg.    0.902     0.103      0.902     0.902     0.902      0.9

=== Confusion Matrix ===
```

```
   a   b   <-- classified as
 705 105 |   a = A
  67 873 |   b = U


=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-V -1 -W 1 -K "weka.classifiers.functions.supportVector.NormalizedPolyKernel -
C 250007 -E 3.5"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

SMO

Kernel used:
  Normalized Poly Kernel: K(x,y) = <x,y>^3.5/(<x,x>^3.5*<y,y>^3.5)^(1/2)

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1180

Number of kernel evaluations: 5950836 (86.731% cached)



Time taken to build model: 17.53 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1361               77.7714 %
Incorrectly Classified Instances       389               22.2286 %
Kappa statistic                          0.5508
Mean absolute error                      0.2223
Root mean squared error                  0.4715
Relative absolute error                 44.7035 %
Root relative squared error             94.5556 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.725     0.177      0.78      0.725     0.751      0.774
A
```

```
                 0.823     0.275      0.776     0.823     0.799      0.774
U
Weighted Avg.    0.778     0.23       0.778     0.778     0.777      0.774

=== Confusion Matrix ===

   a   b   <-- classified as
 587 223 |   a = A
 166 774 |   b = U


=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-V -1 -W 1 -K "weka.classifiers.functions.supportVector.Puk -C 250007 -O 0.9 -
S 6.4"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

SMO

Kernel used:
  Puk kernel

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1126

Number of kernel evaluations: 6529658 (55.816% cached)


Time taken to build model: 18.22 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1431               81.7714 %
Incorrectly Classified Instances       319               18.2286 %
Kappa statistic                          0.6329
Mean absolute error                      0.1823
Root mean squared error                  0.4269
Relative absolute error                 36.6592 %
Root relative squared error             85.6264 %
Total Number of Instances             1750
```

```
=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.794     0.162      0.809     0.794     0.801      0.816
A
                 0.838     0.206      0.825     0.838     0.832      0.816
U
Weighted Avg.    0.818     0.186      0.818     0.818     0.818      0.816

=== Confusion Matrix ===

   a   b   <-- classified as
 643 167 |   a = A
 152 788 |   b = U


=== Run information ===

Scheme:       weka.classifiers.functions.SMO -C 1.0 -L 0.0010 -P 1.0E-12 -N 0
-V -1 -W 1 -K "weka.classifiers.functions.supportVector.Puk -C 250007 -O 0.9 -
S 6.4"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

SMO

Kernel used:
  Puk kernel

Classifier for classes: A, U

BinarySMO

Number of support vectors: 1126

Number of kernel evaluations: 6529658 (55.816% cached)



Time taken to build model: 17.76 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1366               78.0571 %
Incorrectly Classified Instances       384               21.9429 %
Kappa statistic                          0.5588
```

```
Mean absolute error                      0.2194
Root mean squared error                  0.4684
Relative absolute error                 44.1289 %
Root relative squared error             93.946  %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.764     0.205      0.762     0.764     0.763      0.779
A
                 0.795     0.236      0.796     0.795     0.796      0.779
U
Weighted Avg.    0.781     0.222      0.781     0.781     0.781      0.779

=== Confusion Matrix ===

   a   b   <-- classified as
 619 191 |   a = A
 193 747 |   b = U


=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.J48 -- -U -M 2
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

All the base classifiers:

FilteredClassifier using weka.classifiers.trees.J48 -U -M 2 on data filtered
through weka.filters.unsupervised.attribute.Remove -V -R 2,7,6,1,5,10,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data
```

Classifier Model
J48 unpruned tree

```
BUTUNLUK KAYBI = 1
|   KULLANILABILIRLIK KAYBI = 1
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7: A (39.0/6.0)
|   |   |   TEHDIT = T4
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (17.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5: A (0.0)
|   |   |   |   |   VARLIK = V1: U (1.0)
|   |   |   |   |   VARLIK = V4: A (0.0)
|   |   |   |   |   VARLIK = V3: A (2.0)
|   |   |   |   |   VARLIK = V2: A (2.0/1.0)
|   |   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T9: A (12.0)
|   |   |   TEHDIT = T10: A (6.0)
|   |   |   TEHDIT = T2: A (11.0)
|   |   |   TEHDIT = T1: A (4.0)
|   |   |   TEHDIT = T3: A (15.0/1.0)
|   |   |   TEHDIT = T8: A (15.0)
|   |   |   TEHDIT = T5: A (4.0)
|   |   |   TEHDIT = T6: A (6.0)
|   |   GIZLILIK KAYBI = 2: A (9.0/2.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   TEHDIT = T7: A (0.0)
|   |   |   TEHDIT = T4: U (1.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (2.0)
|   |   |   TEHDIT = T1: U (3.0/1.0)
|   |   |   TEHDIT = T3: A (0.0)
|   |   |   TEHDIT = T8: U (1.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: A (1.0)
|   |   GIZLILIK KAYBI = 4: A (0.0)
|   |   GIZLILIK KAYBI = 5: A (2.0)
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5: A (5.0/1.0)
|   |   |   |   VARLIK = V1: A (0.0)
|   |   |   |   VARLIK = V4: A (3.0)
|   |   |   |   VARLIK = V3: U (1.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T4: A (2.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (4.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (7.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0)
|   |   |   TEHDIT = T6: A (0.0)
|   |   GIZLILIK KAYBI = 2
|   |   |   TEHDIT = T7: A (5.0/1.0)
|   |   |   TEHDIT = T4: A (1.0)
|   |   |   TEHDIT = T9: A (0.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: U (1.0)
```

```
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (4.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0/1.0)
|   |   |   TEHDIT = T6: A (3.0/1.0)
|   |   GIZLILIK KAYBI = 3: U (3.0/1.0)
|   |   GIZLILIK KAYBI = 4: A (2.0)
|   |   GIZLILIK KAYBI = 5: U (2.0)
|   KULLANILABILIRLIK KAYBI = 3
|   |   TEHDIT = T7: A (11.0)
|   |   TEHDIT = T4: U (1.0)
|   |   TEHDIT = T9: A (2.0/1.0)
|   |   TEHDIT = T10: A (2.0)
|   |   TEHDIT = T2
|   |   |   VARLIK = V5: A (0.0)
|   |   |   VARLIK = V1: A (2.0/1.0)
|   |   |   VARLIK = V4: A (0.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (2.0)
|   |   TEHDIT = T1: A (0.0)
|   |   TEHDIT = T3: A (0.0)
|   |   TEHDIT = T8: A (0.0)
|   |   TEHDIT = T5: A (0.0)
|   |   TEHDIT = T6: A (2.0/1.0)
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7: U (1.0)
|   |   TEHDIT = T4: A (1.0)
|   |   TEHDIT = T9: A (1.0)
|   |   TEHDIT = T10: U (0.0)
|   |   TEHDIT = T2: U (4.0)
|   |   TEHDIT = T1: U (0.0)
|   |   TEHDIT = T3: U (1.0)
|   |   TEHDIT = T8: U (1.0)
|   |   TEHDIT = T5: U (1.0)
|   |   TEHDIT = T6: A (3.0)
|   KULLANILABILIRLIK KAYBI = 5: U (2.0)
…
…
…
```

Number of Leaves : 707

Size of the tree : 831

```
Time taken to build model: 2.55 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1641               93.7714 %
Incorrectly Classified Instances       109                6.2286 %
Kappa statistic                          0.8744
Mean absolute error                      0.2155
Root mean squared error                  0.2622
Relative absolute error                 43.3454 %
Root relative squared error             52.588  %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===
```

```
               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.911     0.039      0.952     0.911     0.931      0.986
A
                 0.961     0.089      0.926     0.961     0.943      0.986
U
Weighted Avg.    0.938     0.066      0.938     0.938     0.938      0.986

=== Confusion Matrix ===

   a   b   <-- classified as
 738  72 |   a = A
  37 903 |   b = U


=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.J48 -- -U -M 2
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

All the base classifiers:

FilteredClassifier using weka.classifiers.trees.J48 -U -M 2 on data filtered
through weka.filters.unsupervised.attribute.Remove -V -R 2,7,6,1,5,10,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data


Classifier Model
J48 unpruned tree
------------------
```

```
BUTUNLUK KAYBI = 1
|   KULLANILABILIRLIK KAYBI = 1
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7: A (39.0/6.0)
|   |   |   TEHDIT = T4
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (17.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5: A (0.0)
|   |   |   |   |   VARLIK = V1: U (1.0)
|   |   |   |   |   VARLIK = V4: A (0.0)
|   |   |   |   |   VARLIK = V3: A (2.0)
|   |   |   |   |   VARLIK = V2: A (2.0/1.0)
|   |   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T9: A (12.0)
|   |   |   TEHDIT = T10: A (6.0)
|   |   |   TEHDIT = T2: A (11.0)
|   |   |   TEHDIT = T1: A (4.0)
|   |   |   TEHDIT = T3: A (15.0/1.0)
|   |   |   TEHDIT = T8: A (15.0)
|   |   |   TEHDIT = T5: A (4.0)
|   |   |   TEHDIT = T6: A (6.0)
|   |   GIZLILIK KAYBI = 2: A (9.0/2.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   TEHDIT = T7: A (0.0)
|   |   |   TEHDIT = T4: U (1.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (2.0)
|   |   |   TEHDIT = T1: U (3.0/1.0)
|   |   |   TEHDIT = T3: A (0.0)
|   |   |   TEHDIT = T8: U (1.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: A (1.0)
|   |   GIZLILIK KAYBI = 4: A (0.0)
|   |   GIZLILIK KAYBI = 5: A (2.0)
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5: A (5.0/1.0)
|   |   |   |   VARLIK = V1: A (0.0)
|   |   |   |   VARLIK = V4: A (3.0)
|   |   |   |   VARLIK = V3: U (1.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T4: A (2.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (4.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (7.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0)
|   |   |   TEHDIT = T6: A (0.0)
|   |   GIZLILIK KAYBI = 2
|   |   |   TEHDIT = T7: A (5.0/1.0)
|   |   |   TEHDIT = T4: A (1.0)
|   |   |   TEHDIT = T9: A (0.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: U (1.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (4.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0/1.0)
|   |   |   TEHDIT = T6: A (3.0/1.0)
|   |   GIZLILIK KAYBI = 3: U (3.0/1.0)
```

```
|   |   GIZLILIK KAYBI = 4: A (2.0)
|   |   GIZLILIK KAYBI = 5: U (2.0)
|   KULLANILABILIRLIK KAYBI = 3
|   |   TEHDIT = T7: A (11.0)
|   |   TEHDIT = T4: U (1.0)
|   |   TEHDIT = T9: A (2.0/1.0)
|   |   TEHDIT = T10: A (2.0)
|   |   TEHDIT = T2
|   |   |   VARLIK = V5: A (0.0)
|   |   |   VARLIK = V1: A (2.0/1.0)
|   |   |   VARLIK = V4: A (0.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (2.0)
|   |   TEHDIT = T1: A (0.0)
|   |   TEHDIT = T3: A (0.0)
|   |   TEHDIT = T8: A (0.0)
|   |   TEHDIT = T5: A (0.0)
|   |   TEHDIT = T6: A (2.0/1.0)
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7: U (1.0)
|   |   TEHDIT = T4: A (1.0)
|   |   TEHDIT = T9: A (1.0)
|   |   TEHDIT = T10: U (0.0)
|   |   TEHDIT = T2: U (4.0)
|   |   TEHDIT = T1: U (0.0)
|   |   TEHDIT = T3: U (1.0)
|   |   TEHDIT = T8: U (1.0)
|   |   TEHDIT = T5: U (1.0)
|   |   TEHDIT = T6: A (3.0)
|   KULLANILABILIRLIK KAYBI = 5: U (2.0)
BUTUNLUK KAYBI = 2
|   KULLANILABILIRLIK KAYBI = 1: A (36.0/7.0)
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1: A (36.0/2.0)
|   |   GIZLILIK KAYBI = 2: A (80.0/8.0)
|   |   GIZLILIK KAYBI = 3: A (30.0/5.0)
|   |   GIZLILIK KAYBI = 4: U (3.0/1.0)
|   |   GIZLILIK KAYBI = 5: A (3.0)
|   KULLANILABILIRLIK KAYBI = 3
|   |   VARLIK = V5
|   |   |   KURUM ICI EGITIM ALDI MI = H: A (3.0/1.0)
|   |   |   KURUM ICI EGITIM ALDI MI = E: U (3.0)
|   |   VARLIK = V1: A (10.0/1.0)
|   |   VARLIK = V4
|   |   |   GIZLILIK KAYBI = 1: A (1.0)
|   |   |   GIZLILIK KAYBI = 2
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (2.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E: U (3.0/1.0)
|   |   |   GIZLILIK KAYBI = 3: U (6.0/2.0)
|   |   |   GIZLILIK KAYBI = 4: A (0.0)
|   |   |   GIZLILIK KAYBI = 5: A (1.0)
|   |   VARLIK = V3
|   |   |   TEHDIT = T7: A (1.0)
|   |   |   TEHDIT = T4: A (5.0/1.0)
|   |   |   TEHDIT = T9: A (0.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (3.0)
|   |   |   TEHDIT = T1: A (0.0)
|   |   |   TEHDIT = T3: A (2.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: U (1.0)
|   |   VARLIK = V2: A (4.0)
|   |   VARLIK = V6: A (2.0/1.0)
```

```
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7
|   |   |   VARLIK = V5: A (2.0/1.0)
|   |   |   VARLIK = V1: A (0.0)
|   |   |   VARLIK = V4: A (1.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (5.0)
|   |   TEHDIT = T4: A (2.0/1.0)
|   |   TEHDIT = T9: A (1.0)
|   |   TEHDIT = T10: U (0.0)
|   |   TEHDIT = T2: U (5.0)
|   |   TEHDIT = T1: A (2.0/1.0)
|   |   TEHDIT = T3: A (2.0)
|   |   TEHDIT = T8: U (7.0/1.0)
|   |   TEHDIT = T5: A (2.0/1.0)
|   |   TEHDIT = T6: U (5.0/2.0)
|   KULLANILABILIRLIK KAYBI = 5: A (7.0)
…
…
…
```

Number of Leaves : 707

Size of the tree : 831

Time taken to build model: 2.46 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	1376	78.6286 %
Incorrectly Classified Instances	374	21.3714 %
Kappa statistic	0.5688	
Mean absolute error	0.322	
Root mean squared error	0.3891	
Relative absolute error	64.75 %	
Root relative squared error	78.0355 %	
Total Number of Instances	1750	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.746	0.179	0.782	0.746	0.764	0.863	A
	0.821	0.254	0.789	0.821	0.805	0.863	U
Weighted Avg.	0.786	0.219	0.786	0.786	0.786	0.863	

=== Confusion Matrix ===

a	b	<-- classified as
604	206	a = A
168	772	b = U

```
=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E
weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

Bayes Network Classifier
not using ADTree
#attributes=12 #classindex=11
Network structure (nodes followed by parents)
VARLIK(6): RISK DUZEYI (ANKET SONUCU)
TEHDIT(10): RISK DUZEYI (ANKET SONUCU)
ZAYIFLIK(9): RISK DUZEYI (ANKET SONUCU)
RISKIN GERCEKLESME OLASILIGI(5): RISK DUZEYI (ANKET SONUCU)
GIZLILIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
BUTUNLUK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
KULLANILABILIRLIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
SON 1 YILDAKI VAKA ADEDI(5): RISK DUZEYI (ANKET SONUCU)
BILGI BIRIKIMI DUZEYI(5): RISK DUZEYI (ANKET SONUCU)
KURUM ICI EGITIM ALDI MI(2): RISK DUZEYI (ANKET SONUCU)
KURUMDA POLITIKA VB MEVCUT MU(2): RISK DUZEYI (ANKET SONUCU)
RISK DUZEYI (ANKET SONUCU)(2):
LogScore Bayes: -28265.362446309206
LogScore BDeu: -28486.943491377962
LogScore MDL: -28453.89054100045
LogScore ENTROPY: -28091.72304425495
LogScore AIC: -28188.72304425495


Time taken to build model: 0 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1347               76.9714 %
Incorrectly Classified Instances       403               23.0286 %
Kappa statistic                          0.538
Mean absolute error                      0.2525
Root mean squared error                  0.422
Relative absolute error                 50.782  %
Root relative squared error             84.6394 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===
```

```
               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.769     0.23       0.743     0.769     0.756      0.851
A
                 0.77      0.231      0.795     0.77      0.782      0.851
U
Weighted Avg.    0.77      0.23       0.771     0.77      0.77       0.851

=== Confusion Matrix ===

   a   b   <-- classified as
 623 187 |   a = A
 216 724 |   b = U


=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E
weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Bayes Network Classifier
not using ADTree
#attributes=12 #classindex=11
Network structure (nodes followed by parents)
VARLIK(6): RISK DUZEYI (ANKET SONUCU)
TEHDIT(10): RISK DUZEYI (ANKET SONUCU)
ZAYIFLIK(9): RISK DUZEYI (ANKET SONUCU)
RISKIN GERCEKLESME OLASILIGI(5): RISK DUZEYI (ANKET SONUCU)
GIZLILIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
BUTUNLUK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
KULLANILABILIRLIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
SON 1 YILDAKI VAKA ADEDI(5): RISK DUZEYI (ANKET SONUCU)
BILGI BIRIKIMI DUZEYI(5): RISK DUZEYI (ANKET SONUCU)
KURUM ICI EGITIM ALDI MI(2): RISK DUZEYI (ANKET SONUCU)
KURUMDA POLITIKA VB MEVCUT MU(2): RISK DUZEYI (ANKET SONUCU)
RISK DUZEYI (ANKET SONUCU)(2):
LogScore Bayes: -28265.362446309206
LogScore BDeu: -28486.943491377962
LogScore MDL: -28453.89054100045
LogScore ENTROPY: -28091.72304425495
LogScore AIC: -28188.72304425495


Time taken to build model: 0 seconds
```

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1338               76.4571 %
Incorrectly Classified Instances       412               23.5429 %
Kappa statistic                          0.5277
Mean absolute error                      0.2575
Root mean squared error                  0.4273
Relative absolute error                 51.7876 %
Root relative squared error             85.6889 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.764     0.235      0.737     0.764     0.75       0.844
A
                 0.765     0.236      0.79      0.765     0.777      0.844
U
Weighted Avg.    0.765     0.235      0.765     0.765     0.765      0.844

=== Confusion Matrix ===

   a   b   <-- classified as
 619 191 |   a = A
 221 719 |   b = U



=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TAN -- -S ENTROPY -E
weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.1
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

Bayes Network Classifier
not using ADTree
#attributes=12 #classindex=11
Network structure (nodes followed by parents)
VARLIK(6): RISK DUZEYI (ANKET SONUCU) ZAYIFLIK
TEHDIT(10): RISK DUZEYI (ANKET SONUCU) SON 1 YILDAKI VAKA ADEDI
ZAYIFLIK(9): RISK DUZEYI (ANKET SONUCU) TEHDIT
RISKIN GERCEKLESME OLASILIGI(5): RISK DUZEYI (ANKET SONUCU) GIZLILIK KAYBI
GIZLILIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
BUTUNLUK KAYBI(5): RISK DUZEYI (ANKET SONUCU) GIZLILIK KAYBI
```

```
KULLANILABILIRLIK KAYBI(5): RISK DUZEYI (ANKET SONUCU) BUTUNLUK KAYBI
SON 1 YILDAKI VAKA ADEDI(5): RISK DUZEYI (ANKET SONUCU) RISKIN GERCEKLESME
OLASILIGI
BILGI BIRIKIMI DUZEYI(5): RISK DUZEYI (ANKET SONUCU) SON 1 YILDAKI VAKA ADEDI
KURUM ICI EGITIM ALDI MI(2): RISK DUZEYI (ANKET SONUCU) BILGI BIRIKIMI DUZEYI
KURUMDA POLITIKA VB MEVCUT MU(2): RISK DUZEYI (ANKET SONUCU) KURUM ICI EGITIM
ALDI MI
RISK DUZEYI (ANKET SONUCU)(2):
LogScore Bayes: -23638.12415835871
LogScore BDeu: -24870.66039976258
LogScore MDL: -24760.368414763852
LogScore ENTROPY: -22658.30345942655
LogScore AIC: -23221.30345942655


Time taken to build model: 0.03 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1410               80.5714 %
Incorrectly Classified Instances       340               19.4286 %
Kappa statistic                          0.6083
Mean absolute error                      0.2619
Root mean squared error                  0.3726
Relative absolute error                 52.6758 %
Root relative squared error             74.7288 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.773     0.166      0.801     0.773     0.786      0.883
A
                 0.834     0.227      0.81      0.834     0.822      0.883
U
Weighted Avg.    0.806     0.199      0.806     0.806     0.805      0.883

=== Confusion Matrix ===

   a   b   <-- classified as
 626 184 |   a = A
 156 784 |   b = U


=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.TAN -- -S ENTROPY -E
weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.1
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
```

```
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Bayes Network Classifier
not using ADTree
#attributes=12 #classindex=11
Network structure (nodes followed by parents)
VARLIK(6): RISK DUZEYI (ANKET SONUCU) ZAYIFLIK
TEHDIT(10): RISK DUZEYI (ANKET SONUCU) SON 1 YILDAKI VAKA ADEDI
ZAYIFLIK(9): RISK DUZEYI (ANKET SONUCU) TEHDIT
RISKIN GERCEKLESME OLASILIGI(5): RISK DUZEYI (ANKET SONUCU) GIZLILIK KAYBI
GIZLILIK KAYBI(5): RISK DUZEYI (ANKET SONUCU)
BUTUNLUK KAYBI(5): RISK DUZEYI (ANKET SONUCU) GIZLILIK KAYBI
KULLANILABILIRLIK KAYBI(5): RISK DUZEYI (ANKET SONUCU) BUTUNLUK KAYBI
SON 1 YILDAKI VAKA ADEDI(5): RISK DUZEYI (ANKET SONUCU) RISKIN GERCEKLESME
OLASILIGI
BILGI BIRIKIMI DUZEYI(5): RISK DUZEYI (ANKET SONUCU) SON 1 YILDAKI VAKA ADEDI
KURUM ICI EGITIM ALDI MI(2): RISK DUZEYI (ANKET SONUCU) BILGI BIRIKIMI DUZEYI
KURUMDA POLITIKA VB MEVCUT MU(2): RISK DUZEYI (ANKET SONUCU) KURUM ICI EGITIM
ALDI MI
RISK DUZEYI (ANKET SONUCU)(2):
LogScore Bayes: -23638.12415835871
LogScore BDeu: -24870.66039976258
LogScore MDL: -24760.368414763852
LogScore ENTROPY: -22658.30345942655
LogScore AIC: -23221.30345942655


Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1345               76.8571 %
Incorrectly Classified Instances       405               23.1429 %
Kappa statistic                          0.5333
Mean absolute error                      0.2921
Root mean squared error                  0.4088
Relative absolute error                 58.7344 %
Root relative squared error             81.99   %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.73      0.198      0.761     0.73      0.745      0.837
A
                 0.802     0.27       0.775     0.802     0.788      0.837
U
Weighted Avg.    0.769     0.237      0.768     0.769     0.768      0.837

=== Confusion Matrix ===

   a   b   <-- classified as
 591 219 |   a = A
 186 754 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.lazy.IBk -K 12 -W 0 -F -A
"weka.core.neighboursearch.CoverTree -A \"weka.core.EuclideanDistance -R
first-last\" -B 2.7"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

IB1 instance-based classifier
using 12 similarity-weighted nearest neighbour(s) for classification


Time taken to build model: 0.78 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1439               82.2286 %
Incorrectly Classified Instances       311               17.7714 %
Kappa statistic                          0.6423
Mean absolute error                      0.2797
Root mean squared error                  0.3519
Relative absolute error                 56.2436 %
Root relative squared error             70.5775 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.801     0.16       0.812     0.801     0.807      0.913
A
                 0.84      0.199      0.831     0.84      0.836      0.913
U
Weighted Avg.    0.822     0.181      0.822     0.822     0.822      0.913

=== Confusion Matrix ===

   a   b   <-- classified as
 649 161 |   a = A
 150 790 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.lazy.IBk -K 12 -W 0 -F -A
"weka.core.neighboursearch.CoverTree -A \"weka.core.EuclideanDistance -R
first-last\" -B 2.7"
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

IB1 instance-based classifier
using 12 similarity-weighted nearest neighbour(s) for classification


Time taken to build model: 0.8 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1357               77.5429 %
Incorrectly Classified Instances       393               22.4571 %
Kappa statistic                          0.5485
Mean absolute error                      0.3182
Root mean squared error                  0.3967
Relative absolute error                 64.0019 %
Root relative squared error             79.5544 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.759     0.211      0.756     0.759     0.758      0.85
A
                 0.789     0.241      0.792     0.789     0.791      0.85
U
Weighted Avg.    0.775     0.227      0.775     0.775     0.775      0.85

=== Confusion Matrix ===

   a   b   <-- classified as
 615 195 |   a = A
 198 742 |   b = U
```

```
=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.J48 -- -C 0.7 -M 2
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

All the base classifiers:

FilteredClassifier using weka.classifiers.trees.J48 -C 0.7 -M 2 on data
filtered through weka.filters.unsupervised.attribute.Remove -V -R
2,7,6,1,5,10,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data


Classifier Model
J48 pruned tree
------------------

BUTUNLUK KAYBI = 1
|   KULLANILABILIRLIK KAYBI = 1
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7: A (39.0/6.0)
|   |   |   TEHDIT = T4
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (17.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5: A (0.0)
|   |   |   |   |   VARLIK = V1: U (1.0)
|   |   |   |   |   VARLIK = V4: A (0.0)
|   |   |   |   |   VARLIK = V3: A (2.0)
|   |   |   |   |   VARLIK = V2: A (2.0/1.0)
|   |   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T9: A (12.0)
|   |   |   TEHDIT = T10: A (6.0)
```

```
|   |   |   TEHDIT = T2: A (11.0)
|   |   |   TEHDIT = T1: A (4.0)
|   |   |   TEHDIT = T3: A (15.0/1.0)
|   |   |   TEHDIT = T8: A (15.0)
|   |   |   TEHDIT = T5: A (4.0)
|   |   |   TEHDIT = T6: A (6.0)
|   |   GIZLILIK KAYBI = 2: A (9.0/2.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   TEHDIT = T7: A (0.0)
|   |   |   TEHDIT = T4: U (1.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (2.0)
|   |   |   TEHDIT = T1: U (3.0/1.0)
|   |   |   TEHDIT = T3: A (0.0)
|   |   |   TEHDIT = T8: U (1.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: A (1.0)
|   |   GIZLILIK KAYBI = 4: A (0.0)
|   |   GIZLILIK KAYBI = 5: A (2.0)
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5: A (5.0/1.0)
|   |   |   |   VARLIK = V1: A (0.0)
|   |   |   |   VARLIK = V4: A (3.0)
|   |   |   |   VARLIK = V3: U (1.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6: A (1.0)
|   |   |   TEHDIT = T4: A (2.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (4.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (7.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0)
|   |   |   TEHDIT = T6: A (0.0)
|   |   GIZLILIK KAYBI = 2
|   |   |   TEHDIT = T7: A (5.0/1.0)
|   |   |   TEHDIT = T4: A (1.0)
|   |   |   TEHDIT = T9: A (0.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: U (1.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (4.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (2.0/1.0)
|   |   |   TEHDIT = T6: A (3.0/1.0)
|   |   GIZLILIK KAYBI = 3: U (3.0/1.0)
|   |   GIZLILIK KAYBI = 4: A (2.0)
|   |   GIZLILIK KAYBI = 5: U (2.0)
|   KULLANILABILIRLIK KAYBI = 3
|   |   TEHDIT = T7: A (11.0)
|   |   TEHDIT = T4: U (1.0)
|   |   TEHDIT = T9: A (2.0/1.0)
|   |   TEHDIT = T10: A (2.0)
|   |   TEHDIT = T2
|   |   |   VARLIK = V5: A (0.0)
|   |   |   VARLIK = V1: A (2.0/1.0)
|   |   |   VARLIK = V4: A (0.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (2.0)
|   |   TEHDIT = T1: A (0.0)
```

```
|   |   TEHDIT = T3: A (0.0)
|   |   TEHDIT = T8: A (0.0)
|   |   TEHDIT = T5: A (0.0)
|   |   TEHDIT = T6: A (2.0/1.0)
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7: U (1.0)
|   |   TEHDIT = T4: A (1.0)
|   |   TEHDIT = T9: A (1.0)
|   |   TEHDIT = T10: U (0.0)
|   |   TEHDIT = T2: U (4.0)
|   |   TEHDIT = T1: U (0.0)
|   |   TEHDIT = T3: U (1.0)
|   |   TEHDIT = T8: U (1.0)
|   |   TEHDIT = T5: U (1.0)
|   |   TEHDIT = T6: A (3.0)
|   KULLANILABILIRLIK KAYBI = 5: U (2.0)
BUTUNLUK KAYBI = 2
|   KULLANILABILIRLIK KAYBI = 1: A (36.0/7.0)
|   KULLANILABILIRLIK KAYBI = 2
|   |   GIZLILIK KAYBI = 1: A (36.0/2.0)
|   |   GIZLILIK KAYBI = 2: A (80.0/8.0)
|   |   GIZLILIK KAYBI = 3: A (30.0/5.0)
|   |   GIZLILIK KAYBI = 4: U (3.0/1.0)
|   |   GIZLILIK KAYBI = 5: A (3.0)
|   KULLANILABILIRLIK KAYBI = 3
|   |   VARLIK = V5
|   |   |   KURUM ICI EGITIM ALDI MI = H: A (3.0/1.0)
|   |   |   KURUM ICI EGITIM ALDI MI = E: U (3.0)
|   |   VARLIK = V1: A (10.0/1.0)
|   |   VARLIK = V4
|   |   |   GIZLILIK KAYBI = 1: A (1.0)
|   |   |   GIZLILIK KAYBI = 2
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (2.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E: U (3.0/1.0)
|   |   |   GIZLILIK KAYBI = 3: U (6.0/2.0)
|   |   |   GIZLILIK KAYBI = 4: A (0.0)
|   |   |   GIZLILIK KAYBI = 5: A (1.0)
|   |   VARLIK = V3
|   |   |   TEHDIT = T7: A (1.0)
|   |   |   TEHDIT = T4: A (5.0/1.0)
|   |   |   TEHDIT = T9: A (0.0)
|   |   |   TEHDIT = T10: A (0.0)
|   |   |   TEHDIT = T2: A (3.0)
|   |   |   TEHDIT = T1: A (0.0)
|   |   |   TEHDIT = T3: A (2.0)
|   |   |   TEHDIT = T8: A (1.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: U (1.0)
|   |   VARLIK = V2: A (4.0)
|   |   VARLIK = V6: A (2.0/1.0)
|   KULLANILABILIRLIK KAYBI = 4
|   |   TEHDIT = T7
|   |   |   VARLIK = V5: A (2.0/1.0)
|   |   |   VARLIK = V1: A (0.0)
|   |   |   VARLIK = V4: A (1.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (5.0)
|   |   TEHDIT = T4: A (2.0/1.0)
|   |   TEHDIT = T9: A (1.0)
|   |   TEHDIT = T10: U (0.0)
|   |   TEHDIT = T2: U (5.0)
|   |   TEHDIT = T1: A (2.0/1.0)
|   |   TEHDIT = T3: A (2.0)
|   |   TEHDIT = T8: U (7.0/1.0)
```

```
|   |   TEHDIT = T5: A (2.0/1.0)
|   |   TEHDIT = T6: U (5.0/2.0)
|   KULLANILABILIRLIK KAYBI = 5: A (7.0)
BUTUNLUK KAYBI = 3
|   GIZLILIK KAYBI = 1
|   |   KULLANILABILIRLIK KAYBI = 1: A (4.0)
|   |   KULLANILABILIRLIK KAYBI = 2
|   |   |   TEHDIT = T7
|   |   |   |   VARLIK = V5: A (2.0)
|   |   |   |   VARLIK = V1: A (0.0)
|   |   |   |   VARLIK = V4: A (3.0)
|   |   |   |   VARLIK = V3: A (0.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6: U (1.0)
|   |   |   TEHDIT = T4: A (2.0)
|   |   |   TEHDIT = T9: A (2.0)
|   |   |   TEHDIT = T10: A (1.0)
|   |   |   TEHDIT = T2: A (1.0)
|   |   |   TEHDIT = T1: A (1.0)
|   |   |   TEHDIT = T3: A (1.0)
|   |   |   TEHDIT = T8: A (0.0)
|   |   |   TEHDIT = T5: A (0.0)
|   |   |   TEHDIT = T6: U (1.0)
|   |   KULLANILABILIRLIK KAYBI = 3: A (5.0)
|   |   KULLANILABILIRLIK KAYBI = 4: A (2.0)
|   |   KULLANILABILIRLIK KAYBI = 5: U (1.0)
|   GIZLILIK KAYBI = 2
|   |   KULLANILABILIRLIK KAYBI = 1: U (5.0/1.0)
|   |   KULLANILABILIRLIK KAYBI = 2: A (15.0/2.0)
|   |   KULLANILABILIRLIK KAYBI = 3
|   |   |   TEHDIT = T7: A (7.0/3.0)
|   |   |   TEHDIT = T4: A (5.0/1.0)
|   |   |   TEHDIT = T9: A (1.0)
|   |   |   TEHDIT = T10: U (1.0)
|   |   |   TEHDIT = T2: A (2.0)
|   |   |   TEHDIT = T1: A (4.0)
|   |   |   TEHDIT = T3: A (6.0)
|   |   |   TEHDIT = T8: A (2.0/1.0)
|   |   |   TEHDIT = T5: A (2.0/1.0)
|   |   |   TEHDIT = T6: U (3.0/1.0)
|   |   KULLANILABILIRLIK KAYBI = 4: U (3.0)
|   |   KULLANILABILIRLIK KAYBI = 5: A (1.0)
|   GIZLILIK KAYBI = 3
|   |   KULLANILABILIRLIK KAYBI = 1: A (8.0)
|   |   KULLANILABILIRLIK KAYBI = 2
|   |   |   VARLIK = V5: A (4.0)
|   |   |   VARLIK = V1: A (3.0/1.0)
|   |   |   VARLIK = V4: U (3.0/1.0)
|   |   |   VARLIK = V3: A (6.0/1.0)
|   |   |   VARLIK = V2: A (2.0)
|   |   |   VARLIK = V6: A (7.0/2.0)
|   |   KULLANILABILIRLIK KAYBI = 3
|   |   |   TEHDIT = T7: A (45.0/17.0)
|   |   |   TEHDIT = T4
|   |   |   |   VARLIK = V5: A (0.0)
|   |   |   |   VARLIK = V1: U (1.0)
|   |   |   |   VARLIK = V4: A (0.0)
|   |   |   |   VARLIK = V3: A (1.0)
|   |   |   |   VARLIK = V2: A (2.0/1.0)
|   |   |   |   VARLIK = V6: A (3.0/1.0)
|   |   |   TEHDIT = T9: A (4.0/1.0)
|   |   |   TEHDIT = T10: U (2.0)
|   |   |   TEHDIT = T2
|   |   |   |   VARLIK = V5: A (0.0)
|   |   |   |   VARLIK = V1: U (4.0)
```

```
|   |   |   |   VARLIK = V4: A (0.0)
|   |   |   |   VARLIK = V3: A (1.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6: A (7.0/2.0)
|   |   |   TEHDIT = T1: A (6.0/2.0)
|   |   |   TEHDIT = T3
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (6.0/2.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   |   VARLIK = V5: U (3.0/1.0)
|   |   |   |   |   VARLIK = V1: U (3.0/1.0)
|   |   |   |   |   VARLIK = V4: U (0.0)
|   |   |   |   |   VARLIK = V3: A (1.0)
|   |   |   |   |   VARLIK = V2: U (0.0)
|   |   |   |   |   VARLIK = V6: U (0.0)
|   |   |   TEHDIT = T8: A (11.0/4.0)
|   |   |   TEHDIT = T5: A (6.0/3.0)
|   |   |   TEHDIT = T6: A (7.0/2.0)
|   |   KULLANILABILIRLIK KAYBI = 4
|   |   |   VARLIK = V5: A (4.0/2.0)
|   |   |   VARLIK = V1: U (4.0/1.0)
|   |   |   VARLIK = V4: A (6.0/3.0)
|   |   |   VARLIK = V3: U (7.0/2.0)
|   |   |   VARLIK = V2: A (4.0)
|   |   |   VARLIK = V6
|   |   |   |   TEHDIT = T7: U (4.0/1.0)
|   |   |   |   TEHDIT = T4: A (3.0/1.0)
|   |   |   |   TEHDIT = T9: U (0.0)
|   |   |   |   TEHDIT = T10: U (0.0)
|   |   |   |   TEHDIT = T2: U (1.0)
|   |   |   |   TEHDIT = T1: U (0.0)
|   |   |   |   TEHDIT = T3: U (0.0)
|   |   |   |   TEHDIT = T8: U (0.0)
|   |   |   |   TEHDIT = T5: U (1.0)
|   |   |   |   TEHDIT = T6: U (0.0)
|   |   KULLANILABILIRLIK KAYBI = 5: U (2.0)
|   GIZLILIK KAYBI = 4
|   |   TEHDIT = T7
|   |   |   KURUM ICI EGITIM ALDI MI = H
|   |   |   |   VARLIK = V5: A (5.0)
|   |   |   |   VARLIK = V1: A (0.0)
|   |   |   |   VARLIK = V4: U (3.0/1.0)
|   |   |   |   VARLIK = V3: A (1.0)
|   |   |   |   VARLIK = V2: A (0.0)
|   |   |   |   VARLIK = V6
|   |   |   |   |   KULLANILABILIRLIK KAYBI = 1: A (0.0)
|   |   |   |   |   KULLANILABILIRLIK KAYBI = 2: A (2.0)
|   |   |   |   |   KULLANILABILIRLIK KAYBI = 3: U (1.0)
|   |   |   |   |   KULLANILABILIRLIK KAYBI = 4: A (2.0/1.0)
|   |   |   |   |   KULLANILABILIRLIK KAYBI = 5: A (0.0)
|   |   |   KURUM ICI EGITIM ALDI MI = E
|   |   |   |   KULLANILABILIRLIK KAYBI = 1: U (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 2: U (3.0/1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 3: A (6.0/2.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 4: U (2.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 5: U (1.0)
|   |   TEHDIT = T4: A (11.0/3.0)
|   |   TEHDIT = T9
|   |   |   KURUM ICI EGITIM ALDI MI = H: U (3.0)
|   |   |   KURUM ICI EGITIM ALDI MI = E: A (2.0)
|   |   TEHDIT = T10: A (0.0)
|   |   TEHDIT = T2: U (7.0/3.0)
|   |   TEHDIT = T1: U (6.0/2.0)
|   |   TEHDIT = T3
|   |   |   VARLIK = V5: A (2.0/1.0)
|   |   |   VARLIK = V1: A (5.0/1.0)
```

```
|   |   |   VARLIK = V4: A (0.0)
|   |   |   VARLIK = V3: U (1.0)
|   |   |   VARLIK = V2: A (0.0)
|   |   |   VARLIK = V6: A (0.0)
|   |   TEHDIT = T8: U (3.0/1.0)
|   |   TEHDIT = T5: U (6.0/2.0)
|   |   TEHDIT = T6: U (5.0/2.0)
|   GIZLILIK KAYBI = 5
|   |   VARLIK = V5: A (1.0)
|   |   VARLIK = V1: U (3.0)
|   |   VARLIK = V4: U (5.0/1.0)
|   |   VARLIK = V3
|   |   |   TEHDIT = T7: A (2.0/1.0)
|   |   |   TEHDIT = T4: U (1.0)
|   |   |   TEHDIT = T9: U (0.0)
|   |   |   TEHDIT = T10: U (0.0)
|   |   |   TEHDIT = T2: U (3.0)
|   |   |   TEHDIT = T1: U (0.0)
|   |   |   TEHDIT = T3: U (1.0)
|   |   |   TEHDIT = T8: A (3.0/1.0)
|   |   |   TEHDIT = T5: U (0.0)
|   |   |   TEHDIT = T6
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (2.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E: U (2.0)
|   |   VARLIK = V2: U (3.0)
|   |   VARLIK = V6: U (3.0)
BUTUNLUK KAYBI = 4
|   VARLIK = V5: U (41.0/3.0)
|   VARLIK = V1
|   |   GIZLILIK KAYBI = 1: U (2.0)
|   |   GIZLILIK KAYBI = 2
|   |   |   KULLANILABILIRLIK KAYBI = 1: U (0.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: U (4.0/1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (0.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4: A (3.0/1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: U (0.0)
|   |   GIZLILIK KAYBI = 3: U (6.0)
|   |   GIZLILIK KAYBI = 4
|   |   |   TEHDIT = T7: U (0.0)
|   |   |   TEHDIT = T4: U (9.0)
|   |   |   TEHDIT = T9: U (0.0)
|   |   |   TEHDIT = T10: U (0.0)
|   |   |   TEHDIT = T2
|   |   |   |   KULLANILABILIRLIK KAYBI = 1: A (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 2: A (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 3: A (2.0/1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 4: A (5.0/2.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 5: U (1.0)
|   |   |   TEHDIT = T1: A (6.0/2.0)
|   |   |   TEHDIT = T3
|   |   |   |   KURUM ICI EGITIM ALDI MI = H: A (7.0/3.0)
|   |   |   |   KURUM ICI EGITIM ALDI MI = E: U (4.0)
|   |   |   TEHDIT = T8: U (0.0)
|   |   |   TEHDIT = T5: U (0.0)
|   |   |   TEHDIT = T6: U (0.0)
|   |   GIZLILIK KAYBI = 5: U (21.0/2.0)
|   VARLIK = V4
|   |   GIZLILIK KAYBI = 1: U (3.0/1.0)
|   |   GIZLILIK KAYBI = 2: A (4.0/2.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   KULLANILABILIRLIK KAYBI = 1: A (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: U (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (5.0/1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4: U (7.0/1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: U (0.0)
```

```
|   |   GIZLILIK KAYBI = 4: U (58.0/10.0)
|   |   GIZLILIK KAYBI = 5
|   |   |   KULLANILABILIRLIK KAYBI = 1: U (2.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: U (2.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4
|   |   |   |   TEHDIT = T7: A (4.0/1.0)
|   |   |   |   TEHDIT = T4: A (0.0)
|   |   |   |   TEHDIT = T9: U (1.0)
|   |   |   |   TEHDIT = T10: A (2.0/1.0)
|   |   |   |   TEHDIT = T2: A (0.0)
|   |   |   |   TEHDIT = T1: A (0.0)
|   |   |   |   TEHDIT = T3: A (0.0)
|   |   |   |   TEHDIT = T8: A (0.0)
|   |   |   |   TEHDIT = T5: A (0.0)
|   |   |   |   TEHDIT = T6: A (0.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: U (5.0/1.0)
|   VARLIK = V3
|   |   KULLANILABILIRLIK KAYBI = 1: A (2.0/1.0)
|   |   KULLANILABILIRLIK KAYBI = 2: A (6.0/2.0)
|   |   KULLANILABILIRLIK KAYBI = 3
|   |   |   TEHDIT = T7: U (7.0/2.0)
|   |   |   TEHDIT = T4: A (6.0/1.0)
|   |   |   TEHDIT = T9: U (0.0)
|   |   |   TEHDIT = T10: U (0.0)
|   |   |   TEHDIT = T2: U (0.0)
|   |   |   TEHDIT = T1: U (0.0)
|   |   |   TEHDIT = T3: U (2.0)
|   |   |   TEHDIT = T8: A (4.0/2.0)
|   |   |   TEHDIT = T5: U (0.0)
|   |   |   TEHDIT = T6: U (2.0)
|   |   KULLANILABILIRLIK KAYBI = 4
|   |   |   GIZLILIK KAYBI = 1: U (5.0/2.0)
|   |   |   GIZLILIK KAYBI = 2
|   |   |   |   TEHDIT = T7: U (2.0)
|   |   |   |   TEHDIT = T4: U (2.0)
|   |   |   |   TEHDIT = T9: U (0.0)
|   |   |   |   TEHDIT = T10: U (0.0)
|   |   |   |   TEHDIT = T2: U (0.0)
|   |   |   |   TEHDIT = T1: U (0.0)
|   |   |   |   TEHDIT = T3: U (0.0)
|   |   |   |   TEHDIT = T8: U (0.0)
|   |   |   |   TEHDIT = T5: U (0.0)
|   |   |   |   TEHDIT = T6: A (2.0)
|   |   |   GIZLILIK KAYBI = 3: U (6.0)
|   |   |   GIZLILIK KAYBI = 4
|   |   |   |   TEHDIT = T7: A (3.0/1.0)
|   |   |   |   TEHDIT = T4: U (4.0/1.0)
|   |   |   |   TEHDIT = T9: U (0.0)
|   |   |   |   TEHDIT = T10: U (0.0)
|   |   |   |   TEHDIT = T2: U (7.0)
|   |   |   |   TEHDIT = T1: U (0.0)
|   |   |   |   TEHDIT = T3: U (4.0)
|   |   |   |   TEHDIT = T8: U (3.0)
|   |   |   |   TEHDIT = T5: U (0.0)
|   |   |   |   TEHDIT = T6: U (5.0/1.0)
|   |   |   GIZLILIK KAYBI = 5: U (5.0/2.0)
|   |   KULLANILABILIRLIK KAYBI = 5
|   |   |   TEHDIT = T7: U (2.0)
|   |   |   TEHDIT = T4: U (0.0)
|   |   |   TEHDIT = T9: U (0.0)
|   |   |   TEHDIT = T10: U (0.0)
|   |   |   TEHDIT = T2: A (2.0/1.0)
|   |   |   TEHDIT = T1: U (0.0)
|   |   |   TEHDIT = T3: A (2.0)
```

```
|   |   |   TEHDIT = T8: U (3.0/1.0)
|   |   |   TEHDIT = T5: U (0.0)
|   |   |   TEHDIT = T6: U (5.0/2.0)
|   VARLIK = V2
|   |   GIZLILIK KAYBI = 1: A (2.0)
|   |   GIZLILIK KAYBI = 2: U (6.0/1.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   KULLANILABILIRLIK KAYBI = 1: U (0.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: U (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (7.0/2.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4: U (7.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: A (1.0)
|   |   GIZLILIK KAYBI = 4
|   |   |   TEHDIT = T7: U (0.0)
|   |   |   TEHDIT = T4: U (12.0/1.0)
|   |   |   TEHDIT = T9: U (0.0)
|   |   |   TEHDIT = T10: U (0.0)
|   |   |   TEHDIT = T2: U (0.0)
|   |   |   TEHDIT = T1: U (0.0)
|   |   |   TEHDIT = T3: U (0.0)
|   |   |   TEHDIT = T8
|   |   |   |   KULLANILABILIRLIK KAYBI = 1: A (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 2: A (1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 3: A (2.0/1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 4: A (4.0/1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 5: U (1.0)
|   |   |   TEHDIT = T5: U (13.0/4.0)
|   |   |   TEHDIT = T6
|   |   |   |   KULLANILABILIRLIK KAYBI = 1: A (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 2: A (0.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 3: A (4.0/1.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 4: U (8.0/3.0)
|   |   |   |   KULLANILABILIRLIK KAYBI = 5: A (1.0)
|   |   GIZLILIK KAYBI = 5: U (10.0/2.0)
|   VARLIK = V6
|   |   GIZLILIK KAYBI = 1: A (1.0)
|   |   GIZLILIK KAYBI = 2: A (2.0/1.0)
|   |   GIZLILIK KAYBI = 3
|   |   |   KULLANILABILIRLIK KAYBI = 1: A (0.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: A (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (6.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4: A (8.0/1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: A (0.0)
|   |   GIZLILIK KAYBI = 4
|   |   |   KULLANILABILIRLIK KAYBI = 1: A (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 2: U (1.0)
|   |   |   KULLANILABILIRLIK KAYBI = 3: U (8.0/3.0)
|   |   |   KULLANILABILIRLIK KAYBI = 4: U (50.0/12.0)
|   |   |   KULLANILABILIRLIK KAYBI = 5: U (9.0/1.0)
|   |   GIZLILIK KAYBI = 5
|   |   |   KURUM ICI EGITIM ALDI MI = H
|   |   |   |   TEHDIT = T7: U (9.0/2.0)
|   |   |   |   TEHDIT = T4: A (4.0)
|   |   |   |   TEHDIT = T9: U (0.0)
|   |   |   |   TEHDIT = T10: U (0.0)
|   |   |   |   TEHDIT = T2: U (3.0/1.0)
|   |   |   |   TEHDIT = T1: U (0.0)
|   |   |   |   TEHDIT = T3: U (0.0)
|   |   |   |   TEHDIT = T8: U (0.0)
|   |   |   |   TEHDIT = T5: U (2.0)
|   |   |   |   TEHDIT = T6: U (0.0)
|   |   |   KURUM ICI EGITIM ALDI MI = E: U (10.0)
BUTUNLUK KAYBI = 5
|   GIZLILIK KAYBI = 1: U (2.0)
|   GIZLILIK KAYBI = 2: U (6.0/2.0)
```

```
|   GIZLILIK KAYBI = 3: U (14.0/2.0)
|   GIZLILIK KAYBI = 4
|   |   KULLANILABILIRLIK KAYBI = 1: A (3.0)
|   |   KULLANILABILIRLIK KAYBI = 2: U (2.0)
|   |   KULLANILABILIRLIK KAYBI = 3: U (3.0/1.0)
|   |   KULLANILABILIRLIK KAYBI = 4
|   |   |   VARLIK = V5: U (1.0)
|   |   |   VARLIK = V1: U (2.0)
|   |   |   VARLIK = V4: U (2.0)
|   |   |   VARLIK = V3: U (3.0/1.0)
|   |   |   VARLIK = V2: A (1.0)
|   |   |   VARLIK = V6: U (4.0)
|   |   KULLANILABILIRLIK KAYBI = 5
|   |   |   VARLIK = V5: U (3.0)
|   |   |   VARLIK = V1: A (4.0/1.0)
|   |   |   VARLIK = V4: U (7.0/1.0)
|   |   |   VARLIK = V3: U (2.0)
|   |   |   VARLIK = V2: A (2.0/1.0)
|   |   |   VARLIK = V6: U (9.0/1.0)
|   GIZLILIK KAYBI = 5: U (291.0/18.0)
```

Number of Leaves : 413

Size of the tree : 491

FilteredClassifier using weka.classifiers.trees.J48 -C 0.7 -M 2 on data filtered through weka.filters.unsupervised.attribute.Remove -V -R 6,10,5,4,3,1,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R6,10,5,4,3,1,12'

```
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'RISKIN GERCEKLESME OLASILIGI' {1,2,3,4,5}
@attribute ZAYIFLIK {Z9,Z2,Z8,Z1,Z5,Z3,Z7,Z4,Z6}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data
```

Classifier Model
J48 pruned tree

```
Time taken to build model: 104.17 seconds

=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances        1642               93.8286 %
Incorrectly Classified Instances       108                6.1714 %
Kappa statistic                          0.8755
Mean absolute error                      0.216
Root mean squared error                  0.2626
Relative absolute error                 43.4324 %
Root relative squared error             52.6714 %
Total Number of Instances             1750

=== Detailed Accuracy By Class ===
```

```
               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area
Class
                 0.912     0.039      0.952     0.912     0.932      0.986
A
                 0.961     0.088      0.927     0.961     0.944      0.986
U
Weighted Avg.    0.938     0.065      0.939     0.938     0.938      0.986

=== Confusion Matrix ===

   a   b   <-- classified as
 739  71 |   a = A
  37 903 |   b = U


=== Run information ===

Scheme:       weka.classifiers.meta.RandomSubSpace -P 0.5 -S 1 -I 180 -W
weka.classifiers.trees.J48 -- -C 0.7 -M 2
Relation:     ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    1750
Attributes:   12
              VARLIK
              TEHDIT
              ZAYIFLIK
              RISKIN GERCEKLESME OLASILIGI
              GIZLILIK KAYBI
              BUTUNLUK KAYBI
              KULLANILABILIRLIK KAYBI
              SON 1 YILDAKI VAKA ADEDI
              BILGI BIRIKIMI DUZEYI
              KURUM ICI EGITIM ALDI MI
              KURUMDA POLITIKA VB MEVCUT MU
              RISK DUZEYI (ANKET SONUCU)
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

All the base classifiers:

FilteredClassifier using weka.classifiers.trees.J48 -C 0.7 -M 2 on data
filtered through weka.filters.unsupervised.attribute.Remove -V -R
2,7,6,1,5,10,12

Filtered Header
@relation 'ANALIZ2-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-V-R2,7,6,1,5,10,12'

@attribute TEHDIT {T7,T4,T9,T10,T2,T1,T3,T8,T5,T6}
@attribute 'KULLANILABILIRLIK KAYBI' {1,2,3,4,5}
@attribute 'BUTUNLUK KAYBI' {1,2,3,4,5}
@attribute VARLIK {V5,V1,V4,V3,V2,V6}
@attribute 'GIZLILIK KAYBI' {1,2,3,4,5}
@attribute 'KURUM ICI EGITIM ALDI MI' {H,E}
@attribute 'RISK DUZEYI (ANKET SONUCU)' {A,U}

@data


Classifier Model
J48 pruned tree
------------------
```

Number of Leaves : 703

Size of the tree : 826

Time taken to build model: 103.63 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances	1374	78.5143 %
Incorrectly Classified Instances	376	21.4857 %
Kappa statistic	0.5666	
Mean absolute error	0.3222	
Root mean squared error	0.3892	
Relative absolute error	64.7994 %	
Root relative squared error	78.0551 %	
Total Number of Instances	1750	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.746	0.181	0.78	0.746	0.763	0.863	A
	0.819	0.254	0.789	0.819	0.804	0.863	U
Weighted Avg.	0.785	0.22	0.785	0.785	0.785	0.863	

=== Confusion Matrix ===

a	b	<-- classified as
604	206	a = A
170	770	b = U