

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İKTİSAT ANABİLİM DALI
ULUSLARARASI TİCARET VE PARA YÖNETİMİ BİLİM DALI

**İÇ KONTROL SİSTEMİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİNİN
KULLANILMASININ İŞLETMENİN BİLANÇOSUNA ETKİSİ**

Yüksek Lisans Tezi

GÖKÇE ÖZKER DURSUN

İstanbul, 2016

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İKTİSAT ANABİLİM DALI
ULUSLARARASI TİCARET VE PARA YÖNETİMİ BİLİM DALI

**İÇ KONTROL SİSTEMİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİNİN
KULLANILMASININ İŞLETMENİN BİLANÇOSUNA ETKİSİ**

Yüksek Lisans Tezi

GÖKÇE ÖZKER DURSUN

Danışman: Yrd.Doç.Dr. ERDEM TURGAN

İstanbul, 2016



T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ

TEZ ONAY BELGESİ

İKTİSAT Anabilim Dalı ULUSLARARASI TİCARET VE PARA YÖNETİMİ Bilim Dalı TEZLİ YÜKSEK LİSANS öğrencisi GÖKÇE ÖZKER DURSUN'nın İÇ KONTROL SİSTEMİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİNİN KULLANILMASININ İŞLETMENİN BİLANÇOSUNA ETKİSİ adlı tez çalışması, Enstitümüz Yönetim Kurulunun 13.06.2016 tarih ve 2016-21/24 sayılı kararıyla oluşturulan jüri tarafından oy birliği / ~~oy çokluğu~~ ile Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi ..13../07../2016.....

Öğretim Üyesi Adı Soyadı

İmzası

Öğretim Üyesi Adı Soyadı	İmzası
1. Tez Danışmanı Yrd. Doç. Dr. ERDEM TURGAN	
2. Jüri Üyesi Prof. Dr. NURDAN ASLAN	
3. Jüri Üyesi Prof. Dr. SİNAN ASLAN	

ÖZET

Bu çalışma bir işletmenin elindeki karmaşık veri yığınlarının veri madenciliği modellemesi kullanılarak segmentasyon bilgisine dönüştürülmesi sonucu elde edilen bilginin iç kontrol sistemi açısından riskleri ortaya koymak ve bilanço hesapları üzerinde gerçekleşecek etkilerini araştırmak amacıyla yapılmıştır.

Araştırma yöntemi olarak Kümeleme modeli ve k-means algoritması kullanılmıştır. Araştırma verileri kalibrasyon ve danışmanlık alanında faaliyet gösteren bir firmanın 2013, 2014 ve 2015 yıllarına ait ayrıntılı müşteri ve satış verileridir. Araştırma sonucunda %68 ve %32 oranında ayrışan iki küme elde edilmiştir. Kümelerin segmentasyon özellikleri incelendiğinde her iki kümede de müşterilerin kobi nitelikli, eski müşterilerden meydana geldiği ve İstanbul Anadolu Yakası'nda faaliyet gösterdikleri tespit edilmiştir. Bu durum iç kontrol açısından yoğunlaşma ve iş risklerini ortaya çıkarmıştır. 2014 ve 2015 senesi karşılaştırıldığında firmanın müşteri sayısı artmasına rağmen cirosunda artış olmadığı görülmektedir. Firmanın yeni bir satış stratejisi geliştirmesi gerektiği belirlenmiştir. Satışların artması ve iç kontrol aktivitelerinin uygun şekilde yürütülmesi bilanço hesaplarında likiditeyi arttırarak net işletme sermayesini güçlendirir.

Anahtar kelimeler: İç kontrol, Veri Madenciliği, Kümeleme Modeli, Bilanço

ABSTRACT

This study has been analyzed the risk factor involved in final data which has been collected from complicated data stack of the entity's sources through applying data mining methodology resulting in segmented data and it has been also discussed the possible effects of risk factor over balance sheet components in terms of internal control system.

Clustering analysis and k-means clustering algorithm have been used as research methods in this study. Research data has sampled from detailed customer and sales stack data of an entity operating in calibration and consultation sector in 2013, 2014 and 2015. Following the search, two different clusters have identified with rates of %68 and %32 respectfully. When segmentation features of clusters analyzed it has been qualified that both customer clusters have had the characteristic of small and medium sized enterprises and geographically located and operated in İstanbul Anatolian side. In terms of internal controlling, this situation has revealed the possible intensity and employment risk. The comparison between 2014 and 2015 period has been revealed that even though the customer numbers have been increased to certain level of point, there has been no increase in entity's turnover. It has been identified that the entity has needed to demonstrate new effective sale strategy. Thus, the entity can increase liquidity in its balance sheet and strengthen its net working capital by increasing sales and promoting internal controlling activities properly.

Key Words: Internal Control, Data mining, Clustering model, Balance Sheet

İÇİNDEKİLER

	Sayfa No.
ÖZ.....	i
ABSTRACT.....	ii
İÇİNDEKİLER	iii
TABLO LİSTESİ.....	vii
ŞEKİL LİSTESİ.....	viii
KISALTMALAR	ix
1. GİRİŞ	1
2. İŞLETMELERDE İÇ KONTROL SİSTEMİ VE ÖNEMİ	3
2.1. İÇ KONTROL TANIMI VE KAPSAMI.....	3
2.1.1. İç Kontrol Kavramının Tanımı.....	3
2.1.2. İç Kontrolün Kapsamı	3
2.2. İÇ KONTROL SİSTEMİNİN AMAÇLARI	4
2.3. COSO (COMMITTEE OF SPONSORING ORGANIZATIONS OF THE TREADWAY COMMISSION) İÇ KONTROL MODELİ YATAY VE DİKEY BİLEŞENLERİ	4
2.3.1. COSO Süreci.....	4
2.3.2. COSO İç Kontrol Entegre Çerçeve Raporu	5
2.3.3. COSO'ya Göre İç Kontrol Kavramı	5
2.3.4. COSO Modeli İç Kontrol Bileşenleri.....	5
2.4. ETKİN BİR İÇ KONTROL SİSTEMİ İÇİN TEMEL İLKELER.....	7
2.5. İÇ KONTROL VE RİSK YÖNETİMİ	8

3. VERİ MADENCİLİĞİ VE KULLANIM ALANLARI.....	10
3.1. BÜYÜK VERİ KAVRAMI.....	10
3.2. BÜYÜK VERİ VE GELENEKSEL YÖNETİM YAKLAŞIMLARI	11
3.3. VERİ MADENCİLİĞİNİN TANIMI	12
3.4. VERİ MADENCİLİĞİNİN TARİHSEL SÜREÇTEKİ GELİŞİMİ	13
3.5. BİLGİ İŞLEME SİSTEMLERİNE İLİŞKİN TEMEL KAVRAMLAR	15
3.5.1. Verinin Tanımı	15
3.5.2. Veri Tabanı Sistemi.....	15
3.5.3. Veri Ambarı Sistemi	16
3.6. VERİ MADENCİLİĞİ SÜRECİ	17
3.6.1. Verilerin Temizlenmesi.....	18
3.6.2. Verilerin Birleştirilmesi	18
3.6.3. Seçim.....	18
3.6.4. Modelin Kurulması ve Değerlendirilmesi.....	19
3.7. VERİ MADENCİLİĞİNDE KULLANILAN TEKNİKLER.....	19
3.7.1. Yapay Sinir Ağları (YSA).....	19
3.7.2. Genetik Algoritmalar	22
3.7.3 İstatistik Teknikler	22
3.7.4. Karar Ağaçları	22
3.7.5. Bulanık Kümeler Yaklaşımı.....	24
3.8. VERİ MADENCİLİĞİ MODELLEME ÇEŞİTLERİ	25
3.8.1. Sınıflama Modeli.....	25
3.8.2. Lojistik Regresyon Modeli.....	26
3.8.3. Kümeleme Modeli.....	28

3.8.4. Değer Tahmin Modeli	30
3.8.5. Veri Görselleştirme	31
3.8.6. Birliktelik Kuralları	31
3.9. VERİ MADENCİLİĞİNDE KARŞILAŞILAN PROBLEMLER.....	33
4. VERİ MADENCİLİĞİNDE KÜMELEME MODELİ KULLANILARAK MÜŞTERİ SEGMENTASYONU BELİRLENMESİNE DAİR BİR UYGULAMA ÖRNEĞİ.....	35
4.1. ÇALIŞMANIN AMACI	35
4.2. SEÇİLEN MODEL VE ALGORİTMA	35
4.2.1. Kümeleme Modeli.....	35
4.2.2. K-means Algoritması	37
4.3. VERİNİN HAZIRLANMASI, TEMİZLENMESİ VE UYUMLAŞTIRILMASI...	39
4.4. MODELLEME	40
4.5. RAPOR SONUÇLARININ DEĞERLENDİRİLMESİ.....	42
Şekil 9 ve 10’da WEKA yazılımının eldeki verilere göre iki ekran görüntüsü verilmiştir.	42
5. İÇ KONTROL SİSTEMİ ÇERÇEVESİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİ KULLANILMASI VE BİLANÇO ÜZERİNDEKİ ETKİLERİ .	44
5.1. İÇ KONTROL SİSTEMİ ÇERÇEVESİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİ KULLANIMININ ÖNEMİ.....	44
5.2. UYGULAMA ÖRNEĞİ SEGMENTASYON ANALİZİ SONUÇLARININ İÇ KONTROL ÇERÇEVESİNDE DEĞERLENDİRİLMESİ	45
5.3. ETİK STANDARTLAR	49
5.3.1. Etik Kavramının Tanımlanması	49
5.3.2. Satış Stratejileri Etik Unsurları	49
5.4. BİLANÇO KAVRAMI.....	51
5.5. UYGULAMA ÖRNEĞİ SEGMENTASYON ANALİZİ SONUÇLARININ BİLANÇO ÜZERİNE ETKİLERİ.....	52

8. SONUÇ	54
KAYNAKÇA.....	57
EKLER	62
EK-1: SEGMENTASYON ANALİZİ İÇİN DÖNÜŞTÜRME EXCELİ.....	62

TABLO LİSTESİ

	Sayfa No.
Tablo 1: Büyük Veri ve Geleneksel Analitiğin Karşılaştırılması	11
Tablo 2: Veri Madenciliğinin Tarihsel Süreçteki Gelişimi.....	14
Tablo 3: Yıllara Göre Firmanın Ciro Müşteri Sayısı Yıllık Tüfe Oranı	48
Tablo 4: Bilanço Denkliği Tablosu.....	51

ŞEKİL LİSTESİ

	Sayfa No.
Şekil 1: Coso İç Kontrol Sürecini Oluşturan Unsurlar.....	6
Şekil 2: Veri Madenciliği Süreçleri.....	17
Şekil 3: Yapay sinir ağları.....	20
Şekil 4: Karar ağaçları.....	23
Şekil 5: Kümeleme modeli.....	28
Şekil 6: Kümeleme algoritması.....	36
Şekil 7: ARFF formatına dönüştürülmüş ekran görüntüsü	41
Şekil 8: WEKA Giriş Ekranı.....	41
Şekil 9: Weka Sonuç Ekranı 1	42
Şekil 10: Weka Sonuç Ekranı 2	43
Şekil 11: WEKA Müşteri Adres Dağılım Görselleştirme Ekranı	46
Şekil 12: WEKA Müşteri Adres Dağılım Küme Yoğunlaşması.....	47

KISALTMALAR

ARFF	Attribute Relation File Format
CO	Cari Oran
COSO	Committee of Sponsoring Organizations of the Treadway Commission
ERP	Enterprise Resource Planning
GNU	GNU's Not Unix
ISO	International Organization for Standardization
KOBİ	Küçük ve Orta Büyüklükte İşletme
MATLAB	Matrix Laboratory
s.	Sayı
SEC	Securities and Exchange Commission
TEFE	Toptan Eşya Fiyatları Endeksi
TÜFE	Tüketici Fiyatları Endeksi
TÜİK	Türkiye İstatistik Kurumu
YSA	Yapay Sinir Ağları
WEKA	Waikato Environment For Knowledge Analysis

1. GİRİŞ

Bilgisayar sistemlerindeki gelişmeyle birlikte işletmelerde üretilen sayısal bilgi miktarının artması dolayısıyla veri tabanlarının daha fazla veriyi saklayabilecek boyutlara vardığı görülmektedir. Bu durum doğru ve daha ayrıntılı bilgiye ulaşmamızı mümkün hale getirmiş ancak başka bir sorunu ortaya çıkarmıştır. Bu da büyük sayısal veri yığınlarının yönetilmesi ve anlamlı hale getirilmesidir.

Çalışmanın amacı; veri madenciliği uygulaması yoluyla bir işletmenin elindeki dağınık veriyi segmentasyon bilgisine dönüştürerek doğru satış stratejisi sağlamaya yardımcı olmaktır. Bununla beraber, analiz sonuçlarının iç kontrol sisteminde ve etik değerler üzerinde ortaya çıkarabileceği etkileri araştırmak ve firmanın satışlarını arttırması durumunda bilanço hesapları üzerine etkilerini incelemektir.

Çalışmada yöntem olarak veri madenciliği modelleme çeşitlerinden biri olan kümeleme yöntemi ve k-means algoritması uygulanmıştır. Müşteri işletmeye ait veriler Excel ortamında alınmış, dönüştürülmüş ve Genel Kamu Lisanslı açık kaynak kodlu Weka yazılım programına yüklenmiş ve elde edilen segmentasyon raporu sonuçları incelenmiştir.

İşletmeye güven esasının temelinde etkin bir iç kontrol sistemi olması gerektiği açıktır. İkinci bölümde COSO modeli iç kontrol sistemi ve önemi, bileşenleri, işletmeye sağlayacağı faydalar ve risk yönetiminden bahsedilerek iç kontrol ortamının genel çerçevesi çizilmeye çalışılmıştır.

Üçüncü bölümde dünyanın geçirdiği dönüşüm itibarıyla büyük veri kavramından yola çıkılarak veri madenciliği süreci açıklanmıştır. Veri madenciliğinin tarihsel arka planı, süreç tasarımı, kullanılan model, algoritmalar ve bu modellerin işletmeler açısından hangi faydaları sağlayacağı ile veri madenciliği uygulamaları sırasında karşılaşılabilecek problemler açıklanmıştır.

Dördüncü bölümde bir kalibrasyon firmasının devamlı müşterileri ve aldıkları hizmet karakteristiklerini analiz etmek amacıyla bir uygulama örneği yapılarak müşteri

segmentasyonu gerekleřtirilmiřtir. Firmanın 2013, 2014 ve 2015 yılları itibariyle 251 adet müşterisine 6 adet veri bařlığında kümeleme analizi yapılmıř ve müşterileri en ok hizmet satıřı yaptıkları sektörler, en ok faaliyet gösterdikleri bölgesel dağılımları, yıllık kalibrasyon periyodları, eski/yeni müşteri yapısı, sektör büyüklükleri ve satışlarının yoğunlařtığı tutar aralıkları esasında iki kümeye ayrılarak segmente edilmiř ve analiz sonuçları yorumlanmıřtır.

Beřinci bölümde öncelikle iřletmelerin Kurumsal Kaynak Planlama(ERP) sistemlerine ihtiya duymalarını ortaya ıkaran sebepler ve i kontrol sistemi erevesinde veri madenciliğı tekniklerinin ERP sistemlerine uyumlařtırılmasının önemi üzerinde durulmuřtur. Dördüncü bölümde uygulanan segmentasyon analizi sonuçları i kontrol aısından deęerlendirilmiř ve iřletmenin satış stratejisini uygularken dikkat etmesi gereken etik ilkeler üzerinde durulmuřtur. Devamında bilano kavramı aıklanarak segmentasyon analizi sonuçlarının satışları arttırdığı varsayımı altında analiz sonuçlarının bilano hesapları üzerindeki etkisi arařtırılmıřtır.

2. İŞLETMELERDE İÇ KONTROL SİSTEMİ VE ÖNEMİ

2.1. İÇ KONTROL TANIMI VE KAPSAMI

2.1.1. İç Kontrol Kavramının Tanımı

İşletmelerde operasyonel süreçlerde etkinlik ve verimliliğin yüksek düzeyde sağlanabilmesi, işletme varlıklarının hata, hile, suistimal etkilerine karşı korunabilmesi, mali raporların güvenilirliğinin sağlanabilmesi, kurumsallaşma yolunda sağlıklı adımlar atılabilmesi ve tüm süreçlerin kayıt altına alınarak yasal düzenlemelere uyum sağlanabilmesi etkin bir iç kontrol sisteminin varlığını zorunlu kılmaktadır.

Literatürde iç kontrol; “kurumun amaçlarına ve belirlenmiş politikalara uygun olarak faaliyetlerin etkili, ekonomik ve verimli bir şekilde yürütülmesini, varlık ve kaynakların korunmasını, muhasebe kayıtlarının doğru ve tam olarak tutulmasını, mali bilgi ve yönetim bilgisinin zamanında ve güvenilir olarak üretmesini sağlamak üzere, kurum tarafından oluşturulan organizasyon, yönetim, süreç, mali ve diğer kontroller bütünü” şeklinde tanımlanmaktadır. (Özer, 2010, s.69.).

2.1.2. İç Kontrolün Kapsamı

İç kontrol işletmelerin mali ve mali olmayan tüm faaliyetlerini ve işlemlerini kapsamaktadır. İşletmelerde iç kontrol sisteminin tasarlanması, kurulması, uygulanması, izlenmesi ve geliştirilmesi esas olarak üst yönetimin sorumluluğunda ve inisiyatifindedir. Ancak; kontrol faaliyetleri tüm yönetici ve çalışanların ortak sorumluluğundadır ve sistemin başarısı iç kontrol kriterlerinin her düzeydeki çalışan tarafından benimsenmesine bağlıdır.

İç kontrolün, sadece kontrol faaliyetlerini değil, örgütsel yapıyı, işleyişi, görev yetki ve sorumlulukları, karar alma süreçlerini kapsayan ve kuruluş çalışanlarının tamamının rol aldığı dinamik bir süreç olması, sistemin makul bir güvence sağlamak üzere tasarım ve uygulama eksikliklerinin giderilmesi amacıyla sürekli izleme ve değerlendirmeye tabi tutulmasını gerektirmektedir. Bu durum iç denetim faaliyetinin gerekliliğini ortaya koymaktadır.(Özer, 2010, s.70.). Dolayısıyla iç denetimin öncelikli

olarak iç kontrol sistemi üzerine odaklanması ve sistemin etkinliğini ölçmesi gerekmektedir. İç kontrol sisteminin denetlenmesinde görevlerin ayrımı ilkesi, işlemlerin yetkilendirilmiş kişilerce yapılması, uygun belgeleme ve muhasebe kayıt düzeninin varlığı, muhasebe kayıtlarının fiziki olarak korunması ve bağımsız mutabakatların yapılması ilkelerinin nasıl işlediğine bakılması gerekmektedir. (Önder, 2008, s.89.).

2.2. İÇ KONTROL SİSTEMİNİN AMAÇLARI

İç kontrol sisteminin temel amaçları şunlardır: (Pehlivanlı, 2010, s.31.).

1. Süreç ve görev tanımlarının düzenlenmesi ve gözden geçirilmesi yoluyla operasyonların etkinlik ve verimliliğinin sağlanması.
2. İşletme varlıklarının her türlü hata hile ve suistimal işlemlerine karşı korunması ve usulsüzlüklerin önlenmesi.
3. Raporların zamanında ve güvenilir şekilde düzenlenmesi.
4. Kanunlar, yönetmelikler ve yasal düzenlemelerden doğan tüm kurallara uyum sağlanması ve prosedürlerin ilgili mevzuat doğrultusunda gözden geçirilmesi.
5. İşletme faaliyetlerinin kayıt altına alınması.
6. İşletmeye ait bilgilerin doğruluk ve güvenilirliğinin temini.
7. Yönetim ve muhasebe kontrolü.

2.3. COSO (COMMITTEE OF SPONSORING ORGANIZATIONS OF THE TREADWAY COMMISSION) İÇ KONTROL MODELİ YATAY VE DİKEY BİLEŞENLERİ

2.3.1. COSO Süreci

COSO, 1985 yılında Amerika'da faaliyet gösteren beş meslek örgütünün biraraya gelerek oluşturduğu hem özel hem de kamu sektöründe genel kabul gören bir nitelik taşıyan bir kontrol modelidir. Modelin beş meslek örgütü aşağıda belirtilmiştir. (Baskıcı, 2012, s.41.).

1. Amerikan Sertifikalı Kamu Muhasebecileri Enstitüsü (American Institute of Certified Public Accountants)
2. Amerikan Muhasebe Birliđi (American Accounting Associations)
3. Finansal Yöneticiler Enstitüsü (Financial Executives Institute)
4. İç Denetçiler Enstitüsü (Institute of Internal Auditors)
5. Yönetim Muhasebecileri Enstitüsü (Institute of Management Accountants)

2.3.2. COSO İç Kontrol Entegre Çerçeve Raporu

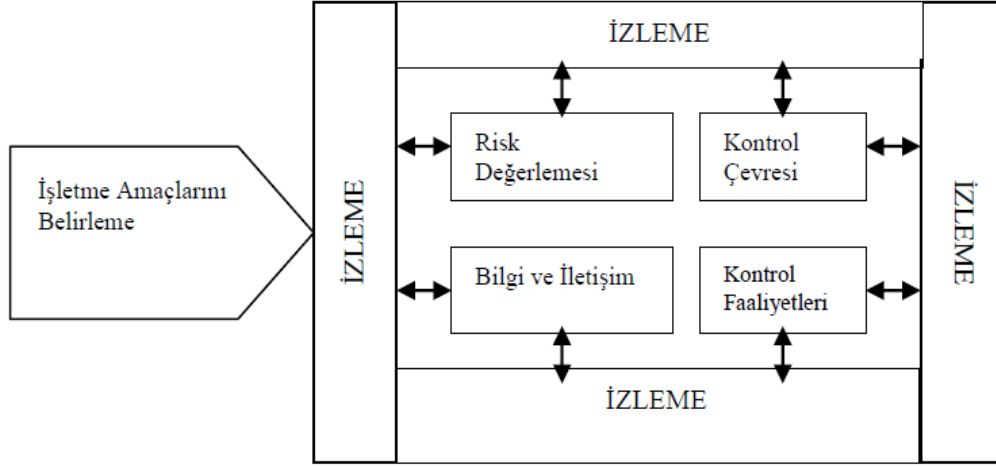
1992 tarihli COSO İç Kontrol Entegre Çerçeve Raporu'nun yayınlanması iç kontrolün yönetsel bir araç olarak kurumsal performansının artırılması, hızla deđişen dünyada iç ve dış risklerin tümüne karşı bir tepki niteliđi taşıması ve tüm dünyada iç kontrole ilişkin ortak bir bakış açısı yakalaması açısından önemlidir. Böylelikle iç kontrol sadece yönetimin temel fonksiyonlarından bir tanesi olma kapsamından çıkartılmış, diđer yönetim fonksiyonları olan “planlama, koordinasyon ve yürütme” nin de önemli bir destekleyicisi haline getirilmiştir. (Kıral, 2014, s.284).

2.3.3. COSO'ya Göre İç Kontrol Kavramı

COSO tarafından yapılan iç kontrol tanımı şu şekildedir. İç kontrol mali tabloların güvenilirliđi, faaliyetlerin ve işlemlerin etkinliđi ve verimliliđi, faaliyetlerin yasa ve yönetmeliklere uygunluđunu sağlama konusunda sınırlı bir güvence vermek üzere, firma üst yönetimi veya yönetim kurulu tarafından önerilen ve kontrol edilen bir yöntemdir. (Aksoy, 2006, s.32.).

2.3.4. COSO Modeli İç Kontrol Bileşenleri

COSO modeli iç kontrol bileşenlerinin işletme amaçlarını belirleme açısından birbirleriyle olan etkileşimleri Şekil 1'de gösterilmiştir.



Şekil 1: Coso İç Kontrol Sürecini Oluşturan Unsurlar

Kaynak: Metin Atmaca, “Muhasebe Skandallarının Önlenmesinde İç Kontrol Sisteminin Etkinleştirilmesi”, Afyon Kocatepe Üniversitesi İİBF Dergisi. (C.XIV, S I,2012), s.197.

1. **Kontrol Çevresi:** Örgüt genelinde kontrol ortamı hazırlamaya yönelik bir unsur olan kontrol çevresinin bileşenleri örgüt yapısı, işletme ortamı ve çalışma tarzı, etik ve ahlaki değerler, kişilerin yeterliliği, yönetim felsefesi, sorumluluk, güvenilirlik, insan kaynakları politikası ve çalışanların gelişimini ifade etmektedir. (Atmaca, 2012, s.197.).
2. **Kontrol Faaliyetleri:** Yönetim faaliyetlerinin bileşenleri yönetim direktiflerini uygulayacak prosedürler, üst yönetim kontrolleri, fonksiyon yada aktivite yönetimi, fiziksel kontroller, uygunluk prosedürleri, kalite programları, şirket amaçlarına ulaşmayı sağlayacak mekanizmalar, görevlerin ayrılığı ilkesidir. (Özer, 2010, s.89).
3. **Risk Değerleme:** Risk değerlendirme bileşenleri riskin işletmeye olan etkisinin tahmini, riskin ortaya çıkma ihtimali ve sıklığının belirlenmesi, riskin nasıl yönetileceği ve hangi önlemlerin alınabileceğinin belirlenmesidir. (Pehlivanlı, 2010, s.78).

4. **Bilgi ve İletişim:** Operasyonel ve finansal bilginin zamanında ve doğru olarak elde edilmesi, iç ve dış bilgiye erişim, organizasyonel iletişim bilgi ve iletişim ortamının bileşenleridir. (Erdoğan, 2006, s.98.).
5. **İzleme:** Denetçinin iç kontrol yapısını tanıma sırasında başvurduğu ve kontrollerin test edilme aşamasında da kanıt toplamak amacı ile kullandığı denetim prosedürlerinden biri olan izleme, önemli bir iç kontrol bileşenidir. Denetçi izleme ile denetleyeceği işletmenin iç kontrol yapısı hakkında genel bir fikir edinmektedir. Kanıt toplamak amacı ile ne tür denetim prosedürleri kullanacağını kararlaştırmaktadır. (Haftacı, 2007, s.142.).

2.4. ETKİN BİR İÇ KONTROL SİSTEMİ İÇİN TEMEL İLKELER

Etkin bir iç kontrol sistemi için temel ilkeler muhasebe organizasyon yapısının kontrolü, görevlerin ayrımı, yetkilendirme, belgeleme ve kayıt düzeni, fiziksel kontroller (yangına karşı koruma sistemleri ve ya iyi düzenlenmiş bir ambar sistemi gibi) ve bağımsız mutabakatlardır.

Görevler ayrılığı ilkesi gereğince, kıymet hareketine neden olunmasının onaylanması ile varlıkların korunması görevleri birbirinden ayrılmış olması önem taşır. Kıymet hareketine neden olunmasını onaylayan kişinin, bu kıymet hareketi ile satın alınan varlıkları koruyan kişiden farklı olması gerekmektedir. Aksi halde yolsuzluk yapma olasılığı artacaktır. Buna göre alış faturası ile ilgili ödemelerin yapılmasını onaylayan kişi hazırlanan ödeme çekini imza yetkisini taşımamalıdır. Varlıkları korumaktan sorumlu olan kişiler bu varlıklarla ilgili muhasebe kayıtlarını yapmamalıdır. Böyle bir ayırım yapılabilecek yolsuzlukların önlenmesine yardımcı olmaktadır. Aksi takdirde bir kişinin varlıkları kendi kişisel kullanımına ayırması ve kendine haksız kazanç sağlaması; aynı zamanda muhasebe kayıtlarında gerekli değiştirmeleri yaparak durumu gizlemesi mümkün olabilmektedir. Örneğin, bir görevli hem müşterilerden tahsilat yapmakta ve hemde alacak hesapları ile ilgili kayıtları tutmakta ise, müşterilerden tahsil ettiği nakdi cebine atabilir ve müşteri hesabına satış kaydını hiç yazmamak veya gerçek olmayan bir alacaklandırma yazmak yoluyla, yaptığı hırsızlığı dengeleyen bir ayarlama yapabilmektedir. (Güredin, 1998, s.180.).

2.5. İÇ KONTROL VE RİSK YÖNETİMİ

İç kontrol süreçlerinin önemli bir bileşeni olan risk değerlemesi denetçiye işletmeyi tanıması ve risk haritalarını yorumlamasında yön göstermektedir. Denetçi kanıtlarını topladıktan sonra sonuçlar ayrıca denetim risk modeli değerlemesi açısından belirtilmektedir. Buna göre; bulgu riski, kabul edilen denetim riski, yapısal risk ve kontrol risklerinin toplamından meydana gelmektedir. (Arens, 2003, s.254).

Şirketin iç kontrol etkinliği ve risk değerlemesi için denetçinin izlemesi gereken adımlar şunlardır: (Carmichael, 1989, s.168.).

- Kontrol çevresi, muhasebe sistemi ve kontrol prosedürlerinin kavranması için gerekli bilgi toplar.
- Herbir önemli hesap bakiyesi, işlem sınıflaması ve mali tablo unsurları için kontrol risk değerlemesi yapar.
- Denetçi açığa çıkan kontrol riskinin maksimumdan düşük seviyede değerlendirilebileceğine inanırsa, süreçlerin unsur ve tasarımlarının etkinliği ve alanında şirketin ilgili olduğu politika ve prosedürler hakkında kanıt toplar.

Karşılaşılabilecek diğer çeşitli risk türleri şunlardır: (Taylor, 1994, s.671).

1. **Kredi riski:** Müşterinin borçlu olduğu tutarı ödemeye gücü olmaması yada ödemeye isteksiz olması, denetçilerin çalışmasının odak noktasıdır.
2. **Teminat riski:** Bazı alacaklar, özellikle senetler, teminatla garanti altına alınmış olabilmektedir. Buradaki risk teminatın yasal durumunun kusurlu olması yada teminat değerinin mevcut bakiyeden düşük olması yani alacağı karşılamamasıdır. Eğer teminatın alacağın tahsilatının belirlenmesi üzerinde önemli bir etkisi varsa denetçi uzmanlık hizmeti alabilmektedir.
3. **Yoğunlaşma riski:** Eğer çoğu şirketin müşterileri aynı endüstri kolunda yada coğrafi bölgedeyse o endüstri yada bölgedeki bir ekonomik durgunluk denetçinin riskini arttıracaktır.
4. **Yönetim riski:** Yönetimin kredi verme üzerine kurulan politikaları, serbest yada sıkı olmasından bağımsız, denetçinin dikkate alması gereken faktörlerdir.

- 5. Operasyonel risk:** Müşterilerden olan potansiyel tahsil edilemeyecek hesapların izlenmesinde yetersiz kalınması ve zamanında etkin şekilde aksiyon alınamaması operasyonel bir risktir. Yetersiz veya başarısız iç süreçlerin tamamını ve bilgi sistemlerini de kapsamaktadır.
- 6. Hile ve içsel risk:** Uygun yetkilendirmeler olmaksızın çalışanların yada ilişkili tarafların suç ortağı olduğu satışlar eğer karşı taraf tamamıyla geri ödeme yapma eğilimi taşııyorsa şirkette zarar ortaya çıkmasına sebep olabilmektedir.
- 7. İş riski:** Yasal düzenlemeler, teknolojik gelişmeler, müşteri davranışları ve rekabet ortamına ilişkin ekonomide ve genel iş koşullarında meydana gelen genel ve beklenmedik değişiklikler sebebiyle kuruluşun maliyet yapısı, ürün konumlandırması, fiyatlandırma veya faaliyetlerindeki adaptasyon sorunları dolayısıyla beklenenden daha düşük gelir elde edilmesi yada daha yüksek maliyetlere katlanılmasından kaynaklanan risktir.
- 8. Piyasa riski:** Piyasadaki dalgalanmalar nedeniyle yatırımcının yada şirketin zarar görmesi yada hedeflerine ulaşamamasıdır. Volatilite olarak da tanımlanan piyasa riski, faiz oranlarındaki değişimi, döviz kurlarındaki değişimi, yada ekonomide düşme (resesyon) dönemine girilmesini içermektedir.
- 9. Likidite riski:** Nakit ihtiyacı olduğu sırada mevcut varlıkları hızlı nakde çevirememeye riski olarak tanımlanır.

3. VERİ MADENCİLİĞİ VE KULLANIM ALANLARI

3.1. BÜYÜK VERİ KAVRAMI

Günümüzde şirketler büyük veriden etkin şekilde faydalanabilmek, bilgiye erişmek ve bilgiyi üretmek konusunda teknolojinin de desteğiyle fark ortaya koyabilmek ve stratejik üstünlük sağlayabilmek için veri madenciliğini bir araç olarak kullanmaktadır. Bu durum yeni yönetsel yaklaşımların temeli olarak büyük veriden büyük değer beklentisini de doğurmaktadır.

Büyük veri, terim olarak tek bir sunucuya sığamayacak ölçüde büyük, satır sütun yapılı veri tabanlarına uyabilecek ölçüde yapılandırılmamış veya statik bir veri ambarına sığamayacak şekilde sürekli akan veriler için kullanılmaktadır. Avantajlı yanı, birçok büyük veri teknolojisinin bedava (açık kaynaklı yazılımlar gibi) veya görece ucuz (ticari sunucular gibi) olmasıdır. Dezavantajı ise, büyük veri teknolojilerinin mimari ve programlama açısından görece emek yoğun bir alan olmasıdır. (Davenport, 2014, s.129.).

Büyük veri, doğru analiz metotları ile yorumlandığında şirketlerin stratejik kararlarını doğru bir biçimde almalarına, risklerini daha iyi yönetmelerine ve inovasyon yapmalarına imkan sağlar. Şirketlerin çoğu, halen konvansiyonel veri ambarı ve veri madenciliği yöntemleriyle elde ettikleri verilerden yola çıkarak, karar almaya devam etmektedirler. Ancak, tüketici eğilimlerini dinamik şekilde öngörebilmek, büyük veriyi analiz edebilmekten ve bu analizlere göre çözüm üretmekten geçmektedir. Büyük veri, geleneksel veritabanı araçları ve algoritmaları ile işlemesi zor olan bu büyük verinin elde edilmesi, saklanması, akışı, analiz edilmesi gibi birçok konuyu içeren bir terim olarak karşımıza çıkmaktadır. Veriler klasik veritabanlarının kaldıramayacağı büyüklükte olduğu gibi verinin büyüme hızı da bir bilgisayar veya bir veri depolama ünitesini aşmaktadır.

Özetle, “büyük veri” olarak adlandırılan bilişim probleminin dört boyutu olduğundan bahsedilmektedir. Bunlar; büyük kütleler, yüksek hızlar, çok çeşitlilik ve verideki tutarsızlıklardır.(Ölmezogulları, 2013, s.1.).

3.2. BÜYÜK VERİ VE GELENEKSEL YÖNETİM YAKLAŞIMLARI

Büyük veri ile alakalı pratikte henüz çözümlenemeyen olgu yeni yönetim yaklaşımlarına ve şirketler için yeni teknoloji mimarisine ihtiyaç olduğudur. Bunun temel sebebi büyük verideki verinin bir havuzdan çok bir nehir gibi aralıksız ve hızla akıyor olmasıdır. Eski dünyada diğer bir deyişle geleneksel karar mekanizmalarında veri analiz için farklı havuzlara ayrılır, model belirlenir ve karar alıcıya destek sağlanırdı. Ancak büyük veriyi hızla akan bir enformasyon olarak düşünürsek bu onun aynı zamanda herhangi bir ayırma yaklaşımını süretle aşabileceği anlamını taşır. Geleneksel yönetim yaklaşımlarında iş ve teknoloji mimarisi işlem ve operasyonlarla başlayıp analiz ve içgörüyeye doğru ilerleyen bir akışı yansıtırken, büyük veri bu yaklaşımı baş aşağı eder; çünkü temel düşünce şudur; veri sürekli bir değişim ve akış halindeyse hızla idrak edip akıllıca tepki veren kuruluşlar avantajlı konumdadır. Bunun anlamı keşif ve çeviklik ön plandadır. (Davenport, 2014, s.129.). Tablo 1’de büyük veri ve geleneksel analitik karşılaştırması verilmiştir.

Tablo 1

Büyük Veri ve Geleneksel Analitiğin Karşılaştırılması

Büyük Veri ve Geleneksel Analitik		
	Büyük Veri	Geleneksel Analitik
Veri tipi	Yapılandırılmış format	Satır sütun şeklinde yapılandırılmış format
Veri hacmi	100 terabayt’tan 1 petabayt’a kadar	100 terabayttan daha az
Veri akışı	Sürekli veri akışı	Statik veri havuzu
Analiz yöntemi	Makine öğrenmesi	Hipoteze dayalı
Birincil amaç	Veriye dayalı ürünler	İç karar desteği ve hizmetleri

Kaynak: Thomas Davenport, “Big Data@Work”, (2014) İstanbul: bzd Yayın ve İletişim Hizmetleri, s.129 (Türk Hava Yolları Yayınları)

LinkedIn'in kurucularından Reid Hoffman güçlü bir büyük veri taraftarı olarak şunu sunmaktadır: (Davenport, 2014, s.129.).

“Tüm bu veri, Web 2.0[sosyal ağlarda ve tüketicilerin web'e katılımındaki patlama] ve artan algılayıcı miktarı nedeniyle var. İnsanlar ve yerler ve bizim için önemli olan her şey ve hayatlarımız etrafında büyük ölçüde semantik olarak indekslenmiş devasa miktardaki bu veriden tonlarca ilginç uygulama çıkacağına inanıyorum...Ürün ve hizmetlerimizin oluşturulma şekli, stratejimizi belirleme ve diğerlerine karşı rekabet avantajı yakalama şeklimiz –eğer veri bunların her biri için güçlü bir unursa ve siz hiçbirşey yapmıyorsanız, bu iş zekası kullanmadan iş yapmaya çalışmaya benzer.”

3.3. VERİ MADENCİLİĞİNİN TANIMI

Genel anlamda veri madenciliği; bilgisayar sistemlerinde depolanan veri yığınları arasından istatistik ve matematik teknikleri kullanılarak verilerdeki gizli örüntüleri çözmeye yarayan, fark edilmesi güç ilişkileri açığa çıkaran, ileriye yönelik tahminler yapılmasını sağlayan ve bu alanda kurallar üreten veri tabanı teknolojisi ve tekniklerinin uygulanmasını ifade etmektedir. (Özkul, Özdemir, 2011, s.108.).

Daha basit ifadeyle veri madenciliği, büyük hacimli veri yığınları içerisinde karar alabilmek için potansiyel olarak faydalı olabilecek, uygulanabilir ve anlamlı bilgilerin çıkarılması işlemleridir. (Pehlivanlı, 2010, s.77.).

Bu konu her geçen gün daha da geliştiği için kullanılan yer ve zamana göre farklı tanımlar karşımıza çıkmaktadır. Bir diğer yaygın tanıma göre veri madenciliği; daha önceden bilinmeyen, geçerli ve uygulanabilir bilgilerin geniş veri tabanlarından elde edilmesi ve bu bilgilerin işletme kararları verirken kullanılmasıdır. Burada altının çizilmesi gereken önemli noktalardan birisi, elde edilecek bilginin 'önceden bilinmeyen' olmasıdır; ulaşılabilecek bilginin önceden bilinmiyor olmasından kasıt, elde edilecek sonucun tahmin edilememesi anlamını taşımaktadır. (Silahtaroglu, 2013, s.12.).

Büyük sayısal verilerin yönetilmesi konusunda karar alıcılar için güçlü bir araç olarak kullanılan veri madenciliği istatistik, matematik disiplinleri, modelleme teknikleri, veritabanı sistemleri ve bilgisayar programları aracılığıyla kullanıcı hedeflerine katkı sağlar. Bu da veri madenciliğinin disiplinler arası bir kavram olduğunu ortaya koymaktadır.

Günümüzde satış pazarlama, biyoloji, finans ve bankacılık, sigortacılık, borsa, perakendecilik, telekomünikasyon, genetik, sağlık, bilim ve mühendislik, kriminoloji, sağlık, endüstri, istihbarat, eğitim, üretim ve planlama, taşımacılık ve benzeri birçok alanda karar alma süreçlerinde veri madenciliği kullanılmaktadır.

3.4. VERİ MADENCİLİĞİNİN TARİHSEL SÜREÇTEKİ GELİŞİMİ

Terim olarak baktığımızda ilk olarak 1960'lı yıllarda istatistikçilerin kullandığı Veri Avlama (Data Fishing) veya Veri Tarama (Data Dredging) adlandırmalarının veri madenciliğinin öncülleri olduğu görülmektedir. Her iki terimde denizcilik literatüründen türetilmiştir. Veri Tarama (Dredging) fiiline baktığımızda denizi ağ ile tarama veya deniz tabanını tarama anlamına gelen özel bir terim olduğu görülmektedir. İstatistikçiler bu terimleri ayrıntılı analizlere başlamadan önce kötü verilerin bulunup ayıklanması manalarını ihtiva etmek için kullanmışlardır. (Pektaş, 2013, s.99.).

1990'lara gelindiğinde Veri Madenciliği (Data Mining) kavramı ortaya atılmış geleneksel istatistik yöntemler yerine veri analizinin algoritmik bilgisayar modülleri tarafından değerlendirildiği süreç başlamıştır.

1990'ların ortalarında dünya çapındaki internet ağının gelişi (www) veriyi, bilgiyi ve bu zamana kadar birikmiş bilgileri etkin olarak yönetmek için büyük bir talep belirlemiştir. Ağ bağlantısı üzerindeki pek çok verinin geleneksel araçlarla yönetilmesi zamanla imkansız hale gelmiştir. Günümüzde veri pek çok organizasyonda kritik bir kaynaktır ve bu kaynağa etkin biçimde erişim sağlamak, veriyi paylaşmak, verinin içinden gerekli bilgiyi ayıklayarak çekmek ve bundan yararlanmak acil ihtiyaç haline gelmiştir. Bu nedenle, depolamanın yanısıra çeşitli veri kaynakları ve sistemler arasında birlikte çalışılabilirliğin sağlanması ve veri ambarları yada veri tabanlarından çeşitli araçlar geliştirilerek verinin çekilmesi amaçlanmıştır.(Thuraisingham, 2003, s.XV.).

2000’li yılların başlarında Veri Tabanı Madenciliği kavramı San Diego’da bir veritabanı firması (o dönemki adı HNC) tarafından patentlenmiş ve piyasaya sunulmuştur.

Günümüzde literatürde veri madenciliğine eş değer olarak kullanılan en yaygın deyim **Veritabanlarında Bilgi Keşfi**’dir. (Pektaş, 2013, s.99.). Tablo 2 veri madenciliğindeki gelişmeleri özetlemektedir.

Tablo 2

Veri Madenciliğinin Tarihsel Süreçteki Gelişimi

Tarih	Basamaklar	Sorular	Kullanılabilir Teknolojiler	İlgili Yazılımlar
1960lar	Veri toplama, Veritabanı Yönetim Sistemleri	Benim son 5 yıldaki toplam kârım nedir?	Bilgisayar, Disk, Düz dosyalar	Fortran
1980ler	Veriye ulaşım, Veri sorgulama	Geçen Mart İstanbul’daki birim satış miktarı nedir?	Daha hızlı ve ucuz bilgisayarlar, daha fazla depolama alanı, ilişkisel veritabanları	Oracle,IBM DB, SQL
1990lar	Veri ambarları, Karar destek sistemleri	Geçen Mart İstanbul’daki birim satış miktarı nedir? Ankara ile karşılaştırmalı olarak görmek istiyorum.	Daha hızlı ve ucuz bilgisayarlar, Daha fazla depolama alanı, İlişkisel veritabanları, OLAP, Çok boyutlu veritabanları, Veri ambarları	SQL Standart, Veri Ambarları, OLAP, Darwin, IBM Intelligent Miner, SPSS Crisp DM, SAS Miner, Angoss Knowledge Seker
1990ların sonu	Veri madenciliği	Ankara’da gelecek ayki birim satışlarım ne durumda olacak? Neden?	Daha hızlı ve ucuz bilgisayarlar, Daha fazla depolama alanı, İlişkisel veritabanları, Gelişmiş bilgisayar algoritmaları	Oracle Data Miner, IBM DB2 UDB Mining, SPSS Clementine, SAS Enterprise Miner
2000ler	Web madenciliği			

Kaynak: Ceyda Durmaz ve Murat Kocamış, “Oracle Data Miner ile Öğrenci Kayıtları Üzerine Bir Veri Madenciliği Uygulaması”, (Mayıs 2008), Bitirme Ödevi. İstanbul: İTÜ Fen Edebiyat Fakültesi,s.4.

3.5. BİLGİ İŞLEME SİSTEMLERİNE İLİŞKİN TEMEL KAVRAMLAR

3.5.1. Verinin Tanımı

Veri ham haldeki işlenmemiş enformasyon parçacığı olarak tanımlanmaktadır. (<https://tr.wikipedia.org/wiki/Veri>) Enformasyonu ise en genel anlamda bir konuya (bağlama) ilişkin ölçüm, deney, gözlem, araştırma, haber toplama gibi bulguların özetlenmesi yöntemleriyle derlenmiş bilgi parçası ortaya koymaktadır. (<https://tr.wikipedia.org/wiki/Enformasyon>) Her iki tanımın da işaret ettiği gibi burada bilgi kavramı önemli bir rol oynamaktadır.

Veri, enformasyon ve bilgi arasındaki ilişkiye baktığımızda veri analiz edilmemiş ve özetlenmemiş ham gerçekler ve gözlemleri, enformasyon, ayrıntıları ile birlikte bir mesajın bir yerden diğer bir yere iletilmesi amacıyla bir iletişim kanalından geçmesi gereken veri miktarını, bilgi ise anlamlı bir forma sokulmuş ve tüm ayrıntıları yerine alıcı tarafından anlamı yeterince anlaşılacak şekilde bir iletişim kanalından geçmesi gereken veriyi ifade etmektedir. Bilgiyi veri ve enformasyondan daha değerli kılan, sözkonusu kavramlardan farklı olarak eyleme yakın olmasıdır. (Daştan, 2008, s.6.).

3.5.2. Veri Tabanı Sistemi

Bilgisayar teknolojisinde son yılların önemli bir gelişmesi olan veri tabanı birbiriyle ilişkili kütüklerin birleştirilerek, birbirinden bağımsız alanlarda ortaklaşa kullanımına imkan veren bir yapılanmadır. Veri tabanı uygulamasıyla bir işletmeye ilişkin tüm bilgiler, ortak kullanıma açık hale gelmektedir. (Erdoğan, 2006, s.140.).

Veri tabanı uygulaması aşağıda belirtilen şekillerde fayda sağlamaktadır; (Erdoğan, 2006, s.140.).

- Verilerin birden fazla yerde gereksizce saklanmasını önlemektedir.
- Belirlenmiş olan bir konudaki tüm bilgileri kapsamaktadır.
- Bilgilerin çelişmesini önleyerek tutarlılığı sağlamaktadır.

- Disk belleğinde tasarruf sağlamaktadır.
- Veriye erişimi hızlandırmaktadır.
- Kullanıcının öğrenmesi kolaydır.

3.5.3. Veri Ambarı Sistemi

Belirli bir döneme ve zaman dilimine ait, yapılacak çalışmaya göre konu odaklı düzenlenmiş, birleştirilmiş ve sabitlenmiş işletmelere ait veri tabanlarına **veri ambarları** denilmektedir. (Silahtaroglu, s.17.). Konu odaklılıktan kasıt; müşteri, ürün, teslimat gibi aynı olay yada varlığı ilgilendiren konulara ait verilerin birbirine bağlanmış olmasıdır. Veri ambarlarının bir diğer özelliği gerekli dönüştürme ve normalizasyon işlemleri sonrasında birden fazla veri tabanının bir araya getirilerek bütünleşik bir yapıya kavuşmuş olmasıdır. Son olarak veri ambarları, geçici ve uçucu değildir; yani veri ambarındaki veriler silinmez ve yeni veri eklenmez. Veri giriş çıkışı yoktur. İşlemsel veri tabanlarından, düz dosyalardan, vs. elde edilmiştir ve yeni bir veri giriş çıkışına uygun mimaride değildir. (Silahtaroglu, 2013, s.17.).

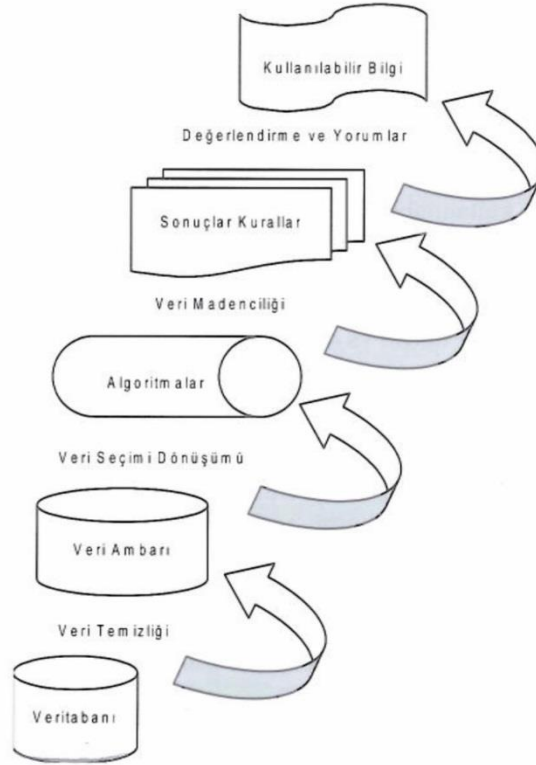
Veri ambarının temel özellikleri ve veri tabanı ile arasındaki farklar şunlardır: (Şeker, 2005, s.1.).

- Veri tabanlarının amacı bütün verileri tutmaktır, veri ambarı ise işlenmiş, özelleştirilmiş verileri tutar.
- Veri tabanları ile veri ambarları arasındaki en önemli farklardan birisi farklı kaynaklardan besleniyor olmalarıdır. Örneğin bir alışveriş merkezinde yazar kasalarından gelen veriler, müşteri yönetimi ile ilgili olan veriler, stok ile ilgili veriler veri tabanlarında tablolar halinde farklı veri kaynakları olarak saklanır. Saklanan bu farklı veri kaynaklarının birleştirilerek, bir amaca yönelik olarak raporlarının hazırlanması ve verinin daha verimli kullanılabilmesi için geliştirilen sistemlere veri ambarı denilmektedir.
- Veri ambarları konu odaklıdır; müşteri, ürün veya satış odaklı özelleştirilebilirler.

- Veri ambarları gidiş (trend) ve deęişim deęerlerini takip ederek zaman içinde kendini g¼ncellemektedir. Veri tabanlarındaki g¼ncellemelerden haberdar olabilmek için çeşitli algoritmalar kullanılmaktadır.
- Yapı olarak veri ambarları doğrudan g¼ncellenemez, veri tabanlarından gerekli g¼ncelleme işlemleri yapılır. Dolayısıyla veri tabanlarından veri ambarlarına tek yönlü bir veri akışı bulunmaktadır.
- Veri ambarları yönetim kademesinde veya karar verme süreçlerinde kullanılan sistemlerdir.

3.6. VERİ MADENCİLİęİ S¼RECI

Veri madencilięi ve bilgi keşfi sürecinde Şekil 2’de gösterilen aşamalar gerçekleştirilmektedir.



Şekil 2: Veri Madencilięi Süreçleri

Kaynak: G¼khan Silahtaroęlu, Veri Madencilięi Kavram ve Algoritmaları, Papatya Yayıncılık Eęitim A.Ş. (2013), s.24.

3.6.1. Verilerin Temizlenmesi

Verilerin hazırlanmasında verilerin birleştirilmeden önceki süreçte temizlenmesi gerekmektedir. Bunun için öncelikli olarak kayıp verilere yapılacak işlemler ve verideki gürültünün ortadan kaldırılması gerekmektedir. Veriyi temizlerken kayıp verinin bulunduğu kaydı veri tabanından çıkarmak, kayıp verileri elle teker teker doldurmak, tüm kayıp verilere aynı bilgiyi girmek, kayıp olan verilere tüm verilerin ortalama değerinin verilmesi gibi çözümler uygulanmaktadır. (Silahtaroglu, 2013, s.23).

3.6.2. Verilerin Birleştirilmesi

Verilerin birleştirilmesi aşaması ise farklı veri tabanlarındaki verilerin tek bir çatı altında (veri ambarında) birleştirilmesi işlemidir. Veriye sahip olma, verinin işlenmesi, iletilmesi ve depolama kapasitesindeki ilerlemeler işletmelerin sahip oldukları çeşitli veri tabanlarının veri ambarlarında birleştirilmesine olanak tanımıştır. En genel anlamında, veri ambarı farklı kaynaklarda tutulan verilerin ortak bir çatı altında birleştirilerek, verilerin zaman boyutunda birbiri ile konuşmasını sağlayan, tutarlı ve doğru verilerin yer aldığı sistemdir. (Tüzüntürk, 2010, s.5.).

3.6.3. Seçim

Bu aşamada yapılacak olan analiz ile ilgili veriler belirlenmektedir. İstatistiksel metodlar belirlendiğinde bunlar erişilebilir veri tabanından ihtiyacımız olan sonuçları sentezlememize yardımcı olmak için bilgisayar hesaplamalarına uygun algoritmalara dönüştürülecektir. Algoritma hakkında çok çeşitli tanımlar mevcuttur ancak bilgisayar bilimlerindeki kullanımını esas alırsak “bir sorunun çözümü için öngörülen işlemlerin mantık ve sembolik anlatımı” olarak tanımlanabilmektedir. (Kurnaz, 2004, s.15.). Veri madenciliği yönetim süreçleri yazılım çözümlerinde olduğu kadar farklı methotlarla ilgili de sağlam bir birikime sahip olmalı; böylece süreci şirketin kendine özgü ihtiyaçlarına adapte edebilme ve karar alma noktasında sonuçları doğru şekilde yorumlayabilme kabiliyeti kazanacaktır. (Giudici, 2003, s.9.).

3.6.4. Modelin Kurulması ve Değerlendirilmesi

Bu aşamada uygun model ve algoritma seçilmektedir. Ancak tanımlanan problem için çok sayıda model denenmekte ve en iyi olduğu düşünülen modelde karar verilene kadar süreç tekrar edilmektedir.(Çınar ve Silahtaroglu, 2012, s.311).

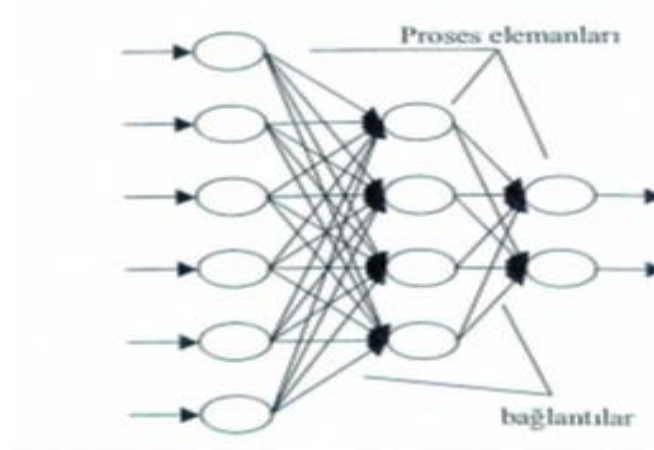
3.7. VERİ MADENCİLİĞİNDE KULLANILAN TEKNİKLER

3.7.1. Yapay Sinir Ağları (YSA)

YSA'lar biyolojik sinir ağlarından esinlenilerek geliştirilmiş, öğrenme yoluyla yeni bilgiler türetebilme ve keşfedebilme gibi yetenekleri hiçbir yardım almadan otomatik olarak gerçekleştirebilmek için tasarlanmış, bilgi işleme sistemleridir. (Silahtaroglu, 2013, s.121). İnsan beyninin öğrenme, ilişkilendirme, sınıflandırma, genelleme, özellik belirleme ve en iyileme (optimizasyon) fonksiyonlarına benzer şekilde tasarlanmaktadır.

Aynen insan beynindeki sinirlerin ölmesi ve yeni snaps bağlantılarının sağlanması gibi YSA'nın da kendi topolojik yapısını yenilemesi mümkündür. Hesaplamalarda kolaylık olması açısından, hücrelerin birbirine düzgün bir geometri ile bağlandıkları ve yapısal değişikliğin olmadığı kabul edilmektedir. (Dursun, 2000, s.14).

YSA'ların birbirine hiyerarşik olarak bağlı ve paralel olarak çalışabilen yapay hücrelerden oluştuklarını belirtebiliriz. Süreç (Proses) elemanları da denilen bu hücrelerin birbirlerine bağlandıkları ve her bağlantının bir değerinin olduğu kabul edilmektedir. Bilgi öğrenme yoluyla elde edilir ve süreç elemanlarının bağlantı değerlerinde saklanan dağınık bir hafıza söz konusudur. Süreç elemanlarının birbirleri ile bağlanmaları sonucu ortaya çıkan ağa YSA denmektedir. Bu ağ ilgili herhangi bir olayın örnekleri ile eğitilerek (öğrenme) genelleme yapabilecek yeteneğe kavuşturulur.(Öztemel, 2003, s.30.). Şekil 3'te bir YSA yapısı gösterilmiştir.



Şekil 3: YSA yapısı

Kaynak: Ercan Öztemel, Yapay Sinir Ağları, (2003) 1. Basım. İstanbul:Papatya Yayıncılık, s.33

YSA'ların başlıca özellikleri şunlardır: (Öztemel, 2003, s.33.).

- YSA'lar makina öğrenmesi gerçekleştirirler.
- Programları çalışma stili bilinen programlama yöntemlerine benzememektedirler.
- Bilginin saklanması sağlarlar.
- YSA'nın sinir ağlarının güvenle çalıştırılabilmesi için önce eğitilmeleri ve verimliliklerinin (performanslarının) test edilmesi gerekmektedir.
- Görülmemiş örnekler hakkında bilgi üretebilirler.
- Algılamaya yönelik olaylarda kullanılabilirler.
- Şekil, örüntü ilişkilendirme ve sınıflandırma yapabilirler.
- Örüntü tamamlama işlemleri gerçekleştirebilirler.
- Kendi kendini organize etme ve öğrenebilme yetenekleri vardır.
- Eksik bilgi ile çalışabilmektedirler.
- Hata toleransına sahiptirler.
- Belirsiz, tam olmayan bilgileri işleyebilmektedirler.
- Dereceli bozulma(graceful degradation) gösterirler.

- Dağıtık belleğe sahiptirler.
- Sadece sayısal (nümerik) bilgiler ile çalışabilmektedirler.

Kullanım yerlerinin veri madenciliği, optik karakter tanıma ve çek okuma, bankalardan kredi isteyen müracaatları değerlendirme, ürünün pazardaki performansını tahmin etme, kredi kartı hilelerini saptama, güvenlik sistemlerinde konuşma ve parmak izi tanıma, mekanik parçaların ömürlerinin ve kırılmalarının tahmin edilmesi ve kalite kontrolü, iş çizelgeleri ve iş sıralaması, üretim planlama ve çizelgeleme gibi alanlar olduğu görülmektedir.

YSA müşteri işletmenin kazançlarını öngörmede de kullanılmaktadır. Denetçi bu öngörülerini gerçek sonuçlarla karşılaştırarak sonuçların uygunluğu için yargıya varmaktadır. Öngörülüş kazançlar, denetçiye müşteri işletmenin başarısının devam edip etmeyeceğini göstermektedir. İç kontrolün zayıf noktalarını araştırmada, denetim kanıtlarını yorumlamada, iflas öngörülerinde bulunmada, satış tahminleri yapmada, kredi riskini tahmin etmede ve birçok muhasebe uygulamasında YSA başarıyla kullanılmaktadır.

Veri tabanının ortaya konulmasında YSA gelişim sürecinin en önemli aşamasıdır; çünkü bu veri tabanı, matematik model kurulmasında kullanılmaktadır. Örneğin, denetçi denetlediği işletme için iflas öngörüsünde bulunmak için öncelikle, iflas etmemiş ve etmiş işletmelere ilişkin(bu işletmelerin sayısı YSA'yı kullananlarca belirlenecektir; ancak çok sayıda işletme kullanılması uygundur) mali oranlar (rasyolar) bağımsız değişkenler olarak sisteme girilmektedir. Bir sonraki aşama kurma adıdır. Yazılım, otomatik olarak oranları (girdileri) ve iflas eden-etmeyen firmaları(çıktıları) gözden geçirerek girdilerle çıktılar arasındaki ilişkileri ve yapıları tanıyan bir model kuracaktır. Sonraki adımda da test verileri kullanılarak YSA'nın tahminleri ile test verilerinin gerçek değerleri karşılaştırılarak farklar ortaya konur. Eğer sonuçlar kabul edilemez düzeyde ise yazılım içindeki parametrelerde değişiklikler yapılmaktadır. Bu değişiklikler, kabul edilebilir bir model elde edilene dek sürdürülür. Kabul edilebilir bir modele ulaşıldığında veri tabanına girilen örneklerin sayısı artırılır veya daha iyi temsil özelliğine sahip verilerin toplanması yoluna gidilir. Eğer sonuçlar tatminkar ise bu durumda YSA'ya iflas öngörüsü yapılacak işletmenin modelde kullanılacak değerleri

(mali oranları) girilerek çıktı değeri alınmaktadır. Bu değer, denetlenen işletmenin iflas etme veya etmeme olasılığıdır. (Erdoğan, 2006, s.182.).

3.7.2. Genetik Algoritmalar

Genetik algoritmalar bilim adamlarının evrimsel yaklaşımı bilgisayar teknolojilerinde kullanmaya başlamasının doğal sonucu olarak genetik kombinasyon, mutasyon, doğal seleksiyon süreçleri gibi evrimsel yaklaşımı temel alan bilgisayar tabanlı problem çözme ve en iyileme (optimizasyon) teknikleri içerisinde en yaygın şekilde kullanılan algoritmalarlardır. Bir problem için muhtemel çözümler birbirleriyle yarışır. Olasılık kurallarına göre çalışan genetik algoritmalar çözüm uzayının tamamını değil belirli bir kısmını taramaktadır. Bu durumda aramanın etkinliğini yükselterek kısa sürede çözüme ulaşmayı sağlamaktadır. Genetik algoritmalar çok aday çeşidi olan en iyileme problemlerinin çözülmesinde (kredi adayları gibi) uygun bir araçtır. Bu algoritmalar, insanın karar alma yeteneğinin bilgisayarlar kullanılarak benzetilmesine dayanan bir tür yapay zeka tekniğidir. (Watson, 2002, s.435.).

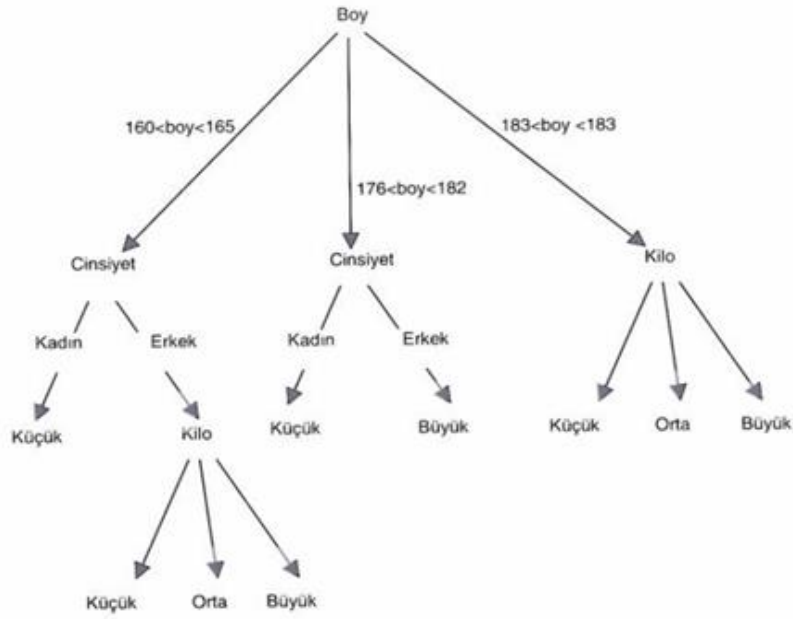
3.7.3 İstatistik Teknikler

Veri madenciliğinde verilerin önceden verilen sınıflara göre ayrılması gelecekte elde edilecek sonuçların tahminidir; yani sınıfların tahminidir. Regresyon, lojistik regresyon, zaman serileri analizi ve Bayes yaklaşımları gibi istatistiksel yöntemler kullanılarak bu sınıflandırma işlemleri gerçekleştirilebilir. (Silahtaroglu, s.97.).

3.7.4. Karar Ağaçları

Sınıflandırma problemlerinde en sık kullanılan algoritmalarından birisi karar ağaçlarıdır. Bu teknikte öncelikle bir ağaç ortaya konulur ve sonrasında veri tabanındaki tüm veriler ağaca uygulanır. Karar ağacı kurulurken en önemli nokta hangi değişkenin ilk düğüm, yani kök düğüm olacağını hesaplanmasıdır. Çünkü kök düğüm belirlendikten sonra boy, kilo ve cinsiyet gibi değişkenler birer düğümü temsil edecek, böylece her bir düğümden sonra dallanma yapılarak diğer düğümlere geçiş sağlanacaktır. Kök düğüm olarak öyle bir değişken seçilmeli ki verilen yanıt ne olursa

olsun, diğer değişkenlerle kıyaslandığında elimizdeki veri tabanını kabaca iki parçaya bölebilmektedir. Şekil 4’de kök düğümü boy değişkeni ile başladığı için cinsiyet ve kilo sorunları birden fazla düğümde sorulmaktadır. Ağaç kurulduktan sonra her kayıt bu ağaç üzerine uygulanır ve böylece her kayıt ulaşılan yaprağa göre sınıflandırılmıştır. Kurulan ağaç aslında bir çok “eğer (if then)” den meydana gelen bir kurallar bütünüdür. (Silahtaroglu, 2013, s.71.).



Şekil 4: Karar ağaçları

Kaynak: Kaynak: Gökhan Silahtaroglu, Veri Madenciliği Kavram ve Algoritmaları, Papatya Yayıncılık Eğitim A.Ş. (2013), s.72.

Öncelikle makine öğrenmesi ve uygulamalı istatistiğin içerisinde çok sayıda karar ağaçları tümevarım algoritmaları tanımlanmıştır. Bunlar girdi-çıkı örneklerinden karar ağacı kurabilen denetimli öğrenme yöntemleridir. Tipik bir karar ağacı öğrenme sistemi araştırma evreninin bir parçası olarak çözüm ararken yukarıdan aşağıya stratejisini de benimsemektedir.

Tümevarım öğrenme yaklaşımlarına dayalı olan bazı yöntemlerin uygulanabilmesi için bir kaç anahtara ihtiyaç vardır: Bunlar nitelik değer tanımı,

önceden tanımlanmış sınıflar (denetimli öğrenme açısından), ayırık sınıflar, yeterli data, mantıksal sınıflandırma modelleridir. (Kantardzic, 2001, s.142.).

3.7.5. Bulanık Kümeler Yaklaşımı

Veri madenciliği uygulamalarında özellikle sınıflama modellemesinde kullanılan bulanık kümeler yaklaşımı belirsizlik kavramı ile ilgilenmektedir ve kaynağını bulanık mantıktan almaktadır. Genel olarak değişik biçimlerde ortaya çıkan karmaşıklık ve belirsizlik gibi tam ve kesin olmayan bilgi kaynaklarına bulanık (fuzzy) kaynaklar adı verilmektedir. (Şen, 2001, s.9).

1970’li yıllardan sonra Doğu dünyasında özellikle Japonya’da bulanık mantık kavramına çok önem verilmiş ve teknoloji alanında kullanılmaya başlanmıştır. Batıda ise gecikmenin ana sebebi Batı kültürünün temelinde ikili Aristo mantığı yatması ve olaylara evet-hayır, siyah-beyaz, 0-1 gibi ikili esasta kesinlikle yaklaşılmasıdır. Batı’da bu iki değer arasında başka seçeneklere kesin değil düşüncesi ile hiç yer verilmezdi. Bunun temel sebebi, Batı’da bulanık (fuzzy) kelimesi güvensizliği ifade ederken Doğu’da ise bu güvensizlikte bile güzelliklerin bulunabileceği düşüncesinin yer almasıdır. Örneğin, insanlar arasındaki gerekli diyalogun bile sağlanması bu tür bulanık(kesin olmayan, oldukça kişisel) görüşlere bağlıdır. (Şen, 2001, s.10).

Aristo mantığına göre insanlar boy bakımından ya uzundur veya değildir. Halbuki Zadeh yaklaşımına göre uzun boyluluğun değişik dereceleri vardır. Aristo mantığına göre çalışan ve şimdiye kadar alışlagelen klasik küme kavramında, bir kümeye giren öğelerin oraya ait oluşları durumunda üyelik dereceleri 1’e, ait olmamaları durumunda ise 0’a eşit varsayılmıştır. İkisi arasında hiçbir üyelik derecesi düşünülemez. Halbuki bulanık kümeler kavramında 0 ile 1 arasında değişen üyelik derecelerinden söz etmek mümkündür. Böylece daha şimdiden bulanık kümelerdeki öğelerin üyelik derecelerinin kesintisiz olarak 0 ile 1 arasında değerler aldığından söz edebilmektedir. Aslında Zadeh küme öğelerinin üyelik derecelerinin 0 ile 1 arasında değişebileceğini ileri sürerek kümeler teorisinde geniş uygulamaya sahip ve doğal hayatla uyumlu olan bulanık küme teorisini geliştirmiştir. (Şen, 2001, s.17.).

Zadeh'in kendi ifadesiyle: (French, 1986, s.361).

“Sık sık, gerçek dünyada karşılaşılan nesnelere sınıfları tam olarak tanımlanmış üyelik kriterlerine sahip değildirler. Örnek olarak, hayvanlar sınıfı açıkça köpekler, atlar, kuşlar, vb. üyelikler içerir. Bununla beraber, denizyıldızları, bakteriler ve benzerleri hayvan sınıflandırmalarına göre müphem bir konumdadır.” (Zadeh, 1965)

Bulanık sistemlerin başlıca özellikleri arasındaki en önemli konu olarak çoklu girdileri, kural tabanı (Eğer-ise türünde yazılabilen) ve çıkarım motoru ile işleyerek tek çıktı haline dönüştürmesi gelmektedir. Bu sistemler kullanılarak mühendislikte görüntü işleme, zaman serileri esaslı tahmin yapmak, kontrol sorunlarını çözmek ve haberleşme yani iletişim konularında uygulamalar yapılmaktadır. Bunun dışında bulanık sistemler mühendislik, tıp, sosyoloji, psikoloji, işletme, uzman sistemler, yapay zeka, sinyal işlenmesi, ulaştırma, kavşak sinyalizasyonu gibi birçok alanda rahatlıkla kullanılabilir.(Şen, 2001, s.26.).

3.8. VERİ MADENCİLİĞİ MODELLEME ÇEŞİTLERİ

3.8.1. Sınıflama Modeli

En çok bilinen veri madenciliği tekniklerinden biri olan sınıflandırma, tahminleyici bir model olup özellikle risk analizi alanında çokça faydalanılan bir tekniktir. (Silahtaroglu, 2013, s.67.). Genel anlamda resim, ses ve örüntü tanıma, hastalık tanıları, dolandırıcılık tespiti, kalite kontrol çalışmaları ve pazarlama konularında bu yöntem sıkça başvurulmaktadır. Bu yöntemde en çok karar ağaçları, YSA ve Bayes sınıflandırıcılar gibi algoritmalar kullanılmaktadır. Kümelemeden farklı olarak denetimli bir öğrenme yöntemidir. (hangi topun hangi sepette olduğunu bildiğimiz gibi) Bu yöntemde her nesnenin sınıf etiketi olarak tanımlanan niteliğinin belirlediği bir sınıfta olduğu varsayılmaktadır.

Sınıflama analizinin şirketler açısından kullanıldığı alanlar ise şu şekildedir; (Alagöz ,2014, s.17).

- Kredi, sigorta işlemlerinde risklerin belirlenmesi
- Kredi verilecek kişilerin çeşitli özellikleri ve mali durumlarına göre risk sınıfının belirlenmesi sonucu müşteriye verilecek kredi tutarına karar verilmesi
- Yaş, meslek, eğitim, araba markası gibi bilgilere göre farklı kasko prim oranlarının uygulanmasının karar ağaçlarının kurulduğu risk faktörleriyle saptanması
- Ürünlerin geleceğe yönelik satış miktar ve fiyatlarının tahmin edilmesi
- Yeni bir ürünün satış imkanının incelenmesi (tahmin edilmesi)
- Ürünlerin satış miktarı, ciro, mevsimsel değişiklikler, şikâyet, arıza, memnuniyet oranı vb. faktörlerin değerlendirilmesi sonucu en iyi satan ürünün saptanması
- Bir ürünün bir bölgede satılma olasılığının incelenmesi
- Müşterinin risk sınıfının belirlenmesi ve risk sınıflarına göre müşterilere risk limiti atanması
- Ekibin eğitim seviyesi, tecrübesi, yaptığı projeler gibi özelliklerine göre sınıflandırılarak ücretlendirilmesi gibi örnek analizler sınıflama algoritmaları veya karar modelleri yoluyla gerçekleştirilebilir.

3.8.2. Lojistik Regresyon Modeli

Veri madenciliği çalışmalarında en çok kullanılan tekniklerden biri lojistik regresyondur. Tıp, ekonomi, meteoroloji, tarım ve veterinerlik alanlarında sıkça kullanılan lojistik regresyon analizinde amaç, kategorik bağımlı değişkenin değerini tahmin etmek olduğundan, aslında burada yapılmaya çalışılan iki yada daha fazla gruba yönelik “üyelik” tahminidir. Buna göre analizin amaçlarından birinin sınıflandırma, diğerinin ise bağımlı ve bağımsız değişkenler arasındaki ilişkileri araştırmak olduğu ifade edilmektedir.(Çokluk, 2010,s.6.).

Lojistik regresyon analizi uygulanırken takip edilmesi gereken adımlar şunlardır: (Keresteci, 2008, s.45.).

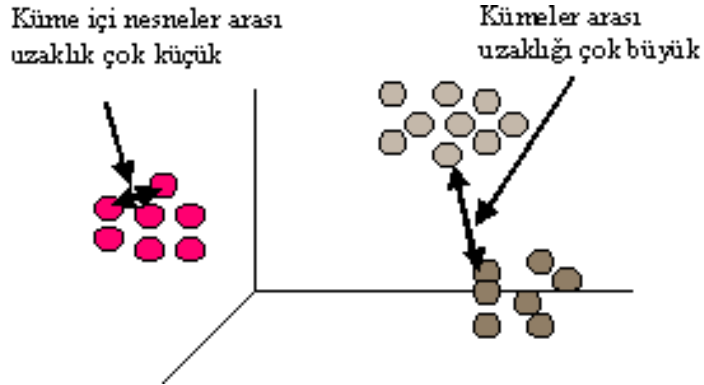
- Önsel grup üyelikleri belirlenmektedir.
- Modele girecek değişkenler belirlenmektedir. Bu amaçla önsel bilgidен yada istatistiksel tekniklerden yararlanılmaktadır.
- Modelin parametreleri Newton-Raphson yöntemi ile tahmin edilir. Ardından modelin tümünün anlamlılığı olabilirlik oranı ile test edilmektedir. Model anlamlı değilse analize son verilmektedir. Eğer model anlamlı bulunursa diğer aşamaya geçilmektedir.
- Tahmin edilen model parametrelerinin tek tek anlamlılığı incelenmektedir. Bu amaçla olabilirlik oranı ya da Wald istatistiği kullanılmaktadır. Her katsayının anlamlılığı incelendikten sonra, teklik oranları incelenerek, açıklayıcı değişkenlerin bağımlı değişken üzerindeki etkileri yorumlanmaktadır.
- Tahmin edilen model parametreleri kullanılarak, herbir gözlemin hangi gruptan geldiği tahmin edilmektedir.
- Modelin uyum iyiliğini incelemek amacıyla doğru sınıflandırma yüzdesi ve yapay R^2 ölçütleri kullanılır. Modelin uyum iyiliği kabul edilebilir düzeyde ise beşinci aşamadaki grup tahminleri kullanılabilir. Aksi halde ikinci aşamaya geçilerek modele girecek değişkenler yeniden gözden geçirilmekte ve işlemler tekrar edilmektedir.

İşletmenin bugünkü durumu ve gelecekte neler olabileceği konusunda tahminlerde bulunmak ve bu bilgileri sektör verileri ile karşılaştırmak için her dönem hızlı ve düşük maliyetle hazırlanan oran (rasyo) analizleri güncelliğini halen korumaktadır. Ancak oran analizinin kolay ve kullanışlı olmasının yanısıra, yöneteme ilişkin eleştiriler bulunmaktadır. Bu eleştirilerden birisi de bir çok araştırmacının savunduğu gibi, finansal analizlerde uygun yöntemin regresyon analizi olduğu yönündedir. Hatta Berry ve Nix (1991) yaptıkları çalışmada iki yöntem arasındaki sıralama korelasyonunda iki yöntem arasındaki farkın yıllar itibari ile belirgin olduğu

sonucuna varmışlardır. Bu farkın karar vericiler için önemli olabileceği vurgulanmıştır. (Karğın, 2010, s.49).

3.8.3. Kümeleme Modeli

Nesneleri birbirlerine olan benzerliklerine göre gruplara ayırma alanlarından biri olan kümeleme (clustering) yöntemi bilgisayar biliminde ses, karakter, resim tanınması, makina öğrenimi gibi alanlarda kullanılmaktadır. (Silahtaroglu, s.154) Makina öğrenimi açısından, herbir küme gizli bir örüntüyü temsil eder ve uygulanan öğrenme ise denetimsiz öğrenmedir. Sınıflandırma işleminde sınıflar önceden belli iken, kümelemede sınıflar önceden belli değildir. Verilerin sadece taşıdıkları değerlere göre değil diğer verilerle olan benzerliğe ve veri tabanındaki diğer verilerin durumuna göre de kümelere ayrılması kümelemeyi sınıflandırmadan ayırmaktadır ve ona dinamik bir yapı kazandırmaktadır. Şekil 5’de kümeleme modelleri gösterilmiştir.



Şekil 5: Kümeleme modeli

Kaynak: Alagöz, A., S. Öge ve M. Ortakarpuz (2014). Bir Kurumsal Zeka Teknolojisi Olarak Veri Madenciliği ile Muhasebe Bilgi Sistemi İlişkisi. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 2014. Vol.1. No.21, s.16.

Kümelemedeki öğrenmenin denetimsiz öğrenme olmasının nedeni önceden belirlenmiş sınıfların bulunmamasıdır. Zaten kümeleme yöntemini sınıflama yönteminden ayıran da budur. Önceden sınıflar belirli iken, yani kadın ve erkek diye iki ayrı sınıf varken yapılan algoritmik öğrenmeye denetimli öğrenme; herhangi bir sınıf

ismi verilmeden yapılan öğrenmeye ise denetimsiz öğrenme denir. Örneğin, veritabanındaki kayıtlarda her kaydın yanında kadın veya erkek bilgisi yazılıyor olsun, bu durumda veri tabanı üzerinde yapılan herhangi bir (kadın veya erkek olduğuna dair) kural çıkarma işlemi denetimli öğrenmedir. Ancak aynı veri tabanında, kayıtların yanında kadın mı erkek mi olduğu bilgisi yok iken yapılan kural çıkarma işlemi denetimsiz öğrenmedir. Bu işlem aynı zamanda veri tabanını iki kümeye ayırma, yani kümeleme işlemidir. Burada kadın/erkek gibi bir etiket yada sınıf olmayacağı için kümeleme kayıtlar arasındaki benzerlik veya mesafe ölçütüne göre yapılmaktadır. (Silahtaroglu, 2013, s.60.).

Kümeleme analizinin başarısına güçlü şekilde etki edecek yöntemlerden biri de algoritma kullanılmadan önce aykırı ve gürültülü verilerin temizlenmesi gerektiğidir. (Tan Steinbach Kumar, 2006, s.570.).

Kümeleme analizinin şirketler açısından kullanıldığı alanlar ise şu şekildedir; (Alagöz, 2014, s.16.).

- Bayii büyüklüklerine, müşteri sayılarına, cirosuna, rakip ürün pazar derinliği veya nüfuz durumuna göre kümeleme yapılabilir.
- Müşteriler, ciroları, sıklıkla aldıkları ürün grubu, ödeme karakteristikleri, yaş ve gelir gruplarına göre bölümlenebilir.
- Birden fazla satış noktası, satılan ürün çeşidi, ciro, müşteri profili vb. özelliklere göre tasnif edilebilir.
- İnsan kaynakları eğitim seviyesi, bir zaman biriminde yaptıkları iş miktarı, yapılan hata miktarı, oturdukları bölge gibi özellikleri kullanılarak belli gruplarda toplanarak analiz edilebilir.
- Ürün ya da hammaddeler için satış/sarf miktarı, ciro, mevsimsel dağılımları gibi özellikleri üzerinden tüketim oranlarına göre tasnifleyerek ABC analizi yapılabilir.

Kümeleme analizi sonuçları ile aşağıdaki konular araştırılabilmektedir. (Alagöz, 2014, s.16.).

- En iyi müşterilerin kimler olduđu.
- Farklı müşteri gruplarının neler olduđu ve bu grupların alışveriş örüntüleri hakkında bilgi.
- Müşterilerin satın alma davranışlarının neler olduđu.
- Hangi gelir grubunun hangi markayı satın almayı tercih ettiđi bilgisi.
- Müşteri sadakat derecesi ve bunun öngörülebilirliđi.

Bunlara benzer birçok konuda, müşteri, tedarikçi, çalışan boyutunda, veri madenciliđi yoluyla veri ambarından cevap aranabilmektedir.

3.8.4. Deđer Tahmin Modeli

Çevremizi kuşatan tüm dünyayı büyük bir veri laboratuvarı olarak düşünürsek, küçük bir çocuđun etrafını gözlemleyerek öğrenmesi, bir nesneyi ona benzeyen diđer unsurlarla ilişkisini kurup defalarca tecrübe etmesi, sınıflandırması, isimlendirmesi ve tekrar benzer bir nesneyle karşılaştığında onu koyacađı kategoriyi dođru bilmesi gibi, bu modelde insandaki tecrübe ederek öğrenme özelliđinden yola çıkılmıştır. (Silahtaroglu, 2013, s.50).

Denetimli (gözetimli) öğrenme, makina öğrenmesinde sınıflandırma veya tümevarımlı (inductive) öğrenme şeklinde ifade edilmektedir. Burada önemli olan nokta sınıf etiketlerinin de verilmiş olmasıdır. Veri madenciliđi algoritmaları veri tabanındaki nesnelerin özelliklerini nesnelerin isimleri ile ilişkilendirerek birbirlerinden farklı yada aynı sınıfta olduklarını bulmaktadır. İleriki aşamada deđişik özellikleri tanıyarak kendisine sorulan nesnenin ismini tahmin etmektedir. Denetimsiz öğrenmede ise nesnelerin özellikleri verilirken tahmin için kullanılacak herhangi bir parametre verilmez; yani nesnelerin isimleri verilmez. Denetimsiz öğrenmede amaç veriler arasındaki ilişki ve örüntülerin ortaya çıkartılmasıdır.

3.8.5. Veri Görselleştirme

Veri görselleştirme karar vericilerin ihtiyaçları doğrultusunda çok boyutlu veriler içersindeki karmaşık bağlantıların grafik araçları yada çizelge formatlarında sergilenmesini sağlamaktadır. Başarılı bir görselleştirme, verinin karakteristiği ve verinin parçaları ve nitelikleri arasındaki ilişkilerin analiz edilip raporlanabilmesi için görsel bir formata dönüştürülmesini gerektirir. Görselleştirmenin amacı görselleştirilen bilginin bir kişi tarafından yorumlanması ve zihinsel bir model olarak biçimlendirilmesidir. (Tan, Steinbach, Kumar, 2006, s.105.).

Hayatın her gününde, grafik yada tablolar gibi görsel teknikler, hava, ekonomi yada siyasi seçim sonuçlarını açıklamak için tercih edilen yöntemlerdir. Benzer şekilde, algoritmik yada matematiksel yaklaşımlar teknik disiplinlerin çoğunda-veri madenciliği de dahil görsel tekniklerin veri analizinde anahtar rol oynayabileceğine vurgu yapmaktadır.

Bu tekniğin kullanılmasında ağır basan motivasyon kişilerin büyük miktardaki görsel bilgiyi kolayca özümseyerek içindeki şekli bulabilmesidir. Bu tekniğin kullanılmasındaki bir diğer genel motivasyon insanların zihnindeki kilitli olan alan bilgisinden faydalanmaktır. Alan bilgisinin kullanımı veri madenciliğinin içinde önemli bir görev olurken, istatistiksel yada algoritmik araçların içindeki benzer bilgilerden tamamen istifade etmek sıklıkla zor yada imkansızdır. Bazı durumlarda, görsel olmayan araçlarla bir analiz yapılabilir ve sonrasında sonuçlar alanın uzmanı tarafından değerlendirme yapmak için görsel olarak sunulabilmektedir. Bir başka durumda, alan uzmanının görsel veriyi test etmesi bağlantıyı ortaya çıkaran şeklin bulunmasında en iyi yol olabilmektedir. Alan bilgisini kullanarak, bir kişi ilgisiz olan şekilleri yok varsayarak doğrudan önemli olan şekle odaklanabilmektedir.(Tan, Steinbach, Kumar, 2006, s.105.).

3.8.6. Birliktelik Kuralları

Birliktelik kuralları veri tabanlarındaki geniş bilgi yığınlarının içersinde yer alan veri parçalarının birlikte gerçekleşme (co-occurrence) bağıntısını açıklamaya çalışan en sık kullanılan yöntemlerdendir. Özellikle satış pazarlama alanında literatürde “Pazar

sepet analizi” olarak geen yntemden ska faydalanılmaktadır. Eėer X malını alanların Y malını da ok yksek olasılıkla aldıklarını biliyorsanız ve bir mşteri X malını alıyor ama Y malını almıyorsa o potansiyel bir Y mşterisidir. (Pektaş, 2013, s.107.).

Burada ama mşterilerin satın alma alışkanlıkları hakkında daha fazla bilgi edinmektir. Bir başka deyişle, “maėazadaki mşteri iin hangi grup yada set ğelerinin birlikte satın alması muhtemeldir?” sorusuna yanıt aramaktır. Sorunun cevabı maėazadaki mşteri işlemlerinin perakende verisi üzerinde gerekleştiren Pazar sepet analizinde saklıdır. Analiz sonrası ulaşılan sonu, katalog tasarımlarında olduėu kadar pazarlama ve reklam stratejilerinin planlanmasında da kullanılabilir. rneėin, Pazar sepet analizi farklı maėazaların dzenlenmesi sırasında yneticilere tasarım anlamında yardımcı olmaktadır.

Bir stratejiye gre, sk sk birlikte satılan rnler benzer rnlerin birarada satışı teşvik etmek iin raflara yan yana yerleştirebilmektedir. rneėin bilgisayar alan bir mşteri aynı anda finansal ynetim yazılımı almaya da eėilim gstermekte, bylece yazılım rnlerini donanım rnlerinin yanında sergilemek mevcut rnlerin her ikisinin de satışını ykseltebilmektedir. Alternatif bir strateji olarak; yazılım ve donanım rnlerini ayrı ayrı deponun karşı ularına yerleştirmek, benzer rnleri satın alan bir mşteriyi yol boyunca başka rnler de toplaması iin ikna ederek genel satışları arttırabilmektedir. rneėin; pahalı bir bilgisayar üzerinde bir mşteri satın alma kararını verdikten sonra, finansal ynetim yazılımı almak iin yazılım rnlerinin olduėu tarafa doėru giderken yolda gvenlik sistemi rnlerini gzlemleyebilmekte ve ev gvenlik sistemi de almaya karar verebilmektedir. (Han ve Kamber, 2001, s.227.).

Pazar sepet analizi ayrıca hangi rnlerin indirimli fiyatlarla satılması gerektiėi noktasında perakendecilere yardımcı olmaktadır. Eėer mşterilerin bilgisayar ve yazıcıyı birlikte satın alma eėilimi varsa yazıcı üzerindeki bir indirim bilgisayarların olduėu kadar yazıcı satışlarını da teşvik edecektir. (Han ve Kamber, 2001, s.227.).

Birliktelik analizinin şirketler açısından kullanıldığı alanlar ise şu şekildedir; (Alagöz, 2014, s.16).

- Müşteri satın alma eğilimlerinin bölgesel, kesim esasında, aylık, günlük ve hatta saatlik periyotlarda tespit edilmesi
- Farklı semtlerde bulunan birden fazla satış şubesi için farklı eğilimler tespit edilmesi ve mağaza esasında doğru stok politikaları geliştirilmesi
- Ürün kataloglarında birlikte satılan ürünlerin aynı sayfaya yerleştirilerek dikkat çekici hale getirilmesi
- Müşteri memnuniyet anketlerinde ve servis kayıtlarında ne tür müşterilerin hangi sıklıkla nelerden şikayet ettikleri ya da memnun olduklarının ortaya çıkarılması
- Birbirini izleyen ödeme ve gecikmelerdeki benzerliklerin belirlenebilmesi
- Üretimdeki insan kaynakları ile ürün kalitesi arasındaki bağıntıların sağlanması

3.9. VERİ MADENCİLİĞİNDE KARŞILAŞILAN PROBLEMLER

Büyük hacimli verilerin bulunduğu veri ortamlarından daha küçük veri kümelerine geçişte benzetim ortamlarında hazırlanmış veri madenciliği sistemleri büyük hacimli, eksik, gürültülü, boş, atık, aykırı veya belirsiz veri kümelerinin bulunduğu ortamlarda yanlış çalışabilmektedir. Problemden istenilen sonucu elde etmek için kullanılan örneklem kümesindeki gereksiz nitelikler “**Artık Veri**” olarak tanımlanır. Veri girişi veya veri toplanması esnasında kurulan sistem dışı hatalara “**Gürültü**” denir. Gürültülü veri aşırı uç değerler yada yanlış girilmiş değerler (mesela bir kişinin doğum tarihinin 1046 olarak girilmiş olması gibi) olabilmektedir. Belirsizlik, yanlışlıkların şiddeti ve verideki gürültünün derecesi ile ilgilidir. Bir veri tabanında, birincil anahtarda yer almayan herhangi bir niteliğin değerine “**Boş Değer**” denir. Boş değer, tanımlı gereği kendisi de dâhil olmak üzere hiçbir değere eşit olmayan değerdir. Veri tabanlarındaki eksik yada kaydı hiç girilmemiş veriler “**Kayıp Veriler(Missing Data)**” olarak tanımlanır. (Savaş, Topaloğlu, Yılmaz, 2012, s.6.).

Kayıp verilerin yaratacağı sorunları ortadan kaldırmak için aşağıdaki yöntemler uygulanabilmektedir; (Silahtaroglu, 2013, s.23.).

- Kayıp verilerin bulunduğu kaydın veri tabanından-veriler kümesinden çıkarılması yada bu gibi kayıtların iptal edilmesi
- Kayıp verilerin elle teker teker doldurulması
- Tüm kayıp verilere aynı bilginin girilmesi
- Kayıp olan verilere tüm verilerin ortalama değerinin verilmesi
- Diğer değişkenlerin yardımı ile kayıp olan verilerin tahmin edilmesi

Bunların dışında, veri tabanının boyutu, kurumsal çevrimiçi veri tabanlarının yapısının dinamik olması ve içeriğinin sürekli değişiyor olması, sınırlı ve eksik veriler, sembolik ve kategorik veri türleri dışında kalan tamsayı, kesirli sayılar, çoklu ortam verisi, coğrafi bilgi içeren veri gibi farklı tipteki veriler üzerinde işlem yapılması gibi faktörler de veri madenciliği çalışmasının başarısını etkilemektedir. (Savaş, Topaloğlu, Yılmaz, 2012, s.6.).

Veri madenciliği çalışması sırasında kullanılan algoritmalarla yapılan hesaplamaların oldukça karmaşık ve fazla sayıda olması uygulamada bilgisayar kullanımını zorunlu kılmıştır. Veri madenciliğinin ilk ortaya çıktığı 1990'lı yıllarda bilgisayarlar çok pahalı olduğu için veri madenciliği kullanımının yaygınlaşması gecikmiştir. Günümüzde bilgisayar ve bilgisayar programlarının fiyatları kabul edilebilir seviyeye inmiştir. Yine de bu konu veri madenciliği uygulamaları için bir kısıt oluşturmaktadır. Uygulamada karşılaşılan bir diğer zorluk ise hem veri madenciliği uzmanlığı olan hem de üzerinde çalışılan sahada uzman olarak yetişmiş insan gücünü aynı anda bulma zorunluluğudur. (Özkan, 2014, s.17.).

4. VERİ MADENCİLİĞİNDE KÜMELEME MODELİ KULLANILARAK MÜŞTERİ SEGMENTASYONU BELİRLENMESİNE DAİR BİR UYGULAMA ÖRNEĞİ

4.1. ÇALIŞMANIN AMACI

Gerçekleştirilen bu çalışmada 2015 yıl sonu itibariyle 600 tane müşterisi olan XYZ* Kalibrasyon Ve Danışmanlık Hizmet Sanayi Tic. Ltd. Şti. firmasının 2013, 2014 ve 2015 yılları itibariyle her üç senede hizmet verdikleri 251 tane müşterisi seçilmiş, devamlı müşterilerini tanımak ve segmente etmek amaçlanmıştır. Araç olarak veri madenciliği tekniklerinin kullanılmasının firmaya iç kontrol açısından ne gibi katkıları olabileceği ve firmanın bilanço hesapları üzerinde hangi etkileri yapabileceği araştırılmıştır.

4.2. SEÇİLEN MODEL VE ALGORİTMA

Çalışmada seçilen model kümeleme analizi ve kullanılan teknik K-means algoritmasıdır. Veri setinin dönüştürülmesi sırasında kullanılan yazılım açık kaynak kodlu GNU lisanslı bir yazılım programı olan Weka programıdır.

4.2.1. Kümeleme Modeli

Kümeleme analizinin amacı kayıtlı veriler üzerindeki birbirine benzer özellikler gösteren verilerin aynı grup içerisinde toplanarak aidiyet kazanacağı birbirine benzemeyen iki farklı küme çıkarmaktır. Bu şekilde çıkan her grup küme olarak adlandırılmakta ve her kayıt tamamen bir kümeye ait olmaktadır. (Ramakrishnan, Gehrke, 2003, s.911).

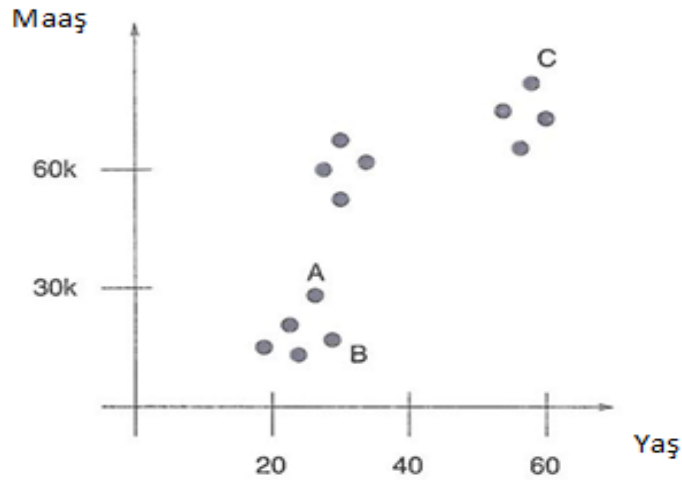
Kayıtlar arasındaki benzerlik mesafe fonksiyonu ile hesaplanarak ölçülmektedir. Bir mesafe fonksiyonu iki kayıt girdisini almakta ve benzerliklerinin ölçüm değerine dönüştürmektedir. Farklı uygulamalar farklı benzerlik kavramlarına

* XYZ Kalibrasyon ve Danışmanlık Hizmetleri Sanayi Tic. Ltd. Şti. Firması 2002 yılında kurulmuş ve 2015 sonu itibariyle 13 çalışanı ile özel sektörde faaliyet göstermektedir.

sahiptir ve hiç bir ölçü bütün alanlar için geçerli değildir. Kümeleme algoritması veri tabanını belirli altkümelere ayırırken birimler arası uzaklıklar için Euclidyen, Standardize Euclidyen, Manhattan Mahalanobis, Kareli Euclidyen, Minkowski veya Canberra gibi alternatif ölçü ve yöntemler kullanmaktadır. (Sarıman, 2011, s.193).

Örnek olarak; şirket bilgilerini yaş ve maaş olarak aldığımızda kayıtları iki boyutlu bir düzlemde Şekil 6'da olduğu gibi incelenebilmektedir. Buna göre yaş ve maaşlar değerlerinden iki düzende kayıtlı değerleri görselleştirebilmekte ve şöyle üç tanımlama yapabilmektedir: Düşük maaşa sahip genç müşteriler, yüksek maaşa sahip genç müşteriler, ve yüksek maaşa sahip yaşlı müşteriler.

Genellikle, kümeleme algoritmasının çıktıları her kümenin özet bir temsilinden meydana gelmektedir. Özet temsil tipleri güçlü bir şekilde bilgisayarla hesaplanan algoritmanın tipine bağlıdır. Örneğin, Şekil 6'da gösterildiği gibi dairesel kümelere sahip olduğunu varsayalım. Aşağıda da gösterildiği gibi, her küme merkezi(sıkça ortalama olarak da geçen) ve yarıçapı tarafından özetlenebilmektedir. (Ramakrishnan, Gehrke, 2003, s.911).



Şekil 6: Kümeleme algoritması

Kaynak: Ramakrishnan, Gehrke, 2003, s.911

Kümeleme analizi ve algoritmaları beş ana başlık altında incelenebilir. Bunlar aşağıda belirtilmiştir: (Silahtaroglu, 2013, s.163).

1. Hiyerarşik yöntemler (SLINK Algoritması, CURE Algoritması, CHAMELEON Algoritması, BIRC Algoritması)
2. Bölümlemeli yöntemler (K-means Algoritması, PAM Algoritması, CLARA Algoritması, CLARANS Algoritması)
3. Yoğunluğa dayalı yöntemler (DBSCAN Algoritması, OPTICS Algoritması, DENCLUE Algoritması)
4. Grid temelli algoritmalar (STING Algoritması, Dalga Kümeleme, CLINQUE Algoritması)
5. Kategorik Verinin Yinelenmesine Dayalı Yöntemler
6. Kısıtlara Dayanan Yöntemler
7. Makine Öğrenmesi Alanında Kullanılan Yöntemler

Gerçekleştirilen segmentasyon analizi çalışmasında kullanılan K-means Algoritması Bölümlemeli Yöntemler içerisinde yer almaktadır. Bölümlemeli kümeleme algoritmalarının çalışma şekli k giriş parametresini alarak n tane nesneyi k tane kümeyle bölerek içiçe bir kümeleme yapısı üzerinde çalışmak yerine tek seviyeli kümeleri bulan işlemler gerçekleştirmektedir. (Sarıman, 2011, s.195.). Bu teknik merkez noktanın kümeyi temsil etmesi esasına dayanmaktadır. Bölümlemeli yöntemlerin uygulamalarının kolay ve verimli olması nedeniyle ve iyi sonuçlar üretebiliyor olmasıyla birlikte K-means algoritmasının en yaygın kullanılan gözetimsiz öğrenme yöntemlerinden biri olduğu bilinmektedir. (Sarıman, 2011, s.195).

4.2.2. K-means Algoritması

En iyi bilinen ve yaygın olarak kullanılan bölümlemeli yöntemlerden olan k-means sürekli olarak kümelerin yenilendiği ve en uygun çözüme ulaşana kadar devam eden döngüsel bir algoritmadır. (Silahtaroglu, 2013, s.186.).

Tarihsel açıdan baktığımızda K-means terimi ilk olarak 1967’de James Mac Queen tarafından kullanılsada algoritmanın ana fikri 1956’da Hugo Steinhaus tarafından ortaya atılmıştır. Algoritmanın diğer algoritmalara göre avantajlı yönü her çeşit veride kabul edilebilir sonuçlar vermesi, dezavantajlı yönü yerel optimumlarda kalarak genel optimumlara ulaşamamasıdır.(Yünel, 2010, s.2).

Çalışmamızda da ağırlık merkezli bir teknik olarak K-means algoritması kullanılmıştır. K-means algoritması n adet veri nesnesinden oluşan bir veri setini girdi parametresi olarak k’yı almakta ve k tane küme içinde kümeiçi benzerlikleri yüksek fakat kümeler arası benzerlikleri düşük olarak sonuçlandırmak için nesnelere kümelere ayırmaktadır. Küme benzerliği kümenin ağırlık merkezi olarak görülebilen içindeki nesnelere ortalama değeriyle ilgili olarak ölçülmektedir. (Han, Kamber, 2001, s.349). Kullanılan matematik yöntem her sınıf için merkez belirlenen noktaya uzaklığa (aynı zamanda bu hata miktarıdır) göre yeni kümelerin yerleştirilmesidir. K-means’ın atama mekanizması her verinin sadece bir kümeye ait olabilmesine izin verir, bu nedenle keskin (bölümlemeli) bir kümeleme algoritmasıdır.

K-means algoritmasının işlem basamakları şöyledir: (Molina,Ullman,Wildom, 2009, s.131).

1. K başlangıç kümeleri seçilerek başlanmaktadır. Bu kümeler tek noktalı, yada küçük nokta setlerinden meydana gelmektedir.
2. Her atanmamış nokta, “en yakın” kümeye yerleştirilmektedir.
3. İsteğe bağlı olarak, bütün noktalar kümelere atandıktan sonra, her kümenin merkezi sabitlenmektedir. Sonra bütün noktalar bu k kümelerine tekrar atanmaktadır. Bazen, ilk atanan noktalar tekrar atanmak durumunda olabilmekte ve böylece başka bir kümede sabitlenmektedir.

4.3. VERİNİN HAZIRLANMASI, TEMİZLENMESİ VE UYUMLAŞTIRILMASI

XYZ Kalibrasyon Ve Danışmanlık. Hizmet. Sanayi .Tic. Ltd Şti firmasının 2013, 2014 ve 2015 yıllarına ait müşteri ismi, adresi, faaliyet gösterdiği sektör, eski/yeni müşteri olup olmadığı, Kobi/büyük firma bilgisi, yıllık kalibrasyon yaptırma periyodu ve 3 yıllık satış rakamları Excel ortamında alınmıştır. Veri setinin içindeki veriler kategorik ve numerik değerlerdir. Kategorik veriler için eksiklik olmamasına rağmen, numerik değerlerdeki eksiklikler, diğer bütün değerlerin ortalaması alınarak temizlenmiştir. Dönüştürme işlemleri olarak da numerik veriler belirli aralıktaki gruplara bölünürken, kategorik veriler için 1,2,3... gibi değerler girilmiştir.

Kalibrasyon şirketinin müşterilerine ait elimizdeki veri setinde gerekli temizleme ve dönüşüm işlemleri gerçekleştirildikten sonra 6 tane nitelik ortaya çıkmıştır. Bunların isimleri ve özellikleri ise aşağıda detaylı olarak açıklanmıştır.

- 1- Adres Sınıfı: Müşteri adreslerinin kayıtlı olduğu semtler belirlenmiş ve sonrasında Anadolu Yakası, Avrupa Yakası, İstanbul dışı olarak üç ana grupta tanımlanmışlardır. Adres sınıfı olarak Anadolu Yakası 1, Avrupa Yakası 2, İstanbul dışı 3 olarak kodlanmıştır.
- 2- Müşteri Tipi: 2015 itibariyle son 5 yılda firmanın kayıtlarına giren müşteriler yeni, diğerleri eski müşteri olarak E ve Y şeklinde kodlanmıştır.
- 3- Sektörü: Müşterilerin faaliyet gösterdikleri sektörler tek tek belirlenmiş ve sonrasında 10 ana sektör sınıfı olarak sınıflandırılmış ve şu şekilde kodlanmıştır: Ambalaj 1, Gıda 2, Hizmet 3, Kimya ve Boya 4, Laboratuvar 5, Makina 6, Mühendislik ve İnşaat 7, Otomotiv 8, Üretim 9, Diğer 10.
- 4- Kobi/Büyük Firma ayrımı: Müşteriler Kobi ise 2, büyük firma kategorisindeyse 1 olarak kodlanmıştır.
- 5- Yıllık kalibrasyon yaptırma periyodu: Numerik veride oluşan eksiklik sektör ortalamaları girilerek temizlenmiştir. Bu değerler gıda, üretim ve diğer sektörler için 5, mühendislik ve inşaat için 1, makina, hizmet, otomotiv, laboratuvar, kimya ve boya ve ambalaj için 2 olarak belirlenmiştir.

- 6- Satış tutarları: 1000TL'nin altındaki tutarlar 0-1000 aralığı, 5000TL'nin altındaki tutarlar 1000-5000 aralığı, 10000TL'nin altındaki tutarlar 5000-10000 aralığı ve 10000TL'nin üzerindeki tutarlar 10000+ olarak kodlanmıştır.

4.4. MODELLEME

Firmanın 251 adet devamlı müşterisine segmentasyon analizi yapmak amaçlanmaktadır ve bu amaç için denetimsiz bir model olarak kümeleme analizi gerçekleştirilmiştir.

Kümeleme algoritmalarından K-means algoritması ve WEKA yazılımı kullanılarak kümeleme işlemine sokulup müşteri segmentasyonu belirlenmiştir.

WEKA açık kaynak kodlu GNU lisanslı bir yazılımdır. Waikato Environment for Knowledge Analysis kelimesinin kısaltılmasından adını alan bu yazılım Waikato Üniversitesinde Java platformu üzerinde geliştirilmiştir. Makina öğrenmesi paketleri açısından ve komut satırı arayüzü bağlantısı açısından en güçlü yazılımdır. Kullanım ve öğrenim kolaylığı programın bir diğer avantajıdır. (Kaya ve Özel, s.4) Yazılımın çalışabilmesi için Java 1.4 bilgisayara yüklenmiştir. WEKA yazılımında ARFF (Attribute-Relation File Format) dosyaları kullanılmaktadır. ARFF dosyaları, değişken tanımlanmasına izin veren ASCII metin dosyalarıdır. ARFF dosyasının başlık kısmında, değişkenler(veri tabanındaki her bir kolonun ismi), bunlar arasındaki ilişkiler ve herbir değişkenin türü ve alacağı değer vs. bulunmaktadır. Veriler @DATA satırından sonra gelmektedir. Şekil 7'de bir ekran görüntüsü verilmiştir.

```

@relation weka
@attribute ADRES {1,3,2}
@attribute ESKI_YENI {E,Y}
@attribute SEKTOR_SINIF {2,7,3,8,9,10,4,5,6,1}
@attribute KOBI_BUYUK {2,1}
@attribute KLBRYSN_PERIOD {5,1,2}
@attribute TUTAR_ARALIGI_15 {0-1000,1000-5000,5000-10000,10000+}
@attribute TUTAR_ARALIGI_14 {0-1000,1000-5000,5000-10000,10000+}
@attribute TUTAR_ARALIGI_13 {1000-5000,0-1000,10000+,5000-10000}

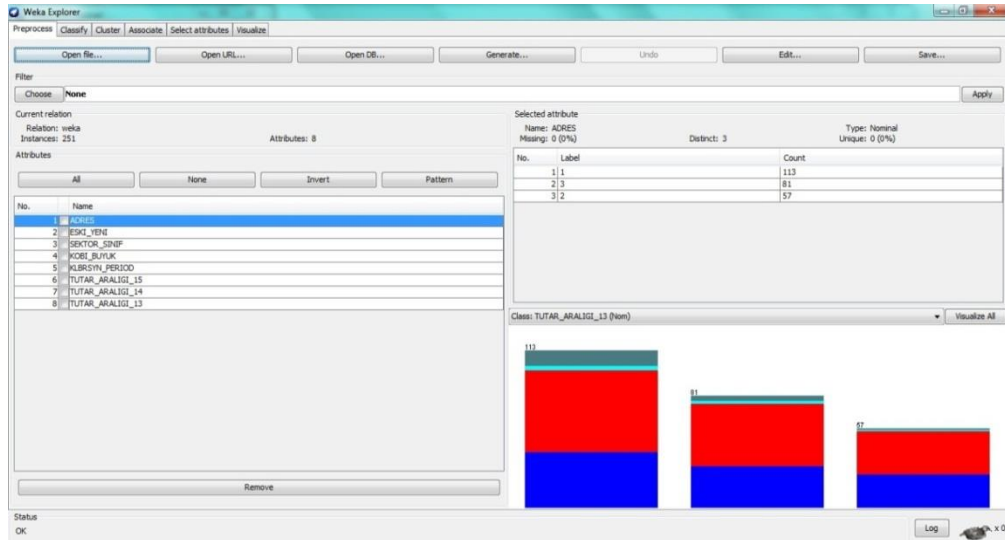
@data
1,E,2,2,5,0-1000,0-1000,1000-5000
3,E,7,2,1,0-1000,0-1000,0-1000
2,E,2,2,5,0-1000,1000-5000,0-1000
2,E,2,1,5,0-1000,0-1000,0-1000
1,E,3,1,5,0-1000,0-1000,10000+
1,E,8,1,2,1000-5000,1000-5000,1000-5000
3,E,9,2,5,0-1000,0-1000,0-1000

```

Şekil 7: ARFF formatına dönüştürülmüş ekran görüntüsü

Kaynak: <https://sourceforge.net/projects/exceltoarffconv/>

ARFF formatına dönüştürülen Excel dosyasının WEKA'ya yüklendiğindeki ekran görüntüsü Şekil 8'de gösterilmiştir.



Şekil 8: WEKA Giriş Ekranı

Kaynak: <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>

Yukarıdaki ekran görüntüsünde veri madenciliği bölümlerinden “cluster”(kümeleme) sekmesi müşteri segmentasyonu çalışması için veri setine en uygun başlık olarak seçilmiştir. “Cluster” sekmesi içinde Simple K-means algoritması seçilmiştir. (K küme sayısı 2 olarak belirlenmiştir.)

Çalışma sırasında K-means dışındaki programda yer alan diğer algoritmalar (Cobweb, ExpectationMaximization, Farthestfirst, MakeDensityBasedClusterer) da test edilmiş ancak en anlamlı sonuç üreten algoritma olduğu için K-means tercih edilmiştir.

4.5. RAPOR SONUÇLARININ DEĞERLENDİRİLMESİ

Şekil 9 ve 10’da WEKA yazılımının eldeki verilere göre iki ekran görüntüsü verilmiştir.

```
=== Run information ===  
  
Scheme:weka.clusterers.SimpleKMeans -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -S 10  
Relation: weka  
Instances: 251  
Attributes: 8  
    ADRES  
    ESKI_YENI  
    SEKTOR_SINIF  
    MOBIL_BUYUK  
    KILBIRSYN_PERIOD  
    TUTAR_ARALIGI_15  
    TUTAR_ARALIGI_14  
    TUTAR_ARALIGI_13  
Test mode:evaluate on training data  
  
=== Model and evaluation on training set ===
```

Şekil 9: Weka Sonuç Ekranı 1

Kaynak: <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>

```
Number of iterations: 4
Within cluster sum of squared errors: 709.0
Missing values globally replaced with mean/mode

Cluster centroids:

Attribute          Full Data          Cluster#
                   (251)             0             1
                   (155)             (96)
-----
ADRES                1                 1             1
ESKI_YENI            E                 E             E
SEKTOR_SINIF        9                 9             4
KOBI_BUYUK           2                 2             2
KLBRSYN_PERIOD      5                 5             2
TUTAR_ARALIGI_15    0-1000            0-1000        1000-5000
TUTAR_ARALIGI_14    0-1000            0-1000        1000-5000
TUTAR_ARALIGI_13    0-1000            0-1000        1000-5000

Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      155 ( 62%)
1       96 ( 38%)
```

Şekil 10: Weka Sonuç Ekranı 2

Kaynak: <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>

Yukarıdaki ekran görüntülerinden de incelenebileceği gibi Weka programı kümeleme modeli k-means algoritmasının ürettiği sonuçlar şöyle yorumlanmaktadır: K-means algoritması dört atama gerçekleştirerek 251 adet müşteri verisi kümesini almış ve 155 verinin ait olduğu %62'lik ve 96 verinin ait olduğu % 38'lik iki kümeye ayırmıştır. Birbirinden farklı bu iki gruptan %62'lik olana bakıldığında müşterilerin Anadolu Yakasında oldukları, eski müşteri oldukları, üretim sektöründe faaliyet gösterdikleri, Kobi sınıfına girdikleri, yılda 5 kez kalibrasyon yaptırdukları ve üç yılda da (2013, 2014 ve 2015) 1000TL'nin altında satış tutarı belirlenmiştir. %38'lik ikinci küme incelendiğinde, müşterilerin yine Anadolu Yakasında ve eski müşteri oldukları, Kimya ve boya sektöründe faaliyet gösterdikleri, Kobi sınıfına girdikleri, yılda 2 kez kalibrasyon yaptırdukları ve üç yılda da (2013, 2014 ve 2015) 1000-5000TL aralığında satış tutarı görülmüştür.

5. İÇ KONTROL SİSTEMİ ÇERÇEVESİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİ KULLANILMASI VE BİLANÇO ÜZERİNDEKİ ETKİLERİ

5.1. İÇ KONTROL SİSTEMİ ÇERÇEVESİNDE BİR ARAÇ OLARAK VERİ MADENCİLİĞİ KULLANIMININ ÖNEMİ

Günümüzde bilişim teknolojileri yeni ekonomi ve yönetim modellerinin etkili aracı haline gelmiştir. Veri madenciliği hem karmaşıklaşan veri yığınlarının yönetilebilmesi açısından hem de karar alıcılar için geleceğe dönük tahmin politikaları yönüyle önemli bir araçtır.

İç kontrol sistem kurulumu değişimi yönetmek ve riskleri fırsata çevirmek açısından çok önemlidir. Ulrich Beck 1992 yılında yayımlanan “Risk Society: Towards a New Modernity” adlı kitabında risk toplumu kavramına vurgu yapmıştır. Beck bu eserinde, 19. Yüzyılda “ya, yada” felsefesinin egemen olduğunu, oysa 20. Yüzyılda “ve” felsefesinin geçerli olduğunu vurgulamaktadır. Buna göre; 19. Yüzyılda birbirinden kopuk yapılar yada birimler vardı. Kopukluk, tekdüzelik, uzmanlaşma, hareketsizlik gibi ifadelerle kendini bulan bu yapı, taraflar arasındaki ayrımları ve çıkmazları açıklamaktadır. 20. Yüzyılda ise “ve” nin egemenliği; yan yana olma, çok boyutluluk, sentez, bulanıklık, belirsizlik gibi kavramları ortaya çıkartmıştır. (Özer, 2010, s.214.).

Toplumsal süreçte yaşanan bu değişim firmalar ölçeğine de yansımış ve pek çok kuruluş bu süreci yönetebilir hale geçmek için Kurumsal Kaynak Planlama (ERP- Enterprise Resource Planning) yazılımlarına ve bütüncül kontrol sistemleri geliştirmeye ihtiyaç duymuştur.

Veri madenciliği kurumsal yazılım projelerinde gelişen teknolojiler içerisinde yer almaktadır. İşletmeyi başarıya götürecek en önemli faktör stratejik hizalama denilen iş stratejisi ile bilişim sistemi arasında sağlanacak uyumdur. Veri madenciliği algoritmaları pek çok kuruluşta işletmenin tüm bölümlerini yöneten kurumsal kaynak planlama yazılımlarına başlangıç kuruluş noktasında ve iç kontrol kriterleri de göz önünde bulundurularak yerleştirilmelidir. ERP sistemlerinin başlangıçta sadece

operasyonel olarak kurulması belki kuruluşun o zamanki ihtiyaçları doğrultusunda yeterli olabilmekte ancak kontrol aktivitelerinin sistem içerisine entegre edilmemesi zamanla öngörülemeyen risklerin ortaya çıkmasına, hata hile suistimal gibi insan faktöründen kaynaklanan zararların artışına sebep olmaktadır. Ayrıca şirketin iç kontrol aktivitelerini sisteme sonradan entegre etme çabası yönetimi ek maliyetlere ve gerçekleşen risklerin sonuçlarına katlanılması gibi durumlarla karşı karşıya bırakmaktadır. Bununla beraber, ERP sistemlerinde üretilen tahmine dayalı modellemelerin özünde istatistik ve zaman serileri bulunması “istatistikçi dün yaptığı tahminlerin bugün neden çıkmadığını yarın bilebilecek kişidir.” ifadesini doğrulamaktadır. (Sümen, 2014, s.56). Başka bir deyişle kirli veri, pazar değişimi, rakiplerin tutumu, ekonomik durum, sosyal olaylar, hatta hava durumu gibi etmenler zaman serileri ile yapılan tahmini bozabilmektedir. Ayrıca, talep normal dağılıma uygun gerçekleşmemiş de olabilmektedir. Bu durum karar alıcılar açısından iç kontrol çerçevesinde belirsizlik riskinin iyi yönetilememesine ve yönetim risklerine yol açmaktadır. Yapılması gereken ERP araçları ile gerçekleştirilen tahminlerin nitel tahminler ile düzeltilmesidir.

5.2. UYGULAMA ÖRNEĞİ SEGMENTASYON ANALİZİ SONUÇLARININ İÇ KONTROL ÇERÇEVESİNDE DEĞERLENDİRİLMESİ

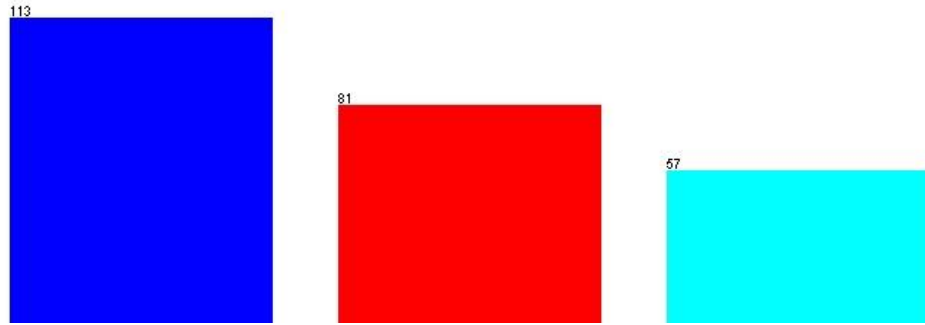
Kimya ve boya sektörlerinde ciro ve potansiyelin yüksek olmasından dolayı bu sektörler kapsamında piyasadaki diğer müşterileri tanıma ve firmanın kendi müşteri portföyüne dahil etmek için çalışması firmaya satışlarında artış ve karlılık getirebilecektir. Ayrıca her iki kümedeki müşterilerin yapısına bakıldığında, eski müşteriler oldukları görülmektedir. Bu durum müşteri sadakati açısından olumlu olarak değerlendirilebilecek olsa da yeni müşterilerin de kalıcılığını sağlamak için doğru bir satış stratejisi belirlemek ve işletme ortamı, yönetim felsefesi ve iç kontrolün başarısı açısından önem taşımaktadır.

Her iki kümede de İstanbul Anadolu Yakası iş potansiyelinin hakim olduğu görülmektedir. Bu durum iç kontrol ve risk değerlendirme açısından yoğunlaşma riskini de beraberinde getirmektedir. Yoğunlaşma riski çoğu müşterinin aynı endüstri kolu veya coğrafi bölgede olmasının o coğrafi bölge yada endüstrideki olası bir ekonomik

durgunluğun iç kontrol riskini de arttıracığına işaret etmektedir. Böyle bir durumda Anadolu Yakasındaki müşteri portföyü iyi yönetilmeli, Avrupa Yakası ve diğer bölgelere satış desteği verilerek riskten korunma yoluna gidilmelidir. Müşterilerin Weka ekranının verinin görselleştirme (visualize) modeli çalıştırılarak genel bölgesel dağılımını görmek amacıyla Şekil 11’de sütun grafikleri ve Şekil 12’de kümeleme yoğunlaşma görseli yer almaktadır.

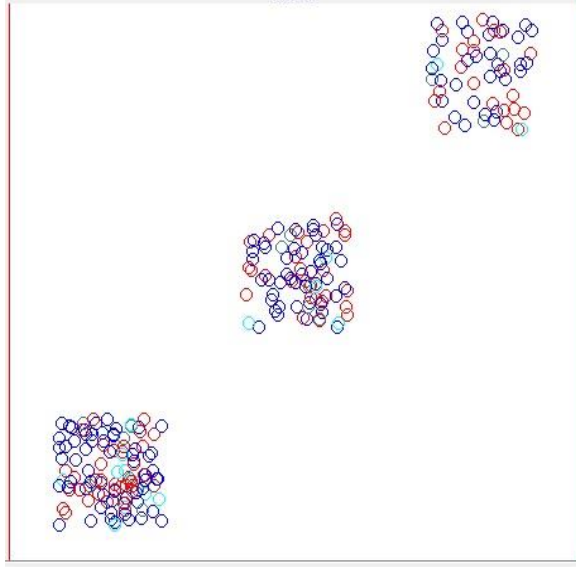
Selected attribute		
Name: ADRES		Type: Nominal
Missing: 0 (0%)		Distinct: 3
		Unique: 0 (0%)
No.	Label	Count
1	1	113
2	3	81
3	2	57

Class: ADRES (Nom) Visualize All



Şekil 11: WEKA Müşteri Adres Dağılım Görselleştirme Ekranı

Kaynak: <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>



Şekil 12: WEKA Müşteri Adres Dağılım Küme Yoğunlaşması

Kaynak: <http://www.cs.waikato.ac.nz/ml/weka/downloading.html>

Bir diğer sonuç, firmanın hizmet verdiği sektörlerin büyük firmalardan ziyade Kobi olarak yoğunlaşmasıdır. Bu durumda rekabet açısından iş riskini arttıran bir unsurdur. İç kontrol açısından etkinliği arttıracak bir satış stratejisi olması açısından büyük firmaların kendi kalibrasyonlarını yaptırmaları yerine işbirliği yapılarak özel laboratuarlara iş aktarımı yönünde çalışmalar yapılması fayda sağlayacaktır. Böylece firma müşteri portföyünü de çeşitlendirmiş olacaktır.

Üretim sektöründeki firmalarda yıllık kalibrasyon yaptırma periyodunun 5, boya ve kimya sektöründe ise 2 olduğu görülmektedir. İşletme sektöründeki firmalara yıllık kalibrasyon hizmeti verme periyodlarını gözlemleyerek teknolojisini bu yönde iyileştirmeli ve geliştirmelidir. Teknolojiyi ortaya çıkaran (teknolojiyi kendisi için geliştiren) bir işletme için teknoloji izleyicisi konumuna göre daha az belirsizlik riski bulunacaktır.

2015 yılı sonu itibarıyla firmanın toplam 597 müşterisi olmasına rağmen analize firmanın devamlı müşterilerinin yapısını tanımak amacıyla her 3 yılda da satış yapılan toplam firma sayısı 251 ile başlanması firmanın toplam müşterilerinin her yıl yaklaşık 1/3' ne düzenli hizmet verdiği ve kalan 2/3'lük müşteri potansiyeline iş yapılamadığını görmemizi sağlamaktadır. Burada 2/3'lük kaçırılan bir potansiyel

bulunmaktadır. İç kontrolün etkinliği açısından bilgi ve iletişim kanalları çalıştırılarak bu durum anketlerle değerlendirilip yıllara göre değişkenlik gösteren müşteriler kazanılmaya ve devamlı müşteri haline getirilmeye çalışılmalıdır.

XYZ Kalibrasyon Ve Danışmanlık. Hizmet. Sanayi .Tic. Ltd Şti firmasının muhasebe bölümünden sağlanan net ciro rakamları ve müşteri sayıları ile TÜİK(Türkiye İstatistik Kurumu) verilerine göre 3 yıllık rakamları Tablo 3’de verilmiştir.

Tablo 3.

Yıllara Göre Firmanın Ciro Müşteri Sayısı Yıllık Tüfe Oranı

Yıllık Ciro (TL)	Müşteri Sayısı	Yıllık Tüfe Oranı(%)
2013 yıl sonu: 412.215	596	7,40
2014 yıl sonu: 458.680	546	8,17
2015 yıl sonu: 442.234	597	8,81

Kaynak: <http://www.kpmgvergi.com/PratikBilgiler/Pages/Ufe-Tufe-Oranlari.aspx>

Firmanın hizmet sektöründe faaliyet gösteriyor olması, satışlarını genellikle yerli firmalara TL esasında yapmasının doğal bir sonucu olarak herhangi bir hammadde alış ve buna bağlı kur riski de bulunmamaktadır. Dolayısıyla veriler analiz edilirken en önemli dış değişken enflasyon olarak düşünülmektedir. 2013 yılından 2014 yılına cirodaki artış oranı %11 olarak gerçekleşmiştir. Firmanın üç yıllık cirosunu yıllık TÜFE oranları ile karşılaştırdığımızda, firmanın cirosundaki artış 2014’de TÜFE’nin çok az üzerinde gözükmemektedir. 2015’de ise firmanın cirosunda %3 oranında azalma vardır. Burada müşteri sayısındaki artışın ciroya orantılı olarak yansımadağı görülmektedir. Bunun sebebi, 2015 yılında daha çok sayıda müşteriye daha küçük rakam ve sıklıklarla satış yapılmış olmasıdır. Bu sonuç da göstermektedir ki, XYZ firması satış stratejisini büyük ve kurumsal firmalardan daha fazla müşteri kazanabilecek şekilde düzenlemelidir. Böylece iç kontrol açısından daha fazla etkinlik sağlanmış olacaktır. Sektörde fazla sayıda hizmet veren kalibrasyon firması olması rekabet riskini de beraberinde getirmektedir. 2015 yılında sektör genelinde rekabete dayalı birim

fiyatlarda düşüş gerçekleşmesi de 2015 cirosunun azalışında bir etkidir. Bu durum da iç kontrol açısından piyasa riskini ortaya çıkarmaktadır.

XYZ firması için gerçekleştirilen segmentasyon analizi sonucunda firmanın satış stratejisini yeniden düzenleyerek müşteri ve satış hacmini arttırması gerekliliği görülmektedir. Ancak bu süreçte firmanın COSO'ya göre iç kontrol sisteminin ana unsurlardan biri olan kontrol çevresi başlığı altında yer alan iş yapış tarzı ve etik unsurlar konusunda dikkat gerektiği gözardı edilmemelidir.

5.3. ETİK STANDARTLAR

5.3.1. Etik Kavramının Tanımlanması

Etik, bireysel kullanımı kişilerin davranışlarının yönetimine verilmiş ahlaki prensipler olarak tanımlanmaktadır. Kısaca, “doğru” yada “yanlış” şeklinde bireysel ayrımlar yoluyla ortaya çıkan kişisel kriterlerdir. Açıkçası, bir toplum etik olmadan yaşamak için kaotik ve tehlikeli bir yer olacaktır. Bu sebeple hükümetler, dinler ve eğitimciler belirli etik standartların tüm toplum üyeleri arasında benimsenmesi ve desteklenmesini sağlamıştır. (Meigs, 1997, s.679).

İşletmelerin ticari kuruluşlar olarak satışlarını arttırıp kâra geçerken aynı zamanda sorumlu organizasyon yapısına uygun hareket etmeleri gerekmektedir. Bunun anlamı; başkalarına hizmet etmenin sorumluluğunu kucaklamak, yanlışları ve başarısızlıkları kabul etmektir. Her organizasyon en az bir topluluk içinde varlığını sürdürmektedir: İster köşedeki bakkal dükkanını çevreleyen mahalle olsun, isterse içinde büyük bir çok uluslu şirketin faaliyet gösterdiği dünya topluluğu olsun. Her bireyin işiyle ev sahipliği yaptığı topluluğa ne kadar iyi hizmet ettiğini sorgulaması gerekmektedir. (Lennick, Kiel, 2005, s.166-209).

5.3.2. Satış Stratejileri Etik Unsurları

COSO'ya göre iç kontrol sisteminin ana unsurlardan biri, Kontrol Çevresi başlığı altında yer alan etik ilkeler uyarınca veri madenciliği kullanılarak satış stratejisi belirlenmesi sürecinde ve satışları arttırmada etik ilkelere ne ölçüde uygun davranıldığı

konusu oldukça önem taşımaktadır. Bu kapsamda aşağıda sıralanan maddeler işletmeye yol gösterme açısından ışık tutacaktır. (Karakaş, 2008, s.23-29).

- Tüketicilere yönelik fiyatlandırmadan ziyade firma amaçlarına ulaşmayı hedefleyen fiyat tespiti.
- Fiyatı değiştirmeksizin ürünün kalite ve/veya miktarını değiştirme.
- Gerekl olmayan talebin uyarılması ve bazı pazarlamacıların zararlı ürünlere karşı kişi ya da işletmeleri tüketici yapma girişimleridir. (Çocuklar ve gençlere sigara, alkol, uyuşturucu satılması gibi).
- Müşteri veri tabanından bilgi satma.
- Rakipleri kötüleme.
- İndirim yapılan ürünleri yeterince bulundurmama.
- Kasırga, deprem gibi doğal afetler sonrası kıt ürünlerin fiyatını arttırma.
- En iyileri olmasa bile, satış elemanlarına en yüksek karlı ürünleri satma noktasında baskı yapma.
- Kişisel özel bilgileri amaç dışı kullanma.
- Yasal olmayan eylemler (sahtekarlık, bilgisayar programlarını ve bilgisayarları bozmak).
- Büyük çaplı şirketlerin pazarlama yöntemleri aracılığıyla az gelişmiş ülkeleri sistematik olarak sömürmesi.
- Yerli endüstrileri uluslararası pazarlara sunulan ürünlerden korumak için serbest ticaret tarifeleri ve kotalar arasındaki değişime odaklanmış diplomatik çatışmalar.
- Kalitesiz ve güvenli olmayan ürünlerin az gelişmiş ya da düzensiz pazarlara satışı.

Satış etiği, satış personeline müşterileri ile yaptıkları günlük kontratlarında yol göstermek amacıyla ahlaki bir çerçeve sağlamaktadır. Etik ikilemler satışta yaygındır

çünkü satış personeli sıklıkla müşterilerin talepleri ve rekabetçi önerilerini cevapladıkları bir alanda karar almak zorunluluğuyla mücadeleye çalışmaktadır. Yönetime karar alma noktasında ikilemin tanımlanması, unsurların ortaya konması, bu unsurların yasallık, haklılık ve fayda açısından değerlendirilmesi ve karar verilmesi gibi dinamikler yardımcı olacaktır. (Dalrymple, Cron, 1982, s. 278).

5.4. BİLANÇO KAVRAMI

Bilanço; belirli bir tarihte bir firmanın varlıklarını (aktif değerlerini) borçlarını ve özkaynaklarını ayrıntılı bir biçimde ortaya koyan şekil şartları taşıyan mali tablodur. (Güçlü, 2011, s.49.). Bilanço denkliği Tablo 4’de verilmiştir.

Tablo 4

Bilanço Denkliği Tablosu

BİLANÇO DENKLİĞİ		
Aktif		Pasif
Döner Varlıklar		Kısa Vadeli Borçlar
		Uzun Vadeli Borçlar
Sabit Varlıklar		Özsermaye
-----		-----
Toplam Aktifler	=	Toplam Pasifler

Kaynak: Müge İşeri, Sorularla İşletme Finansmanı, s.4.

Varlıklar, dönen ve duran olarak adlandırılan iki grupta toplanmaktadır.

Dönen varlıklar grubunda bir yıl veya işletmenin normal faaliyet dönemi içinde paraya çevrileceği veya kullanılacağı tahmin edilen varlıklar yer almaktadır.

Duran varlıklarda ise normal şartlar altında bir yıl içinde elden çıkarılması düşünülmeyen veya yararları bir yılda tükenmeyecek olan varlıklar yer almaktadır.

Kaynaklar kısa ve uzun vadeli yabancı kaynaklarla öz kaynaklardan meydana gelmektedir..

Kısa vadeli yabancı kaynaklar bir yıl içinde ödenmesi gereken; uzun vadeli yabancı kaynaklar ise bir yıldan sonra ödenecek olan borçlardır. Öz kaynaklar ise işletme sahip ve ortakları tarafından işletme kişiliğine tahsis edilmiş kaynaklardır. (Argun, 2003, s.12.).

Bilançoyu bir diğer ifadeyle; bir firmanın sahip olduğu varlıkları ve bu varlıkların üzerindeki hakları gösteren temel tablo olarak da tanımlamamız mümkündür. (Güçlü, 2011, s.49.). İşletme ilk kurulduğu zaman kurucuların işletmeye varlık olarak tahsis ettikleri kaynak sermayedir; dolayısıyla sermaye işletme sahiplerinin varlıklar üzerindeki haklarıdır. Kar; bu haklarda ortaya çıkan artıştır.

İşletme varlıklarını arttırmak için; ya kar alınmayıp işletmeye bırakılır, yada üçüncü kişilerden borç alınır. Borçlar; üçüncü kişilerin varlıklar üzerindeki haklarıdır. İşletme varlıkları için ya sermaye yada borçlanmak suretiyle kaynak sağlar. Dolayısıyla her varlığın bir kaynağı olması kavramı bilanço eşitliği olarak karşımıza çıkmaktadır. (Çanakçıoğlu, 2008, s.51.).

5.5. UYGULAMA ÖRNEĞİ SEGMENTASYON ANALİZİ SONUÇLARININ BİLANÇO ÜZERİNE ETKİLERİ

Yapılan veri madenciliği çalışmasında işletmenin son üç yıllık devamlı müşterileri ve müşteri karakteristikleri belirlenmiştir. İşletmenin bu doğrultuda yeni bir satış stratejisi uygulaması sonucunda satışlarını arttıracacağı varsayımı altında peşin satışlarda likidite ve cari oran, vadeli satışlarda ise ticari alacaklar hesabında bir artış meydana gelecektir. Likidite, anında ve değerinde nakde dönüşebilme yeteneği demektir. Cari Oran (CO), dönen varlıkların firmanın kısa süreli yükümlülüklerini karşılamakta kullanabileceği iktisadi değerlerin kısa süreli borçlara bölünmesi ile hesaplanmaktadır. CO'nın hesaplanmasında amaç, firmanın kısa vadeli borçlarını

ödeme gücünü ölçmek ve net işletme (çalışma) sermayesinin yeterli olup olmadığını saptayabilmektir. Net işletme sermayesi, kısa süreli borçların dönen varlıklardan çıkartılması ile bulunmaktadır. CO, işletmeye kredi verenler tarafından da önem taşır. Ticari alacakların artması aynı zamanda alacak tahsilatı riskini de doğuracaktır. Bu noktada işletmenin nasıl bir tahsilat politikası uyguladığı önem kazanmaktadır.

XYZ firması kurumsal olan ve sürekli çalıştıkları kobilere vadeli satış, düşük tutarlı satış yaptığı ve yeni çalıştıkları firmalara peşin satış politikası uygulamaktadır. İşletme yönetiminin tahsilat risklerinin önüne geçmek için peşin satışlarda uyguladığı kontrol aktivitesinin yerinde olduğu görülmüştür. Kendisine satış yapılan müşteriye borcunu XYZ işletmesinin banka hesabına havale ettikten sonra kalibrasyon sertifikası gönderilmektedir. Bilançoda en likit varlık kasa ve banka hesapları olduğundan müşteri doğru bir tahsilat politikası uygulamaktadır. Vade uygulanan müşterilerde ise alacak tutarı ile ilgili sözleşme imzalanarak sertifikaları gönderilmektedir. Ticari alacaklar için vade farkı uygulaması yapılmamaktadır. Güvenilir olmayan ve 1000TL'nin altında çek tahsilatı kabul edilmemektedir. Firma bu şekilde finansal risklerini dengelemeye çalışmaktadır.

İşletmeler bazen artan satışlar neticesinde düşük oranda bir tahsilat yaparken kendi müşterilerine daha uzun vadeli kredi tanıma yoluna gidebilmektedirler. Diğer bir ifadeyle satışlarını kredili yapmayı tercih etmektedirler. Ticari alacaklardaki bu durumun böyle devam etmesi, işletmenin nakit darboğazı içersine girmesine yol açabilmektedir. Geçmiş veriler ışığında müşterilere ödeme potansiyelleri ile ilgili olarak risk segmentasyonu yapılarak kredilerin buna uygun olarak çıkartılması sonucu ticari alacakların zamanında tahsil edilmesi işletmenin bilançosunda aktifini güçlendirecek ve likiditeyi arttıracaktır. Şüpheli ticari alacak hesap kalemi üzerinde kısmen denetim sağlamış olacaktır. İşletme tahsilatlarını güçlendirmek yoluyla dönen varlıklarını arttırarak kısa vadeli yabancı kaynaklarını rahatlıkla ödeyebilir hale gelecek ve günlük faaliyetlerini karşılayabilmek için yeterli oranda net işletme sermayesine sahip olacaktır.

8. SONUÇ

Teknolojik yoğunlaşmayla beraber 1990'lı yılların sonu ve 2000'lerde hız kazanan veri madenciliği uygulamaları daha gelişmiş algoritmalarla beslenen paket programların işletmelere sunulmasını da beraberinde getirmiştir. Veri madenciliği uygulamaları yapmak için bilgisayar programları gerekmektedir. Veri madenciliği yazılımları ticari ve açık kaynak kodlu olmak üzere iki gruba ayrılmaktadır. Ticari yazılımlara SPSS Clementine, Excel, SPSS, SAS (Enterprise Miner), Angoss, KXEN, MS SQL Server, MATLAB, Oracle'ın geliştirdiği modeller örnektir. Açık kaynak kodlu yazılımlara ücretsiz olarak erişilebilen Keel, Knime, Orange, R, Rapid Miner(Yale) ve WEKA yazılımlar örnektir.

Çalışmanın amacı, müşterinin elindeki dağınık halde bulunan veriyi, veri madenciliği süreci işletilerek segmente etmek, segmentasyon analizi sonuçlarının iç kontrol sistemi açısından etkilerini inceleyerek satış stratejilerini yeniden düzenlemesine yardımcı olarak satışlarını arttırması halinde bilanço hesapları üzerinde gerçekleştirilecek etkileri araştırmaktır.

Verilerin dönüştürülmesi sürecinde GNU genel kamu lisansı altında bildirilmiş açık kaynak kodlu WEKA yazılımı kullanılmıştır. Bu yazılımın tercih edilmesinin sebebi; içerdiği algoritma açısından geniş kapsamlı olması, makine öğrenmesi paketleri açısından en güçlü yazılım olması, istatistiksel kütüphanesinde uygulama örneği için seçilen kümeleme modelini içeriyor olması, kullanım ve öğrenim kolaylığı açısından kullanıcı dostu olması, görselleştirme yönünden güçlü olmasıdır. Segmentasyon analizi çalışması yapmak için kümeleme modeli (clustering) seçilmiş, algoritma olarak sırasıyla EM(Expectation Maximization), FarthestFirst, Cobweb, k-Means, Makedensitybase algoritmaları çalıştırılmış, en anlamlı sonuç üreten rapor K-means algoritmasından sağlandığı için bu algoritma çalışmaya dayanak teşkil etmiştir. Kullanılan algoritma sayesinde verilerden bilgi çıkarımı yapılabilmektedir.

Yapılan çalışmada XYZ Kalibrasyon firmasının son üç yıllık (2013, 2014 ve 2015) satışlarına ilişkin müşteri esasında kaydedilen verileri Excel ortamında alınmış, segmentasyon analizi yapmak için öncelikle devamlı müşterileri belirlenmiştir. 251 adet devamlı müşteriye ait karakteristiklerin belirlenmesi amacıyla öncelikle Excel

ortamında düzenleme yapılmış, kategorik değerler nümerik değerlere dönüştürülmüş, hazırlanan dosya ARFF formatında WEKA programına yüklenmiştir. Kümeleme (clustering) yöntemi ve basit K-means algoritması çalıştırılarak müşteriler iki kümeye ayrılmıştır.

%62 ve %38 oranında ayrılan her iki küme de de müşterilerin Kobi nitelikli ve eski müşterilerin bulunduğu ve Anadolu Yakası'nda faaliyet gösterdikleri tespit edilmiştir. Bu durum iç kontrol açısından yoğunlaşma ve iş risklerini de beraberinde getirecektir. Riski dağıtmak için Avrupa Yakası ve İstanbul dışı firmalara satış desteği verilmelidir. Her iki kümede de ağırlıklı müşterilerin Kobi işletmeleri olduğu göz önüne alınırsa büyük firmalar için dahili kalibrasyon kullanımı ile dışardan hizmet alma konusunda maliyet en iyileme (optimizasyonu) gibi analizlerle işbirliği yapma yoluna gidilerek büyük firmalar da müşteri portföyüne kazandırılabilir. %62'lik kümede üretim sektörlerinin yoğunlaştığı, yılda 5 kez kalibrasyon yaptırdıkları, satış tutarının 1000TL'nin altında olduğu görülmüştür. %38'lik ikinci kümede boya ve kimya sektör firmalarının yoğunlaştığı, yılda 2 kez ölçüm yaptırdıkları ve satış tutarlarının 1000-5000TL aralığında olduğu ortaya çıkmıştır. Boya ve kimya sektöründeki firmalar daha kazançlı olduklarından, bu sektörden daha fazla pay alma yoluna gidilmesi kazanç sağlayacaktır. %62'lik kümenin üretim işletmelerinde yoğunlaşması bu sektörlerde daha büyük bir potansiyelin var olduğunu göstermektedir. XYZ firması, üretim sektöründe faaliyet gösteren firmalara ayrı bir satış stratejisi uygulamalıdır.

XYZ işletmesi 2015 yılı sonunda 597 müşteri sayısına sahipken son üç yılda hizmet verdiği devamlı müşteri sayısının 251 olması müşteri portföyünü ve rekabet riskini daha iyi yönetmesi gerektiğini ortaya koymaktadır. Devamlı olmayan müşteriler anket, vb. yöntemlerle devamlı hale getirilmeye çalışılmalıdır. Ayrıca her iki kümede de eski müşterilerin yoğunlaşması, yeni müşterilere yönelik satış stratejileri geliştirilmesi gerekliliğini ortaya koymaktadır. XYZ işletmesinin satış stratejisini belirleyip satışlarını arttırırken ne ölçüde etik davranacağı da iç kontrol açısından önem taşımaktadır.

XYZ firmasının müşteri segmentasyonu sonuçları ışığında doğru bir satış stratejisi uygulaması ve satışlarını arttırması halinde bilanço hesapları da etkilenecektir. Peşin satışlarda likidite ve cari oran, vadeli satışlarda ticari alacaklar hesabı

etkilenecektir. İ kontrol aısından firmanın rekabet ve tahsilat risklerini aynı anda ynetebilmek iin belirli bir vadeli satıř politikası uygulaması gerekmektedir. Yapılan incelemede firmanın mřterilerini “kurumsallık” ve “devamlı mřteri olma” kriterlerine gre sınıflandırdığı ve bu dođrultuda vadeli satıř politikası uyguladığı grlmřtr. İlk defa alıřılan mřterilere peřin satıř ynteminin uygulanması bilanoda řpheli alacak riskini azaltıcı bir unsurdur. Firmanın mřterilerini yakından tanınması, alacaklarını zamanında tahsil edebilmesi, dnen varlıklarını arttırarak kısa vadeli ykmllklerini rahatlıkla deyebilmesi ve net iřletme sermayesini glendirmesi anlamını tařımaktadır.

Gerekleřtirilen uygulamadan elde edilen sonuların tm iřletmelere ve bu iřletmelerin satıř pazarlama, muhasebe, iř geliřtirme ve i denetim blmlerine yararlı olacağı dřnlmektedir.

Firmalara ait oklu verilere veri madenciliđi teknikleri uygulanması nceden bilinmeyen ok sayıda desen ve rntnn ortaya ıkartılması ve karar alıcılara tahminleme konusunda farklı bakıř aıları sađlayacaktır.

KAYNAKÇA

- Aksoy, T. (2006). Tüm Yönleriyle Denetim. İkinci Baskı. Ankara: Yetkin Yayınları.
- Alagöz, A., S. Öge ve M. Ortakarpuz (2014). Bir Kurumsal Zeka Teknolojisi Olarak Veri Madenciliği ile Muhasebe Bilgi Sistemi İlişkisi. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 2014. Vol.1. No.21, s.16.
- Arens, A.A ve J.K. Loebbecke. (1988). Auditing An Integrated Approach. Prentice-Hall International, Inc.
- Arens, A.A., R.J.Elder ve M.S. Beasley. (2003). Auditing and Assurance Services An Integrated Approach. Pearson Education, Inc.
- Argun, D. ve C. İbiş. Mali Tablolar Analizi Uygulamaları. İstanbul: Mart Matbaacılık Sanatları A.Ş.
- Atmaca, M.(2012). Muhasebe Skandallarının Önlenmesinde İç Kontrol Sisteminin Etkinleştirilmesi. Afyon Kocatepe Üniversitesi İİBF Dergisi. (C.XIV, S I,2012)
- Bakkal, H. ve A. Kasımoğlu. (2012). İç Kontrol Sistemine Karşılaştırmalı Bir Bakış COSO ve COCO Modeli. Mevzuat Dergisi. Cilt.15. Sayı.178., s.11.
- Baskıcı, Ç.(2012). “İç Kontrol Sisteminin Kurumsal Yönetim Anlayışındaki Yeri: İMKB Şirketlerinde Bir Uygulama”, *Yüksek Lisans Tezi*.Ankara: Ankara Üniversitesi SBE
- Carmichael, D.R. ve J.J. Willingham.(1989). Auditing Concepts and Methods. McGraw-Hill, Inc.
- Çanakçıoğlu,M.(2008). Finansal Muhasebe. 1. Basım, İstanbul: ISMMMO Yayınları-107, s.51.
- Çınar, Ayşe ve Gökhan Silahtaroğlu. “Veri Madenciliği Teknikleri ile Müşteri Memnuniyetine Etki Eden Gizli Nedenleri Keşfi”,Marmara Üniversitesi İİBF Dergisi, Vol.33 No.309-333, 2012, s.311
- Çokluk, Ö.(2010). Lojistik Regresyon Analizi: Kavram ve Uygulama. Ankara Üniversitesi Kuram ve Uygulamada Eğitim Bilimleri / Educational Sciences: Theory & Practice. 10 (3) Yaz / Summer 2010 1357-1407
- Durmaz, C. ve M. Kocamış (2008). Oracle Data Miner ile Öğrenci Kayıtları Üzerine Bir Veri Madenciliği Uygulaması (Bitirme Ödevi. İstanbul Teknik Üniversitesi. Fen Edebiyat Fakültesi. Mühendislik Bilimleri.), s.4.

- Dursun, K.(2000). Yapay Sinir Ağları ve Meteorolojiye Uygulanması. İTÜ Uçak ve Uzay Bilimleri Fakültesi. Bitirme Çalışması, s.14.
- Dalrymple, D.J. ve W. L. Cron.(1982). Sales Management. USA: John Wiley & Sons, Inc.
- Daştan, A. (2008). Bilgi ve Eğitim Teknolojilerinde Yaşanan Gelişmelerin Muhasebe Eğitimine Etkisi. Ankara: Sıramat Ltd. Şti.
- Davenport, T.(2014).Big Data@Work. İstanbul: bzd Yayın ve İletişim Hizmetleri, s.129 (Türk Hava Yolları Yayınları)
- Erdoğan, M. (2006). Denetim Kavramsal ve Teknolojik Yapı. 3. Baskı. Ankara: Maliye ve Hukuk Yayınları
- Ertaş, F.C.(2011). Muhasebe Organizasyonu. 3. Baskı, Ankara: Seçkin Yayıncılık San. Ve Tic. A.Ş.
- French, S.(1986). Decision Theory: An Introduction to the Mathematics of Rationality.England: Ellis Horwood Limited
- Genel, H.(2004). Genetik Algoritmalarla Portföy Optimizasyonu.Yüksek Lisans Tezi. Ankara Üniversitesi. Sosyal Bilimler Enstitüsü.
- Giudici, P.(2003). Applied Data Mining Statistical Methods for Business and Industry. England: John Wiley & Sons Ltd.
- Güçlü, F.(2011). Muhasebe Denetimi İlkeler ve Teknikler. 4.Basım. Ankara: Detay Yayınları
- Güredin, E. (1998). Denetim. Sekizinci Baskı. İstanbul: Beta
- Haftacı,V. (2007). Muhasebe Dentimi. Birinci Baskı. İstanbul: Avcı Ofset.
- Han, J. ve M. Kamber.(2001). Data Mining. San Francisco:Morgan Kaufmann Publishers.
- <http://www.kpmgvergi.com/PratikBilgiler/Pages/Ufe-Tufe-Oranlari.aspx>
- İşeri, M.(2005). Sorularla İşletme Finansmanı. İstanbul Kültür Üniversitesi. İİBF Kbi Danışma Birimi. İstanbul, s.4-5
- Kantardzic, M.(2003). Data Mining Concepts, Models, Methods, and Algorithms. IEEE Press, Inc.
- Karakaş, G.H.(2008). Satış Yönetiminde Etik;Satışçıların Etik Algılamalarının Demografik Faktörleri ile İlişisini Ölçmeye Yönelik İlaç Sektöründe Bir

Uygulama. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı. Yüksek Lisans Tezi.

- Karğın, M.(2010). Muhasebe Verileri Olarak Finansal Oranlar ve İstatistiksel Özellikler. Ankara:Özbaran Ofset
- Kaya, M. ve S.A.Özel. Açık Kaynak Kodlu Veri Madenciliği Yazılımlarının Karşılaştırılması. Adana Bilim ve Teknoloji Üniversitesi Bilgisayar Müh. Bölümü. Çukurova Üniversitesi Bilgisayar Mühendisliği Bölümü <http://ab.org.tr/ab14/bildiri/79.pdf>
- Keresteci, Y.S.(2008). Kredi Kartı Kullanımında Sahtecilik Tespit Sistemleri. Gebze Yüksek Teknoloji Enstitüsü Sosyal Bilimler Enstitüsü. Strateji Bilimi Anabilim Dalı. Gebze 2008
- Kıral, H. (Ed.). (2014). İç Denetim. Yönetime Değer Katmak. Ankara: İç Denetim Koordinasyon Kurulu
- Kurnaz, N. ve T. Çetinoğlu. (2010). İç Denetim Güncel Yaklaşımlar. Birinci Baskı. Kocaeli: Umuttepe Yayınları.
- Kurnaz, S.(2004). Veri Yapıları ve Algoritma Temelleri. 1. Basım. İstanbul:Papatya Yayıncılık
- Kurumsal Çözüm Önerileri.(Mayıs 2014). ERP Komitesi Kurumsal Dönüşüm Platformu, s.56.
- Lennick, D. ve F. Kiel.(2005). Etik Zeka İş Performansının Arttırılması ve Liderlik Başarısı. Birinci Basım, İstanbul: CSA Global Publishing, s.23-29
- Meigs, R.F., M. A. Meigs, M. Bettner ve R. Whittington.(1998). Financial Accounting. USA: The McGraw-Hill Companies, Inc.
- Molina, H.G., J. D. Ullman ve J. Wildom.(2009). Database Systems The Complete Book. Pearson Education Inc.
- Ölmezoğulları, E. ve İ. Arı. Büyük Veri Problemlerine Çözüm olarak Veri Akış Madenciliği. Özyeğin Üniversitesi Bilgisayar Mühendisliği
- Önder, F. (2008). Türk Hukukunda İç Denetim ve Uluslararası Standartlara Uyumu. Birinci Baskı. Ankara: Asil Yayın Dağıtım Ltd. Şti.
- Özer, M. (2010). Kuruluşlarda Süreç, Performans ve Risk Analizi/Yönetimi. Birinci Baskı. Ankara: Adalet Yayınevi

- Özkan, M.(2014). Veri Madenciliğinin Finansal Kararlarda Kullanımı. Çankırı Karatekin Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. Vol.4. No.1, s.17.
- Özkul,U. ve F.Z. Özdemir.(2011). İşletmelerde Hile Riski Yönetimi. 1. Baskı. İstanbul:Beta
- Öztemel, E.(2003). Yapay Sinir Ağları. 1. Basım. İstanbul:Papatya Yayıncılık
- Pehlivanlı, D. (2010). Modern İç Denetim. Birinci Baskı. İstanbul: Beta
- Pehlivanlı, D.(2011). Hile Denetimi Metodoloji ve Raporlama. 1. Baskı. İstanbul: Beta Basım A.Ş.
- Pehlivan, P. (Yaz - Sonbahar 2013). Bilgi Teknolojileri Yönetişimi. İç Denetim. No.36, s.61
- Pektaş, A.O. (2013). SPSS ile Veri Madenciliği. 1. Baskı. İstanbul: Dikeyksen Yayın Dağıtım, Yazılım ve Eğitim Hizmetleri San. Ve Tic. Ltd. Şti.
- Ramakrishnan, R. Ve J. Gehrke. (2003). Database Management Systems. McGraw-Hill Companies, Inc., 1221 Avenue of the Americas. Newyork, NY 10020
- Sarıman, G.(2011). Veri Madenciliğinde Kümeleme Teknikleri Üzerine Bir Çalışma K-Means ve K-Medoids Kümeleme Algoritmalarının Karşılaştırılması. Süleyman Demirel Üniveristesesi Fen Bilimleri Enstitüsü Dergisi 15-3(2011),192-202, s.193
- Savaş, S. N. Topaloğlu ve M. Yılmaz. (Bahar 2012). Veri Madenciliği ve Türkiyedeki Uygulama Örnekleri. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi. Sayı.21, s.6.
- Silahtaroglu, G. (2013).Veri Madenciliği Kavram ve Algoritmaları. 2.Baskı. İstanbul: Papatya Yayıncılık Eğitim A.Ş.
- Şeker, S.E.(2015). Veri Ambarı [Electronic Version]. YBS Ansiklopedi. Vol.2, No.4, (03.12.2015)
- Şen, Z.(2001). Bulanık Mantık ve Modelleme İlkeleri. İstanbul: Bilge Sanat Yapım Yay. Tant. Kağ. Turz. San. Tic. Ltd. Şti.
- Tan, P.N., M. Steinbach ve V. Kumar(2006). Introduction to Data Mining. Boston: Pearson Education, Inc., 2006
- Taylor, D.H.ve G.W. Glezen. (1994). Auditing: Integrated Concepts and Procedures. John Wiley&Sons, Inc.

- Terzi,Ö. E. U. Küçüksille, G. Ergin, A. İlker. (Şubat 2011). Veri Madenciliği Süreci Kullanılarak Güneş Işınımı Tahmini. Vol.3. No.2, s.3.
- Thuraisingham, B.M.(2003). Web Data Mining and Applications in Bussiness Intelligence and Counter-Terrorism. CRC Press LLC. Boca Raton. FL, USA, s.XV.
- Tüzüntürk, S. (2010).Veri Madenciliği ve İstatistik. Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. Vol.29. No.1, s.5.
- Watson, R.T. (2002)Data Management Databases and Organizations. NewYork:John Wiley&Sons, Inc.
- WEKA programı <http://www.cs.waikato.ac.nz/ml/weka/>
- Yünel, Y.(Mayıs 2010). k-Means Kümeleme Algoritmasının Genetik Algoritma Kullanılarak Geliştirilmesi. İTÜ Fen Edebiyat Fakültesi Matematik Mühendisliği Bölümü.(Bitirme ödevi), s.2.

EKLER

Ek-1: Segmentasyon Analizi İçin Dönüştürme Exceli

<i>Adres</i>	<i>Eski / Yeni</i>	<i>Sektör Sınıf</i>	<i>Kobi Büyük</i>	<i>Kalibrasyon Periyod</i>	<i>Tutar Aralığı 15</i>	<i>Tutar Aralığı 14</i>	<i>Tutar Aralığı 13</i>
1	E	2	2	5	0-1000	0-1000	1000-5000
3	E	7	2	1	0-1000	0-1000	0-1000
2	E	2	2	5	0-1000	1000-5000	0-1000
2	E	2	1	5	0-1000	0-1000	0-1000
1	E	3	1	5	0-1000	0-1000	10000+
1	E	8	1	2	1000-5000	1000-5000	1000-5000
3	E	9	2	5	0-1000	0-1000	0-1000
3	E	7	2	1	0-1000	0-1000	0-1000
2	E	10	2	5	0-1000	0-1000	0-1000
1	E	4	1	2	5000-10000	5000-10000	5000-10000
2	E	4	2	2	0-1000	0-1000	0-1000
1	Y	5	2	1	1000-5000	1000-5000	1000-5000
2	Y	7	2	1	1000-5000	0-1000	1000-5000
2	Y	10	1	5	0-1000	0-1000	0-1000
1	E	6	2	2	0-1000	0-1000	0-1000
2	E	5	2	1	0-1000	5000-10000	5000-10000
3	E	3	2	5	0-1000	1000-5000	0-1000
1	E	4	2	2	1000-5000	1000-5000	0-1000
1	E	2	2	5	0-1000	0-1000	0-1000
1	E	2	2	5	1000-5000	1000-5000	1000-5000
3	E	7	2	1	0-1000	1000-5000	1000-5000
2	Y	9	2	5	1000-5000	1000-5000	1000-5000
3	E	7	2	1	0-1000	0-1000	0-1000
2	E	4	2	2	0-1000	5000-10000	1000-5000
3	E	4	1	2	1000-5000	0-1000	1000-5000
2	E	5	1	1	1000-5000	1000-5000	0-1000
1	E	7	2	1	1000-5000	1000-5000	0-1000
1	E	6	2	2	1000-5000	0-1000	0-1000
3	Y	9	2	5	1000-5000	0-1000	1000-5000
1	Y	4	2	2	0-1000	1000-5000	0-1000
3	Y	6	2	2	1000-5000	1000-5000	1000-5000
3	Y	9	1	5	0-1000	0-1000	0-1000
2	E	3	2	5	0-1000	0-1000	0-1000
1	E	3	1	5	1000-5000	5000-10000	5000-10000
1	E	6	2	2	0-1000	0-1000	1000-5000
1	E	1	2	2	1000-5000	1000-5000	1000-5000
1	Y	8	2	2	1000-5000	0-1000	0-1000
2	Y	9	1	5	1000-5000	5000-10000	1000-5000
1	E	9	2	5	0-1000	0-1000	0-1000
3	E	3	1	5	1000-5000	1000-5000	1000-5000

3	E	7	1	1	0-1000	1000-5000	1000-5000
2	E	7	1	1	10000+	10000+	1000-5000
2	E	2	1	5	1000-5000	1000-5000	1000-5000
3	E	10	2	5	0-1000	0-1000	0-1000
1	E	4	2	2	0-1000	0-1000	0-1000
3	E	6	2	2	1000-5000	1000-5000	1000-5000
3	E	10	2	5	0-1000	0-1000	0-1000
3	E	8	2	2	0-1000	0-1000	0-1000
1	Y	4	2	2	10000+	10000+	5000-10000
1	E	6	2	2	0-1000	0-1000	0-1000
3	E	10	1	5	1000-5000	1000-5000	1000-5000
3	E	9	1	5	10000+	5000-10000	5000-10000
3	E	4	2	2	0-1000	0-1000	0-1000
3	E	4	2	2	0-1000	0-1000	0-1000
1	Y	10	2	5	1000-5000	1000-5000	1000-5000
2	E	7	2	1	0-1000	0-1000	0-1000
3	E	4	1	2	1000-5000	1000-5000	1000-5000
3	E	8	1	2	1000-5000	0-1000	5000-10000
3	E	2	2	5	1000-5000	0-1000	0-1000
3	E	10	1	5	0-1000	0-1000	0-1000
2	E	7	2	1	0-1000	0-1000	0-1000
2	E	7	2	1	1000-5000	1000-5000	0-1000
1	E	9	2	5	0-1000	0-1000	0-1000
1	E	7	2	1	0-1000	0-1000	0-1000
2	E	10	2	5	0-1000	0-1000	0-1000
1	E	10	2	5	0-1000	0-1000	0-1000
1	E	7	2	1	1000-5000	0-1000	1000-5000
1	Y	3	2	5	10000+	10000+	0-1000
1	E	9	1	5	10000+	10000+	5000-10000
3	E	2	2	5	1000-5000	1000-5000	1000-5000
2	E	9	2	5	0-1000	0-1000	0-1000
1	E	7	2	1	0-1000	0-1000	0-1000
2	E	4	2	2	0-1000	0-1000	0-1000
3	E	8	1	2	0-1000	0-1000	0-1000
3	E	9	2	5	0-1000	0-1000	0-1000
1	E	5	2	1	0-1000	0-1000	0-1000
3	Y	7	2	1	0-1000	0-1000	0-1000
2	E	2	2	5	0-1000	1000-5000	0-1000
1	Y	10	2	5	1000-5000	1000-5000	1000-5000
3	E	2	2	5	0-1000	0-1000	0-1000
1	Y	5	2	1	5000-10000	5000-10000	1000-5000
3	E	9	2	5	0-1000	0-1000	0-1000
2	E	2	2	5	1000-5000	0-1000	0-1000
1	E	9	2	5	1000-5000	1000-5000	0-1000
1	E	9	2	5	0-1000	0-1000	0-1000
3	E	9	2	5	0-1000	0-1000	0-1000
2	E	4	2	2	1000-5000	0-1000	1000-5000
2	E	1	2	2	1000-5000	0-1000	1000-5000
1	E	4	1	2	1000-5000	1000-5000	1000-5000

1	E	1	2	2	0-1000	1000-5000	1000-5000
3	E	1	1	2	1000-5000	0-1000	0-1000
1	E	9	1	5	0-1000	0-1000	1000-5000
1	E	9	1	5	1000-5000	0-1000	1000-5000
1	E	9	2	5	0-1000	0-1000	0-1000
3	E	8	2	2	1000-5000	1000-5000	1000-5000
3	E	9	2	5	0-1000	0-1000	0-1000
1	E	6	2	2	0-1000	0-1000	0-1000
3	Y	3	2	5	0-1000	0-1000	0-1000
1	Y	7	2	1	1000-5000	0-1000	1000-5000
1	E	6	2	2	0-1000	0-1000	0-1000
2	Y	8	2	2	1000-5000	1000-5000	1000-5000
1	E	9	2	5	1000-5000	1000-5000	1000-5000
3	E	2	1	5	0-1000	1000-5000	0-1000
3	Y	9	2	5	1000-5000	1000-5000	1000-5000
1	E	9	2	5	0-1000	0-1000	0-1000
2	E	5	1	1	0-1000	0-1000	1000-5000
2	E	6	2	2	0-1000	0-1000	0-1000
1	Y	3	2	5	0-1000	0-1000	0-1000
3	Y	9	2	5	0-1000	0-1000	0-1000
3	E	2	2	5	1000-5000	1000-5000	1000-5000
1	E	7	2	1	0-1000	0-1000	0-1000
1	E	5	2	1	1000-5000	1000-5000	1000-5000
2	Y	5	1	1	1000-5000	1000-5000	1000-5000
2	Y	10	1	5	0-1000	0-1000	0-1000
1	E	4	2	2	0-1000	1000-5000	1000-5000
1	E	9	2	5	1000-5000	1000-5000	1000-5000
3	E	4	2	2	1000-5000	1000-5000	1000-5000
1	E	9	2	5	0-1000	1000-5000	1000-5000
2	Y	4	2	2	10000+	10000+	10000+
1	E	5	2	1	10000+	10000+	5000-10000
2	E	9	2	5	0-1000	0-1000	0-1000
1	E	4	2	2	0-1000	0-1000	0-1000
3	Y	9	1	5	10000+	10000+	10000+
1	Y	2	2	5	0-1000	0-1000	0-1000
1	Y	3	2	5	1000-5000	1000-5000	1000-5000
1	Y	4	1	2	1000-5000	1000-5000	1000-5000
1	Y	4	2	2	0-1000	1000-5000	1000-5000
3	E	1	2	2	0-1000	0-1000	0-1000
3	E	4	1	2	10000+	5000-10000	5000-10000
3	E	4	2	2	5000-10000	5000-10000	1000-5000
2	E	7	1	1	1000-5000	1000-5000	1000-5000
3	E	4	2	2	0-1000	0-1000	0-1000
1	E	4	2	2	1000-5000	1000-5000	1000-5000
1	Y	4	2	2	1000-5000	0-1000	0-1000
3	E	7	2	1	0-1000	0-1000	0-1000
3	E	3	2	5	10000+	10000+	5000-10000
3	E	9	1	5	1000-5000	1000-5000	1000-5000
3	E	4	1	2	10000+	10000+	10000+

1	E	7	2	1	0-1000	0-1000	0-1000
2	E	5	2	1	0-1000	0-1000	0-1000
1	E	5	2	1	10000+	5000-10000	5000-10000
3	E	4	2	2	1000-5000	0-1000	1000-5000
3	E	10	1	5	1000-5000	1000-5000	1000-5000
2	E	1	2	2	5000-10000	0-1000	1000-5000
2	E	5	2	1	1000-5000	1000-5000	1000-5000
1	E	9	2	5	0-1000	0-1000	0-1000
3	E	5	2	1	1000-5000	1000-5000	1000-5000
1	E	10	1	5	0-1000	0-1000	0-1000
2	E	3	1	5	10000+	1000-5000	1000-5000
3	E	4	1	2	1000-5000	0-1000	1000-5000
2	E	8	1	2	5000-10000	5000-10000	1000-5000
1	E	9	2	5	5000-10000	1000-5000	5000-10000
1	Y	6	2	2	0-1000	0-1000	0-1000
1	Y	9	2	5	0-1000	1000-5000	1000-5000
2	E	3	2	5	0-1000	0-1000	0-1000
3	E	7	2	1	1000-5000	1000-5000	1000-5000
1	E	3	2	5	0-1000	0-1000	0-1000
1	E	4	2	2	1000-5000	1000-5000	1000-5000
1	E	5	2	1	5000-10000	0-1000	1000-5000
1	Y	9	2	5	1000-5000	1000-5000	1000-5000
1	Y	10	2	5	5000-10000	5000-10000	5000-10000
1	E	4	2	2	0-1000	1000-5000	1000-5000
2	E	10	2	5	1000-5000	1000-5000	1000-5000
1	Y	5	2	1	0-1000	0-1000	0-1000
1	Y	4	2	2	1000-5000	0-1000	1000-5000
1	E	9	1	5	1000-5000	1000-5000	1000-5000
2	Y	8	2	2	0-1000	1000-5000	0-1000
2	Y	1	2	2	0-1000	0-1000	0-1000
1	E	4	2	2	1000-5000	1000-5000	1000-5000
3	E	4	2	2	1000-5000	0-1000	0-1000
1	E	8	1	2	0-1000	0-1000	0-1000
3	E	4	2	2	0-1000	0-1000	0-1000
1	E	7	2	1	1000-5000	0-1000	0-1000
1	E	10	2	5	0-1000	0-1000	0-1000
1	E	9	2	5	1000-5000	1000-5000	1000-5000
2	E	9	1	5	0-1000	0-1000	0-1000
3	E	6	2	2	0-1000	1000-5000	0-1000
3	Y	4	2	2	0-1000	0-1000	0-1000
2	E	6	1	2	1000-5000	1000-5000	0-1000
1	Y	9	2	5	0-1000	0-1000	0-1000
1	Y	3	2	5	0-1000	0-1000	0-1000
1	E	4	1	2	0-1000	0-1000	0-1000
1	E	10	2	5	0-1000	0-1000	0-1000
3	E	4	1	2	0-1000	1000-5000	0-1000
1	E	4	1	2	5000-10000	10000+	10000+
1	E	4	1	2	0-1000	0-1000	0-1000
1	E	4	1	2	1000-5000	5000-10000	5000-10000

1	E	7	2	1	0-1000	0-1000	0-1000
1	Y	9	2	5	1000-5000	1000-5000	1000-5000
1	E	2	1	5	5000-10000	5000-10000	0-1000
1	E	9	2	5	5000-10000	5000-10000	1000-5000
1	E	4	2	2	10000+	10000+	10000+
1	Y	2	2	5	0-1000	1000-5000	1000-5000
2	E	9	2	5	0-1000	0-1000	0-1000
1	E	10	1	5	1000-5000	1000-5000	1000-5000
1	E	8	2	2	0-1000	1000-5000	0-1000
3	E	1	2	2	1000-5000	1000-5000	1000-5000
1	Y	6	2	2	0-1000	0-1000	0-1000
1	E	10	2	5	0-1000	0-1000	0-1000
3	E	8	2	2	0-1000	0-1000	0-1000
1	E	7	2	1	10000+	10000+	5000-10000
1	E	9	2	5	0-1000	0-1000	0-1000
3	E	2	2	5	0-1000	0-1000	0-1000
1	Y	7	2	1	0-1000	0-1000	0-1000
2	E	9	2	5	1000-5000	1000-5000	1000-5000
1	Y	3	2	5	5000-10000	5000-10000	5000-10000
3	E	9	2	5	0-1000	0-1000	0-1000
1	E	8	2	2	0-1000	0-1000	0-1000
1	Y	10	2	5	0-1000	0-1000	0-1000
1	E	10	2	5	1000-5000	1000-5000	1000-5000
2	E	9	1	5	1000-5000	1000-5000	1000-5000
3	E	10	2	5	0-1000	0-1000	0-1000
3	E	4	2	2	1000-5000	1000-5000	1000-5000
1	E	7	2	1	0-1000	0-1000	0-1000
3	E	6	2	2	0-1000	0-1000	0-1000
3	E	8	2	2	5000-10000	1000-5000	1000-5000
3	E	9	1	5	0-1000	0-1000	0-1000
1	E	1	2	2	1000-5000	1000-5000	1000-5000
3	Y	7	2	1	1000-5000	1000-5000	1000-5000
2	E	10	2	5	0-1000	0-1000	0-1000
2	E	3	2	5	0-1000	1000-5000	0-1000
1	E	4	1	2	0-1000	1000-5000	1000-5000
1	E	7	2	1	0-1000	0-1000	0-1000
2	E	4	1	2	1000-5000	1000-5000	1000-5000
3	E	8	1	2	0-1000	0-1000	0-1000
3	E	2	2	5	1000-5000	1000-5000	1000-5000
3	Y	7	2	1	0-1000	0-1000	0-1000
3	E	7	2	1	0-1000	0-1000	0-1000
2	E	9	1	5	1000-5000	1000-5000	1000-5000
1	E	3	2	5	0-1000	0-1000	0-1000
3	E	9	2	5	0-1000	0-1000	0-1000
2	E	5	2	1	0-1000	0-1000	0-1000
3	E	2	1	5	5000-10000	5000-10000	1000-5000
2	E	8	2	2	0-1000	0-1000	0-1000
2	E	7	2	1	1000-5000	1000-5000	1000-5000
1	Y	3	2	5	1000-5000	1000-5000	0-1000

1	Y	7	2	1	0-1000	0-1000	0-1000
1	E	4	2	2	1000-5000	0-1000	0-1000
2	E	2	1	5	0-1000	0-1000	1000-5000
2	E	2	1	5	0-1000	0-1000	1000-5000
2	E	10	2	5	0-1000	0-1000	0-1000
3	E	9	2	5	0-1000	0-1000	0-1000
2	E	10	2	5	1000-5000	1000-5000	0-1000
3	E	3	1	5	0-1000	0-1000	0-1000
1	E	10	2	5	0-1000	0-1000	0-1000
3	E	9	1	5	5000-10000	10000+	1000-5000
1	Y	2	2	5	0-1000	0-1000	0-1000
1	E	4	2	2	0-1000	0-1000	0-1000
3	E	7	2	1	0-1000	0-1000	1000-5000
3	E	1	2	2	0-1000	0-1000	0-1000
3	E	7	2	1	1000-5000	1000-5000	1000-5000