

Kredi Kartı Taleplerinin Değerlendirilmesinde Grup ve Bireysel Kredi Puanlama Modellerinin Karşılaştırılmalı Bir Analizi

Hüseyin İNCE*
Bora AKTAN**

Özet

Kredilendirme, bankacılığın en temel işlevi olmakla birlikte aynı zamanda en riskli faaliyetlerinden biridir. Bu nedenle bankaların, kredilendirme faaliyetlerini verimli ve kredilerin geri dönmemesinden doğabilecek zararları en az düzeyde tutulabilecek şekilde yürütmeleri gerekir. Bankalar, bu sebeple, son yıllarda ayrıştırma analizi, logit ve probit modelleri ya da sınıflama ve regresyon ağaçları ve yapay sinir ağları gibi çeşitli teknikler yardımı ile kredilendirme faaliyetlerinde hızlı ve sağlıklı karar verilmesini sağlayan kredi puanlama sistemini kullanmaktadır. Bu çalışmanın amacı, grup modelleri ile bireysel kredi puanlama modellerin performanslarının karşılaştırılmasıdır. Bu amaç doğrultusunda, yapay sinir ağları ve karar ağaçları teknikleri kullanılarak geliştirilmiş bireysel modeller ile Bagging ve Adaboost teknikleri ile elde edilmiş grup modelleri kullanılmıştır. Yapılan analiz ve değerlendirmeler sonucu grup kredi puanlama modellerinin bireysel modellere üstünlük sağladığı görülmüştür.

Anahtar Kelimeler: Bankacılık, Banka Kredileri, Kredi Puanlama, Yapay Sinir Ağları, Sınıflandırma ve Regresyon Ağaçları, Grup Modelleri

JEL Sınıflaması: G21, G32, C44, C45

Abstract - A Comparative Analysis of Individual and Ensemble Credit Scoring Techniques in Evaluating Credit Card Loan Applications

One of the main tasks of a bank is to lend money. As a financial intermediary, one of its roles is to reduce lending risks. Bank lending is an art as well as a science. Success depends on techniques used, knowledge and on an aptitude to assess both credit-worthiness of a potential borrower and the merits of the proposition to be financed. In recent years, banks have increasingly used credit-scoring techniques to evaluate the loan applications they receive from consumers. Credit-scoring techniques are usually based on discriminant models or related techniques, such as logit or probit models or neural networks, in which several variables are used jointly to set up a numerical score for each loan applicant. This study explores the performance of both individual models by using neural networks, and classification and regression trees and ensemble models by using Bagging and Adaboost techniques. Experimental studies using real world data sets have demonstrated that the ensemble models outperform the other credit scoring models.

Keywords: Banking, Bank Lending, Credit-Scoring, Neural Networks, Classification and Regression Trees, Ensemble Models

JEL Classification: G21, G32, C44, C45

* Doç. Dr., Gebze Yüksek Teknoloji Enstitüsü, İşletme Fakültesi, İşletme Bölümü

** Yrd. Doç. Dr. Yaşar Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Finansman Bölümü

1. Giriş

Kredilendirme bankacılığın en temel işlevi olmakla beraber aynı zamanda en riskli faaliyetlerinden biridir. Tasarruf sahiplerinden topladığı mevduatı kredi olarak kullandıran ve ülke ekonomileri açısından büyük önem taşıyan bankaların güvenilir, etkin ve verimli bir şekilde çalışmaları zorunlu olup, tahsis edilen kredilerin yasal faizleriyle birlikte bankalara geri dönüşü sistemin vazgeçilmez bir özelliğidir (Seval, 1990; Berk, 2001; Blochlinger ve Leipold, 2006; Aydın, 2006; Akgüç, 2007). Son yıllarda özellikle ticari bankalar, tüketici kredisi pazarındaki hızlı büyüme ve zorlu rekabet şartları dolayısıyla kredilendirme kararlarında başvuruların potansiyel riskini tespit edebilmek amacıyla gerek finansal kurumların kendileri ve gerekse de araştırmacılar tarafından geliştirilmiş çeşitli değerlendirme modellerini yaygın olarak kullanmaya başlamışlardır. (Aydın, 2006; Vojtek ve Kocenda, 2006; Çinko, 2006; Zhao, 2007; Bodur ve Teker, 2005; Avery, Calem, v.d., 2004; Crook ve Banasik, 2004; Jacobson ve Roszbach, 2003; Karan ve Arslan, 2008; Ince ve Aktan, 2009).

Önceleri, bankalar kredilendirme kararlarını verirken yalnızca finansal analistler tarafından ortaya konmuş ilke ve kurallardan faydalanırlardı. Zaman içinde kredi talebinde bulunan kişi ya da kurum sayısı büyük bir hızla arttığından hem ekonomik hem de insan gücü açısından söz konusu talepleri değerlendirmek hemen hemen imkansız hale gelmeye başlamıştır. Bu şekilde, gerek verimlilik gerekse insan zafiyetleri göz önüne alındığında süratli ve bir o kadar da sağlıklı kararların verilebilmesi için gerek finansal kurumlar gerekse de akademisyenler tarafından birçok sayısal yöntem geliştirilmiştir. Kredi kararlarını desteklemek için geliştirilen istatistiksel metotlar, parametrik olmayan istatistiksel metotlar, yapay zeka yaklaşımlarına dayanan teknikler önerilmiştir (Thomas, 2000). Lee, Chiu, v.d., (2006) bu tekniklerin amacını, kredi için başvuranları borcunu ana para ve faiziyle geri ödeyebilecek durumda olan 'ödeme gücü yüksek kredi' grubu veya büyük olasılıkla geri ödeyemeyeceği için reddedilen 'ödeme gücü zayıf kredi' grubu olarak ayırmak olarak belirtmektedir.

Çinko (2006), kredi kartı değerlendirme işlevinin, geçmiş bilgilerden yararlanılarak kredinin tahsis edilip edilmeyeceği kararının ortaya çıkarılması dolayısıyla, bunun klasik bir karar verme problemi olarak düşünülmesi gerektiği, bu sebepten dolayı da gerçekte kredi değerlemenin bir çeşit veri madenciliği (*data-mining*) olduğunu vurgulamaktadır. Kredi değerlendirme problemleri temel olarak daha genel ve geniş olarak tartışılan sınıflandırma problemleri içinde yer alır (Lee, Chiu, v.d., 2002). Sınıflandırma problemleri karar alma desteği, finansal tahminleme, dolandırıcılık ve sahtekarlıkların saptaması, pazarlama stratejisi, süreç kontrolü ve ilgili diğer alanlarda geniş uygulama olanağına sahip olması nedeniyle uzun zamandır iş hayatında önemli rol oynamaktadır (Chen, Han, v.d., 1996; Fayyad, Piatetsky-Shapiro, v.d., 1996; Lee, Chiu, v.d., 2006).

Sınıflandırma problemleri istatistiksel metotlardan yapay zeka algoritmalarına kadar değişen farklı teknikler kullanarak çözülebilir. Kredi puanlama modellerini oluşturur-

ken regresyon, doğrusal ve doğrusal olmayan ayrıştırma analizleri, logit ve probit modelleri içeren istatistiksel metotlardan geniş ölçüde faydalanılmıştır (Vojtek ve Koãenda, 2006; Lee, Chiu, v.d., 2002; Lee, Chiu, v.d., 2006). Uygulamada en popüler yöntemler doğrusal ayrıştırma analizi, lojistik regresyon ve bunların türevleridir. Bu yöntemler diğerlerine nazaran daha kolaydır ve hali hazırda yorumlanabilir basit sonuçlar verir. Fakat bu yöntemlerin kredi puanlama hesabı uygulanmalarında bazı sınırlamaları söz konusudur. Öncelikle, bu yöntemler yüksek ölçülü girdilerle ve küçük örnek boyutlarıyla ilgili problemlerde etkili değildir. En önemlisi, bu teknikler doğrusal ayrılabilirlik ve normallik varsayımlarına dayanır. Dahası, model sürecine otomasyon uygulamak ve sürekli güncellenen bir akış sağlamak oldukça zordur. Yang'a göre (2007), çevre ve popülasyon zaman içinde değiştikçe statik modeller genelde bu duruma adapte olamazlar. Bu yüzden, bu tür modeller taslaklarından tekrar yapılandırılmalıdır.

Bu klasik yöntemlere ek olarak, yapay zeka teknikleri kredi başvurularını değerlendirme amacıyla uygulanmıştır. Uygulamacılar ve araştırmacılar kredi puanı için, k- en yakın komşu (Henley ve Hand, 1996), karar ağaçları (Lee, Chiu, v.d., 2006; Martens, Baesens, v.d., 2007), yapay sinir ağları (Lee, Chiu, v.d., 2002; Malhotra ve Malhotra, 2002; West, 2000), genetik programlama (Ong, Huang, v.d., 2005) ve destek vektör makinası modellerini (Martens, Baesens, v.d., 2007) içeren değişik teknikler geliştirmişlerdir. Lee, Chiu, v.d. (2006), bu tekniklerin, bağımlı ve bağımsız değişkenlerin karmaşık doğrusal olmayan bağlantılar gösterdiği ayrıştırma analizleri ve lojistik regresyonda alternatif olarak kullanılabileceğini öne sürmüşlerdir.

Diğer taraftan, yukarıda sayılan ve yapay zeka tekniklerini içeren bireysel kredi değerlendirme modellerinin bir kombinasyonu olan gruplama tekniklerinin ardında yatan fikir, her bir bileşen modelinin özelliklerini, veri kümesinde bulunan farklı örnekleri yakalamak için kullanmaktır. Hem teorik hem ampirik çalışmalar, doğrulukları artırmak için gruplanmanın etkili bir yol olabildiğini göstermektedir (Pal, 2007).

Bankacılık Düzenleme ve Denetleme Kurumu (BDDK)'nun Kasım 2008 verilerine göre, dünyada gelişmekte olan piyasalar içinde yer alan ülkemiz bankacılık sektörünün 370.782 milyon TL'ye yükselen toplam kredileri içinde kredi kartları dahil tüketici kredilerinin 118.710 milyon TL (toplam kredilerin %32'si)'lik bir hacime ulaştığı, diğer taraftan, içinde bulunduğumuz ve ABD'de başlayarak küresel bir boyut kazanan kredi krizi (credit crunch) göz önüne alındığında, kredilendirme kararlarının ülke ekonomileri açısından ne denli büyük bir önem arz ettiğini bir kez daha ortaya çıkarmaktadır.

Çalışmanın bu giriş bölümünden sonra İkinci Bölüm'de kısaca kredilendirme kararlarında kullanılan kredi puanlama (*credit scoring*) modelleri ile ilgili bir yazın taraması yapılarak yapay sinir ağları (*neural networks*), karar ağaçları (*decision trees*) ve grup tekniklerinin (*ensemble techniques*) kısa bir özeti gözden geçirilecektir. Ayrış-

tırma analizi (*discriminant analysis*), lojistik regresyon (*logistic regression*), yapay sinir ağları ve sınıflandırma ve regresyon ağaçları (*classification and regression trees*) kullanılarak yapılan kredi puanlama hesaplamalarının bulguları Üçüncü Bölüm'de tartışılmaktadır. Son olarak, Dördüncü Bölüm çalışmanın sonuç bölümünü oluşturmaktadır.

2. Araştırma Yöntemi ve Yazın Taraması

Bireylerin/firmaların finansal yükümlülüğünü beklendiği gibi yerine getirip getiremeyeceklerini tahmin etmek amacıyla nicel ve nitel analizler yardımıyla bir takım kredi değerlendirme modelleri geliştirilmiştir. Kredi değerlendirme modelleri kredi talebinde bulunanları etkileyebilecek öznel ve nesnel faktörleri araştırır. Bu Aşamada, kredi değerlendirme modelleri yazın desteği ile kısaca açıklanacaktır.

Bilgi bilimine (Chen ve Liu, 2004), önemli katkıları olan yapay zeka teknikleri kredi puanlama hesabında da uygulanabilir. Karar ağaçları, yapay sinir ağları, genetik programlama ve k- en yakın komşu gibi bazı yapay zeka teknikleri uygulamaclar ve araştırmacılar tarafından kredi başvurularının değerlendirilmesi için geliştirilmiştir (Malhotra ve Malhotra, 2002; West, 2000; Ong, Huang, v.d., 2005; Lee, Chiu, v.d., 2006; Lee ve Chen, 2005; Lee, Chiu, v.d., 2002). Bu çalışmada, yapay sinir ağları, sınıflandırma ve regresyon ağaçları ve grup tekniklerine dayanan kredi değerlendirme modelleri geliştirip karşılaştırılacaktır.

2.1. Yapay Sinir Ağları

Yapay sinir ağları mühendislik, ticaret, öngörü alanlarıyla ilgili bir çok problemi çözmeye yaygın olarak kullanılmış (Vellido, Lisboa, v.d., 1999; Lee, Chiu, v.d., 2006) finasta kredi puanlama hesapları ile ilgili problemlere de uygulanmıştır (Lee, Chiu, v.d., 2006; Lee, Chiu, v.d., 2002).

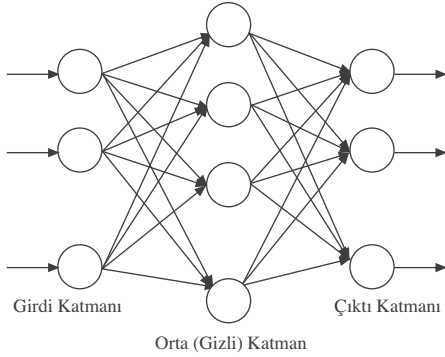
Nispeten basit işlemler elementlerin (nodüllerin) yüksek derecede birbirine bağlı ağlar kullanılarak girdileri istenilen çıktılara dönüştüren algoritmik prosedür olan Yapay Sinir Ağları (YSA), doğrusal olmayan regresyon ve sınıflandırma modelleridir. Yapay sinir ağı hücresinde temel olarak dış ortamdan ya da diğer nöronlardan alınan veriler yani girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktılar bulunmaktadır. Dış ortamdan alınan veri, ağırlıklar aracılığıyla nörona bağlanır ve bu ağırlıklar ilgili girdinin etkisini belirler. Toplam fonksiyonu ise net girdiyi hesaplar; net girdi, girdilerin ağırlıklı toplamına eşittir. Aktivasyon fonksiyonu işlem süresince net çıktıyı hesaplar ve bu işlem aynı zamanda nöron çıkışı verir. Genelde aktivasyon fonksiyonu doğrusal olmayan (non-linear) bir fonksiyondur (Ince, 2006). Ağ topolojisi sinir hücrelerinin organizasyonlarını ve bağlantı türlerini verir. Hücreler bir seri katmanı üzerinde farklı katmanlardaki farklı hücrelerle bağlantılı olacak şekilde ayarlanır. Giriş katmanı denilen ilk katman girdileri alır. Tek gizli katmanlı bir yapay sinir ağı

örneđi Şekil 1’de verilmiştir (Crook, Edelman, v.d., 2007). Uygun ağ topolojisi (ör., gizli katmandaki gizli nöron sayısı) aşağıdaki eşitlikle sağlanabilir:

$$h = \sqrt{n + m} + \alpha, \quad \alpha \in [1,10] \quad (1)$$

h , gizli birimlerin sayısı, n ve m sırasıyla girdi ve çıktılarının sayısıdır.

Şekil 1: Üç Katmanlı Bir YSA



Yapay sinir ağları, ileri beslemeli ve geri beslemeli ağlar olarak farklı kategorilere ayrılabilir. İleri beslemeli ağlarda nöronlar sadece bir önceki katmandan girdi alıp sıradaki katmana çıktı gönderirler. Çok katmanlı yapay sinir ađı (MLP) eğitmek için geri beslemeli algoritma (*back propagation*, BP) kullanır. En uygun ağırlığı bulmak için, BP ağ hatasını en aza indirmeye çalışır. Öğrenme oranı olarak adlandırılan adım büyüklüğü eğitim başlamadan önce belirlenmelidir. Bu oran sıfır ile bir arasında bir değerdir. Öğrenme oranı bu algoritma için çok önemlidir çünkü düşük değerler optimuma ulaşma öncesinde işlem sürecini yavaşlatmaya eğilimliken daha büyük değerler ağda sapmalara sebep olur ve yakınsama yapamaz. Geri besleme algoritmasının bazı varyasyonları yerel minimuma ulaşma, ağır kovaryans ve yapay sinir ađında bulunan detaylı bilginin fazla işlenmesi gibi zorlukların üstesinden gelmek için önerilmiştir (Haykin, 1994; 1998).

2.2. Karar Ağaçları

Karar ağaçları, çeşitli pazarlama uygulamaları (*müşteri ilişkileri yönetimi, doğrudan postalama*), insan kaynakları yönetimi (*geçmişte işletmeye en faydalı olan bireylerin özelliklerini kullanarak işe alma süreçlerinin belirlenmesi*), tıp (*tıbbi gözlem verilerinden yararlanarak en etkin kararların verilmesi*), hangi değişkenlerin satışları etkilediğinin belirlenmesi, üretim yönetimi gibi uygulamalarda kullanılmaktadır (Akpinar, 2000). Bunun yanında Shen, v.d. (2007), kredi kartı verileri ile sahtecilik tespiti için yapılan sınıflandırma modeli çalışmasında karar ağaçlarının etkin olarak kullanıldığını ve olumlu sonuçlar ortaya koyduđu gözlemlemiştir. Yapılan bir başka çalışma

da, bankanın sahtecilik tespit sistemi karar ağacı teknikleri üzerinde geliştirilmiş olup sistem sonuçları oldukça başarılı olmuştur (Lee, Chiu, v.d., 2002; Lee, Chiu, v.d., 2006). Veri madenciliği tekniklerinden biri olan karar ağaçları firmaların bilançolarının incelenmesinde kullanılmış ve tahmin doğruluğu olarak tatmin edici sonuçlar ortaya koymuştur (Kirkos, Spathis, v.d., 2007). Bunun yanında ileride kavramsal olarak açıklanacağı üzere grup, Bagging ve Boosting algoritmalarının uygulamalarında da karar ağaçları teknikleri kullanılmıştır.

Karar ağacı, tümevarımsal muhakeme kullanarak sınıflandırma modeli oluşturmadaki farklı yöntemlerden biridir. Basit kural serileri uygulayarak oluşturulan verinin bölümlenmesini gösteren ağaç şeklinde bir yapı modeli oluşturur. Bu kurallar parçalanmanın tekrar işlemi yoluyla tahminlemede kullanılabilir. Lee, Chiu, v.d. (2005)'ne göre, karar ağacı teorisi kredi değerlendirme için oldukça uygundur ve geniş alanda kullanılır. ID3, C4.5, Sınıflandırma ve Regresyon Ağaçları (CART) ve CHAID karar ağaçları algoritmaları öngörü ve sınıflandırma için kullanılmıştır.

ID3 algoritması Quinlan (1993), tarafından karar ağacı üretmek için önerilmiştir. Bilgi kazancı teorisine dayanır. ID3 karar ağaçlarını branşlamaya nitelik olarak en uygun bilgi kazancını belirler ki bu şekilde oluşturulan ağaçlar basit yapılı olurlar (Zhao, 2007). Edinilen bilgi tüm veri setinin haricinde belli başlı nitelikler kullanarak karar ağacının düğümleri tarafından oluşturulan alt-ağaçların entropi katsayısı ile hesaplanabilir. ID3'ün sakıncası edinilen bilgiyi nitelik seçmek için kural olarak kullanmasıdır çünkü dallandırma daha yüksek nitelik değerleri üzerinde meyilli olacaktır. Bu olumsuzluğu ortadan kaldırmak için ID3'ün bir uzantısı ve revizyonu olan C4.5 önerilmiştir (Chang ve Chen, 2008). C4.5 algoritmaları bilgi kazancı-oranını niteliği bölümlenmek için kullanır. Büyük boyutlardaki veri seti üzerinde çalışmak için C4.5 karar ağacı algoritması revize edilerek C5 algoritması geliştirilmiştir (Chang ve Chen, 2008). Buna ek olarak, hız açısından C4.5'tan daha hızlıdır ve hafıza kullanımında daha etkilidir. (Tso ve Yau, 2007)'ya göre, C5'in C4.5 algoritmasına göre şu avantajları vardır: "(1) nominal parçalar için düğüm birleştirme opsiyonu yoktur, (2) yanlış sınıflandırma maliyeti belirlenebilir; (3) destekleme ve çapraz-değerlendirme özellikleri mevcuttur; ve (4) ağaçlardan kural setleri oluşturmak için algoritmalar daha gelişmiştir".

Bu algoritmaların yanı sıra, bazı araştırmacılar başka karar ağaçları teknikleri sunmuşlardır. Bunlardan biri istatistiksel bir metot olan ve Sınıflandırma ve Regresyon Ağacı olarak bilinen Breiman, Friedman, v.d. (1984), tarafından geliştirilen karar ağacı tekniğidir. Bu hem regresyon problemlerinin çözümü hem de sınıflandırma için kullanılan tekrarlanan bir bölümlenmedir. Temel olarak bir nesneyi iki veya daha çok popülasyonda sınıflandırmak için kullanılır. Sınıflandırma ve Regresyon Ağacı algoritması aşağıdaki gibi üç aşamada özetlenebilir (Chang ve Chen, 2008):

- 1- Bu aşamada, tekrarlanan bölümlenme tekniği değişken seçmek ve ayırma ölçütü kullanarak noktaları ayırmak için kullanılır. Birçok bozukluk ve farklılık ölçü-

sü kullanarak en iyi belirleyici seçilir (Gini, ikişerli, sıralı ikişerli ve en küçük ka-reler sapması). Bu ölçülerin nasıl hesaplanacağı ile ilgili detaylı bilgiye Breiman, Friedman, v.d. (1984)'dan ulaşılabilir ki burada amaç, hedef değışkene bağı olarak olabildiğince homojen veri alt setleri oluşturmaktır.

- 2- Büyük bir ağaç tanımlandıktan sonra, Sınıflandırma ve Regresyon Ağacı mini-mum maliyet karmaşıklığını kapsayacak şekilde fazlalıkları çıkarma prosedürü uygular. Fazlalıkları çıkarma prosedürü, gelişmiş en büyük ağaçtan başlayıp tek bir ağaç düğümü kalana kadar süren bir işlem içinde iç içe geçmiş ağaç alt setleri temin eder.
- 3- Son aşamada, en düşük çapraz-değerli veya hata kriteri kullanarak en uygun ağaç seçilir.

2.3. Gruplama (Ensemble) Teknikleri

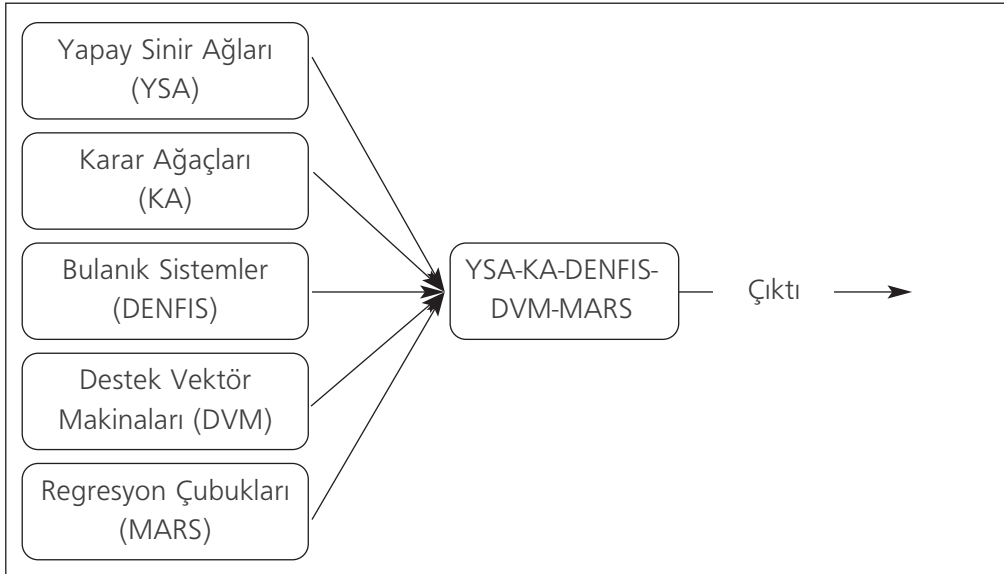
Ensemble'in sözlük anlamı; "birlik, grup, takım"dır. Oza, (2006) ensemble'ı "*birkaç farklı yöntem ile yapılan tahminlerin kombinasyonunu veren fonksiyon*" şeklinde tanımlamaktadır. Bu yöntemde karışık bir hesaplama önce basit hesaplamalara bölünür. Sonra elde edilen çözümlerin birleştirilmesiyle karmaşık sorun çözülür.

Gruplama (*ensemble*) tekniklerinin ardında yatan fikir, her bir bileşen modelinin tek özelliklerini, veri kümesinde bulunan farklı örnekleri yakalamak için kullanmaktır. Bir komisyonun üyelerini seçerken kullanılan temel mantık ile ensemble oluştururken kullanılan aynıdır (Oza, 2006). Komisyonun her bir üyesi kendi alanında mümkün olduğu kadar yetenekli olmalı, fakat üyeler birbirini tamamlayıcı olmalı. Eğer üyeler birbirini tamamlayıcı olmazsa ve her konuda aynı fikirdeyseler, bu durumda bir komisyon oluşturmaya gerek yoktur ve tek bir üye yeterlidir. Eğer üyeler birbirini tamamlayıcı olursa, bu durumda bir veya birkaç üye hata yaptığıında diğer üyelerin bu hatayı düzeltme olasılığı daha fazladır.

Hem teorik hem ampirik çalışmalar, doğrulukları artırmak için gruplanmanın etkili bir yol olabildiğini göstermektedir. Özgün çalışmalarında Bates ve Granger (1969), farklı tekniklerin doğrusal bir kombinasyonunun, bireysel tekniklerden daha küçük bir hata varyansı vereceğini göstermişlerdir. O zamandan bu yana, pek çok araştırmacı gruplama veya bileşik tahminler üzerinde çalışmalarda bulunmuştur. Makridakis, Anderson, v.d. (1982), farklı tek modellerin birleştirilmesinin, tahmin doğruluğunu artırmada genel uygulama haline geldiğini bildirmiştir. Sonrasında, Pelikan, De Groot, v.d. (1992), zaman serileri tahmin doğruluğunu artırmak amacıyla, farklı ileri-beslemeli sinir ağlarını toplamayı önermiştir. Sürekli bağımlı değışkenli tahmin problemleri için grup tekniklerinin bazıları, doğrusal grubu (örn.; basit ortalama ağırlıklı ortalama), yığılmalı regresyon ve doğrusal olmayan grubu (örn, sinir ağı merkezli) içermektedir (Benediktsson, Sveinsson, v.d., 1997; Yu, Wang, v.d., 2005; Perrone ve Cooper, 1993; Breiman, 1996).

Hansen, McDonald, v.d. (1992), birkaç yapay sinir ağından oluşan bir grup kullanılarak bir yapay sinir ağı modelinin genelleştirme kapasitesinin önemli ölçüde artırılabilirliğini bulgulamıştır. Çoğu zaman, sınıflandırma problemlerinden dolayı, yeni örnekleri tasnif etmek için genel anlamda çoğunluk oylamasına (*majority voting*) göre, bir grup sistemi birtakım yollarla bireysel sınıflama kararlarını birleştirmektedir. Temel düşünce, bir takım modelleri (*uzmanları*) eğitmek ve oy kullanmalarına izin vermek. Çoğunluk oylaması yönteminde, bütün bağımsız modellere aynı derecede önem verilmektedir. Modelleri birleştirmenin bir diğer şekli, bağımsız modellerin yetersiz ölçüde önemle eğitildiği ağırlıklı oylama yoludur. Bu, bazı ağırlıkları bağımsız modeller tarafından verilen tahmine bağlayıp birleştirmekle gerçekleşmektedir. Olmeda ve Fernandez (1997), bir genetik algoritmanın bağımsız modellerin optimum kombinasyonunu belirlediği genetik algoritma merkezli bir grup sistemi ortaya koymuşlardır; bu şekilde doğruluk maksimuma çıkarılmıştır. Zhou, Wu, v.d. (2002), gruplanan yapay sinir ağları modeli üzerinde kapsamlı bir çalışma yürütmüş ve bir grup oluşturmak için bir takım sinir ağlarını kullanmanın tüm sinir ağlarını kullanmaktan daha iyi olduğunu öne sürmüştür. Bunun yanı sıra, sinir ağlarının grubun parçası haline gelmesi için sinir ağları mevcut dizisinden seçiminde kullanılabilir bir yaklaşım önermişlerdir. Ayrıca genetik algoritma, ağırlıkları bileşen ağlarına atamak için kullanılmıştır.

Şekil 2: Grup Sistem Tasarımı



Notlar: **YSA:** Neural Networks, **KA:** Decision Trees, **DENFIS:** Dynamic evolving neuro-fuzzy inference system, **DVM:** Support Vector Machines **MARS:** Multivariate adaptive regression splines.

Belirli bir veri kümesi için bir akıllı teknik türü diğerini geçmekte ve farklı bir veri kümesi kullanıldığında sonuçlar tamamen zıt olabilmektedir. Hiçbir çoğunluğu kay-

betmemek ve akıllı tekniklerin avantajlarını toplamak için bir grup, her birine mutlak öncelik düzeyi verilmesiyle bütün bağımsız akıllı tekniklerin çıktılarını kullanmakta ve bir hakem yardımıyla çıktı sağlamaktadır. Bu çalışmada, farklı yapıdaki yapay sinir ağları modelleri, karar ağaçları gruplandırılarak performansları bireysel modeller ile karşılaştırılacaktır.

3. Uygulama

Grup ve yapay sinir ağları, karar ağaçları gibi bireysel teknikler kullanılarak geliştirilen kredi değerlendirme modellerinin performanslarını karşılaştırmak için bir finansal kurumdan alınan kredi kartı verileri kullanılmıştır. Veri seti sekiz tane açıklayıcı değişken ve bir tane bağımlı değişken içermektedir. Açıklayıcı değişkenler sırası ile cinsiyet, yaş, medeni durumu, eğitim durumu, meslek, gelir, müşteri türü ve başka bankalardan aldığı kredi kartı sayısıdır. Bağımlı değişken ise müşterinin kredi durumunu (iyi veya kötü) gösteren değişkendir. Veri setinde 1422 müşterinin bilgileri bulunmaktadır. Bunlar içerisinde, veri setinin yaklaşık %70'i yani 990 müşteri rassal olarak seçilmiş ve model parametrelerini tahmin etmek için eğitim seti olarak geri kalan kısım ise, modellerin performansını değerlendirmek için kullanılmıştır.

İlk olarak yapay sinir ağı ve karar ağacı modeli oluşturulacak ve bunlarla ilgili elde edilen sonuçlar verilecek, daha sonra ise grup teknikleri ile oluşturulan modelden çıkacak sonuçlar ile karşılaştırılacaktır. Çalışmada kullanılan yapay sinir ağı için optimum verimi sağlayacak topolojiyi belirlemedeki en önemli sorun gizli katmanda bulunan nöron sayısının belirlemektir. Lee, Chiu, v.d. (2002)'e göre kompleks sistemler tek gizli katmanlı yapay sinir ağı ile modellenebilir. Crook, Edelman, v.d. (2007), gizli katmandaki nöron sayısını belirlemek için Denklem 1' de verilen formülü önermişlerdir. Ayrıca, optimum topolojiyi oluşturmak için genetik algoritmalarda önerilmektedir. Bu çalışmada gizli katman sayısı ve nöron sayısını belirlemek için genetik algoritmalarından faydalanılmıştır (Ong, Huang, v.d., 2005). Neurosolution yazılımı kullanılarak en uygun topoloji seçilmiştir. Gizli katmandaki nöron sayısı 13 olarak belirlenmiştir. Algoritmanın sonuca ulaşma kriteri olarak hata kareleri ortalamasının karekökü (RMSE) 0.0001 den küçük veya iterasyon sayısı 5000 seçilmiştir.

Tablo 1, eğitim ve test seti için elde edilen tahmin sonuçlarını göstermektedir. Söz konusu tabloda görüldüğü gibi eğitim ve test seti için doğru tahmin etme oranı sırası ile %65.9 ve %66.9 olarak elde edilmiştir. Eğitim setinde iyi kredi puanına sahip 165 müşteri kötü krediye sahip müşteri olarak sınıflandırılmıştır. Ayrıca kötü kredi puanına sahip 173 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır. Test setinde ise, iyi kredi puanına sahip 74 müşteri kötü kredi puanına sahip olarak ve kötü kredi puanına sahip 69 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır.

Tablo 1: Eğitim ve Test Seti için YSA Kullanılarak Elde Edilen Sınıflandırma Sonuçları

Gerçek	Eğitim Seti Tahmin		Test Seti Tahmin	
	İyi	Kötü	İyi	Kötü
İyi	329(%67)	165(%33)	143(%66)	74(%34)
Kötü	173(%35)	323(%65)	69(%33)	146(%67)
Eğitim seti için doğru tahmin etme: %65.9				
Test seti için doğru tahmin etme: %66.9				

Kredi değerlendirme modeli olarak CHAID karar ağacı tekniği seçilmiştir. Bu model en çok kullanılan tekniklerden birisidir. Elde edilen sonuçlar Tablo 2' de verilmiştir. Buna göre, eğitim ve test seti için doğru tahmin etme oranı sırası ile %67.1 ve %62.4' tür. Eğitim setinde iyi kredi puanına sahip 232 müşteri kötü kredi puanına sahip müşteri olarak sınıflandırılmıştır. Ayrıca kötü kredi puanına sahip 105 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır. Test setinde ise, iyi kredi puanına sahip 102 müşteri kötü kredi puanına sahip olarak ve kötü kredi puanına sahip 48 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır.

Tablo 2: Eğitim ve Test Seti için Karar Ağaçları Kullanılarak Elde Edilen Sınıflandırma Sonuçları

Gerçek	Eğitim Seti Tahmin		Test Seti Tahmin	
	İyi	Kötü	İyi	Kötü
İyi	270(%54)	232(%46)	107(%51)	102(%49)
Kötü	105(%20)	416(%80)	48(%20)	142(%80)
Eğitim seti için doğru tahmin etme: %67.1				
Test seti için doğru tahmin etme: %62.4				

Araştırmacılar çeşitli türlerde gruplama modeli önermişlerdir (Zhou, Wu, vd., 2002; West, Dellana, v.d., 2005). Bu çalışmada Adaboost ve Bagging teknikleri kullanılarak geliştirilmiş kredi değerlendirme modelleri kullanılacaktır. Bu teknikler ile ilgili ayrıntılı bilgi (Tsai ve Wu, 2008; Zhou, Wu, v.d., 2002; West, Dellana, v.d., 2005) bulunabilir.¹

¹ Bagging "Bootstrap Aggregating" kelimelerinin birleşiminden oluşmuştur. Uygulanması kolay ve iyi sonuçlar veren bir algoritmadır (bkz Breiman, 1996). Özellikle veri sayısı az ise kullanılır. Farklı eğitim veri alt kümeleri, yenisiyle değiştirme (replacement) yöntemi ile eğitim setinden elde edilir. Her bir veri alt kümesi farklı sınıflandırıcıları eğitmek için kullanılır. Yapay sinir ağları ve karar ağaçlarında kolaylıkla uygulanabilirler. Boosting ise, Schapire (1990)'in tanımladığı bir algoritmadır. Bagging'te olduğu gibi boosting'te de ensemble'ı oluşturan sınıflandırıcılar verilerin tekrar örneklenmesi (resample) yöntemi ile elde edilmekte ve daha sonra çoğunluk oyları ile birleştirilmektedir. Boosting'te tekrar örnekleme ile her bir ardışık sınıflandırıcı için en çok öğretici eğitim seti elde edilmeye çalışılır. 1997 yılında boosting algoritmalarından en çok kullanılanı olan AdaBoost algoritması (bkz. Freund and Schapire, 1997) geliştirilmiştir.

Destek vektör makinaları, yapay sinir ağları ve karar ağaçları kullanılarak oluşturulan Adaboost ve bagging gruplama modelleri oluşturulmuştur. Destek vektör makinaları ile oluşturulan Bagging ve Adaboost modeli diğerlerinden daha iyi sonuç vermiştir. Tablo 3 ve 4 bu modeller kullanılarak elde edilen bulguları göstermektedir.

Tablo 3: Eğitim ve Test Seti için Bagging Metodu Kullanılarak Elde Edilen Sınıflandırma Sonuçları

Gerçek	Eğitim Seti Tahmin		Test Seti Tahmin	
	İyi	Kötü	İyi	Kötü
İyi	439(%89)	55(%11)	156(%72)	61(%28)
Kötü	68(%14)	422(%86)	63(%30)	147(%70)
Eğitim seti için doğru tahmin etme: %87.5				
Test seti için doğru tahmin etme: %70.9				

Tablo 3' de görüldüğü gibi eğitim ve test seti için doğru tahmin etme oranı sırası ile %87.5 ve %70.9' tür. Eğitim setinde iyi kredi puanına sahip müşterilerin %11'i kötü kredi puanına sahip müşteri olarak sınıflandırılmıştır. Ayrıca, kötü kredi puanına sahip müşterilerin %14'ü iyi kredi puanına sahip müşteri olarak sınıflandırılmıştır. Test setinde ise, iyi kredi puanına sahip 61 müşteri kötü kredi puanına sahip olarak ve kötü kredi puanına sahip 63 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır.

Tablo 4: Eğitim ve Test Seti için Adaboost Metodu Kullanılarak Elde Edilen Sınıflandırma Sonuçları

Gerçek	Eğitim Seti Tahmin		Test Seti Tahmin	
	İyi	Kötü	İyi	Kötü
İyi	438(%89)	56(%11)	135(%62)	82(%38)
Kötü	63(%14)	427(%86)	33(%33)	141(%67)
Eğitim seti için doğru tahmin etme: %87.9				
Test seti için doğru tahmin etme: %64.6				

Adaboost metodu da Bagging tekniğine benzer sonuçlar vermiştir. Eğitim ve test seti için doğru tahmin etme oranları sırası ile %87.9 ve %64.6 olarak hesaplanmıştır. Buna ek olarak, eğitim setinde iyi kredi puanına sahip müşterilerin %11'i kötü kredi puanına sahip müşteri olarak sınıflandırılmıştır. Ayrıca kötü kredi puanına sahip müşterilerin %14'ü iyi kredi puanına sahip müşteri olarak sınıflandırılmıştır. Test setinde ise, iyi kredi puanına sahip 82 müşteri kötü kredi puanına sahip olarak ve kötü kredi puanına sahip 33 müşteri iyi kredi puanına sahip olarak sınıflandırılmıştır.

İkinci olarak bireysel modeller ile gruplama modelleri karşılaştırılmıştır. Yukarıda verilen bilgilerde de görüleceği gibi gruplama tekniği ile oluşturulan modeller bireysel tekniklerden daha iyi sonuç vermiştir. Bunun en temel nedeni, gruplama modellerinin birden fazla bireysel teknik içermesinden kaynaklanan ve dolayısıyla hatalı sonuç bulma olasılığındaki düşüştür.

5. Sonuç

Kredi puanlama, sınıflandırma yöntemlerinin yardımı ile kredilendirme sürecinde daha süratli ve sağlıklı karar verilmesini sağlamaktadır. Bu makalede grup teknikleri ile oluşturulan kredi değerlendirme modelleri ile bireysel kredi değerlendirme modelleri karşılaştırılmıştır. Özellikle, ticari bankalar arasındaki rekabetin oldukça yoğun olması sebebiyle kredi değerlendirme bu kuruluşlar için hayati öneme sahiptir. Her geçen gün daha fazla finansal hizmet veren kurumlar daha etkin ve verimli stratejiler geliştirmek için çaba sarf etmekte, bu nedenle, farklı yapıda kredi değerlendirme modelleri geliştirmeye çalışmaktadırlar. Bu amaçla, bilgi teknolojisi ve modelleme tekniklerindeki gelişmeler daha ileri modelleri oluşturmaya olanak sağlamaktadır. Geleneksel tekniklerin yerini, yapay zeka teknikleri veya melez kredi değerlendirme modelleri almıştır. Bu çalışmanın amacı, gruplama modelleri ile bireysel kredi değerlendirme modellerinin performanslarının karşılaştırılmasıdır. Bu amaç doğrultusunda, yapay sinir ağları ve karar ağaçları teknikleri kullanılarak geliştirilmiş bireysel modeller ile Bagging ve Adaboost tekniği ile elde edilmiş gruplama modelleri kullanılmıştır. Yapılan analiz ve değerlendirmeler sonucu grup kredi değerlendirme modellerinin bireysel modellere üstünlük sağladığı bulgulanmıştır.

Kaynakça

1. Akgüç, Ö.. (2007). *Banka Yönetimi ve Performans Analizi*. Arayış Basım ve Yayıncılık: İstanbul.
2. Akpınar, H.. (2000). Veri Tabanlı Bilgi Keşfi ve Veri Madenciliği. *İ.Ü. İşletme Fakültesi Dergisi*, 29 (1): 1-22.
3. Avery, R. B., Calem, P. S. ve Canner, G. B.. (2004). Consumer Credit Scoring: Do Situational Circumstances Matter? *Journal of Banking and Finance*, 28:835-856.
4. Aydın, N. (Ed.). (2006). *Bankacılık Uygulamaları*. Anadolu Üniversitesi Yayını, No. 1711.
5. Bates, J.M. ve Granger, C.W.J.. (1969). The Combination of Forecasts. *Operational Research Quarterly*, 20 (4): 451-468.
6. Bankacılık Denetleme ve Düzenleme Kurumu, (2009) Aylık Bülten Ocak 2009. www.bddk.org.tr
7. Benediktsson, J. A., Sveinsson, J. R., Ersoy, O. K. ve Swain, P. H.. (1997). Parallel Consensual Neural Networks. *IEEE Transactions on Neural Networks*, 8:54-64.
8. Berk, N.. (2001). *Bankacılıkta Pazara Yönelik Kredi Talebi*, Beta Basım Yayın: İstanbul.
9. Blochlinger, A. ve Leippold, M.. (2006). Economic Benefit of Powerful Credit Scoring. *Journal of Banking and Finance*, 30: 851-873.
10. Bodur, C. ve Teker, S.. (2005). Credit Scoring of Companies: Application to the ISEM Companies. *İTÜ Dergisi/b*, 2(1): 25-36.
11. Breiman, L., Friedman, J. H., Olshen, R. A. ve Stone, C. J.. (1984). *Classification and Regression Trees*. Wadsworth and Brooks/Cole, Monterey.
12. Breiman, L.. (1996). Bagging Predictors. *Machine Learning*, 24 (3):123-140.
13. Chang, C. L. ve Chen, C. H.. (2008). Applying Decision Tree and Neural Network to Increase Quality of Dermatologic Diagnosis. *Expert Systems with Applications*, 3(6): 4035-4041.
14. Chen, M. S., Han, J. ve Yu, P. S.. (1996). Data Mining: An Overview From a Database Perspective. *IEEE Trans. Knowledge Data Engineering*, 8(6): 866-883.
15. Chen, S. Y. ve Liu, X.. (2004). The Contribution of Data Mining to Information Science. *Journal of Information Science*, 30(6): 550-558.
16. Çınko, M.. (2006). Kredi Kartı Değerlendirme Tekniklerinin Karşılaştırılması. *İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi*, 5 (9): 143-153.
17. Crook, J. ve Banasik, J.. (2004). Does Reject Inference Really Improve the Performance of Application Scoring Models? *Journal of Banking and Finance*, 28: 857-874.
18. Crook, J. N., Edelman, D. B. ve Thomas, L. C.. (2007). Recent Developments in Consumer Credit Risk Assessment. *European Journal of Operational Research*, No.183.

19. Fayyad, U., Piatetsky-Shapiro, G. ve Smyth, P.. (1996). The KDD Process for Extracting Useful Knowledge From Volumes of Data. *Communications of the ACM*, 39: 27–34.
20. Freund, Y. ve Schapire, R. E.. (1997). Decision-Theoretic Generalization of on-line Learning and application to Boosting. *Journal of Computer and System Sciences*, 55 (1): 119-139.
21. Hansen, J., McDonald J. ve Slice, J.. (1992). Artificial Intelligence and Generalized Qualitative Response Models: an Empirical Test on Two Audit Decision-Making Domains. *Decision Science*, 23:708–723.
22. Haykin, S.. (1994). *Neural Networks*. Macmillan College Publishing Company Inc, New York.
23. Haykin, S.. (1998). *Neural Networks: A Comprehensive Foundation*. 2e. Prentice Hall, Upper Saddle River, NJ.
24. Henley, W. E. ve Hand, D. J.. (1996). Ak-Nearest Neighbor Classifier for Assessing Consumer Credit Risk. *Statistician*, 44 (1): 77–95.
25. Ince, H.. (2006). Yapay Sinir Ağlarının Portföy Yönetiminde Kullanılması. *İktisat İşletme ve Finans*, 21 (4): 114-126.
26. Ince, H. ve Aktan, B.. (2009). A Comparison of Data Mining Techniques for Credit Scoring in Banking: A Managerial Perspective. *Journal of Business Economics and Management*, 10(3): 233-240.
27. Jacobson, T. ve Roszbach, K.. (2003). Bank Lending Policy, Credit Scoring and Value at Risk. *Journal of Banking and Finance*, 27: 615-633
28. Karan, M. B. ve Arslan, O.. (2008). Consumer Credit Risk Factors of Turkish Households, *Bank and Bank Systems*, 3(1): 42-57.
29. Kirkos, E., Spathis, C. ve Manolopoulos, Y.. (2007). Data Mining Techniques for the Detection of Fraudulent Financial Statements. *Expert Systems with Applications*, 32 (4): 995-1003.
30. Lee, H., Jo, H. ve Han, I.. (1997). Bankruptcy Prediction Using Case-Based Reasoning, Neural Networks and Discriminant Analysis. *Expert Systems with Applications*, 13: 97–108.
31. Lee, T. S., Chiu, C. C., Lu, C. J. ve Chen, I.F.. (2002). Credit Scoring Using the Hybrid Neural Discriminant Technique. *Expert Systems with Applications*, 23: 245–254.
32. Lee, T. S. ve Chen, I. F.. (2005). A Two-Stage Hybrid Credit Scoring Model Using Artificial Neural Networks and Multivariate Adaptive Regression Splines. *Expert Systems with Applications*, 28: 743–752.
33. Lee, T. S., Chiu, C. C., Chou, Y. C. ve Lu, C. J.. (2006). Mining the Customer Credit Using Classification and Regression Tree and Multivariate Adaptive Regression Splines. *Computational Statistics and Data Analysis*, 50: 1113 – 1130.
34. Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibdon, M., Lewandowski, R., Newton, J., Parzen, E. ve Winkler, R.. (1982). The Accuracy of Extrapolation (time series) Methods: Results of a Forecasting Competition. *Journal of Forecasting*, 1 (2): 111-153.

35. Malhotra, R. ve Malhotra, D. K.. (2002). Differentiating Between Good Credits and Bad Credits Using Neuro-Fuzzy Systems. *European Journal of Operational Research*, 136(1): 190–211.
36. Martens, D., Baesens, B., Van Gestel, T. ve Vanthienen, J.. (2007). Comprehensive Credit Scoring Models Using Rule Extraction From Support Vector Machines. *European Journal of Operational Research*, 183(3): 1466 - 1476.
37. Olmeda, I. ve Fernandez, E.. (1997). Hybrid Classifiers for Financial Multicriteria Decision Making: The Case of Bankruptcy Prediction. *Computational Economics*, 10: 317–335.
38. Ong, C. S., Huang, J. J. ve Tzeng, G. H.. (2005). Building Credit Scoring Models Using Genetic Programming. *Expert Systems with Applications*, 29(1): 41–47.
39. Oza, N. C.. (2006). Ensemble Data Mining Methods, Encyclopedia of Data Warehousing and Mining. *Idea Group Reference*, pp.448-452.
40. Quinlan, J. R.. (1993). *C4.5: Programs for machine learning*. Morgan Kaufman, San Francisco, CA.
41. Pal, M.. (2007). Ensemble Learning with Decision Tree for Remote Sensing Classification. *Proceedings of World Academy of Science Engineering and Technology*, 26(December): 735-737.
42. Pelikan, E., De Groot, C. ve Wurtz, D.. (1992). Power Consumption in West-Bohemia: Improved Forecasts Decorrelating Connectionist Networks. *Neural Network World*, No.2, 701-712.
43. Perrone, M. P. ve Cooper, L. N. (1993). When Networks Disagree: Ensemble Methods for Hybrid Neural Networks. *Neural Networks for speech and Image Processing*, Chapman Hall, 126–142.
44. Schapire, R. E.. (1990). The Strength of Weak Learnability. *Machine Learning*, 5(2): 197-227.
45. Shen, A., Tong, R. ve Deng, Y.. (2007). Application of Classification Models on Credit Card Fraud Detection. *School of Management, Graduate University of the Chinese Academy of Sciences, China*, 1-4
46. Seval, B.. (1990). *Kredilendirme Süreci ve Kredi Yönetimi*. İ.Ü. İşletme Fakültesi, Muhasebe Enstitüsü Yayın No.59: İstanbul.
47. Thomas, L. C.. (2000). A Survey of Credit and Behavioral Scoring: Forecasting Financial Risk of Lending to Consumers. *International Journal of Forecasting*, 16: 149–172.
48. Tsai, C. F. ve Wu, J. W.. (2008). Using Neural Network Ensembles for Bankruptcy Prediction and Credit Scoring. *Expert Systems with Applications*, 34(4): 2639–2649.
49. Tso, K. F. G. ve Yau, K. K. W.. (2007). Predicting Electricity Energy Consumption: A Comparison of Regression Analysis, Decision Tree and Neural Networks. *Energy*, 32: 1761–1768.
50. Vellido, A., Lisboa, P. J. G. ve Vaughan, J.. (1999). Neural Networks in Business: A Survey of Applications (1992–1998). *Expert Systems with Applications*, 17: 51–70.

51. Vojtek, M. ve Koãenda, E.. (2006). Credit Scoring Methods. *Finance a úvûr-Czech Journal of Economics and Finance*, 56: 152-167.
52. West, D.. (2000). Neural Network Credit Scoring Models. *Computers and Operational Research*, 27: 1131–1152.
53. West, D., Dellana, S. ve Qian, J.. (2005). Neural Network Ensemble Strategies for Financial Decision Applications. *Computers & Operations Research*, 32: 2543–2559.
54. Yang, Y.. (2007). Adaptive Credit Scoring with Kernel Learning Methods. *European Journal of Operational Research*, 183(3): 1521-1536.
55. Yu, L., Wang, S. Y. ve Lai, K. K.. (2005). A Novel Non-Linear Ensemble Forecasting Model Incorporating GLAR and ANN for Foreign Exchange Rates. *Computers and Operations Research*, 32 (10): 2523–2541.
56. Zhao, H.. (2007). A Multi-Objective Genetic Programming Approach to Developing Pareto Optimal Decision Trees. *Decision Support Systems*, 43: 809–826.
57. Zhou, Z. H., Wu, J. ve Tang, W.. (2002). Ensembling Neural Networks: Many Could be Better Than All. *Artificial Intelligence*, **137** (1-2):239-263.