T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KREDİ SKORLAMA

AYŞE NİLGÜN KARAKAYA

YÜKSEK LİSANS TEZİ

İSTATİSTİK ANA BİLİM DALI

KONYA, 2010

ÖZET

YÜKSEK LİSANS TEZİ

KREDİ SKORLAMA

Ayşe Nilgün KARAKAYA

Selçuk Üniversitesi Fen Bilimleri Enstitüsü

İstatistik Ana Bilim Dalı

Danışman: Doç.Dr.M.Fedai

Jüri: Doç.Dr.M.Fedai KAYA

Doç.Dr.Halil AYDOĞDU

Yrd.Doç.Dr.Hasan KÖSE

Gelişmekte olan bankacılık sektöründe kredi kartları gündelik hayatta önemli bir ürün olarak yerini almıştır. Müşteriye kredi kartı verilmesi-verilmemesi ise banka açısından

dikkatle incelenerek alınması gereken bir karardır ve kredi kartı talebi arttıkça başvuruların değerlendirilmesi daha da karmaşık bir hal almaktadır. Değerlendirmeyi yapan kişiler farklı kriterleri dikkate alabilecekleri için alınan kararlar sübjektif olabilir. Bu durumda, hem artan başvuru sayısına doğru zamanda cevap vermek, hem de sübjektif kriterlerden arınıp objektif kararlar alabilmek için çeşitli istatistiksel ve istatistiksel olmayan teknikler kullanılmaktadır. Bu çalışmada, bir bankanın kredi kartı müşterilerine ait on üç değişken ile iyi müşteri-kötü müşteri ayırımı yapılmaya çalışılmıştır ve uygulama sonuçları birbirleri ile kıyaslanmıştır.

Uygulamada kullanılan bazı istatistiksel teknikler; diskriminant analizi, lojistik regresyon, en yakın komşu yaklaşımı yapay sinir ağlarıdır.Uygulamalarına yer verilmeden teorik bilgi olarak sunulan teknikler ise doğrusal programlama ve tam sayılı programlama konularına değinilmiştir.

Modellerin uygulamasında çeşitli avantajlar-dezavantajlar mevcut olmasıyla birlikte kullanılan veri setine göre tahmin başarısı en yüksek olan modelin lojistik regresyon olduğu söylenebilir. İlgili veri setine göre tüm modellere giren AYLIK NET GELİR değişkeni ise bu çalışma için kredi skorlamada etkili değişkenler olarak gözlemlenmiştir.

Anahtar Kelimeler:Kredi skorlama, lojistik regresyon, diskriminant analizi, en yakın komşu yaklaşımı, yapay sinir ağları

ABSTRACT

Master Thesis

CREDIT SCORING

Ayşe Nilgün KARAKAYA

Selcuk University

Graduate School of Natural and Applied Sciences

Department os Statistics

Supervisor:Assist Doc.Dr.M.Fedai KAYA

Jury:Assist.Doc.Dr.M.Fedai KAYA

Assist.Doc.Dr.Nimet Pehlivan YAPICI

Assist.Doc.Dr.Halil AYDOĞDU

Nowadays,credit cards in daily life become the most important product in developing finance sector.On behalf of banks, to issue or not to issue credit cards to clients should be decided after carefully investigating their backgrounds. By increasing credit card demands, the evaluation of credit card applications become more complex system. Because credit cards specialist consider different criteria, the decisions can be subjective. So, the statistical and non-statistical methods are used to both respond increasing credit card applications at right time and make objective decisions by getting rid of subjectivity.

In this work, good-bad clients are tried to be separated by regarding thirteen variables belongs to credit card clients of a specific bank and the results of aplication are compared between each other.

The statistical methodsused to in this work are discriminant analysis,logistic regression,nearest neighbour approach and neural networks.Moreover, the subjects, lineer programming and integer programming, are theoretically mentioned in this study without giving their applications.

Wheares there are some advantages and disadvantages of implementing those methods, the estimation rate of logistic regression is the highest among them in accordance with used data sets. Because MONTHLY NET INCOME is common variables in all models regarding to related data set, those are observed as the most effective variables in the credit scoring in this work.

Keywords: Credit scoring, logistic regression, discriminant analysis, nearest neighbour, neural networks.

TEŞEKKÜR

Bu tez çalışmamda fikirleriyle yol gösteren bütün anlayışı ve özellikle toleransıyla cesaretimi kırmadan bana yardımcı olan sevgili hocam Doç.Dr.M.Fedai KAYA'ya

Hayatım boyunca bana verdikleri güven ve destekten dolayı sevgili babam Cemal PEKER'e, özellikle tez çalışmamı bilgisayara dökerken kızım Özüm'den sevgi ve şevkatlerini esirgemeden ilgi gösterdikleri annem Şükran PEKER, kardeşim Hidayet PEKER'e, tez çalışmamda kaynak bulmama yardımcı olan ve beni destekleyen eşim Murat Barış KARAKAYA'ya ve çalışmanın düzen ve tertibinde bana yardımcı olan kardeşim Bekir PEKER'e teşekkürü bir borç bilirim.

İÇİNDEKİLER

ÖZET III
ABSTRACT V
TEŞEKKÜR VII
İÇİNDEKİLER VIII
KISALTMA LİSTESİ X
KREDİ RİSKİNİN HESAPLANMASINDA SKORLAMA YAKLAŞIMI 1
1.GİRİŞ 1
2. KREDİ SKORLAMA'NIN TANIMI TARİHÇESİ VE OLUŞTURULMASI 2
2.1. Kredi Skorlama (KS) 2
2.2. Kredi Skorlamanın Tarihsel Gelişimi ve Finans Sektöründeki Rolü 3
2.3. Kredi Skorlama Yönteminin Diğer Yöntemlere Göre Üstünlükleri 4
2.4. Kredi Skorlama ve Veri Madenciliği 6
2.5. Örnekleme Seçimi 7
2.6. İyi Müşteri-Kötü Müşteri Tanımlaması 7
2.7. Veri Kaynakları 8
2.8. Alt Ana Kütle Belirlemek (Determining Subpopulation) 9
2.9. Karakteristiklerin Genel Olarak Sınıflandırılması 10
3. KREDİ SKORLAMADA KULLANILAN YÖNTEMLER 10
3.1. Dikstiriminant Analizi 10
3.2. İki Grup Karesel Diskriminant Analizi 11
3.2.1. Kovaryans Matrislerinin Eşitliğinin Sınanması 14
3.2.2. Yeni Gözlemlerin Sınıflandırılması 17
3.3. Lojistik Regresyon 18
3.3.1. Logit Modelinin Tahmin Edilmesi 21
3.4. En Yakın Komşu Yaklaşımı (Nearest Neighbour Approach) 22
3.5. Yapay Sinir Ağları (Neural Networks) (YSA) 23
3.5.1. YSA Avantajları 24
3.5.2. Tek Katmanlı Sinir Ağları (Single Layer Neural Networks) 26
3.5.3. Çok Katmanlı Sinir Ağları (Multilayer Perceptrons) 27
3.6. Doğrusal Programlama (Lineer Programming) 28
3.7. Tamsayılı Programlama (Integer Programming) 29
4. UYGULUMALAR 31
4.1. BETİMSEL İSTATİSTİKLER 33
4.2. Kredi Kartlarında En Yakın Komşu Tekniği ile Skorlama 37
4.2.1. SONUÇ 39

4.3. Kredi kartlarında Diskriminant Analizi ile Skorlama 40
4.3.1. SONUÇ 44
5. SONUÇLAR VE GENEL DEĞERLENDİRME 45
KAYNAKLAR 46
ÖZGEÇMİŞ 48

KISALTMA LİSTESİ

ADA	Adımsal diskriminant analizi
DS	Davranış skorlaması
KS	Kredi Skorlama
MMD	Maksimum sapmanın minimizasyonu
MSD	Sapmaların mutlak değerler toplamı
YSA	Yapay sinir ağları

KREDİ RİSKİNİN HESAPLANMASINDA SKORLAMA YAKLAŞIMI

1.GİRİŞ

Finansal kuruluşlar uzun yıllardan beri çeşitli nedenlerden dolayı çeşitli nedenlerden dolayı güçlüklerle karşılaşmaktadırlar. Finansal sistemlerin temel aktörleri olarak görülen bankalar, yapılan işin doğası gereği birçok riski karşılamak durumunda kalmışlardır. Ekonomide lokomotif görevini gören bu sektörde yaşanan sorunların belli başlı nedeni ise zayıf risk yönetimini ya da fon ihtiyacı olan kişilerin kredibilitesinde çeşitli nedenlerden dolayı yaşanan azalmanın önceden tahmin edilememesidir. Bu yüzden ekonomik döngüde tasarruf fazlası olanlar ile fon ihtiyacı olanları bir araya getiren bankaların uygun parametreler yardımı ile maruz kalınabilecek riskleri belirlemeleri ve risk ayarlı getirilerini maksimize etmeleri gerekmektedir.

Bankalar açısından kredi riskinin temel kaynağı, kullandırdıkları krediler olsa da bankalar; faaliyetleri doğrultusunda oluşturdukları diğer enstrümanlarda da kredi riski ile karşı karşıya kalmaktadırlar. Geleneksel yaklaşımda bankalar ilk aşamada iyi kredi, "ileride ödenmeme olasılığı düşük olan kredi" verme üzerine yoğunlaşmışlardır. Bu aşamada bankalar bünyelerinde barındırdıkları istihbarat elemanlarının görüşleri doğrultusunda kredi talebinde bulunan firmaları ya da kişileri değerlendirmiştir. Piyasada iyi itibara sahip olanların kredileri onaylanır iken diğer başvurular için bilgi eksikliğinden dolayı risk değerlendirmesi doğru yapılamadığından kredi tahsisi gerçekleşmemiştir. Geleneksel yaklaşımda müşterinin krediyi ödeyememe olasılığı rastlantısal bir olay olarak algılandığından bu konu üzerinde istatistiksel bir yaklaşım yapılmamıştır.Fakat ekonomik düzende yaşanan iflaslardaki artış, kredi faiz marjlarındaki rekabet ortamı doğrultusundaki oynaklık, kredinin güvencesi olarak alınan teminatlardaki değer kaybı ve faaliyet yapısı gereği kredinin yanı sıra türev ürünlerde yaşanan artış kredi riskinin ölçümünün önemini artırmıştır.

Bu tez çalışmasında, kredi değerlendirmesinde sübjektifliği ortadan kaldıran ve çok sayıda başvuruyu kısa zamanda değerlendirmeyi saplayan yöntemlerden biri olan kredi skorlaması üzerinde durulmuştur. Değişik tipteki kredi borçlularının ortak özelliklerden ve seçilen faktörlere verilen ağırlıklardan yola çıkarak kredi kullandırılan bireylerin kredilerini ödeyememe olasılığı ve kredilerin ödenmemesinde hangi faktörlerin etkili olduğu nümerik olarak hesaplanmıştır. Ayrıca bu çalışmada kredi skorlama modelinin kurulumu, aşamaları, istatistiksel skorlama yöntemleri ve konuyla ilgili uygulamalara değinilecektir.

2. KREDİ SKORLAMA'NIN TANIMI TARİHÇESİ VE OLUŞTURULMASI

2.1. Kredi Skorlama (KS)

Kredi Skorlama başvuru sahibinin kredibilitesinin istatistiksel yöntemler kullanılarak ölçülmesi yöntemidir. Kredi skorlamada amaç bir kredi başvurusunda, müşterinin krediyi geri ödeyememe olasılığının hesaplanmasıdır. Bugün kredi riski yönteminin en önemli enstrümanlarından biri olan skorlama, başvuru sahibinin verilerini istatistiksel yöntemlerle analiz ederek gelecekte nasıl bir ödeme performansı sergileyeceğini tahmin etmeye yarayan bir yöntemdir(Lawrence and Solomon 2002). Bu yöntemle yüksek riskli müşterilere kredi vermeyi reddetmek finansal kurumun olası zararını azaltacak, düşük riskli müşterilere kredi vermek karını artıracak aynı zamanda müşterilerin ödeyemeyeceği kredileri almasını bir bakıma engelleyecektir. Kredi skorlaması sonucu, başvuruya dair bir skor elde edilir. Bu skor başvuru sahibinin başvuru formunda belirttiği bilgiler doğrultusunda katsayılar ile ağırlıklandırılmış modelden elde edilir. Bu model kredi kayıplarını bireysel başvuru seviyesinde öngörmeyi sağlamaktadır.

Geleneksel kredi değerlendirme yönteminde, kredi kararı bir değerlendirme uzmanının kredi ile ilgili parametrelere olumlu ya da olumsuz görüşünü verdikten sonra genel toplama bakılarak verilirken, kredi skorlamada her kritere belli bir skor verilerek bunun sonucunda istatistiksel bir skor hesaplanır (Sounders, 1996). Yeni başvuruya kredi verilmesi kararı başvuru skorlama adı altında incelenirken, mevcut müşterinin ileriye dönük davranışlarının takibi, limit artırım talepleri davranışsal skorlama adı altında incelenir.

Skorlama yöntemi, tüketicinin krediyi ödeyip ödeyemeyeceğini tahmin ederken neden belli bir davranışı göstereceğini açıklamakla ilgilenmez. Skorlamadaki diğer kabul ise tüketicinin bulunduğu çevrenin, tüketicinin talep ettiği kredinin sorunlu hale gelip gelmemesini öngörmeye yardımcı olmasıdır (Thomas, Edelman and Crook, 2002). Tüketicinin ödeme performansını etkileyen tüm özellikleri tahmin modelinde kullanılmaktadır. Ancak bir özellik istatistiki olarak anlamlı değil ise modele katılmaz. Bir diğer nokta istatistiki olarak anlamlı olabilecek bazı verilerin, kullanımı skorlamanın yapıldığı ülkenin yasalarınca yasaklanmış olduğundan dolayı, kullanılmamsıdır. Örneğin ABD'de din, ırk, cinsiyet gibi karakteristiklerin skorlamada kullanılması ayrımcılık yapabileceği gerekçesiyle yasaklanmıştır.

2.2. Kredi Skorlamanın Tarihsel Gelişimi ve Finans Sektöründeki Rolü

Kurumların müşterilerle ilgili gizlilik ilkesi, değişken ve gelişen teknoloji koşulları, müşterilerin artan hizmet beklentileri ve gerekse kurumların riski kontrol altında tutmak istemeleri aynı zamanda artan rekabet koşullarına ayak uydurmak zorunda olmaları, kredi veren kuruluşları daha hızlı ve daha doğru kredi vermelerini sağlayacak arayışlara itmiştir.

Kredi vermenin geçmişinin 5000 yıla dayanmasına karşın kredi skorlamanın geçmişi son 50-60 yıla dayanmaktadır. 1941 yılında Durand, topluluk içindeki grupların ayrıştırılmasında aynı karakteristik değişkenlerin ayrıştırıcı etkiye sahip olduğundan yola çıkarak iyi ve kötü kredilerin belirlenebileceğini ortaya koymuştur. Yaptığı bu çalışma Amerika Ekonomik Araştırmalar Bürosu için yapılan bir projede kullanılmıştır (Thomas, Edelman and Crook, 2002).

Kredi riskinin yönetiminde asıl sorunlar İkinci Dünya Savaşı'nın başlamasıyla ortaya çıkmıştır. Finans kuruluşlarının kredi uzmanlarının askere gitmesi bu konuda uzman sıkıntısının yaşanmasına neden olmuş bunun sonucunda, kredi kuruluşları ve posta ile satış yapan şirketler, analistlerden kredi kararlarını hangi kriterlere göre verdiklerini yazmalarını istemişlerdir.Bunun üzerine Speigel Firmasının yöneticisi Henry Wells, istatistiksel yöntemlere dayanan kredi skorlama sistemini oluşturmuştur. Bu uygulamada, kredilere daha önceden belirlenen kararteristiklere dayanarak puan verilmesi esas alınmış fakat karakteristikler arasındaki ilişkiler gözetilmemiştir (Fabozzi, 1999).

1950'lerde San Fransisco'da Bill Fair ve Earl Isaac tarafından bu konudaki ilk danışmanlık şirketi kurulmuştur. Bu şirket finans kuruluşlarına, perakendecilere ve posta ile satış yapan şirketlere danışmanlık hizmeti vermiştir (Fabozzi,1999). 1960'larda kredi kartlarının da ekonomiye dahil olması, kredi sektöründe gelişmeler sonucunda gün geçtikçe gelen kredi başvurularının artması, bu başvuruların insan gücüyle uygulanabilirliğinin oldukça zor olması ve bilgisayar teknolojisindeki gelişmeler kredi skorlamanın çok iyi bir tahmin edici ve karar verme aracı olarak görülmesini sağlamıştır. 1970'lerde Visa ve Master rekabet koşullarından dolayı kredi kartlarını ücretsiz hale getirmişlerdir. Kredi kartı başvurusunun fazla olması ve kredi riskinin geleneksel yöntemde deneyimlere dayalı olan ekspertiz modeliyle yapılması sonuçların güvenilirliği konusunda sorunlara yol açmıştır. Aynı zamanda hem kredi tahsis edenler hem de kredi başvurusu yapanlar açısından değerlendirmenin makul sürede gerçekleşmesi isteği kredi kuruluşlarını kredi skorlamaya

teşvik etmiştir(Lewis,1992). 1980'e gelindiğinde bu sistem kredi kartı tahsisinde tam olarak kullanılmaya başlanmıştır.1975'de Amerika'da çıkarılan Eşit Kredi Fırsatları Kanunu (Equal Credit Oppurtunity) ile birlikte kredi skorlama güvenli ve ampirik olarak üretilmiş yasal bir teknik olmuştur(Thomas, Edelman and Crook, 2002).

Ülkemizde ise kredi skorlama özellikle bankacılık sektöründe kullanım alanı bulmaktadır. Hızla tüketim ekonomisi çizgisine kayan ekonomimizde bankalarında "kumbara" ile simgelenen yaklaşım yerine taşıt kredileriyle, konut kredileriyle ve kredi kartlarıyla bireysel tüketimi özendiren bir modaya ayak uydurduğu gözlenmektedir(Esener, 1999). Klasik bankacılık anlayışı ürün bazlı karlılığa dayanırken günümüz bankacılığı müşteri bazlı karlılığa dayanmaktadır. Amaç müşterilerin borç ödemelerini minimize etmek yerine bu müşterilerden daha fazla nasıl gelir elde edileceği temeli üstüne kurulmuştur. Bunun sonucunda müşterilere sınırsız sayıda hizmet sunmak ve taleplere doğru ve hızlı cevap vermek için profesyonel uygulamalara gidilmiştir. Günümüzde bazı bankalar "dışsal skorlama" yı yani kredi skorlama kuruluşlarının hazırladıkları skorkartları kullanırken bazıları ise "içsel skorlama" olarak adlandırılan kendi bünyelerindeki risk yönetimi birimlerinin hazırladıkları skorkartları kullanmaktadırlar. Aynı şekilde ilgili firmalardan sistemlerin güncellenmesinde danışmanlık hizmeti de alınmaktadır.

2.3. Kredi Skorlama Yönteminin Diğer Yöntemlere Göre Üstünlükleri

Her geçen gün artan kredi talepleri beraberinde değerlendirmeyi yapacak işgücünün istihdam sorununu da getirmektedir. İnsan algısının zamana, iş yoğunluğuna, çalışan ortama göre kişiden kişiye değişmesi kredi kararlarında tutarsızlığa yol açmaktadır. Kredi skorlama, kredi değerlendirme sürecinde sübjektifliği ortadan kaldırmakta objektifliği sağlamaktadır. Kredi skorlama yöntemi ile kredi karar yetkisinin şubelerde olduğu durumlarda merkezin politikalarını uygulatmasında etkinlik kayıplarını önlemekte ve değişik noktalardan değişik sayılardaki başvurular, kim tarafından alınırsa alınsın tek bir karar modeline tabi tutulmakta bu şekilde bireyler arasında eşitlik sağlanmaktadır. Kredi politikalarında yapılacak değişikliklerle şube ve merkezdeki çok sayıda personeli detaylı bir eğitime tutmaya gerek kalmaksızın yalnızca skorlama sisteminde değişikliklere gidilerek yeni politikaların adaptasyon süreci kısaltılmaktadır(Lewis, 1992). Geleneksel yöntemde çeşitli meslek gruplarına ya da yaş gruplarına kredi verilmez gibi önyargılar vardır ve bu başvurulara olumsuz yanıt verilir. Oysa kredi skorlamada, temerrüde düşme olasılığı yüksek kabul edilen gruplara bile diğer karakteristiklerin değerlendirilmesi sonucu kredi tahsisi mümkündür.

Geleneksel yönteme kıyasla kredi skorlamada daha fazla karakteristik göz önünde bulundurulup kredi kararına etkide bulunur(Mays, 2004).

Kredi skorlama sayesinde başvuru başına harcanan zaman azalmaktadır. Başvuru verileri önceden bilgisayara yüklenmiş derece algoritmaları sonucu değerlendirilmekte kısa sürede analiz edilip karar verilmektedir. Çok sayıda başvurunun kısa sürede değerlendirilmesi ise daha fazla başvuru çekmekte ve kabul sayılarının mutlak değer olarak da artması müşteri memnuniyetini beraberinde getirmektedir. Banka açısından bakıldığında, eğer içsel skorlama uygulanmaya başlanmadıysa ve dışarıdan destek alınıyorsa skorlama için ilk başta ödenen maliyet yıllara dağıtıldığında, bankanın bir yılda topladığı yüz binlerce kredi talebinin tahsisi, takibi ve diğer destek prosesleri için gerek duyulan personele ödenen tutardan daha düşük kalmaktadır(Mays ,2004).

Hem geleneksel yöntem hem de kredi skorlama tek başına sorunlu krediyi önleyemez. Yapılan değerlendirme sonucu yüksek skor elde eden müşterinin daha sonra ciddi bir hastalığa yakalanması, doğal afette hayatını kaybetmesi ya da işini bir anda kriz sonucu kaybetmesi bu metotlarla öngörülemez. Fakat yapılan çalışmalar sonucu kredi skorlamanın geleneksel yönteme göre sorunlu kredi oranını azaltmada daha etkili olduğu görülmüştür. Öte yandan kredi tahsisinin ardından kullanılan davranışsal skorlama, kredi takibinde birçok kolaylık sunmaktadır. Bu yöntemle müşteri önceden belirlenen aksiyonları beklemeden alma şansını yakalayacaktır.Örneğin kişi temerrüde düşmüşse ya da limit aşımınında ise, müşterinin hesaplanmış risk profiline göre limit artışı yapılmakta ya da limit aşımının giderilmesi sağlanmaktadır.

Yaşanan rekabet ortamı, bankaları klasik dağıtım kanalları dışına çıkarak alternatif dağıtım kanallarına yöneltmektedir. Her kurum müşterinin olduğu yerde kredi sunarak müşterileri çekmek istemektedir. Bu durumda hızlı ve yaygın kredi sürecine yol açmaktadır. Bankalar ve kuruluşlar bu prosesi kredi skorlama sisteminin elektronik ve hızlı başvuru işleme ve iş akışı entegrasyonuyla sağlanmaktadır. Kredi skorlama müşterilere özel imkanlar geliştirmeye de imkan vermektedir. Örneğin yapılan skorlama sonucu yüksek derece elde etmiş bir müşteriye düşük risk priminden dolayı daha düşük kredi faizi teklif edilebilmektedir(Mays, 2004).

2.4. Kredi Skorlama ve Veri Madenciliği

Veri madenciliği, verilerde anlamlı ilişkileri ve yapıları tespit etmek için veri analizi ve araştırma tekniğidir. Madenciliğe benzer şekilde, bu teknikte de gerekli olan verinin nereden ve nasıl bulunacağı belirlenmeye çalışılır. Son yıllarda, şirketler özellikle de bankalar ve perakendeciler, müşterileri hakkında sahip oldukları bilginin tanımlanmasının değerini kavramışlardır. Elektronik fon transferinin ve bağlılık kartlarının yaygın olarak kullanılmasıyla birlikte, bu tür şirketler müşterileri hakkında bilgileri kolaylıkla toplayabilmektedirler. Bilgisayar teknolojisi de toplanan büyük miktarlardaki verilerin analizinde kolaylık sağlamaktadır. Rekabetin artması, ikame ürünler ve internet gibi kolay iletişim kanalları müşterilerin kolaylıkla yer değiştirmesine neden olmaktadır. Böylece, müşteri davranışlarını anlamak ve analiz etmek büyük bir önem arz etmektedir. Bu nedenle şirketler veri ambarları oluşturmak ve veri madenciliği gibi teknikleri kullanmak için çok büyük miktarda paralar harcamaktadırlar. Veri madenciliğinin ana tekniklerine göz atıldığında, bu tekniğin kredi skorlamada çok başarılı neticeler alınmasını sağladığı görülmüştür. Temel veri madenciliği teknikleri arasında veri özeti (data summary), değişken düşürme (variable reduction), gözlem kümeleme (observation clustering), tahmin ve açıklama (prediction and explanation) yer almaktadır. Frekans, ortalama, varyans ve çapraz tablolama gibi standart tasvir edici istatistikler verilerin özetlenmesi için kullanılır. Aynı zamanda, kesikli sınıflar için sürekli değişkenlerin kategorize edilmesi için de çok kullanışlıdır. Betimsel istatistikler, KS'de çok kullanılan kaba sınıflama tekniğidir. Hangi değişkenlerin en önemli olduğunu tespit etmek, gereksiz olanları analizden çıkarmak KS uygulamalarında çok sık kullanılan teknikler olduğu gibi veri madenciliği uygulamalarında da kullanılmaktadır. Müşterileri satın aldığı farklı ürünlere göre veya diğer özelliklere göre gruplara ayırmak diğer bir veri madenciliği aracıdır. KS aynı zamanda müşterilerin davranışlarına göre farklı gruplar oluşturur ve her bir grup için ayrı bir skor kart düzenlenir. Bu fikir alt ana kütlelerin segmentasyonu anlamına gelir ve böylece her bir alt ana kütle için bir skor kart profili oluşturulur. Esasında KS'de kullanmak için geliştirilen, örneğin gelecek yıl hangi müşterinin hangi finansal aracı kullanacağını tahmin eden teknikler veri madenciliği için de çok önemlidir. Gerçekte, veri madenciliğinin kullandığı segmentasyon analizi belirli tip davranışlara sahip segmentleri göstermek için kullanılır. Böylece, veri madenciliği KS için çok önemli olmazsa olmaz bir teknik ve teknoloji olup,daha geniş bir alanda uygulanması gerekir. KS'yi uygulamadaki hatalardan, eksikliklerden korunmak ve diğer alanlarda

uygulayabilmek için, veri madenciliğini KS ile birlikte kullananlar çalışmalarında çok daha fazla başarıya ve gelişmeye ulaşacaklardır(Thomas, Edelman and Crook, 2002).

2.5. Örnekleme Seçimi

Kredi Skorlama (KS) ve Davranış Skorlaması (DS) ile ilgili tüm metotlar skorlama (scoring) sistemini geliştirmek için müşterilere ait geçmiş verilere ve onların hikayelerine ihtiyaç duymaktadır. Örnekleme seçiminde dikkat edilmesi gereken iki husus mevcuttur. Birincisi; örnekleme gelecekte başvurusu mümkün olan kişileri temsil etmelidir. İkincisi; örnekleme ödeme alışkanlığının iyi ya da kötü olduğunu yansıtabilecek kadar yeterli bilgiyi içine almalıdır. Bunu en iyi açıklayacak örnekleme, mümkün olan en yakın geçmişte borç almış kişilerin bilgilerinin yer aldığı veri tabanıdır. Başvuru skorlaşmasında (Application Scoring) bu son 12 ay, davranış skorlamasında ise son 18-24 aydır(Thomas, Edelman and Crook, 2002). Örnekleme seçiminde dikkat edilmesi gereken diğer bir nokta; örnekleme büyüklüğünün ne kadar olacağı ve iyi-kötü kredilerin hangi unsurlara göre ayrılacağıdır. Örneklemedeki iyi müşteri-kötü müşteri oranı eşit mi olmalıdır, yoksa ana kütlede olduğu sekliyle mi temsil edilmelidir? Ana kütledeki oranına göre iyi-kötü oranı belirlendiğinde kötü kredilerin açıklanmasına yetecek kadar veri örneklemede bulunmayacağından dolayı genelde bu oran 50:50 olarak kabul edilir. Örneklemede iyi-kötü değişkenlerinin dağılımları aynı değilse buna izin verecek örneklemeyi elde etmek için sonuçların düzeltilmesi gerekir. Lewis, örnekleme büyüklüğü ve iyi-kötü kredi oranı için; 1500 iyi ve 1500 kötü sayısını yeterli bulmuştur(Lewis, 1992). Pratikte ise, çok daha büyük örneklemeler kullanılır. Örnekleme mevcut ana kütleden rastgele seçilmişse bunun gerçekten rastgele olduğuna emin olmak gerekir. Eğer ana kütle listesindeki iyilerden her 10 tanede bir tanesi seçilirse örneklemede bunun temsili %10 olacaktır. Bununla birlikte, şubelere gidilmesi zorunlu olduğunda ilk olarak yapılması gereken iyilerin kırsal ve kentsel olarak rastgele seçildiğinden emin olmaktır. Ya da örneklemenin seçildiği tarihte belirli ürünlerin pazarlaması yapılıp yapılmadığına veya belirli kitlelere hitap edilip-edilmediğine bakılmalıdır. Örneğin belirli bir ay örnekleme olarak alındığında o ayda bir ürünün pazarlaması yapılıyor olabilir ve bu ürün gençlere hitap ediyorsa örneklemedeki genç oranı daha yüksek çıkması kaçınılmazdır.

2.6. İyi Müşteri-Kötü Müşteri Tanımlaması

Skor kart geliştirilmesinde aşamalardan biri iyi müşteri-kötü müşteri sınıflamasının ne şekilde yapılacağıdır. Bazı müşterileri kötü olarak tanımlama diğer tüm müşterilerin iyi

olduğu anlamına gelmez. Müşterilerin iyi-kötü olması dışında en az 2 seçenek daha vardır. Birincisi“tanımlanamayanlar”, ikincisi “yeterli gözlem değeri olmayanlar”.Genellikle KS geliştirmede ödemeler arasında 3 dönem sorun yaşananlar sorunlu kredi olarak nitelendirilir. Tanımlanamayanlar 2 dönem problem yasayıp 3. döneme düşmeyenlerdir. İyi-kötü ayrımı ne şekilde yapılırsa yapılsın KS tekniği etkilenmez. Örneklemeden“tanımlanamayanlar” ve“yeterli gözlem değeri olmayanlar” ı çıkarıp sadece iyi-kötü sınıflaması için skor kart geliştirmek gerekir. Tabii ki iyi-kötünün farklı sınıflandırılması farklı skor kart sonuçları geliştirecektir. Diğer bir sorun da kötü kredinin ekstrem şekilden tanımlanmasından kaynaklanmaktadır. Bu durumda çok az olan kötü kredi nedeniyle modelin güvenilirliği sarsılabilir(Thomas, Edelman and Crook, 2002)

2.7. Veri Kaynakları

İyi-kötü kredilerin ayırt edilebilmesi için karakteristikler gereklidir ve bu karakteristikler farklı yollarla elde edilebilir. Bunlardan en önemlisi kredi kartı başvuru formudur. Yapılacak analiz çeşidine uygun karakteristiklerin elde edilebilmesi için çeşitli formatlarda kredi kartı başvuru formu düzenlenebilir. Bu şekilde verilen cevaplar şıklara bağlı olarak aynı formatta olacaktır ve analiz aşamasına kolaylıkla geçilebilecektir. Başvuru formundaki karakteristiklerin ne olacağına karar verilirken meşru cevaplar alınabilecek şekilde format oluşturulmalıdır. Örneğin, ev durumu karakteristiğine bakıldığında “size ait”, “kira”, “aile fertlerinden birine ait”, lojman” seçenekleri olduğu varsayılsın. Tabi bu daha geniş seçenekli de olabilirdi. Sorun belirlenen seçeneklere “hayır” cevabı verildiğinde ne yapılacağıdır. Genel yaklaşım her karakteristik için “diğer” seçeneği eklenmesidir. Fakat cevabı gerçekten “hayır” olan karakteristikler de vardır. Bu da çelişki yaratabilir. İş karakteristiği yüzlerce iş seçeneği olduğu durumlarda kodlanması zor bir alandır. Unvanlarla ilgili bilgiler çok açıklayıcı olamamaktadır. Örneğin başvuru yapan kişi yönetici seçeneğini işaretlediğinde; çok büyük bir fabrikada yönetici olduğu anlamı çıkabileceği gibi, küçük bir işletmede yönetici olduğu anlamı da çıkabilir. Daha genel yaklaşımlarla bu asılmak istenmektedir. Gelir karakteristiği de net bir şekilde ifade edilmediği takdirde karışıklığa neden olabilir.Aylık gelir mi, toplam gelir mi, ya da hane halkının toplam geliri mi? Genellikle başvuru formlarında bunların her biri ayrı ayrı istenilmektedir ve analize farklı değişkenler olarak girmektedirler.Sonuç olarak, verilerin geçerliliğini denetlemek gerekir. Mesela 18 yaş altı ya da 100 yaş üstü değerler işaretlenebilir. Farklı karakteristikler için

karakteristiğin dağılımı incelenerek uç olan durumlar gözlemlenebilir. Bu şekilde veri doğrulama hilekarlık tespitinde denenen bir yoldur.

Verilen cevaplardaki uyuşmazlık çok sık gözlemlenebilmektedir. Diğer bir problem de kredi kartı başvurusu yapan kişinin formu şube yetkilisiyle birlikte doldurup doldurmayacağıdır. Şube yetkilisi müşterinin başvurusunun kabul edilebilmesi için ne tür cevaplar vermesi gerektiği şeklinde müşteriyi yönlendiriyor olabilir. Bu durumda müşterinin gerçek bilgilerinden ziyade oluşturulan skor kartta başvurusunun onaylanması için gerekli olan yanlış bilgiler girilmiş olacaktır ve bu da müşteri hakkında onay aşamasında yanlış karar verilmesi anlamına gelir.

2.8. Alt Ana Kütle Belirlemek (Determining Subpopulation)

Bundan sonraki iki alt bölüm skorlama sisteminde hangi değişkenlerin ne şekilde kullanılması gerektiğiyle ilgilidir. Değişkenlerin belirlenmesinde önemli bir teknik de ana kütlenin alt ana kütlelere ayrılarak her bir sınıf için ayrı skor kart geliştirilmesidir. Bu teknik istatistiksel nedenlerle olduğu kadar kredi politikası açısından da uygulanır. Örneğin DS'de ilgili banka ile son dönemlerde çalışan müşteriler olduğu gibi uzun süredir çalışan müşteriler de mevcuttur. Örneğin 6 aylık ortalama bakiyelere bakıldığında ikinci tip müşteriler için ulaşılabilir bir veridir fakat birinci tip müşteriler için ulaşılabilir bir veri değildir. Bankanın uygulayacağı bir başka politika da genç müşterileri yaşlı müşterilerden farklı olarak değerlendirmek olabilir. Bu durumda da 25 yaş altı ve üstü için ayrı skor kart oluşturma yoluna gidilebilir. Ana kütlenin alt ana kütlelere ayrılmasının istatistiksel nedeni ise bir karakteristik ve diğer karakteristikler arasında çok fazla etkileşim olmasıdır. Bundan dolayı karakteristiklerin farklı davranışları için farklı skor kart geliştirilmektedir. Uygulamacılar diğer karakteristiklerle en çok etkileşim halinde olan karakteristik için Sınıflandırma Ağacı (Classification Tree) kullanarak skor kart oluşturabilirler. Sınıflandırma ağacındaki en iyi ayrılma noktası farklı skor kart oluşturulması için uygun alt ana kütleleri göstermektedir. Alt ana kütlelere bölmede politik nedenler istatistiksel nedenlere göre daha fazla kullanılmaktadırlar. Banasik'in belirttiği gibi alt ana kütlelere bölme her zaman tahmini geliştirmez. Alt ana kütleler farklı değilse, daha küçük örneklemeyle çalışmak skor kartın performansını, alt ana kütlelere bölmenin sağladığı avantaja göre daha çok etkileyecektir. Bu durumda alt ana kütleye bölmenin bir anlamı yoktur(Banasik ,vd., 1996).

2.9. Karakteristiklerin Genel Olarak Sınıflandırılması

Her bir karakteristik verilecek cevapların birbirleriyle ilişkilerine göre alt ana kütlelere ayrılmalıdır. Alt ana kütlelere bölme işlemi cevapların kategorik-sürekli olmasına bağlı olarak iki farklı şekilde yapılabilir. Kategorik veriler için; çok fazla farklı cevap şıkkı olabilir ve her bir cevapla ilgili yeterli çoğunluk sağlanamadığı için güvenli bir analiz yapılamaz.

3. KREDİ SKORLAMADA KULLANILAN YÖNTEMLER

3.1. Dikstiriminant Analizi

Diskriminant analizi belirli aralıklarda tanımlanmış kategorik olarak belirlenen bağımlı (kriter) değişken ile bağımsız (kestirimci) değişkenler arasında fonksiyonel bir ilişki kurarak gruplar arası ayırımı sağlayan bir data analiz tekniği olarak tanımlanabilir(Malhotra ve Naresh, 1993).Diskriminant analizi, değişkenlerin değerlerinin gruplar içinde olabildiğince az, gruplar arasında olabildiğince fazla dağıldığı lineer bileşimler bulmaya çalışır(Yıldız, 1995). Diskriminant Analizi, iki ve daha fazla sayıda grubun çok sayıda değişkene göre karsılaştırılmasını sağlayan bir tekniktir. Analizin amacı, grupların hangi değişkenler açısından birbirinden farklılaştığının ortaya çıkarılmasıdır. Diğer bir ifadeyle, grupları ayırıcı özelliklerinin belirlenmesidir.

Diskriminant analizinin temel görevleri;

- Grupların birbirlerinden ayrılmasını sağlayan bir model kurmak
- Yeni gözlemleri uygun sınıflara yerleştirmek

Diskriminant analizi varsayımları;

- X veri matrisi çok değişkenli normal dağılım göstermelidir.
- Varyans kovaryans matrisleri homojen olmalıdır.
- Değişkenlerin ortalamaları ve varyansları arasında anlamlı korelasyon bulunmamalıdır.
- Degişkenler arasında çoklu bağımlılık (multicollinearity) bulunmamalıdır.
- X matrisi grupların birbirlerinden ayrılmasında rol oynamayacak gereksiz değişkenler içermemelidir (redundancy).

Araştırmacının amacı doğrultusunda farklı çözüm teknikleri geliştirilmiştir. Bunlardan bazılarının aşağıdaki gibi özetlenmesi mümkündür.

- Bireylerden ya da birimlerden oluşan iki veya daha çok sayıda grubun ortalama özellikleri açısından ,istatistiksel olarak anlamlı bir ayrım gösterip göstermediklerinin belirlenmesi
- Grupların ayrımını yapmada her bir kestirimci değişkenin etkisinin belirlenmesi
- Herhangi bir birimin hangi gruba dahil olabileceğinin kestirilmesi
- Grup içi değişime oranla gruplar arası ayırımı maksimize eden kestirimci değişkenlerin bunları en iyi yansıtan doğrusal (veya doğrusal olmayan) bir fonksiyona dönüştürülmesi. Burada söz konusu olan bireyleri hatalı sınıflandırma olasılığını minimum yapacak biçimde bir fonksiyon elde etmektir(Yıldız,1995).

Diskriminant analizi, grup kovaryans matrislerinin eşit olup olmamasına göre farklı biçimlerde uygulanmaktadır. Grup kovaryans matrisleri benzer olduğunda Dogrusal Diskriminant Analizi, benzer olmadığında da Karesel Diskriminant Analizi kullanılır(Özdamar,2002b).

3.2. İki Grup Karesel Diskriminant Analizi

İki gruplu diskriminant analizinin temeli bireyler hakkında toplanan, çoklu ve korelasyonlu verilerden hareketle, iki grup öğeleri arasında maksimum ayırımı sağlayan tek bir karesel bileşim modeline geçmeye dayanır. Böylece çok değişkenli profiller tek değişkenli veriler kümesine dönüştürülür.

Kestirimci değişkenlerin dağılımlarına ilişkin belirli varsayımların yapılması halinde iki grubun ortalamalarından oluşan vektöre ilişkin ortalamalara karşılık gelen odak noktaların eşitliği hipotezi sınanabilir. Eğer farklı oldukları sonucuna varılırsa, söz konusu iki gruba yeni çok değişkenli profiller ekleyerek, bunların hangi gruba dahil olacaklarını diskriminant analizi ile belirlemek mümkün olur. Bu da her profil için bir diskriminant puanı (skoru) hesaplayarak ve bu değeri örneklemeden elde edilen karesel bilesen değeriyle kıyaslayarak yapılabilir. Eğer söz konusu değer sınır değerinin bir yakasındaysa 1. gruba, diğer yakasındaysa 2. gruba sokulur. Sınır değerine eşitse herhangi bir gruba sokulabilir(Yıldız, 1995).

Diskriminant analiz fonksiyonlarını oluştururken ele alınan yapının varsayımlarını belirlemek genel yapının formüle edilmesi açısından önem kazanmaktadır. Her şeyden önce kriter (bağımlı-nitel) *Y* değişkeninin, gösterge değişkenle kodlanacağı (0,1) belirtilir. Örnekleme büyüklükleri (gruplardaki gözlem sayıları) eşit olarak alındığında, ortalamalar vektörünü oluşturan elemanların toplamı sıfır olacaktır. Burada grup değerleri genel ortalamadan sapmalar olarak alındığında, birinci grubun değerleri negatif, ikinci grubun pozitif sapmalardan oluşacak ve toplamları sıfır olacaktır(Yıldız,1995).

A ve B isimli iki toplum düşünüldüğünde bu toplumlarda *n* birimlik *p* tane birbirleri ile ilişkili gözlemler var ise *X* veri matrisinde gözlemler toplumdan topluma farklılıklar sergileyecektir. *X* matrisi *A* toplumundan alınan X_1 gözlem matrisi ve *B* toplumundan alınan X_2 gözlem matrisinden oluşmaktadır. Fisher, çok değişkenli *X* gözlem matrisini *A* ve *B* toplumlarına ilişkin tek değişkenli *y* değerlerine dönüştürmeyi ve bu *y* değerlerinin toplumlara göre birbirlerinden olabildiğince farklı olmalarını sağlamayı amaçlamıştır. *y* 'leri ise *X* gözlem matrisinin karesel bileşenleriyle elde etmiştir. X_1 ve X_2 gözlem matrislerine ait kovaryans matrisleri S_1 ve S_2 'dir. Bazen katsayılar ölçeklendirilerek (scaled) ya da normalize edilerek kullanılmaktadır. Normalizasyonun temel amacı katsayıları genel katsayılar içinde ağırlıklandırarak elemanların kolay yorumlanmasını sağlamaktır.

İki tür normalizasyon yaklaşımı vardır;

$$b^* = \frac{b}{\sqrt{b'b}}$$

$$b^* = \frac{b}{b_1}$$

burada b_1 en büyük karesel bileşendir.

Gruplar arasındaki farklılığı maksimize edecek bir diskriminant fonksiyonu aracılığı ile grupları birbirinden ayırmak mümkün olacaktır. Bu nedenle ortak diskriminant fonksiyonu belirlenir.*i* ve *j* grupları arasındaki diskriminant fonksiyonu;

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p$$

şeklinde yazılır. Bu fonksiyondaki b_i karesel bileşenleri de ortalama fark vektörü ile aşağıdaki gibi olur.

$$b = (S_1^{-1} - S_2^{-1})\left(\overline{x_1} - \overline{x_2}\right)$$

$A = (S_1^{-1} - S_2^{-1})$ olmak üzere her bir grubun b_i katsayılar vektörü

$$b_i = A\left(\overline{x_1} - \overline{x_2}\right)' \qquad i = 1,2,...,g \qquad j = 1,2,...,p$$

biçiminde hesaplanır.

Sabit değer b_0 katsayısı ise;

$$b_0 = -(1/2)\overline{x}(S_1^{-1} - S_2^{-1})\overline{x}$$

Diskriminant analizi her bir grup için birer diskriminat fonksiyonu hesaplar.

$$Y_i = b_{0i} + b_{1i}X_1 + b_{2i}X_2 + ... + b_{pi}X_p \quad i = 1,2,...,p$$

Bu fonksiyonda b_{0i} sabit değeri, b_{ij} ise karesel bileşenleri (kanonik değişkenler) belirtmektedir.

S_1 ve S_2 kovaryans matrisleri ve $\bar{x}_i$ *i.* grup ortalama vektörü olmak üzere her bir grubun b_i katsayılar vektörü;

$$b_{ij} = (S_1^{-1} - S_2^{-1})(\bar{x}_i) \qquad i = 1,2,...,g \qquad j = 1,2,...,p$$

biçiminde hesaplanır.

Sabit değer ise;

$$b_{i0} = -(1/2)\overline{x}_i\left(S_1^{-1} - S_2^{-1}\right)\overline{x_i}$$

biçiminde hesaplanır.

Gruplara göre belirlenen sabit ve kanonik katsayılar değişken değerleri ile çarpılarak karesel diskriminant fonksiyonları hesaplanır.

1. grup için diskriminant fonksiyonu;

Y_i=b_{10}+$b_{11}X_{11}$+b_{12} X_{12}+…+b_{1p} X_{1p}

eşitliğinde değerler yerine konularak hesaplanır.

2. grup için diskriminant fonksiyonu;

$Y_2=b_{20}+b_{21}X_{21}+b_{22}X_{22}+...+b_{2p}X_{2p}$

eşitliğinde değerler yerine konularak hesaplanır.

p değişkenli *g* grup arasındaki karesel uzaklığı veren ve Mahalonobis tarafından ileri sürülen D^2 uzaklığı aşağıdaki gibi hesaplanır.

$$D_{ij}^2 = \left(\overline{x_i} - \overline{x_j}\right) S^{-1} \left(\overline{x_i} - \overline{x_j}\right)$$

D^2 uzaklığının *i* ve *j* gruplarını birbirinden ayırmada etkin rol oynayıp oynamadığı test edilebilir.Bu amaçla Hotelling T^2 yaklaşımı ile D^2'nin önemliliği test edilebilir.

$$T^2=\left(\frac{n_1 * n_2}{n_1 + n_2}\right) D^2$$

$T^{2'}$nin önemliliğinin test edilmesi için *F* yaklaşımından yararlanılır.

$$F=\frac{(n_1 + n_2 - p - 1)}{p(n_1 + n_2 - 2)}\ T^2$$

F'nin önemliliği $p,(n_1+n_2-p-1)$ serbestlik dereceli *F* dağılımının kritik değerleri kullanılarak belirlenir. [$F(\alpha, p, (n_1 + n_2 - p - 1)$] (Yıldız, 1995)

3.2.1. Kovaryans Matrislerinin Eşitliğinin Sınanması

Eğer diskriminant fonksiyonunu hesaplamadan ya da Hotelling T^2 testini uygulamadan önce kovaryans matrisleri homojenligi test edilmek istenirse, başka bir test gerekli olacaktır.

Barlett, iki ya da daha fazla kovaryans matrisinin eşitliğini test etmek için bir χ^2 yaklaşımı geliştirmiştir. Daha sonra Box, F testini temel alan daha gelişmiş ve karmaşık bir yöntem ortaya atmıştır.

İki toplumdan oluşan ve *p* değişkeninin kullanıldığı bir örnekleme düşünülsün.Ayrıca *G* kovaryans matrisi S_g=(*g*=1,2,…….*G*) ve toplam gözlem sayısı $\sum_{g=1}^{G} n_g = n$ olan toplanmış gruplar için kovaryans matrisi *S* göz önünde tutulsun.Bu durumda aşağıdaki sıfır hipotezi yazılabilir.

$$H_0 = \sum\nolimits_1 = \sum\nolimits_2 = ... = \sum\nolimits_g = ... = \sum\nolimits_G$$

$$H_1 = \sum\nolimits_1 \neq \sum\nolimits_2 \neq ..... \neq \sum\nolimits_g \neq ... \neq \sum\nolimits_G$$

Burada $\sum_g$, *g* grup için ana kütle kovaryans matrisini göstermektedir. Barlett tarafından geliştirilen kovaryans matrislerinin eşitliğini testinde kullanılacak, iki ve daha çok grup için genel formül aşağıdaki gibidir.

$$B = (n - G)\ln|S| - \sum_{g=1}^{G}(n_g - 1)\ln|S|$$

Bulunan sonuç $s.d = \frac{1}{2}[(G-1)(n)(n+1)]$'li χ^2 değerleriyle karşılaştırıldığında eşitlik reddedilirse lineer diskriminant fonksiyonu uygulanamaz. Eğer kovaryans matrisleri eşit olmazsa odak noktaların eşitliğinin sınanmasında sistematik bir hata ortaya çıkar. İki gruplu durumda kovaryans matrisleri eşit olmaması halinde Hotelling T^2 testindeki sıfır hipotezinin daha sık kabul edildiği görülecektir.

Dahası, sınıflandırma hataları terimleriyle birbirine eşit olmayan kovaryans matrisleri halinde lineer bir diskriminant fonksiyonunun kullanımı, bu matris en çok toplanmış gruplar içi kovaryans matrisine katkıda bulunduğu için, büyük kovaryans matrisine sahip gruba pek çok noktanın atanmasına yol açacaktır.

Cooley ve Lohnes'in değindiği gibi dikkate değer ölçüdeki büyük örneklemelerde, kovaryans matrislerinin eşitliği sıfır hipotezinin, reddedilmesinin mümkün olduğu varsayımı

genel olarak kabul görmüştür. Böylece odak noktalarının eşitliğinin testinde Hotelling T^2 testinin yeterince kuvvetli olduğu varsayımı ile çoğu araştırmacı kovaryans matrislerinin eşitliği üzerinde durmaz. Sonuçta kuvvetlilik (robustness) konusunda yeni gelişmeler olmadan bu ilksel testlerin kabul edilmesi önerilir(Yıldız, 1995).

Buraya kadar diskriminant analizi için en iyi ayırıcı değişkenlerin bilindiği varsayımıyla modeller incelenmiştir. Tabii ki potansiyel ayırıcı değişkenlerin olduğu fakat diskriminant analizi için en iyi ayırımı sağlayan değişken setinin bilinmediği durumlar da söz konusu olabilir. Adımsal Diskriminant Analizi (ADA) diskriminant fonksiyonu için en iyi değişken setini seçmede kullanılan faydalı bir tekniktir (Sharma, 1996). ADA ileri doğru seçme (forward selection), geriye doğru seçme (backward selection) ve adımsal seçme (stepwise selection) olmak üzere üç başlık altında incelenmektedir.

Adımsal seçme tekniği ileri doğru seçme ve geri doğru seçme tekniklerinin kombinasyonu şeklinde çalışmaktadır. Diskriminant fonksiyonunda hiçbir değişken yokken analiz başlar ve her adımda yeni değişkenler eklenir ya da elenir. Diskriminant fonksiyonundaki değişkenler istatistiksel kriter olarak belirlenen ayırma gücünden (discriminating power) daha düşükse modelden çıkarılır ya da tam tersi şekildeyse modelden elenir. Süreç hiçbir değişkenin modele girmediği ve çıkmadığı anda durdurulur (Sharma, 1996).

Değişkenler arasında ilişki yoksa üç teknik de aynı sonucu verecektir. Bununla birlikte verilerde ciddi bir çoklu doğrusal bağlantı (multicollinearity) problemi varsa sonuçlar birbirlerinden oldukça farklı çıkacaktır. Araştırmacılar veri setlerinde çoklu doğrusal bağlantı olduğuna inanırlarsa genellikle adımsal diskriminant analizini seçerler. Genellikle kullanılan seçme kriteri Wilks' Λ, Roo's V, Mahalanobis Kareli Uzaklıklar (Squared Distance), Gruplararası F Oranıdır (Between Group F Ratio). Bunlardan uygulamalarda kriter olarak daha sık seçilen Wilks'in Λ'sı "Λ" açıklanacaktır.

"Λ"

$$W = \frac{SS_w}{SS_b + SS_w}$$

"Λ" grup içi kareler toplamının kareler toplamına oranıdır. Her adımda diskriminant modelindeki mevcut değişkenlerden sonra en küçük "Λ"değerine sahip olan değişken modelden çıkarılır. "Λ" değeri gruplar arası ayırma ve grup içi homojenlige dayanmaktadır. "Λ" değeri değişkenin en iyi ayırım yapacağını gösterse de adımsal seçme tekniği değişkenler arasında çoklu doğrusal bağlantı var ise bu değişkenlerden birini modele almaz. Çünkü iki değişkenin de modelde olması gereksizdir (Sharma, 1996).

3.2.2. Yeni Gözlemlerin Sınıflandırılması

Fisher *X* veri matrisinde, *A* toplumunda genel örneklerin gözlemlerinden hesaplanan *Y*'lerin ortalamasını μ_{1Y} ve *B* toplumundan gelen örneklerin gözlemlerinden hesaplanan *Y*'lerin ortalamasını da μ_{2Y} ile gösterilsin ve μ_{1Y} ile μ_{2Y} arasındaki karesel uzaklığı maksimum yapacak bir karesel bileşen bulmayı amaçlamıştır.

Her iki toplum aynı kovaryans matrisine sahip toplumlar olmak üzere;

$\mu_1 = E(X/A)$: 1.toplumun çok değişkenli gözlemlerin beklenen değeri

$\mu_2 = E(X/B)$: 2.toplumun çok değişkenli gözlemlerin beklenen değeri

olarak bulunur.

Karesel dönüştürme;

$Y = \left(\overline{x_1} - \overline{x_2}\right)' Ax$ biçiminde hesaplanır.

Y değerlerinin ortalamaları;

$$\mu_{1Y} = E(Y/A) = E(b'X/A) = b'\mu_1$$

$$\mu_{2Y} = E(Y/B) = E(b'X/B) = b'\mu_2$$

biçiminde hesaplanır.

Karesel diskriminant fonksiyonu, *A ve B* toplumlarından alınan çok değişkenli gözlemleri ortak toplum varyansına göre birbirlerinden olabildiğince farklı olan tek değişkenli

ortalamalar biçiminde gösterecek bir forma dönüştürür. İki toplumda Y dönüştürülmüş değerler ortalamaları arasındaki orta nokta;

$$k = \left(\frac{1}{2}\right)\ln\left(\frac{|S_1|}{|S_2|}\right) + \frac{1}{2}(\overline{x_1'}S_1^{-1}\overline{x_1} + \overline{x_2'}S_2^{-1}\overline{x_2})$$ biçiminde hesaplanır.

k değeri yeni gözlemin hangi sınıfta yer alabileceğini belirlemeye yarayan bir kriterdir. x_0 gözlem vektörüne sahip bir yeni gözlemin 1. ya da 2. toplumdan hangisine ait bir birim olduğuna karar vermek için önce y_0 gibi tek değişkenli bir değerle ifade edilmesi gerekir.

$$y_0 = -\frac{1}{2}x_0' A x_0 + (x_1'S_1^{-1} - x_2'S_2^{-1})x_0 - k$$

y_0 değerinin hangi topluma ait olduğu aşağıdaki gibi belirlenir.

Eğer $y_0 - k \geq 0$ ise x_0 1.topluma ait bir gözlemdir.

Eğer $y_0 - k \leq 0$ ise x_0 2.topluma ait bir gözlemdir.

Genelde toplumlara ilişkin kovaryans matrisleri $\sum$ ve ortalama vektörleri μ_i bilinmez. Bu durumda karesel diskriminant analizi veri matrisinden elde edilen ortalama vektörlerine ve kovaryans matrisine göre yürütülür.

3.3. Lojistik Regresyon

Lojistik regresyon; cevap değişkenin kategorik, ikili, üçlü ve çoklu kategorilerde gözlendiği durumlarda açıklayıcı değişkenlerle-sonuç ilişkisini belirlemede yararlanılan bir tekniktir. Lojistik regresyon ikili (binary) değişkenlerin sürekli değişkenlerle bağımlılığını tanımlamak için ilk defa Cox tarafından geliştirilmiştir(Anderson,1994).

Bağımsız değişken kategorik olduğunda değişken ortalaması durumun hangi kategoriye daha yakın olduğunu gösteren olasılık fonksiyonudur(Menard, 1995). Değişkenleri 0 ve 1 olarak kodlama; değişken ortalamasının değişkenin kategorilerinden hangisine ait olacağı olasılığını göstermektedir. y'nin tahmin değeri iki kategoriden hangisine ait olduğunun tahmin olasılığıdır (predicted probability)(Menard, 1995). Basit ve çoklu regresyon analizleri bağımlı değişken ile açıklayıcı değişken ya da değişkenler arasındaki matematiksel bağıntıyı analiz

etmekte kullanılmaktadır. Bu tekniklerin uygulanabileceği veri setlerinde bağımlı değişkenin normal dağılım göstermesi, bağımsız değişkenlerin normal dağılım gösteren toplum ya da toplumlardan çekilmiş olması ve hata varyansının $\varepsilon \sim N(0,\sigma^2)$ parametreli normal dağılım göstermesi gerekmektedir. Bu ve benzeri koşulların yerine getirilemediği veri setlerine basit ya da çoklu regresyon analizi uygulanamaz (Özdamar, 2002a).

Bağımlı değişken ikili gözlemler içeriyorsa normal olmayan hata ve sabit olmayan varyans gibi problemlere sahiptir(Kunter vd.,1996). Sınıflama ve regresyon modelleri arasındaki temel fark tahmin edilen bağımlı değişkenin kategorik ya da süreklilik gösteren bir değere sahip olmasıdır. Ancak çok terimli lojistik regresyon (Multinomial Logistic Regression) gibi kategorik değerlerin de tahmin edilmesine olanak sağlayan tekniklerle, her iki model giderek birbirlerine yaklaşmakta ve bunun sonucu olarak aynı tekniklerden yararlanılması mümkün olmaktadır(Akpınar, 2000).

Lojistik regresyon analizi, temelde bir regresyon çözümlemesi olmakla birlikte bir ayırıcı çözümleme tekniği olma özelliği göstermektedir(Akkuş vd., 2005). Normal dağılım varsayımı, süreklilik varsayımı ön koşulu yoktur. Varsayım kısıtlaması olmaması, bu tekniğe olan ilgiyi arttırmaktadır. Bununla birlikte değişkenler arasında çok değişkenli normal dağılım mevcutsa çözüm daha uygun olacaktır. Ayrıca diğer regresyon tekniklerinde olduğu gibi değişkenler arasında çoklu doğrusal bağıntı olması yanlı tahmincilere ve standart hatada artışa neden olabilmektedir. Bundan dolayı tahmin edici değişkenler kategorik değişkenler olduğu takdirde süreç daha etkili olacaktır. Bağımlı değişken üzerinde açıklayıcı değişkenlerin etkileri olasılık olarak elde edilerek risk faktörlerinin olasılık olarak belirlenmesi sağlanır. Diskriminant analizi de verilerin belirli olasılıklara göre belirli sınıflara atanmasını sağlar ve çok değişkenli normal dağılım varsayımını kabul etmektedir. Lojistik regresyonun böyle bir varsayım gerektirmemesi ayırıcı bir özelliğidir. Lojistik regresyon, lojistik modellere göre parametre tahminleri yapar. Bağımlı değişkenin tahmini değerlerini olasılık olarak hesaplayarak, olasılık kurallarına uygun sınıflama yapma imkanı veren bir istatistiksel tekniktir(Özdamar , 2002a).

$$P_i = E(Y = 1 / X_i) = \beta_1 + \beta_2 X_i$$

Burada $Z_i = \beta_1 + \beta_2 X_i$ 'dir. Bu fonksiyon lojistik dağılım fonksiyonu olarak adlandırılır.P_i *0* ile 1 arasında değerler alır ve Z_i ile dolayısıyla X_i ile ilişkisi doğrusal

değildir.P_i yalnız X ile değil, β 'larla da doğrusal ilişki içerisinde değildir. Ama bu görüntüsel bir durumdur.Model doğrusallaştırılabilir(Gujarati, 1995).

Şöyle ki;

$$1-P_i=\frac{1}{1+e^{Z_i}}$$

$$\frac{P_I}{1-P_I}=\frac{1+e^{Z_i}}{1+e^{-Z_i}}=e^{Z_i}$$

Bu durumda $\frac{P_i}{1-P_i}$ bir olayın gerçekleşmesinin bahis **(odds)** oranıdır ve doğal logaritması alındığında;

$$L_i=ln\left[\frac{P_i}{1-P_i}\right]=Z_i$$

$$=\beta_1+\beta_2 X_i$$

Yani bahis oranı logaritması L,yalnız X'e göre değil, ana kütle katsayılarına göre de doğrusaldır. L'ye **logit**, bu modellere de **logit model** denir.

Logit modeli şu özelliklere sahiptir.

- P, 0 'dan 1'e giderken (Yani Z_i,$-\infty$ 'dan $+\infty$ 'a doğru değişirken), logit L'de - $-\infty$ 'dan $+\infty$ 'a doğru değişir. Olasılıklar 0 ile 1 arasında yer alırken, logitler için sınırlama söz konusu değildir.
- L, X 'e göre doğrusal olmakla birlikte olasılıkların kendileri böyle değildir.
- Yorumu şöyle yapılır. β_2, eğim, X 'teki bir birim değişmeye karşılık L 'deki değişmeyi ölçer. Sabit terim β_1 ise X 0 olursa bağımlı değişkenin log-bahis oranıdır. Regresyonda sabit terimlerde her zaman görüldüğü gibi bu da fiziksel bir anlam taşımayabilir.

- Logit modeli log-bahis oranının X ile doğrusal ilişki içinde olduğunu varsayar.
- Belli bir X^* veriyken, bağımlı değişkenin bahis oranı değil de kendi olasılığı tahmin edilmek istenirse, β_1 ile β_2 tahminleri bir kez elde edildikten sonra doğrudan bulunabilir.

β_1 ile β_2 nin nasıl tahmin edileceği ise aşağıda anlatılmaktadır.

3.3.1. Logit Modelinin Tahmin Edilmesi

Tahmin modeli;

$$L_i = ln\left[\frac{P_i}{1-P_i}\right] = \beta_1 + \beta_2 X_i + u_i$$

ana kütle katsayılarını tahmin edebilmek için en yüksek olabilirlik (maximum likelihood) tekniğine başvurulabilir. X_i düzeyindeki N_i gözlemden n_i tanesinde olayın gerçekleştiği varsayılırsa ($n_i \leq N_i$);

$$\hat{P}_i = \frac{n_i}{N_i}$$

değeri, her X_i düzeyine karşılık gelen P_i tahmininde bu kullanılabilir. Eğer N_i yeteri kadar büyükse, $\hat{P}_i$ P_i'nin iyi bir tahmini olacaktır. Tahmin edilmiş P_i ile tahmin edilmiş logit modeli;

Bu da her X_i'deki gözlem sayısı N_i yeteri kadar büyükse, gerçek logit L_i'nin iyi bir tahmincisi olacaktır.

Bilinmeyen P_i'yi $\hat{P}_i$ ile σ^2'nin bir tahmin edicisi olan şu ifade kullanılabilir.

$$\sigma^2 = \frac{1}{N_i \hat{P}_i(1-\hat{P}_i)}$$

3.4. En Yakın Komşu Yaklaşımı (Nearest Neighbour Approach)

En yakın komsu tekniği, ilk defa Fix ve Hudges tarafından geliştirilen sınıflandırma probleminde standart, parametrik olmayan bir yaklaşımdır. Bu teknik KS'de ilk defa Chatterjee ve Barcun ve daha sonra Henley ve Hand tarafından uygulanmıştır. Bu tekniğin dayandığı mantık, herhangi iki başvurunun birbirinden ne kadar uzak olduğunu ölçmek için başvuru veri uzayında bir mesafeyi seçmeye dayanmaktadır. Geçmiş başvuruların bir örnekleme temsili standart alınır. Yeni bir başvuru temsili örneklemede *k* kadar yakın başvurular arasındaki (yeni başvurunun en yakın komsuları) iyi-kötü oranına dayanarak iyi veya kötü olarak sınıflanır(Thomas, vd., 2002).

En yakın komsu yaklaşımı eşit iyi-kötü oranı olan örneklemelerde optimum performans gösterir (http://www.fractalanalytic.com). Bu yaklaşımı uygulamak için üç tane parametreye ihtiyaç vardır: mesafe, en yakın komsular setini oluşturan kaç tane başvuru sayısı olduğu (k), ve bir başvurunun iyi olarak sınıflanabilmesi için başvuru iyi oranının ne olması gerektiği. Normal olarak, eğer komşuların çoğunluğu iyi ise, başvuru iyi olarak sınıflandırılır. Aksi taktirde başvuru kötü olarak sınıflanır. Ortalama var olan maliyet *M*, ve iyiyi reddetmenin ortalama kayıp karı *K* olarak tanımlansın. Eğer en yakın komşuların en azından $M/M+K$ tanesi iyi ise, yeni bir başvuru iyi olarak sınıflandırılır. Eğer yeni bir başvurunun iyi olma olasılığı iyi olan komşuların oranı ise, bu kriter beklenen kaybı minimize edecektir.

Mesafenin seçimi oldukça önemlidir. Fukanaga ve Flick genel bir mesafe tanımı yapmıştır:

$$d(x_1, x_2) = (x_1 - x_2)\,A(x_1)\left((x_1 - x_2)^T\right)^{1/2}$$

A(x), pxp simetrik sonlu pozitif bir matristir. Eğer *x*'e bağlı ise, *A(x)* yerel mesafe olarak adlandırılır, eğer *x*'ten bağımsız ise global mesafe olarak adlandırılır. Yerel mesafenin eksikliği,genelde uygun olmayan deneme setinin özelliklerini dikkate alır. Bu nedenle bir çok araştırmacı global mesafeye odaklanır. KS'de en yakın komşu yaklaşımının en detaylı uygulaması Henley ve Hand tarafından yapılmıştır. Bu teknik, öklit uzunluğu ve iyi ile kötüyü en iyi ayıran yönün uzunluğu karışımına odaklanır. Eğer *w;p* boyutlu yön vektörü ise, Henley ve Hand'in mesafe ifadesi şöyledir(Thomas, vd., 2002).

$$d(x_1, x_2) = \left\{(x_1 - x_2)^T (I + D_w.w^T)(x_1 - x_2)\right\}^{1/2}$$

KS'de lineer ve lojistik regresyon yaklaşımları kadar çok sık kullanılmamasına karşın, en yakın komşu tekniği gerçek uygulamalar için bazı önemli özelliklere sahiptir. Dinamik olarak yeni olaylar eklemeyerek deneme setini güncellemek çok kolaydır ve eklenenin iyi veya kötü olduğu bilindiğinde kolayca o olay örneklemeden çıkarılabilir. İlk seferde iyi bir mesafe bulmak, skor kart oluşturmada regresyon tekniğiyle hemen hemen eşdeğer bir netice verir. Böylece çoğu uygulayıcı bu noktada ilerlemeyi durdurup geleneksel bir skor kart kullanmayı tercih ederler. Sınıflandırma ağacı yaklaşımı ile kıyasladığımızda, en yakın komşu yaklaşımı her bir başvuranın özelliği için bir skor üretmez, uygulayıcılar için bir denge noktası belirler ve onların gerçekte sistemin ne yaptığını anlamasını sağlar.

3.5. Yapay Sinir Ağları (Neural Networks) (YSA)

Sinir ağları aslında, insan beynindeki bilginin işlenmesini ve iletilmesini modelleme denemelerinden geliştirilmiştir. Beyinde, çok sayıda dendrit elektrik sinyallerini nörona (neuron) taşır, nöronlar sinyalleri nabız atışına çevirirler ve aksona (axon) oradan da diğer nöronların dendritlerine (dendrites) bilgiyi ileten çok sayıda uyarım (synapse) gönderirler. insan beyni tahmini 10 milyar nörona sahiptir. Beyine benzer bir şekilde, bir sinir ağı çok sayıda girdiden (değişkenlerden) oluşmaktadır. Her bir girdi bir ağırlıklı katsayı ile çarpılmaktadır. Bu beyindeki dendrit yapısına benzemektedir. Ürünler toplanıp bir nöronda dönüştürülür. Elde edilen sonuç diger nöron için bir girdi değeri olur. Bağlantısal Mimariler (Connectionist Architectutures), Adaptif Sistemler (Adaptive Systems) veya Paralel Dağıtılmış işlemciler (Parallel Distributed Processing) olarak da adlandırılan YSA'lar, oldukça fazla bağlantı içeren ve paralel yapılandırılmış beyin işlevinden esinlenen bir bilgi işlem paradigmasıdır. Farklı isimlerle anılmaları, farklılık sağlayan bazı temel özelliklerinden kaynaklanmaktadır. Bağlantısal Mimari (veya Bağlantısal Sistem) olarak anılmalarının temel sebebi, bireysel işlem elemanları (processing nodes) arasındaki bağlantılardır. Ayrıca, bu bağlantıların ağırlıkları değişebildiğinden YSA'lar çalışma sistemlerini daha da etkinleştirebilmektedirler ve bu yüzden Adaptif Sistem olarak da adlandırılmaktadırlar. Paralel Dağıtılmış işlemciler olarak adlandırılmalarının sebebi ise ağ içinde çok sayıdaki nod veya nöronların hepsinin birbirlerine paralel olarak çalışmalarıdır. Bu yapı, eşanlı bir çözüm üretebilme yeteneği sağlamaktadır(Yurtoğlu, 2005). YSA'lar, tanımlanmamış girdi veriler hakkında karar verirken genelleme yapabildikleri için iyi birer yapı tanımlayıcısı (pattern recognition engine) ve sağlam sınıflayıcılardırlar (robust classifier).

YSA'ların çok sayıda farklı çeşitleri vardır. Bu farklılıkların kaynağı, mimarisi, öğrenme tekniği, bağlantı yapısı vb. olabilmektedir. Genel olarak, YSA'lar üç ana kritere göre sınıflandırılmaktadırlar. Bu kriterlerden biri öğrenme tekniğidir. Temel olarak iki çeşit öğrenme algoritması vardır: yönlendirmeli öğrenme ve yönlendirmesiz öğrenme. Her tekniğin kullandığı öğrenme kuralı değişebilmekteyse de, YSA'lar bu iki algoritmaya göre sınıflandırılırlar.

İkinci bir sınıflandırma, ağın kullandığı veriye göre yapılmaktadır. Temel olarak, kalitatif ve kantitatif olmak üzere iki tür veri vardır. Kalitatif verilerle çalışan ağlar, ister yönlendirmeli ister yönlendirmesiz öğrenme kullansın, sınıflandırma ağları olarak bilinirler. Kantitatif veriler kullanan yönlendirmeli eğitme ise regresyon olarak adlandırılmaktadır. Son sınıflandırma kriteri ise ağın yapısıdır. Bazı ağlar ileri besleme seklinde yapılandırılırken, bazı ağlar ise geri besleme yapısı içermektedir. İleri besleme sinir ağlarında, işlem elemanları arasındaki bağlantılar bir döngü oluşturmazlar ve bu ağlar girdi veriye genellikle hızlı bir şekilde karşılık üretirler. Geri beslemeli ağlarda (Recurrent Networks) ise bağlantılar döngü içerirler ve hatta her seferinde yeni veri kullanabilmektedirler. Bu ağlar, döngü sebebiyle girdinin karşılığını yavaş bir şekilde oluştururlar. Bu yüzden, bu tür ağların eğitme süreci daha uzun olmaktadır. Ayrıca, hem ileri besleme hem de geri yayılma olarak tanımlanabilecek ağ yapıları da mevcuttur(Yurtoğlu, 2005).

3.5.1. YSA Avantajları

Teknolojik gelişme olarak da görülmesi gereken yapay sinir ağları tekniği, özellikleri ve yapabildikleri sayesinde önemli avantajlar sunmaktadır. YSA'ların farklılık ve avantajları; Doğrusal Olmayan Yapı; YSA'ların en önemli özelliklerinden birisi gerçek hayattaki olası doğrusal olmayan yapıları da dikkate alabilmesidir. Analiz konusunun içerdiği veri setinin doğrusal veya doğrusal olmayan yapı içeriyor olması, analiz sonuçlarını etkileyecek önemli bir faktördür. Bu yüzden, doğrusal olmayan yapıları dikkate alabilmesi YSA'ların önemli bir özelliğidir.

Öğrenme; YSA'ların diğer bir önemli avantajı en önemli özelliğinden kaynaklanmaktadır. Esin kaynağı insan beyninin çalışma sistemi olan bu teknik, eğitme veya başlangıç tecrübesi sayesinde veriyi kullanarak öğrenme yeteneğine sahiptir. Bu özelliği sayesinde ise geleneksel teknikler için çok karmaşık kalan problemlere çözüm

sağlayabilmektedirler. Ayrıca, insanların kolayca yapabildiği ama geleneksel metotların uygulanamadığı basit işlemler için de oldukça uygundurlar.

Yerel İşlem ve Esneklik; YSA'lar, geleneksel işlemcilerden farklı şekilde işlem yapmaktadırlar. Geleneksel işlemcilerde, tek bir merkezi işlem elemanı her hareketi sırasıyla gerçekleştirir. YSA modelleri, her biri büyük bir problemin bir parçası ile ilgilenen çok sayıda basit işlem elemanlarından oluşma ve bağlantı ağırlıklarının ayarlanabilmesi gibi özelliklerinden dolayı önemli derecede esnek bir yapıya sahiptirler. Bu esnek yapı sayesinde ağın bir kısmının zarar görmesi modelde sadece performans düşüklüğü yaratır. Modelin işlevini tamamen yitirmesi söz konusu olmaz. Ayrıca, toplam işlem yükünü paylasan işlem elemanlarının birbirleri arasındaki yoğun bağlantı yapısı sinirsel hesaplamanın temel güç kaynağıdır. Bu yerel işlem yapısı sayesinde, YSA tekniği en karmaşık problemlere bile uygulanabilmekte ve tatminkar çözümler sağlayabilmektedir.

Gerçek Zamanlı İşlem; YSA hesaplamaları paralel olarak yürütülebildiğinden gerçek zamanlı işlem yapabilir.

Genelleme; öğrenme yeteneği sayesinde bilinen örnekleri kullanarak daha önce karşılaşılmamış durumlarda genelleme yapabilmektedir. Yani, hatalı (noisy) veya kayıp veriler için çözüm üretebilmektedir

Hafıza; Bunlara ek olarak, işlem elemanları arasındaki ağırlıklı bağlantılar sayesinde dağıtılmış hafızada bilgi saklayabildikleri söylenebilir.

Kendi İlişkisini Oluşturma; YSA, bilgilere (verilere) göre kendi ilişkilerini oluştururlar, denklem içermezler.

Sınırsız Sayıda Değişken ve Parametre; YSA modelleri sınırsız sayıda değişken ve parametre ile çalışabilmektedir. Bu sayede mükemmel bir öngörü doğruluğu ile genel çözümler sağlanabilmektedir(Yurtoğlu, 2005).

YSA kullanımının sağladığı avantajlar doğrultusunda sonuçlarının başarısı, bu uygulamayı kullanan şirketlerin rakamsal verileriyle de özetlenmek istenirse; kredi skorlamada YSA kullanan HNC şirketi karlılığını %27 artırmıştır. Bu sistem daha sonra mortgage kredileri için de kullanılmıştır (Stergious ve Siganos, 2005). Finans şirketlerinde YSA kullanan uygulayıcılar birçok faydalarını görmektedirler. American Express ve Security

Pasific Bankası hilekarlık tespiti (fraud detection) ve küçük isletmelere verilen krediler için YSA modeli kullanmaktadırlar. Lloyds Bowmaker Motor Finance, otomobil finansmanı için kredi skorlamada YSA kullanmaktadır. Ve YSA uyguladıktan sonra model %10 daha verimli hale gelmiştir(Yegorova, 2001).

3.5.2. Tek Katmanlı Sinir Ağları (Single Layer Neural Networks)

Bir tek katmanlı sinir ağı, yukarıda belirtilen bileşenlerden, yani birden fazla girdiden oluşur. Ancak burada, bir nöronda elde edilen sonucun diğer nöron için girdi değeri olması özelliği yoktur, zaten araştırılan değer dönüştürülen değerdir.

Matematiksel ifadesi ise şöyledir:

$$u_k = w_{k0}x_0 + w_{k1}x_1 + \ldots + w_{kp}x_p = \sum_{q=0}^{p} w_{kq}x_q$$

$$y_k = F(u_k)$$

Her bir $x_1,\ldots,x_p$ kredi kartı başvurusunun bir karakteristiğini ifade eden bir değişkendir. Her biri sinyal olarak adlandırılan bir değer alır. Sinaptik ağırlıklar, eğer pozitif iseler ilgili değişkeni artırdıkları için uyarıcı olarak bilinirler; eğer negatif iseler pozitif değişkenler için u_k değerini düşürdükleri için yasaklayıcı olarak adlandırılırlar. Her bir ağırlığın gösterimindeki (k, p) harflerinin sırasına dikkat edilirse; *k* ağırlığın uygulandığı sinir indisini, p ise değişken indisini gösterir. Tek katmanlı sinir ağında *k*=1'dir.Çünkü yalnızca bir tane sinir vardır. Dikkat edilirse x_0 değerine 1 atanır. Yukarıda vermiş olduğumuz $w_{k0}x_0$ terimi yalnızca w_{k0} ile gösterilir. Bu ifade sabit değerle artan veya azalan bir u_k fonksiyonuna sahiptir(Thomas,vd., 2002).

u değeri bir aktivasyon fonksiyonu kullanılarak dönüşüme tabi tutulur. Eski ağlarda, bu fonksiyon lineerdi ve sadece ilgili ağların sınıflandırma problemleri ile sınırlıydı. Çeşitli alternatif transfer-dönüşüm fonksiyonları kullanılmaktadır.

Nöron davranışını belirleyen önemli etmenlerden biri nöronun aktivasyon fonksiyonudur. Biyolojik nöronlarda, toplam belli bir değeri astığında nöronun kısa süreli bir

darbe gönderdiği bilinmektedir. Bu davranışa benzer bir davranışı yapay nöronlarla elde etmek için kullanılan aktivasyon fonksiyonlarından ikisi;

- Eşik değeri fonksiyonu

$F(u) = 1$ eğer $u \geq 0$

$= 0$ eğer $u < 0$

$u = 0$ için en büyük sinir çıktı değeri 1 olur. Aksi taktirde, çıktı değeri 0 olur.

- lojistik fonksiyon

$$F(u) = \frac{1}{1 + e^{-au}}$$

İkili karar mekanizması gerektiren durumlarda eşik değeri fonksiyonu ile tanımlanan sert geçişli aktivasyon fonksiyonları tercih edilmektedir(Efe ve Kaynak ,2004).

3.5.3. Çok Katmanlı Sinir Ağları (Multilayer Perceptrons)

Çok katmanlı algılayıcı; sinyallerin bir girdi katmanından, çıktı sinyallerine ait (farklı y_v değerleri) bir çıktı katmanından, ve bu katmanlar arasında olan gizli katman olarak adlandırılan çok sayıda sinir tabakasından meydana gelir. Gizli katmandaki her bir nöron, bu katmandaki farklı bir nöron için aynı girdilerden farklı girdilerin meydana geldiği bir ağlık setine sahiptir. Gizli katmandaki her bir nörondaki çıktılar ağırlıklara sahiptir ve bunlar eğer varsa sonraki gizli katmandaki nöronlar için girdi olur. Aksi durumda, çıktı katmanına girdi olurlar. Çıktı katmanı her bir nöron için değerleri verir. Bu değerler, her bir olayı kesme değeri ile karsılaştırarak sınıflamada kullanılır. Matematiksel ifadesi ise;

$$y_k = F_1(\sum_{q=0}^{p} w_{kq} x_q$$

F'in alt indisindeki 1, girdi katmanından sonraki ilk katman olduğunu gösterir. y_k (k=1,...,r)

birinci gizli katmandan elde edilen çıktı değerleridir. Bir katmanın çıktısı ondan sonra gelen katmanın girdisi olduğundan, aşağıdaki ifade yazılabilir;

$$z_v = F_2\left(\sum_{k=1}^{r} K_{vk} y_k\right) = F_2\left(\sum_{k=1}^{r} K_{vk}(F_1(\sum_{q=0}^{p} w_{kq} x_q))\right)$$

z_v (v=1,…,s) çıktı katmanındaki nöron v'nin çıktı değerini göstermektedir. F_2 çıktı katmanındaki aktivasyon fonksiyonunu ve K_{vk} da gizli katmandaki nöron k ve çıktı katmanındaki nöron v'yi birleştiren y_k katmanına uygulanan ağırlık değerlerini göstermektedir.

Ağırlıkların vektörel hesaplanması deneme olarak adlandırılır. Birçok teknik vardır, fakat en çok bilineni geriye doğru yayılma algoritmasıdır. Deneme çiftleri; durumun her bir girdi değişkeni için değerlerden ve olayın bilinen sınıflamasından oluşur ve hata fonksiyonunu minimize etmek için ağırlıklar tekrar tekrar hesaplanıp ağ üzerinde gösterilir (Thomas, vd., 2002).

3.6. Doğrusal Programlama (Lineer Programming)

Mangasarian iki ayrı grubun yer aldığı sınıflama problemlerinde doğrusal programlamanın kullanabileceğini belirtmiştir. Freed ve Glover ve Hand, Mutlak Hatalar Toplamını (MSAE) ya da Maksimum Hataların Minimizasyonunu (MME) amaç fonksiyonu alan lineer programlama modellerini lineer olması gerekmeyen iki grubu ayırmak için tanımışlardır.

Uygulama değişkeni, $\underline{X} = \left(X_1, X_2, ..., X_P \right)$ değerlerinin bütün kombinasyonları A_G ve A_B olarak adlandırılan iki gruba ayrılabilir: A_G , iyi olarak cevaplanan soruları; A_B , kötü olarak cevaplanan soruları göstermektedir. Bir örneklemenin n adet başvuru aldığı kabul edilsin. Örneklemenin n_G adeti iyiler ise, notasyon kolaylığı için, bunların örneklemedeki birinci n_G 'ler olduğu kabul edilsin. Örneklemenin (Örnekleme $i = n_G + 1, ... n_G + n_B$ geri kalan n_B adedi "kötü" olur. Başvuru i'nin başvuru değişkeni X'e $\left(X_1, X_2, ..., X_P \right)$olan cevabında $(x_{i1}, x_{i2}, ..., x_{ip})$ karakteristiklerine (özelliklerine) sahip olduğu kabul edilsin. Cevapların ağırlıklı ortalama toplamının $w_1 X_1 + w_2 X_2 + ... + w_p X_p$ iyi başvurular için belirli bir kesme

değerinin c (cutoff) üzerinde olduğu ve kötüler için kesme değerinin altında olduğu ağırlıkları veya skorları ($w_1, w_2, ..., w_p$) seçilmek istenir. Eğer başvuru değişkenleri ikili (binary) değişkene dönüştürülürse, verilen farklı cevaplar için ağırlıklar (w_i'leri) onların skoru olarak düşünülebilir. Genellikle iyiler ve kötüler biri birinden kusursuz bir şekilde ayrılamaz. Dolayısıyla, muhtemel hatalar için pozitif veya sıfır değeri alan a_i değişkenleri tanımlanır. Böylece, örneklemede başvuru i; iyi değeri alırsa, $w_1 x_{i1} + w_2 x_{i2} + ... + w_p x_{ip} \geq c - a_i$ kısıdı tanımlanır. Buna karsın, eğer başvuru kötü olursa, $w_1 x_{j1} + w_2 x_{j2} + ... + w_p x_{jp} \leq c + a_j$ kısıdı tanımlanır. Buradaki sapmaların mutlak değerler toplamını (MSD) minimize eden ağırlıklarını

($w_1, w_2, ..., w_p$) bulmak için, aşağıdaki lineer programlama modeli çözülmelidir:

Min ($a_1 + a_2 + ... + a_{n_g + n_B}$)

Su kısıtlara göre: $w_1 x_{i1} + w_2 x_{i2} + ... + w_p x_{ip} \geq c - a_i$, $\quad 1 \leq i \leq n_G$

$w_1 x_{i1} + w_2 x_{i2} + ... + w_p x_{ip} \leq c + a_i$, $\quad n_G + 1 \leq i \leq n_G + n_B$

$a_i \geq 0 \quad 1 \leq i \leq n_G + n_B$

Maksimum Sapmanın Minimizasyonu (MMD) yerine, her bir kısıttaki aynı hatalar teriminin minimizasyonu yapılırsa, aşağıdaki model çözülür:

Minimize a

Şu kısıtlara göre: $w_1 x_{i1} + w_2 x_{i2} + ... + w_p x_{ip} \geq c - a \quad 1 \leq i \leq n_G$

$w_1 x_{i1} + w_2 x_{i2} + ... + w_p x_{ip} \leq c + a \quad n_G + 1 \leq i \leq n_G + n_B$

$a \geq 0, \quad 1 \leq i \leq n_G + n_B$ (Thomas,vd., 2002)

3.7. Tamsayılı Programlama (Integer Programming)

Lineer programlama modelleri yanlış sınıflanan başvuru kredi skorlarındaki standart sapmanın toplamını minimize eder. Buna karşın, daha pratik bir yol yanlış sınıflama sayılarını, ya da eğer bir kötüyü iyi olarak yanlış sınıflamanın maliyeti M, bir iyiyi kötü olarak yanlış sınıflama maliyetinden K, çok farklıysa yanlış sınıflamanın toplam maliyetini, minimize etmektir. Bu kriterleri dikkate alarak skor kartlar oluştururken lineer programlama kullanılabilir, fakat bazı değişkenler tamsayı olmalıdır {0,1} gibi. Bu teknik tamsayılı

programlama olarak adlandırılır. Koehler ve Erengue aşağıdaki tamsayılı programlama modelini geliştirmişlerdir.

Minimize $K(d_1 + ... + d_{n_G}) + M(d_{n_{G+1}} + d_{n_{G+B}})$

Şu kısıtlara göre: $w_1 x_{i1} + ... + w_p x_{ip} \geq c - Md_i$, $1 \leq i \leq n_G$

$w_1 x_{i1} + ... + w_p x_{ip} \leq c - Md_i$, $n_G + 1 \leq i \leq n_G + n_B$,

$0 \leq d_i \leq 1$, d_i tamsayı.

Böylece d_i, eğer örneklemede müşteri i yanlış sınıflanırsa 1'e eşit sınıflanmaz ise 0'a eşit olan bir değişken olacak.c=0, w_i=0, i=0,…,p olduğunda yukarıda belirttiğimiz programlama denklemi tekrardan minimize edilirse, aşağıdaki gibi bir normalizasyon kısıdı eklenmelidir.

$$\sum_{j=1}^{p} (s_j^+ + s_j^-) = 1,$$

$0 \leq s_j^+$, $s_j^- \leq 1$ ve s_j^+, s_j^- tamsayı, j=1,…,p

$$-1 + 2s_j \leq w_j \leq 1 - 2s_j, \quad j=1,\ldots,p$$

Bu kısıtlar seti $s_i^+ ves_i^-$'den birinin 1'e eşit olmasını ve ilgili w_i değerinin 1'den büyük veya 1'den küçük olmasını gerektirir.(c'yi negatif veya pozitif değere eşitlemeye zorlar). Bu durumda, aşağıdaki eşitlik yazılır.

$\sum_{j=1}^{p} w_i = 1$ ve c=+1 ya da -1

Koehler ve Erenguc tamsayılı programlama modelinin lineer programlama modellerine göre daha iyi bir sınıflama yaptığını göstermişlerdir. Fakat, iki büyük dezavantaj vardır. Birincisi, tamsayılı programlama modellerinin çözümü lineer modellere göre çok daha fazla zaman almakta ve böylece yüzlerce olay setinden yalnızca küçük bir örnekleme ile çözüm yapılabilmektedir. İkincisi, deneme seti üzerinde aynı yanlış sınıflandırma sayısı ile çok sayıda optimal çözüm vardır.

Çeşitli araştırmacılar hem incelenen örneklemde güçlü olan bir sınıflamayı garanti etmek hem de deneme örneklemesindeki yanlış sınıflamaların sayısını azaltmak için tamsayılı programlama modelinin denklemine ekstra kısıtlar ve ikincil amaç fonksiyonları eklemişlerdir. Şu ana kadar tam sayılı programlama minimum yanlış sınıflandırma kriterine sahip sınıflandırma problemlerinin çözümü için kullanılmıştır, fakat hataların mutlak toplamının minimizasyonunda lineer programlama yaklaşımında ortaya çıkan iki güçlüğün üstesinden tamsayılı programlama gelir(Thomas,vd., 2002).

4. UYGULUMALAR

Kredi kartı skorlaması için bir bankanın kredi kartı müşterilerine ait 1806 ve on üç değişken ile çalışılmıştır. Değişkenlerin kategorileri kredi kartı başvuru formundaki kategorileri baz alınarak kodlanmıştır. AYLIK NET GELİR, MAAŞ gibi parasal değerlerin alt sınırlarının küçük olmasının nedeni kredi kartı müşterilerinin bilgilerinin güncellenmemiş olmasıdır. İyi müşteri-kötü müşteri ayrımı banka tarafından yapılmış olup veriler üzerinde analizler uygulanmıştır. Bu uygulamada SPSS paket programı kullanılmıştır.

Uygulamada kullanılan değişkenler aşağıdaki çizelgede gösterilmektedir.

DEĞİŞKENLER

İYİ-KÖTÜ	Müşterilerin iyi ya da kötü olusu
1	İyi
2	Kötü
MÜSTERİ TİPİ:	Müşterilerin hangi kampanyalarla banka müşterisi oldukları
1	B (Markette kazanılan müşteriler)
2	C (Mektup1)
3	D (Mektup 2)
4	G (Grup Firmaların Çalışanları)
5	J (Özel Müşteriler)
6	M (Mail Order)
7	N (Normal)
8	O (Okul Kampanyası)
9	R (Alışveriş Merkezi)
10	T (Toplu Pazarlama)
YAS:	Müşteri yası
1	18-30 yas arası
2	30-39 yas arası
3	39+
CİNSİYET:	Müşteri cinsiyeti
1	Erkek
2	Bayan
MEDENİ DURUM:	Müşteri medeni durumu

1	Evli
2	Dul
3	Bekar
ÖGRENİM DURUMU:	Müşterinin öğrenim durumu
1	Okula Gitmemiş
2	İlkokul Mezunu
3	Ortaokul Mezunu
4	Lise Mezunu
5	Yüksekokul Mezunu
6	Üniversite Mezunu
7	Master/Doktora
ARABA SAHİBİ:	Müşterinin arabası olup-olmaması
1	Var
2	Yok
MAAS:	Müşterinin maaşı
1	100-750 YTL
2	750-1000 YTL
3	1000-1900 YTL
4	1900-3000 YTL
5	3000 YTL+
AYLIK NET GELİR:	Müşterinin aylık net geliri
1	700 YTL altı
2	700 – 1000 YTL
3	1000 – 1800 YTL
4	1800 – 3000YTL
5	3000 YTL ve üzeri
BASKA BAN.KRD.KARTI:	Müşterinin başka banka kredi kartı olup olmaması
1	1 tane var
2	2 tane var
3	Yok
BASKA BAN.KART.LİM.:	Müşterinin başka banka kart limiti
1	Yok
2	<300 YTL
3	300 YTL+
EK KART DURUMU:	Müşterinin ek kartı olup olmaması
1	Var
2	Yok
ÇALISMA SEKLİ:	Müşterinin isindeki unvanı
1	Diğer
2	Emekli
3	Ev hanımı
4	Kamu sektörü
5	Serbest meslek
6	Özel sektör
7	Öğrenci

4.1. BETİMSEL İSTATİSTİKLER

Çalışmada ileri analiz tekniklerine yer verilmesine rağmen ön bilgi sağlaması açısından bazı betimsel istatistikler incelenmiş ve bazı yorumlar yapılmıştır. Müşteri yaşı, maaş, aylık net gelir, başka banka kredi kart limiti gibi sürekli değişkenler için betimsel istatistikler aşağıda verilmiştir.

Betimsel istatistikler

	İYİ-KÖTÜ			**İSTATİSTİK**	**STANDART HATA**
MÜŞTERİ YAŞI	İYİ	**Ortalama**		**34,99**	0,79
		Ort.için %95 güv.ara.	Alt Sınır	33,43	
			Üst Sınır	36,56	
			Medyan	33	
			Varyans	100,1	
			Stan. Sap.	10	
			Aralık	44	
	KÖTÜ	**Ortalama**		**36,13**	0,74
		Ort.için %95 güv.ara.	Alt Sınır	34,65	
			Üst Sınır	37,6	
			Medyan	34	
			Varyans	96,97	
			Stan. Sap.	9,847	
			Aralık	53	
MAAŞ	İYİ	**Ortalama**		**1.832,00**	188,1
		Ort.için %95 güv.ara.	Alt Sınır	1461	
			Üst Sınır	2204	
			Medyan	1000	
			Varyans	6,00E+06	
			Stan. Sap.	2372	
			Aralık	19774	
	KÖTÜ	**Ortalama**		**3.331,00**	276,6
		Ort.için %95 güv.ara.	Alt Sınır	2785	
			Üst Sınır	3877	
			Medyan	2000	
			Varyans	1,00E+07	
			Stan. Sap.	3638	
			Aralık	YOK	
AYLIK NET GELİR	İYİ	**Ortalama**		**1.744,00**	189,1
		Ort.için %95 güv.ara.	Alt Sınır	1.370,00	
			Üst Sınır	2.117,00	
			Medyan	1000	
			Varyans	6,00E+06	
			Stan. Sap.	2384	
			Aralık	21.900,00	
	KÖTÜ	**Ortalama**		**3.118,00**	295,5

		Ort.için %95 güv.ara.	Alt Sınır	2535	
			Üst Sınır	3702	
			Medyan	2000	
			Varyans	2,00E+07	
			Stan. Sap.	3887	
			Aralık	26850	
BAŞ.BAN.KR.K.LİM.	İYİ	**Ortalama**		**2337**	305,6
		Ort.için %95 güv.ara.	Alt Sınır	1734	
			Üst Sınır	2941	
			Medyan	1500	
			Varyans	1,00E+07	
			Stan. Sap.	3854	
			Aralık	44750	
	KÖTÜ	**Ortalama**		**3608**	239,7
		Ort.için %95 güv.ara.	Alt Sınır	3135	
			Üst Sınır	4082	
			Medyan	2600	
			Varyans	1,00E+07	
			Stan. Sap.	3153	
			Aralık	19880	

Yukarıdaki Betimsel istatistikler tablosu sürekli, nümerik değişkenler için özet istatistikleri göstermektedir. Özet istatistik tablosu merkezi eğilim ölçümlerini (ortalama, medyan) de içerir. İyi-Kötü müşterilerin farklı değişkenler için ortalamalarının farkları incelendiğinde; iyi müşterilerle kötü müşterilerin ikamet süreleri arasında çok fazla bir farklılık olmadığı gözlenmektedir. Müşteri yaşları açısından incelediğimizde; kötü müşterilerin yaş ortalamasının daha yüksek olduğu görülmektedir. Maaş ve buna bağlı olarak aylık net gelir, başka banka kredi kartı limiti açısından da kötü müşterilerin ortalaması ciddi anlamda yüksektir.

Kategorik değişkenlerin frekans tabloları incelendiğinde;

MÜŞTERİ TİPİ	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
B	164	17,58	54,49	137	15,09	45,51
C	65	6,97	63,11	38	4,185	36,89
D	16	1,71	57,14	12	1,322	42,86
G	6	0,64	42,86	8	0,881	57,14
J	50	5,36	32,05	106	11,67	67,95
M	85	9,11	37,44	142	15,64	62,56
N	401	42,98	57,7	294	32,38	42,3
O	2	0,21	66,67	1	0,11	33,33
R	136	14,58	45,03	166	18,28	54,97
T	8	0,86	66,67	4	0,441	33,33
TOPLAM	933	100	50,68	908	100	49,32

Yukarıdaki frekans tablosu incelendiğinde; J kampanyasıyla kazanılan müşteriler %67,95 oranıyla kötü ve %32,05 oranıyla iyi müşteridirler. Bu da J (özel müşteriler) kampanyasının başarısı hakkında soru işareti oluşturmaktadır. Bu kampanyada kazanılan müşteriler büyük oranda ödeme problemi yaratan müşteriler olmuştur. Bundan dolayı C kampanyasını başarılı bir kampanya olduğu söylenebilir.

İyi-Kötü Müşteriler için cinsiyet dağılımı

CİNSİYET	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
ERKEK	856	91,75	50,8	829	91,3	49,2
KADIN	76	8,15	49.35	78	8,59	50,65

Yukarıdaki tablo incelendiğinde; kredi kartı müşterilerinin büyük oranda erkek müşterilerden oluştuğu görülmektedir. İyi ve kötü müşteriler içerisinde kadın-erkek oranları birbirine yakındır. Bu da cinsiyetin iyi müşteri-kötü müşteri ayırımında belirleyici bir etkisi olmadığı seklinde yorumlanabilir.

İyi-kötü müşteriler için medeni hal dağılımı

MEDENİ HAL	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
EVLİ	677	72,56	51,6	635	69,93	48,4
DUL	1	0,11	20	4	0,44	80
BEKAR	254	27,22	48,66	268	29,52	51,34
TOPLAM	932	99,89	50,68	907	99,89	49,32

Yukarıdaki tablo incelendiğinde; müşterilerin evli, dul, bekar olmalarının iyi-kötü müşteri ayrımı yapılmasında etkili bir bilgi olmadığı söylenebilir.

İyi-Kötü Müşteriler için öğrenim durumu dağılımı

ÖĞRENİM DURUMU	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
Eğitim almamış	464	49,73	54,59	386	42,51	45,41
İlkokul	1	0,11	25	3	0,33	75
Ortaokul	348	37,3	50	348	38,33	50
Ön Lisans	106	11,36	42,91	141	15,53	57,09
Üniversite	1	1,39	33,33	26	2,86	66,67
Y.Lisans/Doktora	1	0,11	50	1	0,11	50
Toplam	933	100	50,68	908	100	49,3

Yukarıdaki tablo incelendiğinde; farklı öğrenim durumu seviyeleri için iyi-kötü müşteri oranlarının yakın olduğu gözlemlenmektedir. Öğrenim durumunun ayırıcı bir özellik taşımadığı söylenebilir.

İyi-Kötü müşteriler için araba sahibi dağılımı

ARABA SAHİBİ	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
VAR	108	11,58	42,35	147	16,18	57,65
YOK	825	88,42	52,02	761	83,81	47,98
TOPLAM	933	100	50,68	908	100	49,32

Yukarıdaki tablo incelendiğinde; araba sahibi olma kriterinin de ayırıcı bir özellik taşımadığı söylenebilir.

İyi-Kötü müşteriler için başka banka kredi kartı dağılımı

BAŞKA BAN.KRD.KARTI	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
1 TANE VAR	105	11,25	50,48	103	11,34	49,52
2 TANE VAR	80	8,57	50	80	8,81	50
YOK	748	80,17	50,78	725	79,84	49,22
TOPLAM	933	100	50,68	908	100	49,32

Yukarıdaki tablo incelendiğinde; müşterilerin başka bankalara ait kredi kartlarının olup olmaması iyi müşteri-kötü müşteri ayırımı için anlamlı gözükmemektedir.

İyi-Kötü müşteriler için ek kart durumu dağılımı

EK KART	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
VAR	22	2,36	50	22	2,42	50
YOK	911	97,64	50,7	886	97,57	49,3
TOPLAM	933	100	50,68	908	100	49,32

Yukarıdaki tablo incelendiğinde; ek kart durumunun da müşterileri iyi müşteri-kötü müşteri olarak ayırmada etkili bir kriter olmadığı söylenebilir.

İyi-Kötü müşteriler için çalışma şekli

ÇALIŞMA ŞEKLİ	FREKANS					
	İYİ	İyiler içerisinde%	%	KÖTÜ	Kötüler içerisinde%	%
Diğer	264	28,3	53,66	228	25,11	46,34
Emekli	102	10,93	62,96	60	6,61	37,04
Ev Hanımı	446	47,8	46,07	522	57,49	53,93
Kamu Sektörü	46	4,93	85,19	8	0,88	14,81
Serbest Meslek	3	0,32	75	1	0,11	25
Öğrenci	70	7,5	98,59	1	0,11	1,41
Toplam	931	99,79	92,18	79	8,7	7.82

Çalışma sekli frekans tablosu incelendiğinde; ev hanımı olan müşterilerin kötü müşteri olma eğilimlerinin, örgenci ve kamu sektörü çalışanlarının ise iyi müşteri olma eğilimlerinin yüksek olduğu görülmektedir. Öğrenciler ise %98,59 oranıyla iyi, %1,41 oranıyla kötü müşteriler içerisinde yer almaktadırlar. Buna göre örgencilerin büyük oranda iyi müşteri olma eğilimi gösterdikleri söylenebilir.

4.2. Kredi Kartlarında En Yakın Komşu Tekniği ile Skorlama

Kredi kartı müşterilerine ait 1806 gözlemden müşteri tipi, çalışma şekli, öğrenim durumu, aylık net gelir, maaş, medeni hal, başka banka kredi kartı, başka banka kart limiti, cinsiyet, ikamet süresi,araba sahibi, yaş değişkenlerine ait son beş veri (1802,1803,1804,1805,1806) X1, X2, X3, X4, X5 olarak tanımlanarak en yakın komsularının iyi ya da kötü olmalarına göre sınıflandırılmaya çalışılmıştır. Verilere aşamalı kümeleme tekniklerinden en yakın komsu tekniği uygulanmıştır ve uzaklık ölçü birimi olarak öklityen uzaklık seçilmiştir. Analizler SPSS paket programlamada yapılmıştır.

X1 için dendogramda bulunan sonuçta en yakın komsular;

X1 müşterisinin sınıflandırılması

1 379
1 930
1 1259
X1 1801
1 725

1259, 930, 379, 725 inci durumlardır. Bu komsular da 1 (iyi) koduyla kodlanmış oldukları için X1'in iyi olduğu yönünde karar alabiliriz ki gerçekte de X1 iyi bir müşteridir.

X2 için dendogramda bulunan sonuçta en yakın komsular;

X2 müşterisinin sınıflandırması

2 1686
2 1116
2 1506
2 449
2 1612
2 1748
2 959
X2 1802
1 1606
1 348
2 1475
2 1649
2 1160
2 845

1606, 348, 1475, 1649, 959, 1748 inci durumlardır. Bu komsular da hem 2 (kötü) koduyla hem de 1 (iyi) koduyla kodlanmış oldukları için X2'nin durumu için net bir yorum yapılamaz. Ama iki tane iyi kodlanmış komsu dışındakiler kötü kodlanmış olduğu için kötü olmaya eğilimli olduğunu söyleyebiliriz ki gerçekte de X2 müşterisi kötü bir müşteridir.

X3 için dendogramda bulunan sonuçta en yakın komsular;

X3 müşterisinin sınıflandırması

2 1788
2 740
2 542
X3 1803
2 551

2 1233
2 1542

551, 1233, 1542, 542, 740, 1788 inci durumlarıdır. Bu komsular da 2 (kötü) koduyla kodlanmış oldukları için X3'ün kötü olduğu yönünde karar alabiliriz ki gerçekte de X3 kötü bir müşteridir.

X4 için dendogramda bulunan sonuçta en yakın komsular;

X4 müşterisinin sınıflandırması

2 182
2 1747
2 1751
2 1162
X4 1804
2 1469
2 1653

1469, 1653, 1162, 1751, 1747, 182 inci durumlardır. Bu komsular da kötü koduyla kodlanmış oldukları için X4'ün kötü olduğu yönünde karar alabiliriz ki gerçekte de X4 kötü bir müşteridir.

X5 için dendogramda bulunan sonuçta en yakın komsular;

X5 müşterisinin sınıflandırması

2 1051
X5 1805
2 1344
2 1625
2 534
2 1423
2 1258
2 1584

1344, 1625, 534, 1423, 1258, 1584, 1051 inci durumlarıdır. Bu komsular da kötü koduyla kodlanmış oldukları için X5'in kötü olduğu yönünde karar alabiliriz ama gerçekte X5 kötü değil iyi bir müşteridir.

4.2.1. SONUÇ

Kredi kartı müşterilerine ait müşteri tipi, çalışma şekli, öğrenim durumu, aylık net gelir, maaş, medeni hal, başka banka kredi kart limiti, cinsiyet, ikamet süresi, araba sahibi, yaş değişkenleriyle yapılan kümeleme analizi sonucunda en yakın komsular incelenmiş, ilgili

kişilerin yüksek oranda doğru sınıflandıkları gözlemlenmiştir. Burada en yakın komsu sayısı k seçiminde gerekli olan K ve M değerleri bilinmediği için rastgele belirlenen sayıda yakın komsulara bakılarak tahminlerin doğru olup olmadığı hakkında yorumda bulunulmuştur. Uygulama hakkında genel bilgi olması açısından bu teknik izlenmiştir. Fakat en yakın komsu teorisine göre k sayısı belirlenmeden bu tekniğin uygulanması yetersizdir. Sonuçların yorumlanmasında bu eksiklik göz önünde bulundurulmalıdır.

4.3. Kredi kartlarında Diskriminant Analizi ile Skorlama

Kredi kartı müşterilerini iyi müşteriler-kötü müşteriler olarak ayırmak için diskriminant analizi uygulanmıştır.

Kayıp veriler

VERİLER	N	Yüzde
Geçerli Veri	1426	79
Kayıp Veri	379	21
Toplam	1805	100

Kredi kartı müşterileri için başka bankya ait kart limiti, aylık net gelir, müşteri yaşı değişkenlerine ait 1805 gözlem değeri ile diskriminant analizi yapılmak istenmiştir. Tabloya bakıldığında; gözlemler içinde grup kodu (iyi–kötü) tanımlı olmayan hiçbir gözlem değeri bulunmadığı, 379 gözlem değeri için değişkenlerden en az birinde kayıp veri bulunduğu görülmektedir. Hem grup kodu tanımlı olmayan hem de en az bir değişken değerinin eksik olduğu hiçbir gözlem değeri yoktur. SPSS'de iyi–kötü olarak tanımlanan bu gözlem değerlerine uygulanan diskriminant analizi sonuçları incelenmiştir.

Değişkenler kovaryans-korelasyon matrisi

Kovaryans/Korelasyon Matrisi		Müşteri Yaşı	Aylık Net Gelir	Başka Banka Kart Limiti
Kovaryans	MÜŞTERİ YAŞI	98,46	2961,3	3096,28
	AYLIK NET	2961,3	10596179	4251575,12
	BAŞKA BANKA KART LİMİTİ	3096,28	4251575,12	12293630,3
Korelasyon	MÜŞTERİ YAŞI	1	0,09	0,08
	AYLIK NET	0.091	1	0,37
	BAŞKA BANKA KART LİMİTİ	0,08	0,37	1

Yukarıdaki tablo incelendiğinde; modele giren değişkenler arasında yüksek korelasyon olmadığı gözlenmektedir.

Değişken yapıları doğrultusunda, iyi müşteriler ve kötü müşterilere ait kovaryans matrisi;

İYİ-KÖTÜ		**Müşteri Yaşı**	**Aylık Net Gelir**	**Başka Banka Kart Limiti**
İYİ	MÜŞTERİ YAŞI	100,08	1.378,75	1.089,44
	AYLIK NET	1.378,08	5.684.511,24	3.455.191,36
	BAŞKA BANKA KART LİMİTİ	1.089,44	3.455.191,36	14,852.727,08
KÖTÜ	MÜŞTERİ YAŞI	96,97	4.415,06	4.983,78
	AYLIK NET	4.415,06	15.108.059,89	4.983.136,96
	BAŞKA BANKA KART LİMİTİ	4.939,78	4.983.136,96	9.942.832,18
TOPLAM	MÜŞTERİ YAŞI	98,49	3.342,42	3.447,59
	AYLIK NET	3.342,42	11.037.280,51	4.676.189,11
	BAŞKA BANKA KART LİMİTİ	3.447,59	4.676.189,11	12.660.979,70

Kovaryans matrislerinin homojenliğine bakarken iki grupta da ilgili kovaryans matrislerinin çok farklılık gösterip göstermediğini incelenir. İlgili değişken çiftlerinin iki grupta da farklı kovaryans değerlerine sahip olması kovaryans matrislerinin homojen olmadığı seklinde yorumlanabilir.

Box-M test sonucu

Box-M testi		51,45
F değeri	Serbestlik derecesi	6
	Serbestlik derecesi 2	774036,34
	Anlamlılık Seviyesi	0,00

Box's M testi " H_0 : Gruplar arası kovaryans matrisleri homojendir" hipotezini test etmek için kullanılır. Test sonucu elde edilen p (Sig)=0,000 belirlenen anlamlılık seviyesinden (0.05) küçük olduğundan dolayı H_0 hipotezi reddedilebilir. Yani gruplar arası kovaryans matrisi homojen değildir ve diskriminant analizi varsayımlarından biri gerçekleşmemiş olur. Bu durumda homojenlik varsayımına gerek duymayan Karesel Diskriminant Analizi uygulanarak sonuçlar aşağıda incelenmiştir.

Grup Ortalamaları Eşitlik Testi

Grup Ort.Eşitlik testi	"Λ"	F	Sd.1	Sd.2	Anlamlılık Sev.(p)
MÜŞTERİ YAŞI	1,00	1,08	1	330	0,03
AYLIK NET GELİR	0,96	14,78	1	330	0,00

BAŞKA BANKA KART LİMİTİ	0,97	10,89	1	330	0,00

Wilk's Lambda çok değişkenli bir anlamlılık testidir. "V" 0 ile 1 arasında değişmektedir. Küçük değerler ilgili değişkenin grupları ayırmada güçlü bir etkisi olduğunu, büyük değerler ise ilgili değişkenin grupları ayırmada daha az etkili olduğunu göstermektedir. Diğer bir istatistik olan F degeri için hesaplanan p belirlenen anlamlılık seviyesinden (0,05) küçük olduğu için aylık net gelir, öğrenim durumu, müşteri tipi değişkenlerinin iyi müşterileri kötü müşterilerden ayırmada kullanılan modele anlamlı katkısı olduğu söylenebilir.

Özdeğer

Fonksiyon	Özdeğer	Varyans %	Kümülatif %	Kanonik Korelasyon
1	0,06	100,00	100,00	0,23

Yukarıdaki tabloda her bir kanonik değişkene ait öz değerin varyansın ne kadarını açıkladığını, kümülatif açıklanan varyansı, ve kanonik korelasyon değerini göstermektedir. Tablo değerlerine bakıldığında varyansın tek bir kanonik değişkenle %100 oranında açıklandığı görülmektedir. Özdeler, gruplar arası kareler toplamının grup içi kareler toplamına oranıdır. Kanonik korelasyon değeri de gruplarla diskriminant analizi skorları arasındaki başarıyı göstermektedir. 1'e yakın kanonik korelasyon değerleri gruplar ve diskriminant skorları arasında güçlü bir iliksiyi göstermektedir. Buradaki kanonik korelasyon değeri incelendiğinde modelin tahmin başarısının düşük olduğu söylenebilir.

Wilks'Lambda değeri

Fonksiyon Testi	Wilks'Lambda	Ki-Kare Değeri	Serbestlik Derecesi	Anlamlılık Sev.
1	0,95	18,54	3,00	0,00

"H_o: Fonksiyonların gruplar arası ortalamaları eşittir" varsayımını test eder. Wilk's Lambda değeri gruplar arasındaki fark ile açıklanamayan diskriminant fonksiyonlarının toplam varyansının oranını göstermektedir ve 0 ile 1 arasında değişmektedir. 0'a yakın değerler grup ortalamalarının farklı olduğunu, 1'e yakın değerler ise grup ortalamaları arasında fark olmadığını göstermektedir. Wilk's Lambda'nın chi-squre dönüşümü verilen serbestlik derecesinde anlamlılığı test eder. Anlamlılık seviyesi (p) küçük ise grup ortalamaları farklıdır, büyük ise de farklı değildir diye yorumlanabilir. Elde edilen anlamlılık seviyesi belirlenen anlamlılık düzeyinden (0,05) küçük olduğu için grupların ortalamaları farklıdır denilebilir.

Standardize edilmiş kanonik diskriminant fonksiyonu katsayısı

Standardize edilmiş kanonik diskriminant fonksiyonu katsayısı	Fonksiyon
MÜŞTERİ YAŞI	0,13
AYLIK NET GELİR	0,68
BAŞKA BANKA KART LİMİTİ	0,48

Değişkenler farklı ölçüm birimleriyle ölçüldükleri için standardize edilmiş kanonik diskriminant fonksiyonu katsayıları hesaplanır. Diskriminant modelindeki ayırıcı değişkenlerin bağımlı değişken üzerindeki göreceli etkisini gösteren standardize ayırma fonksiyonu katsayılarına göre, gruplar arasındaki en fazla ayırıcı etkide bulunan değişkenlerin önem sırası; aylık net gelir, başka banka kart limiti, müşteri yaşı biçimindedir. Skor=+0,0,68*aylık net gelir+0,13*müşteri yaşı+0,48*başka banka kart limiti seklinde ifade edilebilir

Yapı matrisi

	Fonksiyon
Yapı Matrisi	1
AYLIK NET GELİR	0,878
BAŞKA BANKA KART LİMİTİ	0,754
MÜŞTERİ YAŞI	0,238

Her bir tahmin edici değişkenin kanonik fonksiyonla grup içi korelasyonlarını içerir. Diskriminant fonksiyonunda her bir değişkenin katkısını ölçmenin bir yoludur. Diskriminant fonksiyonuna katkı sırasına göre aylık net gelir 0,878 ile en büyük katkıyı yapan değişkendir ve müşteri yaşı 0,238 ile an az katkıyı yapan değişkendir.

İyi-kötü müşteriler için diskriminant fonksiyonu katsayıları

Sınıflandırma Fonksiyonu Katsayıları	İYİ – KÖTÜ	
	İYİ	**KÖTÜ**
MÜŞTERİ YAŞI	0,35	0,36
AYLIK NET GELİR	0,00	0,00
BAŞKA BANKA KART LİMİTİ	0,00	0,00
Sabit	-6,98	-7,65

İyi ve kötü müşteri gruplarının sınıflama fonksiyonlarını gösterir.

Sınıflandırma

Sınıflandırma Sonuçları		İYİ-KÖTÜ	TAHMİN		TOPLAM
			İYİ	**KÖTÜ**	
Orijinal	Sayı	İyi	820	80	900
		Kötü	731	174	905
	%	İyi	91,11	8,89	100
		Kötü	80,77	19,23	100

Bu modelde 900 iyi müşterilerden 820 tanesi (%91,1) iyi olarak sınıflandırılmış, 80 (%8,9) tanesi kötü olarak sınıflandırılmıştır. 905 tane kötü gözlemden 174 tanesi (%19,23) kötü olarak sınıflandırılmış, 731 tanesi (%80,7) iyi olarak sınıflandırılmıştır.

4.3.1. SONUÇ

Kredi kartları verileriyle skorlamada diskriminant analizi için modele aylık net gelir, müşteri yaşı, başka banka kart limiti değişkenleri girmiştir. Analiz SPSS paket programında yapılmıştır. Gruplar arası kovaryans matrisi homojenligi sağlanamamıştır ve bu nedenle Karesel Diskriminant Analizi uygulanmıştır. Modele giren değişkenlerden modele en yüksek katkıyı aylık net gelir değişkeninin yaptığı görülmektedir.Varyans tek bir kanonik değişkenler %100 oranında açıklanmıştır ve modelin tahmin başarısı düşüktür. İyi müşteriler %91,11 oranı ile, kötü müşteriler ise %19,23 oranı ile tahmin edilebilmektedirler. Model kötü müşterileri tahmin etmede başarılı değildir.

5. SONUÇLAR VE GENEL DEĞERLENDİRME

Bu çalışmada, bir bankanın kredi kartı müşterilerine ait veriler bazı kredi skorlama teknikleri ile analiz edilmiştir. Uygulamada, istatistiksel tekniklerden Diskriminant Analizi, En Yakın Komsu tekniği kullanılmıştır. Çalışmanın amacı kredi skorlama ile ilgili kullanılabilecek bazı teknikleri aynı örneğe uygulayarak sonuçları kıyaslamak olduğu için seçilen bu tekniklerin teorilerine derinlemesine inilmemiştir. Kredi skorlamada kullanılabilecek yöntemlerin bir arada verilmesi açısında Lineer Programlama ve Tamsayı Programlama uygulamalarına yer verilmeksizin teorilerine kısaca değinilmiştir.

İlk uygulanan teknik; En yakın komsu yaklaşımı olup; bu teknikte varsayım kısıtlaması olmadığı için tüm değişkenlerle analiz yapılmış, analizde iyi müşteri-kötü müşteri oldukları bilinen fakat bilinmiyor olarak kodlanan beş müşterinin dendogramdaki en yakın komsularına bakılarak tahminde bulunulmuştur. En yakın komsu kümeleme tekniği ve öklit uzaklıklara dayanarak dendogram elde edilmiştir. Burada modelin eksikliği incelenecek en yakın komsu sayısının belirlenememesidir. Bu sayı *M* ve *K* bilgilerine bağlıdır ve bu bilgiler çalışma sırasında elde edilemediği için en yakın komsularla ilgili bir sayı belirlenememiştir. Bu durumda; uygulamanın ne şekilde gelişeceğinin gösterilmesi açısından rastgele en yakın komsular incelenerek bazı yorumlar yapılmıştır ama bunun sadece gösterim amaçlı olduğu unutulmamalıdır. İleride yapılacak çalışmalarda *M, K* değerlerinden yola çıkarak belirlenen sayı ile sağlıklı analizler yapılabilir.

İkinci uygulanan teknik; teknik varsayım kısıtları olan diskriminant analizidir.aylık net gelir, müşteri yaşı, başka banka kart limiti değişkenleri ile model kurulmak istendiğinde homojenlik varsayımının gerçeklenmediği görülmektedir ve bu nedenle Karesel Ayırma Analizi kullanılmıştır. Diskriminant modeline aylık net gelir 0,878 ile en büyük katkıyı yapan değişkendir. Müşteri yaşı değişkeni ise 0,238 ile en az katkıyı yapan değişkendir ve de adımsal diskriminant analizi modeline girmemiştir. Modelin genel basarı oranı %55,15'tir. Bu basarı oranı iyi müşterileri tahminde %91,1'e kadar çıkarken kötü müşterilerin tahmininde %19,2'dir.

Modellere giren değişkenler incelendiğinde; tüm modellere aylık net gelir değişkeninin büyük katkısı olduğu gözlemlenmektedir. Mevcut veri yapısı ve bu çalışma için bahsedilen değişkenlerin kredi skorlaşmada etkili değişkenler olduğu söylenebilir.

KAYNAKLAR

1. Agresti, A., (2001), Categorical Data Analysis, Jhon Willey & Sons, Florida
2. Doğan, İ., (2002) "Kümeleme Analizi İle Seleksiyon", Türk J.Vet.Anim.Sci.
3. Efe Ö.ve Kaynak, O.,(2004), Yapay Sinir Ağları ve Uygulamaları, Boğaziçi Üniversitesi Yayınları ,İstanbul
4. Kunter, M.H., Nachtsheim,C.,J., Neter, J. ve Wasserman, W., (1996), Applied Linear Statistical Models, The Mc.Graw – Hill Campanies, New York
5. Özdamar , K., (2002a), Paket Programlar ile İstatistiksel Veri Analizi 1, Kağan Kitapevi,Eskişehir
6. Özdamar , K., (2002b), Paket Programlar ile İstatistiksel Veri Analizi 2, Kağan Kitapevi,Eskişehir
7. S.P.S.S., (2001),Answer Tree 3.0 User's Guide S.P.S.S..Inc.,Chicago.
8. S.P.S.S., (2001), Regression Models 11.0,S.P.S.S., Inc Chicago.
9. Tatlıdil, H.,(1996), Uygulamalı Çok Değişkenli İstatistiksel Analiz, Cem Web Ofset,Ankara
10. Yıldız,D., (1995), Diskriminant Analizine İlişkin Bazı Yöntemler,İstanbul
11. Şakar,H., (2001),Bankalarda Kredilendirme Teknikleri,İstanbul
12. Akgüç,Ö.,(2000),Kredi Taleplerinin Değerlendirilmesi,İstanbul
13. Seval,B.,(1990),kredilendirme Süreci ve kredi Yönetimi,İstanbul
14. Andersen, E.B., (1994), The Statistical Analysis of Categorical Data, Springer-Verlag, Berlin-Heidelberg.
15. Gujarati, D.N., (1999), Temel Ekonometri (Çev., Ü., Senesen & G, Senesen), Literatür Yayınları,İstanbul
16. Menard,S., (1995), Applied Logistic Regression Analysis,Sage Publications,California
17. Malhodra, Naresh, K., (1993), Marketing Research an Applied Orientation,Prentic Hall İnternational
18. Gujarati, D.N., (1999), Temel Ekonometri (Çev., Ü., Senesen & G, Şenesen), Literatür Yayıcılık, İstanbul
19. Menard, S., (1995), Applied Logistic Regression Analysis,Sage Publications,California

20. Thomas,L.C., Edelman, D.B.ve Crook, J.N., (2002),Credit Scoring and Its Applications, Society for Industrial and Applied Mathematics, Philadelhia
21. Sharma, S., (1996), Applied Multivariate Techniques, John Willey & Sons, Canada
22. Mays, E., (2004), Credit scoring for risk managers: the Handbook for Lenders, Thomsan/South-Western, Ohio
23. Lewis M. E., (1992), An Introduction to Credit Scoring, Fair Isaac Co., San Rafael
24. Yegorova, I. (2001), A Succesful Neural Network-based Methodology for Predicting Small Business Loan Default
25. Yurtoğlu, H., (2005), Yapay Sinir Ağları ile Öngörü Modellemesi :Bazı Makroekonomik Değişkenler için Türkiye Örneği

ÖZGEÇMİŞ

Doğum Tarihi	01.01.1979	
Doğum Yeri	Konya	
Lise	1993 - 1997	Konya Lisesi (Yabancı Dil Ağırlıklı Lise)
Lisans	1998 - 2002	Selçuk Üniversitesi Fen-Edeb.Fak.İstatistik Bölümü

Çalıştığı Kurumlar ve Görev Seviyesi

2004 -	Akbank T.A.Ş. Nalçacı Şubesi Bireysel Müşteri İlişkileri Yönetici Yardımcısı