

**ÇOK DEĞİŞKENLİ
UYARLAMALI REGRESYON UZANIMLARI**

MULTIVARIATE ADAPTIVE REGRESSION SPLINES

BİRCE ÜNAL

Hacettepe Üniversitesi
Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin
İSTATİSTİK Anabilim Dalı için öngördüğü
YÜKSEK LİSANS TEZİ
olarak hazırlanmıştır.

2009

ÇOK DEĞİŞKENLİ UYARLAMALI REGRESYON UZANIMLARI

Birce Ünal

ÖZ

Bu çalışmada Çok Değişkenli Uyarlamalı Regresyon Uzanımları(Multivariate Adaptive Regression Splines/MARS) yöntemi ve bu yöntemle kurulan modellerin oluşturulma süreci ayrıntılı olarak araştırılmıştır. Son yıllarda değişkenler arasındaki ilişkilerin araştırılmasında karmaşık algoritmaların kullanılabilirliği artmıştır. Böylece modelleme yaparken değişkenler arasındaki karmaşık ilişkilerin daha detaylı bir şekilde incelenebilmesi mümkün kılınabilmiş ve kurulan modellerin geleneksel istatistik yöntemlere göre oldukça başarılı olduğu gösterilmiştir.

Yapılan araştırmada MARS modelleme yönteminin gelişimi, literatürdeki uygulamaları, MARS yazılımı ile gerçekleştirilen çözümleme süreci ve gerçek verilerle yöntemin uygulanabilirliği gösterilmiş, sonuçlar tartışılmıştır.

Çalışmada uygulamadan önce kullanılan MARS yazılımının özellikleri, kullanım aşamaları özetlenmiştir. Daha sonra yöntemin uygulanabilirliğinin gösterilebilmesi ve geleneksel istatistiksel yöntemlerle kıyaslanabilmesi amacıyla gerçek veri uygulaması gerçekleştirilmiştir. Uygulama için Türkiye'deki bir finans kuruluşundan elde edilen bireysel kredi başvuru verileri kullanılmıştır. Kredi kararının kabul ya da reddedilmesi durumuna göre modelleme yapılmıştır. Kurulan MARS modelinin geleneksel yöntemlerle karşılaştırılabilmesi için ikili lojistik regresyon yöntemi ile modelleme yapılmıştır. MARS modeli doğru sınıflandırma oranı, I. Tip Hata ve II. Tip Hata'ya göre ikili lojistik regresyon modelinden daha iyi sonuçlar vermiştir. Kurulan her iki model ile kredi kararını etkileyen faktörler araştırıldığında MARS modelinin daha detay yorumlar yapabilmeye olanak sağlayan sonuçlar verdiği tespit edilmiştir.

Anahtar Kelimeler: Çok Değişkenli Uyarlamalı Regresyon Uzanımları, MARS, Regresyon Uzanımları, Temel Fonksiyon, Kredi Değerlendirmesi

Danışman: Doç. Dr. Meral ÇETİN, Hacettepe Üniversitesi, İstatistik Bölümü

MULTIVARIATE ADAPTIVE REGRESSION SPLINES

Birce Ünal

ABSTRACT

In this study, Multivariate Adaptive Regression Splines (MARS) method and the modeling process of the MARS models studied in detail. The usability of complex algorithms in analysis of relations between variables has increased in recent years. Thus, it has become possible to analyse complex relations between variables in detail when modelling and it has been shown that the constructed models are considerably successful than the traditional statistical methods.

The development of MARS modelling method, its applications in litterateur, analysis process with MARS software and applicability of the method with real data are shown and results are discussed in this study.

Before the application, the features of MARS software and its utilization steps are summarized. Then a real data application is performed to be able to show the applicability of the method and to compare it with traditional statistical methods. Personal loan application data that is obtained from a financial institution in Turkey is used for application. Modelling is made according to acceptance and refusal of credit decision. In order to compare constructed MARS method with traditional methods, modelling is made with binary logistic regression method. MARS model has given better results than binary logistic regression according to classification success ratio, Type I and Type II Errors. With both of the constructed methods, it is observed that the MARS modelling gives better results which allow to make more detailed interpretation when the factors that affect credit decision is analysed.

Keywords: Multivariate Adaptive Regression Splines, MARS, Regression Splines, Basis Function, Credit Scoring.

Advisor: Doç. Dr. Meral ÇETİN, Hacettepe University, Department of Statistics

TEŞEKKÜR

Tez çalışması için seçtiğim konuda beni destekleyen, çalışma süresi boyunca isabetli yönlendirmeleriyle bana yardımcı olan tez danışmanım Sayın Doç. Dr. Meral ÇETİN'e, tez çalışmam boyunca benimle birlikte yorulan, hiçbir özveriyi esirgemeyen, bu günlere gelmemi sağlayan anneme ve aileme, bu süre boyunca bana sonsuz destek olan, moral veren, her seferinde devam etmemi sağlayan ev arkadaşşıma ve diğer tüm dostlarıma, iş hayatımdaki tüm yoğunluğa rağmen beni sürekli motive eden, destek olan yöneticilerim ve iş arkadaşşılarıma tüm içtenliğimle teşekkür ederim.

İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZ	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER DİZİNİ.....	iv
ŞEKİLLER DİZİNİ.....	vi
ÇİZELGELER DİZİNİ.....	viii
1. GİRİŞ	1
2. GENEL BİLGİLER	5
2.1. Çok Değişkenli Uyarlamalı Regresyon Uzanımları Yöntemi Kullanılarak Oluşturulan Modellerin Kullanım Alanları ve Gelişimi.....	5
2.2. Çok Değişkenli Uyarlamalı Regresyon Uzanımlarının Temeli.....	8
2.2.1. Regresyon uzanımı	8
2.2.2. Temel fonksiyon ve düğüm noktası	10
2.2.2.1. Temel fonksiyon algoritması	15
2.2.2.2. Süreklilik.....	18
2.3. Çok Değişkenli Uyarlamalı Regresyon Uzanımları Yöntemi Kullanılarak MARS Modelinin Kurulması	19
2.3.1. Varyans analizi ayrıştırması.....	21
2.3.2. Model seçimi.....	24
2.3.3. Çok değişkenli uyarlamalı regresyon uzanımları yönteminde kayıp değerlerin incelenmesi.....	26
2.4. Deneysel Bayes Yöntemiyle Temel Fonksiyon ve Düğüm Noktası Seçimi	26
3. UYGULAMA	31
3.1. MARS Yazılımı.....	32
3.2. Çözümlemede Kullanılan Veriler	43
3.3. MARS Çözümlemesi	46
3.3.1. Model seçimi.....	48
3.3.2. İleriye doğru adım.....	51
3.3.3. Geriye doğru adım.....	53
3.3.4. Varyans analizi	54
3.3.5. Temel fonksiyonlar	61

3.3.6. Modelin yorumlanması	65
3.4. İkili Lojistik Regresyon Çözümlemesi	87
3.5. Genel Sonuçlar	93
4. SONUÇ.....	96
KAYNAKLAR	98
EKLER.....	104

ŞEKİLLER DİZİNİ

Şekil 2.1. Regresyon uzanımları.....	9
Şekil 2.2. $y=f(x)$ fonksiyonu için düğüm noktaları.....	11
Şekil 2.3. MARS tarafından kullanılan $(x-t)_+$ ve $(t-x)_+$ temel fonksiyonları.....	12
Şekil 2.4. MARS ileriye doğru model kurma şeması.....	14
Şekil 2.5. Temel fonksiyonların oluşturulmasının ikili ağaç diyagramı ile gösterilmesi	17
Şekil 3.1. MARS yazılımı veri açma penceresi	33
Şekil 3.2. MARS yazılımında verilerin görünüşü.....	34
Şekil 3.3. MARS yazılımından elde edilebilecek grafik çeşitleri	35
Şekil 3.4. MARS yazılımında model kurulumu.....	36
Şekil 3.5. MARS yazılımında etkileşim matrisi.....	38
Şekil 3.6. MARS yazılımında serbestlik derecesinin belirlenmesi	39
Şekil 3.7. MARS yazılımında temel fonksiyon sayısının ve etkileşime girecek değişken sayısının belirlenmesi.....	40
Şekil 3.8. MARS yazılımı kullanılarak elde edilen $y=f(x)$ fonksiyon dağılımı	42
Şekil 3.9. Kredi kararı değişkeni histogram grafiği (MARS yazılımı).....	46
Şekil 3.10. MARS yazılımında en iyi modelin gösterildiği ekran	49
Şekil 3.11. MARS yazılımı model özet bilgilerinin bulunduğu ekran	50
Şekil 3.12. MARS yazılımı ANOVA ayrıştırma tablosunun yer aldığı ekran.....	56
Şekil 3.13. MARS yazılımında modeldeki değişkenlerin nispi önem düzeylerinin sıralandığı ekran.....	57
Şekil 3.14. MARS yazılımında sınıflandırma başarısının gösterildiği ekran.....	66
Şekil 3.15. BORCLULUK_ORANI ve kredi kararı arasındaki ilişki.....	68
Şekil 3.16. Şube görüşünün bulunduğu durumda BORCLULUK_ORANI ve kredi kararı arasındaki ilişki	69
Şekil 3.17. BELGELENEN_TOP_GELIR ve kredi kararı arasındaki ilişki	72
Şekil 3.18. Borçluluk oranı varlığında BELGELENEN_TOP_GELIR ve kredi kararı arasındaki ilişki	73
Şekil 3.19. Borçluluk oranı varlığında BEYAN_EDILEN_GELIR ve kredi kararı arasındaki ilişki	74
Şekil 3.20. Borçluluk oranı varlığında ve KKB güncel en kötü statü bilgisinin bulunmadığı koşulda BEYAN_EDILEN_GELIR ve kredi kararı arasındaki ilişki ..	75

Şekil 3.21. Borçluluk oranı varlığında KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki	76
Şekil 3.22. Borçluluk oranı varlığında, KKB'de ödemelerine ait gecikmenin olmadığı durumda KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki	77
Şekil 3.23. Borçluluk oranı varlığında, KKB'de ödemelerine ait gecikmenin olmadığı ve şube görüşünün belirtildiği durumda KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki	78
Şekil 3.24. Borçluluk oranı varlığında ve KKB güncel en kötü statü bilgisinin bulunmadığı durumda YENIDEN_DEG_SAYISI ve kredi kararı arasındaki ilişki..	79
Şekil 3.25. Borçluluk oranı varlığında KKB güncel en kötü statü bilgisinin bulunmadığı durumda KKB_TOP_BORC ve kredi kararı arasındaki ilişki	80
Şekil 3.26. Borçluluk oranı varlığında, şube görüşünün belirtilmediği durumda KKB_TOP_BORC ve kredi kararı arasındaki ilişki	81
Şekil 3.27. Borçluluk oranı varlığında, şube görüşünün belirtilmediği durumda SON_1AYDAKI_RET_ADEDI ve kredi kararı arasındaki ilişki.....	83
Şekil 3.28. Borçluluk oranı varlığında TALEP_TL ve kredi kararı arasındaki ilişki	84
Şekil 3.29. Borçluluk oranı varlığında ve KKB'de güncel en kötü statü bilgisinin bulunmaması durumunda TALEP_VADE ile kredi kararı arasındaki ilişki	85
Şekil 3.30. Borçluluk oranı varlığında ve KKB'deki son 6 aya ait en kötü ödeme bilgisinin ve güncelde en kötü ödemesine ait bir bilginin bulunmadığı durumda BANKAYLA_CAL_SURESİ ve kredi kararı arasındaki ilişki....	86
Şekil 3.31. Borçluluk oranı varlığında ve KKB'deki güncel en kötü ödeme bilgisinin bulunmadığı durumda YAS ve kredi kararı arasındaki ilişki	87
Şekil 3.32. SPSS'de kurulan modelin önem düzeyi	88
Şekil 3.33. SPSS'de kurulan modelin R ² değeri	88
Şekil 3.34. SPSS'de kurulan modelin sınıflandırma başarısı	88

ÇİZELGELER DİZİNİ

Çizelge 3.1. Veri setindeki değişkenler	44
Çizelge 3.2. Model kurulumunda seçilen değişkenler	46
Çizelge 3.3. Varyans ayrıştırma tablosu	55
Çizelge 3.4. Modele giren değişkenlerin nispi önem dereceleri	58
Çizelge 3.5. Kurulan modellere ait I. Tip Hata ve II. Tip Hata	94

BİRİNCİ BÖLÜM

1. GİRİŞ

Değişkenler arasındaki ilişkilerin modellenmesinde regresyon yöntemleri sıkça kullanılmaktadır. Kurulan modeller doğrusal ya da doğrusal olmayan yapıda olabilmektedir.

Kurulan modellerde değişkenler arasındaki ilişkilerin gerçek yapısının daha iyi anlaşılabilmesi için çok fazla değişken ve değişkenlerin etkileşimleri bir arada kullanılmaktadır. Ancak geleneksel yöntemler kullanılarak bu yapıda modellerin kurulabilmesi zordur. Son yıllarda bilgisayar teknolojisindeki gelişmelere paralel bir şekilde birçok değişkenin bir arada değerlendirildiği algoritmaların kullanılabilirliği artmıştır. Bu karmaşık algoritmaları kullanan regresyon yöntemlerinden birisi de “Çok Değişkenli Uyarlamalı Regresyon Uzanımları” (Multivariate Adaptive Regression Splines / MARS) tekniğidir. Doğrusal ve doğrusal olmayan geleneksel yöntemlerin aksine, kullanılacak bağımlı ve bağımsız değişkenler için herhangi bir varsayım olmadan çalışan bu teknikte, bağımsız değişkenlerin belirlenen farklı aralık değerleri için farklı katsayılar üretilir ve modele etkileşim terimlerinin de katılmasıyla gerçek yapıyı daha iyi yansıtan modeller kurulabilmektedir.

Y bağımlı değişken ve X açıklayıcı değişkenleri arasındaki ilişki Eş. 1.1’de verilmiştir. Burada ε ; Y ve X değişkenleri arasındaki ilişkiden, bağımlı değişkenin sapmasını göstermektedir.

Hata! Yer işareti tanımlanmamış.

$$Y = f(X) + \varepsilon \quad (1.1)$$

Parametrik olmayan modellemenin temelinde yatan görüş, Eş. 1.1’deki gibi bir model yapısındaki X değişkenleri arasında doğrusal olmayan bir ilişkinin bulunabileceği olasılığıdır. Bu modelde f(X), doğrusal, lojistik veya doğrusal olmayan bir ilişki içinde olabilmektedir.

Çok Değişkenli Uyarlamalı Regresyon Uzanımları yöntemi, 1991 yılında Stanford Üniversitesinden Jerome H. Friedman(1991a) tarafından geliştirilmiştir. Literatürde

bu yöntem, yabancı dildeki kendi orijinal adının baş harflerinin kısaltılmasıyla (**M**ultivariate **A**daptive **R**egression **S**plines: **MARS**) anılmaktadır. MARS'ın temeli bir bağımlı değişken ile bir açıklayıcı değişkenler seti arasındaki olası ilişkiyi belirlemekte “düzleştirme uzanımları” (smoothing splines) kullanılmasına dayanmaktadır. Yani bağımlı ve bağımsız değişkenler arasındaki doğrusal olmayan ilişkileri doğrusal yapıya dönüştürme amacıyla uygun dönüşümler bularak çalışır ve bağımsız değişkenler arasındaki etkileşimleri belirlemek için kullanılan bir yeniliktir. MARS tekniği bağımlı ve bağımsız değişkenler arasında oluşturulan fonksiyonel ilişkiye ait hiçbir varsayım içermeyen parametrik olmayan bir yöntemdir (Temel vd., 2005).

Yöntemde; değişkenler arasındaki ilişkideki değişimlerin meydana geldiği noktalar belirlenmektedir ve bu noktalar arasında her bir tanım aralığında uygun dönüşümlerle düz bir doğru elde edilmeye çalışılmaktadır. MARS yönteminde bu tanım aralıkları arasındaki kaymalar kontrol altına alınarak sağlanan süreklilikle birlikte düzleştirme yapılmaktadır. Söz konusu kaymalar, “düğüm noktası” olarak adlandırılan konumlarda meydana gelmektedir ve rejimler arasında düz bir geçişi sağlamaktadır. Düğüm noktaları arasında tanımlı olan ilişki nasıl bir yapıda olursa olsun (karesel, kübik, parabolik vb) amaç bu ilişkiyi doğrusal hale getirmektir. MARS yönteminin algoritması, tüm değişkenler karşısında olduğu kadar tüm değişkenler arasındaki olası tüm etkileşimleri içeren her düğüm konumunu belirlemeye yöneliktir (Sephton, 1994).

Olası her düğüm konumunun belirlenmesi, farklı değişken kombinasyonlarının kullanımı yoluyla yapılmaktadır ve başlıca bileşenlerin analiz edilmesiyle oluşturulan değişken kombinasyonlarına “temel fonksiyonlar” (basis functions) adı verilmektedir. Tüm uygun temel fonksiyon sayısı ve düğüm konumları belirlendiğinde, seçilen temel fonksiyonlar üzerinden uygun modelin tahminlerini yapabilmek için en küçük kareler regresyonuna başvurulmaktadır.

Tek bir bağımsız X_i değişkeni ile Y_i bağımlı değişkeni arasındaki ilişki modellendiğinde, Eş. 1.2'de görülen yapıda bir genel model yazılabilmektedir:

$$Y_i = \beta_0 + \sum_{k=1}^K \beta_k B_k(X_i) + \varepsilon_i \quad (1.2)$$

Eş. 1.2’de $B_k(X_i)$: X_i ’nin k ’inci temel fonksiyonudur. Temel fonksiyonlar, X_i ’nin yüksek oranda doğrusal olmayan dönüşümleri olabilir.

Sephton (2001, s.41) β_k parametrelerinin tahminini şu şekilde açıklamıştır:

“ β_k parametrelerinin tahmini, görülen eşitliğin tahmin edilmesiyle ortaya çıkan artıkların kareleri toplamını minimum kılacak şekilde yapılır”.

MARS yönteminin avantajı, açıklayıcı değişkenlerin çeşitli kombinasyonlarından oluşan temel fonksiyonları tahmin ederek açıklayıcı değişkenlerin hem kendi başına katkılarını hem de birbirleriyle etkileşimlerini bağımlı değişkeni açıklamak için kullanmasıdır.

MARS, artıkların karesi toplamını azaltan düğüm konumlarını tanımlamaktadır. Modeller iki aşamalı bir süreçle inşa edilmektedir. İlk aşamada değişkenleri ve potansiyel düğümleri test ederek aşırı doymuş (overfit) bir modele ulaşılır. İkinci aşamada araştırmaya uymayan gereksiz faktörler ve bileşenler elenmektedir. Bu yöntemde model seçimi, Craven ve Wabha (1979) tarafından geliştirilen “genelleştirilmiş çapraz geçerlilik” (generalized crossvalidation/ GCV) kriterini temel almaktadır. Ayrıca bu kriter dışında Friedman (1991a), düzeltilmiş $R^2 (\bar{R}^2)$ ’yi de bir karşılaştırma ölçütü olarak önermektedir. Etkileşim terimleri içeren bir model, sadece düzeltilmiş R^2 önemli ölçüde yüksekse tercih edilebilmektedir.

MARS tahminleri, modelin varyans analizinden (ANOVA) kolayca yorumlanabilir. Modelde yer alan her bir açıklayıcı değişkenin bireysel katkısı da belirlenebildiğinden değişkenlerin katkısı, belirli bir ANOVA fonksiyonu (değişken) modelden soyutlanarak, modelin düzeltilmiş R^2 ’si tahmin edilerek belirlenmektedir. Bu, her bir ANOVA fonksiyonunun (değişkenin) anlamlılığını yorumlamaya imkan vermektedir (Tunay, 2001).

MARS geleneksel yöntemlere göre son derece zor olan, çok boyutlu verilerin içinde gizlenmiş karmaşık veri yapısını, optimal veri dönüşümlerini ve verilerin karşılıklı etkileşimlerini belirleyebilme avantajı ile regresyon modellemesinde yeni bir yaklaşımdır. Küçük veri tabanları söz konusu olduğunda bile iyi bir regresyon modelinin geliştirilmesi çok zaman almakta ve emek istemektedir. Bununla birlikte, MARS yöntemiyle geniş veri tabanları ve çok karmaşık veri yapıları için regresyon modelleri kolayca geliştirilebilmektedir.

Son dönemde, ekonometri tahminlerinde doğrusal ve parametrik olmayan modelleme yaklaşımları geniş kullanım alanı bulmaktadır. Bunun başlıca nedeni, sözü geçen modern tahmin yöntemlerinin gerçek hayattaki karmaşık ekonomik ilişkilerin analiz edilmesinde önemli bir üstünlük taşımasıdır.

MARS yönteminde bağımlı ve bağımsız değişkenlerin dağılımları ile ilgili herhangi bir varsayım bulunmamakta, değişkenler kategorik veya sürekli olabilmektedir. Bağımlı ve bağımsız değişkenler arasında doğrusallık varsayımı bulunmaz. MARS hızlıdır ve diğer doğrusal modellere göre daha az veri gerektirmektedir (Lee and Chen, 2005). MARS yöntemi kayıp değerlerden ve aşırı uç değerlerden çok az etkilenmektedir (Salford Systems, 2001). Ayrıca Çoklubağlantıdan çok az etkilenir ve modelinin yorumlanması oldukça kolaydır.

Yukarıda kısaca tanımlanan MARS yöntemini konu alan bu çalışmada “Çok Değişkenli Uyarlamalı Regresyon Uzanımları” yönteminin teorik temelleri incelenecektir. Daha sonra Türkiye’de bir bankadan alınan kredi değerlendirmesiyle ilgili veri seti üzerinde yöntemin uygulanabilirliği Salford System yazılımı (MARS 3.0) kullanılarak gösterilecektir.

İKİNCİ BÖLÜM

2. GENEL BİLGİLER

1990'ların başında Stanford Üniversitesinden fizikçi ve istatistikçi Jerome H. Friedman tarafından geliştirilen “Çok Değişkenli Uyarlamalı Regresyon Uzanımları” (**M**ultivariate **A**daptive **R**egression **S**plines / **MARS**) yöntemi, bir bağımlı değişken ile bir açıklayıcı değişkenler seti arasındaki olası ilişkiyi belirlemekte “düzleştirme uzanımları” (smoothing splines) tekniği kullanılmasına dayanmaktadır. Veri setinden otomatik bir şekilde tanımlanan parametrelerin her biri ile düğüm konumlarında ilişkilendirilmiş temel fonksiyonların genişletilmesi yöntemine dayanarak, tıpkı tekrarlamalı parçalı regresyon gibi çalışan bir yöntem olarak sunulmuştur. Kullanılan algoritma ile değişkenlerin birbirleriyle olan kombinasyonlarından oluşturulan temel fonksiyonlardan meydana gelen modellerde; karmaşık etkileşimlerin birlikte ve tek tek yorumlanabilmesinde büyük bir kolaylık sağladığı gösterilmiştir (Friedman, 1991a).

2.1. Çok Değişkenli Uyarlamalı Regresyon Uzanımları Yöntemi Kullanılarak Oluşturulan Modellerin Kullanım Alanları ve Gelişimi

Çok değişkenli uyarlamalı regresyon uzanımları yöntemi ekonomi, biyoloji, tıp, ziraat, pazarlama alanlarında kullanılmıştır. Özellikle uygulamadaki kolaylığı sayesinde veri madenciliği alanında popüler olmuştur. MARS bir düzleştirme algoritmasıdır. Bu nedenle de her türlü veri analizinde rahatlıkla kullanılabildiği gibi diğer analizler için de verilerin istenilen şekilde düzenlenmesine imkan vermektedir. Aynı zamanda diğer analizlere başlamadan önce önemli değişkenlerin saptanmasına yardımcı olan bir araç olarak da kullanılabilmektedir.

Modellemelerde MARS'ın kullanılmasına örnek olarak K. Batu Tunay'ın çalışması gösterilebilmektedir. Türkiye'de 1987-2000 döneminde farklı parasal büyüklüklere göre paranın gelir dolaşım hızlarının tahminleri MARS yöntemi ile yapılmıştır (Tunay, 2001).

Devam eden yıl, MARS modellemesi doğrudan pazarlama yöntemi kullanılarak elde edilen sipariş oranları üzerinde uygulanmıştır. Çalışmada lojistik regresyon, doğrusal MARS temel fonksiyonları ve lojistik MARS temel fonksiyonları modellenmiş ve yorumlanmıştır (Deichmann et al., 2002).

Yang et al. (2003), toprak yoluyla böcek ilaçlarının bir yerden başka bir yere taşınmasıyla ilgili bir araştırmada MARS yöntemini kullanmışlardır. Çalışmadaki sonuçlar, standart hata ve korelasyon terimleri açısından Yapay Sinir Ağları kullanılarak elde edilen doğrusal regresyon ile karşılaştırılmıştır ve yöntemin ziraat verilerinde kullanımının potansiyel gücü gösterilmiştir.

MARS yöntemi, bilgi teknolojileri (IT, Information Technology) yatırımları ve firma performansları arasındaki ilişkinin açıklanmasında kullanılmıştır. Bilgi teknolojileri ve birimlerin performansı arasındaki ilişkinin, bu konudaki diğer çalışmalarda gösterildiğinden daha karmaşık olduğu bulunmuştur. Bu sayede yöneticiler için yeni bir bakış açısı getirilmiştir (Osei-Bryson and Ko, 2004). Aynı yıl içerisinde yine aynı araştırmacılar tarafından bilgi teknolojileri ile ilgili benzer bir çalışma da sağlık hizmetleri alanında yapılmıştır. Sağlık hizmetleri bilgi teknolojilerine yapılan yatırımlar ile organizasyonel verimlilik arasındaki karmaşık ilişki MARS yöntemi kullanılarak araştırılmış ve yeni görüşlere ulaşılmıştır (Ko and Osei-Bryson, 2004).

Veri madenciliğinde de oldukça popüler olan MARS, kredi skorlarının araştırılması konusunda da kullanılmıştır. Lee et al. (2006), çok değişkenli uyarlamalı regresyon uzanımları yöntemleriyle banka kredi kartı verilerine yapılan uygulamada doğru sınıflandırma oranı açısından diğer yöntemlere göre (bu yöntemler, Sınıflama ve Regresyon Ağaçları yöntemi -**Classification and Regression Tree, CART**-, diskriminant analizi, lojistik regresyon ve sinir ağları) daha iyi sonuçlar verdiği gösterilmiştir. Ayrıca 2. tip hata (kötü kredi etiketiyle sınıflandırılmış müşterinin iyi kredi etiketine sahip olması) açısından da araştırıldığında, düşük 2. tip hata oranına sahip olduğu gösterilerek hatalı sınıflandırma maliyetleri açısından bakıldığında daha iyi kredi skora yeteneğine sahip olduğu gösterilmiştir.

2007 yılında MARS yöntemi mühendislikteki yüksek boyutlu modellemede kullanılan yapay sinir ağları modeline alternatif olarak sunulmuştur (Cervellera et

al., 2007). Çalışmada dokuz boyutlu envanter tahmini probleminde ve sekiz boyutlu su rezervi probleminde uygulama yapılarak her iki yöntem karşılaştırılmıştır.

MARS izleyen yıllar içerisinde, tıp literatüründe de giderek daha fazla popüler olmaya başlamıştır. Kuhnert et al. (2000), çalışmalarında motosiklet kazaları sonucunda oluşan yaralanmalar üzerine yapılan epidemiyolojik (epidemiological) vaka-kontrol (case-control) verilerinde MARS yöntemini kullanarak yorumlama yapmışlardır. Sonuçlar Lojistik Regresyon ve CART (classification and regression trees) yöntemleriyle de karşılaştırılarak MARS'ın sadece bir modelleme aracı olmadığı, aynı zamanda lojistik regresyon gibi çok iyi bilinen yöntemlere göre daha detay analizlere imkan verdiği gösterilmiştir. Sonraki yıllarda kimya verileri üzerine yapılan MARS yöntemi uygulamasında karmaşık ilişkilerin modellenenebilirliği bazı yazarlarca gösterilmiştir (Xu et al., 2004). Devam eden yıllarda da mide ve bağırsaklardaki ilaç emilimleri hakkında moleküllerin emilim verileri kullanılarak modelleme çalışmaları gerçekleştirilmiş ve yöntemin bu veriler üzerine uygulanabilirliği gösterilmiştir (Deconinck et al., 2005; Deconinck et al., 2007).

MARS yöntemi kendi başına bir modelleme tekniği olmasının yanında aynı zamanda diğer modelleme yöntemleri için de bir araç olarak kullanılabilir. Gooijer et al. (1998), veri setindeki istatistiksel olarak anlamlı olan değişkenleri MARS yöntemi ile tespit edip, bu değişkenler için dummy değişkenler oluşturmuşlardır. Dolayısıyla MARS yönteminin modellemede doğrusallığın sağlanması ve dummy değişkenler oluşturmada kullanılması gösterilmiştir.

Bayesci Çok Değişkenli Regresyon Uzanımları (Bayesian Multivariate Adaptive Regression Splines) yaklaşımının kullanılarak, yaşam verilerinin (survival data) tek değişkenli ve çok değişkenli olarak modellenmesi, MARS yönteminin araç olarak kullanılmasının bir başka örneği olarak sunulmuştur (Mallick et al. 1999). Bu çalışmada MARS, hazard fonksiyonlarının oluşturulmasında kullanılmıştır.

Daha sonra Francis (2000) çalışmasında MARS ve Sinir Ağlarının karşılaştırdığı çalışmada, MARS ile ön süreç olarak kayıp verilerin ve kategorik değişkenlerin

temel fonksiyonlarını oluşturarak diğer süreçlere hazırlık yapılabileceğini göstermiştir.

2003 yılında benzer bir çalışma da TMARS olarak adlandırılan iki adımlı çok değişkenli regresyon uzanımlarının (two-step MARS, TMARS) kullanılmasıyla gerçekleştirilmiştir. Bu yöntemde doğrusal modellenen bir yapıda, modeldeki uyum eksikliğini arttıran değişkenler MARS algoritmasıyla temel fonksiyonlara dönüştürülerek elde edilen model yorumlanmıştır (Xu et al., 2003).

Devam eden yıl bazı araştırmacılar tarafından sınıflandırma probleminin modellenmesinde önce MARS kullanılmış ve daha sonra da bu yöntemle tespit edilen istatistiksel olarak anlamlı olan değişkenler kullanılarak yapay sinir ağı modeli kurulmuştur (Chou, 2004).

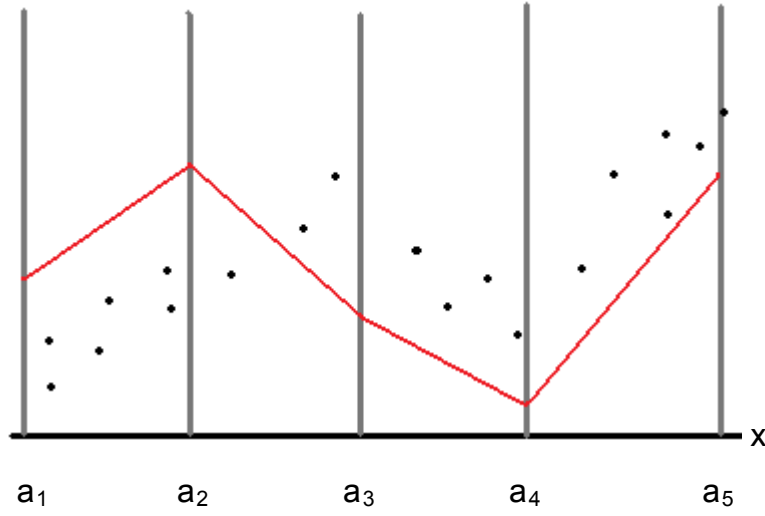
Lee and Chen (2005)'de MARS, yapay sinir ağı modeline destekleyici yöntem olarak kullanılmıştır. Önce MARS kredi derecelendirme modeli uygulanmış, daha sonra burada belirlenen önemli değişkenler yapay sinir ağı modeline girdi olarak alınmıştır. Sonuçlar, iki aşamalı melez modelleme yöntemiyle diğer yöntemlere göre en az hatalı sınıflama maliyeti elde edilmiştir.

2.2. Çok Değişkenli Uyarlamalı Regresyon Uzanımlarının Temeli

2.2.1. Regresyon uzanımı

MARS'ın temeli bağımlı değişken ile açıklayıcı değişkenler seti arasındaki olası ilişkiyi belirlemekte düzleştirme uzanımları kullanılmasına dayanmaktadır. Uzanımlar parçalı polinomlar olarak tanımlanmış özel fonksiyonlardır. Başka bir deyişle bir değişkenin dağılımını sadece noktalarla değil de o noktaların etkileşimini baz alarak oluşturulan çizgilerle göstermektir. Dolayısıyla MARS yöntemi, bağımlı ve bağımsız değişkenler arasındaki doğrusal olmayan ilişkileri doğrusal yapıya dönüştürme amacıyla uygun dönüşümler bularak çalışmaktadır ve bağımsız değişkenler arasındaki etkileşimleri belirlemek için kullanılır. Şekil 2.1'de $a_1, a_2, \dots, a_5$ düğüm noktaları arasında oluşturulan regresyon uzanımları grafiği

verilmiştir. Bu grafikten doğrusal olmayan ilişkinin nasıl doğrusallaştırıldığı açıkça görülmektedir.



Şekil 2.1. Regresyon uzanımları

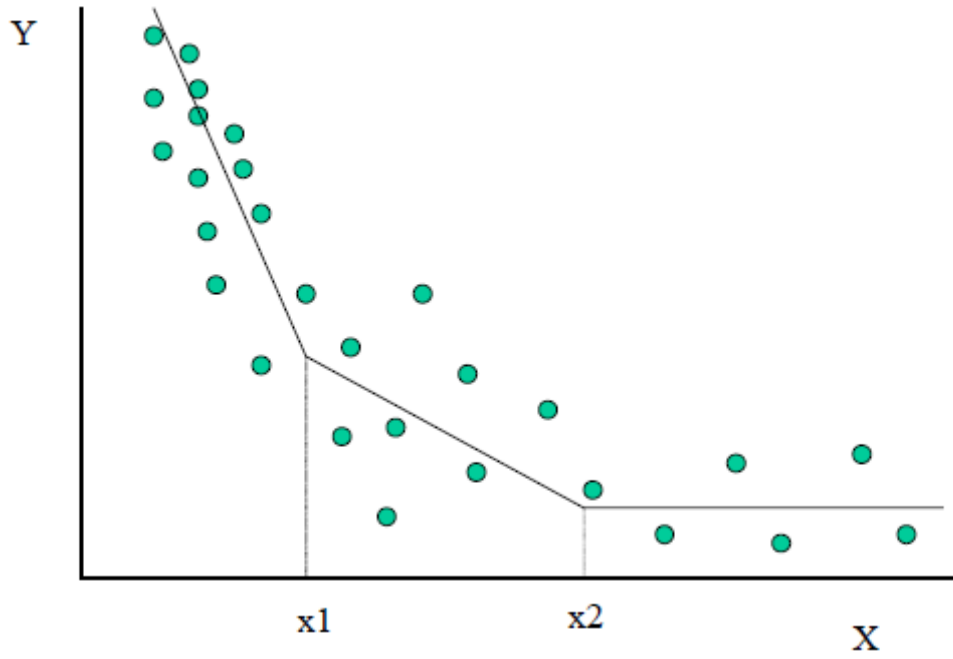
Regresyon eşitliklerinde bağımlı değişken ve açıklayıcı değişken arasındaki ilişki tek bir fonksiyon kullanılarak (örneğin; doğrusal, karesel, kübik ya da log linear) modellenmeye ve her bir bağımsız değişkenin bağımlı değişken üzerindeki etkisi tek bir katsayı ile açıklanmaya çalışılmaktadır. Doğrusallığın olmadığı durumlarda yüksek dereceden terimler (x^2 vb.) kullanılabilir, ancak bu terimlerin katsayıları tüm veri seti kullanılarak hesaplandığından, gerçek fonksiyonun lokal özellikleri (bağımlı değişkeni her noktada aynı biçimde etkilemediğinden) açıklanamayabilmektedir (Hastie and Tibshirani, 1990). Regresyon uzanımları yaklaşımı ise modelleme yaparken sonucu parçalı polinom olarak düşünmektedir. Bağımsız değişken parçalı sürekli doğrusal fonksiyon (doğrusal uzanım) olarak ya da bağımsız değişkenin türevi sürekli parçalı kübik fonksiyon olarak modellenebilmektedir. Parçalı polinom fonksiyonu $f(x)$, her bağımsız değişkenin tanım aralığının bir ya da birden fazla aralığa bölünmesiyle elde edilmektedir ve her aralıkta ayrı bir polinomla f fonksiyonunu temsil etmektedir (Hastie et al., 2001). Böylece uzanımlar, düğüm noktalarında parçaları düz bir şekilde birbirleriyle birleştirilmiş olan parçalı polinomlar olarak tanımlanabilmektedir (Eubank, 1988). Burada, düğüm noktalarının veri setindeki bir bölgenin sonu ve bir diğerinin başlangıcı olduğu unutulmamalıdır (Salford Systems, 2001). Başka bir deyişle regresyon uzanım fonksiyonu; düğüm noktalarında birbirleriyle birleşmiş

parçalı polinom temel fonksiyonlarının doğrusal kombinasyonlarıdır ve bu temel fonksiyonların katsayıları hata kareler toplamının minimize edilmesiyle elde edilmektedir. Bu süreç regresyon tahmin süreciyle hemen hemen aynı olduğundan, tahmin edilen uzanım regresyon uzanımı olarak adlandırılmaktadır (Osei-Bryson and Ko, 2004).

2.2.2. Temel fonksiyon ve düğüm noktası

MARS yönteminde değişkenlerin doğrusal kombinasyonları yerine, tek tek değişkenlerin doğrusal olmayan fonksiyonlarının doğrusal kombinasyonlarıyla modelleme yapılmaktadır. Bunun için de temel fonksiyonlara başvurulur. Seçilen temel fonksiyonlarda yapılan genişletmeler kullanılarak uygun modelin tahminleri elde edilmektedir.

Düğüm noktası; veri kümesindeki tanım aralıklarından birinin sonu, bir diğerinin başlangıcıdır. Dolayısıyla düğüm noktası fonksiyonun hareketinin değiştiği yerdir (Salford Systems, 2001). MARS yönteminde bağımsız değişken değerinin tanımlı olduğu aralıkta, doğrunun eğiminin değişmediği en son değer düğüm noktası olarak kabul edilmektedir. MARS modeli kurulurken, düğüm noktalarında tanımlı temel fonksiyonlardan yararlanılmaktadır. Şekil 2.2'de bağımsız değişken X ve bağımlı değişken Y arasındaki ilişkinin dağılımı görülmektedir. Burada X bağımsız değişkeninin düğüm noktalarındaki dağılımına dikkat edilirse; X1. düğüm noktasında $y=f(x)$ fonksiyonunun eğiminin değiştiği görülmektedir. Benzer şekilde X2. düğüm noktasında da fonksiyonun eğimi için aynı şey söylenebilmektedir.



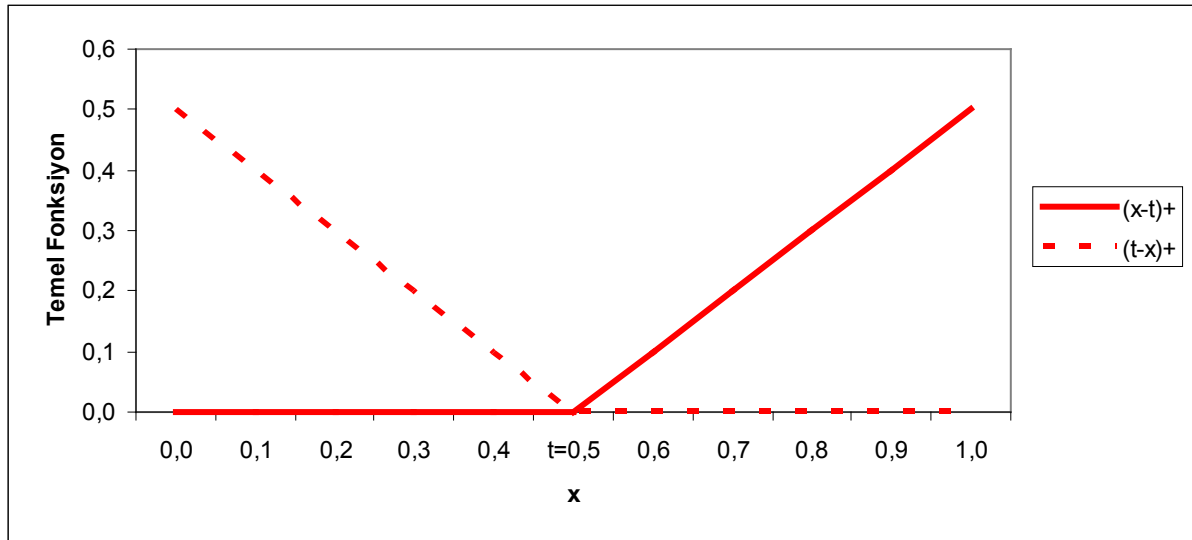
Şekil 2.2. $y=f(x)$ fonksiyonu için düğüm noktaları (Briand et al, 2004)

Temel fonksiyonlar ise x değişkeni parçalı doğrusal yapıda olacak şekilde türetilmektedir. Eş. 2.1. ve Eş. 2.2.'de $(x-t)_+$ ve $(t-x)_+$ formundaki parçalı doğrusal temel fonksiyonlar görülmektedir. (Buradaki “+”nın anlamı pozitif taraftır.)

$$(x-t)_+ = \begin{cases} x-t, & x > t \\ 0, & \text{ö.d.} \end{cases} \quad (2.1)$$

$$(t-x)_+ = \begin{cases} t-x, & x < t \\ 0, & \text{ö.d.} \end{cases} \quad (2.2)$$

Örnek olarak $(x-0.5)_+$ ve $(0.5-x)_+$ temel fonksiyonları Şekil 2.3’de gösterilmiştir.



Şekil 2.3. MARS tarafından kullanılan $(x-t)_+$ ve $(t-x)_+$ temel fonksiyonları (Hastie et al., 2001)

Şekil 2.3'deki her bir fonksiyon $t=0,5$ değerindeki düğüm noktasında parçalı doğrusaldır. Bunlar doğrusal uzanımlardır ve bu iki fonksiyon yansıma çifti (a reflected pair) olarak adlandırılır. Amaç her X_j için, her x_{ij} gözlem düğüm noktasında yansıma çiftleri oluşturmaktır. Dolayısıyla temel fonksiyonların toplamı,

$$C = \{(X_j - t)(t - X_j)\} \quad t \in \{x_1, x_2, \dots, x_{N_j}\} \quad j = 1, 2, \dots, p. \quad (2.3)$$

şeklinde olacaktır. Eğer bütün X değerleri farklı ise, $2Np$ adet temel fonksiyon olacaktır. Her temel fonksiyon sadece X_j 'ye bağlı olsa da (örneğin, $B(X) = (X_j - t)_+$, bu temel fonksiyon tüm R^p uzayındaki bir fonksiyondur.

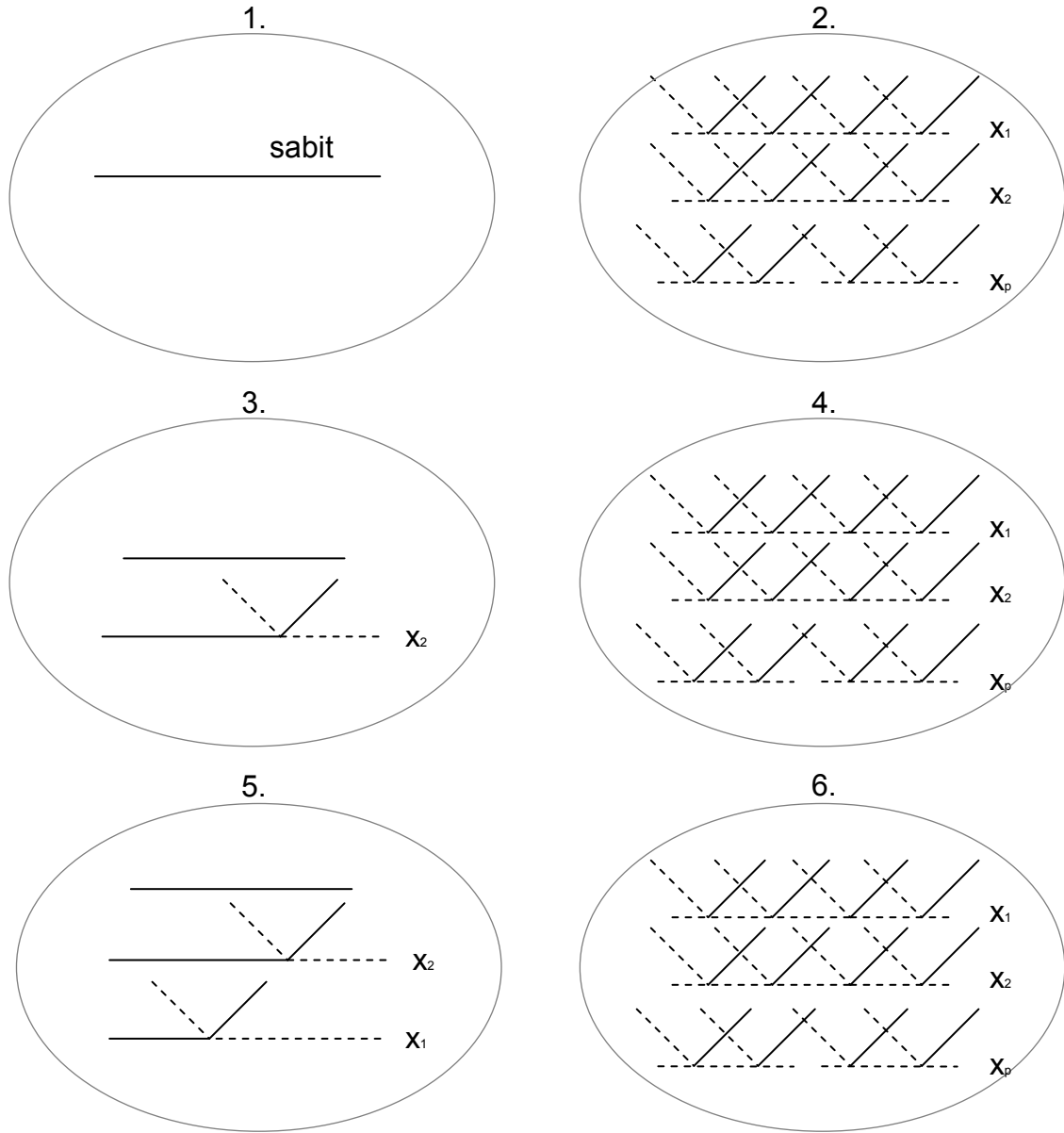
Model kurma yöntemi ileriye doğru adımsal doğrusal regresyon gibidir. Ancak orijinal X değişkenleri yerine burada C kümesindeki ve ondan türetilen kümelerdeki fonksiyonlar kullanılmaktadır. Dolayısıyla model Eş. 2.4'deki gibi olmaktadır.

$$f(x) = \beta_0 + \sum_{m=1}^M \beta_m B_m(X) \quad (2.4)$$

Eş. 2.4'deki $B_m(X)$ temel fonksiyonu C kümesindeki herhangi bir fonksiyon ya da onun gibi iki ya da daha fazla fonksiyonun ürünü olabilmektedir. β_m 'ler standart doğrusal regresyondaki gibi hata kareler toplamının enküçüklenmesi yöntemiyle tahmin edilmektedir.

Her bir bağımsız değişken X_j 'nin t düğüm noktasında gözlemlenmiş x_{ij} parçalı doğrusal temel fonksiyon ikilileri modele alınacak temel fonksiyon adaylarıdır. Model kurulumuna $B_m(X)=1$ sabit fonksiyonu ile başlanmaktadır. Diğer tüm temel fonksiyonlar modelin kurulmasında kendi içlerinde karşılaştırılacaklardır. Her adımda temel fonksiyon adayları (aday ikilileri) modeldeki diğer temel fonksiyonlarla da karşılaştırılmaktadır ve artık hatayı en çok azaltan, o anda modele alınmaktadır.

Şekil 2.4'de temel fonksiyonların modele nasıl alındığı ile ilgili bir şema görülmektedir. Her aşamada tüm B_m fonksiyon çiftleri yeni bir temel fonksiyon olarak C'deki yansıma çiftlerinden biri ile birlikte model kümesi M'de dikkate alınır (Hastie et al., 2001). Şekil 2.4'te görülen ve 1'den 6'ya kadar sıralanmış kümeler içerisinde yansıma çiftlerinin nasıl modele girdiği sırasıyla görülmektedir. 1. adımda sabit terim oluşturulmaktadır. 2. adımda ise tüm yansıma çiftleri araştırılır ve 3. adımda X_2 değişkenine ait bir yansıma çifti modele girmiştir. 4. adımda yeniden C kümesindeki tüm yansıma çiftleri araştırılır ve 5. adımda X_1 değişkenine ait yansıma çifti modele girmiştir ve süreç 6. adımla birlikte bu şekilde devam etmektedir.



Şekil 2.4. MARS ileriye doğru model kurma şeması (Hastie et al., 2001)

Düğüm noktalarının elde edilmesi de temel fonksiyonların işleyişine benzemektedir. Sürecin başında düğüm noktaları sayısı ve bu noktaların konumları bilinmemektedir. Adımsal süreç kullanılarak uzanım düğüm noktalarını yerleştirmek için en iyi noktalar araştırılmaktadır. Bağımsız değişkenin her değeri, olası bir düğüm noktası yeri olarak test edilmektedir. Sadece modele önemli bir etkisi olanlar bırakılmaktadır. Düğüm noktalarının belirlenmesinde de genelleştirilmiş çapraz geçerlilik kullanılmaktadır (Francis, 2000).

2.2.2.1. Temel Fonksiyon Algoritması

MARS yönteminde bağımsız değişkenlerin yerine, bu değişkenlerin temel fonksiyonları kullanılarak modelleme yapılmaktadır. Dolayısıyla model kurulurken amaç sadece model için en uygun katsayıları tahmin etmek değil, aynı zamanda elimizde bulunan veri setinden en iyi temel fonksiyonlar kümesini (alt bölümleri) elde etmektir.

MARS modeli, $x \in R_m$ olduğunda Eş. 2.5'de görüldüğü gibi olmaktadır(Friedman, 1991a).

$$f(X) = \sum_{m=1}^M a_m B_m(x) \quad (2.5)$$

Burada:

a_m : $\{a_m\}_1^M$ olarak tanımlı modeli en iyi kılacak katsayıları

R_m : $\{R_m\}_1^M$ olarak tanımlı, D uzayının bir bölümünü temsil eden ayırık kümeler (D uzayının alt bölümleri) olarak kabul edilmiştir.

Amacımız B_m temel fonksiyonlarını $B_m(x) = I[x \in R_m]$ elde etmektir. Burada I , gösterge(indicator) değişkenidir. I gösterge değişkeninin varsayımı(hipotezi) doğruysa bir, öteki durumlarda sıfır değerini almaktadır. Olumlu varsayımı gösteren adım fonksiyonu $H(\eta)$ 'dir ve Eş. 2.6'da görülmektedir.

$$H(\eta) = \begin{cases} 1 & \eta \geq 0 \\ 0 & \text{ö.d.} \end{cases} \quad (2.6)$$

LOF(g); $g(x)$ fonksiyonunun veri setine uyum eksikliğini hesaplayan bir süreçtir. Tüm bunlara göre ileriye doğru adımsal regresyon sürecini temsil eden, bir parçalama stratejisi yazılabilir. Bu parçalama süreci Algoritma 1'de görülmektedir.

Algoritma 1. (Friedman, 1991a)

```

 $B_1(x) \leftarrow 1$ 
For  $M-2$  to  $M_{\max}$  do:  $lof^* \leftarrow \infty$ 
For  $m = 1$  to  $M-1$  do:
For  $v = 1$  to  $n$  do:
For  $t \in \{x_{vj} | B_m(x_j) > 0\}$ 
 $g \leftarrow \sum_{i \neq m} a_i B_i(x) + a_m B_m(x) H[+(x_v - t)] + a_M B_M(x) H[-(x_v - t)]$ 
 $lof \leftarrow \min_{a_1, \dots, a_M} LOF(g)$ 
if  $lof < lof^*$  then  $lof^* \leftarrow lof$ ;  $m^* \leftarrow m$ ;  $v^* \leftarrow v$ ;  $t^* \leftarrow t$  end if
end for
end for
end for
 $B_M(x) \leftarrow B_{m^*}(x) H[-(x_{v^*} - t^*)]$ 
 $B_{m^*}(x) \leftarrow B_{m^*}(x) H[+(x_{v^*} - t^*)]$ 
end for
end algorithm

```

Algoritma 1'in birinci satırı, tanım aralığının ilk bölümünün oluşturulmasına eşittir. İlk For döngüsü ile parçalama süreci, tanım aralıklarının (temel fonksiyonların) en son sayısı $M_{\max}$ olana kadar tekrarlanmaktadır. Diğer görülen iç içe üç tane for döngüsü, temel fonksiyon B_{m^*} 'nin (modelde var olan), tahmin edici değişken x_{v^*} 'nin ve "parçalama noktası" t^* 'nin seçilmesi için bir optimizasyon gerçekleştirmektedir. Yani burada parçalama noktası t^* 'daki v^* değişkeninin R_{m^*} 'da benzer bölümlere parçalanmasının karşılığı görülmektedir. Temel fonksiyon kümesinin doğrusal regresyonu olan g fonksiyonunun uyum eksikliğini gösteren $LOF(g)$ fonksiyonu, 7. satırda minimize edilmektedir. (Yani artıkların karesi min. kılınacak şekilde katsayılar hesaplanmaktadır.) Algoritma 1 ile elde edilen temel fonksiyon Eş. 2.7.;

$$B_m(x) = \prod_{k=1}^{K_m} H[s_{km} \cdot (x_{v(k,m)} - t_{km})] \quad (2.7)$$

şeklinde yazılabilmektedir. Burada;

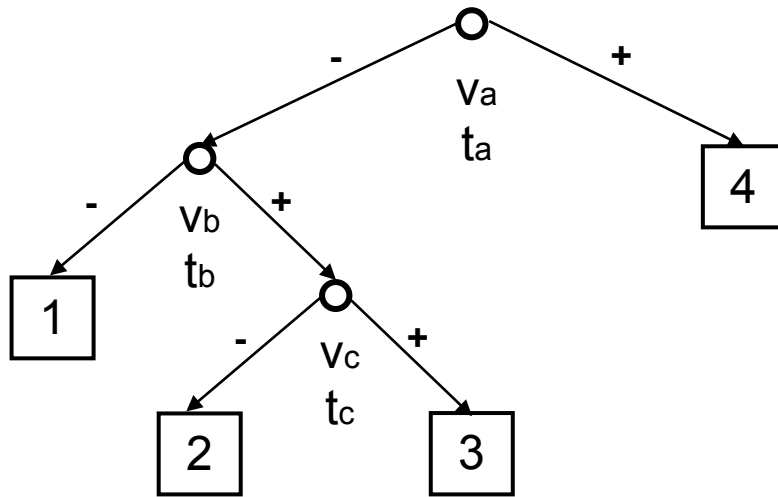
K_m : m. temel fonksiyonda çarpılan doğrusal fonksiyon sayısını,

s_{km} : ± 1 değerlerini(adım fonksiyonunun sağ tarafı mı yoksa sol tarafı mı anlayabilmek için yer almaktadır.),

$x_{v(k,m)}$: m. temel fonksiyondaki tahmin edici değişkenleri

t_{km} : Temel fonksiyonların parçalama noktasını(yöntemin en başında belirtilen düğüm noktasını) göstermektedir (Friedman,1991a).

İleriye doğru adımsal sürecin yapısı gereği tüm temel fonksiyonların oluşturulmasını göstermek için ikili ağaç diyagramı çizilebilir (Breiman et al., 1984). Bu ağaç diyagramını incelemek temel fonksiyonların oluşturulması mantığının kavranması için yararlı olacaktır.



Şekil 2.5. Temel fonksiyonların oluşturulmasının ikili ağaç diyagramı ile gösterilmesi (Friedman, 1991a).

Şekil 2.5’de görülen şematikten aşağıdaki temel fonksiyonlar yazılabilmektedir.

$$B_1 = H[-(x_{v_a} - t_a)]H[-(x_{v_b} - t_b)]$$

$$B_2 = H[-(x_{v_a} - t_a)]H[+(x_{v_b} - t_b)]H[-(x_{v_c} - t_c)]$$

$$B_3 = H[-(x_{v_a} - t_a)]H[+(x_{v_b} - t_b)]H[+(x_{v_c} - t_c)]$$

$$B_4 = H[+(x_{v_a} - t_a)]$$

2.2.2.2. Süreklilik

Düğüm noktalarında değişkenler için oluşturulan temel fonksiyonlar kullanılarak modelleme yapılmaktadır. Oluşturulan temel fonksiyon ikilileri (Eş. 2.1. ve Eş. 2.2.) geçiş noktalarında sürekli olmadığından, düğüm noktaları arasındaki sürekliliğin nasıl sağlanacağı problemi ortaya çıkmaktadır. Sürekliliğin sağlanması mümkündür. Algoritma 1'de ufak değişiklikler yapılarak sürekli modellerin oluşturulması mümkün kılınabilmektedir. Eş. 2.6'da görülen ve Algoritma 1'de kullanılan adım fonksiyonun yapısı gereği modelde süreklilik bulunmamaktadır. Eğer adım fonksiyonu görüldüğü her yerde (yani 6., 12. ve 13. satırlarda) kendisinin sürekli bir fonksiyonu ile yer değiştirilerek, Algoritma 1 sürekli modeller üretir bir duruma getirilir (Friedman,1991a).

Eş. 2.8'de q . derece uzanımları temsil eden tek taraflı kesikli güç (one-sided truncated power) temel fonksiyonlarının tanımlanması görülmektedir. Burada t , düğüm noktası; q , uzanım derecesi ve "+" da pozitif tarafı göstermektedir. $q>0$ için uzanım yaklaşımı sürekli ve $q-1$ sürekli türeve sahiptir (Friedman,1991a).

$$b_q(x-t) = (x-t)_+^q \quad (2.8)$$

İki taraflı kesikli güç temel fonksiyonları da Eş. 2.9.;

$$b_q^\pm(x-t) = [\pm(x-t)]_+^q \quad (2.9)$$

biçiminde tanımlanmaktadır.

Algoritma 1'deki adım fonksiyonları, $q=0$ uzanımları için iki taraflı kesikli güç temel fonksiyonlarına benzemektedirler. Dolayısıyla çok değişkenli uzanım temel fonksiyonları Eş. 2.10'daki gibi gösterilebilmektedir.

$$B_m^{(q)}(x) = \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]_+^q \quad (2.10)$$

Algoritma 1. aynı deęişken üzerinde, ikili ağaç diyagramında olduęu gibi çoklu parçalamaya izin vermektedir. Dolayısıyla model için belirlenen en son temel fonksiyonlar, çarpımlarında aynı deęişkenleri içeren bir çok faktöre sahip olabilmektedirler (Friedman,1991a).

2.3. Çok Deęişkenli Uyarlamalı Regresyon Uzanımları Yöntemi Kullanılarak MARS Modelinin Kurulması

Çok deęişkenli uyarlamalı regresyon uzanımları algoritması ileriye doęru ve geriye doęru olmak üzere iki adımdan oluşmaktadır.

- 1. Adım:** İleriye doęru adımda temel fonksiyonlar ve/veya bu fonksiyonların çarpımları en büyük modele ulaşılıncaya kadar eklenmektedir. Bir başka deyişle modelin karmaşıklığı maksimum seviyeye ulaşılıncaya kadar eklenen temel fonksiyonlarla model büyümektedir. Burada temel fonksiyonlar bağımsız deęişkenleri en uygun düğüm noktalarıyla aralıklara bölen parçalı doğrusal regresyon eğrileridir (Yerlikaya vd., 2008).
- 2. Adım:** Bu adımda MARS bir geriye doęru adım algoritması gerçekleştirir. Yani önemli bağımsız deęişkenler ve bu deęişkenlerin etkileşimleri belirlenerek, hata kareler toplamı minimum olacak şekilde oluşturulan optimum modele ulaşabilmek için temel fonksiyonlar modelden çıkarılmaktadır.

Aşağıda görülen 1. ve 2. algoritmalar MARS yöntemindeki ileriye ve geriye doęru adım süreçlerini göstermektedir.

Algoritma 2. MARS İleriye Doğru Adım Algoritması (Friedman, 1991a)

```

 $B_1(x) \leftarrow 1; M \leftarrow 2$ 
Loop until  $M > M_{\max} : lof^* \leftarrow \infty$ 
For  $m=1$  to  $M-1$  do:
  For  $v \notin \{v(k, m) | 1 \leq k \leq K_m\}$ 
    For  $t \in \{x_{vj} | B_m(x_j) > 0\}$ 
       $g \leftarrow \sum_{i=1}^{M-1} a_i B_i(x) + a_M B_m(x) [+(x_v - t)]_+ + a_{M+1} B_m(x) [-(x_v - t)]_+$ 
       $lof \leftarrow \min_{a_1, \dots, a_{m+1}} LOF(g)$ 
      if  $lof < lof^*$  then  $lof^* \leftarrow lof; m^* \leftarrow m; v^* \leftarrow v; t^* \leftarrow t$  end if
    end for
  end for
end for
 $B_M(x) \leftarrow B_{m^*}(x) [+(x_{v^*} - t^*)]_+$ 
 $B_{M+1}(x) \leftarrow B_{m^*}(x) [-(x_{v^*} - t^*)]_+$ 
 $M \leftarrow M + 2$ 
end loop
end algorithm

```

Algoritma 2 MARS stratejisinin ileriye doğru adımını göstermektedir. Bu algoritma, daha önce verilen Algoritma 1'in geliştirilmiş biçimidir. Algoritma 1'deki 6., 12. ve 13. satırlardaki adım fonksiyonlarının yerine kesikli güç temel fonksiyonları yazılmıştır. Ana temel fonksiyonlar 6. satırda modifiye edilmiş olan modele dahil olmuştur ve 12-14 satırlarının gereği güncelleştirilmiş modelde aynen kalmıştır. 4. satırda değişkenler boyunca kontrol döngüsü ile temel fonksiyon çarpımları farklı değişkenlerle ilgili faktörleri içermek durumunda bırakılmıştır (Eş. 2.7'de görüldüğü gibi). Bu algoritma ayrı marjinal veri kümesi değişkenlerinde bulunan düğüm noktalarıyla tüm temel fonksiyonların alt kümesi olan $M_{\max} q=1$ kesikli güç uzanım temel fonksiyonlarını üretmektedir. Temel fonksiyonların bu şekilde parçalanması adımından sonra geriye doğru silme adımı stratejisi temel fonksiyonların final serisini oluşturur. Bu yaklaşımla belirlenen düğüm noktaları kullanılarak final modele ulaşılır.

Algoritma 3. MARS Geriye Doğru Adım Algoritması (Friedman, 1991a)

$J^* = \{1, 2, \dots, M_{\max}\}; K^* \leftarrow J^*$
 $lof^* \leftarrow \min_{\{a_j | j \in J^*\}} LOF\left(\sum_{j \in J^*} a_j B_j(x)\right)$
For $M = M_{\max}$ to 2 do: $b \leftarrow \infty; L \leftarrow K^*$
For $m=2$ to M do: $K \leftarrow L - \{m\}$
 $lof \leftarrow \min_{\{a_k | k \in K\}} LOF\left(\sum_{k \in K} a_k B_k(x)\right)$
if $lof < b$ then $b \leftarrow lof; K^* \leftarrow K$ end if
if $lof < lof^*$ then $lof^* \leftarrow lof; J^* \leftarrow K$ end if
end for
end for
end Algorithm

Algoritma 3'ün birinci satırında, Algoritma 2'den elde edilmiş J^* bütün temel fonksiyon kümesini kapsayan model yer almaktadır. Algoritma 3'de dıştaki For döngüsünün her tekrarında bir temel fonksiyon silinmektedir. İçteki for döngüsü de silinen bu temel fonksiyonu seçmektedir. Silinmesiyle model uyumunda en çok artışı sağlayan ya da en az azalışa sebep olanlar silinir. $B_1(x)=1$ sabit fonksiyonu silim işleminde seçilebilir değildir. Algoritma 3, $M_{\max}-1$ model serisi oluşturur ve her bir model kendisinden bir önceki modelden en az bir temel fonksiyon eksikliğine sahiptir (Friedman, 1991a).

2.3.1. Varyans analizi ayrıştırması

Uygulanan Algoritma 2 ve Algoritma 3'ün neticesinde ulaşılan MARS modeli,

$$f(x) = a_0 + \sum_{m=1}^M a_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]_+ \quad (2.11)$$

şeklinde yazılabilmektedir.

Algoritma 2 ile tüm temel fonksiyonlar üretilmektedir ve modelin bütünü görülebilmektedir. Ancak modelin bu hali, yaklaşımın doğasını daha iyi

anlayabilmek için yeterli değildir. Eş. 2.11' deki terimler düzenlenerek, y ve x arasındaki tahmini bir ilişki hakkında önemli bir bilgiyi ortaya çıkaran bir model yeniden şekillendirilebilir. Yani Eş. 2.11'deki model Eş. 2.12'deki gibi şekillendirilmiştir:

$$f(x) = a_0 + \sum_{K_m=1} f_i(x_i) + \sum_{K_m=2} f_{ij}(x_i, x_j) + \sum_{K_m=3} f_{ijk}(x_i, x_j, x_k) + \dots \quad (2.12)$$

Eş. 2.12'deki modelde yer alan ilk terim sabit terimdir. Modelde görülen 2. terim, tek değişken içeren tüm temel fonksiyonların toplamı olarak yer almıştır. Modeldeki 3. terim, değişken ikilileri ve varsa bu değişkenlerin etkileşimlerini gösteren tüm temel fonksiyonlar üzerinden alınmış toplam olarak tanımlanmaktadır. Benzer şekilde, modeldeki 4. terim üç değişkenli etkileşimleri temsil eden temel fonksiyonları içermektedir ve modelin yapısı bu şekilde devam etmektedir.

Tek tek bu toplamlar ele alınırsa; $V(m) = \{v(k, m)\}_1^{K_m}$ 'nin Eş. 2.11'deki B_m m. temel fonksiyon ile birleştirilmiş değişken kümesi olması varsayımıyla, Eş. 2.12'deki ilk toplamda yer alan her bir fonksiyon Eş. 2.13'de görüldüğü gibi gösterilebilmektedir. Eş. 2.13'de sadece x_i 'yi içeren tüm tek değişkenli temel fonksiyonların toplamı görülmektedir ve bu da tek değişkenli fonksiyonun $q=1$ uzanım gösterimidir.

$$f_i(x) = \sum_{\substack{K_m=1 \\ i \in V(m)}} a_m B_m(x) \quad (2.13)$$

Her biri iki değişkenli fonksiyon olan Eş. 2.12'deki ikinci toplam da aşağıdaki gibi olmaktadır. Eş. 2.14, belirli x_i ve x_j değişken ikililerini içeren iki değişkenli temel fonksiyonların toplamıdır.

$$f_{ij}(x_i, x_j) = \sum_{\substack{K_m=2 \\ (i,j) \in V(m)}} a_m B_m(x_i, x_j) \quad (2.14)$$

Eş. 2.14'e tek değişkenlilerin etkileri Eş. 2.13 eklendiğinde, x_i ve x_j 'nin modele birlikte katkısını gösteren iki değişkenli vektör(tensor) çarpım uzanım yaklaşımı olan $q=1$ 'i vermektedir ve Eş. 2.15' de görüldüğü gibi yazılabilmektedir.

$$f_{ij}(x_i, x_j) = f_i(x_i) + f_j(x_j) + f_{ij}(x_i, x_j) \quad (2.15)$$

Benzer şekilde, belirli üçlü değişkeni içeren tüm temel fonksiyonları bir araya getirerek sağlanabilen üçüncü toplamdaki her bir fonksiyon da Eş. 2.16'da görülmektedir.

$$f_{ijk}(x_i, x_j, x_k) = \sum_{\substack{K_m=3 \\ (i,j,k) \in V(m)}} a_m B_m(x_i, x_j, x_k) \quad (2.16.)$$

x_i , x_j ve x_k 'yı içeren tek ve iki değişkenli temel fonksiyonlar toplamına, üç değişkenli olan da sağlanarak modele eklenmektedir. Daha çok değişken içeren terimler de hep birlikte toplanabilir ve benzer şekilde gösterilebilir. Çapraz tablolar için varyans analizi (analysis of variance for contingency tables) ile elde edilen ayrıştırmalara(decomposition) benzerliği sebebiyle, Eş. 2.12 MARS modelinin ANOVA ayrıştırması(decomposition) olarak adlandırılmaktadır.

MARS modelinin yorumları, kendi ANOVA ayrıştırması ile oldukça kolaylaştırılmıştır. Eş. 2.12'deki gösterim ile modele giren tek değişkenlerin kendi başlarına katkılarını ya da diğer değişkenlerle etkileşimlerini içeren bir şekilde modele girmişlerse bu etkileşimleri ve etkileşim derceleriyle katkıları açıklanabilmektedir. Yorum, ANOVA ayrıştırmasının grafik şeklinde gösterilmesiyle geliştirilmiştir. Eklemeli terimler Eş. 2.13, $f_i(x_i)$ 'nin x_i karşısındaki saçılım grafiği ile yorumlanabilir. Benzer şekilde iki değişkenli durumlarda $f_{ij}(x_i, x_j)$ 'nin x_i ve x_j karşısındaki saçılım grafikleri elde edilerek yorum yapılabilir. Yüksek dereceli etkileşimlerin olduğu modeller diğer değişkenlerin çok sayıda değerleri için değişken çiftleri üzerinde saçılım grafikleriyle görüntülenebilmektedir (Friedman, 1991a).

2.3.2. Model seçimi

Çok değişkenli uyarlamalı regresyon uzanımları yönteminde model seçimi için Craven ve Wabha (1979) tarafından geliştirilmiş “genelleştirilmiş çapraz geçerlilik” (generalized crossvalidation/ GCV) kriteri temel alınmaktadır. MARS hangi temel fonksiyonların modele alınacağını belirleyebilmek için GCV’yi kullanmaktadır ve GCV, ortalama kare hatanın model karmaşıklığına bağlı olan cezaya (penalty) bölümü olarak tanımlanabilmektedir (Xu et al., 2004). GCV kriteri,

$$LOF(f_M) = GCV(M) = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - f_M(x_i))^2}{\left(1 - \frac{C(M)}{n}\right)^2} \quad (2.17)$$

şeklinde gösterilebilmektedir.

Algoritma 2 ve 3’deki LOF fonksiyonu GCV kriterine dayanmaktadır. Eş. 2.17’deki f_M fonksiyonu, temel fonksiyonlara dayanan MARS tahmin modelidir. y bağımlı değişkendir ve n veri setindeki gözlem sayısını göstermektedir. $C(M)$, modele giren temel fonksiyonların sayısı ile artan karmaşıklık cezasıdır ve Eş. 2.18,

$$C(M) = (M + 1) + d \cdot M \quad (2.18)$$

biçiminde tanımlanmaktadır. Burada M modele giren temel fonksiyon sayısıdır. d ise; model içerisine eklenen her temel fonksiyon için cezadır. Aynı zamanda düzleştirme parametresi (smoothing parameter) olarak da adlandırılabilir. d ’nin büyük değerleri daha az temel fonksiyonlara götürür ve böylece daha düz fonksiyonlar tahmin edilmektedir (Xu et al., 2004).

Yapılan tüm çalışmalarda en iyi d değerinin $2 \leq d \leq 4$ aralığında olduğu görülmüştür (Friedman, 1991a). Bazı matematiksel ve simülasyon sonuçları, parçalı doğrusal regresyonda nokta seçimi için üç parametrenin bedelinin ödenmesi gerektiğini önermektedir. Dolayısıyla $d=3$ olarak belirlendiğinde Eş.

2.18; $M+1$ adet doğrusal bağımsız temel fonksiyona sahip olan bir modelde, ileriye doğru adımda M adet düğüm noktası seçilmiştir diye yorumlanabilir. Modelin sınırlı eklemeli olduğu durumlarda $c=2$ cezası kullanılabilmektedir. Tüm bunlar kullanılarak geriye doğru adımda GCV'yi minimum kılan model seçilmektedir (Hastie et al., 2001).

MARS modelleri kurulurken GCV kriteri temel fonksiyon sayısının belirlenmesinde şöyle kullanılmaktadır; MARS algoritmasıyla oluşturulan modeller arasında GCV değerinin minimum olduğu model tercih edilir. Bir başka deyişle, GCV'nin minimum olduğu değerdeki temel fonksiyon sayısı kadar modele temel fonksiyon girmiş olur.

Model kurulurken ileriye doğru adımda olası tüm temel fonksiyonlar ve düğüm noktaları belirlenmektedir. İşte geriye doğru adımda modele girecek düğüm noktaları da yine GCV kriterini minimum yapacak şekilde seçilmektedirler.

MARS algoritmasının ana adımları Xu et al.(2004) tarafından aşağıdaki gibi özetlenmiştir:

0. Modelin hesaplanabilen en fazla karmaşıklığı seçilir ve d parametresi tanımlanır.

İleriye doğru adımsal seçim:

1. En basit modelle, sadece sabit katsayı ile, başlanır,.
2. Her açıklayıcı değişken için temel fonksiyonların kümesi araştırılır.
3. Tahmin hatasını minimize eden temel fonksiyonlar ikilileri belirlenir ve modele dahil edilir.
4. Modelin önceden belirlenen tanımlanmış karmaşıklığa ulaşana kadar adım 2'ye gidilir.

Geriye doğru silme adımı:

5. Tüm temel fonksiyonların kümesi (sabit hariç) araştırılır ve GCV kriteri kullanılarak modele en az katkıda bulunan temel fonksiyonlar silinir.
6. GCV kendi minimumuna ulaşana kadar 5. adım tekrarlanır.

2.3.3. Çok deęişkenli uyarlamalı regresyon uzanımları yönteminde kayıp deęerlerin incelenmesi

Büyük veri setleriyle çalışırken sıkça kayıp deęerlere rastlanılabilmektedir. Veri setinde herhangi bir deęişkendeki herhangi bir verinin kayıp olması durumunda o kaydın silinmesi ya da kayıp deęişkenin o deęeri yerine deęişken ortalaması kullanılması gibi yöntemlerle istatistiksel paket programlarda uygulamalar kolaylıkla yapılabilmektedir. Kayıp veriler için geliştirilmiş yöntemler olsa da istatistiksel yazılım paketlerinde uygulamak her zaman uygun ve kolay deęildir. Kayıp veri ile birlikte ortaya çıkan iki önemli problem Francis(2000, s15) tarafından aşığıdaki gibi sıralanmıştır.

1. Bir çok istatistiksel paket programda herhangi bir deęişkendeki kayıp deęerli bir kaydın silinmesi nedeniyle bir çok verinin analizde kaybolma problemi mevcuttur.
2. Analizin doğru olabilmesi için bağımlı ve bağımsız deęişkenlerin, mevcut kayıp deęerlerden bağımsız olmasının saęlanması gerekmektedir.

MARS'ın kayıp veri problemini ele alış şekli; kayıp verisi olan deęişken için bir temel fonksiyon oluşturmaktır. Oluşturulan temel fonksiyon, deęişkende kayıp veri varsa "1" deęerini, yoksa "0" deęerini alır. Daha sonraki adımda araştırma süreci yardımıyla veri setindeki kayıp veri temel fonksiyonları ve dięer deęişkenler arasında bir etkileşimin varlığı ve bu etkileşimin bağımlı deęişkenin tahmin edilmesinde istatistiksel olarak önemli olup olmadığı araştırılır. Bu ilişkilerin belirlenmesiyle ortaya çıkan temel fonksiyonlar aracılığıyla, dięer deęişkenler kayıp deęişkenler için bir vekil gibi rol yapabilmektedirler (Francis, 2000).

2.4. Deneysel Bayes Yöntemiyle Temel Fonksiyon ve Düşüm Noktası Seçimi

MARS yönteminin dięer modelleme tekniklerine göre sahip olduęu en önemli farklılık temel fonksiyon ve düşüm noktalarını kullanmasıdır. Sakamoto(2007), Bayesci yaklaşımın avantajlarını kullanarak deneysel Bayes yöntemi ile MARS'taki temel fonksiyonlar ve düşüm noktaları seçimine alternatif sunmuştur. Temel fonksiyonların seçimi ve düşüm noktalarının pozisyonlarının seçimi için marjinal

olabilirlik Laplace yaklaşımını ve temel fonksiyon sayılarının belirlenmesi için de Akaike Bayes bilgi kriterini(Akaike Bayes information criterion, AIBC) kullanmıştır.

MARS temel fonksiyonların eklenmesi ve daha sonraki adımda da modelden çıkartılması için genelleştirilmiş çapraz geçerlilik kriterini(GCV) kabul etmektedir. Bazı uygulamalarda çok fazla sayıda temel fonksiyon seçildiğinden model yapısını yorumlamak kolay olmayabilir. Ayrıca, kullanılan kriter tahmin edici değişkenler uzayındaki uzaklıklar kullanılarak elde edildiğinden model yapısının araştırılması için uygun olmayabilir düşüncesinden yola çıkarak başka bir kriterin araştırılması gerektiği düşünülmüştür.

Belirlenen temel fonksiyonların temel katsayılarının tahmin edilmesi için karmaşıklık cezası olarak gösterilen temel katsayıların önseli ile tanımlanan ceza olabilirlik yaklaşımı kullanılmaktadır. Bayes yaklaşımının gereği olarak kullanılan önseller(priors), temel fonksiyon sayılarına, her temel fonksiyondaki değişkenlere, düğüm noktalarının konumlarına, ana katsayılar dayanmaktadır ve sonra tahmin edilen fonksiyonun sonsal(posterior) örnekleme Tersinir Sıçramalı Markov Zinciri Monte Carlo(reversible jump Markov Chain Monte Carlo, RJMCMC) algoritmasıyla yapılmaktadır(Green, 1995).

Eş. 2.19'da $B(x), k = 1, \dots, K$ olan temel fonksiyonların doğrusal kombinasyonları olarak tanımlanmış f regresyon fonksiyonu görülmektedir. Burada katsayılar $a_0, a_1, \dots, a_K$ tahmin edilecektir.

$$f(x) = a_0 + \sum_{k=1}^K a_k B(x) \quad (2.19)$$

Eş. 2.19'daki her bir temel fonksiyonun güç temel fonksiyonları olarak nasıl tanımlandığı Eş. 2.20'de görülmektedir.

$$B_k(x) = \prod_{l=1}^{L_k} [s_{kl}(x_{v(k,l)} - c_{kl})]_+^m \quad (2.20)$$

Önceki bölümlerde belirtildiği gibi güç temel fonksiyonlarının yapısı gereği $[x]^+$, x 'in pozitif taraflarını ve m tam sayısı uzanımlar derecesini göstermektedir. K , temel fonksiyonlar sayısı, L_k etkileşim derecesi, $s_{kl} + 1$ ya da -1 , temel fonksiyonlara giren değişken sayısı $v(k,l)$ ve düğüm noktasının pozisyonu c_{kl} olarak verilerden temel fonksiyon algoritması bölümünde belirtildiği gibi uyum eksikliği kriteri ile seçilmiştir.

Bayesci yaklaşımla temel fonksiyon seçimini anlayabilmek için $B_k(x)$ temel fonksiyonları varlığında a katsayısındaki normal önsel Eş. 2.21'deki gibi olmaktadır.

$$p(a|B_K, \lambda) = (\lambda / 2\pi)^{(K+1)/2} \exp\left(-\frac{\lambda}{2} a^T a\right) \quad (2.21)$$

Burada λ pozitif sayı fonksiyonun karmaşıklığını kontrol eden yüksek(hyper) parametredir. Ayrıca tanımlanan önselin bazı Bayesci yaklaşımlarda kabul edildiği bilinmektedir (örneğin Denison et al., 2002).

Bayes teoreminden yola çıkıldığında a 'nın sonsal yoğunluğu

$$p(a|y, B_K, \sigma^2, \lambda) \propto p(y|B_K, a, \sigma^2) p(a|B_K, \lambda) \quad (2.22)$$

şeklinde olacaktır. Burada görülen $p(y|B_K, a, \sigma^2)$, verilen B_K ve a için y 'nin yoğunluğudur ve y dağılımının varyansı σ^2 olmaktadır. Dolayısıyla hata varyansı σ^2 ile normal dağılan y 'nin varlığında, a 'nın sonsal dağılımı Eş. 2.23'deki ortalama ve hata varyansı σ^2 ile normal olacaktır.

$$\hat{a} = (B_K^T B_K + \lambda * I)^{-1} B_K^T y \quad (2.23)$$

Burada $\lambda^* = \lambda \sigma^2$ ve I birim matrisdir. Buna göre varyans matrisi $\sigma^2 (B_K^T B_K + \lambda^* I)^{-1}$ şeklinde tanımlanacaktır.

a 'nın sonsal yoğunluğu olan $\hat{a}$, ceza log-olabilirliğin (log-likelihood) enbüyüklenmesine eşit olmaktadır ve Eş. 2.24'te görülmektedir:

$$lp(a|y, B_K, \sigma^2, \lambda) = \log p(y|B_K, a, \sigma^2) - \frac{\lambda}{2} a^T a + \frac{K+1}{2} \log \lambda + \text{sabit} \quad (2.24)$$

Normal dağılımın varlığında maksimum ceza olabilirlik doğrudan hesaplanabilir. Dağılımlarda normalliğin olmadığı durumlarda ise maksimum olabilirliğin tekrarlanması yoluna dayanan Fisher skorumaya yöntemi gibi algoritmalarla elde edilebilmektedir (örnekler McCullagh and Nelder, 1989; Green and Silverman, 1994).

$c = \{c_{kl}\}$ düğüm noktası pozisyonunun, σ^2 varyansının ve λ yüksek parametresinin tahmin edilebilmesi için uyum eksikliği kriteri olarak marjinal olabilirliği dikkate alınmaktadır. a 'nın sonsal yoğunluğundan integrallenmiş marjinal olabilirliği Eş. 2.25'deki gibi yazılabilmektedir. Marjinal olabilirlik fonksiyonunun enbüyüklenmesi yöntemi, Bayes yöntemindeki önsellerde bilgi olmadığında (non-informative priors) σ^2 ve λ 'nın varsayımları gibi yapılacaktır. Eş. 2.25'de görülen integralin doğrudan hesaplanması normal dağılım dışındaki durumlarda mümkün olamamaktadır.

$$p(y|B_K, \sigma^2, \lambda) = \int p(y|B_K, a, \sigma^2) p(a|B_K, \lambda) da = \int \exp\{l_p(a|y, B_K, \sigma^2, \lambda)\} da. \quad (2.25)$$

Laplace yaklaşımı kullanılarak (Tierney and Kadane, 1986; Davison, 1986), ki buradaki yaklaşım ceza log-olabilirliğin Taylor açılımı $\tilde{a}$ 'da maksimum olmaktadır, marjinal log-olabilirliğin yaklaşık hesaplaması

$$l_M(c, \sigma^2, \lambda|y) = \log p(y|B_K, \sigma^2, \lambda) \approx lp\left(\hat{a}|y, B_K, \sigma^2, \lambda - \frac{1}{2} \log |Hp(\hat{a})| + \text{sabit}.\right) \quad (2.26)$$

şeklinde elde edilmektedir.

Burada $Hp(\hat{a})$, ceza log-olabilirliğin negatif Hessian matrisi olarak tanımlanmaktadır ve

$$Hp(\hat{a}) = \left(-\frac{\partial^2 lp}{\partial a \partial a^T} \right)_{\hat{a}} \quad (2.27)$$

biçiminde verilir. Normal yaklaşım olduğunda Eş. 2.26'daki eşitlik, $Hp(\hat{a}) = \sigma^{-2} (B_K^T B_K + \lambda * I)$ şeklinde olacaktır. Benzer bir tanım da normal olmayan bir durum için elde edilir.

Düğüm noktaları pozisyonu $c = \{c_{kl}\}$, σ^2 varyans ve λ yüksek parametre elemanları, $\{v(k, l)\}$ değişkenlerini içeren temel fonksiyonlar ve hatta temel fonksiyon kombinasyonları hakkındaki yaklaşık marjinal log-olabilirlik Eş. 2.26'da maksimize edilmektedir. Sakamoto(2007), temel fonksiyon sayısı K arttıkça, marjinal log-olabilirliğin(Eş. 2.26.) de artmasına rağmen, ileriye doğru adım aşamasında Akaike Bayes Bilgi Kriterini(ABIC) minimize eden K ve temel fonksiyon kombinasyonlarının seçilmesi düşüncesini savunmuştur. Bu kriter Eş. 2.28;

$$ABIC = -2l_M(c, \sigma^2, \lambda | y) + 2(q + 2) \quad (2.28)$$

şeklinde tanımlanmaktadır(Akaike, 1980). Bu eşitlikte q, temel fonksiyonlardaki $c = \{c_{kl}\}$ düğüm noktalarının sayısıdır. Geriye doğru adım aşamasında, c düğüm noktaları tahmin edildiğinden, sadece marjinal log-olabilirlik $l_M(c, \sigma^2, \lambda | y)$ maksimize edilmektedir.

ÜÇÜNCÜ BÖLÜM

3. UYGULAMA

Finansal kurumlar ellerindeki kaynakları en verimli ve karlı bir şekilde değerlendirmek istemektedirler. Öte yandan finansal sistemin devamlılığı ve ekonominin daha güçlü olabilmesi için kredilerin önemi büyüktür. Doğru kişilere kullandırılan ve böylece geri ödenebilmesi sağlanan krediler finansal sistemin devamlılığı için gereklidir. Dolayısıyla kredilerin verilmesi bir finansal probleme dönüştüğünden, bu problemin en doğru şekilde çözülerek, kredi kararının verilmesi gerekmektedir. İşte bu yüzden finansal kurumlar, kredilerin değerlendirilmesinde belirli stratejilere göre bir karar mekanizması oluşturmaktadırlar. Bu çalışmadaki amaç finansal kurumların karar mekanizmasına benzer bir kredi modellemesi yapılmasıdır.

Çalışmanın bu bölümünde şimdiye kadar açıklanan çok değişkenli uyarlamalı regresyon uzanımları yönteminin gerçek bir veri üzerinde uygulanabilirliğinin gösterilebilmesi için Türkiye'deki bir finans kurumundan elde edilen bireysel kredi değerlendirme verileri üzerinde çalışılacaktır. Gerçek veri uygulamasından önce MARS yazılımı ve kullanımı özetlenecek, çalışmada kullanılan veriler ve özelliklerinden bahsedilecektir. Uygulamanın 3. bölümünde bağımlı değişkenin kredi kararı olduğu durumda bir MARS modellemesi yapılacak ve yönteminin uygulanabilirliği gösterilmeye çalışılacaktır. Kredi kararının onay / ret olması durumuna göre kurulacak modelin yorumları yine bu bölümde verilecektir. Uygulamanın 4. bölümünde ise Lojistik Regresyon yöntemi ile çözümleme yapılacak ve sonuçlar kısmında MARS yöntemiyle karşılaştırılacaktır. MARS modellemesi yöntemi için MARS 3.0 yazılımı, Lojistik Regresyon yöntemi için ise SPSS 15.0 yazılımı kullanılacaktır.

3.1. MARS Yazılımı

Çalışmanın bu bölümünde MARS yazılımı kullanılarak modellemenin nasıl gerçekleştirileceğinin gösterilmesinin yanı sıra MARS yazılımında verilerin açılması, elde edilebilen grafiklerin gösterilmesi, bu yazılım ile yapılabileceklerin özetlenmesi de amaçlanmıştır. Çözümlemede kullanılacak olan MARS yazılımı Salford Systems web adresinden temin edilmiştir(<http://www.salfordsystems.com>).

Kullanıcının isteğine bağlı olarak, MARS modellerinin kurulması, bu yöntem kullanılarak verilerin istenilen şekilde düzenlenmesi, verilerden istenilen bilgilerin elde edilebilmesi için kullanılabilecek olan MARS yazılımı Salford System tarafından yapılmıştır. MARS yazılımı ile hızlı bir şekilde model kurulumu ve geliştirilmesi yapılabilmektedir. Yazılımla yapılabilecekler aşağıdaki gibi özetlenmiştir (Salford Systems, 2001):

- Tahmin edilecek değişkenlerden hangilerinin araştırılan konuyla ilgili, hangilerinin ilgisiz olduğunun belirlenmesi,
- Bağımlı değişken ile doğrusal olmayan bir ilişki sergileyen tahmin edici değişkenlerin dönüşümü,
- Tahmin edici değişkenler arasındaki etkileşimlerin tespit edilmesi,
- Yeni değişken teknikleri kullanılarak kayıp değerlerin ele alınması,
- Aşırı doymuşluğa (Overfitting) karşı koruma için gelişmiş bireysel-testler kullanılması gerçekleştirilebilmektedir.

MARS yazılımı her seviyeden teknik bilgiye sahip kullanıcılar için kullanması oldukça kolay bir programdır. Süratli bir şekilde tüm olası modeller araştırılabilmekte ve hızlıca en optimal çözüme ulaşılabilmektedir.

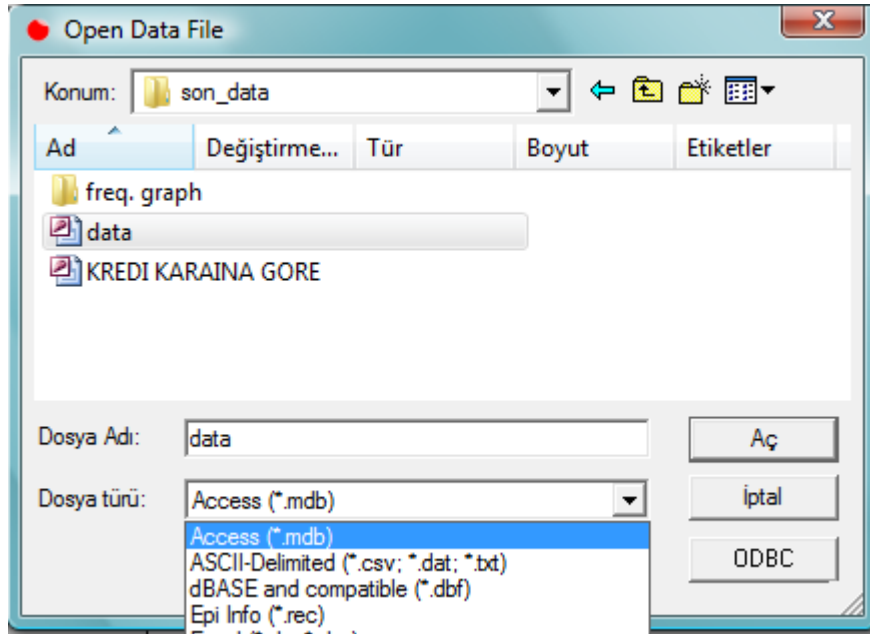
Windows işletim sisteminde tüm görsel özellikleri ile birlikte rahatlıkla kullanılabilmektedir. UNIX işletim sistemli bilgisayarlarda ise tüm görsel grafik sonuçlarını açabilmek için MARS model görüntüleyicisi (MARS model viewer) kullanılabilmektedir.

MARS yazılımı ile BASIC programlama dili kullanılabilir. Dolayısıyla yeni değişkenler tanımlayabilme, varolan değişkenlerin modifiye edilmesi, matematiksel, istatistiksel ve olasılıksal dağılımların girilebilmesi, tanımlanabilmesi mümkün olabilmektedir.

Minimum sistem özellikleri olan bir bilgisayarda çalıştırılabilir. Elbette veri setindeki kayıt sayısına bağlı olarak hafızada gerekli boş alan artabilmektedir.

Verilerin Hazırlanması

MARS yazılımında veri girişinin yapılmasına gerek duyulmamakta ve bir çok formatta(Systat, Excel, SAS, ASCII, vb.) veri dosyası bu yazılımda açılabilir. Bu çalışma için elde edilen veriler Microsoft Access dosyası formatındadır ve Şekil 3.1’de görüldüğü gibi **File** menüsünden **Open** seçeneği seçilerek ve dosya türü Access olarak belirlenerek verilere ulaşılabilmektedir. Şekil 3.2’de MARS yazılımında verilerin görünümü bulunmaktadır.



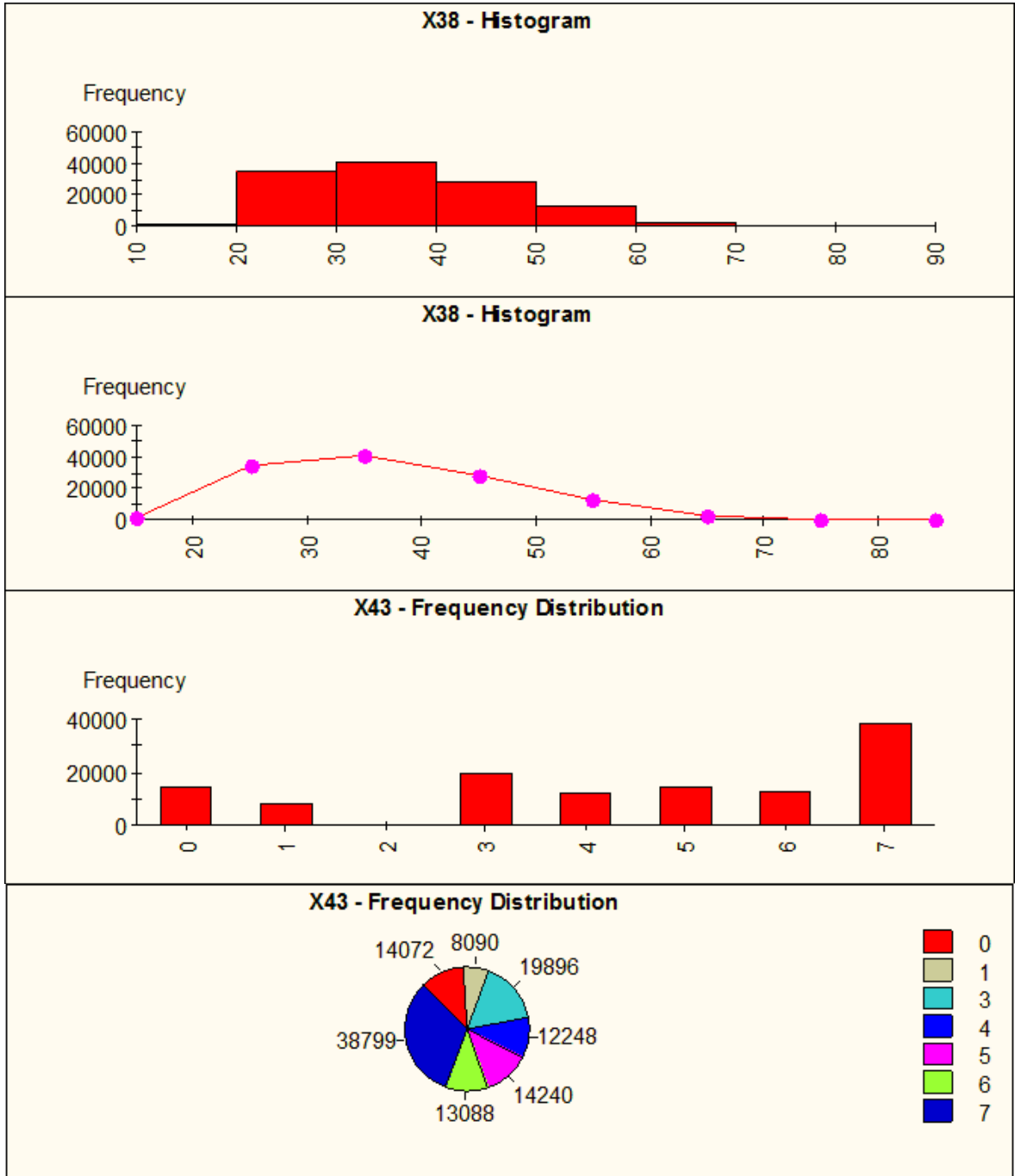
Şekil 3.1. MARS yazılımı veri açma penceresi

Metin şeklinde tanımlanmış değişkenler bu yazılımda direkt algılanmakta ve bu değişkenler için sayısal tanımlama yapılmasına gerek duyulmamaktadır.

	X1	X2\$	X3\$	X4	X5	X6\$	X7	X8\$	X9	X10\$	X11	X12\$
1	1		D	2.00812E+007	6 Y		1 M		0 Y		103 Y	
2	2		D	2.00812E+007	6 Y		1 A		0 Y		144 Y	
3	3		D	2.00812E+007	6 Y		1 A		0 Y		11 N	
4	4		D	2.00812E+007	6 Y		1 A		0 Y		23 N	
5	5		D	2.00812E+007	6 Y		1 M		0 Y		83 N	
6	6		D	2.00812E+007	6 Y		1 A		0 Y		2 Y	
7	7		D	2.00812E+007	1 Y		1 M		0 Y		124 Y	
8	8		D	2.00812E+007	6 N		1 M		0 Y		11 Y	
9	9		D	2.00812E+007	6 Y		1 A		1 Y		124 Y	
10	10		D	2.00812E+007	11 N		1 M		0 Y		169 N	
11	11		D	2.00812E+007	14 Y		1 M		0 Y		33 N	
12	12		D	2.00812E+007	6 Y		1 K		1 Y		109 N	
13	13		D	2.00812E+007	11 Y		1 M		1 Y		110 Y	
14	14		D	2.00812E+007	6 Y		1 K		0 Y		45 Y	
15	15		D	2.00812E+007	6 Y		1 K		0 Y		16 Y	
16	16		D	2.00812E+007	6 Y		1 M		0 Y		60 Y	
17	17		D	2.00812E+007	11 Y		1 M		0 Y		182 Y	
18	18		D	2.00812E+007	5 Y		1 M		0 Y		184 Y	
19	19		D	2.00812E+007	6 Y		1 A		0 Y		109 Y	
20	20		D	2.00812E+007	6 Y		1 M		0 Y		207 Y	

Şekil 3.2. MARS yazılımında verilerin görünüşü

Verilerin sıklık grafiklerine ulaşabilmek için yazılımda **Explore** menüsünden **Frequency Distribution** seçeneği seçilir ve hangi değişkenlerin grafiklerinin oluşturulması isteniyorsa o değişkenler seçilerek **Generate Charts** butonu tıklanır. Değişkenleri detaylı inceleyebilmek için grafiklerin yer aldığı ekrandaki seçenekler kullanılarak grafik seçenekleri değiştirilebilir. Bar, nokta, çizgi, pasta grafikleri elde edilebilmektedir. Ayrıca elde edilen grafikler MARS yazılımından kolaylıkla başka formatlara çevrilerek oluşturulacak raporlarda kolaylıkla kullanılabilme özelliğine sahiptirler. Şekil 3.3’de MARS yazılımı ile elde edilebilecek grafik çeşitlerinin örnekleri yer almaktadır, çalışmadaki tüm değişkenlerin grafikleri ise Ek2’de verilmiştir.

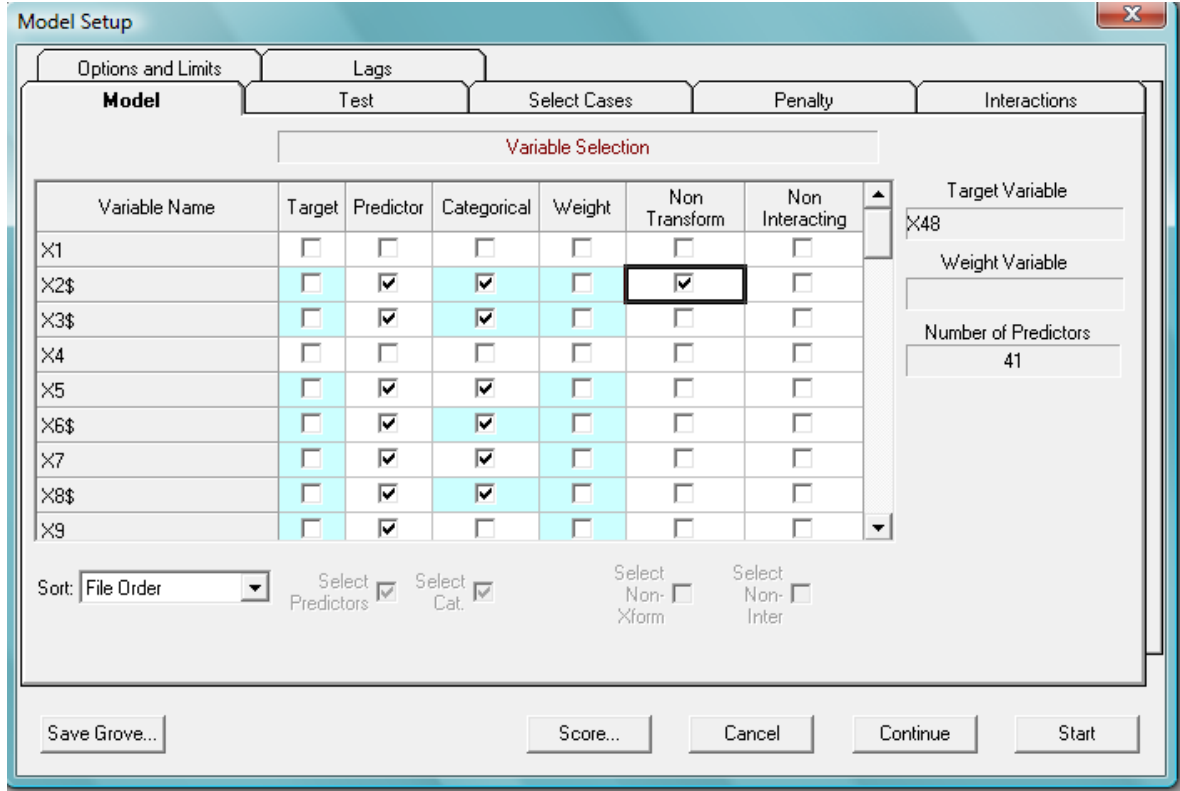


Şekil 3.3. MARS yazılımından elde edilebilecek grafik çeşitleri

Model Kurulumu

Model kurulumu aşamasında **Model** menüsünden **Construct Model** seçeneği ile **Model Setup** ekranına girilir. Bu ekranda **Model** kısmında sırasıyla değişkenler yer almaktadır. Bağımlı ve bağımsız değişkenlerin belirlenebildiği bu ekranda aynı zamanda kategorik değişkenler de seçilerek tanımlanabilmektedir. Bağımlı değişken için **Target**, bağımsız değişkenler için **Predictor** ve kategorik değişkenler için **Categorical** kısımları seçilmektedir. Ayrıca belirlenecek

değişkenler için ağırlıklandırılma yapılabilmesi için **Weight** kısmı işaretlenmektedir. Belirlenecek değişken/değişkenler için dönüşüm uygulanmasının istenmediği, o değişkenin olduğu gibi modelde yer almasının tercih edildiği durumda **Non Transform** kısmı seçilmelidir. İlgili tanımlamaların yapıldığı ekran görüntüsü Şekil 3.4’de görülmektedir.

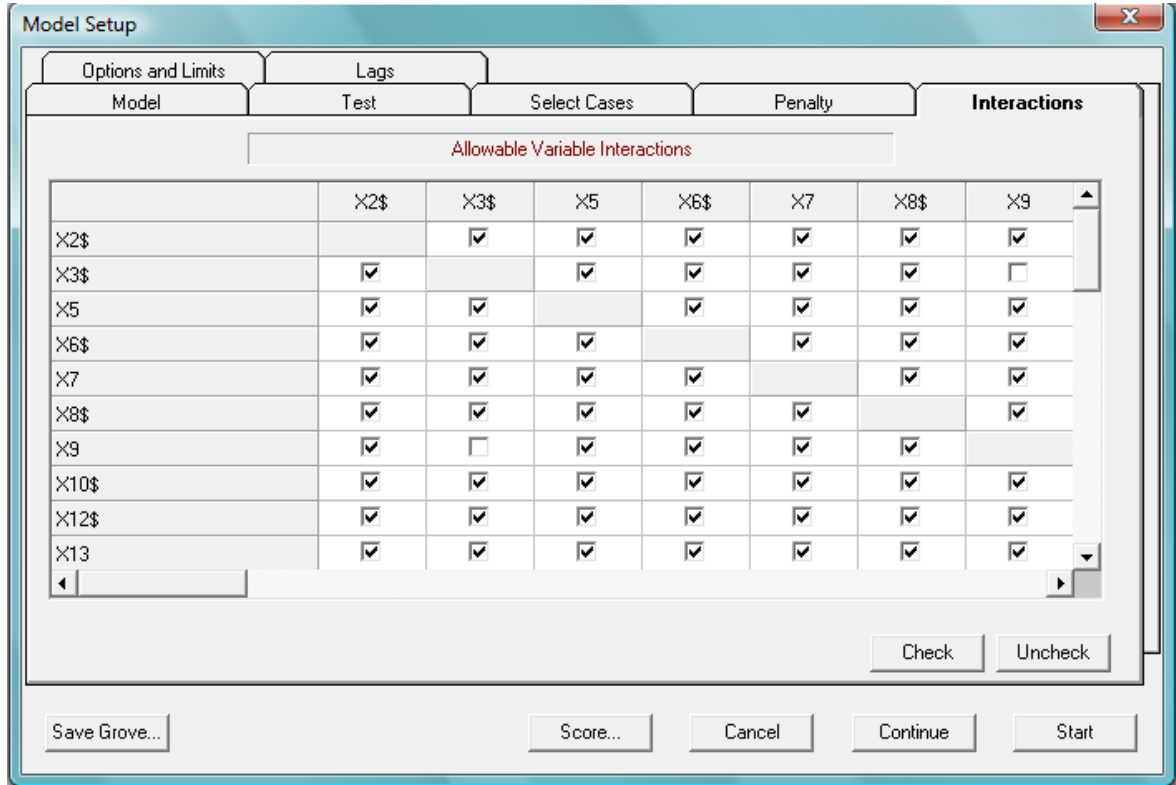


Şekil 3.4. MARS yazılımında model kurulumu

Başka bir model kullanılarak tahmin edilmiş bir değer ya da herhangi bir skor değeri olan bir değişkenin modelde hiç bir dönüşüme maruz kalmadan, olduğu gibi yer alması ve modele olan etkisinin araştırılmak istenildiği durumlar söz konusu olabilmektedir. İşte böyle durumlarda seçilen değişkenler için dönüşümlerin engellenmesi istenebilir ve bu da o değişkenler için bulunacak düğüm noktalarının engellenmesine karşılık gelmektedir. Dolayısıyla problem düğüm noktalarının engellenmesi ile çözülebilecektir. Böyle bir durumda herhangi bir değişken belirlenirse, o değişkenin minimum değerinde sözde bir düğüm noktası belirlenecektir. MARS yönteminin geriye doğru adımından sonra en son modelde bu değişkenin bulunmama olasılığı da söz konusudur.

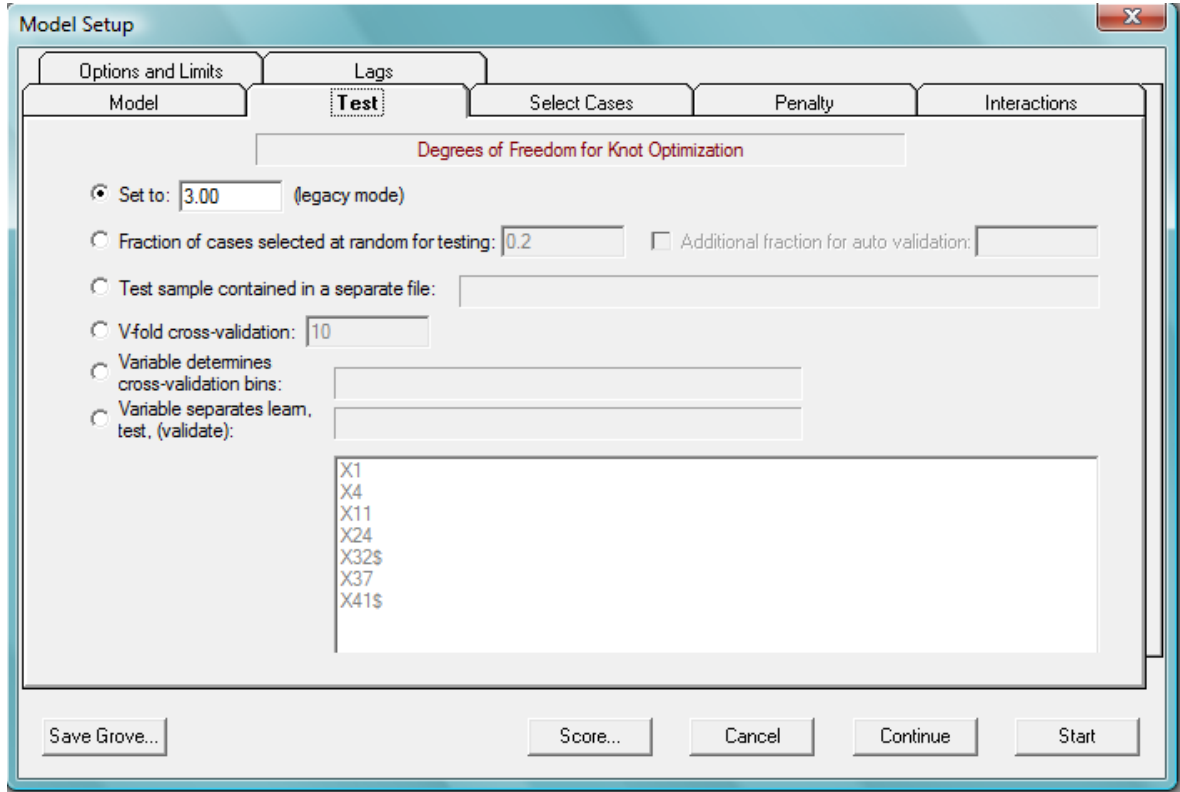
Eğer bütün değişkenler için dönüşümler engellenirse, MARS yazılımı adimsal regresyonun bir varyasyonu şeklinde çalışacaktır. Bu da o modelin referans alınarak dönüşümlerin getirilerini ölçmede kullanılabilir (Salford Systems, 2001).

MARS yazılımı ile veri setindeki belirlenebilecek maksimum derecedeki etkileşimlerin ve aynı zamanda ikili etkileşimlerin kontrolü de yapılabilir. Araştırmacının problem çeşidine bağlı olarak belirlenen bazı etkileşimler engellenebilir. Ayrıca herhangi bir etkileşimden istenen bir değişkeni hariç tutmak ya da tam tersi istenen bir değişkeni etkileşime dahil etmek mümkün olmaktadır. Bu yazılımda iki yolla bunu yapabilmek mümkündür. Birincisi, model kurulum aşamasında Şekil 3.4’de görülmektedir. Görülen **Model** menüsünden **Non interacting** seçeneği seçilerek, o değişkenin hiç bir etkileşime girmemesi sağlanabilir. İkincisi, model kurulumunda etkileşim sayısı belirlendikten sonra **Model Setup** penceresinde aktifleşen **Interactions** bölümünde tüm tahmin edici değişkenlerin etkileşimlerinin belirlenebileceği satırlar ve sütunlardan oluşan bir matris kullanılarak olmaktadır. Şekil 3.5’de görülen bu matriste yer alan hücreler seçilerek/seçilmeyerek değişkenler arasındaki lokal etkileşimler belirlenebilir/engellenebilir.



Şekil 3.5. MARS yazılımında etkileşim matrisi

Model Setup ekranından **Test** kısmına geçildiğinde kurulacak model için serbestlik derecesinin belirlenebildiği ekrana ulaşılmaktadır. Bu çözümleme için serbestlik derecesi 3 olarak belirlenmiştir ve Şekil 3.6'da görülmektedir.

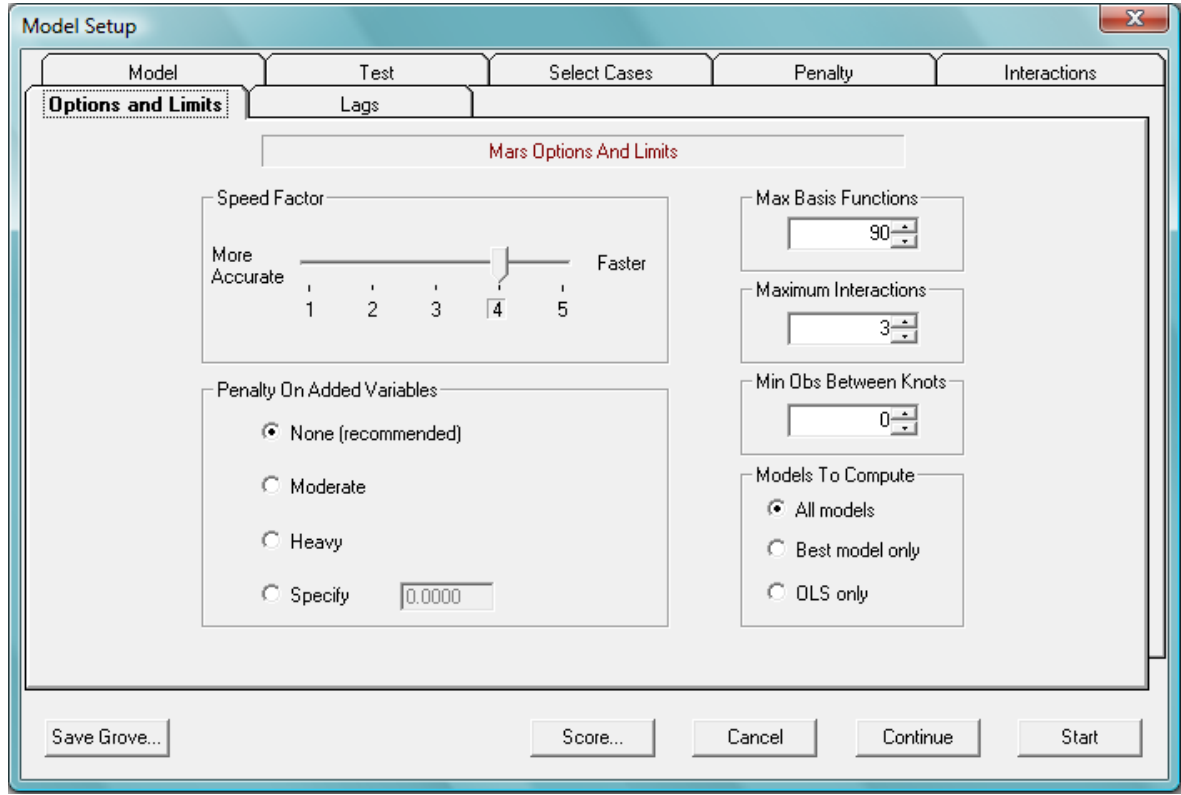


Şekil 3.6. MARS yazılımında serbestlik derecesinin belirlenmesi

Model Setup ekranından **Options and Limits** kısmı seçildiğinde oluşturulabilecek temel fonksiyon ve etkileşim sayısının belirlenebildiği ekrana ulaşılmaktadır. **Maximum Interactions** kısmına manuel etkileşim sayıları girilerek, 2 yönlü ya da 3 yönlü etkileşim ayarlanabilmektedir. Böylece herhangi bir değişkenden ya da değişkenlerden kurulan temel fonksiyonlar arasındaki etkileşimler incelenebilecektir. Bu bölümde bir değişiklik yapılmadığı takdirde herhangi bir etkileşimin bulunmadığı modeller kurulacaktır. Bu ekran Şekil 3.7.'de görülebilmektedir.

Model kurulum aşamasına herhangi bir etkileşime izin vermeden başlanması ve böylece ana etkilerin yer aldığı bir modelin oluşturulması yorumlama açısından hem kolay olacak hem de ana etkiler daha iyi anlaşılacaktır. Böyle bir model için MARS yazılımı her değişkenin optimal dönüşümleri ile her birinin temel fonksiyonlarını oluşturacaktır ve en son model her bir temel fonksiyon çarpımları toplamıyla oluşacaktır. Ana etkilerin yer aldığı model yardımıyla oluşturulan modelin neyi anlattığı, modelin araştırılan konu için uygun olup olmadığı gibi

soruların yanıtı bulunabilecektir. Bu aşamadan sonra etkileşimlerin denenmesi ve modelin buna göre geliştirilmesi uygun olacaktır (Salford Systems, 2001).



Şekil 3.7. MARS yazılımında temel fonksiyon sayısının ve etkileşime girecek değişken sayısının belirlenmesi

Model kurulurken yöntemin ileriye doğru adımında eklenecek temel fonksiyon sayısı, MARS yazılımında var olan ayarlar değiştirilmezse en çok 15 ya da daha az sayıda olacaktır. Şekil 3.7'deki ekranda temel fonksiyonların maksimum sayısının belirlendiği kısım görülebilmektedir. Temel fonksiyonların maksimum sayısı, tahmin edici değişken sayısının en azından iki katı kadar olacak şekilde seçilmelidir (Salford Systems, 2001).

Maksimum temel fonksiyon sayısı belirlenirken değişkenler arasındaki etkileşim sayısı da dikkate alınmalıdır. Hem etkileşimler hem de ana değişkenlerin etkilerinin araştırılması için maksimum sayının artırılmasında fayda vardır.

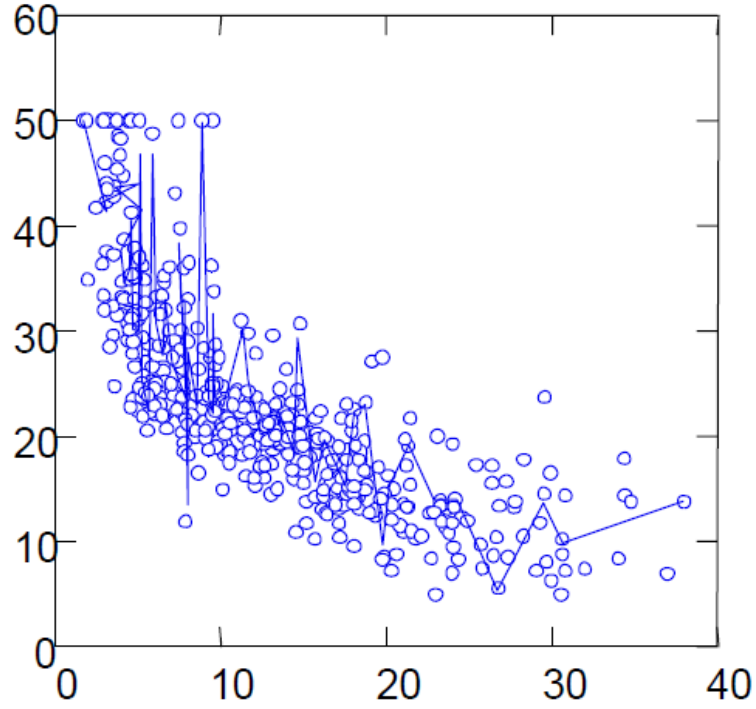
MARS yazılımının Şekil 3.7'de görülen bir başka özelliği de eklenen değişkenler için ceza belirlenmesidir. Eklenen değişkenler üzerindeki ceza, zaten modelde var

olan deęiřkenlerin eklenen yeni deęiřkenler olarak yeniden kullanılmasını saęlayan bir durumdur. Ceza arttırıldıkça, MARS var olan deęiřkenlerde yeni düęüm noktaları yaratacak ya da var olan deęiřkenleri içeren etkileřim terimleri oluřturma eğilimi gösterecektir.

Ceza kavramı aslında çoklubaęlantı (multicollinearity) sorunu ile mücadele etmek için sunulmuřtur. X_1 , X_2 ve X_3 deęiřkenlerinin yüksek derecede korelasyona sahip olduęu varsayılıں. Eğer X_1 ilk olarak modele dahil edilirse ve eklenen deęiřkenler için bir ceza varsa, MARS X_1 , X_2 ve X_3 'ün kombinasyonlarını kullanmak yerine sadece X_1 'i kullanma eğiliminde olmaktadır. Eğer deęiřkenler arasındaki korelasyon biraz yüksekse, model uyumunda bir miktar kayıp olacaktır. Ceza aynı zamanda çok az deęiřkene sahip olmasına karřın çok fazla temel fonksiyona sahip olabilecek modellerin kurulmasında da kullanılmaktadır (Salford Systems, 2001).

MARS yazılımı veri setindeki her deęerde düęüm noktası oluřturmaya izin vermektedir. Yani veride dikte edilen her yerde MARS regresyonunun eğiminin ya da yönünün deęiřmesine izin verilmektedir. Yöntemin geriye doęru adımında gereksiz olan noktalar silinmektedir.

řekil 3.8'de MARS yazılımı kullanılarak tek bir tahmin edici ile kurulan basit bir model için oluřturulan fonksiyon görölmektedir. İki deęiřken arasında çok keskin iniř ve çıkıřların olduęu bir iliřkiye sahip olunduęu görölmektedir. Eğer MARS düz çizgilerle belirtilmiř böyle bir model seęerse, bu modelin veri setini çok iyi yansıtacaęı ařıkardır. Ancak bu derece bir esneklik bazı uygulamalarda tercih edilmeyebilir ve daha az lokal iliřkileri temsil eden, daha düz bir model istenebilir. Dolayısıyla düęüm noktalarının sayısı bu anlamda önemli olmaktadır.



Şekil 3.8. MARS yazılımı kullanılarak elde edilen $y=f(x)$ fonksiyon dağılımı (Salford Systems, 2001)

Düğüm noktalarını sınırlamanın, başka bir deyişle MARS modelini daha az lokal uyarlamalı yapmanın en etkili yolu, minimum genişliği makul bir oranda olabildiğince büyük tanımlamak ya da başka bir deyişle düğüm noktaları arasındaki minimum gözlemlerin sayısını arttırmak olacaktır. Örneğin; minimum sayı 100 olarak tanımlandığında, düğüm noktaları arasında en az 100 gözlem olmak zorundadır. Şekil 3.7’de yer alan **Model Setup** penceresinin **Options and Limits** bölümünde **Min Obs Between Knots** kutusuna düğüm noktaları arasında bulunması istenen gözlem sayısı manuel olarak girilerek belirlenebilmektedir.

Model kurulumuyla ilgili ölçütler belirlendikten sonra **Start** tuşuna basıldığında model kurulum aşaması başlatılmaktadır.

3.2. Çözümlemede Kullanılan Veriler

Çözümlemede kullanılan veriler Türkiye'deki bir finans kurumuna Aralık 2008 – Ocak 2009 tarihleri arasında, internet, şube ve sms gibi kaynaklar aracılığıyla yapılan, belirli bir portföye bağlı olmayan, ihtiyaç kredisi başvurularını içermektedir. Veri setinde yapılan değerlendirmeler sonucunda verilen onay ya da ret kararları, verilen bu karar doğrultusunda oluşturulan onaylanmış limit, onaylanmış vade ve onaylanan kredilerin kullandırılıp, kullandırılmadıklarıyla ilgili bilgiler de yer almaktadır. Tüm finans kurumlarının verdiği bilgilerle oluşturulan bir bilgi bankası olan Kredi Kayıt Bürosu (KKB) verileri de her başvuru yapan birey için uygulamada yer almıştır. Çalışmada 48 adet değişkenin olduğu toplam 120.433 adet başvuru bulunmaktadır. Bu değişkenlerin 21 tanesi nitel ve 27 tanesi niceldir.

Çalışmanın esas amacı kredi kararının verilmesinde diğer değişkenlerin etkilerinin araştırılması olduğundan, 9 düzeyli olan KREDI_KARARI değişkenindeki kabul/ret dışındaki diğer karar kriterleri(değerlendirilmesi tamamlanmamış, manuel değerlendirilmeye yönlendirilmiş, iade edilmiş vb.) çıkartılarak, sadece farklı kabul ve ret kriterleri(kabul ve ret nedenleri farklı olanlar) birleştirilmiş, böylece değişken 2 düzeyli hale getirilmiştir. Dolayısıyla yeni başvuru sayısı 65.000 olmuştur. Tüm değişkenler ve açıklamaları Çizelge 3.1'de görülmektedir. Çalışmadaki değişkenlerin özellikleri, tanımlayıcı istatistikleri Ek1'de ve değişkenlere ait grafikler Ek2'de detaylı bir şekilde verilmiştir.

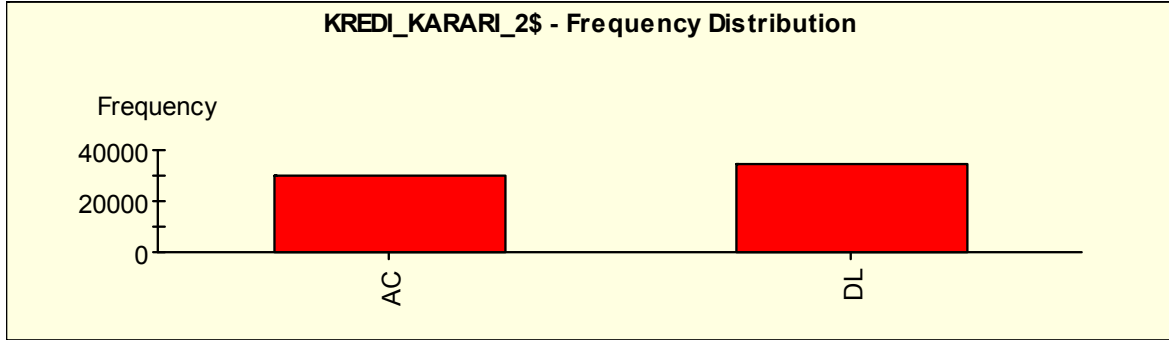
Sıra	Değişkenler	Değişkenin Açıklaması
1	TEMİNAT	Teminat
2	EGITIM_DURUMU	Eğitim durumu
3	KULL_AMACI	Kredi kullanım amacı
4	EV_TELEFON_BILG	Ev telefon bilgisi var/yok
5	ODEME_PLANI_TIPI	Ödeme planı tipi
6	OTURULAN_EV_DURUMU	Oturduğu evin durumu
7	YENIDEN_DEG_SAYISI	Yeniden değerlendirme sayısı
8	YENI_MEVCUT_MUSTERI	Yeni / Mevcut müşteri ayrımı
9	TOPLAM_CAL_SURESI	Toplam çalışma süresi (ay)
10	MAAS_MUSTERISI	Maaş müşterisi olup / olmama
11	BANKAYLA_CAL_SURESI	Bankayla çalışma süresi (gün)
12	SIMDIKI_ISYERI_CAL_SURESI	Şimdiki iş yeri çalışma süresi (ay)
13	MUSTERI_TIPI	Müşteri tipi
14	ILISKI_SKORU	İlişki skoru
15	SEKTOR_TIPI	Sektör tipi
16	MUSTERI_BANKA_SEGMENTI	Müşterinin banka segmenti
17	KOTU_KAYIT_TURU	Müşterinin daha önceden almış olduğu kötü kayıt türü
18	CALISMA_SEKLI	Çalışma şekli
19	SUBE_GORUSU	Şube görüşü
20	SOSYAL_GUVENCE	Sosyal güvencesi
21	BEYAN_EDILEN_GELIR	Müşterinin başvuru sırasında beyan ettiği gelir
22	MESLEK	Meslek
23	SON_1AYDAKI_RET_ADEDI	Son 1 ay içerisindeki kredi başvurularında aldığı ret adedi
24	KREDI_KARAR_KODU	Kredi karar kodu
25	SON_1AYDAKI_ONAY_ADEDI	Son 1 ay içerisindeki kredi başvurularında aldığı onay adedi
26	KKB_AYLIK_TAKSIT	KKB sorgusu sonucunda görüntülenen tüm bankalardaki toplam taksit tutarı
27	KKB_GUNCEL_ENKOTU_STATU	KKB'deki güncel en kötü statüsü
28	KKB_GECMIS_ENKOTU_STATU	KKB'deki geçmiş en kötü statüsü
29	MEDENI_DURUM	Medeni durum
30	KREDI_KARARI	Kredi kararı
31	KKB_TOP_BORC	KKB sorgusu sonucunda görüntülenen tüm bankalardaki açık kredilerinin bakiye toplamı
32	COCUK_SAYISI	Çocuk sayısı
33	KKB_SON_6AYDAKI_EN_KOTU_ODEME	KKB sorgusu sonucunda görüntülenen tüm bankalardaki açık kredilerinin son 6 ay ödemesinde gecikme var ise en kötüsü
34	CINSIYET	Cinsiyet
35	ONAYLANAN_VADE	Onaylanan vade
36	YAS	Yaş
37	BELGELENEN_TOP_GELIR	Belgelenen toplam gelir
38	BORCLULUK_ORANI	Borçluluk oranı (borçlar/gelir)
39	KULLANDIRIM_DURUMU	Kredinin kullandırım durumu
40	BASVURU_SKORU	Başvuru skoru
41	KKB_SKOR	Büro skoru - KKB'nin ürettiği skor
42	TALEP_TL	Müşterinin talep ettiği TL tutar
43	TALEP_DOVIZ	Müşterinin talep ettiği tutar (döviz)
44	TALEP_VADE	Müşterinin talep ettiği vade
45	KREDI_PARA_BIRIMI	Kredinin para birimi
46	ONAYLANAN_LIMIT	Onaylanan limit

Çizelge 3.1. Veri setindeki değişkenler

Önceki bölümlerde MARS yazılımı ile model kurma aşamasında kayıp değişkenler için yeni bir temel fonksiyon üretilerek o temel fonksiyonla birlikte model kurulabildiğinden bahsedilmiştir. Bu nedenle model kurulurken, kayıp değer oranları kabul edilebilir seviyede olan değişkenler için bu yeni oluşturulan temel fonksiyonlar kabul edilecek ve sonuçlar buna göre incelenecektir. Buna göre Ek1'de yer alan değişkenlere ait kayıp değer oranları incelendiğinde, kayıp değer oranı en yüksek olan değişkenler sırasıyla **TEMİNAT**, **SUBE_GORUSU** ve **KOTU_KAYIT_TURU** olarak karşımıza çıkmaktadır. **TEMİNAT** değişkenine ait kayıp değer oranı %100 olarak çıkmıştır. Kredi kararı değişkeni için düzenleme yapılmadan önce bu değişkenin kayıp değer oranı %96 olarak bulunmuştur. Dolayısıyla veri setindeki teminat oranı sadece %4'tür ve ilgili düzenlemelerden sonra bu değişkene ait %4'lük bilgi de silinmiştir. Değişken çözümlemeye dahil edilemeyecektir. **SUBE_GORUSU** değişkeni için kayıp oranı %96'dır. Bu çalışmada farklı kanallardan(şube, sms, internet, atm vb.) yapılan kredi başvuruları kullanıldığından, her kredi başvurusu için şube görüşü beklenmediğinden ve bu şube görüşünün kredi kararına olan etkisinin araştırılmak istenmesi sebepleriyle bu değişken için kayıp oranı kabul edilebilir bir seviye olarak belirlenmiştir. 4 düzeyli olan **SUBE_GORUSU** değişkeni modele dahil edildiğinde, oluşturulacak dummy değişken ile kredi kararına etkisi araştırılacaktır. Böylece şube görüşünün 4 farklı durumunda nasıl etkilendiğine ilave olarak, şube görüşünün bulunduğu ya da bulunmadığı durumlarda kredi kararına olan etkileri görülebilecektir. **KOTU_KAYIT_TURU** değişkeni için kayıp değer oranı da %95 olarak gerçekleşmiştir. Başvuruda bulunan kişilere ait o finans kurumundaki mevcut kredilerine ya da kredi kartlarına ait ödemelerindeki sakıncalı durum bilgilerini gösteren bu değişkende, sakıncalı bir durum yoksa giriş yapılmadığından kayıp oranı kabul edilebilir bir düzeydedir. Başvuran kişilerin %95'inde kötü ödeme bulunmadığı söylenebilir. Bu durumda 4 farklı kötü kayıt türü için ve ilave olarak kötü kaydın varlığında kredi kararının nasıl etkilendiği gözlemlenebilecektir. Diğer değişkenlerin hesaplanan kayıp değer oranları kabul edilebilir seviyede gerçekleşmiştir.

3.3. MARS Çözümlemesi

Çalışmanın bu bölümünde kurulacak MARS modelinin bağımlı değişkeni **KREDI_KARARI** olarak belirlenmiştir. Bu değişkene ait histogram grafiği Şekil 3.9'da verilmiştir. Görüldüğü gibi çalışılan veri setindeki kabul oranı %46,46 ve ret oranı %53,54'tür.



Şekil 3.9. Kredi kararı değişkeni histogram grafiği (MARS yazılımı)

Çözümlemede Çizelge 3.2'de görülen toplam 35 adet bağımsız değişken seçilmiştir.

1	EGITIM_DURUMU	19	SON_1AYDAKI_ONAY_ADEDI
2	KULL_AMACI	20	KKB_AYLIK_TAKSIT
3	EV_TELEFON_BILG	21	KKB_GUNCEL_ENKOTU_STATU
4	ODEME_PLANI_TIPI	22	KKB_GECMIS_ENKOTU_STATU
5	OTURULAN_EV_DURUMU	23	MEDENI_DURUM
6	YENIDEN_DEG_SAYISI	24	KKB_TOP_BORC
7	YENI_MEVCUT_MUSTERI	25	COCUK_SAYISI
8	TOPLAM_CAL_SURESI	26	KKB_SON_6AYDAKI_EN_KOTU_ODEME
9	MAAS_MUSTERISI	27	CINSIYET
10	BANKAYLA_CAL_SURESI	28	YAS
11	SIMDIKI_ISYERI_CAL_SURESI	29	BELGELENEN_TOP_GELIR
12	MUSTERI_TIPI	30	BORCLULUK_ORANI
13	KOTU_KAYIT_TURU	31	BASVURU_SKORU
14	CALISMA_SEKLI	32	KKB_SKOR
15	SUBE_GORUSU	33	TALEP_TL
16	SOSYAL_GUVENCE	34	TALEP_DOVIZ
17	BEYAN_EDILEN_GELIR	35	TALEP_VADE
18	SON_1AYDAKI_RET_ADEDI		

Çizelge 3.2. Model kurulumunda seçilen değişkenler

Bu bölümdeki çözümlemeye dahil edilmeyen değişkenler sırasıyla **ILISKI_SKORU**, **SEKTOR_TIPI**, **MUSTERI_BANKA_SEGMENTI**, **MESLEK**, **KREDI_KARAR_KODU**, **KULLANDIRIM_DURUMU**, **ONAYLANAN_VADE**, **ONAYLANAN_LIMIT**'tir.

ILISKI_SKORU değişkeni müşteri ve finansal kurum arasındaki bağlantı nedenini göstermektedir. Ek1'de de görüleceği üzere müşterinin banka ile olan ilişki türünü gösreran 1.263 farklı değer bulunmaktadır ve bu 1.263 adet farklı ilişki düzeyi çok yüksek bulunduğundan ve sadece kendi başına kurulacak modelin karmaşıklığını arttıracığından çözümlemeye dahil edilmemiştir.

SEKTOR_TIPI değişkeni de başvuruda bulunan kişilerin sektör tipini göstermektedir ve 52 düzeylidir. Bu nedenle de bu değişken model kurulumunda yer almayacaktır.

MUSTERI_BANKA_SEGMENTI değişkeni müşteri banka segmentini göstermektedir ve her finansal kurum için farklılık gösterdiğinden çalışmaya dahil edilmemiştir.

MESLEK değişkeni başvuru yapan bireylerin öğretmen, doktor, şöför gibi meslek bilgilerini içermektedir. Çözümlemeye meslek bilgisi için **CALISMA_SEKLI** değişkeni dahil edilmiştir. **MESLEK** değişkeni **CALISMA_SEKLI**'nin bir alt kümesi olduğundan ve daha ayrıntılı bilgiler içerdiğinden çözümlemede yer almamıştır.

KREDI_KARAR_KODU değişkeni değerlendirmeler sonucunda onaylanan / ret edilen kredilerin ne sebeple onaylandığını / reddedildiği gibi bilgileri içermektedir. Ancak bu değişkenin modele dahil edilmesi kredi kararının onay / ret durumunu %100 etkilemektedir ve modeldeki diğer değişkenlerin önem düzeyi sıfır çıkmaktadır. Bu nedenle de çalışmaya dahil edilmemiştir.

KULLANDIRIM_DURUMU değişkeni, onaylanan kredinin kullandırıp kullandırılmadığı ile ilgili bilgileri içermektedir. Dolayısıyla bu modelleme çalışmasına dahil edilmemiştir.

ONAYLANAN_VADE ve **ONAYLANAN_LIMIT** deęişkenleri de onaylanan kredi tutarı ve onaylanan kredinin vade bilgisi olduğundan, bağımlı deęişkene olan etkileri yüksek olacağından ve dięer deęişkenleri incelemek açısından engel teşkil edeceğinden çalışmaya alınmamıştır.

Kredi kayıt bürosundan elde edilen verilerin kredi kararını büyük ölçüde etkilemesi beklenmektedir. Çünkü bu veriler, tüm finansal kurumlar tarafından oluşturulan ortak bir veri bankasından elde edilmiş, kişilerin risk bilgileri ile ilgili bilgilerdir. Dolayısıyla kredi kararı verilirken başvuruda bulunan kişilerin güncel risk tutarı, geçmiş en kötü ödemesi, aylık taksit tutarı gibi bilgileri göz önüne alınması beklenmektedir.

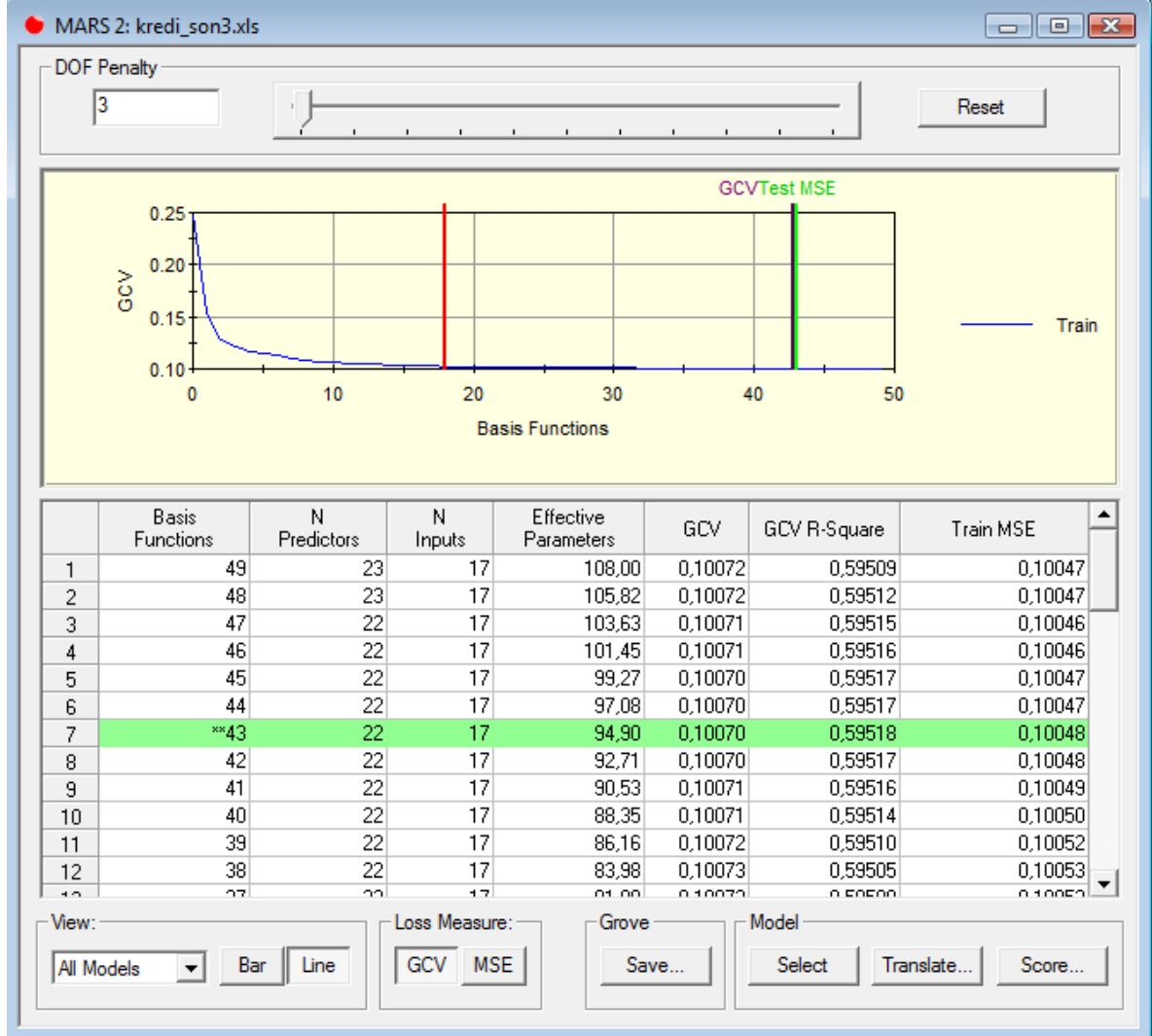
Deęişken seçimi yapılmış, serbestlik derecesi 3 ve etkileşim sayısı da 0 olarak belirlenmiştir. Bağımsız deęişken sayısı 35 olduğundan minimum temel fonksiyon sayısı 70 olmalıdır ve bu çalışma için temel fonksiyon sayısı 70 olarak belirlenmiştir.

3.3.1. Model seçimi

MARS yazılımı bu kriterler doğrultusunda toplam 50 farklı model üretmiştir. Yazılımın özellięi gereęi, program tarafından kurulan modellerin yer aldığı ekranda en iyi model kullanıcıya farklı bir renk ile belirtilerek gösterilmektedir. Bu çalışmada da en iyi model 7 numaralı model olarak belirlenmiştir ve Şekil 3.10'da verilen MARS yazılımı ekran görüntüsünde oluşturulan modeller içerisinde en iyi modelin koyu renk ile belirtildięi görölmektedir.

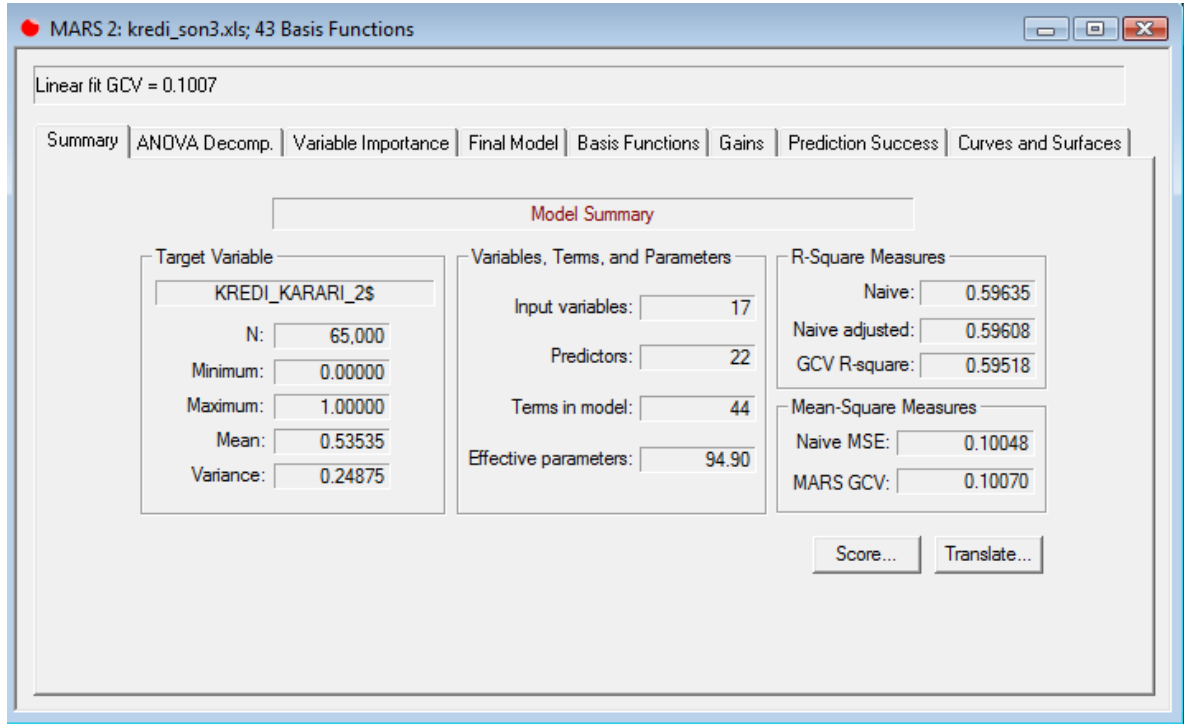
Şekil 3.10'da görülebilen GCV ve Temel Fonskiyon grafięine bakıldığında, temel fonksiyon sayısı arttıkça GCV deęerinin kendi minimumuna ulaştığı görölmektedir. Modeldeki temel fonksiyon sayısı, bir başka ifade ile modele giren terim sayısı(sabit terim dışında) 17'den sonra GCV deęeri minimumuna ulaşmıştır. Ancak temel fonksiyon sayısının 43 olduğ u, yani sabit terim dışında modeldeki terim sayısının 43 olduğ u model en iyi model olarak belirlenmiştir ve grafikte en iyi modelin bulunduę u yer koyu renk ile gösterilmiştir. Model kurulum aşamasında maksimum temel fonksiyon sayısı 70 olarak belirlenmiş, en iyi modele giren temel

fonksiyon sayısı da 44 olarak gerçekleşmiştir. Bunun nedeni ileriye doğru adımdan sonra sadece modele dahil edilmesi anlamlı olan temel fonksiyonların seçilerek diğerlerinin silinmesidir.



Şekil 3.10. MARS yazılımında en iyi modelin gösterildiği ekran

En iyi modelde temel fonksiyonların oluşturulmasında kullanılan bağımsız değişken sayısı 17'dir. Ancak kayıp değerler içeren değişkenler için oluşturulan dummy değişkenlerle birlikte temel fonksiyonların oluşturulmasında kullanılan toplam değişken sayısı 22 olmaktadır. GCV değeri 0,10070 olarak hesaplanmış ve $GCV-R^2 = 0,59518$ olarak bulunmuştur. Bu bilgilere MARS yazılımındaki Model Özet tablosundan da ulaşılabilmektedir. Şekil 3.11'de MARS yazılımı model özet bilgilerinin bulunduğu ekran görüntüsü yer almaktadır.



Şekil 3.11. MARS yazılımı model özet bilgilerinin bulunduğu ekran

Ek3'de yer alan MARS modeli çıktı bilgilerinde model kurulum aşamasında belirlenen tüm parametrelerin standart hata, t-Ratio ve p değerlerinin bulunduğu özet bilgiler yer almaktadır. Bu özet bilgilerin bulunduğu tablodaki p değerleri incelendiğinde, sabit terim dışındaki tüm değişkenlerin p değerlerinin 0,05'den büyük olduğu görülmektedir. Dolayısıyla değişkenlerinin kredi kararına bir etkileri bulunmamaktadır. Tüm bu değişkenlerin bulunduğu bir modelleme yapılsaydı modelin F değeri 0,80685 ve p değeri de 0,84337 olacaktı. Dolayısıyla kurulacak model istatistiksel olarak anlamsız olacaktı. MARS modelinde bu noktada ileriye doğru adım süreci gerçekleştirilerek gerekli dönüşümler yapılacak ve temel fonksiyonlar oluşturulacaktır. Oluşturulan bu temel fonksiyonlarla doğrusal bir model elde edilmeye çalışılacaktır. Daha sonra geriye doğru adım süreciyle bu temel fonksiyonlar arasından istatistiksel olarak anlamsız bulunanlar modelden çıkartılarak en iyi modele ulaşılabacaktır.

3.3.2. İleriye doğru adım

MARS modelinin ileriye doğru adımında modele giren temel fonksiyonlar Ek4'te verilmiştir. Ek4'te verilen rapor incelendiğinde modelin ileriye doğru adımında oluşturulan olası tüm temel fonksiyonlar, bu temel fonksiyonları oluşturan değişkenler ve o değişkenler için oluşturulan düğüm noktaları görülebilmektedir. Model kurulum aşamasında maksimum temel fonksiyon sayısı 70 olarak belirlenmişti ve buna göre MARS yazılımı 70 tane temel fonksiyon üretmiştir. Aşağıda MARS yönteminin ileriye doğru adım aşamasında modele dahil edilen Temel Fonksiyonların(TF) açıklamaları verilmiştir. İkili olarak verilen temel fonksiyonlar çalışmanın başında anlatılan yansıma çiftleridir. Temel Fonksiyonlar TF olarak kısaltılmıştır ve çalışmanın bundan sonraki bölümlerinde bu şekilde ifade edilecektir.

Model kurulum aşamasında etkileşim sayısı belirlenmemiştir. Dolayısıyla kurulan modelde sadece değişkenlerin ana etkileri görülebilecektir. Ancak kayıp değerler içeren bağımsız değişkenler için üretilen dummy fonksiyonların diğer değişkenlerle etkileşim içerisinde olduğu görülmektedir. Bunun nedeni; kayıp değerleri olan bağımsız değişkenlerin bağımlı değişken üzerindeki etkisinin görülebilmesidir.

Ek4'te verilen ileriye doğru adımda modele giren temel fonksiyonlar tablosunun her satırında temel fonksiyonların ikili olarak yer aldığı görülmektedir. Bunun nedeni tez çalışmasının başında bahsedilen temel fonksiyonların yapısı gereği yansıma çifti olarak oluşturulmasıdır. Örnek olarak TF45 ve TF46 temel fonksiyonlarına bakılırsa, bu temel fonksiyonlar **BELGELENEN_TOP_GELIR** değişkeninden oluşmaktadır. **BELGELENEN_TOP_GELIR** değişkeni modele 1150.00000 düğüm noktasında TF45 ve TF46 temel fonksiyonlarıyla girmiştir. TF45 ve TF46 yansıma çiftidir. Sürekli bir değişken olan **BELGELENEN_TOP_GELIR** için oluşturulan temel fonksiyonlar $TF45 = (BELGELENEN_TOP_GELIR - 1150)$ ve $TF46 = (1150 - BELGELENEN_TOP_GELIR)$ olarak tanımlanmaktadır, buna göre TF45 belgelenen toplam gelir tutarının 1.150.-TL'den büyük olduğu tutarları ve TF46'da bunun tam tersi belgelenen toplam gelir tutarının 1.150.-TL'den küçük olan tutarları göstermektedir. Kategorik bir değişken olan **BORCLULUK_ORANI dummy**

değişkeni modele 0 ve 1 düğüm noktasında TF1 ve TF2 temel fonksiyonlarıyla girmiştir. Değişken kategorik olduğunda temel fonksiyonlar o koşulların sağlandığı durumu göstermektedir. Dolayısıyla $TF1 = (BORCLULUK_ORANI \text{ dummy} = 0)$ ve $TF2 = (BORCLULUK_ORANI \text{ dummy} = 1)$ olmaktadır. TF1 borçluluk oranının bulunduğu durumları, TF2’de borçluluk oranının bulunduğu durumları göstermektedir.

Aşağıda Ek4’te verilen ileriye doğru adımda modele giren temel fonksiyonların bazılarının nasıl yorumlanabildiğiyle ilgili açıklamalar bulunmaktadır.

TF0: Sabit terim modele her zaman TF0 olarak, yani ilk sırada ve sıfır temel fonksiyon olarak girmektedir.

TF3-TF4: **BORCLULUK_ORANI** 15.00000 düğüm noktasında yine **BORCLULUK_ORANI** dummy değişkeninin 1 düğüm noktasıyla birlikte modele girmiştir. Yani buradaki yansıma çiftleri $(BORCLULUK_ORANI - 15.00000)$ ve $(15.00000 - BORCLULUK_ORANI)$, $(BORCLULUK_ORANI \text{ dummy} = 1)$ ile etkileşim içerisinde olmaktadır. TF3, borçluluk oranı varlığında hesaplanan bu borçluluk oranının 15’den büyük olduğu durumları göstermektedir.

TF13-TF14: **KKB_GUNCEL_ENKOTU_STATU** dummy değişkeni 0 ve 1 düğüm noktalarında **BORCLULUK_ORANI** dummy değişkeninin 1 düğüm noktasıyla birlikte modele dahil edilmiştir. Buna göre TF13 borçluluk oranı varlığında, kredi kayıt bürosundan elde edilen günceldeki en kötü gecikme bilgisinin bulunduğu durumu göstermektedir.

TF15- TF16: **KKB_GUNCEL_ENKOTU_STATU** 0011111000 düğüm noktalarında yine bu değişken için oluşturulan dummy değişkeninin 1 ve 3 düğüm noktalarında birlikte modele girmiştir. MARS kategorik değişkenleri modele dahil ederken ayrıca bir kodlama kullanmaktadır. Şöyleki; **KKB_GUNCEL_ENKOTU_STATU** değişkeni 0, 1, 2, 3, 4, 5, 6, 8, L ve U olmak üzere toplam 10 düzeyli bir değişkendir. Her düzey için 0 ve 1’ler kullanılarak ayrı bir kodlama yapılarak, modele hangi düzeyin etkili olduğu gösterilmektedir. 0011111000 düğüm noktası 10 basamaklı bir ifadedir ve sıfırlar etkili olmayan düzeyi ve 1’ler de etkili olan düzeyleri

göstermektedirler. Dolayısıyla 2, 3, 4, 5, 6 düzeyleri 1'ler ile ve 0, 1, 8, L ve U düzeyleri de 0'larla ifade edilmiştir. TF15 ve TF16 yansıma çiftleri olduğundan ve birbirlerini tümleyen bir yapıda olmaları gerektiğinden, TF15 = (KKB_GUNCEL_ENKOTU_STATU = 0011111000) şeklinde tanımlanırsa, TF16 = (KKB_GUNCEL_ENKOTU_STATU = 1100000111) olacaktır. Sonuç olarak TF15 = (KKB_GUNCEL_ENKOTU_STATU = 2, 3, 4, 5, 6) ve TF16 = (KKB_GUNCEL_ENKOTU_STATU = 0, 1, 8, L, U) yansıma çiftleri KKB_GUNCEL_ENKOTU_STATU = 1, 3 ile birlikte etkileşim içerisinde modele dahil edilmiştir.

TF25-TF26: **KKB_SON_6AYDAKI_EN_KOTU_ODEME** değişkeni 001111110000 düğüm noktalarında **KKB_GUNCEL_ENKOTU_STATU** dummy değişkeninin 1 ve 3 düğüm noktalarında birlikte modele dahil edilmiştir. **KKB_SON_6AYDAKI_EN_KOTU_ODEME** değişkeni 0, 1, 2, 3, 4, 5, 6, 8, D, L, U ve X olmak üzere 12 düzeyli bir değişkendir. TF25 ve TF26 yansıma çiftleri (KKB_SON_6AYDAKI_EN_KOTU_ODEME = 2, 3, 4, 5, 6, 8) ve (KKB_SON_6AYDAKI_EN_KOTU_ODEME = 0, 1, D, L, U, X) olarak (KKB_GUNCEL_ENKOTU_STATU dummy = 1, 3) ile birlikte etkileşim içerisinde dir.

3.3.3. Geriye doğru adım

Bu aşamada ileriye doğru adımda oluşturulan 70 temel fonksiyon araştırılmış ve GCV kriteri kullanılarak modele en az katkıda bulunan temel fonksiyonlar silinmiştir ve 44 adet temel fonksiyon kalmıştır. Geriye doğru adım sürecinden sonra modele giren temel fonksiyonlar Ek5'de verilmiştir. Geriye doğru adım sürecinden sonra TF0 sabit terim ile birlikte modele giren temel fonksiyonlar aşağıda sıralanmıştır.

TF0 TF1 TF3 TF4 TF5 TF6 TF7 TF9 TF12 TF13 TF15 TF17
TF18 TF19 TF20 TF21 TF22 TF25 TF27 TF28 TF31 TF35 TF36 TF37
TF38 TF39 TF41 TF43 TF45 TF46 TF47 TF49 TF50 TF51 TF53 TF54
TF56 TF59 TF61 TF63 TF64 TF66 TF67 TF68

3.3.4. Varyans analizi

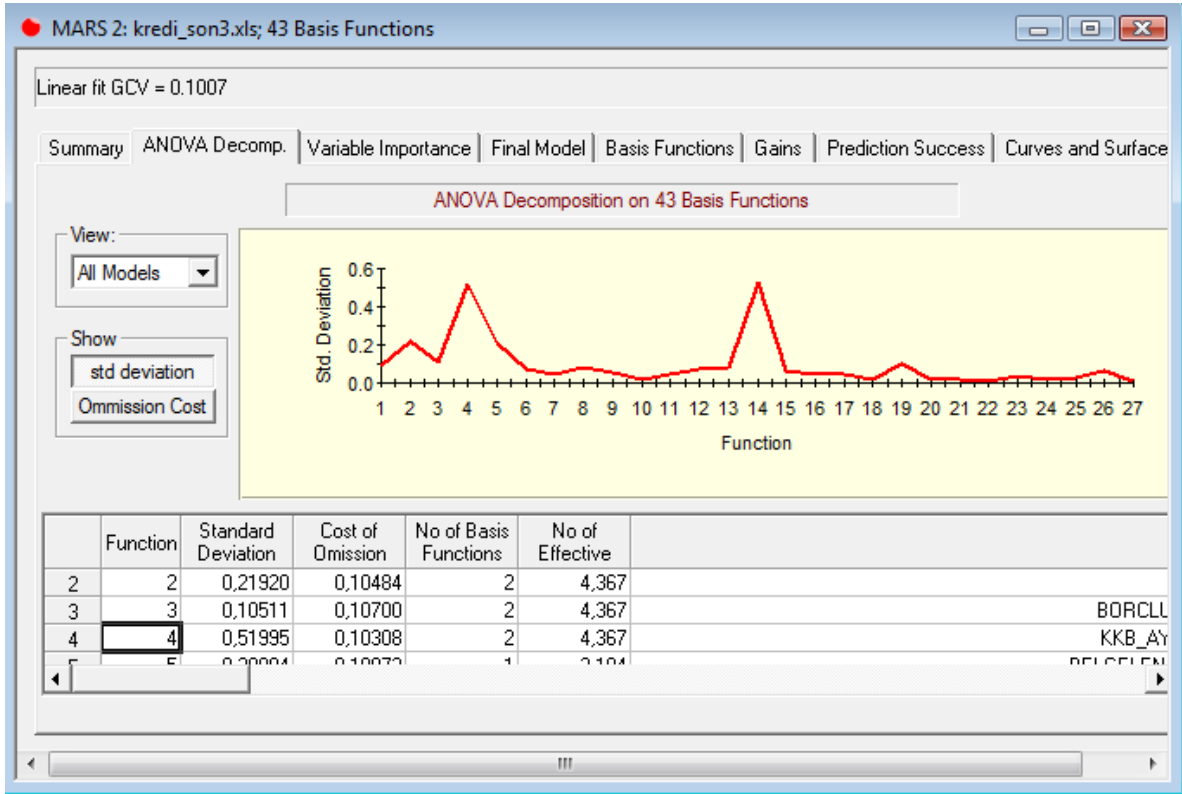
MARS yazılımı ile değişken/değişkenlerin birlikte modele olan katkıları ANOVA ayrıştırma tablosu yardımıyla görülebilmektedir. Çizelge 3.3’de görülen varyans ayrıştırma tablosunda modele giren 43 temel fonksiyonu oluşturan değişken/değişkenlerin standart sapma değeri, çıkarmadaki kayıp, temel fonksiyon sayısı, efektif parametre sayısı ve o temel fonksiyonu oluşturan değişkenler görülmektedir.

Çizelge 3.3’deki “Çıkarmadaki Kayıp” sütunu söz konusu değişken ya da bir araya getirilen değişkenlerin modelden çıkartılması durumunda, tahminlerde meydana gelecek kaybı, başka bir ifade ile modeldeki uyum eksikliğini göstermektedir. Temel fonksiyon sayısı o değişken ya da bir araya getirilmiş değişkenlerin kaç temel fonksiyonla modele dahil olduklarını ve efektif parametre sayısı o değişken/değişkenler birleşiminin yüklendiği toplam serbestlik derecesini göstermektedir(Salford Systems, 2001).

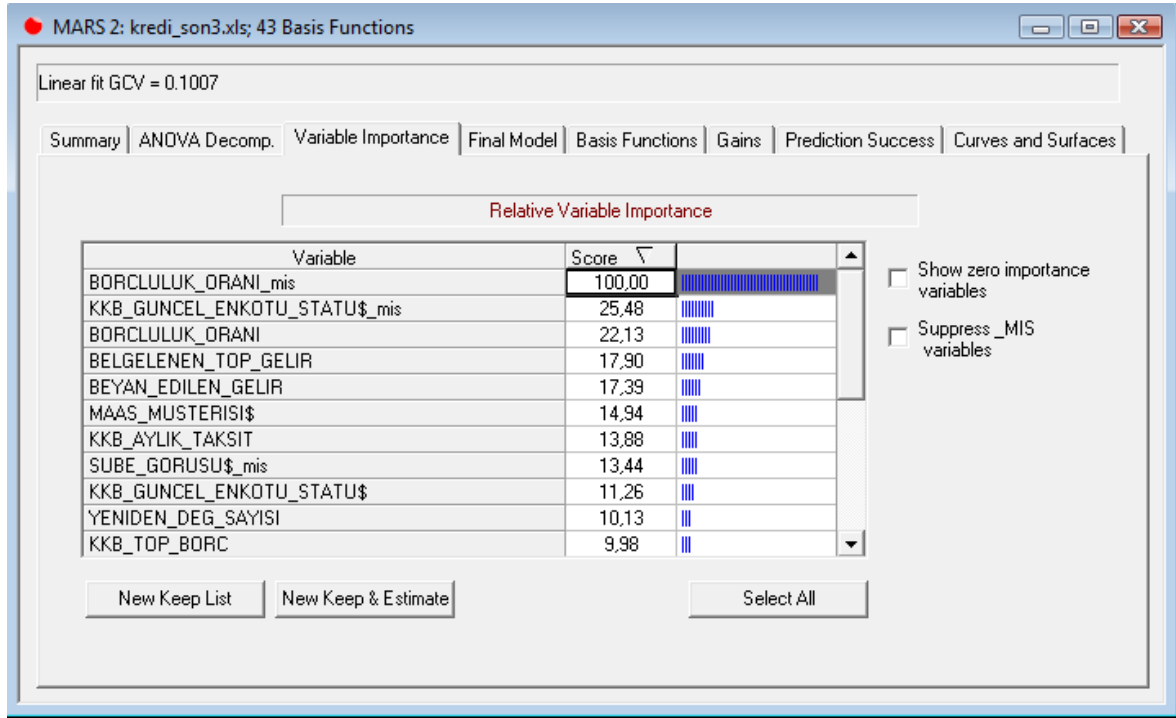
Fonksiyon	Standart Sapma	Çıkarmadaki Kayıp	Temel fonksiyon sayısı	Efektif parametre sayısı	Değişkenler
1	0,08708	0,10073	1	2,184	BORCLULUK_ORANI_mis
2	0,2192	0,10484	2	4,367	BELGELENEN_TOP_GELIR
3	0,10511	0,107	2	4,367	BORCLULUK_ORANI, BORCLULUK_ORANI_mis
4	0,51995	0,10308	2	4,367	KKB_AYLIK_TAKSIT, BORCLULUK_ORANI_mis
5	0,20994	0,10072	1	2,184	BELGELENEN_TOP_GELIR, BORCLULUK_ORANI_mis
6	0,07438	0,10359	1	2,184	MAAS_MUSTERISI\$, BORCLULUK_ORANI_mis
7	0,04184	0,10098	1	2,184	BEYAN_EDILEN_GELIR, BORCLULUK_ORANI_mis
8	0,0827	0,1008	1	2,184	KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
9	0,05321	0,10106	2	4,367	TALEP_TL, BORCLULUK_ORANI_mis
10	0,01979	0,10099	1	2,184	CALISMA_SEKLI, BORCLULUK_ORANI_mis
11	0,04737	0,10227	1	2,184	KKB_GUNCEL_ENKOTU_STATU\$, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
12	0,07598	0,10203	2	4,367	YENIDEN_DEG_SAYISI, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
13	0,0788	0,10089	2	4,367	KKB_TOP_BORC, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
14	0,527	0,10183	2	4,367	KKB_AYLIK_TAKSIT, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
15	0,06217	0,10185	3	6,551	BEYAN_EDILEN_GELIR, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
16	0,04638	0,1012	2	4,367	SON_1AYDAKI_RET_ADEDI, SUBE_GORUSU\$_mis, BORCLULUK_ORANI_mis
17	0,04501	0,10094	2	4,367	KKB_TOP_BORC, SUBE_GORUSU\$_mis, BORCLULUK_ORANI_mis
18	0,0176	0,10098	1	2,184	KULL_AMACI, SUBE_GORUSU\$_mis, BORCLULUK_ORANI_mis
19	0,09662	0,10089	1	2,184	SUBE_GORUSU\$_mis, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
20	0,01793	0,10091	2	4,367	BORCLULUK_ORANI, SUBE_GORUSU\$_mis, BORCLULUK_ORANI_mis
21	0,01478	0,10077	1	2,184	TALEP_VADE, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
22	0,01144	0,10076	2	4,367	YAS, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
23	0,03202	0,10145	2	4,367	KKB_SON_6AYDAKI_EN_KOTU_ODEME\$, KKB_GUNCEL_ENKOTU_STATU\$_mis, KKB_SON_6AYDAKI_EN_KOTU_ODEME\$_mis, BORCLULUK_ORANI_mis
24	0,0197	0,101	1	2,184	KKB_GECMIS_ENKOTU_STATU\$, KKB_GUNCEL_ENKOTU_STATU\$_mis, KKB_GECMIS_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
25	0,028	0,10114	2	4,367	KKB_GUNCEL_ENKOTU_STATU\$, SUBE_GORUSU\$_mis, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
26	0,0674	0,10081	1	2,184	KKB_AYLIK_TAKSIT, SUBE_GORUSU\$_mis, KKB_GUNCEL_ENKOTU_STATU\$_mis, BORCLULUK_ORANI_mis
27	0,00973	0,10076	2	4,367	BANKAYLA_CAL_SURESI, KKB_GUNCEL_ENKOTU_STATU\$_mis, KKB_SON_6AYDAKI_EN_KOTU_ODEME\$_mis, BORCLULUK_ORANI_mis

Çizelge 3.3. Varyans ayrıştırma tablosu

MARS yazılımında varyans ayrıştırmasının bulunduğu ekran görüntüsü Şekil 3.12’de verilmiştir. Burada yer alan Standart Sapma ve fonksiyonlar grafiğinde Çizelge 3.2’de sıralanan 27 adet fonksiyondaki değişken/değişkenler birleşimlerinin standart sapma değerleri görülmektedir. Bu grafiğe göre 2, 4, 14, 19 numaralı fonksiyonların standart sapmaları yüksektir. Bir başka ifadeyle standart sapmanın yüksek olması o değişkenin modelin tüm açıklama gücüne olan katkısının da yüksek olduğunu göstermektedir. Dolayısıyla 14 numaralı fonksiyonda gösterilen ve standart sapması en yüksek olan **KKB_AYLIK_TAKSIT**, **KKB_GUNCEL_ENKOTU_STATU** dummy değişken ve **BORCLULUK_ORANI** dummy değişkenlerinin birlikte kredi kararı modellemesinde önem düzeylerinin yüksek olduğu söylenebilir. Benzer şekilde 4 numaralı fonksiyonda yer alan **KKB_AYLIK_TAKSIT**, **BORCLULUK_ORANI** dummy değişkenlerinin birlikte kredi kararına olan etkisi yüksektir.



Şekil 3.12. MARS yazılımı ANOVA ayrıştırma tablosunun yer aldığı ekran



Şekil 3.13. MARS yazılımında modeldeki değişkenlerin nispi önem düzeylerinin sıralandığı ekran

MARS yazılımında en iyi modeli oluşturan değişkenlerin nispi önem düzeylerine göre aldıkları skor değeriyle birlikte kullanıcıya sunulmaktadır. Bu skor değerleri bağımsız değişkenlerin modeldeki diğer değişkenlere göre nispi katkılarıdır. Şekil 3.13’de MARS yazılımında en iyi modelde yer alan değişkenlerin en yüksek önem derecesinden en düşük önem derecesine göre sıralanarak verildiği ekran görüntüsü yer almaktadır. Çizelge 3.4’de modele giren tüm değişkenlerin önem dereceleri görülebilmektedir.

BORCLULUK_ORANI dummy değişkeninin önem düzeyi %100 çıkmıştır. Bir başka ifadeyle borçluluk oranının bulunup/bulunmaması durmu kredi kararını %100 etkilemektedir. Kredi talebinde bulunan kişilerin borçluluk oranlarının yüksek olmasının ya da tersi bu oranın düşük olmasının kredi kararına direkt olarak etkisinin olması beklenen bir durumdur. **BORCLULUK_ORANI**’nın kredi kararına etkisi %22,13 seviyesinde gerçekleşmiştir. Borçluluk oranı başvuru sahibinin mevcut tüm risklerinin gelirlerine oranı ile hesaplanmaktadır. Dolayısıyla bu oranın kredi kararı verilirken önemli olması zaten beklenen bir durumdur.

KKB_GUNCEL_ENKOTU_STATU dummy değişkeni önem düzeyi %25,48 olarak hesaplanmıştır ve KKB’de güncel en kötü statü bilgisinin bulunup/bulunmamasının kredi kararına etkisi %25,48’dir. Bu kabul edilebilir, mantıklı bir durumdur. Kredi kayıt bürosunda olumsuz bilgilerin bulunmamasının kredi kararı verilirken önemli olması beklenmektedir. **KKB_GUNCEL_ENKOTU_STATU** değişkenin önem düzeyi de %11.26’dır. Dolayısıyla kredi kayıt bürosundaki güncel olumsuz bilgilerin kredi kararına bir etkileri olduğu söylenebilir ve daha detaylı yorumlar model aşamasında yapılabilecektir.

DEĞİŞKEN	Skor
BORCLULUK_ORANI_dummy	100
KKB_GUNCEL_ENKOTU_STATU_dummy	25,48
BORCLULUK_ORANI	22,13
BELGELENEN_TOP_GELIR	17,9
BEYAN_EDILEN_GELIR	17,39
MAAS_MUSTERISI	14,94
KKB_AYLIK_TAKSIT	13,88
SUBE_GORUSU_dummy	13,44
KKB_GUNCEL_ENKOTU_STATU	11,26
YENIDEN_DEG_SAYISI	10,13
KKB_TOP_BORC	9,98
KKB_SON_6AYDAKI_EN_KOTU_ODEME_dummy	7,81
KKB_SON_6AYDAKI_EN_KOTU_ODEME	7,6
SON_1AYDAKI_RET_ADEDI	6,2
TALEP_TL	5,25
KKB_GECMIS_ENKOTU_STATU_dummy	4,81
KKB_GECMIS_ENKOTU_STATU	4,81
CALISMA_SEKLI	4,72
KULL_AMACI	4,63
TALEP_VADE	2,35
BANKAYLA_CAL_SURESI	2,15
YAS	2,12

Çizelge 3.4. Modele giren değişkenlerin nispi önem dereceleri

BELGELENEN_TOP_GELIR ve **BEYAN_EDILEN_GELIR** değişkenlerinin kredi kararına olan etkileri %17 düzeyinde olmuştur. Gelir bilgisi kredi kararı verilirken önemli bir bilgidir. Belgelenen gelirin kredi kararına göre yapılan modellemede beyan edilene göre daha önemli olması beklenmiştir. Bu beklentiyle model kurulum aşamasında her iki değişken de seçilmiştir. Sonuç olarak belgelenen

toplam gelir %17,9 önem düzeyi ile beyan edilen toplam gelire göre daha etkili bir değişkendir.

MAAS_MUSTERISI değişkeni %14,94 önem düzeyi ile modele girmiştir. Kredi kararı verilirken o başvuruyu yapan kişinin maaş müşterisi olup olmasının önemli çıkması kabul edilebilir bir sonuçtur. Kredi geri ödemelerinde, zaten maaş ödemelerinin o finans kurumundan geçiyor olması bir avantaj olarak düşünülebileceği gibi, maaş ödemelerinin o kurumdan geçmesi kişinin daha iyi tanınıyor olmasını ve bu da kişinin kredi kullanım ihtiyaçlarını, ödeme prensiplerini, olumsuz ya da olumlu tüm bilgilerinin biliniyor olmasını beraberinde getirmektedir. Modelin verildiği aşamada bu değişken için daha detay yorum yapılabilecektir.

KKB_AYLIK_TAKSIT değişkeninin modeldeki önem derecesi %13,88'dir. Başvuru yapan bireylerin kredi kayıt bürosundan elde edilen tüm finansal kurumlardaki aylık taksit miktarının kredi kararına önemli bir etkisinin bulunduğu söylenebilir.

SUBE_GORUSU dummy değişkeni önem düzeyi %13,44 olarak gerçekleşmiştir. Bir başvuruda şube görüşünün bulunması, olumlu/olumsuz kredi kararını etkilemektedir. Bununla ilgili daha detay yorumlar model incelendiğinde yapılabilecektir.

YENIDEN_DEG_SAYISI değişkeni %10,13 ile önemli bulunmuştur. Bir kredi teklifinin birden fazla değerlendirilmesinin kredi kararı verilirken etkili bir değişken olması beklenmektedir. Model yorumlarında bu etkinin olumlu mu olumsuz mu olduğu görülecektir.

KKB_TOP_BORC değişkeni önem düzeyi % 9,98 olarak hesaplanmıştır. Başvuru yapan bireyin kredi kayıt bürosundan elde edilen tüm finansal kurumlardaki toplam borç tutarının kredi kararı verilirken önemli olması beklenen bir durumdur.

KKB_SON_6AYDAKI_EN_KOTU_ODEME dummy değişkeni ve **KKB_SON_6AYDAKI_EN_KOTU_ODEME** değişkeninin modele etki düzeyleri % 7'dir. Kredi kayıt bürosunda son 6 aydaki en kötü ödeme bilgisinin

bulunup/bulunmaması durumu ve bu kötü ödeme kaydının tipi kredi kararı verilirken önemlidir.

SON_1AYDAKI_RET_ADEDI değişkeni % 6,2 ile önemli bulunmuştur. Yapılan başvurunun reddedilmesi durumunun kredi kararı verilirken önemli bir kriter olmasını bekleriz. Reddedilme nedenine göre verilecek yeni karar değişebilir. Dolayısıyla bu değişken için çıkan önem düzeyi de kabul edilebilir bir durumdur.

TALEP_TL değişkeninin önem düzeyi % 5,25 olarak gerçekleşmiştir. Başvuru sırasında talep edilen tutar bilgisinin önemli bulunması kabul edilebilir bir durumdur. Kredi kararı verilirken düşük ya da çok yüksek limit taleplerinde bulunulmasının önemli bir kriter olması beklenmektedir.

KKB_GECMIS_ENKOTU_STATU dummy değişkeni ve **KKB_GECMIS_ENKOTU_STATU** değişkeninin önem düzeyi %4,81 olarak hesaplanmıştır. Kredi kayıt bürosundan elde edilen başvuru yapan bireylerin geçmişteki en kötü olumsuz bilgilerinin kredi kararı verilirken sorgulanıyor olması beklenen bir durumdur. Her iki değişkenin kredi kararına etkileri aynı çıkmıştır. Model kısmında daha detay yorumlar yapılacaktır.

CALISMA_SEKLI değişkeninin kredi kararına olan etki düzeyi %4,72'dir. Başvuru yapan bireylerin özel sektör çalışanı ya da kamu çalışanı vb. olması durumunun verilen kredi kararlarında önemli olması zaten beklenen bir durumdur. Kamu çalışanlarının düzenli gelirleri, özel sektöre göre daha tercih edilesi bir durum olabilir. Bu değişkenin gerçek etkisi model kısmında incelenecektir.

KULL_AMACI % 4,63 önem düzeyi ile modele girmiştir. Talep edilen kredinin kullanım amacının belirtilmesinin kredi kararı verilirken önemli bir kriter olması kabul edilebilir bir durumdur. Daha önce reddedilmiş bir başvuru, kullanım amacına göre tekrar değerlendirilmeye alınabilir. Bu değişkenle ilgili daha detay yorumlar model incelenirken yapılacaktır.

TALEP_VADE değişkeninin önem düzeyi % 2,35'dir. Finansal kurumlar verilecek kredinin vadesinin kısa olmasını isteyecektir. Bunun nedeni de bu koşullarda

verilen kredilerin daha abuk geri donecek olmasıdır. Dolayısıyla kredi kararı verilirken talep edilen vadenin ok yuksek ya da ok düşek olması kredi kararı verilirken nemli bir kriterdir.

BANKAYLA_CAL_SURESI deėiřkeninin nem düzeyi %2.15'dir. Uzun süredir o finansal kurumla alıřan bir mřterinin olumlu olumsuz tüm zellikleriyle tanınmasından, rneėin; daha nceki kredi risklerini dzenli demelerle kapatıp/kapatmaması, kredi kartı harcamaları vb bilgilere sahip olmanın kredi kararı verirken etkisinin olması beklenen bir durumdur.

YAS deėiřkeni de % 2,12 nem düzeyi ile modele girmiřtir. Bu deėiřken ile ilgili detay yorum model kısmında yapılacaktır.

3.3.5. Temel fonksiyonlar

Geriye doėru adım sürecinden sonra modele en az katkısı olan temel fonksiyonlar ıkarılmıř ve en iyi modele giren temel fonksiyonlar belirlenmiřtir. Bu blmde en son modele giren temel fonksiyon eřitlikleri, bu temel fonksiyonların tanım kmeleri ve en son model tanımlanacaktır.

Daha nce kategorik ve dummy deėiřkenler iin ayrı bir kodlama yapıldıėından bahsedilmiřti. MARS yazılımı kategorik ve dummy deėiřkenlerin hangi dzeylerinin nemli olduėunu tanımlanan alt kmeler ile gstermektedir. Kategorik deėiřkenlerin dzeylerinin belirli gruplara ayrılmasıyla oluřan bu alt kmeler aracılıėıyla model yorumlaması basit bir řekilde yapılabilmektedir. Dummy deėiřkenler iin ise 0 ve 1'ler ile oluřturulan temel fonksiyon yapısı olduėu anlatılmıřtı. Deėiřkende kayıp deėer varsa 1, yoksa 0 olarak tanımlanan temel fonksiyon iin verilen bir alt kmede kayıp deėerlerin var olması durumunda o temel fonksiyonun modele etkisinin yorumlanması ya da benzer řekilde o temel fonksiyonun var olan deėiřkenlerin varlıėında baėımlı deėiřkene olan etkisinin yorumlanması saėlanmaktadır. Bu alt kmeler sırasıyla Eř. 3.1, Eř. 3.2, Eř. 3.3, Eř. 3.4, Eř.3.5, Eř. 3.6, Eř. 3.7, Eř. 3.8, Eř. 3.9, Eř. 3.10 ve Eř. 3.11'de verilmiřtir.

BORCLULUK_ORANI_dummy;
Alt Küme1 = { BORCLULUK_ORANI var } (3.1)

MAAS_MUSTERISI;
Alt Küme1 = { "Y" } (3.2)

KKB_GUNCEL_ENKOTU_STATU_dummy;
Alt Küme1 = { KKB_GUNCEL_ENKOTU_STATU var } (3.3)

KKB_GUNCEL_ENKOTU_STATU;
Alt Küme1 = { "2", "3", "4", "5", "6" }
Alt Küme 2 = { "2", "3", "4", "5", "6", "U" }
Alt Küme 3 = { "0", "2", "3", "4", "5", "U" } (3.4)

KKB_SON_6AYDAKI_EN_KOTU_ODEME_dummy;
Alt Küme 1 = { KKB_SON_6AYDAKI_EN_KOTU_ODEME var } (3.5)

KKB_SON_6AYDAKI_EN_KOTU_ODEME;
Alt Küme1 = { "2", "3", "4", "5", "6", "8" }
Alt Küme 2 = { "0", "1", "2", "4", "5", "6", "L" } (3.6)

KKB_GECMIS_ENKOTU_STATU_dummy;
Alt Küme1 = { KKB_GECMIS_ENKOTU_STATU var } (3.7)

KKB_GECMIS_ENKOTU_STATU;
Alt Küme1 = { "L" } (3.8)

SUBE_GORUSU_dummy;
Alt Küme1 = { SUBE_GORUSU yok }
Alt Küme2 = { SUBE_GORUSU var } (3.9)

KULL_AMACI;
Alt Küme1 = { "12", "13" } (3.10)

CALISMA_SEKLI;

Alt Küme1 = { "1010", "1040", "1080" } (3.11)

En son modele giren temel fonksiyon eşitlikleri Eş. 3.12'de verilmiştir.

TF1 = (BORCLULUK_ORANI var);
TF3 = max(0, BORCLULUK_ORANI - 15) * TF1;
TF4 = max(0, 15 - BORCLULUK_ORANI) * TF1;
TF5 = max(0, KKB_AYLIK_TAKSIT - 1492) * TF1;
TF6 = max(0, 1492 - KKB_AYLIK_TAKSIT) * TF1;
TF7 = max(0, BELGELENEN_TOP_GELIR - 1786) * TF1;
TF9 = (MAAS_MUSTERISI in ("Y")) * TF1;
TF12 = max(0, 2232.67 - BEYAN_EDILEN_GELIR) * TF1;
TF13 = (KKB_GUNCEL_ENKOTU_STATU var) * TF1;
TF15 = (KKB_GUNCEL_ENKOTU_STATU in ("3", "2", "4", "6", "5")) * TF13;
TF17 = max(0, YENIDEN_DEG_SAYISI - 1) * TF13;
TF18 = max(0, 1 - YENIDEN_DEG_SAYISI) * TF13;
TF19 = max(0, KKB_TOP_BORC - 29821) * TF13;
TF20 = max(0, 29821 - KKB_TOP_BORC) * TF13;
TF21 = max(0, KKB_AYLIK_TAKSIT - 252) * TF13;
TF22 = max(0, 252 - KKB_AYLIK_TAKSIT) * TF13;
TF23 = (KKB_SON_6AYDAKI_EN_KOTU_ODEME var) * TF13;
TF25 = (KKB_SON_6AYDAKI_EN_KOTU_ODEME in ("3", "2", "4", "8", "6", "5"))
* TF23;
TF27 = max(0, BEYAN_EDILEN_GELIR - 858) * TF13;
TF28 = max(0, 858 - BEYAN_EDILEN_GELIR) * TF13;
TF29 = (KKB_GECMIS_ENKOTU_STATU var) * TF13;
TF31 = (KKB_GECMIS_ENKOTU_STATU in ("L")) * TF29;
TF33 = (SUBE_GORUSU yok) * TF1;
TF34 = (SUBE_GORUSU var) * TF1;
TF35 = max(0, SON_1AYDAKI_RET_ADEDI - 1) * TF33;
TF36 = max(0, 1 - SON_1AYDAKI_RET_ADEDI) * TF33;
TF37 = max(0, KKB_TOP_BORC - 10782) * TF33;
TF38 = max(0, 10782 - KKB_TOP_BORC) * TF33;

$$\begin{aligned}
TF39 &= (KULL_AMACI \text{ in } (12, 13)) * TF33; \\
TF41 &= (KKB_GUNCCEL_ENKOTU_STATU \text{ var }) * TF34; \\
TF43 &= (KKB_GUNCCEL_ENKOTU_STATU \text{ in } ("3", "2", "U", "4", "6", "5")) * TF41; \\
TF45 &= \max(0, BELGELENEN_TOP_GELIR - 1150); \\
TF46 &= \max(0, 1150 - BELGELENEN_TOP_GELIR); \\
TF47 &= \max(0, BEYAN_EDILEN_GELIR - 7000) * TF13; \\
TF49 &= \max(0, TALEP_TL - 20000) * TF1; \\
TF50 &= \max(0, 20000 - TALEP_TL) * TF1; \\
TF51 &= (CALISMA_SEKLI \text{ in } (1040, 1080, 1010)) * TF1; \\
TF53 &= \max(0, BORCLULUK_ORANI - 15) * TF34; \\
TF54 &= \max(0, 15 - BORCLULUK_ORANI) * TF34; \\
TF56 &= \max(0, 6043 - KKB_AYLIK_TAKSIT) * TF41; \\
TF57 &= (KKB_GUNCCEL_ENKOTU_STATU \text{ var }) * TF33; \\
TF59 &= (KKB_GUNCCEL_ENKOTU_STATU \text{ in } ("0", "3", "2", "U", "4", "5")) * TF57; \\
TF61 &= (KKB_SON_6AYDAKI_EN_KOTU_ODEME \text{ in } ("1", "0", "2", "4", "6", "L", "5")) * TF23; \\
TF63 &= \max(0, BANKAYLA_CAL_SURESI - 28) * TF23; \\
TF64 &= \max(0, 28 - BANKAYLA_CAL_SURESI) * TF23; \\
TF66 &= \max(0, 36 - TALEP_VADE) * TF13; \\
TF67 &= \max(0, YAS - 23) * TF13; \\
TF68 &= \max(0, 23 - YAS) * TF13.
\end{aligned} \tag{3.12}$$

En son modeli oluşturan temel fonksiyonların tahmin değerleri, standart hataları, T-Ratio değerleri ve p değerleri Ek 6'da verilmiştir. Bu tabloya göre tüm temel fonksiyonların p değerleri $p < 0,05$ olarak hesaplanmıştır. Dolayısıyla modele giren bu temel fonksiyonlar anlamlıdır.

3.3.6. Modelin yorumlanması

Eş. 3.12’de verilen temel fonksiyonlarla kurulan en son modelin kapalı gösterimi de Eş. 3.13’de verilmiştir.

$$\begin{aligned} \text{MODEL KREDI_KARARI} = & \text{TF1 TF3 TF4 TF5 TF6 TF7 TF9} \\ & \text{TF12 TF13 TF15 TF17 TF18 TF19 TF20 TF21} \\ & \text{TF22 TF25 TF27 TF28 TF31 TF35 TF36 TF37} \\ & \text{TF38 TF39 TF41 TF43 TF45 TF46 TF47 TF49} \\ & \text{TF50 TF51 TF53 TF54 TF56 TF59 TF61 TF63} \\ & \text{TF64 TF66 TF67 TF68} \end{aligned} \quad (3.13)$$

Kabul edilen en son model Eş. 3.13;

$$\begin{aligned} Y = & 0.735671 - 0.18742 * \text{TF1} - 0.00462571 * \text{TF3} + 0.0369069 * \text{TF4} \\ & - 0.000174136 * \text{TF5} - 0.0001282 * \text{TF6} - 5.26293\text{E-}005 * \text{TF7} \\ & - 0.170077 * \text{TF9} + 7.24259\text{E-}005 * \text{TF12} + 0.174057 * \text{TF13} \\ & + 0.297651 * \text{TF15} + 0.0245812 * \text{TF17} - 0.151998 * \text{TF18} \\ & + 1.12298\text{E-}006 * \text{TF19} - 5.799\text{E-}006 * \text{TF20} + 0.000175163 * \text{TF21} \\ & + 0.00045513 * \text{TF22} + 0.122874 * \text{TF25} - 5.14336\text{E-}005 * \text{TF27} \\ & + 0.000748412 * \text{TF28} + 0.104 * \text{TF31} + 0.0157406 * \text{TF35} \\ & - 0.0937448 * \text{TF36} - 1.22924\text{E-}006 * \text{TF37} - 9.44826\text{E-}006 * \text{TF38} \\ & + 0.0779919 * \text{TF39} - 0.674358 * \text{TF41} - 0.416217 * \text{TF43} \\ & + 5.2668\text{E-}005 * \text{TF45} + 0.000230043 * \text{TF46} \\ & + 4.83984\text{E-}005 * \text{TF47} + 1.8785\text{E-}006 * \text{TF49} \\ & - 7.15759\text{E-}006 * \text{TF50} - 0.0552926 * \text{TF51} - 0.00383828 * \text{TF53} \\ & - 0.0325104 * \text{TF54} + 8.84942\text{E-}005 * \text{TF56} - 0.0371366 * \text{TF59} \\ & - 0.0344338 * \text{TF61} - 3.35616\text{E-}005 * \text{TF63} + 0.000959064 * \text{TF64} \\ & + 0.0014413 * \text{TF66} + 0.00100831 * \text{TF67} - 0.0107273 * \text{TF68} \end{aligned} \quad (3.13)$$

olarak bulunmuştur.

Çalışmanın önceki bölümlerinde temel fonksiyonların oluşturulmasında diğer temel fonksiyonların kullanılmasından bahsedilmiştir. Eş. 3.12’de görülen temel

fonksiyonlardan TF23, TF29, TF33, TF34 ve TF57 Eş. 3.13'deki final model eşitliğinde direkt olarak yer almamışlar ve diğer temel fonksiyonların oluşturulmasında kullanılmışlardır.

Kurulan bu model için 3 serbestlik derecesinde hesaplanan F değeri, p değeri sırasıyla Eş. 3.14 ve Eş. 3.15'de verilmiştir. Görüldüğü gibi model için hesaplanan F değeri oldukça yüksek bulunmuştur, bu da modelin anlamlı olduğunu göstermektedir. Benzer bir yorum p değerine bakılarak da yapılabilmektedir ($p < 0,05$).

$$F\text{-STATISTIC} = 2231.74254 \quad (3.14)$$

$$P\text{-VALUE} = 0.00000 \quad (3.15)$$

Kurulan modelin anlamlı olduğu Eş. 3.14 ve Eş. 3.15'deki istatistik değerleriyle kanıtlanmıştır ve modelin sınıflandırma başarısının da yüksek olması beklenmektedir. Bunun için Şekil 3.14'de görülen, MARS yazılımında verilen tahmin başarı tablosunun ekran görüntüsü incelenecektir.

MARS 2: kredi_son3.xls; 43 Basis Functions

Linear fit GCV = 0.1007

Summary | ANOVA Decomp. | Variable Importance | Final Model | Basis Functions | Gains | Prediction Success | Curves and Surfaces

Prediction Success Table

Actual Class	Total Cases	Percent Correct	Predicted Class	
			Class AC N=38219	Class DL N=26781
Class AC	30.202	97,18	97	3
Class DL	34.798	74,51	25	75
Total:	65.000.00			
Average:		85,84		
Overall % Correct:		85,04		

Threshold: 0.6032 Show Table... Count Row % Column %

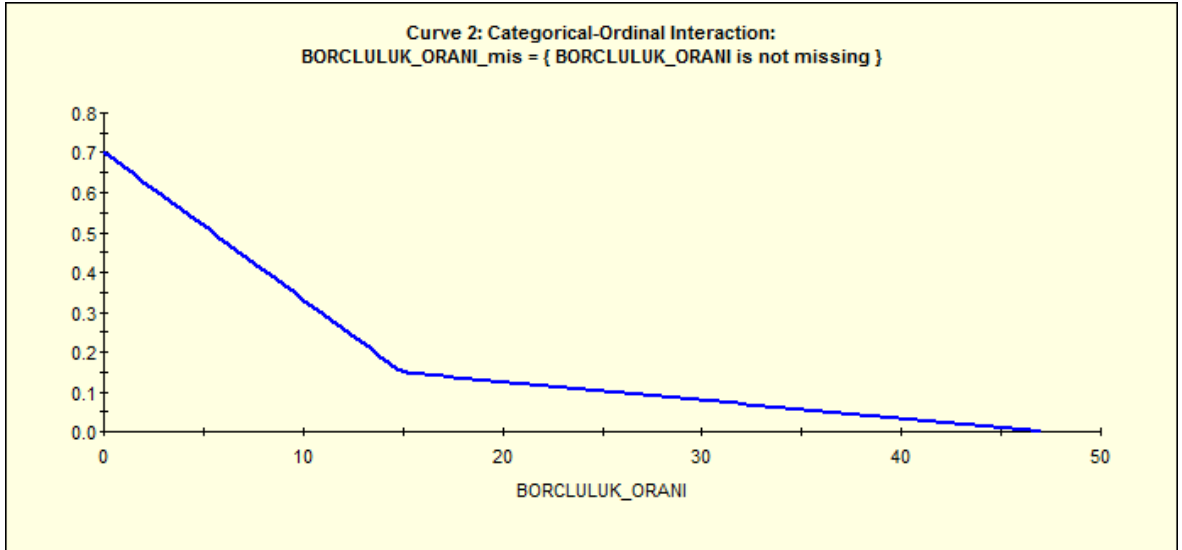
Şekil 3.14. MARS yazılımında sınıflandırma başarısının gösterildiği ekran

Şekil 3.14’de de görüldüğü gibi 65.000 gözlem değeri bulunan veri setindeki kabul edilen başvuru sayısı 30.202 ve ret edilen başvuru sayısı da 34.798’dir. Bağımlı değişkenin **KREDI_KARARI** olarak belirlenerek kurulan modelin doğru sınıflandırma oranı % 85,04 olarak bulunmuştur. Kurulan modelin kredi kararının kabul olması için doğru sınıflandırma oranı %97 olarak gerçekleşirken, ret olması durumu için bu oran %75 olarak bulunmuştur. Dolayısıyla kurulan modelin başarılı olduğu kabul edilmektedir.

Çalışmanın bu bölümünde kurulan ve geçerliliği kanıtlanan modeldeki değişkenlerin **KREDI_KARARI** bağımlı değişkenine olan etkileri ve bu değişkeni nasıl açıkladıklarıyla ilgili yorumlardan bahsedilecektir. Doğrusal modellerde katsayı değerleri, bağımsız değişkenin tüm değerleri için sabittir. Ancak Friedman (1991b) MARS modellerinde katsayı değerlerinin düğüm değerlerine göre değişim gösterdiğini söylemiştir (Temel vd., 2005). Bu durum mikro açıklamalar için bir avantajdır ancak modelde bağımlı ve bağımsız değişken arasındaki ilişkinin genel yorumları için oluşturulan grafiklerden yararlanılacaktır. Aşağıda sırasıyla önemli bulunan değişkenlerin kurulan modelde kredi kararına olan etkileri ve bağımlı değişken ile arasındaki ilişkileri yorumlanacaktır. Yorumlamalarda Ek 6’da verilen final modeline ait çıktı sonuçları, Eş. 3.13’de verilen kurulmuş MARS modeli, bu modeli oluşturan ve Eş. 3.12’de verilen temel fonksiyon tanımlamaları, bu temel fonksiyonlara ait Eş. 3.1, Eş. 3.2, Eş. 3.3, Eş. 3.4, Eş.3.5, Eş. 3.6, Eş. 3.7, Eş. 3.8, Eş. 3.9, Eş. 3.10 ve Eş. 3.11’de verilen alt kümeler birlikte kullanılacaktır.

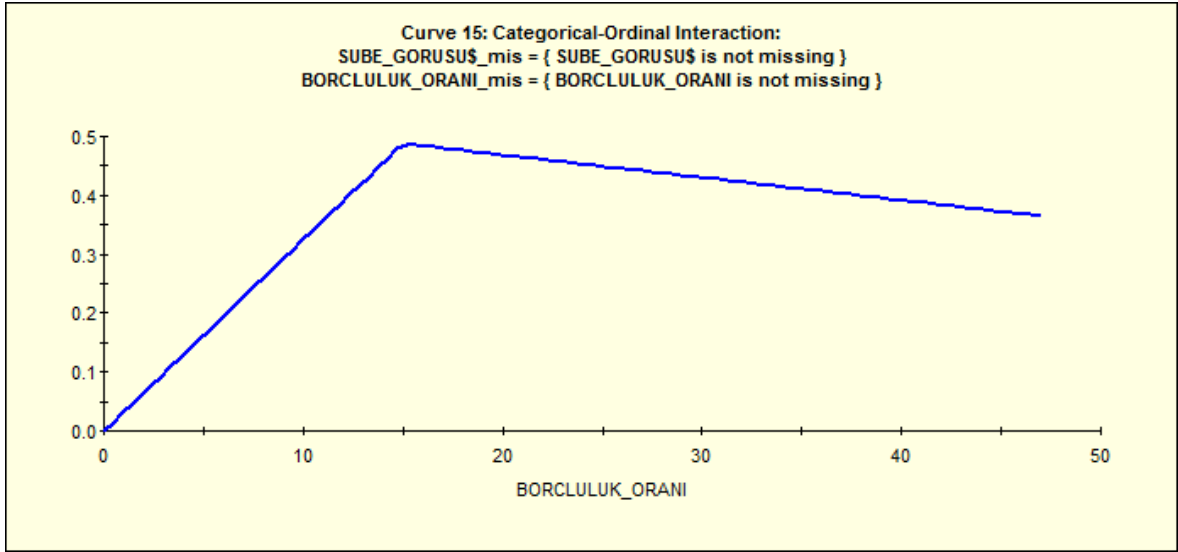
BORCLULUK_ORANI değişkeni için dummy değişken oluşturulmuştur. Çünkü her bir başvuru için borçluluk oranı bulunmamaktadır. Kişinin var olan tüm risklerinin geliri oranından oluşan borçluluk oranı, kredi kayıt bürosunda kaydı olmayan kişiler için hesaplanamamıştır. Bu durumda hiç bir kaydı olmayan bireylerin borçluluk oranı da olmayacaktır. Kurulan MARS modelinde **BORCLULUK_ORANI_dummy** TF1 fonksiyonu ile modelde yer almıştır. Bu dummy değişkenin modeldeki anlamı borçluluk oranının bulunması durumunun kredi kararını negatif yönde etkilemesidir. Kurulan modeldeki katsayısı -0,1874’dir.

BORCLULUK_ORANI değişkeni modele tek bir düğüm noktasında TF3-TF4 temel fonksiyonlarıyla girmiştir. Tek bir düğüm noktasında yansıma çiftleri ile modele girmesinin anlamı kredi kararı ile arasındaki ilişkinin doğrusal olmasıdır ve bu sebeple bu değişkene herhangi bir dönüşüm uygulanmamıştır. Bahsedilen bu ilişki negatif yönlüdür. Borçluluk oranının 15 olduğu nokta bu değişkenin düğüm noktasıdır ve borçluluk oranı ve kredi kararı arasındaki eğim 15 düğüm noktasından sonra değişmektedir. Şekil 3.15’de kredi kararı ve borçluluk oranı arasındaki ilişki görülmektedir. Görüldüğü gibi borçluluk oranı arttıkça kredi kararına olan etki de azalmaktadır. Borçluluk oranının 15’den büyük olduğu durum için model katsayısı -0,0046 olurken 15’den küçük olduğu durum için ise 0,0369 olarak gerçekleşmiştir. Kurulan modelde borçluluk oranının kredi kararına negatif yönlü bir etkisi bulunmaktadır ancak borçluluk oranı dummy değişkenin varlığında %15’den küçük borçluluk oranının etkisi, %15’den büyük olma durumu kadar etkili çıkmamıştır. Kredi kararı verilirken borçluluk oranı kriterinin önemli olması kabul edilebilir bir durumdur. Bunun sebebi verilecek kredinin geri dönüş olasılığının kişinin borçluluk oranı arttıkça azalacak olması beklentisinden kaynaklanmaktadır. Borçluluk oranı %15’e kadar kredi kararını olumlu etkileyebilirken, %15’den yüksek borçluluk oranına sahip olunması kredi kararını olumsuz etkilemektedir.



Şekil 3.15. BORCLULUK_ORANI ve kredi kararı arasındaki ilişki

BORCLULUK_ORANI değişkeni şube görüşünün bulunduğu durumda modele TF53-TF54 fonksiyonlarıyla girmiştir. Bir önceki paragrafta borçluluk oranı ve kredi kararı arasındaki ilişkinin ters yönlü olduğu ancak borçluluk oranının %15 seviyesine kadar modelde etkisinin olumlu olduğu söylenmiştir. Burada ise bir önceki duruma ilave olarak şube görüşü de bu değişkenle etkileşime girerek modelde yer almıştır. Şube görüşüyle birlikte kredi kararına olan etki Şekil 3.16'dan da görüldüğü gibi borçluluk oranının %15 seviyesine kadar pozitif ve borçluluk oranının %15'den yüksek olduğu durumlar için de negatif yönlü olarak gerçekleşmiştir. Bir önceki durumla benzer bir sonuç elde edilmiştir. Model katsayıları incelendiğinde; %15'den küçük borçluluk oranı kredi kararında -0,00383828'lik bir değişime neden olurken, %15'den yüksek borçluluk oranı kredi kararında -0,0325104 düzeyinde bir değişime neden olmaktadır. Buradan ulaşılabilecek sonuç; borçluluk oranı kabul edilebilir seviyede dahi olsa şube görüşünün o kişinin borçluluk durumu hakkında olumsuz bir görüş belirtmiş olmasıdır.



Şekil 3.16. Şube görüşünün bulunduğu durumda BORCLULUK_ORANI ve kredi kararı arasındaki ilişki

KKB_GUNCEL_ENKOTU_STATU ve onun için oluşturulan dummy değişken, diğer değişkenlerle birlikte modele dahil edilmiştir. Bu değişken kategorik olduğundan modele sadece ilgili koşullar sağlandığında dahil olmuştur.

KKB_GUNCEL_ENKOTU_STATU_dummy değişkeni borçluluk oranı ile birlikte TF13 temel fonksiyonuyla modele girmiştir. Borçluluk oranı varlığı altında başvurusu bulunan bireyin kredi kayıt bürosunda güncel en kötü statü bilgisinin bulunması kredi kararını 0.174057 etkilemektedir. KKB'den alınan bu bilgi, kişinin açık tüketici kredisi, kredi kartları, hesaplarına tanımlı ek para limitlerine ait geri ödemelerindeki aylık gecikme durumunu göstermektedir. Bu dummy değişkenin kredi kararına negatif bir etkisi olması beklenmektedir. Ancak güncel en kötü kayıt bilgisinde bu kötü kaydın düzeyi henüz bilinmediği için bu modelde etkisi negatif olmamıştır. Bu değişkene ait modele giren diğer terimler ile bu değişkenin kredi kararına toplam etkisi daha net görülebilecektir. Dolayısıyla borçluluk oranı gibi bu kriter de kredi kararının verilmesinde bir veridir ve modelde yer alması kabul edilebilir bir durumdur.

KKB_GUNCEL_ENKOTU_STATU değişkeni modele borçluluk oranı ve **KKB_GUNCEL_ENKOTU_STATU_dummy** değişkeni ile birlikte TF15 olarak modele dahil edilmiştir. Başvuru yapan bireyin borçluluk oranı varsa, kredi kayıt bürosunda güncel en kötü statüsü hakkında bilgisinin de bulunması beklenmektedir. Eğer bu her iki durum sağlandıysa ve kredi kayıt bürosundaki güncel en kötü statüsü "2", "3", "4", "5" ve "6" ise bu durum kredi kararında 0.297651'lik bir artışa sebep olmaktadır. Kredi kayıt bürosundaki bu statüler kişinin finansal kurumlardaki geri ödemelerine ait gecikme bilgisini ay olarak vermektedir. Yani modelde 6'ya kadar olan gecikmeler için kredi kararına olan etkinin negatif yönlü olmadığı görülmektedir.

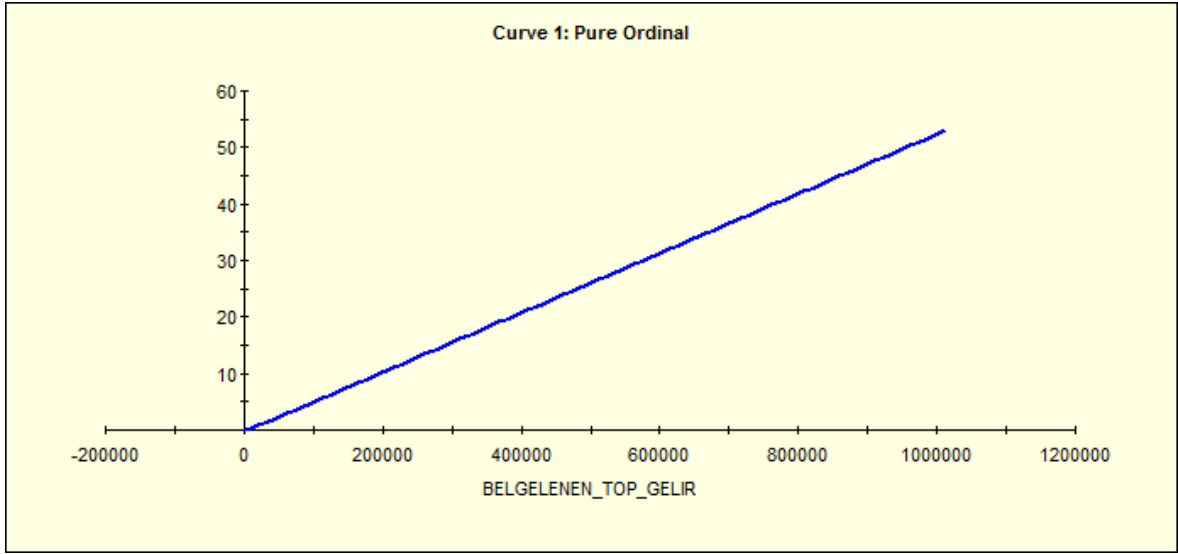
KKB_GUNCEL_ENKOTU_STATU_dummy değişkeni borçluluk oranı varlığında ve şube görüşünün belirtildiği durumda modele TF41 ile girmiştir. Kredi kayıt bürosunda güncel durumda en kötü statü bilgisi ve şube görüşünün bulunması kredi kararını -0.674358 etkilemektedir. Bu değişkenin kredi kararını negatif yönde etkilemesinin nedeni girilen şube görüşüdür. Şube görüşü olumsuz olarak belirtildiğinden bu koşullar altında modele giren değişkenin kredi kararına etkisi negatif olarak gerçekleşmiştir.

KKB_GUNCEL_ENKOTU_STATU değişkeni "2", "3", "4", "5", "6" ve "U" düzeylerinde borçluluk oranı ve şube görüşünün bulunduğu durumda modele

TF43 ile tekrar girmiştir. Değişkenin bu düzeyleri kişinin güncelde mevcut kredi taksitleri ya da kredi kartı borç ödemelerindeki aylık gecikmeyi, "U" ise belirsizlik durumunu göstermektedir. Dolayısıyla kredi kayıt bürosunda güncel riskleri içerisinde ödemelerini 2, 3, 4, 5, 6 ay geciktiren ve ya güncel en kötü kayıt bilgisinin belirsiz olması durumlarının kredi kararına etkisi olumlu gerçekleşmemiştir. Daha önce ödemelerdeki gecikmeler 6 aya kadar kabul edilebilir bir durumken, burada girilen şube görüşü ile birlikte kredi kararını olumsuz etkileyen bir durum haline gelmiştir. Şube görüşü olumsuz olduğundan kurulan modelde de bu değişkenin katsayısı -0,4162 olarak hesaplanmıştır.

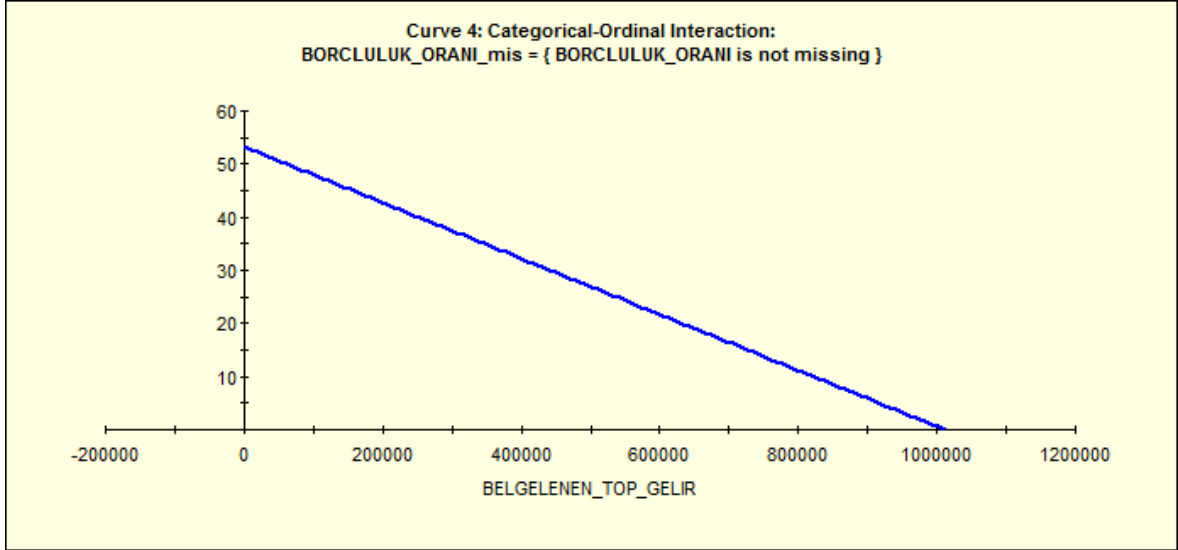
KKB_GUNCEL_ENKOTU_STATU değişkeni "0", "2", "3", "4", "5" ve "U" düzeylerinde borçluluk oranı varlığında ve şube görüşünün bulunmadığı durumda modele TF59 ile girmiştir. Bu değişkenin bu koşullar altında belirtilen düzeyleri kurulan modelde kredi kararında -0.0371366'lık bir değişime neden olmaktadır. Modele giren bu terimde şube görüşü bilgisi bulunmamaktadır. Daha önceki durumlarda da şube görüşü varlığında yine bu değişkenin kredi kararını olumsuz etkilediği gözlemlenmiştir. Burada bir önceki durumdan farklı olarak şube görüşünün bulunmamasının etkisi görülmüştür. Şube görüşü bulunmasa da borçluluk oranı ve kredi kayıt bürosundan alınan güncel en kötü kayıt bilgisi kredi kararını olumsuz etkilemektedir. Bir önceki durumdan tek fark, şube görüşünün bulunmamasının kredi kararına olan negatif etkinin daha az bir düzeyde gerçekleşmiş olmasıdır. Bahsedilen durum model katsayılarından açıkça görülmektedir.

BELGELENEN_TOP_GELIR ile kredi kararı arasında pozitif yönlü bir ilişki olması beklenmektedir. Çünkü başvuru yapan bireylerin gelirleri arttıkça kredibiliteleri de aynı oranda artmaktadır. Bu ilişkinin gösterildiği grafik Şekil 3.17'de verilmiştir. Bu değişken kurulan modele 1.150,0000 düğüm noktasında TF45-TF46 ile girmiştir. Tek bir düğüm noktasıyla modele dahil edilmesi bu değişken ile kredi kararı arasındaki ilişkinin doğrusal olduğunu ve bu değişkenin de dönüşüme uğramadan modelde yer aldığını göstermektedir. Kurulan modelde belgelenen gelir tutarı 1.150.-TL'den büyük olduğunda kredi kararına etkisi 0,0001 (5.2668E-005) ve 1.150.-TL'den küçük olduğunda 0.000230043 olmaktadır. Kurulan modelde de belgelenen gelir ile kredi kararı arasındaki ilişki pozitif yönlüdür.



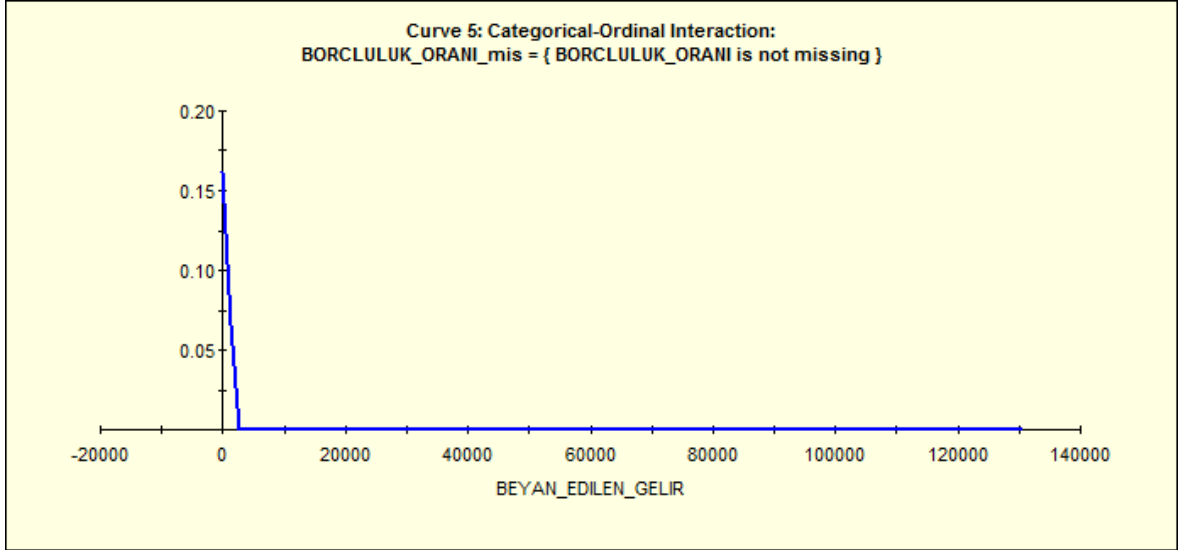
Şekil 3.17. BELGELENEN_TOP_GELIR ve kredi kararı arasındaki ilişki

BELGELENEN_TOP_GELIR değişkeni borçluluk oranının varlığında modele 1.785,9999 düğüm noktasında TF7 ile dahil edilmiştir. Borçluluk oranı bulunan kişilerin belgelenen toplam gelirleri 1.786.-TL'den fazla olduğunda kredi kararına olan etkisi -0,0001'dir (5.26293E-005). Belgelenen gelir ve kredi kararı arasındaki ilişki pozitif yönlü iken aksine burada belgelenen gelirin artması durumunda kredi kararına olacak etki negatif yönlü çıkmıştır. Bunun nedeni borçluluk oranının varlığıdır. Gelir ne olursa olsun borçluluk oranı değişmeyecektir, geliri 1.000.-TL olan kişi için de 10.000.-TL olan kişi için de borçluluk oranı %50 çıkabilir ancak borç tutarı değişecektir. Dolayısıyla verilerin elde edildiği o dönemde, kişinin borç miktarının belirli bir seviyesinden sonra, belgelenen gelir ile kredi kararı arasındaki ilişkiyi negatif yönde etkilenmektedir. Şekil 3.18'de borçluluk oranı koşulu altında belgelenen toplam gelir ile kredi kararı arasındaki ilişki görülmektedir.



Şekil 3.18. Borçluluk oranı varlığında BELGELENEN_TOP_GELİR ve kredi kararı arasındaki ilişki

BEYAN_EDILEN_GELİR değişkeni borçluluk oranı ile birlikte 2.232,6699 düğüm noktasında TF12 ile modele dahil edilmiştir. Buna göre; başvuruyu yapan bireyin başvuru sırasında beyan ettiği toplam gelir 2.232,67.-TL'den küçükse bunu kredi kararında olan etkisi pozitif yönlü ve 0,0001 (7.24259E-005) düzeyindedir. Şekil 3.19'den de görülebileceği gibi bu değişkenin 2.232,67 düğüm noktasından sonraki değerleri için kredi kararı ile arasındaki ilişki düzeyi sıfırdır ve beyan edilen gelirin bu tutardan sonra kredi kararına olan bir etkisi bulunmamaktadır. Beyan edilen gelir kişilerin kredi başvuruları sırasında beyan ettikleri toplam gelir tutarını göstermektedir. Dolayısıyla bu bilgilerin belgelenen gelire göre ne kadar güvenilir olduğu tartışılır. Kurulan bu modelde de beyan edilen gelir 2.232,67.-TL'den daha büyükse dikkate alınmamıştır. Beyan edilen gelirden çok belgelenen gelir bilgisinin kredi kararı verilirken önemli olması gerekmektedir.



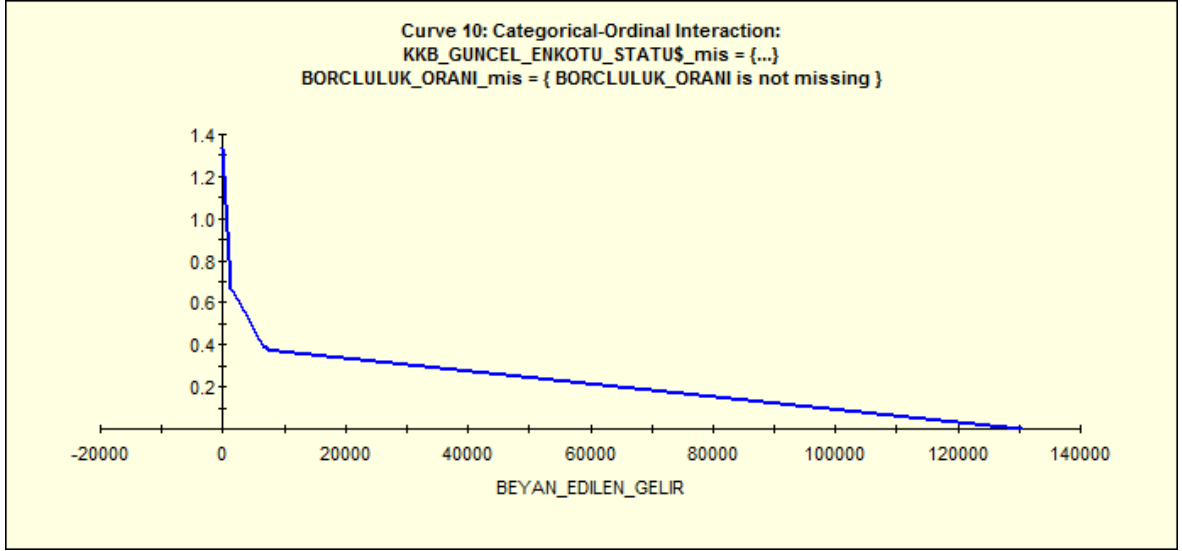
Şekil 3.19. Borçluluk oranı varlığında BEYAN_ EDILEN_ GELİR ve kredi kararı arasındaki ilişki

BEYAN_EDILEN_GELİR değişkeni 858,0000 düğüm noktasında borçluluk oranı varlığında ve KKB güncel en kötü statü bilgisinin bulunmadığı durumda modele TF27-TF28 ile dahil edilmiştir. Beyan edilen gelir ve kredi kararı arasındaki ilişkinin eğimi 858 düğüm noktasında değişmektedir. Bu değişim de kurulan modelde kredi kararında beyan edilen gelir 858.-TL'den büyük bir tutar olduğunda -0,0001 (5.14336E-005) ve bu tutardan küçük olduğu durumda da 0,0007 kadar olmaktadır.

BEYAN_EDILEN_GELİR değişkeni borçluluk oranı varlığında ve KKB güncel en kötü statüsü bulunmadığı durumda 7.000,0000 düğüm noktasında modele TF47 ile dahil edilmiştir. Buna göre bu koşullar altında beyan edilen gelir 7.000.-TL'den büyükse kurulan modelde bunun kredi kararına olan etkisi yok denecek kadar az, 0,0000 (4.83984E-005) düzeyinde olmaktadır.

Şekil 3.20'de borçluluk oranı varlığında beyan edilen gelir ve kredi kararı arasındaki ilişki görülmektedir. Yukarıda belirtilen koşullar altında bu grafikten beyan edilen gelir ve kredi kararı arasındaki ilişkinin negatif yönlü olduğu ve 828 ve 7.000 noktalarında bu ilişkinin eğiminin değiştiği görülmektedir. Borçluluk oranı hesaplanan bir müşteri için kredi kayıt bürosundan güncelde en kötü statüsü hakkında elde edilen bilgiler bulunmuyorsa, o kişinin ödemeleriyle ilgili mevcut

durumda bir problemin olmadığı söylenebilir. Ancak borçluluk oranı yüksek olduğu için ilave kredi alması ileriye dönük kredi geri ödemelerinde sorun teşkil edebileceğinden, beyan edilen gelirin artması kredi kararının olumlu yönde verilmesi için yeterli değildir. Burada müşterinin krediyi hangi amaçla kullanacağı daha çok önem teşkil edecektir.

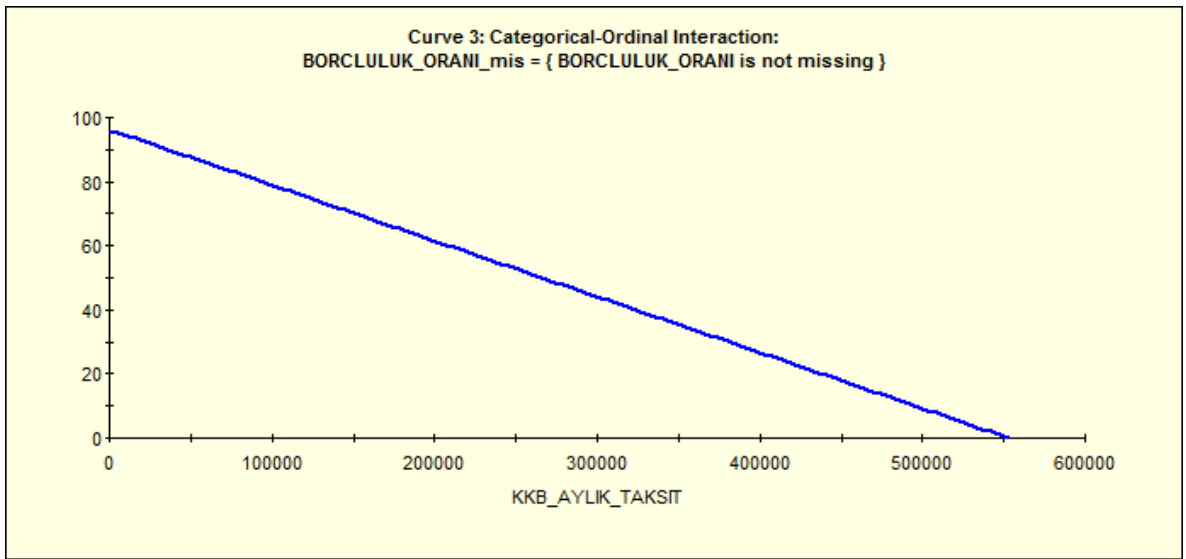


Şekil 3.20. Borçluluk oranı varlığında ve KKB güncel en kötü statü bilgisinin bulunmadığı koşula BEYAN_ EDILEN_ GELIR ve kredi kararı arasındaki ilişki

MAAS_MUSTERISI değişkeni borçluluk oranı varlığında modele “Y= evet” düğüm noktasında TF9 ile dahil edilmiştir. Bu değişken kurulan modelde kredi kararında -0.170077 değişikliğe sebep olmaktadır. Başvuru yapan bireyin maaş müşterisi olması finansal kurumun o müşteri hakkında diğer kurumlara göre daha fazla bilgi sahibi olması anlamına gelmektedir. Bunun dışında ilave hiçbir etki yaratmamıştır. Bunun nedeni müşteri o finansal kurumdan maaş alsa bile başka kurumlardan kredi kullanabileceğinden, müşterinin sabit gelirin o finansal kurum için bir güvencesi yoktur. Dolayısıyla başka bir finansal kurumda olası bir olumsuzluğun ortaya çıkması durumunda bankanın sorumluluğu diğer finansal kurum lehine maaş müşterisin varlıklarına bloke koymaktır.

KKB_AYLIK_TAKSIT değişkeni borçluluk oranı varlığında 1.491,9999 düğüm noktasında TF5-TF6 ile modele dahil edilmiştir. Şekil 3.21’den de görülebileceği gibi bu değişkenin kredi kararı ile olan ilişkisi negatif yönlü ve doğrusaldır. Kurulan

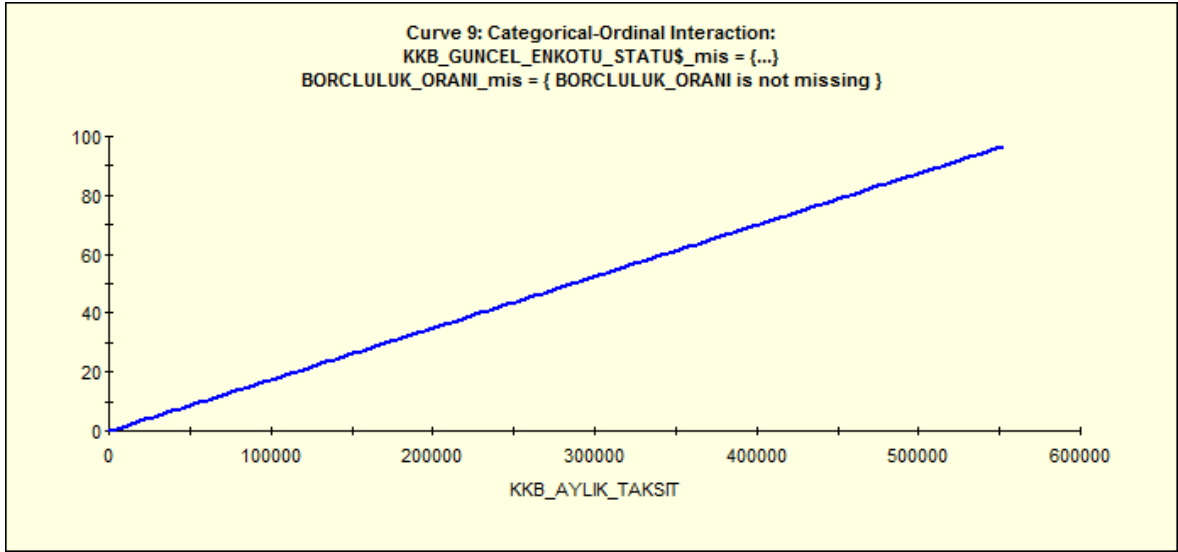
modelde 1.492'den büyük olması kredi kararında negatif yönlü -0.000174136 bir değişime neden olmaktadır. Benzer şekilde bu değişkenin 1.492'den küçük olduğu durumlarda yine kredi kararına etkisi negatif yönlü olarak gerçekleşmektedir (-0.0001282). Bu ilişkinin eğimi KKB aylık taksit tutarının 1.492.-TL'den küçük ve büyük olduğunda değişmektedir. KKB aylık taksit tutarı, başvuru yapan o müşterinin zaten borçlu olduğunu gösterdiğinden ve aylık ödemesi gereken taksit tutarının kredi kararına olumsuz olarak etkisinin bulunması beklenen bir durumdur ve kredi kararını olumsuz yönde etkilemesi beklenmektedir. Bu durum modelden de kanıtlanmış olmaktadır.



Şekil 3.21. Borçluluk oranı varlığında KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki

KKB_AYLIK_TAKSIT değişkeni borçluluk oranının bulunduğu ve KKB güncel en kötü statüsünün var olmadığı durumda 252,0000 düğüm noktasında TF21-TF22 temel fonksiyonlarıyla modelde yer almıştır. Belirtilen bu şartlar altında KKB'deki aylık taksit tutarı 252.-TL'den büyük olduğunda kredi kararını 0.000175163 ve küçük olduğunda da 0.00045513 etkilemektedir. Şekil 3.22'den de görülebileceği üzere kurulan modelde borçluluk oranı bulunan ancak kredi kayıt bürosunda güncel ödemesine ait bir gecikme bilgisi yer almayan başvuruların kredi kararına olan etkisi kredi kayıt bürosundaki aylık taksit toplamı arttıkça çok düşük de olsa artmaktadır. Normal şartlar altında ödenmesi gereken aylık taksit tutarı, yani borç artarken bunun kredi kararına olan etkisinin negatif olması beklenirken, burada

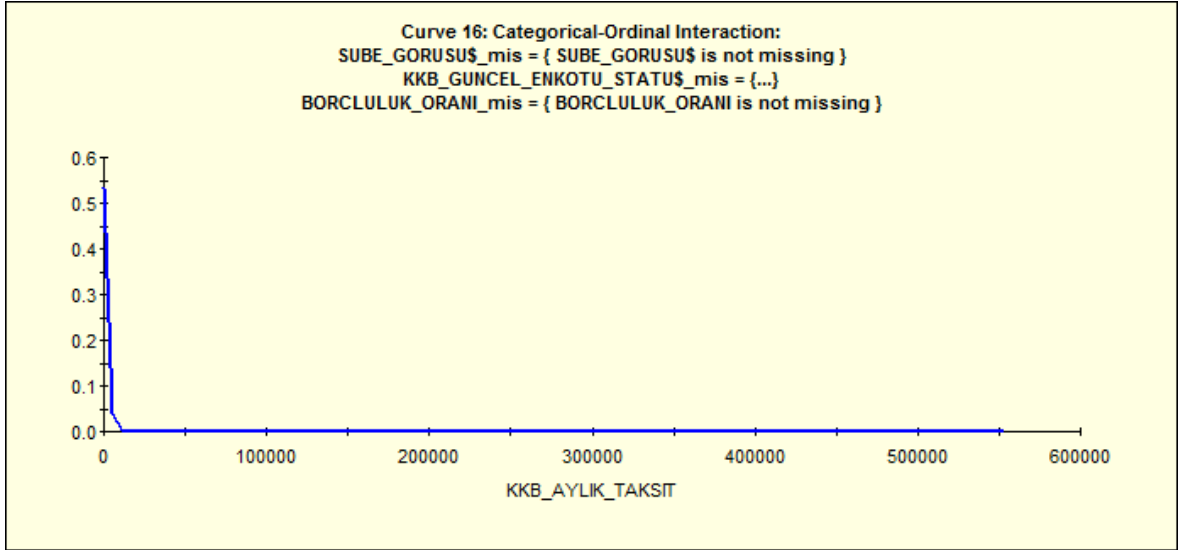
bahsedilen koşullar altında tam tersi gerçekleşmiştir. Birey borçlu olabilir ancak ödemeleriyle ilgili güncelde herhangi bir kötü kaydı bulunmuyorsa bu durumun kredi kararına olumlu bir etkide bulunmasının sebebi talep edilen kredilerin kullanım amacıyla ilgili olabilmektedir. Eğer kurulan modelde kullanım amacı borçların yeniden yapılandırılması ya da diğer banka borçlarının finanse edilmesi ise diğer koşulların da uygun olduğu durumlar için toplam borç tutarı arttıkça kredi kararına olan etki de artabilir yorumu yapılabilmektedir.



Şekil 3.22. Borçluluk oranı varlığında, KKB’de ödemelerine ait gecikmenin olmadığı durumda KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki

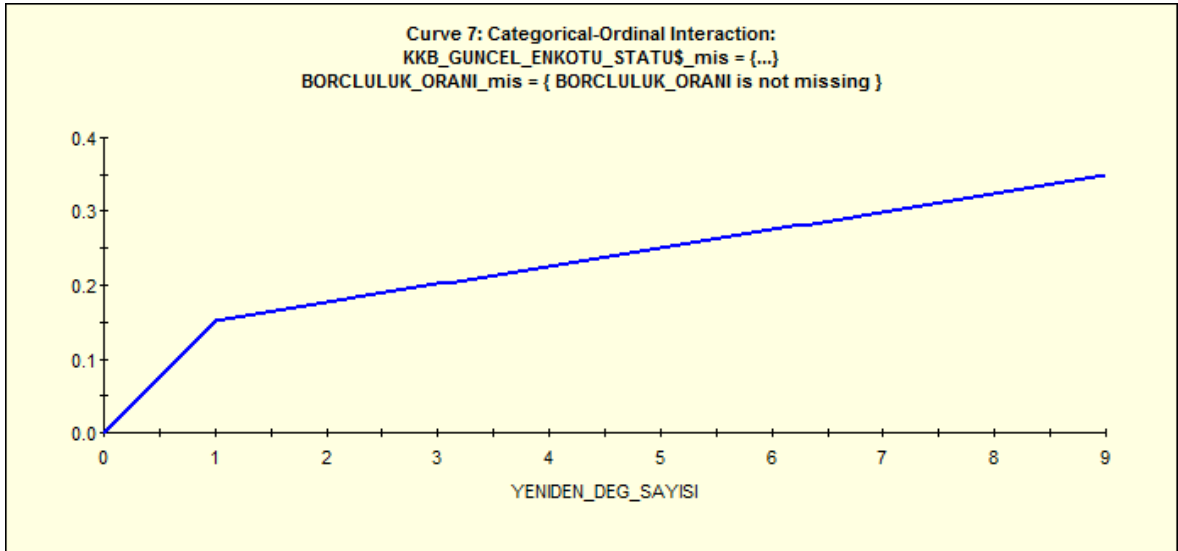
KKB_AYLIK_TAKSIT değişkeni borçluluk oranı varlığında, şube görüşünün bulunduğu ve KKB güncel en kötü statü bilgisinin bulunmadığı koşulda 6.042,9995 düğüm noktasında TF56 ile modele dahil edilmiştir. Bu koşullar altında kredi kayıt bürosundan alınan aylık taksit bilgisinin 6.042.-TL’den küçük olması kredi kararını 0,0001 (8.84942E-005) etkilemektedir. Burada da bir önceki durumdan farklı olarak şube görüşü bu değişkenle birlikte modele girmiştir. Ayrıca güncel en kötü ödeme bilgisinin bulunmaması durumunun etkisi de görülmektedir. Şube görüşünün de bulunması aylık taksit tutarı arttıkça kredi kararına olan olumlu etkinin bir önceki durumdan farklı olarak sınırlandırılmasına neden olmuştur. Dolayısıyla Şekil 3.23’den de görülebileceği gibi şube görüşünün de yer aldığı bu koşullar altında aylık taksit tutarı 6.042.-TL’ye kadar kredi kararını olumlu etkilemekteyken, bu tutardan büyük değerler için ise kredi kararına bir etkisi

kalmamıştır. Nedeni de kredi kararı için kabul edilebilecek maksimum toplam aylık taksit miktarının 6.042.-TL olarak kabul edilmiş olmasıdır.



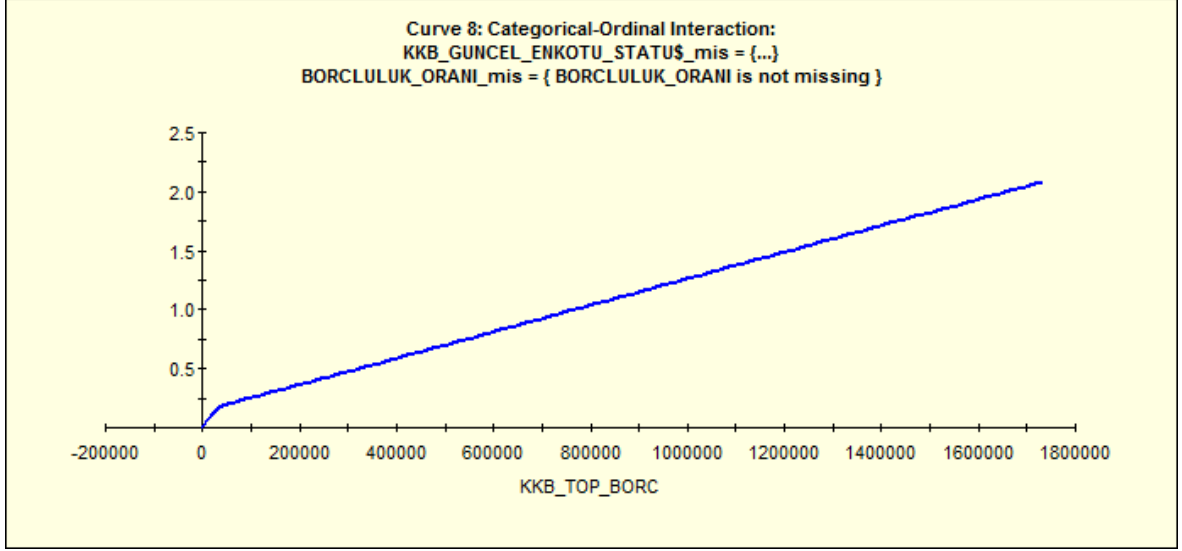
Şekil 3.23. Borçluluk oranı varlığında, KKB’de ödemelerine ait gecikmenin olmadığı ve şube görüşünün belirtildiği durumda KKB_AYLIK_TAKSIT ve kredi kararı arasındaki ilişki

YENIDEN_DEG_SAYISI değişkeni borçluluk oranı varlığında ve KKB güncel en kötü statü kaydının bulunmadığı durumda 1,0000 düğüm noktasında TF17-TF18 temel fonksiyonlarıyla modele dahil edilmiştir. Şekil 3.24’den de görülebileceği gibi bu değişken ve kredi kararı arasındaki ilişki doğrusaldır. Yeniden değerlendirme sayısı arttıkça kredi kararına olan etki artmaktadır ve bu ilişkinin eğimi yeniden değerlendirme sayısının 1 olduğu durumdan sonra değişmektedir. Kurulan modelde ise yeniden değerlendirme sayısı 1’den fazla olan başvuruların kredi kararına olan etkisi 0.0245812 olarak gerçekleşirken, yeniden değerlendirme sayısı 1’den azsa, yani yoksa bu durumda yeniden değerlendirme sayısı ile kredi kararı arasındaki ilişki -0.151998 olmaktadır. 1’den fazla değerlendirmede bu ilişkinin pozitif çıkmasının nedeni güncelde kredi kayıt bürosundaki en kötü ödeme bilgisinin yeniden değerlendirme süresi boyunca değişmemesi olarak yorumlanabilir. 1 defadan fazla yapılan değerlendirmeler için daha önce reddedilen başvurular ya da iade edilen başvuruların bu koşullar altında kredi kararına olan etkileri olumlu olmaktadır.



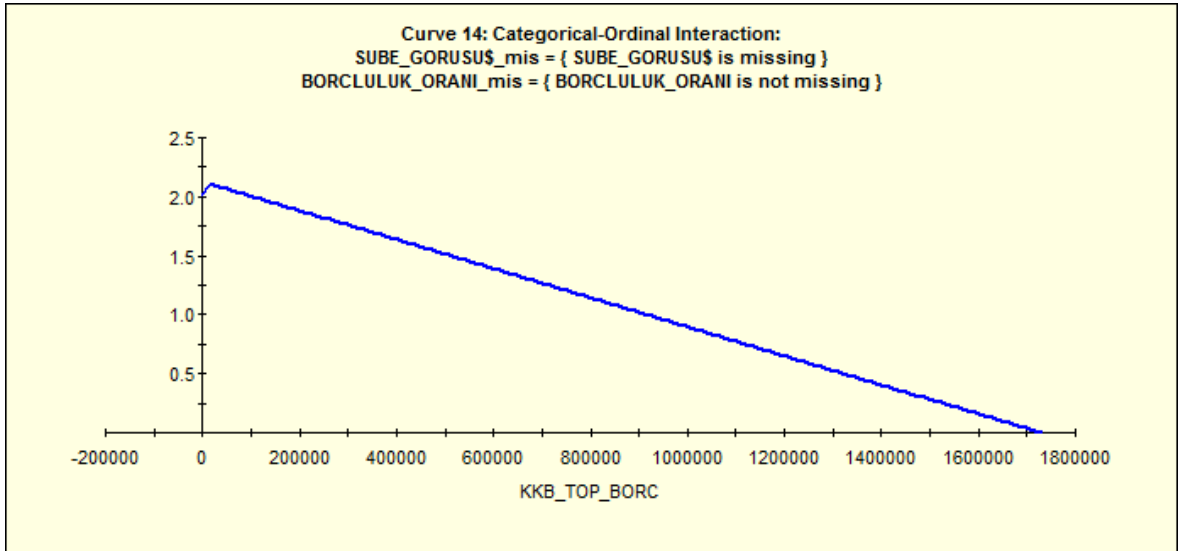
Şekil 3.24. Borçluluk oranı varlığında ve KKB güncel en kötü statü bilgisinin bulunmadığı durumda YENIDEN_DEG_SAYISI ve kredi kararı arasındaki ilişki

KKB_TOP_BORC değişkeni borçluluk oranı bilgisinin bulunduğu ve KKB güncel en kötü statü bilgisinin bulunmadığı durumda 29.821,0000 düğüm noktasında TF19-TF20 temel fonksiyonlarıyla modele dahil edilmiştir. Bu değişkenin kredi kararı ile arasındaki ilişki doğrusaldır ve modele hiç bir dönüşüme maruz kalmadan dahil edilmiştir. Şekil 3.25’den de görülebildiği gibi belirlenen bu koşullar altında kredi kayıt bürosundan alınan toplam borç tutarı ile kredi kararı arasındaki ilişki pozitif yönlüdür. Normal şartlar altında borç artarken kredi kararına olan etkisinin negatif olması beklenirken, bunun tam tersi gerçekleşmiştir. Bunun nedeni kredi kayıt bürosundan alınan güncel en kötü statü bilgisinin bulunmamasıdır. Kişi borçlu olabilir ancak güncelde ödeme durumuyla ilgili olumsuz bir bilgisi bulunmuyorsa toplam borç arttıkça kredi kararına olan olumlu etki de artabilir. Bu durumda bu kişilerin talep ettikleri kredinin kullanım amacı borçların yeniden yapılandırılması ya da diğer banka borçlarının finanse edilmesi olabilir. Kabul edilen modele bakıldığında ise bu koşullar altında KKB’deki toplam borcun 29.821.-TL’den büyük olması durumunda kredi kararına olan etkisi 0,0000 (1.12298E-006) ve 29.821.-TL’den küçük olduğu durumda da -0,0000 (- 5.799E-006) düzeylerinde, yok denecek kadar azdır. Dolayısıyla kabul edilen modelde belirtilen şartlar altında kredi kayıt bürosundaki toplam borcun kredi kararına etkisi çok düşüktür.



Şekil 3.25. Borçluluk oranı varlığında KKB güncel en kötü statü bilgisinin bulunmadığı durumda KKB_TOP_BORC ve kredi kararı arasındaki ilişki

KKB_TOP_BORC değişkeni borçluluk oranı varlığında ve şube görüşünün bulunmadığı durumda 10.782,0000 düğüm noktasında TF37-TF38 ile modele dahil edilmiştir. Şekil 3.26'dan da görülebileceği gibi bu koşullar altında kredi kayıt bürosundan alınan toplam borç değişkeni ile kredi kararı arasındaki ilişki negatif yönlü çıkmıştır. Kurulan modelde kredi kayıt bürosundaki tüm borçlar toplamının 10.782.-TL'den büyük olması kredi kararını -0,0000 (-1.22924E-006) etkilemektedir. Bu tutardan düşük olması durumunda ise kredi kararını yine negatif yönde -0,0000 (9.44826E-006) etkilemektedir. Burada kredi kararının verilebilmesi için kişinin borçluluk oranının bulunması dışında hiçbir yardımcı bilginin bulunmadığı bir durumla karşılaşmıştır. Şekil 3.26'da borçluluk oranı varlığında, şube görüşünün belirtilmediği durumda KKB_TOP_BORC ve kredi kararı arasındaki ilişki verilmiştir. Kişi borçludur ancak geri ödeme performansı ile ilgili bir belirsizlik söz konusudur. Dolayısıyla toplam borç arttıkça kredi kararına olan etkisi olumsuz olarak ortaya çıkmıştır.



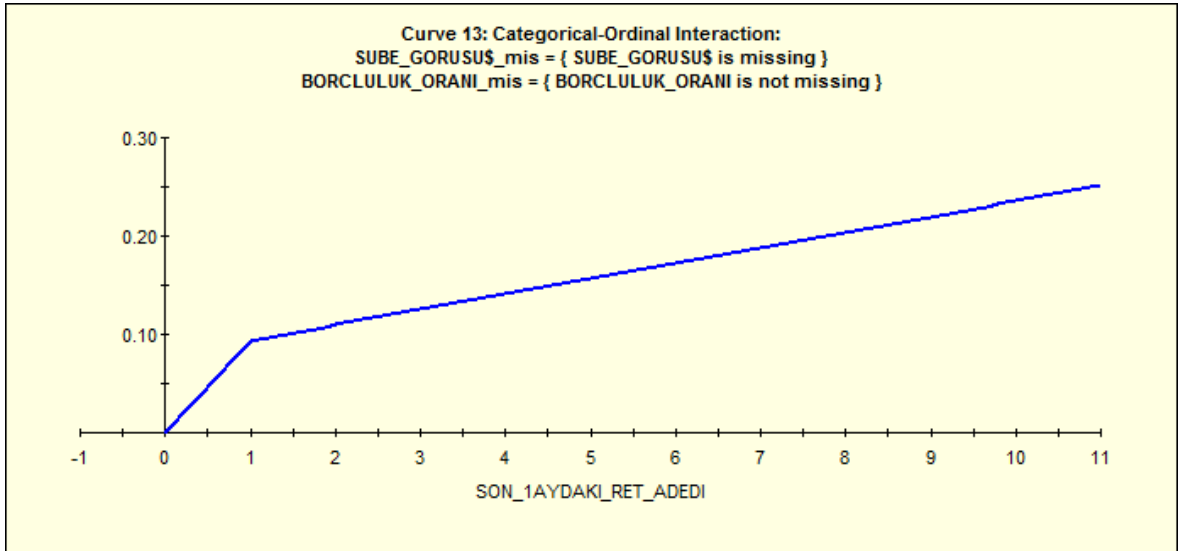
Şekil 3.26. Borçluluk oranı varlığında, şube görüşünün belirtilmediği durumda KKB_TOP_BORC ve kredi kararı arasındaki ilişki

KKB_SON_6AYDAKI_EN_KOTU_ODEME değişkeni borçluluk oranı varlığında, güncelde en kötü ödemesi bilgisinin var olması ve son 6 aydaki en kötü ödemesinin 2, 3, 4, 5, 6, 8 olduğu durumda TF25 ile gösterilmiş ve modele girmiştir. Bu terimin modeldeki katsayısı 0,122874 olarak bulunmuştur. Kredi kararına olan bu pozitif etki günceldeki en kötü ödemenin belli olmamasıdır. Kişi eğer özel bir nedenden ötürü borcunu ödemiyeceyse, geçmişten gelen kötü kayıtlar güncelde de var olacaktır. Şöyleki; küçük tutarlı ya da unutulmuş bir kredi kartı borcu varsa kişi farkında olmadığından güncelde de kötü kaydı olacaktır ya da finansal kurumla ortaya çıkan anlaşmazlıktan ötürü ödenmeyen bir borç varsa yine benzer şekilde bu durum da kötü kayda sebep olacaktır. Böyle bir durum varlığında kredi kararının olumlu etkilenmesi kabul edilebilir bir durumdur.

KKB_SON_6AYDAKI_EN_KOTU_ODEME değişkeni borçluluk oranı varlığında ve KKB'deki güncel en kötü statüsünün bulunduğu durumda "0", "1", "2", "4", "5", "6", "L" düzeylerinde TF61 ile modele girmiştir. Buna göre borçluluk oranı var olan ve güncelde kredi kayıt bürosunda kötü statüsü bulunan ve son altı ayda da kredi taksitlerine ait gecikmelerin 6'ya kadar olduğu ve hatta idari takibe düştüğü durumun kredi kararına etkisi, kurulan modelde -0,0344 olarak hesaplanmıştır. Az önceki durumdan farklı olarak burada ödenmeyen bir kredi borcunun idari takibe

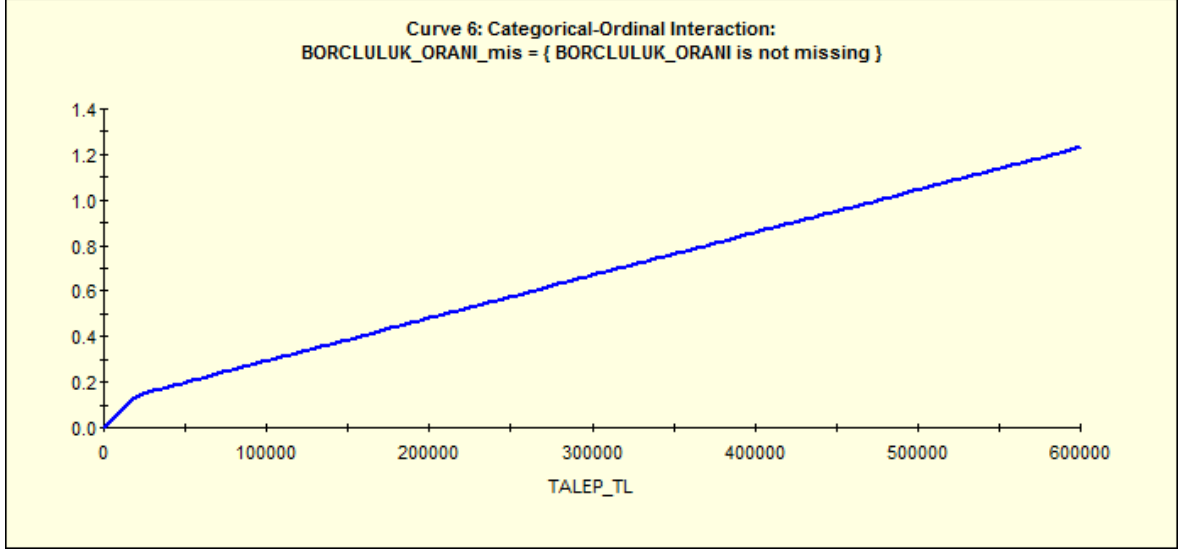
düşmesi (“L=idari takip”) durumu bulunmaktadır ve bu da kredi kararını bir önceki duruma göre negatif etkilemesine neden olmuştur.

SON_1AYDAKI_RET_ADEDI değişkeni borçluluk oranı varlığında ve şube görüşünün bulunmadığı durumda 1,0000 noktasında TF35-TF36 temel fonksiyonlarıyla modele dahil edilmiştir. Buna göre şubenin görüş bildirmediği ancak son bir aydaki ret adedi birden fazla olması durumu kredi kararında 0.0157406 kadar bir değişime sebep olmaktadır. Burada daha önce ret edilen başvuruların kurulan modelde kredi kararına etkisinin pozitif yönde gerçekleştiği görülmektedir. Son 1 aydaki ret adedi 1’den azsa -0.0937448 düzeyinde bir etkisi olmaktadır. Daha önce ret edilen başvuruların tekrar değerlendirilip olumlu bir sonuçla sonuçlanması mümkündür. Burada önemli olan daha önce neden ret edildiğidir. Başvuruların elde edildiği o zaman dilimi içerisinde müşterinin gelir düzeyi gibi kriterlerinde değişiklikler olabileceği gibi finansal kurumun da o dönemde politikası değişmiş ve ürünlerini yönlendirdiği müşteri profili daha önce ret edilen müşteri profiline yaklaşmış olabilir. Bu başvuruların kriterleri artık uygun koşulları sağladığından tekrar değerlendirilmeye alınmış olabilir. Şekil 3.27’de belirtilen koşullarda kredi kararı ve SON_1AYDAKI_RET_ADEDI arasındaki ilişki görülmektedir. Her müşteri için şube görüşü oluşturulmak zorunda değildir. Sisteme girilen başvurular belli kriterleri sağlanmadığında, otomatik olarak ret edilebilir. Böyle bir durumda şube görüşü oluşmamaktadır. Şartlar değiştiğinde de sistem tekrar otomatik olarak bu başvuruyu olumlu olarak değerlendirebilmektedir.



Şekil 3.27. Borçluluk oranı varlığında, şube görüşünün belirtilmediği durumda SON_1AYDAKI_RET_ADEDI ve kredi kararı arasındaki ilişki

TALEP_TL değişkeni borçluluk oranı varlığında 19.999,9980 düğüm noktasında TF49-TF50 ile modele dahil edilmiştir. Şekil 3.28'den de görülebileceği gibi bu değişkenin kredi kararına olan etkisi doğrusaldır ve pozitif yönlüdür. İlişkinin eğimi talep edilen tutarın 20.000.-TL olmasından sonra değişmektedir. Kurulan modelde ise bu değişkenin 20.000.-TL'den küçük olduğu tutarlarda kredi kararına olan etkisi -0,0000 (-7.15759E-006), büyük olduğu tutarlarda ise 0,0000 (1.8785E-006) olarak gerçekleşmiştir. İhtiyaç kredisi talepleri genellikle bireylerin aylık gelirleriyle doğru orantılıdır. İhtiyaç kredisine olan talebin yüksek bir tutarda olması o bireyin yüksek gelir düzeyine sahip olduğu şeklinde yorumlanabilir. Dolayısıyla talebi yüksek olan kişinin bu krediyi ödeyebilecek miktarda maddi güce sahip olduğu düşünülerek kredi kararına talep tutarının kredi kararına olumlu bir katkı yaratacağı kabul edilebilir bir durumdur.



Şekil 3.28. Borçluluk oranı varlığında TALEP_TL ve kredi kararı arasındaki ilişki

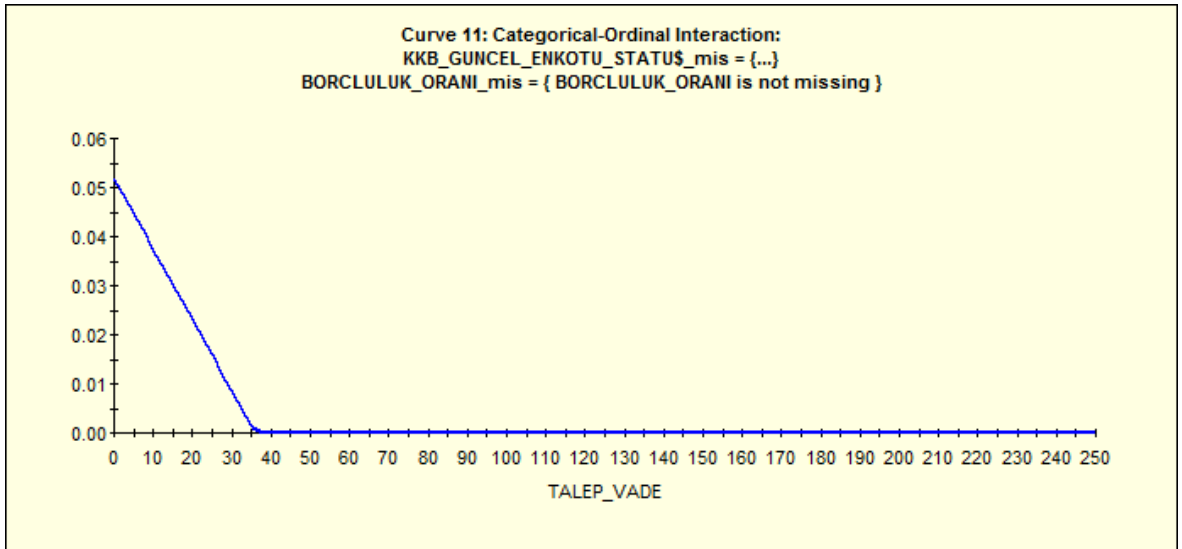
KKB_GECMIS_ENKOTU_STATU değişkeni borçluluk oranı varlığında, KKB’de güncel en kötü statüsünün de bulunduğu durumda “L=kanuni takip” düğüm noktasında TF31 ile modele dahile dirmiştir. Bu koşullar altında kredi kayıt bürosundan alınan geçmiş en kötü statü değişkeni kredi kararını 0.104 etkilemektedir. Daha önce kanuni takibe düşen bir kredi borcunun bulunmasının kredi kararına olumlu bir etkisinin bulunması ancak ve ancak o kötü kaydın kapanmış bir krediye ait olması durumunda gerçekleşmiş olabilir

CALISMA_SEKLI değişkeni borçluluk oranının bulunduğu durumda "1010=Çalışmıyor", "1040=Kamu Ücretli Çalışan", "1080=Yan Gelir Sahibi" noktalarında modele TF51 ile girmiştir. Bu değişkenin kurulan modelde kredi kararına etkisi -0,0553 düzeyinde gerçekleşmiştir. Varolan bir borcu bulunan ve çalışmayan ya da kamuda ücretli çalışan ya da yan gelir sahibi olduğu belirtilen başvurularda çalışma şekilleri kredi kararını olumsuz yönde etkilemektedir. Burada kamu ücret düzeylerinin düşüklüğü göz önüne alındığında bu durum kabul edilebilir bir durumdur.

KULL_AMACI değişkeni borçluluk oranı varlığında ve şube görüşünün bulunmadığı durumda "12=borç yapılandırma" ve "13=motorsiklet alımı" düğüm noktalarında TF39 ile modele dahil edilmiştir. Kurulan modelde kredinin kullanım amacının borç yapılandırma ve motorsiklet alımı olması kredi kararını 0.0779919

etkilemektedir. Kurulan modelde borç yapılandırmasının önemli çıkması daha önce mevcut taksit ve borç tutarlarının yüksek olması durumlarına da bir açıklık getirmiştir.

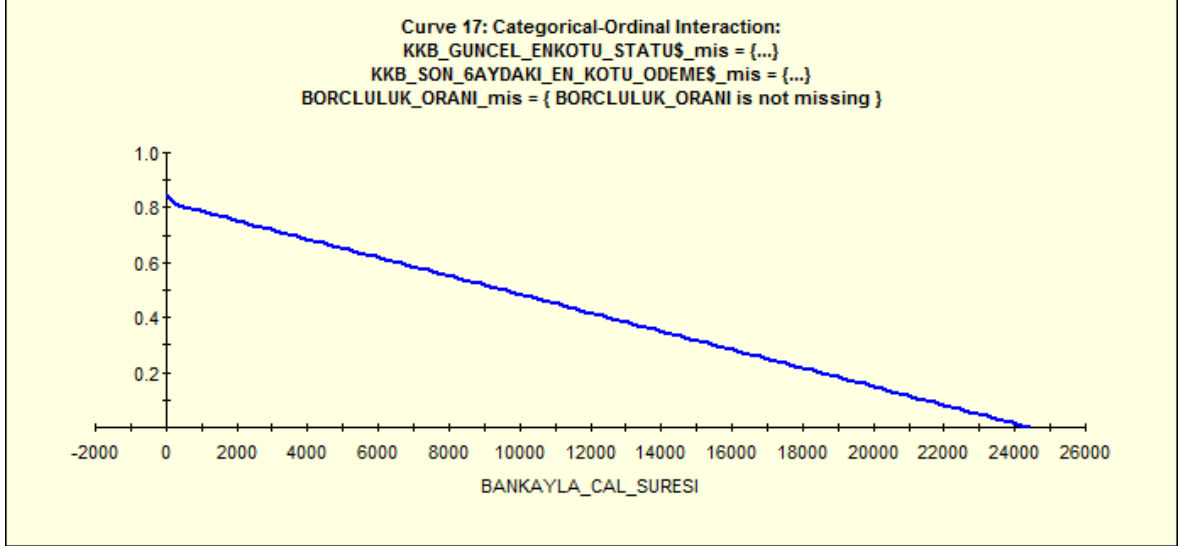
TALEP_VADE değişkeni borçluluk oranı varlığında ve KKB'deki güncel en kötü statüsünün bulunmadığı durumda 36,0000 düğüm noktasında TF66 temel fonksiyonu ile modele dahil edilmiştir. Talep edilen vadenin 36 aydan düşük olması kurulan modelde kredi kararında 0,0014 değişikliğe sebep olmaktadır. 36 aydan büyük vadelerde bu ilişki sıfır çıkmıştır. Bunun anlamı mevcut koşullar altında talep edilen vadenin 36 aya kadar değerlendirmeye alınmasıdır. Şekil 3.29'da borçluluk oranı varlığında ve KKB'de güncel en kötü statü bilgisinin bulunmaması durumunda TALEP_VADE ile kredi kararı arasındaki ilişki görülebilmektedir.



Şekil 3.29. Borçluluk oranı varlığında ve KKB'de güncel en kötü statü bilgisinin bulunmaması durumunda TALEP_VADE ile kredi kararı arasındaki ilişki

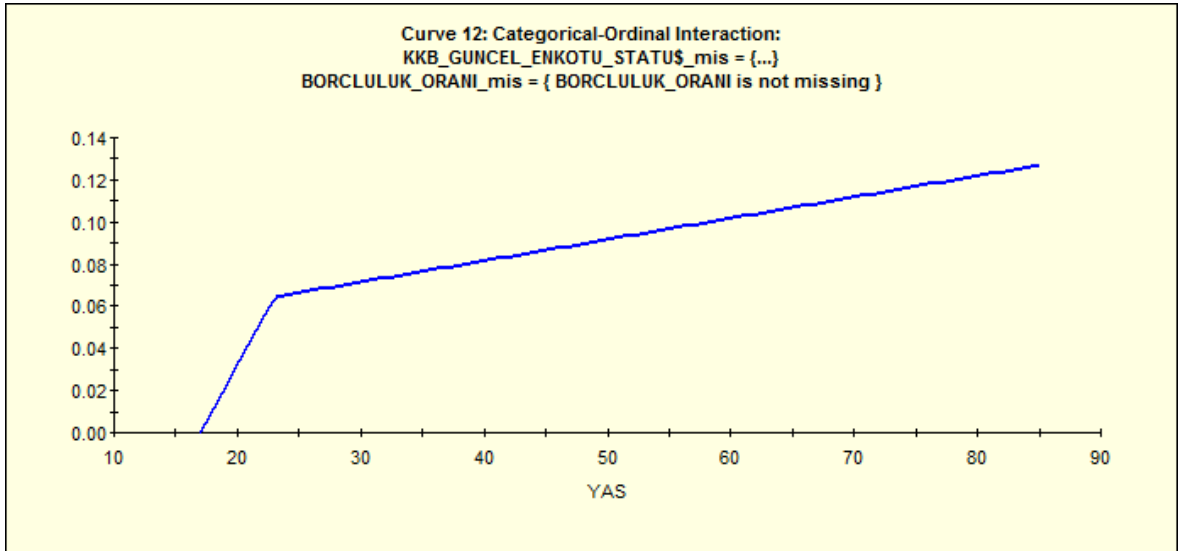
BANKAYLA_CAL_SURESI değişkeni borçluluk oranı varlığında ve KKB'deki son 6 aya ait en kötü ödeme bilgisinin bulunmadığı ve güncelde de yine en kötü ödemesine ait bir bilginin bulunmadığı durumda 28,0000 düğüm noktasında TF63-TF64 ile modele dahil edilmiştir. Şekil 3.30'da bu koşullar altında bankayla çalışma süresi ile kredi kararı arasındaki ilişki görülmektedir. Bu ilişki negatif yönlü ve doğrusaldır. Bankayla çalışma süresi 28 günden sonra bu ilişkinin eğimi değişmektedir. Uzun süre banka ile çalışan müşterilerin sürekli banka kaynaklarını

kullanarak hayatlarını devam ettirdikleri, yani dış bir kaynağa sürekli olarak bağımlı olduklarını göstereceği yorumu yapılabilir. Dolayısıyla ilave bir kredi taleplerinin olumlu karşılanmaması söz konusu olmaktadır.



Şekil 3.30. Borçluluk oranı varlığında ve KKB'deki son 6 aya ait en kötü ödeme bilgisinin ve günceldeki en kötü ödemesine ait bir bilginin bulunmadığı durumda BANKAYLA_CAL_SURESI ve kredi kararı arasındaki ilişki

YAS değişkeni borçluluk oranı varlığında ve güncelde KKB'deki en kötü ödeme bilgisinin bulunmadığı durumda 23,0000 düğüm noktasında modele dahil edilmiştir. Kurulan modelde bu koşullar altında bu değişkenin kredi kararına olan etkisi yaşı 23'ten büyük olması durumunda 0,0010 ve küçük olması durumunda ise -0,0107 olarak gerçekleşmiştir. Şekil 3.31'den de görülebileceği gibi yaşı artması ile kredi kararı arasındaki ilişki doğrusal ve pozitifdir. Bu ilişki 18 yaş üstünde ortaya çıkmıştır ve 23'den sonra eğimi değişmiştir. Yani 18-23 yaş arası kredi başvurusu yapılabilse bile bu kredilerin geri ödemesi büyük bir olasılıkla kişinin kendi geliri ile değil, ailesinden aldığı finansal destekle yapılacaktır. Finansal kurum için bu durum güvenilir bir kaynak değildir.



Şekil 3.31. Borçluluk oranı varlığında ve KKB'deki güncel en kötü ödeme bilgisinin bulunmadığı durumda YAS ve kredi kararı arasındaki ilişki

3.4. İkili Lojistik Regresyon Çözümlemesi

Çalışmanın bu bölümünde bir önceki bölümde MARS modelinin kurulumunda kullanılan verilere İkili Lojistik Regresyon Analizi uygulanacaktır. Lojistik regresyon analizinde değişkenlerin normal dağılım göstermeleri varsayımı gerekmektedir, ayrıca açıklayıcı değişkenler kesikli veya sürekli, nicel veya kategorik değişkenler olabilmektedir. Bu özelliklerinden dolayı lojistik regresyon yöntemi ile kurulan model, MARS modeli ile karşılaştırılabilecektir. Bu çalışmada bağımlı değişkenin ikili cevap içeren bir yapıda olması (“kredi kararı: kabul/ret”) nedeniyle verilere ikili lojistik regresyon analizi uygulanacaktır.

MARS çözümlemesi bölümünde belirlenen 35 bağımsız değişken ikili lojistik regresyon çözümlemesine de aynen dahil edilmiştir. MARS yazılımından farklı olarak SPSS’de bir değişkende hem sayısal hem de alfabetik değerlerin bulunmasına izin verilmediğinden, sayısal kodlama yapılarak çözümlemeye yer alması sağlanmıştır. İkili lojistik yöntemi “ENTER” seçilerek program çalıştırılmıştır ve buna göre elde edilen modelin önem düzeyi Şekil 3.32’de görülebilmektedir. Burada $p < 0,05$ olarak hesaplandığından kurulan model anlamlıdır. Kurulan modelin Cox & Snell R^2 değeri 0,221 ve Nagelkerke R^2 değeri de 0,311 olarak hesaplanmıştır. Bu R^2 değerleri Şekil 3.33’de görülebilmektedir.

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10631,828	112	,000
	Block	10631,828	112	,000
	Model	10631,828	112	,000

Şekil 3.32. SPSS’de kurulan modelin önem düzeyi

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	41998,758 ^a	,221	,311

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Şekil 3.33. SPSS’de kurulan modelin R² değeri

Kurulan modelin kredi kararına göre sınıflandırmadaki başarısı Şekil 3.34’de verilmiştir. Kredi kararının kabul olduğu durumlar için %92,6 ve ret edildiği durumlar için de %43,9 oranında doğru sınıflandırma yapılmıştır ve buna göre toplam doğru sınıflandırma oranı %77,6 olarak gerçekleşmiştir.

Classification Table

Observed			Predicted		
			KREDİ KARARI 2		Percentage Correct
			AC	DL	
Step 1	KREDİ_KARARI_2	AC	27313	2190	92,6
		DL	7362	5760	43,9
Overall Percentage					77,6

a. The cut value is ,500

Şekil 3.34. SPSS’de kurulan modelin sınıflandırma başarısı

Kurulan model Ek7’de verilmiştir. Aşağıda modele giren değişkenlerin kredi kararına olan etkileri ve ilgili yorumlar yer almaktadır.

BELGELENEN_TOP_GELIR değişkeni modele -0,0008 katsayısı ile girmiştir. Bu değişkenin kredi kararına etkisi olumlu olmamıştır. Belgelenen gelir kredi kararı verilirken tek başına değerlendirilmemelidir. Bunun sebebi gelir arttıkça borçluluk oranının da artabileceği gerçeğidir. Bu nedenle de kurulan modelde bu değişkenin etkisi negatif olarak gerçekleşmiştir.

YENIDEN_DEG_SAYISI’nın kurulan modelde kredi kararına olan etkisi 0,6439’dür. Yeniden değerlendirme sayısındaki artışın, kredi kararına olumlu bir etkisi bulunmaktadır. Bunun 2 nedeni olabilmektedir. Birincisi yeniden değerlendirilen başvurularda bazı kriterlerin kabul edilebilir duruma gelmesi, örneğin kişinin gelir durumunun düzelmesi olabilir. İkinci neden ise finansal kurumun o dönemdeki kredi politikasının değişmesi olabilir. Dolayısıyla bu değişkenin kredi kararına olan pozitif yönlü etkisi kabul edilebilir bir durumdur.

MAAS_MUSTERISI(1) değişkeni, yani o finansal kurumun maaş müşterisi olmama durumu, maaş müşterisi olma durumuna göre kredi kararında 0,7963’lük bir değişime sebep olmaktadır. Verilerin temin edildiği zaman aralığı içerisinde eğer başka banka risklerinin alınması gibi bir durumun varlığı sözkonusu ise bu kabul edilebilir bir durumdur.

BORCLULUK_ORANI değişkeni kurulan modelde, kredi kararında -0,0824 düzeyinde bir değişikliğe sebep olmaktadır. Borçluluk oranının artması durumunda verilecek kredi kararının olumsuz etkilenmesini göstermektedir.

KKB_TOP_BORC değişkeninin kredi kararına etkisi çok düşüktür(0,0000). Kurulan modelde toplam borç miktarından ziyade diğer değişkenlerin önemli çıkması beklenmektedir. Eğer borç miktarının etkisi çok düşükse, geçmiş ödeme bilgilerinin daha önemli olması beklenmektedir. Birey borçlu olabilir ancak ödemeleri düzgünse borç miktarının kredi kararına etkisi olmayabilir.

BEYAN_EDILEN_GELIR, kurulan modelde kredi kararında -0,0003 değişime neden olmaktadır. Beyan ettiği gelir yüksek olan kişilerin borçluluk durumları yüksekse ya da borçlarını ödeme performansları kötüyse böyle bir durum ortaya çıkabilir.

SON_1AYDAKI_RET_ADEDI değişkeni modelde kredi kararını 0,3438 etkilemektedir. Bu durumun açıklanabilmesi için o başvurunun neden ret edildiğinin bilinmesi gerekmektedir. Başvuru yapan kişinin kriterlerindeki iyileşme ya da finansal kurumun kredi verme stratejilerinde herhangi bir değişiklik olduysa o başvurunun kabul edilmesi durumları ortaya çıkabilmektedir.

TALEP_TL değişkeni kredi kararını 0,0000 etkilemektedir. Talep edilen TL tutarın kredi kararında önemli derecede bir değişikliğe sebep olmamaktadır. Talep edilen tutar, kişiye ait diğer kriterler sağlandığında onaylanmakta ya da reddedilmektedir yorumu yapılabilir.

KKB_AYLIK_TAKSIT değişkeni modele 0,0003 katsayısı ile dahil edilmiştir. Bireyin aylık taksit tutarındaki artış kredi kararında olumsuz bir etki yaratmamıştır. Bireyin gelir durumu, geri ödeme performansı, talep ettiği krediyi kullanım amacı gibi kriterler göz önüne alındığında yapılan başvuru onaylanabilmektedir. Bu değişken tek başına olumsuzluk göstergesi değildir.

KKB_SON_6AYDAKI_EN_KOTU_ODEME2 değişkeni kişinin son 6 ayda ödemelerindeki gecikme miktarını göstermektedir. Bu değişken sırasıyla 0, 1, 2, 3 ve 4 aylık gecikmeler olarak modele -0,3651, -0,3433, 0,2990, 0,5205 katsayıları ile girmiştir. Bireye ait son 6 aydaki ödeme gecikme sayısı arttıkça modelde kredi kararına olan etki de pozitif yönlü ve daha güçlü olmuştur.

KKB_SKOR değişkeni de kredi kararı verilirken önemli bir kriter olarak çıkmıştır. Bu değişken kurulan modele 0, 1, 3 ve 4 düzeylerinde sırasıyla -0,1972, -0,4237, -0,2608, -0,1445 katsayıları ile girmişlerdir. KKB'den alınan skor değerleri 0-7 arasında değişmektedir ve en iyi skor değeri 7'dir. Dolayısıyla kurulan modelde skor değerleri 0, 1, 3, ve 4 olan başvurularda kredi kararına olumsuz bir etkilerinin bulunması kabul edilebilir bir durumdur.

YAS değişkeni kredi kararında 0,0112 değişikliğe sebep olmaktadır.

OTURULAN_EV_DURUMU değişkeni kurulan modele dahil edilmiştir. “K = kiralık” ise -0,1642 etkilemektedir.

KKB_GUNCEL_ENKOTU_STATU değişkeni bu değişkenin sırasıyla 2, 3, 4, 5 ve 6 olduğu durumlarda 1,6818, 1,4877, 1,6439, 1,9303, 1,8865 katsayıları ile modele girmiştir. Kredi kayıt bürosundan elde edilen ödemelerdeki en kötü gecikme miktarı 2 ay’dan 6 aya çıktıkça kredi kararına olan etkisi de pozitif yönlü olmaktadır ve artış göstermektedir.

BASVURU_SKORU değişkeni kurulan modele 1, 2, 3, 4 ve 5 olduğu durumlarda -0,2167, -0,3521, -0,1283, 0,1072, 0,1196 katsayıları ile dahil edilmiştir. Başvuru sırasında üretilen bu skor değeri 1-6 arasında değişmektedir ve en iyi skor değeri 6’dır. Dolayısıyla skor değeri arttıkça kredi kararına olan etkinin de artması beklenmektedir.

KKB_GECMIS_ENKOTU_STATU değişkeninin kurulan modelde kredi kararına olan etkisi olumsuzdur. Geçmişteki en kötü ödeme bilgisinin 1-6 arasında olması kredi kararını sırasıyla -0,5641, -0,4766, -0,5521, -0,6251, -0,5735, -0,8754, -0,7562 etkilemektedir.

CINSIYET değişkeninin modeldeki katsayısı -0,1366 olarak çıkmıştır. Dolayısıyla kredi kararına kadınlara göre erkeklerin etkisi negatif olarak bulunmuştur.

SOSYAL_GUVENCE değişkeni de kurulan modelde yer almıştır. Bu değişken için “Y = yeşil kart” düzeyi referans olarak belirlenmiştir. Başvuruda sosyal güvence bilgisinin bulunmaması, yeşil kart sahibi olunmasına göre kurulan modelde kredi kararını -1,6817 ile olumsuz etkilemektedir. Eğer sosyal güvence “B = bağkur” ise bu durumun kurulan modelde kredi kararına etkisi -0,2403 olmaktadır. Sosyal güvenci tipi “E=emekli” olan başvurular kredi kararında -0,2730 değişime neden olmaktadır. Sosyal güvencesi “S = sığorta” olan başvuruların, kurulan modelde kredi kararına olan etkisi -0,3378 olarak gerçekleşmiştir. Görüldüğü gibi kredi

kararını olumsuz olarak en çok etkileyen sosyal güvence durumu sosyal güvencenin bulunmamasıdır.

TALEP_VADE değişkeni kurulan modelde kredi kararında -0,0047 değişime neden olmaktadır. Talep edilen vade arttıkça kredi kararı bu durumdan olumsuz etkilenmektedir. Vade arttıkça kredinin geri dönüş süresi o kadar artacağından, bu kabul edilebilir bir durumdur.

KULL_AMACI değişkeni 12 düzeyinde 0,4201 katsayısı ile modele dahil edilmiştir. Dolayısıyla kullanım amacı borç yapılandırma olarak çıkmıştır. Bu durumda kullanım amacının borç yapılandırma olması kredi kararını olumlu etkilemektedir. Başvuruların yapıldığı dönemde kişilerin mevcut borçlarının yapılandırılması ya da başka başka finansal kurum risklerinin alınması gibi durumlar söz konusudur.

SON_1AYDAKI_ONAY_ADEDI ile kredi kararı arasındaki ilişki pozitif yönlüdür. Kurulan modele göre son 1 ayda verilen onay adedi arttıkça kredi kararında 0,4233 değişime neden olmaktadır.

EGITIM_DURUMU değişkeni modele -1,2017 katsayısı ile dahil edilmiştir. Kurulan modelde eğitim durumunun “D = doktora” olması, “L = lise” olmasına göre kredi kararında -1,2017 değişime neden olmaktadır.

EV_TELEFON_BILG değişkeni modele -0,0973 katsayısı ile birlikte dahil edilmiştir. Buna göre ev telefon bilgisini finansal kurumla paylaşmayan bireylerin, başvurularında, bu bilgiyi paylaşanlara göre kredi kararında -0,0973 değişime neden olmaktadır. Ev telefon bilgisinin bulunması, bir güven kriteri olarak önemli sayılmış olabilir.

MUSTERI_TIPI değişkeni modelde 0,1515 katsayısı ile yer almaktadır. Buna göre, müşteri tipi bilgisinin bulunmaması durumu kredi kararında 0,1515 değişime neden olmaktadır. Borç yapılandırması nedeniyle bankaya gelen yeni müşteriler için kredi başvurusu yapıldığı durumda bu değişken ile kredi kararı arasındaki ilişki kabul edilebilir olmaktadır.

CALISMA_SEKLI “1050” düzeyinde modele girmiştir. Bu değişkenin referans düzeyi yan gelir sahibi olunmasıdır. Dolayısıyla çalışma şeklinin öğrenci olarak belirtilmesi, yan gelir sahibine göre kredi kararında 1,7381 kat değişime neden olmaktadır.

BANKAYLA_CAL_SURESI değişkeni kurulan modelde -0,0007 katsayısı ile yer almaktadır. Bankayla olan çalışma süresindeki artış kredi kararına olumsuz yönde bir etkiye neden olmaktadır.

3.5. Genel Sonuçlar

Uygulamanın 3. ve 4. bölümlerinde ihtiyaç kredisi başvuru verileri kullanılarak kredi kararının kabul/ret olma durumlarına göre MARS modellemesi ve İkili Lojistik Regresyon modellemesi yapılmıştır. Bu bölümde kurulan bu iki model sonuçları karşılaştırılacaktır.

Her iki model de istatistiksel olarak anlamlı çıktığından kabul edilmiştir. Bu aşamada modellerin karşılaştırılması için doğru sınıflandırma oranına ve I. Tip hata ile II. Tip hata oranlarına bakılarak hangi modelin tercih edilebileceğine karar verilecektir.

Kredi kararı değişkeni içerisindeki kabul sayısı 30.202 ve ret sayısı da 34.798’dir. Önceki bölümlerde çalışılan veri setindeki kabul oranının %46,46 ve ret oranı %53,54 olarak hesaplandığı söylenmiştir. Kurulan MARS modelinin doğru sınıflandırma oranı %85,04 olarak hesaplanmıştır. İkili lojistik regresyon modelinde ise toplam doğru sınıflandırma oranı %77,60 olarak gerçekleşmiştir. Buna göre MARS modeli toplam doğru sınıflandırma oranına göre daha başarılı bir model olarak kabul edilmiştir.

I. Tip Hata, doğru bir yokluk hipotezinin yanlışlıkla reddedilmesi olasılığıdır ve $(1 - \text{I. Tip Hata})$ testin güvenilirlik düzeyini göstermektedir. II. Tip Hata, yokluk hipotezi yanlış olduğu zaman, bu hipotezin reddedilme olasılığıdır. Bu tez çalışmasında incelenen durum için I. Tip Hata onaylanan bir kredinin kurulan

modelde yanlışlıkla reddedilmesini, II. Tip Hata ise reddedilen bir kredinin yanlışlıkla kabul edilmesini göstermektedir. Finansal kurumlar için I. Tip Hata oranı daha kabul edilebilir bir durumdur. En kötü ihtimalle kredi verilmesi gereken bir müşteriye kredi verilmeyecektir, ancak daha sonra yapılacak değerlendirmelerde ya da başka kampanyalarda o kişiye kredi kartı ya da kredi gibi ürünler verilerek o müşteriyle olan ilişkiler güçlendirilebilecektir. Ancak II. Tip Hata finansal kurumlar için kabul edilemeyecek bir durumdur. II. Tip Hata'sı yüksek olan modellerde kredi verilmemesi gereken kişilere kredi verilmesi durumu söz konusudur. Bu da o finansal kurum için kaynakların boşa harcanması anlamına gelmektedir. Çizelge 3.5'de kurulan her iki model için I. Tip ve II. Tip hatalar verilmiştir.

MARS Modeli					İkili Lojistik Regresyon Modeli			
	Kabul		Ret		Kabul		Ret	
	N	%	N	%	N	%	N	%
Kabul	29.349	0,9718	853	0,0282	27.313	0,9258	2.190	0,0742
Ret	8.870	0,2549	25.928	0,7451	7.362	0,5610	5.760	0,4390

Çizelge 3.5. Kurulan modellere ait I. Tip Hata ve II. Tip Hata

Buna göre MARS modelinin I. Tip Hata oranı 0,0282 iken ikili lojistik regresyon modelinin I. Tip Hata oranı 0,0742 olarak bulunmuştur. Mars modelinin II. Tip Hata oranı 0,2549 ve ikili lojistik regresyon modelinde ise bu 0,5610 olarak hesaplanmıştır. Yapılan başvurunun kabul edilmesi için I. Tip Hata MARS modelinde daha düşük çıkmıştır. Böylece MARS modelinin kredi kararının kabul edilmesi durumunu daha iyi modellediği söylenmektedir. Kurulan bu modelin kredi kararını kabul etmede güvenilirlik düzeyi (1-0,2549) olarak gösterilebilir ve MARS modelinin kredi kararını kabul etmedeki güvenilirlik düzeyi ikili lojistik regresyon modeline göre daha iyi olduğu söylenebilir. II. Tip Hata oranına göre de MARS modeli lojistik regresyon modeline göre daha başarılı bir sonuç vermiştir. Dolayısıyla Lojistik regresyon modeline göre MARS modeli kredi kararının reddedilmesi için daha iyi karar vermektedir.

Kurulan her iki modeldeki değişkenlere bakıldığında genelde benzer değişkenlerin modellere dahil edildiği ve bu değişkenlerin kredi kararına olan etkilerinin benzer

şekilde olduğu söylenebilir. İkili lojistik regresyon modeline MARS modelinden farklı olarak **KKB_SKOR**, **OTURULAN_EV_DURUMU**, **BASVURU_SKORU**, **CINSIYET**, **SOSYAL_GUVENC**, **SON_1AYDAKI_ONAY_ADEDI**, **EGITIM_DURUMU**, **EV_TELEFON_BILG**, **MUSTERI_TIPI** değişkenleri dahil edilmiştir. Başvuru skoru, kredi kayıt bürosunun ürettiği skor bilgilerinin modelleme yapılırken kullanılıyor olması beklenen bir durumdur. Ancak MARS modelinde bu değişkenlerin kredi kararına olan etkileri görülememiştir.

MARS modelinde kayıp değerlere sahip olan değişkenlerin de kredi kararına etkileri görülebilmektedir. Ancak bu çalışmada ikili lojistik regresyon modelinden kayıp değerlerin kredi kararına ne tür bir etkiye bulunduğu hakkında herhangi bir yorum yapılamamış, sadece var olan değerler üzerinden değişkenlerin etkileri hakkında genel bir yorum elde edilmiştir. Lojistik regresyon yönteminde değişkende bulunan kayıp veri yerine o değişkenin ortalamasının konulması ya da başka geliştirilmiş yöntemler bulunmaktadır. Ancak MARS yöntemi hiç ilave bir çaba sarfetmeden kayıp verileri ele almakta ve daha kolay bir şekilde bağımlı değişkene olan etkilerini göstermektedir.

MARS modelinde değişkenler ve kredi kararı arasındaki ilişkilerin hangi noktalarda değişim gösterdiği hakkında detay bilgilere ulaşılabilmektedir. Bu durum da çalışan kredi kararı mekanizmasının değişkenin hangi düzeylerinde nasıl çalıştığı hakkında daha çok bilgi sahibi olunabilmesini sağlamaktadır. YAS değişkeni üzerinden bir örnek verilirse; MARS modelinde bu değişken 23 düğüm noktasında iki temel fonksiyonla modele dahil edilmiştir ve kurulan modeldeki katsayısı 23'den büyük olduğu durumlar için 0,0010 ve küçük olduğu durumlar için ise -0,0107'dir. Bu bilgiler kullanılarak kredi kararı verilirken yaşın etkisinin 23'den büyük olması ve küçük olması durumlarında değişiklik gösterdiği, hatta o finansal kurumun 23 yaş altı bireylere kredi vermek istemediği şeklinde yorumlanabilir. İkili lojistik regresyon modelinde YAS değişkeni 0,0112 katsayısı ile yer almıştır. Buna göre sadece yaş büyüdükçe kredi kararına olan etkisinin olumlu olduğu yorumuna ulaşılmaktadır.

DÖRDÜNCÜ BÖLÜM

4. SONUÇ

Hazırlanan tez çalışmasında Çok Değişkenli Uyarlamalı Regresyon Uzanımları (MARS) yönteminin gelişimini, kuramsal ve matematiksel yapısı, uygulama alanları, literatürdeki yeri ve bilgisayar ortamındaki kullanımı açısından incelenmiştir.

Çalışmada çok değişkenli uyarlamalı regresyon uzanımları yönteminin son dönemde hızla popülerleşmesine, ekonomi, biyoloji, tıp, ziraat, pazarlama gibi çok farklı alanlarda kullanıldığına değinilmiştir. Yöntemin kullanımı üzerine literatüre yakın zamanda yapılmış araştırmalardan bahsedilmiş, elde edilen bulgular konusunda notlar düşülmüştür.

Çok değişkenli uyarlamalı regresyon uzanımları yönteminin teorik yapısı detaylı olarak incelenmiştir. Yöntemin temelinde yer alan regresyon uzanımları, temel fonksiyonlar gibi kavramlar açıklanmıştır. Model kurulumunun nasıl gerçekleştiği adım adım anlatılmış, yöntemin algoritmaları verilmiş ve modellemelerin bu algoritmalara göre nasıl gerçekleştirildiği gösterilmiştir.

Çok değişkenli uyarlamalı regresyon modellerinin uygulaması için kullanılan MARS yazılımı, özellikleri ve kullanımı özetlenmiştir. MARS yazılımının kullanım aşamaları, modelleme sonuçlarının elde edilmesi ve elde edilen sonuçların nasıl yorumlanabileceği konusunda detay bilgi verilmiştir.

Yöntemin uygulanabilirliğinin gösterilebilmesi ve geleneksel istatistiksel yöntemlerle kıyaslanabilmesi amacıyla gerçek veri uygulaması gerçekleştirilmiştir. Uygulama için Türkiye'deki bir finans kuruluşundan elde edilen 65.000 adet bireysel kredi başvuru verisi kullanılmıştır. Kredi kararının kabul ya da reddedilmesi durumuna göre kategorik modelleme yapılmıştır. Kurulan MARS modelinin kıyaslanabilmesi için ikili lojistik regresyonu ile de modelleme yapılmıştır. Veri setindeki toplam 48 değişkenden 35 tanesi alınarak kredi kararı

modellenmiştir. Oluşturulan MARS modelinin doğru sınıflandırma oranı ve I. Tip Hata açısından ikili lojistik regresyon modeline göre daha başarılıdır.

Kurulan MARS modelinde kredi kararını etkileyen değişkenler borçluluk oranı, kredi kayıt bürosundan alınan güncel en kötü statü, belgelenen toplam gelir tutarı, beyan edilen gelir tutarı, maaş müşterisi olma durumu, kredi kayıt bürosundaki toplam aylık taksit tutarı, şube görüşü, kredi kayıt bürosundaki güncel en kötü kayıt bilgisi, yeniden değerlendirme sayısı, kredi kayıt bürosundaki toplam borç tutarı, kredi kayıt bürosundaki son altı aya ait en kötü ödeme bilgisi, son 1 aydaki ret adedi, talep edilen TL tutar, kredi kayıt bürosundaki geçmiş en kötü kayıt bilgisi, çalışma şekli, kredinin kullanım amacı, talep edilen vade, bankayla çalışma süresi ve yaş olarak belirlenmiştir. MARS modelinde ikili lojistik regresyon yöntemi ile kurulan modelden farklı olarak şube görüşü değişkeni yer almıştır. Şube görüşünün kredi kararının belirlenmesindeki etkisi önemlidir. Müşterileri daha iyi tanıyan şubenin görüşleri kredi kararını etkilemektedir ve nasıl etkilediği gösterilmiştir.

İkili lojistik regresyon modelinde de MARS modelinde farklı olarak başvuru skoru, kredi kayıt bürosu skoru, cinsiyet, ev telefon bilgisi, müşteri tipi, oturulan evin durumu ve sosyal güvence değişkenleri modele girmiştir. Skor değerlerinin kredi kararı verilirken dikkate alınan önemli bir kriter olması beklenmektedir. Ancak bu değişkenlerin kurulan en iyi MARS modelinde yer almaması, kayıp değişkenler için oluşturulan temel fonksiyonların değişkenlerle etkileşime girerek en son modelin buna göre oluşturulmasından kaynaklanmaktadır. Modelde oluşturulan dummy değişkenler önemli çıkmıştır ve bu değişkenlerin önem dereceleri skor bilgilerine göre daha yüksektir. Buna ek olarak, ikili lojistik regresyonda da kayıp değerlerin kredi kararına olan etkileri gözlemlenememiştir.

Kurulan her iki modelde de yer alan değişkenlerin kredi kararına etkileri benzer şekilde gerçekleşmiştir. Ancak MARS modeli kullanılarak bağımsız değişkenlere ait daha detay yorumlar yapılabildiği gösterilmiştir.

Tez çalışmasına, elde edilen tüm bulgular işlenmiş ve sonuçlar paylaşılmıştır.

KAYNAKLAR

- Akaike, H., 1980, Likelihood and Bayes procedure. In Bayesian Statistics (Bernardo J. M. et al. eds.). University Press, Valencia, Spain, 143-166.
- Breiman L., Friedman J.H., Olshen R.A., and Stone C.J., 1984, Classification and Regression Trees, Wadsworth, Belmont, CA, 358p.
- Briand L.C., Freimut B., Vollei F., 2004, Using multiple adaptive regression splines to support decision making in code inspections. The Journal of Systems and Software, 73, 205–217.
- Cervellera C., Wen A., Chen V.C.P., 2007, Neural network and regression spline value function approximations for stochastic dynamic programming. Computers & Operations Research, 34, 70–90.
- Chou S.M., Leeb T.S., Shao Y.E., Chen I.F., 2004, Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications 27 (2004) 133–142.
- Craven, P. and Wahba, G., 1979, Smoothing noisy data with spline functions. Estimating the correct degree of smoothing by the method of generalized cross-validation. Numerische Mathematik, 31, 317-403.
- Davison A.C., 1986, Approximate predictive likelihood, Biometrika, 73, 323-332.
- Deconinck E., Coomans D., Vander Heyden Y., 2007, Exploration of linear modelling techniques and their combination with multivariate adaptive regression splines to predict gastro-intestinal absorption of drugs. Journal of Pharmaceutical and Biomedical Analysis, 43(1), 119-130.

- Deconinck E., Xu Q.S., Put R., Coomans D., Massart D.L., Vander Heyden Y., 2005, Prediction of gastro-intestinal absorption using multivariate adaptive regression splines. *Journal of Pharmaceutical and Biomedical Analysis*, 39, 1021–1030.
- Deichmann J., Eshghi A., Haughton D., Sayek S. and Teebagy N., 2002, Application of multiple adaptive Regression splines (MARS) in Direct response modeling. *Journal Of Interactive Marketing*, 16(4), 15-27.
- Denison D.G.T., Holmes C.C., Mallick B.K., Smith A.F.M., 2002, *Bayesian Methods for Nonlinear Classification and Regression*, John Wiley and Sons, ISBN 0471490369, 9780471490364, 227p.
- Eubank R.L., 1988, *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, ISBN 0824778693, 9780824778699, New York, 438p.
- Francis L., 2000, *Martian Chronicles: Is MARS better than Neural Networks?* Salford Systems, Resources, Publications, <http://salford-systems.com/doc/03wf027.pdf>.
- Friedman J.H., 1991a, Multivariate Adaptive Regression Splines. *Annals of Statistics*, 19(1), 1-141.
- Friedman J.H., 1991b, Discussion. *Technometrics*, 33, 145-148.
- Gooijer J.G. De, Ray B.K., Krager H., 1998, Forecasting exchange rates using TSMARS. *Journal of International Money and Finance*, 17, 513-534.
- Green P.J. and Silverman B.W., 1994, *Non-parametric Regression and Generalized Linear Models*, Chapman and Hall/CRC Press, ISBN 0412300400, 9780412300400, London, 182p.
- Green P.J., 1995, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, 82, 711-732.

- Hastie T., Tibshirani R., Friedman J.H., 2001, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, ISBN 0387952845, 9780387952840, 533p.
- Hastie T.J. and Tibshirani R.J., 1990, *Generalized Additive Models*, Chapman & Hall/ CRC Press, ISBN 0412343908, 9780412343902, London, 335p.
- Ko M., Osei-Bryson K., 2004, The productivity impact of information technology in the healthcare industry: an empirical study using a regression spline-based approach. *Information and Software Technology*, 46, 65–73.
- Kuhnert P.M., Do K., McClure R., 2000, Combining non-parametric models with logistic regression: an application to motor vehicle injury data. *Computational Statistics & Data Analysis*, 34, 371-386.
- Lee T., Chiu C., Chou Y., Lu C., 2006, Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, 50, 1113 – 1130.
- Lee T.S., 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.
- Mallick B.K., Denison D.G.T., and Smith A.F.M., 1999, Bayesian Survival analysis Using a MARS model. *Biometrics*, 55(4), 1071-1077.
- McCullagh P. and Nelder J.A., 1989, *Generalized linear models*, CRC Press, ISBN 0412317605, 9780412317606, London, 511p.
- Osei-Bryson K., Ko M., 2004, Exploring the relationship between information technology investments and firm performance using regression splines analysis. *Information & Management*, 42, 1–13.

- Sakamoto W., 2007, MARS: Selecting Basis and Knots With the Empirical Bayes Method, *Computational Statistics*, 22(4), 583-597.
- Salford Systems, 2001, MARS User Guide, San Diego, CA,
<http://www.salfordsystems.com/mars.php>.
- Sephton P., 2001, Forecasting Recessions: Can We Do Better on MARS? Federal Reserve Bank Of St. Louis, March/April 2001, 39-49.
- Sephton P., 1994, Cointegration Tests on MARS. *Computational Economics*, 7(1), 23-35.
- Temel G.O., Çamdeviren H., Yazıcı A.C., 2005, Regresyon Modellerine Alternatif Bir Yaklaşım: MARS. VIII. Ulusal Biyoistatistik Kongresi, 20-22 Eylül 2005, Bursa, s.105-123.
- Tierney L. and Kadane J.B., 1986, Accurate approximations for posterior moments and marginal densities., *Journal of the American Statistical Association*, 81(393), 82-86.
- Tunay K.B., 2001, Türkiye’de paranın gelir dolaşım hızlarının MARS yöntemiyle tahmini. *ODTÜ Gelişme Dergisi*, 28 (3-4), 431-454.
- Xu Q.S., Daszykowski M., Walczak B., Daeyaert F., Jonge de M.R., Heeres J., Koymans L.M.H., Lewi P.J., Vinkers H.M., Janssen P.A., Massart D.L., 2004, Multivariate adaptive regression splines—studies of HIV reverse transcriptase inhibitors. *Chemometrics and Intelligent Laboratory Systems*, 72, 27– 34.
- Xu Q.S., Massart D.L., Liang Y.Z., Fang K.T., 2003, Two-step multivariate adaptive regression splines for modeling a quantitative relationship between gas chromatography retention indices and molecular descriptors. *Journal of Chromatography A*, 998, 155–167.

Yang C., Prasher S.O., Lacroix R., Kim S.H., 2003, A Multivariate Adaptive Regression Splines Model for Simulation of Pesticide Transport in Soils. Biosystems Engineering, 86(1), 9–15.

Yerlikaya F., Weber G.W., Taylan P., Batmaz İ., Köksal G., 2008, Mars Algoritmasında Tikhonov Düzenlemesi Ve Çok Amaçlı Optimizasyon Kullanımı.

Proceedings of Operational Research and Industrial Engineering Annual Conference, Galatasaray University, Istanbul, Turkey,
<http://www3.iam.metu.edu.tr/iam/images/0/09/Marstikh2008.pdf>

EKLER DİZİNİ

- Ek 1. Çalışmada Kullanılan Verilerin Tanımlayıcı İstatistikleri
- Ek 2. Çalışmada Kullanılan Verilerin Sıklık Grafikleri
- Ek 3. Mars Modelindeki Parametrelerin Özet Bilgileri
- Ek 4. Mars Yöntemi İleriye Doğru Adımında Modele Giren Temel Fonksiyonlar
- Ek 5. Mars Final Modeli
- Ek 6. MARS Final Modelini Oluşturan Temel Fonksiyonların Özellikleri
- Ek 7. İkili Lojistik Regresyon Modeli

EK 1: Çalışmada Kullanılan Verilerin Tanımlayıcı İstatistikleri

	N	N Kayıp Değer	N = 0	N ≠ 0	N Farklı Değer
ID	65000	0	0	65000	Many values
EGITIM_DURUMU	65000	0			5
KULL_AMACI	65000	0	0	65000	13
EV_TELEFON_BILG	65000	0			2
ODEME_PLANI_TIPI	65000	0	0	65000	8
OTURULAN_EV_DURUMU	65000	0			4
YENIDEN_DEG_SAYISI	65000	0	56140	8860	10
YENI_MEVCUT_MUSTERI	65000	0			2
TOPLAM_CAL_SURESI	65000	0	4845	60155	477
MAAS_MUSTERISI	65000	0			2
BANKAYLA_CAL_SURESI	65000	0	8745	56255	272
SIMDIKI_ISYERI_CAL_SURESI	65000	0	2008	62992	438
MUSTERI_TIPI	58993	6007			2
ILISKI_SKORU	65000	0	14551	50449	1263
SEKTOR_TIPI	65000	0	6513	58487	52
MUSTERI_BANKA_SEGMENTI	58972	6028			15
KOTU_KAYIT_TURU	3495	61505			5
CALISMA_SEKLI	65000	0	0	65000	10
SUBE_GORUSU	1843	63157			4
SOSYAL_GUVENCE	64952	48			6
BEYAN_EDILEN_GELIR	65000	0	7	64993	22543
SON_1AYDAKI_RET_ADEDI	65000	0	58169	6831	12
KREDI_KARAR_KODU	65000	0			69
SON_1AYDAKI_ONAY_ADEDI	65000	0	64693	307	4
KKB_AYLIK_TAKSIT	65000	0	30211	34789	3159
KKB_GUNCEL_ENKOTU_STATU	57896	7104			10
KKB_GECMIS_ENKOTU_STATU	60824	4176			9
MEDENI_DURUM	65000	0			4
KREDI_KARARI	65000	0			6
KKB_TOP_BORC	65000	0	12630	52370	20446
COCUK_SAYISI	65000	0	55670	9330	7
KKB_SON_6AYDAKI_EN_KOTU_ODEME	60824	4176			12
CINSIYET	65000	0			2
ONAYLANAN_VADE	65000	0	0	65000	42
YAS	65000	0	0	65000	66
BELGELENEN_TOP_GELIR	65000	0	57726	7274	1509
BORCLULUK_ORANI	44508	20492	1758	42750	19
KULLANDIRIM_DURUMU	65000	0	43447	21553	2
BASVURU_SKORU	65000	0	0	65000	6
KKB_SKOR	65000	0	9430	55570	7
TALEP_TL	65000	0	154	64846	471
TALEP_DOVIZ	65000	0	64984	16	12
TALEP_VADE	65000	0	154	64846	47
ONAYLANAN_LIMIT	65000	0	0	65000	503
KREDI_KARARI_2	65000	0			2

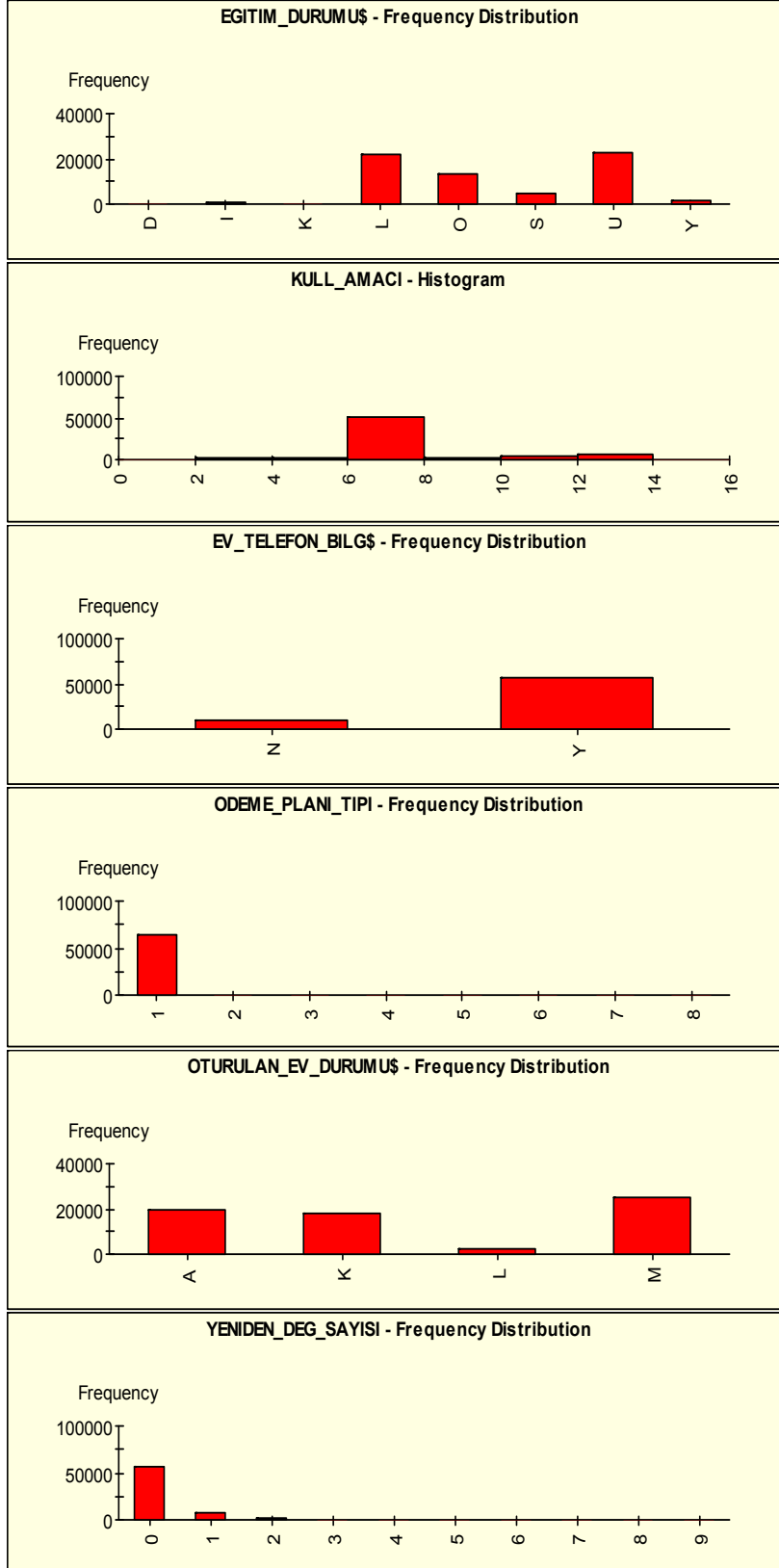
EK 1: Çalışmada Kullanılan Verilerin Tanımlayıcı İstatistikleri (devam)

	Ortalama	Std Deviation	Skewness	Coeff Variation	Cond. Mean
ID	34034	19651	-0.001376	0.57739	34034
EGITIM_DURUMU					
KULL_AMACI	67.038	22.311	13.835	0.33281	67.038
EV_TELEFON_BILG					
ODEME_PLANI_TIPI	10.015	0.063107	68.552	0.063015	10.015
OTURULAN_EV_DURUMU					
YENIDEN_DEG_SAYISI	0.17945	0.51868	40.176	28.905	13.165
YENI_MEVCUT_MUSTERI					
TOPLAM_CAL_SURESI	78.626	81.778	16.085	10.401	84.958
MAAS_MUSTERISI					
BANKAYLA_CAL_SURESI	44.225	141.29	158.54	31.948	51.1
SIMDIKI_ISYERI_CAL_SURESI	123.33	96.753	0.9286	0.78449	127.26
MUSTERI_TIPI					
ILISKI_SKORU	194.77	229.54	15.106	11.785	250.94
SEKTOR_TIPI	2974.1	1002.6	-25.643	0.33711	3305.3
MUSTERI_BANKA_SEGMENTI					
KOTU_KAYIT_TURU					
CALISMA_SEKLI	1053.4	14.56	-11.417	0.013822	1053.4
SUBE_GORUSU					
SOSYAL_GUVENCE					
BEYAN_EDILEN_GELIR	1508.1	2117.7	17.818	14.043	1508.2
SON_1AYDAKI_RET_ADEDI	0.14915	0.50799	47.681	34.058	14.193
KREDI_KARAR_KODU					
SON_1AYDAKI_ONAY_ADEDI	0.0049692	0.07975	30.722	16.049	10.521
KKB_AYLIK_TAKSIT	399.21	3036	128.58	76.049	745.89
KKB_GUNCCEL_ENKOTU_STATU					
KKB_GECMIS_ENKOTU_STATU					
MEDENI_DURUM					
KREDI_KARARI					
KKB_TOP_BORC	8357.5	22450	23.045	26.862	10373
COCUK_SAYISI	0.41572	38.266	24.748	92.047	28.962
KKB_SON_6AYDAKI_EN_KOTU_ODEME					
CINSIYET					
ONAYLANAN_VADE	22.5	10.704	0.95286	0.47572	22.5
YAS	36.829	10.498	0.59127	0.28504	36.829
BELGELENEN_TOP_GELIR	202.39	4035.9	242.49	19.941	1808.5
BORCLULUK_ORANI	27.574	93.778	-0.82433	0.3401	28.708
KULLANDIRIM_DURUMU	0.33158	0.47079	0.71547	14.198	1
BASVURU_SKORU	36.418	16.978	-0.10627	0.46621	36.418
KKB_SKOR	42.668	24.986	-0.47318	0.58559	49.908
TALEP_TL	5545.5	5431.4	27.383	0.97943	5558.7
TALEP_DOVIZ	3.48	291.83	116.57	83.859	14138
TALEP_VADE	22.05	10.707	11.096	0.48557	22.102
ONAYLANAN_LIMIT	5084.5	5103.1	32.66	10.037	5084.5
KREDI_KARARI_2					

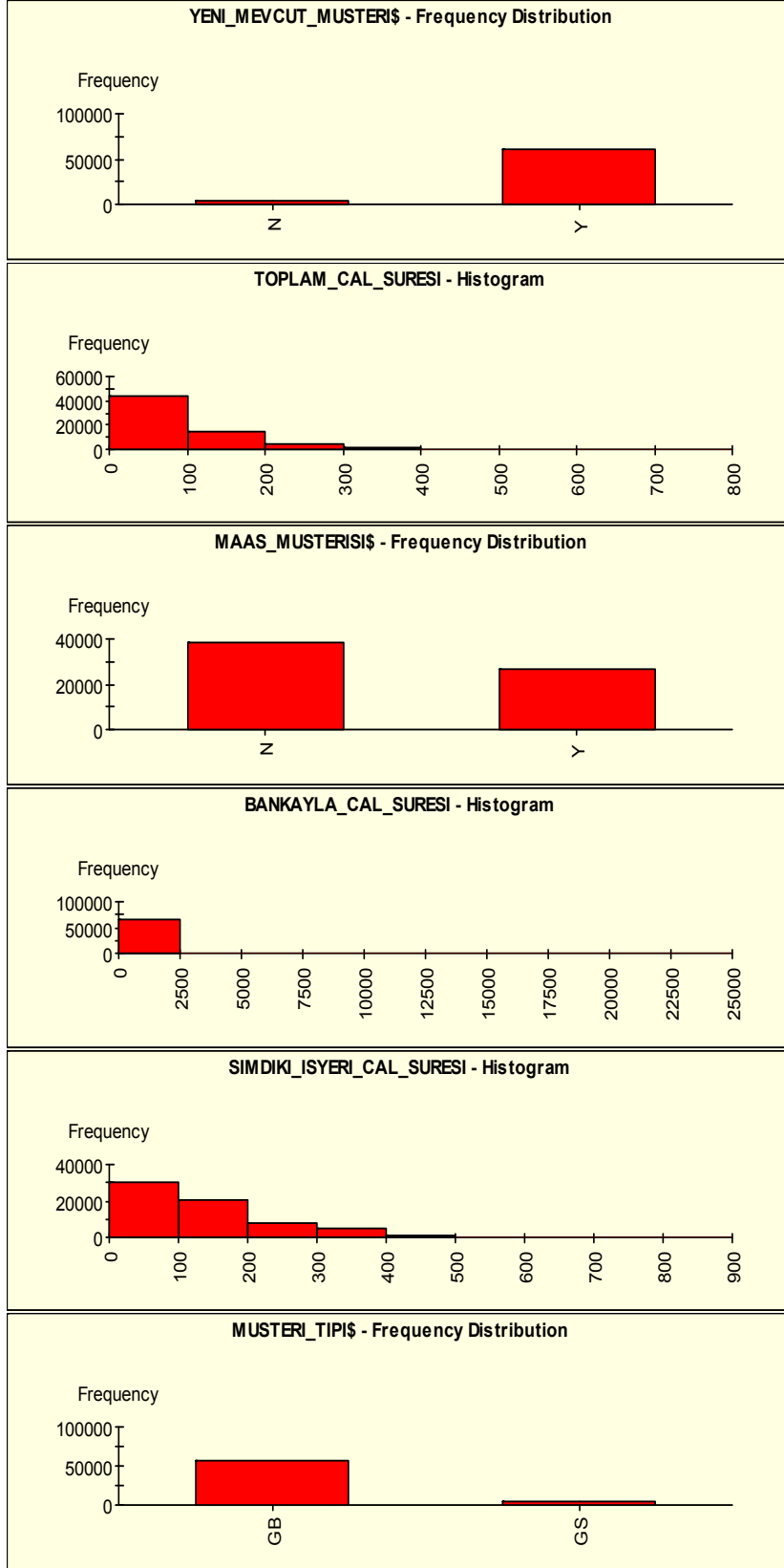
EK 1: Çalışmada Kullanılan Verilerin Tanımlayıcı İstatistikleri (devam)

	Sum of Weights	Sum	Variance	Kurtosis	Std Error Mean
ID	65000	2212230451	3,86E+12	-11.971	77.078
EGITIM_DURUMU					
KULL_AMACI	65000	435747	49.777	19.645	0.008751
EV_TELEFON_BILG					
ODEME_PLANI_TIPI	65000	65095	0.0039825	5771	0.00024753
OTURULAN_EV_DURUMU					
YENIDEN_DEG_SAYISI	65000	11664	0.26903	24.088	0.0020344
YENI_MEVCUT_MUSTERI					
TOPLAM_CAL_SURESI	65000	5110674	6687.7	2.877	0.32076
MAAS_MUSTERISI					
BANKAYLA_CAL_SURESI	65000	2874649	19964	27376	0.55419
SIMDIKI_ISYERI_CAL_SURESI	65000	8016632	9361.2	0.35405	0.3795
MUSTERI_TIPI					
ILISKI_SKORU	65000	12659894	52687	33.241	0.90031
SEKTOR_TIPI	65000	193316060	1,01E+10	47.824	39.325
MUSTERI_BANKA_SEGMENTI					
KOTU_KAYIT_TURU					
CALISMA_SEKLI	65000	68472638	212.01	0.3878	0.057111
SUBE_GORUSU					
SOSYAL_GUVENCE					
BEYAN_EDILEN_GELIR	65000	9,80E+11	4,48E+10	611.9	83.063
SON_1AYDAKI_RET_ADEDI	65000	9695	0.25805	32.874	0.0019925
KREDI_KARAR_KODU					
SON_1AYDAKI_ONAY_ADEDI	65000	323	0.00636	1954.6	0.0003128
KKB_AYLIK_TAKSIT	65000	25948604	9,22E+09	20258	11.908
KKB_GUNCEL_ENKOTU_STATU					
KKB_GECMIS_ENKOTU_STATU					
MEDENI_DURUM					
KREDI_KARARI					
KKB_TOP_BORC	65000	543239629	5,04E+12	1222	88.057
COCUK_SAYISI	65000	27022	14.643	634.12	0.015009
KKB_SON_6AYDAKI_EN_KOTU_ODEME					
CINSIYET					
ONAYLANAN_VADE	65000	1462520	114.57	0.98722	0.041984
YAS	65000	2393885	110.2	-0.20955	0.041176
BELGELENEN_TOP_GELIR	65000	13155276	1,63E+11	60775	15.83
BORCLULUK_ORANI	44508	1227248	87.943	0.85899	0.044451
KULLANDIRIM_DURUMU	65000	21553	0.22164	-14.881	0.0018466
BASVURU_SKORU	65000	236714	28.827	-12.036	0.0066595
KKB_SKOR	65000	277341	62.429	-11.053	0.0098002
TALEP_TL	65000	3,60E+12	2,95E+11	2465.1	21.304
TALEP_DOVIZ	65000	226200	85165	16428	11.447
TALEP_VADE	65000	1433232	114.63	4.251	0.041995
ONAYLANAN_LIMIT	65000	3,30E+12	2,60E+11	3171	20.016
KREDI_KARARI_2					

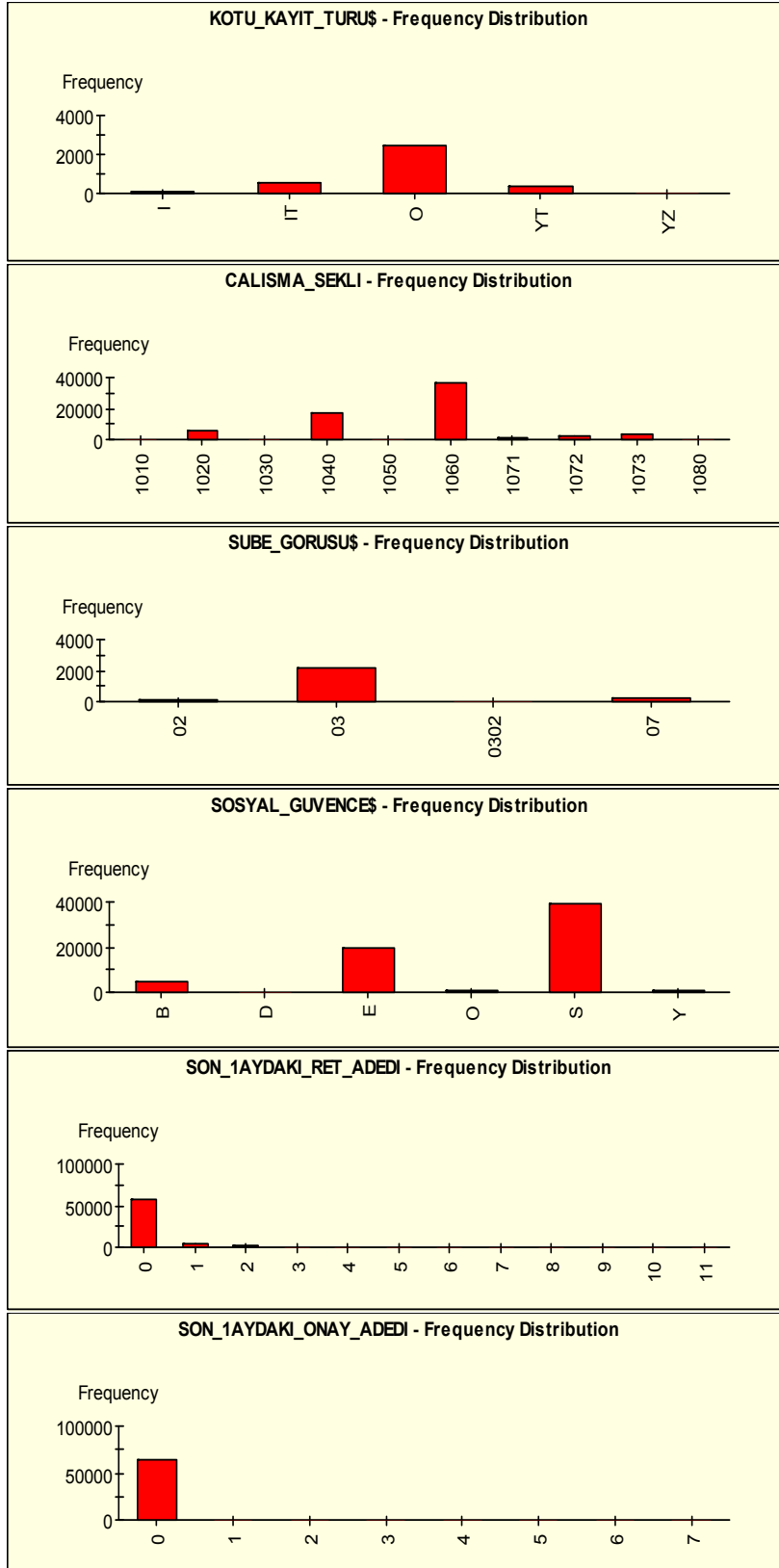
EK 2: Çalışmada Kullanılan Verilerin Sıklık Grafikleri



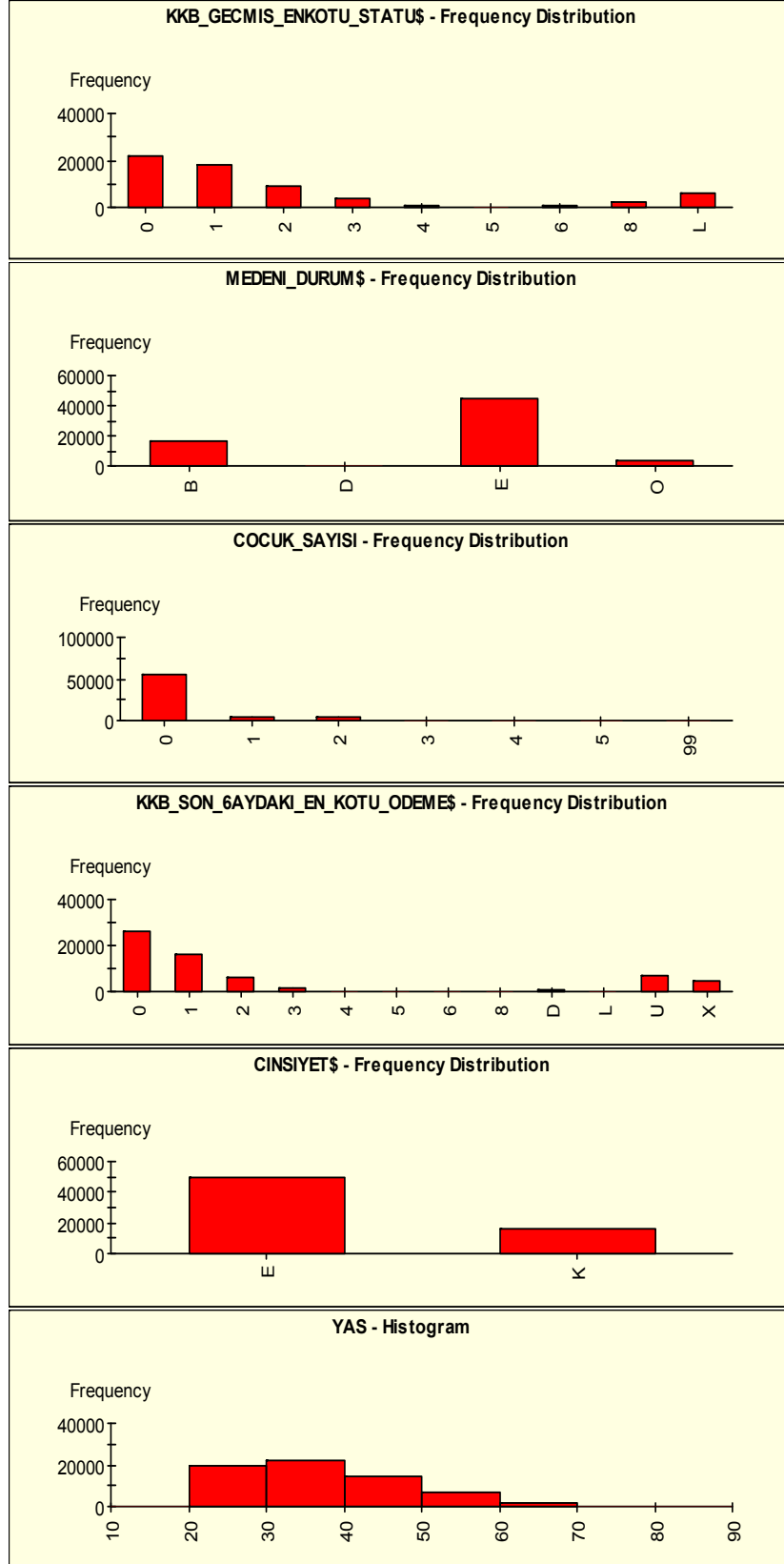
EK 2: Çalışmada Kullanılan Verilerin Sıklık Grafikleri (devam)



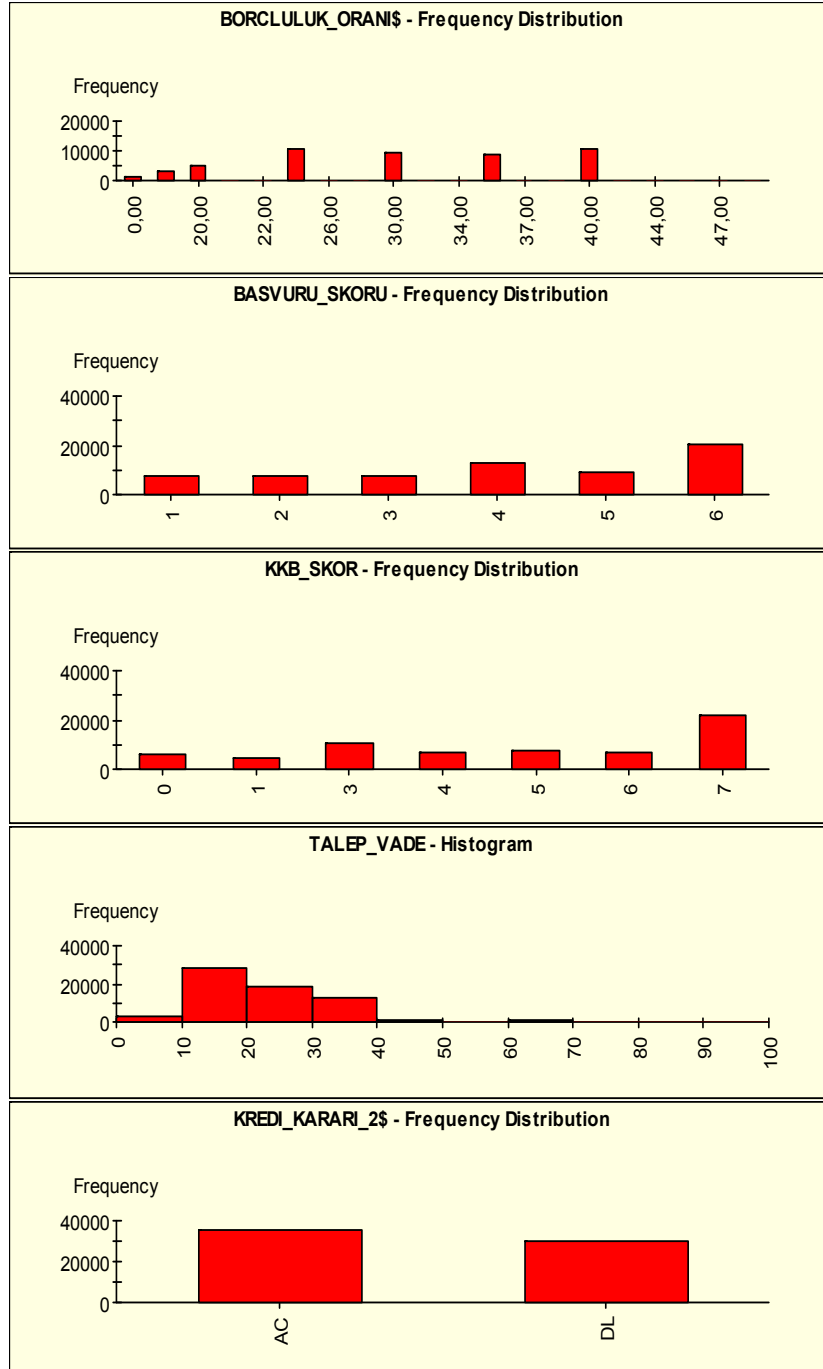
EK 2: Çalışmada Kullanılan Verilerin Sıklık Grafikleri (devam)



EK 2: Çalışmada Kullanılan Verilerin Sıklık Grafikleri (devam)



EK 2: Çalışmada Kullanılan Verilerin Sıklık Grafikleri (devam)



EK 3: MARS modelindeki parametrelerin özet bilgileri

PARAMETRELERİN ÖZET BİLGİLERİ				
Parametreler	Tahmin	Standart Hata	T değeri	P değeri
Constant	-557.478	235.763	-236.457	0.01998
EGITIM_DURUMU = "D"	0.00000			
EGITIM_DURUMU = "E"	0.00000			
EGITIM_DURUMU = "I"	0.01393	0.12664	0.11000	0.91263
EGITIM_DURUMU = "K"	0.00000			
EGITIM_DURUMU = "L"	Reference			
KULL_AMACI = 1	-0.46630	0.42609	-109.436	0.27643
KULL_AMACI = 2	0.00000			
KULL_AMACI = 3	-0.65840	0.53995	-121.936	0.22557
KULL_AMACI = 4	-0.57764	0.80795	-0.71494	0.47631
KULL_AMACI = 5	0.00000			
KULL_AMACI = 6	-0.47227	0.42861	-110.186	0.27317
KULL_AMACI = 8	-0.64434	0.57037	-112.969	0.26131
KULL_AMACI = 9	0.00000			
KULL_AMACI = 10	-197.280	121.002	-163.039	0.10617
KULL_AMACI = 11	-142.814	0.77611	-184.012	0.06872
KULL_AMACI = 12	0.00000			
KULL_AMACI = 13	0.00000			
KULL_AMACI = 14	Reference			
EV_TELEFON_BILG = "N"	0.02076	0.26373	0.07873	0.93741
EV_TELEFON_BILG = "Y"	Reference			
ODEME_PLANI_TIPI = 1	0.00000			
ODEME_PLANI_TIPI = 2	0.00000			
ODEME_PLANI_TIPI = 3	0.00000			
ODEME_PLANI_TIPI = 4	0.00000			
ODEME_PLANI_TIPI = 5	0.00000			
ODEME_PLANI_TIPI = 6	0.00000			
ODEME_PLANI_TIPI = 7	0.00000			
ODEME_PLANI_TIPI = 8	Reference			
OTURULAN_EV_DURUMU = "A"	0.10996	0.26763	0.41088	0.68204
OTURULAN_EV_DURUMU = "K"	0.13513	0.28977	0.46631	0.64201
OTURULAN_EV_DURUMU = "L"	0.03674	0.27120	0.13549	0.89250
OTURULAN_EV_DURUMU = "M"	Reference			
YENI_MEVCUT_MUSTERI = "N"	0.00000			
YENI_MEVCUT_MUSTERI = "Y"	Reference			
MAAS_MUSTERISI = "N"	-0.09710	0.16952	-0.57281	0.56806
MAAS_MUSTERISI = "Y"	Reference			
MUSTERI_TIPI = "GB"	0.35744	0.22903	156.063	0.12177
MUSTERI_TIPI = "GS"	Reference			
KOTU_KAYIT_TURU = "I"	0.31306	0.31342	0.99885	0.32028
KOTU_KAYIT_TURU = "IT"	0.62541	0.46585	134.252	0.18247
KOTU_KAYIT_TURU = "O"	0.64733	0.35502	182.338	0.07123
KOTU_KAYIT_TURU = "YT"	0.00000			
KOTU_KAYIT_TURU = "YZ"	Reference			
CALISMA_SEKLI = 1010	0.98261	0.63843	153.910	0.12694
CALISMA_SEKLI = 1020	122.045	0.64464	189.325	0.06122
CALISMA_SEKLI = 1030	0.87860	0.70838	124.030	0.21777
CALISMA_SEKLI = 1040	134.458	0.69738	192.803	0.05669

EK 3: MARS modelindeki parametrelerin özet bilgileri (devam)

PARAMETRELERİN ÖZET BİLGİLERİ				
Parametreler	Tahmin	Standart Hata	T değeri	P değeri
CALISMA_SEKLI = 1050	0.00000			
CALISMA_SEKLI = 1060	134.798	0.68246	197.517	0.05101
CALISMA_SEKLI = 1071	123.851	0.74100	167.141	0.09777
CALISMA_SEKLI = 1072	0.00000			
CALISMA_SEKLI = 1073	0.00000			
CALISMA_SEKLI = 1080	Reference			
SUBE_GORUSU = "02"	0.76048	111.494	0.68208	0.49676
SUBE_GORUSU = "03"	0.00000			
SUBE_GORUSU = "0302"	139.044	790.830	0.17582	0.86079
SUBE_GORUSU = "07"	Reference			
SOSYAL_GUVENCE = "B"	0.35010	0.54251	0.64534	0.52019
SOSYAL_GUVENCE = "D"	0.38434	0.55548	0.69191	0.49060
SOSYAL_GUVENCE = "E"	0.32845	0.58305	0.56334	0.57446
SOSYAL_GUVENCE = "O"	0.00000			
SOSYAL_GUVENCE = "S"	-0.68890	0.93522	-0.73662	0.46308
SOSYAL_GUVENCE = "Y"	Reference			
KKB_GUNCCEL_ENKOTU_STATU = "0"	0.83845	0.97679	0.85837	0.39274
KKB_GUNCCEL_ENKOTU_STATU = "1"	0.63394	0.85206	0.74401	0.45862
KKB_GUNCCEL_ENKOTU_STATU = "2"	0.65740	0.86857	0.75687	0.45091
KKB_GUNCCEL_ENKOTU_STATU = "3"	0.78575	0.92708	0.84755	0.39871
KKB_GUNCCEL_ENKOTU_STATU = "4"	0.49156	0.89359	0.55010	0.58348
KKB_GUNCCEL_ENKOTU_STATU = "5"	161.526	118.781	135.986	0.17693
KKB_GUNCCEL_ENKOTU_STATU = "6"	0.00000			
KKB_GUNCCEL_ENKOTU_STATU = "8"	249.861	132.359	188.774	0.06196
KKB_GUNCCEL_ENKOTU_STATU = "L"	0.14495	113.259	0.12798	0.89842
KKB_GUNCCEL_ENKOTU_STATU = "U"	Reference			
KKB_GECMIS_ENKOTU_STATU = "0"	-0.25963	0.63876	-0.40646	0.68527
KKB_GECMIS_ENKOTU_STATU = "1"	0.11905	0.48694	0.24448	0.80736
KKB_GECMIS_ENKOTU_STATU = "2"	0.11879	0.48425	0.24530	0.80673
KKB_GECMIS_ENKOTU_STATU = "3"	0.01205	0.46966	0.02565	0.97959
KKB_GECMIS_ENKOTU_STATU = "4"	0.08465	0.46127	0.18350	0.85477
KKB_GECMIS_ENKOTU_STATU = "5"	0.19615	0.61208	0.32046	0.74929
KKB_GECMIS_ENKOTU_STATU = "6"	-0.91629	0.86907	-105.434	0.29427
KKB_GECMIS_ENKOTU_STATU = "8"	192.531	131.442	146.477	0.14612
KKB_GECMIS_ENKOTU_STATU = "L"	Reference			
MEDENI_DURUM = "B"	110.815	0.69483	159.485	0.11390
MEDENI_DURUM = "D"	139.879	0.72477	192.999	0.05644
MEDENI_DURUM = "E"	114.778	0.72032	159.344	0.11422
MEDENI_DURUM = "O"	Reference			
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "0"	0.96276	0.80072	120.237	0.23206
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "1"	0.89518	0.81603	109.699	0.27528
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "2"	0.76467	0.86021	0.88893	0.37618
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "3"	130.474	0.85639	152.354	0.13078
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "4"	0.55937	103.791	0.53894	0.59112
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "5"	0.83527	0.85623	0.97552	0.33166
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "6"	104.707	0.84137	124.449	0.21623

EK 3: MARS modelindeki parametrelerin özet bilgileri (devam)

PARAMETRELERİN ÖZET BİLGİLERİ				
Parametreler	Tahmin	Standart Hata	T değeri	P değeri
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "8"	0.58859	105.249	0.55924	0.57725
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "D"	0.00000			
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "L"	108.682	0.98695	110.120	0.27345
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "U"	0.00000			
KKB_SON_6AYDAKI_EN_KOTU_ODEME = "X"	Reference			
CINSİYET = "E"	-0.12493	0.15924	-0.78456	0.43457
CINSİYET = "K"	Reference			
BASVURU_SKORU = 1	0.33843	0.39812	0.85007	0.39732
BASVURU_SKORU = 2	0.18634	0.37894	0.49174	0.62398
BASVURU_SKORU = 3	0.33546	0.33367	100.535	0.31715
BASVURU_SKORU = 4	0.21324	0.29037	0.73437	0.46444
BASVURU_SKORU = 5	0.57417	0.27821	206.377	0.04163
BASVURU_SKORU = 6	Reference			
KKB_SKOR = 0	0.43347	0.54982	0.78839	0.43233
KKB_SKOR = 1	0.37683	0.68114	0.55324	0.58134
KKB_SKOR = 3	0.60532	0.66785	0.90637	0.36692
KKB_SKOR = 4	0.54776	0.60563	0.90444	0.36793
KKB_SKOR = 5	0.34981	0.55555	0.62967	0.53035
KKB_SKOR = 6	0.54103	0.59343	0.91169	0.36412
KKB_SKOR = 7	Reference			
YENIDEN_DEG_SAYISI	-0.07268	0.10027	-0.72487	0.47022
TOPLAM_CAL_SURESI	-0.00166	0.00086	-191.621	0.05819
BANKAYLA_CAL_SURESI	-0.00038	0.00157	-0.24121	0.80989
SIMDIKI_ISYERI_CAL_SURESI	0.00127	0.00100	127.194	0.20634
BEYAN_EDILEN_GELIR	-0.00001	0.00004	-0.25979	0.79556
SON_1AYDAKI_RET_ADEDI	0.14629	0.13745	106.432	0.28975
SON_1AYDAKI_ONAY_ADEDI	0.00000			
KKB_AYLIK_TAKSIT	0.00008	0.00010	0.79128	0.43065
KKB_TOP_BORC	-0.00000	0.00000	-0.25466	0.79951
COCUK_SAYISI	-0.00453	0.06426	-0.07055	0.94390
YAS	0.00504	0.00803	0.62824	0.53128
BELGELENEN_TOP_GELIR	-0.00003	0.00005	-0.62174	0.53553
BORCLULUK_ORANI	-0.00719	0.01900	-0.37847	0.70589
TALEP_TL	0.00000	0.00001	0.04165	0.96686
TALEP_DOVIZ	0.00000			
TALEP_VADE	0.00679	0.00712	0.95410	0.34233

EK 4: MARS yöntemi ileriye doğru adımında modele giren temel fonksiyonlar

İLERİYE DOĞRU ADIMINDA MODELE GİREN TEMEL FONKSİYONLAR						
TF	GCV	Efektif Parametre Sayısı	Değişken 1	Düğüm Noktası	Değişken 2	Düzey
0	0.24876	1.0				
2 1	0.14940	2.0	BORCLULUK_ORANI_mis	0-1		
4 3	0.13071	6.0	BORCLULUK_ORANI	15.00000	BORCLULUK_ORANI_mis	1
6 5	0.11897	10.0	KKB_AYLIK_TAKSIT	1491.99988	BORCLULUK_ORANI_mis	1
8 7	0.11501	14.0	BELGELENEN_TOP_GELİR	1785.99988	BORCLULUK_ORANI_mis	1
10 9	0.11311	17.0	MAAS_MUSTERİSİ\$	0-1	BORCLULUK_ORANI_mis	1
12 11	0.11099	21.0	BEYAN_EDİLEN_GELİR	2232.66992	BORCLULUK_ORANI_mis	1
14 13	0.11098	22.0	KKB_GUNCEL_ENKOTU_S TATU\$ mis	0-1	BORCLULUK_ORANI_mis	1
16 15	0.10890	25.0	KKB_GUNCEL_ENKOTU_S TATU\$	0011111000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
18 17	0.10778	29.0	YENİDEN_DEĞ_SAYISI	1.00000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
20 19	0.10680	33.0	KKB_TOP_BORC	29821.00000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
22 21	0.10597	37.0	KKB_AYLIK_TAKSIT	252.00000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
24 23	0.10597	37.0	KKB_SON_6AYDAKI_EN_K OTU_ODEME\$ mis	0-1	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
26 25	0.10532	40.0	KKB_SON_6AYDAKI_EN_K OTU_ODEME\$	00111110000	KKB_SON_6AYDAKI_EN_K OTU_ODEME\$ mis	2-3
28 27	0.10481	44.0	BEYAN_EDİLEN_GELİR	858.00000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
30 29	0.10481	44.0	KKB_GECMIS_ENKOTU_S TATU\$ mis	0-1	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
32 31	0.10438	47.0	KKB_GECMIS_ENKOTU_S TATU\$	000000001	KKB_GECMIS_ENKOTU_S TATU\$ mis	2-9
34 33	0.10386	50.0	SUBE_GORUSU\$ mis	1-0	BORCLULUK_ORANI_mis	1
36 35	0.10338	54.0	SON_1AYDAKI_RET_ADE DI	1.00000	SUBE_GORUSU\$ mis	3-3
38 37	0.10296	58.0	KKB_TOP_BORC	10782.00000	SUBE_GORUSU\$ mis	3-3
40 39	0.10263	61.0	KULL_AMACI	Report Later	SUBE_GORUSU\$ mis	3-3
42 41	0.10262	62.0	KKB_GUNCEL_ENKOTU_S TATU\$ mis	0-1	SUBE_GORUSU\$ mis	3-4
44 43	0.10230	65.0	KKB_GUNCEL_ENKOTU_S TATU\$	0011111001	KKB_GUNCEL_ENKOTU_S TATU\$ mis	4-1
46 45	0.10203	69.0	BELGELENEN_TOP_GELİR	1150.00000		
48 47	0.10180	72.0	BEYAN_EDİLEN_GELİR	7000.00000	KKB_GUNCEL_ENKOTU_S TATU\$ mis	1-3
50 49	0.10154	76.0	TALEP_TL	19999.99805	BORCLULUK_ORANI_mis	1
52 51	0.10129	79.0	CALISMA_SEKLI	1001000001	BORCLULUK_ORANI_mis	1
54 53	0.10117	83.0	BORCLULUK_ORANI	15.00000	SUBE_GORUSU\$ mis	3-4
56 55	0.10107	87.0	KKB_AYLIK_TAKSIT	6042.99951	KKB_GUNCEL_ENKOTU_S TATU\$ mis	4-1

EK 4: MARS yöntemi ileriye doğru adımında modele giren temel fonksiyonlar (devam)

İLERİYE DOĞRU ADIMINDA MODELE GİREN TEMEL FONKSİYONLAR (devam)						
TF	GCV	Efektif Parametre Sayısı	Değişken	Düğüm Noktası	Değişken 2	Düzey
58 57	0.10107	87.0	KKB_GUNCEL_ENKOTU_STATU\$_mis	0-1	SUBE_GORUSU\$_mis	3-3
60 59	0.10100	90.0	KKB_GUNCEL_ENKOTU_STATU\$	1011110001	SUBE_GORUSU\$_mis	5-7
62 61	0.10089	93.0	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$	111011100100	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$_mis	2-3
64 63	0.10083	97.0	BANKAYLA_CAL_SUR ESI	28.00000	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$_mis	2-3
66 65	0.10077	101.0	TALEP_VADE	36.00000	KKB_GUNCEL_ENKOTU_STATU\$_mis	1-3
68 67	0.10071	105.0	YAS	23.00000	KKB_GUNCEL_ENKOTU_STATU\$_mis	1-3
70 69	0.10072	106.0	SOSYAL_GUVENCE\$_mis	0-1	BORCLULUK_ORANI_mis	1

EK 5: MARS final modeli

GERİYE DOĞRU ADIM SÜRECİNDEN SONRA MODELE GİREN TEMEL FONKSİYONLAR (FİNAL MODEL)				
TF	Katsayı	Değişken	Değişken 2	Düğüm noktası
0				
1	-0.18742	BORCLULUK_ORANI_mis		0-1
3	-0.00463	BORCLULUK_ORANI	BORCLULUK_ORANI_mis	15.00000
4	0.03691	BORCLULUK_ORANI	BORCLULUK_ORANI_mis	15.00000
5	-0.00017	KKB_AYLIK_TAKSIT	BORCLULUK_ORANI_mis	1491.99988
6	-0.00013	KKB_AYLIK_TAKSIT	BORCLULUK_ORANI_mis	1491.99988
7	-0.00005	BELGELENEN_TOP_GELİR	BORCLULUK_ORANI_mis	1785.99988
9	-0.17008	MAAS_MUSTERİSİ\$	BORCLULUK_ORANI_mis	0-1
12	0.00007	BEYAN_EDİLEN_GELİR	BORCLULUK_ORANI_mis	2232.66992
13	0.17406	KKB_GUNCEL_ENKOTU_STATU\$ mis	BORCLULUK_ORANI_mis	0-1
15	0.29765	KKB_GUNCEL_ENKOTU_STATU\$	KKB_GUNCEL_ENKOTU_STATU\$ mis	0011111000
17	0.02458	YENİDEN_DEĞ_SAYISI	KKB_GUNCEL_ENKOTU_STATU\$ mis	1.00000
18	-0.15200	YENİDEN_DEĞ_SAYISI	KKB_GUNCEL_ENKOTU_STATU\$ mis	1.00000
19	0.00000	KKB_TOP_BORC	KKB_GUNCEL_ENKOTU_STATU\$ mis	29821.00000
20	-0.00001	KKB_TOP_BORC	KKB_GUNCEL_ENKOTU_STATU\$ mis	29821.00000
21	0.00018	KKB_AYLIK_TAKSIT	KKB_GUNCEL_ENKOTU_STATU\$ mis	252.00000
22	0.00046	KKB_AYLIK_TAKSIT	KKB_GUNCEL_ENKOTU_STATU\$ mis	252.00000
25	0.12287	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$ mis	001111110000
27	-0.00005	BEYAN_EDİLEN_GELİR	KKB_GUNCEL_ENKOTU_STATU\$ mis	858.00000
28	0.00075	BEYAN_EDİLEN_GELİR	KKB_GUNCEL_ENKOTU_STATU\$ mis	858.00000
31	0.10400	KKB_GECMİS_ENKOTU_STATU\$	KKB_GECMİS_ENKOTU_STATU\$ mis	000000001
35	0.01574	SON_1AYDAKI_RET_ADEDİ	SUBE_GORUSU\$ mis	1.00000
36	-0.09374	SON_1AYDAKI_RET_ADEDİ	SUBE_GORUSU\$ mis	1.00000
37	-0.00000	KKB_TOP_BORC	SUBE_GORUSU\$ mis	10782.00000
38	-0.00001	KKB_TOP_BORC	SUBE_GORUSU\$ mis	10782.00000
39	0.07799	KULL_AMACI	SUBE_GORUSU\$ mis	0000000000110
41	-0.67436	KKB_GUNCEL_ENKOTU_STATU\$ mis	SUBE_GORUSU\$ mis	0-1
43	-0.41622	KKB_GUNCEL_ENKOTU_STATU\$	KKB_GUNCEL_ENKOTU_STATU\$ mis	0011111001
45	0.00005	BELGELENEN_TOP_GELİR		1150.00000
46	0.00023	BELGELENEN_TOP_GELİR		1150.00000
47	0.00005	BEYAN_EDİLEN_GELİR	KKB_GUNCEL_ENKOTU_STATU\$ mis	7000.00000
49	0.00000	TALEP_TL	BORCLULUK_ORANI_mis	19999.99805
50	-0.00001	TALEP_TL	BORCLULUK_ORANI_mis	19999.99805
51	-0.05529	CALISMA_SEKLI	BORCLULUK_ORANI_mis	1001000001
53	-0.00384	BORCLULUK_ORANI	SUBE_GORUSU\$ mis	15.00000
54	-0.03251	BORCLULUK_ORANI	SUBE_GORUSU\$ mis	15.00000
56	0.00009	KKB_AYLIK_TAKSIT	KKB_GUNCEL_ENKOTU_STATU\$ mis	6042.99951
59	-0.03714	KKB_GUNCEL_ENKOTU_STATU\$	KKB_GUNCEL_ENKOTU_STATU\$ mis	1011110001
61	-0.03443	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$ mis	111011100100
63	-0.00003	BANKAYLA_CAL_SURESI	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$ mis	28.00000
64	0.00096	BANKAYLA_CAL_SURESI	KKB_SON_6AYDAKI_EN_KOTU_ÖDEME\$ mis	28.00000
66	0.00144	TALEP_VADE	KKB_GUNCEL_ENKOTU_STATU\$ mis	36.00000
67	0.00101	YAS	KKB_GUNCEL_ENKOTU_STATU\$ mis	23.00000
68	-0.01073	YAS	KKB_GUNCEL_ENKOTU_STATU\$ mis	23.00000

EK 6: MARS Final Modelini Oluşturan Temel Fonksiyonların Özellikleri

PARAMETRE	KATSAYI	ST. HATA	T-RATIO	P-VALUE
Constant	0.73566	0.00691	10.643.219	0.00000
TF1	-0.18758	0.03733	-502.509	0.00000
TF3	-0.00463	0.00029	-1.613.265	0.00000
TF4	0.03691	0.00062	5.906.852	0.00000
TF5	-0.00017	0.00002	-723.072	0.00000
TF6	-0.00013	0.00002	-533.307	0.00000
TF7	-0.00005	0.00001	-361.351	0.00030
TF9	-0.17008	0.00394	-4.318.580	0.00000
TF12	0.00007	0.00001	1.358.921	0.00000
TF13	0.17408	0.02133	815.978	0.00000
TF15	0.29765	0.00934	3.188.359	0.00000
TF17	0.02458	0.00638	385.210	0.00012
TF18	-0.15200	0.00577	-2.634.080	0.00000
TF19	0.00000	0.00000	349.057	0.00048
TF20	-0.00001	0.00000	-1.119.153	0.00000
TF21	0.00018	0.00002	727.362	0.00000
TF22	0.00046	0.00003	1.420.313	0.00000
TF25	0.12287	0.00614	2.000.545	0.00000
TF27	-0.00005	0.00000	-2.149.652	0.00000
TF28	0.00075	0.00004	1.752.617	0.00000
TF31	0.10400	0.00740	1.404.790	0.00000
TF35	0.01574	0.00660	238.440	0.01711
TF36	-0.09375	0.00615	-1.524.406	0.00000
TF37	-0.00000	0.00000	-374.095	0.00018
TF38	-0.00001	0.00000	-1.024.298	0.00000
TF39	0.07799	0.00576	1.352.940	0.00000
TF41	-0.67437	0.05984	-1.126.899	0.00000
TF43	-0.41622	0.02723	-1.528.772	0.00000
TF45	0.00005	0.00001	362.049	0.00029
TF46	0.00023	0.00001	4.041.202	0.00000
TF47	0.00005	0.00000	1.706.745	0.00000
TF49	0.00000	0.00000	460.701	0.00000
TF50	-0.00001	0.00000	-1.432.371	0.00000
TF51	-0.05529	0.00401	-1.377.636	0.00000
TF53	-0.00384	0.00120	-319.448	0.00140
TF54	-0.03251	0.00271	-1.199.373	0.00000
TF56	0.00009	0.00001	850.119	0.00000
TF59	-0.03714	0.00431	-861.349	0.00000
TF61	-0.03443	0.00386	-892.894	0.00000
TF63	-0.00003	0.00001	-265.083	0.00803
TF64	0.00096	0.00016	590.689	0.00000
TF66	0.00144	0.00020	709.279	0.00000
TF67	0.00101	0.00017	579.264	0.00000
TF68	-0.01073	0.00423	-253.311	0.01131
F-STATISTIC = 2231.74254		S.E. OF REGRESSION = 0.31698		
P-VALUE = 0.00000		RESIDUAL SUM OF SQUARES = 6526.54552		
[MDF,NDF] = [43, 64956]		REGRESSION SUM OF SQUARES = 9642.21134		

EK 7: İkili lojistik regresyon modeli

Modele Giren Değişkenler ve Düzeyleri	Katsayı	Standart Hata	Serbestlik derecesi	P değeri
EGITIM_DURUMU			4	0,0510
EGITIM_DURUMU(1)	-1,2017	0,4274	1	0,0049
EGITIM_DURUMU(2)	-0,2898	0,9404	1	0,7579
EGITIM_DURUMU(3)	0,0259	0,0285	1	0,3632
EGITIM_DURUMU(4)	0,4276	0,5507	1	0,4374
KULL_AMACI			12	0,0000
KULL_AMACI(1)	-0,1202	0,2274	1	0,5971
KULL_AMACI(2)	-0,2380	0,1628	1	0,1438
KULL_AMACI(3)	-0,1213	0,2283	1	0,5954
KULL_AMACI(4)	-0,3291	0,1858	1	0,0764
KULL_AMACI(5)	-0,1177	0,1931	1	0,5420
KULL_AMACI(6)	-0,0618	0,1422	1	0,6641
KULL_AMACI(7)	0,0635	0,1707	1	0,7101
KULL_AMACI(8)	-0,3094	0,3101	1	0,3184
KULL_AMACI(9)	-0,0201	0,2131	1	0,9248
KULL_AMACI(10)	-0,2280	0,1525	1	0,1348
KULL_AMACI(11)	0,4201	0,1469	1	0,0042
KULL_AMACI(12)	0,7843	0,8058	1	0,3304
EV_TELEFON_BILG(1)	-0,0973	0,0383	1	0,0111
ODEME_PLANI_TIPI			6	0,6282
ODEME_PLANI_TIPI(1)	-21,4791	40.073,4517	1	0,9996
ODEME_PLANI_TIPI(2)	-20,8419	40.073,4517	1	0,9996
ODEME_PLANI_TIPI(3)	-23,0546	40.073,4517	1	0,9995
ODEME_PLANI_TIPI(4)	-0,1918	56.756,9943	1	1,0000
ODEME_PLANI_TIPI(5)	-42,0578	46.269,4289	1	0,9993
ODEME_PLANI_TIPI(6)	-20,6058	40.073,4517	1	0,9996
OTURULAN_EV_DURUMU			3	0,0000
OTURULAN_EV_DURUMU(1)	-0,1937	0,0332	1	0,0000
OTURULAN_EV_DURUMU(2)	-0,1642	0,0321	1	0,0000
OTURULAN_EV_DURUMU(3)	-0,0552	0,0732	1	0,4512
YENIDEN_DEG_SAYISI	0,6439	0,0274	1	0,0000
YENI_MEVCUT_MUSTERI(1)	0,0931	0,0774	1	0,2294
TOPLAM_CAL_SURESI	0,0002	0,0002	1	0,3826
MAAS_MUSTERISI(1)	0,7963	0,0362	1	0,0000
BANKAYLA_CAL_SURESI	-0,0007	0,0003	1	0,0498
SIMDIKI_ISYERI_CAL_SURESI	0,0000	0,0002	1	0,9712
MUSTERI_TIPI			1	0,0196
MUSTERI_TIPI(2)	0,1515	0,0649	1	0,0196
KOTU_KAYIT_TURU			5	0,0377
KOTU_KAYIT_TURU(1)	21,0735	39.425,0006	1	0,9996
KOTU_KAYIT_TURU(2)	20,0683	39.425,0006	1	0,9996
KOTU_KAYIT_TURU(3)	21,5048	39.425,0006	1	0,9996
KOTU_KAYIT_TURU(4)	20,8655	39.425,0006	1	0,9996
KOTU_KAYIT_TURU(5)	21,1913	39.425,0006	1	0,9996

EK 7: İkili lojistik regresyon modeli (devam)

Modele Giren Değişkenler ve düzeyleri	Katsayı	Standart Hata	Serbestlik derecesi	P değeri
CALISMA_SEKLI			9	0,0000
CALISMA_SEKLI(1)	0,0230	0,7192	1	0,9745
CALISMA_SEKLI(2)	0,2441	0,2603	1	0,3484
CALISMA_SEKLI(3)	0,7569	0,6567	1	0,2491
CALISMA_SEKLI(4)	-0,0416	0,2603	1	0,8731
CALISMA_SEKLI(5)	1,7381	0,7605	1	0,0223
CALISMA_SEKLI(6)	0,3456	0,2583	1	0,1810
CALISMA_SEKLI(7)	0,3217	0,2727	1	0,2381
CALISMA_SEKLI(8)	0,3361	0,2652	1	0,2049
CALISMA_SEKLI(9)	0,2654	0,2634	1	0,3137
SUBE_GORUSU			4	0,0000
SUBE_GORUSU(1)	19,6673	39.668,5393	1	0,9996
SUBE_GORUSU(2)	16,9684	39.668,5393	1	0,9997
SUBE_GORUSU(3)	18,7670	39.668,5393	1	0,9996
SUBE_GORUSU(4)	15,2203	39.668,5393	1	0,9997
SOSYAL_GUVENCE			6	0,0008
SOSYAL_GUVENCE(1)	-1,6817	0,6591	1	0,0107
SOSYAL_GUVENCE(2)	-0,2403	0,1187	1	0,0429
SOSYAL_GUVENCE(3)	-0,3601	0,2489	1	0,1480
SOSYAL_GUVENCE(4)	-0,2730	0,1168	1	0,0194
SOSYAL_GUVENCE(5)	0,0413	0,1756	1	0,8139
SOSYAL_GUVENCE(6)	-0,3378	0,1123	1	0,0026
BEYAN_EDILEN_GELIR	-0,0003	0,0000	1	0,0000
SON_1AYDAKI_RET_ADEDI	0,3438	0,0252	1	0,0000
SON_1AYDAKI_ONAY_ADEDI	0,4233	0,1499	1	0,0047
KKB_AYLIK_TAKSIT	0,0003	0,0000	1	0,0000
KKB_GUNCEL_ENKOTU_STATU2			9	0,0000
KKB_GUNCEL_ENKOTU_STATU2(1)	0,1661	0,2801	1	0,5531
KKB_GUNCEL_ENKOTU_STATU2(2)	0,4408	0,2826	1	0,1188
KKB_GUNCEL_ENKOTU_STATU2(3)	1,6818	0,2912	1	0,0000
KKB_GUNCEL_ENKOTU_STATU2(4)	1,4877	0,3145	1	0,0000
KKB_GUNCEL_ENKOTU_STATU2(5)	1,6439	0,5049	1	0,0011
KKB_GUNCEL_ENKOTU_STATU2(6)	1,9303	0,7153	1	0,0070
KKB_GUNCEL_ENKOTU_STATU2(7)	1,8865	0,4786	1	0,0001
KKB_GUNCEL_ENKOTU_STATU2(8)	0,2451	0,3367	1	0,4666
KKB_GUNCEL_ENKOTU_STATU2(9)	0,2641	0,3108	1	0,3955
KKB_GECMIS_ENKOTU_STATU2			8	0,0003
KKB_GECMIS_ENKOTU_STATU2(1)	-0,5641	0,1451	1	0,0001
KKB_GECMIS_ENKOTU_STATU2(2)	-0,4766	0,1441	1	0,0009
KKB_GECMIS_ENKOTU_STATU2(3)	-0,5521	0,1453	1	0,0001
KKB_GECMIS_ENKOTU_STATU2(4)	-0,6251	0,1512	1	0,0000
KKB_GECMIS_ENKOTU_STATU2(5)	-0,5735	0,1877	1	0,0022
KKB_GECMIS_ENKOTU_STATU2(6)	-0,8754	0,2603	1	0,0008
KKB_GECMIS_ENKOTU_STATU2(7)	-0,7562	0,2204	1	0,0006
KKB_GECMIS_ENKOTU_STATU2(8)	-0,4236	0,2341	1	0,0703

EK 7: İkili lojistik regresyon modeli (devam)

Modele Giren Değişkenler ve düzeyleri	Katsayı	Standart Hata	Serbestlik derecesi	P değeri
MEDENI_DURUM			3	0,5205
MEDENI_DURUM(1)	-0,0331	0,0693	1	0,6329
MEDENI_DURUM(2)	0,1522	0,1356	1	0,2616
MEDENI_DURUM(3)	-0,0291	0,0625	1	0,6419
KKB_TOP_BORC	0,0000	0,0000	1	0,0000
COCUK_SAYISI	0,0011	0,0034	1	0,7473
KKB_SON_6AYDAKI_EN_KOTU_ODEME2			11	0,0000
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (1)	-0,3651	0,0486	1	0,0000
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (2)	-0,3433	0,0579	1	0,0000
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (3)	0,2990	0,0786	1	0,0001
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (4)	0,5205	0,1147	1	0,0000
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (5)	0,2909	0,2469	1	0,2387
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (6)	0,2421	0,3898	1	0,5345
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (7)	0,2196	0,2758	1	0,4260
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (8)	0,7283	0,8797	1	0,4078
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (9)	-0,0444	0,1141	1	0,6973
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (10)	-0,8923	0,6237	1	0,1526
KKB_SON_6AYDAKI_EN_KOTU_ODEME2 (11)	-0,0169	0,0561	1	0,7637
CINSIYET(1)	-0,1366	0,0338	1	0,0001
YAS	0,0112	0,0019	1	0,0000
BELGELENEN_TOP_GELIR	-0,0008	0,0000	1	0,0000
BORCLULUK_ORANI	-0,0824	0,0041	1	0,0000
BASVURU_SKORU			5	0,0000
BASVURU_SKORU(1)	-0,2167	0,0902	1	0,0163
BASVURU_SKORU(2)	-0,3521	0,0742	1	0,0000
BASVURU_SKORU(3)	-0,1283	0,0639	1	0,0446
BASVURU_SKORU(4)	0,1072	0,0496	1	0,0306
BASVURU_SKORU(5)	0,1196	0,0474	1	0,0116
KKB_SKOR			6	0,0000
KKB_SKOR(1)	-0,1024	0,0954	1	0,2832
KKB_SKOR(2)	-0,1972	0,0988	1	0,0460
KKB_SKOR(3)	-0,4237	0,0673	1	0,0000
KKB_SKOR(4)	-0,2608	0,0593	1	0,0000
KKB_SKOR(5)	-0,1445	0,0476	1	0,0024
KKB_SKOR(6)	0,0077	0,0420	1	0,8554
TALEP_TL	0,0000	0,0000	1	0,0000
TALEP_DOVIZ	0,0000	0,0000	1	0,9692
TALEP_VADE	-0,0047	0,0016	1	0,0032
Constant	-0,8102	0,0105	1	0,0000

ÖZGEÇMİŞ

Adı Soyadı: Birce Ünal

Doğum Yeri: Ankara

Doğum Yılı: 1983

Medeni Hali: Bekar

Eğitim ve Akademik Durumu:

Lise: 1997 - 2001 Kurtuluş Yabancı Dil Ağırlıklı Lisesi, Ankara

Lisans: 2001 - 2005 Hacettepe Üniversitesi İstatistik Bölümü, Ankara

Yabancı Dil: İngilizce

İş Tecrübesi:

2007- 2007 : Heraeus Electro-Nite, İstatistik Uzmanı

2007- : Türkiye Garanti Bankası A.Ş., Kurumsal ve Ticari Kredileri
Müdürlüğü, Kredi Değerlendirme Yetkilisi