

HASAR REZERVİ İÇİN RİSK SINIFLANDIRMASI

RISK CLASSIFICATION FOR LOSS RESERVING

DAMLA BARLAS

Hacettepe Üniversitesi
Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin
AKTÜERYA BİLİMLERİ Anabilim Dalı için Öngördüğü
YÜKSEK LİSANS TEZİ
Olarak Hazırlanmıştır.

2010

Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Bu çalışma jürimiz tarafından **AKTÜERYA BİLİMLERİ ANABİLİM DALI** 'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan :.....
Doç.Dr. İbrahim ÖZKAN

Üye (Danışman) :.....
Prof. Dr. Ömer ESENSOY

Üye :.....
Doç.Dr. Meral SUCU

ONAY

Bu tez/.../2010 tarihinde Enstitü Yönetim Kurulu 'nca kabul edilmiştir.

Prof.Dr. Adil DENİZLİ
Fen Bilimleri Enstitüsü Müdürü

HASAR REZERVİ İÇİN RİSK SINIFLANDIRMASI

Damla BARLAS

ÖZ

Çalışmada, karar ağaçlarının sigortacılıkta risk sınıflandırmasında kullanımı ve hasar rezervine etkisi üzerinde durulmuştur. Risk sınıflandırması, benzer hasar özelliklerine sahip sigortalıların aynı sınıfta toplanmasıdır. Hasar rezervi ise, sigorta şirketinin henüz ödenmemiş hasarlar ile meydana gelmiş, ancak sigorta şirketine bildirilmemiş hasarlar için ayırmış olduğu karşılıklardır.

Çalışmada 2005-2008 yılları arasında zorunlu trafik sigortasına ilişkin dört yıllık hasar verisi, ödenmiş hasar tutarı, il plaka kodu, araç türü, araç modeli, yakıt türü, sürücünün cinsiyeti ve sürücünün yaşı değişkenleri dikkate alınarak türetilmiştir. Ödenmiş hasar tutarı bağımlı değişken, diğer değişkenler ise bağımsız değişkenler olarak alınmıştır. Bu veri kümesi, yıllar itibariyle CHAID ve CRT algoritmaları kullanılarak sınıflandırılmış ve sınıflandırmanın sonuçları analiz edilerek karşılaştırılmıştır. CRT algoritmasının sonuçları kullanılarak risk grupları belirlenmiş ve gruplar arası farklılıklara değinilmiştir.

Çalışmanın sonraki aşamasında, oluşturulan risk gruplarının hasar rezervine etkisi incelenmiştir. Bir portföyde bulunan sigortalılar için risk sınıflandırması yapılmadan tahmin edilen hasar rezervi ile risk sınıflandırması yapılarak tahmin edilen hasar rezervi karşılaştırılmıştır. Hasar rezervi hesaplamasında Mack Zincir Merdiven Yöntemi kullanılmıştır.

Sonuç olarak, risk sınıflandırması ile gruplar arasında hasar tutarları bakımından farklılık olduğu ve sınıflandırma yapılarak tahmin edilen hasar rezervinin, risk sınıflandırması yapılmadan tahmin edilen hasar rezervinden daha yüksek olduğu görülmüştür.

Anahtar Kelimeler: Veri Madenciliği, Risk Sınıflandırması, CHAID Algoritması, CRT Algoritması, Hasar Rezervi, Mack Zincir Merdiven Yöntemi.

Danışman: Prof.Dr. Ömer ESENSOY, Hacettepe Üniversitesi, Aktüerya Bilimleri Bölümü.

RISK CLASSIFICATION FOR LOSS RESERVING

Damla BARLAS

ABSTRACT

Decision trees is one of the important models of data mining and it is used in a widespread area. This thesis is concern with the usage of decision trees in insurance risk classification system and the effects of risk classification on loss reserving. Risk classification means categorizing insurers according to their similar risk characteristics for the purpose of a workable insurance system. Loss reserve is a figure that represents an insurance company's best estimate of what future losses will be.

Simulated data set including city, vehicle and fuel category etc. between the years 2005-2008 is used in application. CHAID and CRT algorithms are used for classification and their results are analysed and compared to each other. According to results of decision trees, there are differences between risk groups in terms of their risk characteristics and loss amounts.

The next phase of the study is to investigate the effects of risk classification on loss reserving.

Loss reserves are calculated using Mack Chain Ladder Method before and after risk classification. It is concluded that total loss reserves of risk groups is greater than loss reserves which is calculated without risk classification.

Key Words: Data Mining, Risk Classification, CHAID Algorithm, CRT Algorithm, Loss Reserving, Mack Chain Ladder

Advisor: Prof.Dr. Ömer ESENSOY, Hacettepe University, Department of Actuarial Sciences.

TEŐEKKÜR

Tez alıőmamda deęerli fikir ve önerileriyle bana yol gösteren ve destek olan danıőmanım Sayın Prof.Dr. Ömer ESENSOY 'a,

alıőma süresince yardım ve hoşgörülerini esirgemeyen deęerli hocalarıma ve alıőma arkadaşlarıma,

alıőmamda bana destek olan ok deęerli arkadaşlarım Mehmet PIRILDAK 'a, Sema TÜZEL 'e, Uęur KARABEY 'e ve Tuęba TUN 'a,

Her koőulda ve her zaman yanımda olan canım aileme,

Teőekkürlerimi sunarım.

İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZ.....	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER DİZİNİ.....	iv
ŞEKİLLER DİZİNİ.....	v
ÇİZELGELER DİZİNİ.....	vi
1. GİRİŞ.....	1
2. VERİ MADENCİLİĞİ.....	4
2.1. Veri Tabanlarında Bilgi Keşfi.....	4
2.1.1. Veri Tabanlarında Bilgi Keşfi Süreci.....	5
2.1.1.1. Problemin Tanımlanması.....	6
2.1.1.2. Verinin Hazırlanması.....	6
2.1.1.3. Modelin Kurulması ve Değerlendirilmesi.....	7
2.1.1.4. Modelin Kullanılması ve İzlenmesi.....	8
2.2. Veri Madenciliği Modelleri.....	9
2.2.1. Sınıflama ve Regresyon Modelleri.....	10
2.3. Veri Madenciliği'nin Sigortacılık Sektörü'nde Kullanım Alanları.....	11
3. RİSK SINIFLANDIRMASI.....	14
3.1. Risk Sınıflandırmasında Kullanılan Yöntemler.....	15
3.1.1. Karar Ağaçları.....	16
3.1.2. Karar Ağacının Yapısı.....	16
3.1.3. CHAID Algoritması.....	19
3.1.4. CRT Algoritması.....	20
4. HASAR REZERVİ.....	22
4.1. Hasar Gelişim Üçgeni.....	22
4.2. Hasar Rezervi Tahmininde Kullanılan Yöntemler.....	26
4.2.1. Zincir Merdiven Yöntemi.....	27
4.2.1.1. Mack Zincir Merdiven Yöntemi.....	28
4.2.1.2. Münih Zincir Merdiven Yöntemi.....	29
4.2.2. Bornhuetter-Ferguson Yöntemi.....	29
4.2.3. Cape-Cod Yöntemi.....	31
4.2.4. Hasar Gelişim Yöntemi.....	31
4.2.5. Toplamsal Yöntem.....	32
4.3. Hasar Rezervi Tahmininde Kullanılan Yöntemlerin Karşılaştırılması.....	33
5. UYGULAMA.....	36
5.1. Trafik Sigortası (Karayolları Motorlu Araçlar Zorunlu Mali Sorumluluk Sigortası).....	36
5.2. Veri Kümesinin Oluşturulması.....	37
5.3. Trafik Sigortası İçin Risk Sınıflandırması.....	45
5.4. Hasar Rezervi Tahmini.....	49
6. SONUÇLAR VE ÖNERİLER.....	57
KAYNAKLAR.....	60
EKLER DİZİNİ.....	63
ÖZGEÇMİŞ.....	79

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 2.1. VTBK süreci.....	6
Şekil 2.2. Başlıca VM modelleri	10
Şekil 3.1. Karar ağacının genel yapısı.....	17
Şekil 3.2. Bağımlı değişken türün göre karar ağaçları.....	18
Şekil 5.1. Lognormal dağılımın olasılık yoğunluk fonksiyonu grafikleri.....	39
Şekil 5.2. Kaza yıllarına göre hasar rezervi tutarı.....	56

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 2.1. 1960-2000 yılları arasında veri tabanı teknolojisinin gelişim süreci.....	5
Çizelge 4.1. Aşamalı hasar tutarına göre hasar gelişim üçgeni.....	23
Çizelge 4.2. Birikimli hasar tutarına göre hasar gelişim üçgeni.....	24
Çizelge 4.3. Birikimli hasar tutarlarına göre tamamlanmış hasar gelişim üçgeni.....	25
Çizelge 4.4. Hasar rezervi tahmininde kullanılan yöntemlerin karşılaştırılması.....	33
Çizelge 5.1. Yıllara göre TRAMER'den elde edilen toplam poliçe adedi ve ödenmiş hasar tutar.....	38
Çizelge 5.2. 2005 yılı için illere göre lognormal dağılımın parametre değerleri.....	40
Çizelge 5.3. 2006 yılı için illere göre lognormal dağılımın parametre değerleri.....	40
Çizelge 5.4. 2007 yılı için illere göre lognormal dağılımın parametre değerleri.....	41
Çizelge 5.5. 2008 yılı için illere göre lognormal dağılımın parametre değerleri.....	41
Çizelge 5.6. Yıllara göre türetilen toplam poliçe adedi ve ödenmiş hasar tutarı.....	42
Çizelge 5.7. Bağımlı değişken ve bağımsız değişkenler.....	42
Çizelge 5.8. İl plaka kodu oranları.....	43
Çizelge 5.9. Araç türü oranları.....	44
Çizelge 5.10. Sürücünün cinsiyeti oranları.....	44
Çizelge 5.11. Yakıt türü oranları.....	45
Çizelge 5.12. Yıllar itibariyle CHAID ve CRT algoritmalarının veri kümesi ile doğrusal ilişki sonuçları.....	46
Çizelge 5.13. 2005, 2006 ve 2007 yılları için yapılan risk sınıflandırmasında ilk derinlikteki sınıflar.....	47
Çizelge 5.14. 2008 yılı için yapılan risk sınıflandırmasında ilk derinlikteki sınıflar.....	47
Çizelge 5.15. Yıllara göre CRT algortimasının kök düğüm ve ilk derinlikteki düğüm bilgileri.....	48
Çizelge 5.16. Risk sınıflandırması yapılmadan hasar gelişim üçgeni (aşamalı hasar tutarına göre) (TL).....	50
Çizelge 5.17. Birinci risk grubu için hasar gelişim üçgeni (aşamalı hasar tutarına göre) (TL)	51
Çizelge 5.18. İkinci risk grubu için hasar gelişim üçgeni (aşamalı hasar tutarına göre) (TL)	51
Çizelge 5.19. Risk sınıflandırması yapılmadan hasar gelişim üçgeni (birikimli hasar tutarına göre) (TL)	52
Çizelge 5.20. Birinci risk grubu için hasar gelişim üçgeni (birikimli hasar tutarına göre) (TL).....	53
Çizelge 5.21. İkinci risk grubu için hasar gelişim üçgeni (birikimli hasar tutarına göre) (TL)	53

ÇİZELGELER DİZİNİ (Devam)

	<u>Sayfa</u>
Çizelge 5.22. Mack Zincir Merdiven Yöntemi 'ne göre risk sınıflandırması yapılmadan tahmin edilen rezerv miktarı (TL)	54
Çizelge 5.23. Mack Zincir Merdiven Yöntemi'ne göre birinci risk grubu (kamyon, minibüs, otobüs) için tahmin edilen rezerv miktarı (TL)	55
Çizelge 5.24. Mack Zincir Merdiven Yöntemi'ne göre ikinci risk grubu (kamyonet, otomobil, diğer araçlar) için tahmin edilen rezerv miktarı (TL)	55

1. GİRİŞ

Teknolojik gelişmeler bir çok işlemin elektronik olarak kayıt altına alınmasını, bu kayıtların kolayca saklanabilmesini, gerektiğinde erişilebilmesini kolaylaştırmakta ve bu işlemlerin her geçen gün daha ucuza mal edilmesini sağlamaktadır. Otomatik veri toplama araçları ve veri tabanı teknolojilerindeki gelişmeye bağlı olarak, veri tabanlarında, veri ambarlarında ve diğer bilgi depolarında çok miktarda bilgi depolanmaktadır. Ancak veri tabanlarında saklanan birçok veriden anlamlı çıkarımlar yapabilmek için bu verilerin, uzmanlarca özel araçlar ve yöntemler kullanılarak analiz edilmesi gerekir. Analizlerde verilen kararların doğruluğu, kararı veren kişinin yeteneklerine ve deneyimine olduğu kadar sahip olduğu bilginin yeterliliğine de bağlıdır. Bu nedenle “bilgi”, mal ve hizmetin yanında üçüncü üretim faktörü olarak değerlendirilir. Doğru analizler için bilginin güvenilir koşullarda depolanması, bilimsel bir metodoloji ile işlenmesi ve elde edilen sonuçların yansız yorumlanması gerekir. Bu ihtiyaçlar doğrultusunda bilgi teknolojilerinde yaşanan gelişmeler ile depolanan verinin boyutlarının artması, Veri Tabanlarında Bilgi Keşif (VTBK) Süreci ve bu süreç içinde yer alan Veri Madenciliği (VM) 'nin gelişmesini sağlamıştır.

VM 'de yaşanan gelişmeler nedeniyle son yıllarda aktüerya literatüründe de VM kullanılarak çeşitli çalışmalar yapılmıştır. Bu çalışmalardan konuyla ilgili bazıları aşağıda verilmiştir.

Kahane et al.(2007) tarafından yapılan çalışmada, araba sigortası için sigorta priminin hesaplanmasında, tarifelerin oluşturulması ve gelecekte oluşacak hasar tutarının tahmin edilmesinde VM teknikleri kullanılmıştır. Doğrusal regresyon modeli (lineer regression model) ile yaşam analizi modeli (survival analysis model) kullanılarak sigortalılar risk seviyelerine göre sınıflandırılmışlardır. Sınıflar arasındaki farklılıklar incelenerek, prim tutarı hesaplanmıştır. Çalışmada sigortalı bilgileri (adresi, yaşı vb.), araç bilgileri (modeli, güvenlik sistemi, araç değeri vb.), poliçe bilgileri (poliçenin başlangıç tarihi, bitiş tarihi, yenileme tarihi, prim tutarı vb.) ile hasar bilgileri (hasar türü, hasar tarihi vb.) kullanılmıştır. Bu değişkenler analizde kullanılarak risk sınıflarına ilişkin prim tutarı modellenmiştir. Elde edilen tahmini prim tutarları gerçek değerleri ile karşılaştırılmış ve sonuçları

yorumlanmıştır. Yapılan çalışmada, her risk grubu için prim tutarına etki eden değişkenler risk sınıflandırması ile ortaya koyulmuştur.

Guo (2003) tarafından yapılan çalışmada mal / kaza sigortası (property / casualty insurance) ele alınmıştır. Sürücü ve araç özelliklerine ilişkin değişkenler (sürücünün cinsiyeti, yaşı, eğitim durumu, kredi skoru, aracın modeli, hasar sayısı vb.) kullanılarak, risk faktörleri belirlenmiş ve her poliçeye ilişkin hasar frekansına göre sigortalılar sınıflandırılmıştır. Bu çalışmanın amacı, kaza sigortalarında VM yaklaşımı kullanılarak risk modellenmesidir. K ortalamalar kümeleme (K means clustering) ile sigortalılar için risk sınıflandırması yapılmıştır. Etkileşimler ve korelasyon veri madenciliği yöntemlerinden Keşfedici Veri Çözümlemesi (Exploratory Data Analysis) ve karar ağaçları ile açıklanmış, hasar frekansı ise lojistik regresyon ile modellenmiştir.

Wang (2003) kaza sigortalarına ilişkin sigortalıları karar ağaçları, sinir ağları ve lojistik regresyon kullanılarak sınıflandırıp, yapılan sınıflandırma sonuçlarını karşılaştırmıştır. Klasik yöntemler ile yapılan analizlerde önemsiz gibi görünen değişkenlerin, VM yöntemleriyle yapılan analizlerde önemli olduğunu belirtmiştir. Çalışmada 23 değişken ve 300.000 poliçeye ait bilgi kullanılmıştır. Araç modeli, araç türü, poliçe başlangıç ve bitiş tarihi, sürücü cinsiyeti, sürücü yaşı vb. değişkenler analizde kullanılmıştır.

Amacı, VM modellerinden sınıflandırma kullanılarak, sigortacılıkta risk sınıflandırması yapmak ve sınıflandırmanın hasar rezervi üzerindeki etkisini incelemek olan bu çalışmanın Birinci Bölümünde, VM 'ye giriş yapılmış ve literatür bilgisine yer verilmiştir. İkinci Bölümde, VM ele alınarak VM modellerine değinilmiştir. Risk sınıflandırması ve karar ağaçlarına Üçüncü Bölümde yer verilmiştir. Karar ağaçlarından yaygın olarak kullanılan CHAID (Chi-Squared Automatic Interaction Detector) ve CRT (Classification and Regression Trees) algortimaları bu bölümde açıklanmıştır. Dördüncü Bölümde, hasar rezervi ve tahmin yöntemlerine değinilmiştir. Beşinci Bölümde, trafik sigortası için türetilen poliçe bilgilerine CHAID ve CRT algoritmaları uygulanıp, risk sınıfları belirlenmiştir. Risk sınıfları, portföydeki paylarına ve ortalama hasar tutarlarına göre karşılaştırılmış ve yorumlanmıştır. Risk sınıfları belirlendikten sonra risk sınıflarına ilişkin hasar rezervi Mack Zincir Merdiven Yöntemi ile tahmin edilerek,

sınıflandırmanın rezerve etkisi incelenmiş ve sınıflandırmadan önce tahmin edilen rezerv ile sınıflandırmadan sonra tahmin edilen rezerv karşılaştırılmıştır. CHAID ve CRT algoritmaları SPSS Clementine 11.1 ve Mack Zincir Merdiven Yöntemi R 2.11.0 programında Chain-Ladder Package kullanılarak uygulanmıştır. Altıncı Bölüm ise sonuçlar ve önerilere ayrılmıştır.

İKİNCİ BÖLÜM

2. VERİ MADENCİLİĞİ

2.1. Veri Tabanlarında Bilgi Keşfi

Veri, kendi başına değersiz olan ancak bir amaç doğrultusunda kullanılan bilgidir. Bilgi ise bir amaca yönelik işlenmiş veridir. Bir diğer ifade ile bilgi, bir soruya yanıt vermek için veriden çıkarılan sonuçtur. Veri tabanı ise, sistematik erişim imkanı olan, yönetilebilir, güncellenebilir, taşınabilir, birbirleri arasında tanımlı ilişkiler bulunabilen bilgiler kümesidir. Kısaca belirli bir amaca yönelik düzenlenmiş kayıt ve dosyaların tümüdür. Gelişen teknoloji ile veri tabanı sistemlerinin aşamalı kullanımı ve hacimlerdeki artış, organizasyonları elde toplanan bu veriden nasıl faydalanılabileceği problemi ile karşı karşıya bırakmıştır. Geleneksel sorgu (Query) veya raporlama araçlarının veri yığınları arasında yetersiz kalması Veri Tabanlarında Bilgi Keşfi-VTBK (Knowledge Discovery in Databases) adı altında sürekli ve yeni arayışlara neden olmaktadır. Veri tabanlarında bilgi keşfi, verilerin doğru, faydalı ve anlaşılır modeller ve örüntüler elde etmede kullanılan özel bir süreçtir. VTBK sürecinde modelin kurulması ve değerlendirmesi aşaması VM olarak adlandırılır (Larose, 2005). VM, çok büyük veri tabanlarındaki ya da veri ambarlarındaki veriler arasında bulunan ilişkiler, örüntüler, değişiklikler, sapma ve eğilimler, belirli yapılar gibi bilgilerin ortaya çıkarılması ve keşfi işlemidir. VM, VTBK uygulamaları ile birlikte faaliyet alanına yönelik karar destek mekanizmaları için gerekli ön bilgileri temin etmek için kullanılır. Sürecin en önemli kısmını oluşturan VM ile VTBK terimleri birçok araştırmacı tarafından eş anlamlı olarak da kullanılmaktadır.

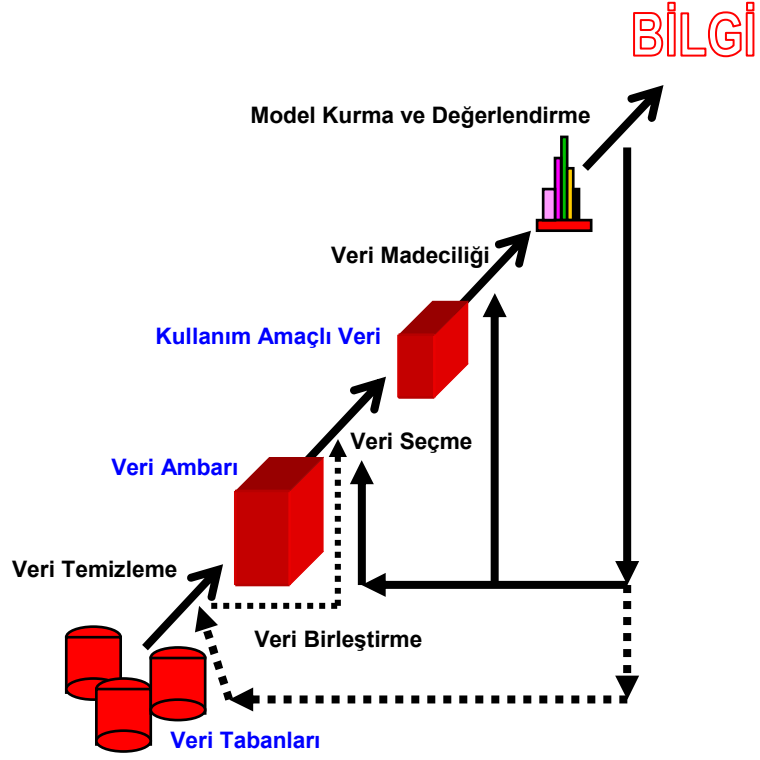
Geçmişten bugüne her zaman veri yorumlanıp, veriden bilgi edinilmeye çalışılmış ve buna bağlı olarak çeşitli donanımlar oluşturulmuştur. Bu donanımlar bilginin taşınmasını sağlamıştır. Veri tabanı teknolojisi ilk kez 1960 'lı yıllarda ortaya çıkmış ve 1990 'lı yıllardan itibaren bilgisayar teknolojisindeki gelişmeye bağlı olarak bilgi teknolojileri tabanlı birçok alanın değişmez bir parçası olmuştur. Zamanla her alanda bilgi toplanmaya başlanarak veri tabanı teknolojisi yıldan yıla gelişim göstermiştir. Veri tabanı teknolojisinin 1960-2000 yılları arasında gelişim süreci kronolojik olarak Çizelge 2.1 'de özetlenmiştir.

Çizelge 2.1. 1960-2000 yılları arasında veri tabanı teknolojisinin gelişim süreci

Gelişim Adımları	Kullanılan Teknolojiler	Ürün Sağlayıcıları	Özellikler
Veri Toplama (1960'lar)	Bilgisayarlar, teypler, diskler	IBM, CDC	Geriye dönük, statik veri dağıtımı
Veri Erişimi (1980'ler)	İlişkisel veri tabanları, SQL, ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Kayıt düzeyinde geriye dönük, dinamik veri dağıtımı
Veri Ambarlama ve Karar Destek Sistemleri (1990'lar)	OLAP, çok boyutlu veri tabanı sistemleri, veri ambarları	Pilot, Comshare, Arbor, Cognos, Microstrategy	Çoklu düzeylerde, geriye dönük, dinamik veri dağıtımı
Veri Madenciliği (2000'ler)	İleri düzeyde algoritmalar, çok işlemcili bilgisayarlar, büyük veri tabanları	Pilot, Lockheed, IBM, SGI, SPSS, SAS, Microsoft vs.	Geleceğe dönük, proaktif enformasyon dağıtımı

2.1.1. Veri Tabanlarında Bilgi Keşfi Süreci

Veri tabanlarından bilgi üretme sürecinde dikkat edilmesi gereken unsurlar bulunmaktadır. Öncelikle analizlerin güvenilir veriye dayandırılması, yapılacak herhangi bir analizin başkaları tarafından da tekrarlanabilir ve sonuçların çözüm odaklı olması gerekmektedir. Bu nedenle, veri tabanındaki verinin hangi süreç ile ve nasıl toplanacağına karar verme aşaması VTBK sürecinin en önemli parçasını oluşturmaktadır. VTBK süreci, problemin tanımlanması, verilerin hazırlanması, modelin kurulması ve değerlendirilmesi ve modelin kullanılması ve izlenmesi adımlarından oluşmaktadır (Han & Kamber, 2006). VTBK sürecinin adımları Şekil 2.1 'de verilmektedir.



Şekil 2.1. VTBK süreci (Han & Kamber, 2006).

2.1.1.1. Problemin Tanımlanması

VM çalışmalarının başarılı olabilmesi için uygulamanın hangi amaçla yapılacağına açık bir şekilde tanımlanması, bir başka deyişle problem üzerine odaklı olması gerekmektedir. Bu aşamada elde edilecek sonuçların nasıl yorumlanacağı açıklanmalı ve sonuçların tahminlerine yer verilmelidir. Problem tanımlandıktan sonra, uygulama adımlarına geçilebilir (Akpınar, 2000).

2.1.1.2. Verinin Hazırlanması

Bu aşama başlangıç verisinin, modeller için kullanılacak veriye dönüştürülmesi aşamasıdır. Bu aşamanın belirgin bir sırası veya tekrar sayısı yoktur. Modelin kurulması aşamasında ortaya çıkacak sorunlar, bu aşamaya sık sık geri dönülmesine ve verilerin yeniden düzenlenmesine neden olabilmektedir. Bu aşama karar vericinin veri keşfi sürecinin toplamı içerisindeki enerji ve zamanının %50 'sinden fazlasını harcamasına neden olmaktadır. Verinin hazırlanması

aşaması toplama, temizleme, birleştirme ve dönüştürme adımlarından oluşmaktadır (Dolgun, 2006).

Veri Toplama: Mevcut olan ve elde edilebilecek verinin oluşturduğu veri tabanları ya da veri ambarlarından çalışma için kullanılacak verinin seçilmesidir.

Veri Temizleme: Tutarsız, yanlış girilen ya da aykırı verinin, değerlendirmeden çıkarılması tercih edilir. Bazı durumlarda çok büyük veri tabanı ile çalışmak yerine örnekleme yapılması uygun olabilir (seçilen örneklerin tüm popülasyonu temsil etmesi önemlidir). Verinin temizleme aşaması, sadece veri çıkarmak gibi düşünülmemelidir, bazı durumlarda eksik verinin tamamlanması için modellemeler yapılmasını da kapsamaktadır (Dolgun, 2006; Han & Kamber, 2006).

Veri Birleştirme: Verinin farklı kaynaklardan toplanması, doğal olarak veri uyumsuzluklarına neden olmaktadır. Bu uyumsuzlukların başlıcaları verinin farklı zamanlara ait olması, güncelleme hataları, veri formatlarının farklı olması, kodlama farklılıkları, farklı ölçü birimleri ve varsayım farklılıklarıdır. Bu adımda farklı kaynaklardan toplanan verinin uyumsuzlukları mümkün olduğu ölçüde giderilerek, tek bir veri tabanında toplanması amaçlanır (Dolgun, 2006; Han & Kamber, 2006).

Veri Dönüştürme: Çözümleme için kullanılması düşünülen veriye ilişkin değişkenlerin uygun biçime dönüştürülmesi aşamasıdır.

2.1.1.3. Modelin Kurulması ve Değerlendirilmesi

Model kurma VM 'nin en önemli yapı taşıdır. Modeli doğru bir biçimde kurabilmek için yapılacak çalışmanın amacı çok iyi kavranmış olmalıdır. Her amaç ile ilgili birden fazla model mevcuttur. Bu durumda eldeki veri kümesine ilişkin modellerin hepsi çalıştırılır ve en uygun model seçilir. Sonuç olarak tanımlanan problem için en uygun modelin bulunabilmesi, olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir (Westphal & Blaxton, 1998; Guo, 2003).

2.1.1.4. Modelin Kullanılması ve İzlenmesi

Modelin kullanılması, kurulan ağ yapının tamamlanması ve sonuçların incelenmesi sürecidir. Modelin izlenmesi, kullanılan modelin test sonuçları ile hata yüzdesinin uygunluğunun yorumlanması ve kullanılan modelin farklılaşan verilere ne gibi tepkiler verdiğinin gözlenmesi ve elde edilen bilgilerin kullanıcılara sunum sürecidir (Berry & Linoff, 2004; Larose, 2005).

VTBK sürecinin en önemli parçası olan VM, eldeki veri kümesinden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı ve anlamlı bilginin çıkarılması sürecidir. VM, veri analiz teknikleri bütünüdür ve tek başına bir çözüm değildir. Mevcut problemleri çözmek, kritik kararları almak veya geleceğe yönelik tahminleri yapmak için gerekli olan bilgileri elde etmeye yarayan bir araçtır. VM süreci, veri yığınlarını elden geçirmekle başlayarak, analiz sonucunda ortaya çıkan sonuçların uzman gözüyle yorumlanması ile tamamlanır. VM 'nin yaygın olmasının sebebi, çok büyük miktardaki veriye kolay erişim ve verinin kullanılabilir bilgiye dönüştürülmesindeki ihtiyaçtır (Berry & Linoff, 2004).

VM teknikleri kullanarak; maliyetleri azaltmak, gelirleri artırmak, verimliliği artırmak, yeni fırsatları ortaya çıkarmak, yeni keşifler yapmak, emek yoğun faaliyetleri otomatikleştirmek, sahtekarlıkları belirlemek ve müşteri deneyimini geliştirmek mümkündür (Han & Kamber, 2006).

VM 'nin başlıca özellikleri özellikleri aşağıdaki biçimde sıralanabilir :

- Büyük ve karmaşık veri kümeleri ile çalışır.
- Her türlü veriyi kullanarak çözümler üretebilir.
- İstatistik, yapay zeka, makine öğrenmesi, veri tabanlarında bilgi keşfi, bilgisayar bilimi, yapı tanıma vb. gibi disiplinlerden faydalanır.
- Daha önceden bilinmeyen, doğrulanabilir, etkinleştirilebilir enformasyon arar.
- Otomatik veya yarı otomatik olarak çalışan çözüm araçları kullanır.
- Sorunlara göre değişen çözüm araçları vardır.

- Hızla büyümekte olan bir sektördür.

Günümüzde VM teknikleri özellikle işletmelerde ve çeşitli alanlarda (pazarlama, bankacılık, borsa vb.) başarı ile kullanılmaktadır (Westphal & Blaxton, 1998; Akpınar, 2000).

2.2. Veri Madenciliği Modelleri

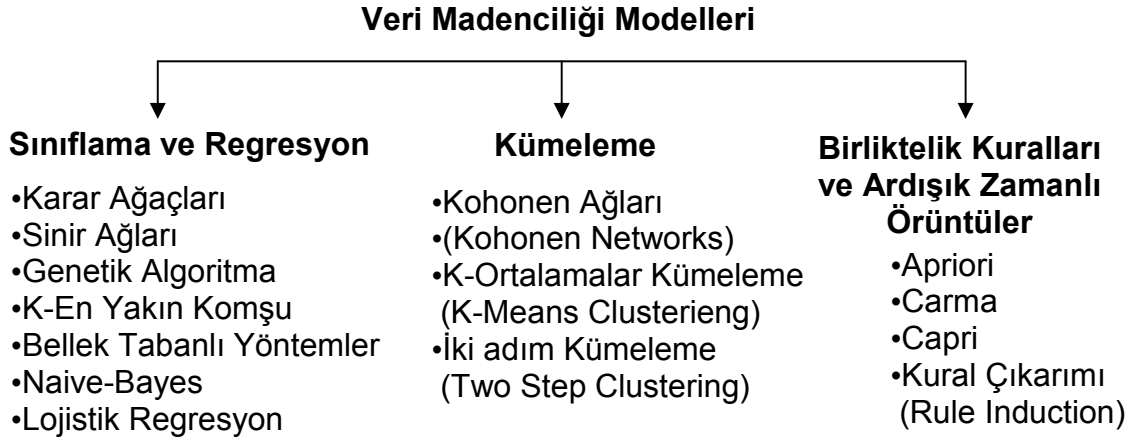
Kuralcı (normative), tahmin edici (predictive) ve tanımlayıcı (descriptive) olmak üzere üç temel matematiksel model vardır. Normative modeller bir teoriye dayanır ve gözlemler üzerinde tasarım yapılır. Tahmin edici modeller, sonuçları bilinen veriden hareket edilerek bir model geliştirilmesi ve kurulan bu modelden yararlanılarak, sonuçları bilinmeyen veri kümeleri için sonuç değerlerin tahmin edilmesi amaçlamaktadır (Dolgun, 2006). Tanımlayıcı modeller ise karar vermeye rehberlik etmede kullanılacak mevcut verideki bilgilerin tanımlanmasını sağlamaktadır. VM 'de kullanılan modeller, tahmin edici (predictive) ve tanımlayıcı (descriptive) modellerdir. VM modelleri kuralcı değildir, bir teoriye dayanmamaktadır. Ancak sonuçlar sadece istatistiksel olarak değerlendirilmeyip teoriye göre de sonuçlar değerlendirilebilir.

VM modellerinin önemli işlevlerinden bazıları aşağıda verilmektedir:

1. Sınıflama (Classification) ve Regresyon,
2. Kümeleme (Clustering),
3. Birliktelik Kuralları (Association Rules) ve Ardışık Zamanlı Örüntüler (Sequential Patterns),

Sınıflama ve regresyon modelleri tahmin edici, kümeleme, birliktelik kuralları ve ardışık zamanlı örüntü modelleri ise tanımlayıcı modellerdir (Han & Kamber, 2006).

Şekil 2.2 'de VM 'nin başlıca modelleri verilmektedir.



Şekil 2.2. Başlıca VM modelleri (Larose, 2005).

2.2.1. Sınıflama ve Regresyon Modelleri

Mevcut veriden hareket ederek geleceğin tahmin edilmesinde de yararlanılan ve VM teknikleri içerisinde yaygın kullanıma sahip olan sınıflama ve regresyon modelleri arasındaki temel fark, tahmin edilen bağımlı değişkenin kategorik veya süreklilik gösteren bir değere sahip olmasıdır. Ancak çok terimli lojistik regresyon gibi kategorik değerlerin de tahmin edilmesine olanak sağlayan tekniklerle, her iki model birbirine yaklaşmaktadır (Carvalho & Freitas, 2004). Çalışmada sınıflama ve regresyon modellerinden karar ağaçları kullanılmıştır.

Karar Ağaçları: Bir değere ya da sınıfa ulaşan kurallar dizisiyle ifade edilen bir yoldur. Karar ağaçlarının karar düğümleri, dallar ve yapraklar olmak üzere üç bileşeni vardır. Bir karar ağacında ilk bileşen, en üstte yer alan test için belirlenmiş olan kök düğümüdür. Test sonuçlarına göre kök düğüm dallara ayrılır. Karar ağaçlarının büyümesi durdurma kuralları ile kontrol edilebilir. Ağacın büyüebildiği maksimum derinlik ya da her düğümdeki kayıt sayısı için bir limit belirlenebilir (Ciarapica & Giacchetta, 2009).

Karar ağaçlarına ayrıntılı olarak Üçüncü Bölümde yer verilmiştir.

2.3. Veri Madenciliği'nin Sigortacılık Sektörü'nde Kullanım Alanları

Sigortacılık, ortaya çıkması muhtemel risklerin tanım ve hesabının yapılması ve bu riskler için fiyatların oluşturulması ile poliçe şart ve koşullarının belirlenmesi tekniğine dayanır. Sigorta primi içerisinde; maruz olunan riske yönelik hesaplanmış risk primi, çeşitli vergi, komisyonlar ve belirli varsayımlara dayanan şirketin diğer giderleri yer almaktadır. Risk primi, şirket harcamaları, komisyonlar ile çeşitli varsayımsal giderler dışında, teminat altına alınan riskin beklenen değeridir.

Risk priminin doğru ve adil olarak hesaplanabilmesi için bilgi işlem altyapılarının iyi olması, yeterli istatistiki bilginin olması gereklidir. Örnek olarak, risk primi hesabına yönelik, bir trafik sigortasında ana konu itibarıyla riske maruz olan araç olup ayrıca işletenin riskleri ile kullanım şekli de yapılacak analize dahil edilmelidir. Dolayısıyla risk priminin hesaplanmasında aracın kaza yapma ve kaza neticesinde hasar ödemesinin ortaya çıkması önem arz etmektedir. Sigorta şirketleri genellikle risk priminin hesaplanmasında belli risk faktörlerini dikkate almaktadırlar. Bu faktörler; aracın cinsi, yaşı, kullanım şekli, donanımı, işletene ait bilgiler gibi olabilmektedir. Risk faktörlerinin sayısını arttırmak mümkündür. Risk faktörlerinin alt ve üst limitlerini belirlemek şirketin bu branştaki yeterli teknik ve idari donanımı ile orantılıdır (A SAS White Paper, 2002).

VM, sigortacılık sektöründe risk sınıflandırmasında, fiyatlandırmada (ratemaking), risk kabul sürecinde (underwriting), reasürans uygulamalarında, sigorta dolandırıcılıklarının tespitinde, riskli müşterilerin ve risk faktörlerin belirlenmesinde kullanılır (Guo, 2003; A SAS White Paper, 2002).

Aktüerya uygulamaları açısından kullanım alanlarının başlıcaları aşağıda verilmektedir:

Risk Sınıflandırması (Risk Classification):

Risk sınıflandırması benzer risk özelliklerine sahip sigortalıların gruplandırılmasıdır. Sınıflandırma bir riskin hangi prim düzeyinde kabul edilebilirliğinin belirlenmesinde temeldir. Sigortacı kendi deneyimlerine göre sınıf sayısını değiştirebilir, azaltabilir veya arttırabilir. Ancak her sınıf kendi içinde

homojen özellikte olmalıdır. Risk sınıflandırmasında kullanılan başlıca VM yöntemleri karar ağaçları, kümeleme analizi ve sinir ağlarıdır. Bu yöntemler sayesinde sigortalılar risk özelliklerine göre sınıflandırılır. Sigorta şirketleri açısından uygun sınıflandırma risklerin, primlerin ve hasar tutarlarının tahmin edilmesinde önemlidir (Apte et al., 1999; Guo, 2003).

Fiyatlandırma (Ratemaking):

Rizikoya katılım (iştirak) payını ifade eden tutardır. Bir başka ifade ile sigortacının üzerine aldığı rizikoya karşılık, sigorta şirketine veya teminatı veren kuruluşa para olarak ödenen bedeldir. Primlerin kayıplara, yapılan harcamalara ve kara göre belirlenmesi süreci ise fiyatlandırma değildir. VM modelleri ile prim tutarları sigortalıların risk özelliklerine göre modellenebilir (Kahane et al., 2007; Werner & Modlin, 2010).

Risk Kabul Süreci (Underwriting):

Sigortacı tarafından yapılan risk değerlendirmesi sonucunda sigortalanması teklif edilen riskin, sigorta poliçesi ile teminat altına alınıp alınmamasına veya alınacaksa da hangi şartlarda alınması gerektiğine karar verme işlemine denir. VM modellerinden ilişki analizi (association analysis) ile hangi poliçenin hangi müşteriye uygun olacağı tespit edilebilir. Sigortalının ek olarak alacağı diğer sigorta ürünleri belirlenmesinde pazar sepeti analizi kullanılır (A SAS White Paper, 2002).

Reasürans:

Sigorta edilmiş riskin belli bir kısmı veya tamamının yeniden sigorta edilmesidir. Sigorta şirketi üstlendiği risklerin bir kısmını ülke içindeki veya dışındaki reasürans şirketlerine kendilerini güvenceye almak için devreder. Reasürans, sigorta şirketlerine tek başlarına yüklenmeleri mali açıdan mümkün olmayan riskleri, sigortalama imkanı vermektedir. Birtakım mali ölçülerle, branş esasıyla tespit edilen saklama payları, sedan şirketin risk üzerindeki sorumluluk miktarını göstermekte olup, aşan kısımlar ise, çeşitli reasürans anlaşmalarıyla reasürans şirketlerine devredilmektedir. Reasürans uygulamalarında bir grubun ödenmiş hasar tutarları, diğer bir grubun beklenen hasar tutarlarının modellenmesinde kullanılmaktadır. Reasürans için poliçelerin seçimi risklere bağlıdır ve bu

poliçelerin seçimi VM modelleri ile yapılabilir (Guo, 2003; A SAS White Paper, 2002).

Sigorta şirketlerinin karşı karşıya buldukları en önemli sorunlardan biri sigorta dolandırıcılıklarıdır. VM modellerinden ilişki analizi (link analysis) ve sapma tespiti (deviation detection) ile beklenen hasar tutarları ve oluşan hasar tutarları karşılaştırılarak, yüksek sapmaya sahip poliçeler ortaya kolayca çıkarılabilir ve incelenebilir (A SAS White Paper, 2002).

Ayrıca riskli müşterilerin ve risk faktörlerinin belirlenmesinde VM modelleri kullanılabilir.

VM 'nin sigorta sektöründe kullanılmasının en önemli nedeni, VM modelleri ile aktüeryal uygulamaların çok yönlü olarak incelenebilmesidir. Çalışmada sigortacılıkta risk sınıflandırmasında VM modellerinden karar ağaçları kullanılmış ve risk sınıflandırmasının sonuçları incelenmiştir. Üçüncü Bölümde sınıflandırma yöntemlerinden karar ağaçlarına değinilmiştir. CHAID ve CRT algoritmaları incelenmiştir.

ÜÇÜNCÜ BÖLÜM

3. RİSK SINIFLANDIRMASI

Riskin Tanımı : Risk, bir olay sonucunda meydana gelen kayıpların olasılığı olarak açıklanabilir. Aktüeryal açıdan risk ise meydana gelen olay sonucunda oluşan olası ekonomik kayıpları ifade etmektedir.

Sigorta, aynı türden tehlikeyle karşı karşıya olan kişilerin, belirli bir miktar para ödemesi yoluyla toplanan tutarın, sadece o tehlikenin gerçekleşmesi sonucu fiilen zarara uğrayanların zararını karşılamada kullanıldığı bir risk transfer sistemidir. Sigortacılıkta adil fiyatlandırma yapılabilmesi için, hasarların olasılıkları, zamanı ve büyüklükleri incelenmelidir. Bu incelemelerin yapılmasında geçmiş bilgiler kullanılmaktadır. Risk sınıflandırması bir riskin hangi prim düzeyinde kabul edilebilir olduğunun belirlenmesinde önemli rol oynar. Risk sınıflandırması, benzer hasarların aynı sınıfta toplanmasıdır (Antonio & Beirlant, 2008; A SAS White Paper, 2002).

Bir sigorta sisteminin kalbi uygun risk sınıflandırması ve fiyatlandırmanın yapılmasıdır. Risk fiyatlandırmasının doğru yapılabilmesi için benzer risk özelliklerine sahip risklerin uygun biçimde sınıflandırması gerekmektedir.

Uygun sigorta sistemini kurmak ve yürütebilmek için risk sınıflandırması üç amaca hizmet etmelidir. Bu amaçlar,

1. Sigorta sisteminin finansal sağlamlığını koruması,
2. Adil olması,
3. Sigorta sisteminin faaliyet gösterebilmesi için ekonomik teşviklere izin vermesi ve böylece daha geniş sigorta kapsamı elde edebilir olmasıdır.

Risk sınıflandırma sistemi aşağıdaki özelliklere sahip olmalıdır:

- Sistem beklenen hasar farklarını ortaya koyabilmelidir,
- Riskler arasındaki farkı, hasarlara dayanarak ayırt edebilmelidir,

- Sistem objektif olarak uygulanabilir olmalıdır,
- Sistem gerçekçi ve uygun maliyetli olmalıdır,
- Sistem toplum için uygun olmalıdır.

Risklerin sigortalanabilmesi için bazı koşulların sağlanması gerekmektedir. Yeterli örneklem büyüklüğü ve homojenlik başlıca koşullardır. Bu iki koşul sağlandığında prim ve hasar tutarlarına ilişkin daha doğru tahminler yapılabilir. Yeterli örneklem büyüklüğü Büyük Sayılar Yasası 'na dayanmaktadır. Çünkü örneklem ne kadar büyükse sonuçlar kitleyi o derece iyi yansıtmaktadır. Homojenlik varsayımı her risk grubunun kendi içinde benzer özelliğe sahip bireylerden oluşmasına dayanmaktadır.

Sigorta şirketleri, sigortalıları bölgeler (territory), demografik özellikler (yaş, cinsiyet, medeni hal gibi) ve diğer değişkenlere (teminat tutarı, sigorta süresi gibi) göre risk sınıflarına ayırırlar. Bu bilgiler hasar tutarlarının tahmininde ve fiyatlandırmada göz önünde bulundurulur (Apte et al., 1999).

3.1. Risk Sınıflandırmasında Kullanılan Yöntemler

Bölüm 2.3.1 'de belirtildiği gibi karar ağaçları sınıflandırmada yaygın olarak kullanılan yöntemlerden biridir. Sınıflama ve regresyon modellerinde kullanılan başlıca sınıflandırma teknikleri aşağıda verilmiştir:

1. Karar Ağaçları (Decision Trees)
2. Yapay Sinir Ağları (Artificial Neural Networks)
3. Genetik Algoritmalar (Genetic Algorithms)
4. K-En Yakın Komşu (K-Nearest Neighbour)
5. Bellek Temelli Nedenleme (Memory Based Reasoning)
6. Naive-Bayes

(Dolgun, 2006; Antonio & Beirlant, 2008).

Çalışmada, sigortacılıkta risk sınıflandırmasında karar ağaçlarından CHAID ve CRT algoritmaları kullanıldığından, bu algoritmalar üzerinde durulmuştur.

3.1.1. Karar Ağaçları

Bir ağaç diyagramı biçiminde olan karar ağaçları, her bir dal ve düğümü bir sınıflandırma sorgusu olacak biçimde dallanan bir yöntem olup, tüm değişkenlere uygulanabilen algoritmalara sahiptir. Tahmin edici ve tanımlayıcı özelliklere sahip olan karar ağaçları, VM 'de kurulum maliyetlerinin düşük olması, yorumlanmalarının kolay olması, veri tabanı sistemleri ile kolayca entegre edilebilmeleri ve daha güvenilir olmaları nedeniyle ile sınıflama modelleri içerisinde yaygın kullanıma sahip olan bir tekniktir. Karar ağaçlarının hedefi bağımlı değişkendeki farklılıkları maksimize edip, veriyi sıralı bir biçimde gruplara ayırmaktır.

Karar ağaçlarının kullanım amaçları,

- Belirli bir sınıfın muhtemel üyesi olacak elemanların belirlenmesi,
- Çeşitli olayların yüksek, orta, düşük risk grupları gibi yorumlanabilecek kategorilere ayrılması,
- Gelecekteki olayların öngörüsü için kurallar oluşturulması,
- Parametrik modellerin kurulmasında kullanılmak üzere çok sayıda değişken ve veri kümesinden gerekli olacakların seçilmesi

olarak açıklanabilir.

3.1.2. Karar Ağacının Yapısı

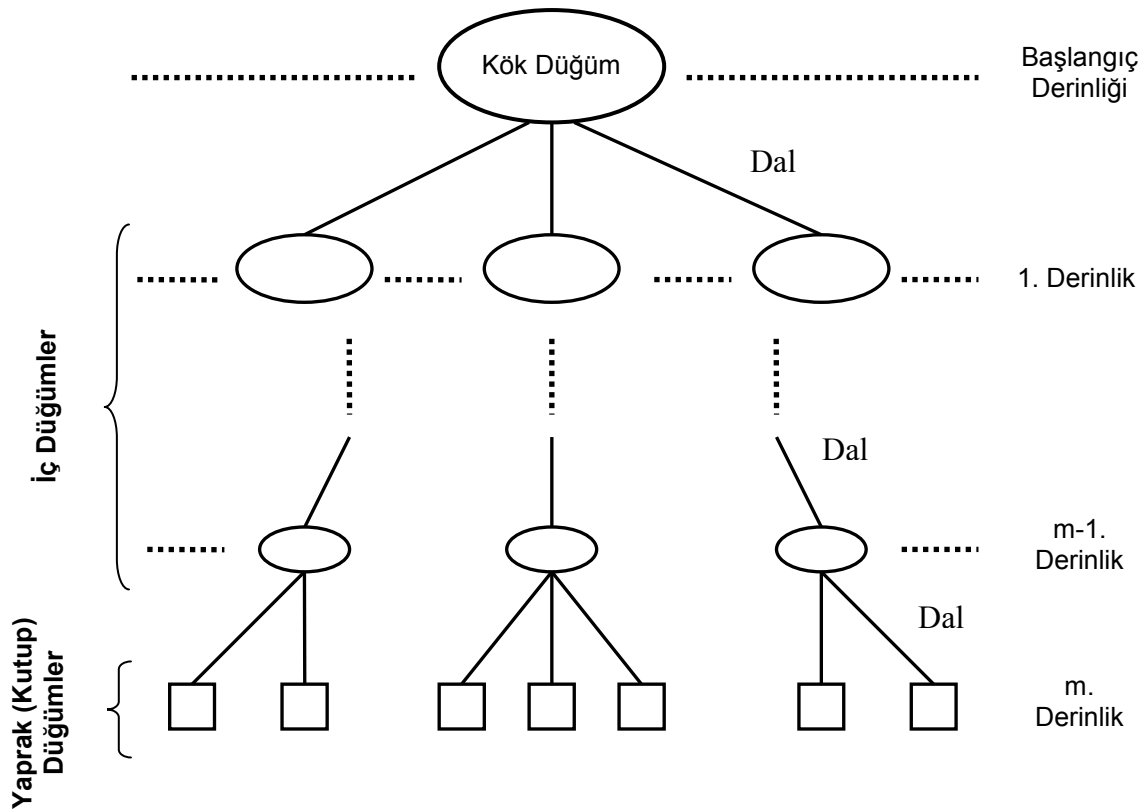
Karar ağaçlarının, karar düğümleri, dallar ve yapraklar olmak üzere üç tür bileşeni vardır. Karar düğümleri de üçe ayrılmaktadır.

Kök düğüm: Kendisinden önce bir dal olmayan ve kendisinden bir veya daha fazla dal çıkabilen düğümdür. Kök düğüm sınıflandırmanın hangi değişkene göre yapıldığını gösterir. Kök düğüm bağımlı değişkeni gösterir.

İç düğümler: Kendisinden önce olup kendisine doğru gelen sadece bir dal olan ve kendisinden en az iki veya daha fazla dal çıkan düğümlerdir.

Yaprak veya kutup (terminal) düğümler: Kendisinden önce olup kendisine doğru gelen sadece bir dal olan ve kendisinden hiç dal çıkmayan düğümlerdir.

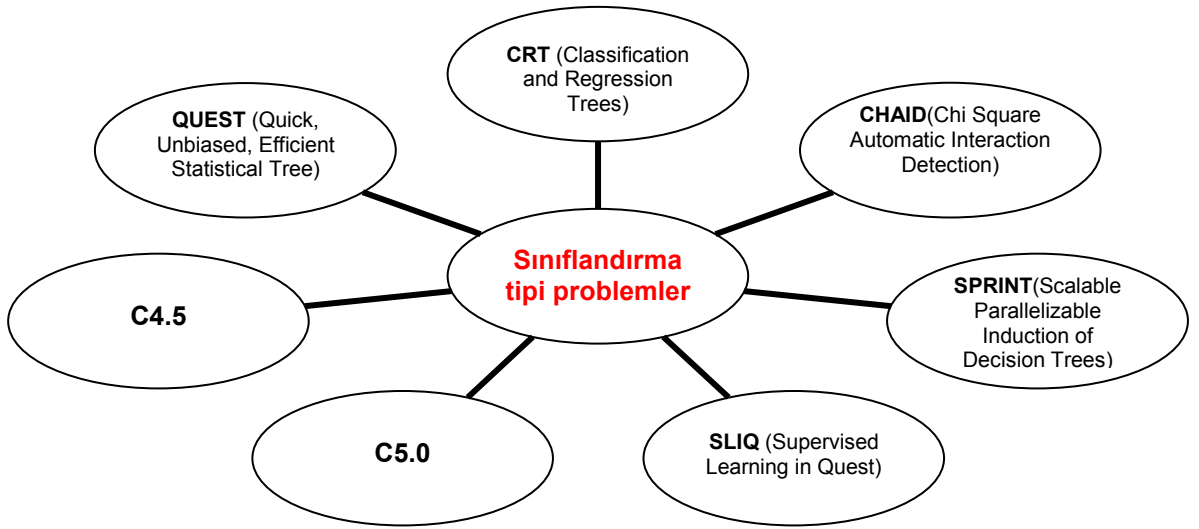
Düğümler arasındaki testin sonucunu gösteren ve tanımlanacak sınıfın belirlenmesini sağlayan yapı dal olarak adlandırılır. Dalın sonucunda sınıflandırma tamamlanamıyorsa tekrar bir karar düğümü oluşur. Karar ağacında her bir dal sonucunda oluşan düğümlerin bulunduğu yer derinliktir. Derinlik sayısını araştırmacı, karar ağaçlarının veri kümesine uygunluğunun analizini yaparak belirleyebilir. Bir karar ağacında derinlik ile oluşan sınıf sayısı doğru orantılıdır. Ancak dikkat edilmesi gereken nokta yaprak düğümlerde bulunan sınıflardaki veri sayısıdır. Eğer yaprak düğümlerinde yeteri kadar veri yoksa yapılan sınıflandırma veri kümesine uygun olmayabilir (Carvalho & Freitas, 2004; Han & Kamber, 2006).



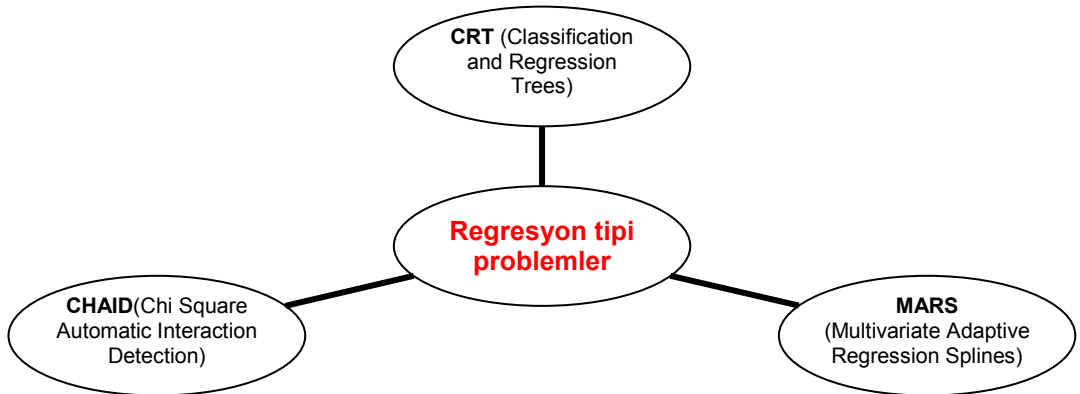
Şekil 3.1. Karar ağacının genel yapısı (Safavian & Landgrebe, 1991).

Yaygın olarak kullanılan karar ağaçları, CHAID (Chi-Squared Automatic Interaction Detector), CRT (Classification and Regression Trees), QUEST (Quick, Unbiased, Efficient Statistical Tree), C4.5, C5.0, MARS (Multivariate Adaptive Regression Splines), SLIQ (Supervised Learning in Quest), SPRINT (Scalable Parallelizable Induction of Decision Trees) vb. olarak verilebilir. Bağımlı değişkenin yapısına göre uygun karar ağaçları ile sınıflandırma yapılır. Şekil 3.2 'de bağımlı değişken yapısına göre uygun karar ağacı örnekleri verilmektedir.

Bağımlı Değişken Kategorik



Bağımlı Değişken Sürekli



Şekil 3.2. Bağımlı değişken türüne göre karar ağaçları

Çalışmada bağımlı değişken olan ödenmiş hasar tutarı sürekli olduğundan, CHAID ve CRT algoritmaları incelenip veri kümesine uygun olan karar ağacı belirlenerek risk sınıflandırması yapılmıştır. CHAID ve CRT algoritmalarının analizleri SPSS CLEMENTINE 11.1 programında yapılmıştır.

3.1.3. CHAID Algoritması

1980 yılında G.V. Kass tarafından geliştirilen CHAID algoritmasında, bağımlı değişkeni en fazla etkileyen bağımsız değişken, bağımlı değişkenin sürekli olması durumunda F testi, kategorik olması durumunda Ki Kare testi kullanılarak belirlenir. Kategorik ve sürekli değişkenler üzerinde çalışabilmesinden dolayı günümüzde de tercih edilen bir algoritmadır.

CHAID algoritması, her bir düğümü iki veya ikiden fazla sınıfa ayırmayı temel alan sınıflandırma yöntemidir. Bu süreç sınıflandırmanın önemli olmadığı düğümlere kadar devam etmektedir.

CHAID algoritmasında kategorik, sürekli ya da sıralı veri kümesi kullanılabilir.

CHAID algoritmasının adımları,

1. Kategorileri bağımlı değişkenle olan ilişki benzerliklerine göre birleştirmek,
2. Tüm düğümleri en iyi değişkene göre sınıflandırmak,
3. Eğer durdurma kriteri sağlanmıyorsa ya da hala sınıflandırılacak alt gruplar varsa birleştirme adımına geri dönmek

biçimindedir (<http://www.kdnuggets.com/software/classification-tree-rules.html>). CHAID algoritmasının her düğümünde kategorisi, düğümün örneklem içindeki yüzdesi ve düğümde bulunan veri sayısı bilgileri bulunmaktadır.

CHAID algoritması bağımlı değişken ile tahmin edici değişkenler arasındaki ilişkiyi görsel olarak kolaylıkla vermektedir ve analizde kullanılan değişkenlerinin hepsinin aynı tür veri olması koşulu yoktur. Bağımlı değişken ile tahmin edici değişkenler sıralı ya da kesikli olabilir. Ayrıca bağımlı değişken için normallik koşulu yoktur (Sugumaran et al., 2007; Zhou & Chen, 2002).

3.1.4. CRT Algoritması

CRT algoritması, her bir düğümü iki sınıfa ayırmayı temel alan sınıflandırma yöntemidir. CRT algoritmasında bağımlı değişken kesikli olması durumunda Gini ya da Twoing, bağımlı değişken sürekli olması durumunda en küçük kareler sapması (least square deviation) kriterlerine göre sınıflandırma yapılır. Bu kriterleri maksimum yapan sınıf sayısı en iyi sınıf sayısı olarak seçilir.

$$P_L = \frac{t_L \text{ 'deki veri sayısı}}{\text{toplam veri sayısı}}$$

$$P_R = \frac{t_R \text{ 'deki veri sayısı}}{\text{toplam veri sayısı}}$$

$$P(j|t_L) = \frac{t_L \text{ 'deki } j \text{ sınıftaki veri sayısı}}{\text{toplam veri sayısı}}$$

$$P(j|t_R) = \frac{t_R \text{ 'deki } j \text{ sınıftaki veri sayısı}}{\text{toplam veri sayısı}}$$

ve t_L , t. düğümün solundaki ara düğüm ve t_R , t. düğümün sağındaki ara düğüm olmak üzere CRT algoritmasında bir düğüm aşağıdaki kriterlere göre belirlenir.

Gini Kriteri

$\Phi(s|t)$, t. düğümden s. düğümün oluşmasının uygunluğunun ölçütüdür. T. düğüm için gini katsayısı;

$$i(t) = \sum_{i,j=1}^k C(i|j)P(i|t)P(j|t) \quad (3.1)$$

biçimindedir. Burada $C(i|j)$, j. sınıfın i. sınıfa bölünmesi durumundaki kaybı (the cost of miss-classifying a class j case as a class i case) gösterir.

$$\Phi(s|t) = i(t) - P_L i(t_L) - P_R i(t_R) \quad (3.2)$$

eşitliğini maksimum yapan değere göre t. düğüm s sınıfa ayrılır.

Twoing Kriteri

$\Phi(s|t)$, t. düğümde s. düğümün oluşmasının uygunluğunun ölçütüdür.

$$\Phi(s|t) = P_L P_R \left[\sum_{j=1}^k |P(j|t_L) - P(j|t_R)| \right]^2 \quad (3.3)$$

eşitliğini maksimum yapan değere göre t. düğüm s sınıfa ayrılır (Larose, 2005, <http://www.kdnuggets.com/software/classification-tree-rules.html>).

En Küçük Kareler Sapması

N, toplam durum sayısı, w_n , n. durum için ağırlık, f_n , n. durum için frekans olmak üzere, bağımlı değişken (Y) sürekli olduğu durumda t. düğümde s. düğümün oluşmasının uygunluğunun ölçütü olan $\Phi(s|t)$ değeri ;

$$\Phi(s|t) = i(t) - P_L i(t_L) - P_R i(t_R) \quad (3.4)$$

biçiminde elde edilir. Burada;

$$i(t) = \frac{\sum_{n=1}^N w_n f_n [y_n - \bar{y}(t)]^2}{\sum_{n=1}^N w_n f_n} \quad \text{ve} \quad \bar{y}(t) = \frac{\sum_{n=1}^N w_n f_n Y_n}{\sum_{n=1}^N w_n f_n} \quad \text{biçimindedir.} \quad (3.5)$$

CRT algoritmasında, CHAID algoritmasında olduğu gibi her düğümde kategorisi, düğümün örneklem içindeki yüzdesi ve düğümde bulunan veri sayısı bilgileri bulunmaktadır.

CRT algoritması tahmin edici değişkenlerin dağılımına ilişkin herhangi bir varsayım yoktur. Yüksek oranda çarpık dağılıma sahip verilerin sınıflandırmasında kullanılabilir (Larose, 2005; Safavian & Landgrebe, 1991).

DÖRDÜNCÜ BÖLÜM

4. HASAR REZERVİ

Sigorta veya reasürans şirketlerinin, henüz ödenmemiş hasarlar ile meydana gelmiş, ancak henüz sigorta veya reasürans şirketinin bildirilmemiş hasarlar için ayırmış oldukları karşılıklara hasar karşılığı ya da rezerv denilmektedir.

Sigorta şirketleri, hayat dışı sigorta branşlarında henüz ödenmemiş yükümlülükleri için karşılık ayırmak zorundadırlar. Bu karşılıklar hasar karşılıkları, prim karşılıkları ve diğer karşılıklardır.

Çalışmada hasar karşılıkları ele alınmıştır. Hayat dışı sigorta branşında hasar karşılıkları aşağıdaki biçimdedir:

IBNR (Incurred but not reported): Oluşmuş ancak ihbar edilmemiş hasarlar.

RBNS (Reported but not settled): Dosya muallak hasarları bir diğer ifade ile ihbar edilmiş ancak kapatılmamış hasarlar.

IBNER (Incurred but not enough reserved): Oluşmuş ancak yeteri kadar karşılık ayrılmamış hasarlar.

RBNER (Reported (and Reserved) but not enough reserved): Raporlanmış veya karşılık ayrılmış ancak karşılıkların yetersiz kaldığı hasarlar.

IBNFR (Incurred but not fully reported): Oluşmuş ancak tam olarak ihbar edilmemiş hasarlar (Kaas et al., 2001).

Hasar karşılıkları hesaplanırken hasar gelişim üçgeninden (run-off triangle) yararlanır. Hasar gelişim üçgeni, gelecek hasar sayılarını ve hasar tutarlarını tahmininde kullanılmaktadır.

4.1. Hasar Gelişim Üçgeni

Hasar gelişim üçgeni, satırları kaza yıllarını, sütunları ise hasar gelişim yıllarını gösteren, hücrelerinde ise hasara ilişkin bilginin bulunduğu üçgendir.

Hasar gelişim üçgenindeki,

Kaza Yılı (Accident Year): Hasarın gerçekleştiği ve sigortacının risk altında olduğu yılları göstermekte,

Hasar Gelişim Yılı (Development Year: The number of years until a payment is made): Hasar ödemelerinin hasar gerçekleştikten kaç yıl sonra yapıldığını göstermektedir.

Hasar gelişim üçgeninin verisi kullanılan yöntemle göre değişiklik göstermektedir. Genel olarak hasar gelişim üçgeninde, aşamalı hasar tutarları (incremental losses), birikimli hasar tutarları (cumulative losses), ödenmiş hasar tutarlarının oluşan hasar tutarlarına oranı ve hasar sayıları kullanılmaktadır. Oluşan ya da ödenmiş hasar tutarları üzerinden hesaplamalar yapılmaktadır (Schmidt, 2006; Friedland, 2009).

$Z_{i,j}$, i. kaza, j. hasar gelişim yılında oluşan aşamalı hasar tutarını göstermek üzere, aşamalı hasar tutarına göre hasar gelişim üçgeni Çizelge 4.1'de verilmiştir.

Çizelge 4.1. Aşamalı hasar tutarına göre hasar gelişim üçgeni

Kaza Yılı	Hasar Gelişim Yılı								
	0	1	...	j	...	n-i	...	n-1	n
0	$Z_{0,0}$	$Z_{0,1}$...	$Z_{0,j}$...	$Z_{0,n-i}$...	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$...	$Z_{1,j}$...	$Z_{1,n-i}$...	$Z_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$Z_{i,0}$	$Z_{i,1}$...	$Z_{i,j}$...	$Z_{i,n-i}$			
⋮	⋮	⋮		⋮					
n-j	$Z_{n-j,0}$	$Z_{n-j,1}$...	$Z_{n-j,j}$					
⋮	⋮	⋮							
n-1	$Z_{n-1,0}$	$Z_{n-1,1}$							
n	$Z_{n,0}$								

Çizelge 4.1 'de n kaza ve n gelişim yılı için gözlemlenmiş aşamalı hasar tutarı (incremental losses) verilmektedir. Bu değerler kullanılarak, gözlemlenmemiş aşamalı hasar tutarını (non-observed incremental losses) tahmin edilir.

$S_{i,j}$, i. kaza, j. hasar gelişim yılında oluşan birikimli hasar tutarını göstermek üzere, birikimli hasar tutarına göre hasar gelişim üçgeni Çizelge 4.2 'de verilmiştir.

Çizelge 4.2. Birikimli hasar tutarlarına göre hasar gelişim üçgeni

Kaza Yılı	Hasar Gelişim Yılı								
	0	1	...	j	...	n-i	...	n-1	n
0	$S_{0,0}$	$S_{0,1}$...	$S_{0,j}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$...	$S_{1,j}$...	$S_{1,n-i}$...	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$S_{i,0}$	$S_{i,1}$...	$S_{i,j}$...	$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
n-j	$S_{n-j,0}$	$S_{n-j,1}$...	$S_{n-j,j}$					
⋮	⋮	⋮							
n-1	$S_{n-1,0}$	$S_{n-1,1}$							
n	$S_{n,0}$								

Çizelge 4.2 'de n kaza ve n gelişim yılı için birikimli hasar tutarı (cumulative losses) verilmektedir. Bu değerler kullanılarak, gözlemlenmemiş birikimli hasar tutarı (non-observed cumulative losses) tahmin edilir.

Çizelge 4.3 'te ise birikimli hasar tutarına göre tamamlanmış hasar gelişim üçgeni verilmektedir.

Çizelge 4.3. Birikimli hasar tutarlarına göre tamamlanmış hasar gelişim üçgeni

Kaza Yılı	Hasar Gelişim Yılı								
	0	1	...	j	...	n-i	...	n-1	n
0	$S_{0,0}$	$S_{0,1}$...	$S_{0,j}$...	$S_{0,n-i}$...	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$...	$S_{1,j}$...	$S_{1,n-i}$...	$S_{1,n-1}$	$S_{1,n}$
⋮	⋮	⋮		⋮		⋮		⋮	⋮
i	$S_{i,0}$	$S_{i,1}$...	$S_{i,j}$...	$S_{i,n-i}$...	$S_{i,n-1}$	$S_{i,n}$
⋮	⋮	⋮		⋮		⋮		⋮	⋮
n-j	$S_{n-j,0}$	$S_{n-j,1}$...	$S_{n-j,j}$...	$S_{n-j,n-i}$...	$S_{n-j,n-1}$	$S_{n-j,n}$
⋮	⋮	⋮		⋮		⋮		⋮	⋮
n-1	$S_{n-1,0}$	$S_{n-1,1}$...	$S_{n-1,j}$...	$S_{n-1,n-i}$...	$S_{n-1,n-1}$	$S_{n-1,n}$
n	$S_{n,0}$	$S_{n,1}$...	$S_{n,j}$...	$S_{n,n-i}$...	$S_{n,n-1}$	$S_{n,n}$

$i + j \leq n$ ise $S_{i,j}$ gözlemlenmiş birikimli hasar tutarını,

$i + j > n$ ise gözlenlenmemiş ya da gelecekte gerçekleşecek birikimli hasar tutarını,

$i + j = n$ ise mevcut (current) birikimli hasar tutarını,

$j = n$ ise son birikimli hasar tutarını göstermektedir.

Aşamalı hasar tutarları kullanılarak birikimli hasar tutarları

$$S_{i,j} = \sum_{l=0}^j Z_{i,l} \quad (4.1)$$

biçiminde elde edilir (Schmidt, 2006).

Birikimli hasar tutarları kullanılarak aşamalı hasar tutarları,

$$Z_{i,j} = \begin{cases} S_{i,j} & , \quad j = 0 \\ S_{i,j} - S_{i,j-1} & , \quad \text{diğer durumlar} \end{cases} \quad (4.2)$$

biçiminde elde edilir.

Hasar gelişim üçgeninin kullanılmasının temel amaçları:

- Gelecekte beklenen birikimli hasar tutarları olan $S_{i,j}$ değerlerini,
- Gelecekte beklenen aşamalı hasar tutarları olan $Z_{i,j} = S_{i,j} - S_{i,j-1}$ değerini,
- Toplam rezerv miktarını olan $\sum_{l=1}^n \sum_{k=n-l+1}^n Z_{l,k}$ değerini,
- Takvim yılı rezervi olan $\sum_{l=p-n}^n Z_{l,p-l}$ değerini,

tahmin etmektir. Burada ;

$$i + j \geq n + 1 \text{ ve } p = n + 1, n + 2, \dots, 2n$$

biçimindedir (Schmidt, 2006).

Hasar gelişim üçgeninde, hasar tutarlarının bulunduğu hücreler arasındaki ilişkiyi tespit edebilmek için bağlantı oranları (link ratio) hesaplanmalıdır. Bağlantı oranları arasında dikey, yatay, diagonal ya da bunların kombinasyonu biçiminde bir model yakalanması gerekmektedir. Hücrelerin belli bir modeli göstermediği hasar gelişim üçgeninde genellikle zincir merdiven yöntemi kullanılır.

4.2. Hasar Rezervi Tahmininde Kullanılan Yöntemler

Gelecekteki hasar tutarlarını tahmin etmede iki türlü yaklaşım vardır. Bunlar deterministik ve stokastik yaklaşımlardır. Deterministik ile stokastik yaklaşımlar arasındaki fark, stokastik modellerin rassal parçalarının olmasıdır. Stokastik modeller ile tahmin yapılırken ortalama ve varyans değeri elde edilir ve hasar rezervi için güven aralığı verilebilir. Ayrıca stokastik modellerde kullanılan yöntemlere bağlı olarak hasar tutarlarına ilişkin dağılım varsayımı da yapılır. Deterministik modellerde ise hasar rezervlerinin tahmini doğrudan gelişim faktörleri kullanılarak hesaplanır. Çalışmada hasar rezervi Mack Zincir Merdiven Yöntemi ile tahmin edilecektir. Mack Zincir Merdiven Yöntemi'nde hasar rezervine ilişkin standart hata değeri de hesaplandığından bu yöntem stokastik bir yöntemdir (Friedland, 2009; Struzziere & Hussian, 1998).

Hasar rezervini tahmin etmekte yaygın olarak kullanılan yöntemlerden bazıları aşağıda verilmiştir (Schmidt & Zoher, 2008):

- Zincir Merdiven Yöntemi (Chain Ladder Method)
- Bornhuetter-Ferguson Yöntemi (Bornhuetter-Ferguson Method)
- Cape-Cod Yöntemi (Cape-Cod Method)
- Hasar Gelişim Yöntemi (Loss-Development Method)
- Toplamsal Yöntem (The Additive Method)

Bu yöntemler hasar gelişim üçgeni kullanılarak uygulanmaktadır.

4.2.1. Zincir Merdiven Yöntemi

Zincir Merdiven Yöntemi, hasar gelişim üçgenine dayalı ödenmiş hasar tutarlarının projeksiyonunu yapan bir yöntemdir. Zincir Merdiven Yöntemi, ödenmiş hasar tutarlarını karşılaştırarak, gelecekte ödenmesi gereken hasar tutarlarını tahmin etmekte kullanılır. Bu yöntemle, bir önceki yılın birikimli hasar tutarına göre mevcut yılın birikimli hasar tutarının tahmin edilebileceği ortalama gelişim faktörü (development factor) hesaplanmaktadır. Gelişim faktörleri, bağlantı oranlarının ortalaması olarak hesaplanabileceği gibi, tüm kolonlardaki toplam hasarı bulup, bir öncekine bölerek de hesaplanabilir (Dahl, 2003).

Zincir Merdiven Yöntemi'nin amacı:

- Rezerv miktarının yeterli düzeyde ayrılıp ayrılmadığını kontrol etmek,
- Rezerv miktarının gelecekte muhtemel ödemeleri karşılamada yeterli olup olmadığını ölçmek,
- Şirketin poliçe sahiplerine karşı olan yükümlülüklerini yeterli düzeyde karşılayıp karşılamadığının kontrol etmektir.

Zincir Merdiven Yöntemi kolay uygulanabilir ve anlaşılabilir olduğundan dolayı tercih edilen yöntemlerden birisidir. Ancak sadece gelişim faktörlerine bağlı bir yöntem olması gibi bazı problemleri bulunmaktadır.

Yaygın olarak kullanılan başlıca Zincir Merdiven Yöntemleri Mack Zincir Merdiven Yöntemi ve Münih Zincir Merdiven Yöntemidir (Sahasrabuddhe, 2008).

4.2.1.1. Mack Zincir Merdiven Yöntemi

Mack Zincir Merdiven Yöntemi 'nde hasar rezervi birikimli hasarlar kullanılarak hesaplanır. Bu yöntemde j. kolondan j+1. kolona büyüme faktörü (f_j);

$$\hat{f}_j = \frac{\sum_{i=1}^{n-j} S_{i,j+1}}{\sum_{i=1}^{n-j} S_{i,j}} = \sum_{i=1}^{n-j} \frac{S_{i,j}}{\sum_{i=1}^{n-j} S_{i,j}} \times \frac{S_{i,j+1}}{S_{i,j}} \quad (4.3)$$

biçiminde hesaplanır.

Eşitlikte \hat{f}_j büyüme faktörü, gözlemlenmiş gelişim faktörlerinin $\left(\frac{S_{i,j+1}}{S_{i,j}}\right)$ ağırlıklandırılmış ortalamasıdır.

Birikimli hasar değerlerinin tahminlerinin toplamı, bir başka deyişle rezerv tutarı,

$$\hat{S}^{MCL} = \sum_{i=2}^n S_{i,n-i+1} \times (\hat{f}_{n-i+1} \times \dots \times \hat{f}_{n-1} - 1) \quad (4.4)$$

biçimindedir.

Mack Zincir Merdiven yönteminin varsayımları aşağıda verilmektedir:

$$E \left[\frac{S_{i,j+1}}{S_{i,j}} \middle| S_{i,1}, S_{i,2}, \dots, S_{i,j} \right] = f_j \quad (4.5)$$

$$\text{Var} \left(\frac{S_{i,j+1}}{S_{i,j}} \middle| S_{i,1}, S_{i,2}, \dots, S_{i,j} \right) = \frac{\sigma_j^2}{S_{i,j}}$$

$\{S_{i,1}, \dots, S_{i,n}\}$ ve $\{S_{k,1}, \dots, S_{k,n}\}$ $i \neq k$ durumunda bağımsızdır (Kubrusly et al., 2009).

4.2.1.2. Münih Zincir Merdiven Yöntemi

Münih Zincir Merdiven Yönteminde birikimli ödenmiş hasar tutarları ve oluşan hasar tutarları kullanılarak rezerv miktarı tahmin edilmektedir. Bu yöntemde, her hücrede ödenmiş hasar tutarının oluşan hasar tutarlarına oranı kullanılarak hesaplamalar yapılır. Hücrelerde kullanılan ödenmiş hasar tutarının oluşan hasar tutarına oranı;

$$(P / I)_{i,j} = \frac{P_{i,j}}{I_{i,j}} \quad (4.6)$$

biçimindedir. Her kaza yılı için tahmin edilen bu oran gözlemlenmiş oranların ortalama oranlarına bağlıdır.

Bu yöntemin Mack Zincir Merdiven Yöntemi 'nden en önemli farkı ödenmiş ve oluşan hasarlar arasındaki ilişkiyi kullanarak hasar rezervinin tahmin edilmesidir (Quarg & Mack, 2008).

4.2.2. Bornhuetter-Ferguson Yöntemi

Bornhuetter-Ferguson Yöntemi 'nin varsayımı α_i ve υ_j parametrelerine bağlıdır.

$\hat{Z}_{i,j}$: i. kaza yılı ve j. gelişim yılı için tahmin edilen aşamalı hasar tutarı,

$\hat{S}_{i,j}$: i. kaza yılı ve j. gelişim yılı için tahmin edilen birikimli hasar tutarı,

α_i : i. kaza yılı için beklenen birikimli son hasar tutarının beklenen değeri,

$\hat{\alpha}_i$: i. kaza yılı için beklenen birikimli son hasar tutarının beklenen değerinin önsel kestirimi,

υ_j : j. gelişim yılı için aşamalı kotaların gelişim örüntüsü (development pattern for incremental quotas),

$\hat{\upsilon}_j$: j. gelişim yılı için aşamalı kotaların gelişim örüntüsünün önsel kestirimi,

γ_j : j. gelişim yılı için birikimli kotaların gelişim örüntüsü (development pattern of cumulative quotas),

$\hat{\gamma}_j$: j. gelişim yılı için birikimli kotaların gelişim örüntüsünün önsel kestirimi (development pattern of cumulative quotas),

göstermektedir. Tüm $i, j \in \{0, 1, \dots, n\}$ olmak üzere;

$$\alpha_i = E(S_{i,n}) \quad (4.7)$$

$$v_j = \frac{E(Z_{i,j})}{E(S_{i,n})}, \quad \sum_{l=0}^n v_l = 1, \quad v_j = \begin{cases} \gamma_0 & , k = 0 \\ \gamma_j - \gamma_{j-1} & , \text{diğer durumlarda} \end{cases} \quad (4.8)$$

$$\gamma_j = \frac{E(S_{i,j})}{E(S_{i,n})}, \quad \gamma_n = 1 \quad (4.9)$$

biçimindedir.

Aşamalı hasar tutarlarına göre Bornhuetter-Ferguson Yöntemi 'nde gözlemlenmemiş hasar tutarı olan $Z_{i,j}$ ($i + j \geq n + 1$) değerinin tahmini;

$$\hat{Z}_{i,j} \cong \hat{\alpha}_i \hat{v}_j \quad (4.10)$$

biçiminde hesaplanır. Benzer biçimde birikimli hasar tutarlarına göre Bornhuetter-Ferguson Yöntemi 'nde gözlemlenmemiş hasar tutarı olan $S_{i,j}$ ($i + j \geq n + 1$) değerinin tahmini;

$$\hat{S}_{i,j} \cong \hat{\alpha}_i \hat{\gamma}_j \quad (4.11)$$

biçiminde hesaplanır. Buradan ;

$$\hat{S}_{i,n}^{BF} = S_{i,n-i} + \hat{\alpha}_i (\hat{\gamma}_j - \hat{\gamma}_{n-i}) \quad (4.12)$$

biçiminde elde edilir (Dahl,, 2003).

4.2.3. Cape-Cod Yöntemi

Cape-Cod Yöntemi birikimli hasarların gelişim modelinin önsel kestirimlerine (γ_j) ve π_i ($\pi_i \in (0, \infty)$, $i = 1, \dots, n$) ile gösterilen prim tutarlarına bağlıdır. Her kaza yılı için prim tutarına (π_i) bağlı olarak son birikimli hasar oranı;

$$\kappa_i = E\left(\frac{S_{i,n}}{\pi_i}\right) \quad (4.13)$$

biçiminde ve tahmini birikimli hasar tutarı için hasar oranı tahmini;

$$\hat{\kappa}_i^{CC} = \frac{\sum_{l=0}^n S_{i,n-l}}{\sum_{l=0}^n \hat{\gamma}_{n-l} \pi_l} \quad (4.14)$$

eşitliği ile verilir.

Cape-Cod Yöntemi 'nde gözlemlenmemiş hasar tutarı olan $S_{i,j}$ ($i+j \geq n+1$) değerinin tahmini;

$$\hat{S}_{i,j}^{CC} = S_{i,n-i} + (\hat{\gamma}_j - \hat{\gamma}_{n-i}) \pi_i \hat{\kappa}^{CC} \quad (4.15)$$

eşitliği ile elde edilir.

Bornhuetter-Ferguson Yöntemi 'nde beklenen hasar tutarları üzerinden rezerv miktarı tahmin edilir. Cape-Cod Yöntemi 'nde ise son hasar tutarları kullanılarak rezerv miktarı tahmin edilir.(Barnett, 2007; Schmidt & Zocher, 2008).

4.2.4. Hasar Gelişim Yöntemi

Hasar Gelişim Yöntemi, Cape-Cod Yöntemi 'nde olduğu gibi birikimli hasarların gelişim modelinin önsel kestirimlerine(γ_j) bağlıdır. Hasar Gelişim Yöntemi 'nde gözlemlenmemiş hasar tutarı olan $S_{i,j}$ ($i+j \geq n+1$) değerinin tahmini;

$$\hat{S}_{i,j}^{LD} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} \quad (4.16)$$

biçiminde elde edilir (Schmidt & Zocher, 2008).

4.2.5. Toplamsal Yöntem

Bu yöntem aşamalı hasar oranı (incremental loss ratio) olarak da adlandırılmaktadır. Bu yöntem π_i prim tutarlarına ek olarak, ζ_j ile ifade edilen ve hasar gelişim yılında beklenen hasarın o yılın primlerine oranı olarak tanımlanan bir parametreye bağlıdır.

π_i , prim tutarlarını ya da kaza yıllarına ilişkin herhangi bir ölçü olmak üzere her hasar gelişim yılı için aşamalı hasar oranlarının beklenen değeri ;

$$\zeta_j = E\left(\frac{Z_{i,j}}{\pi_i}\right) \quad (4.17)$$

ve

$$\hat{\zeta}_j^{AD} = \frac{\sum_{l=0}^{n-j} Z_{l,j}}{\sum_{l=0}^{n-j} \pi_l} \quad (4.18)$$

eşitliği ile verilir.

Toplamsal Yöntem 'de gözlemlenmemiş hasar tutarı olan $Z_{i,j}$ ve $S_{i,j}$ ($i+j \geq n+1$) değerlerinin tahmini;

$$\hat{Z}_{i,j}^{AD} = \hat{\zeta}_j^{AD} \pi_i \quad (4.19)$$

$$\hat{S}_{i,j}^{AD} = S_{i,n-i} + \sum_{l=n-i+1}^j \hat{Z}_{i,l}^{AD} \quad (4.20)$$

eşitlikleriyle tanımlanır.

Toplamsal Yöntem de Bornhuetter-Ferguson Yöntemi 'nde önsel kestirimlerin kullanıldığı özel bir durumdur (Schmidt & Zocher, 2008).

4.3. Hasar Rezervi Tahmininde Kullanılan Yöntemlerin Karşılaştırılması

Hasar rezervi tahmininde yaygın olarak kullanılan ve yukarıda açıklanan yöntemler kullanılan parametrelere ve veriye bağlı olarak farklılıklar göstermektedir. Bu farklılıklar Çizelge 4.4 'te özetlenmiştir.

Çizelge 4.4. Hasar rezervi tahmininde kullanılan yöntemlerinin karşılaştırılması

Yöntemler	Kullanılan Veri	Veri Türü	Önsel Kestirim için Gerekli Veri	Yöntem Türü
Mack Zincir Merdiven	Oluşan, toplam ya da ödenmiş hasar tutarı	Kümülatif	-	Stokastik
Münih Zincir Merdiven	Ödenmiş hasar tutarının oluşan hasar tutarına oranı	Kümülatif	-	Stokastik
Bornhuetter-Ferguson	Oluşan, toplam ya da ödenmiş hasar tutarı	Kümülatif / aşamalı	Beklenen hasar tutarı / net prim tutarı	Deterministik
Cape-Cod ¹	Bildirilmiş ya da ödenmiş hasar tutarı	Kümülatif	Prim tutarı	Deterministik
Hasar Gelişim	Oluşan, toplam ya da ödenmiş hasar tutarı	Kümülatif	-	Deterministik
Toplamsal	Oluşan, toplam ya da ödenmiş hasar tutarı	Kümülatif / aşamalı	Beklenen hasar prim oranı	Deterministik

(Kaynak: Casualty Loss Reserve Seminar, 1992; Schmidt & Zocher, 2008; Friedland, 2009).

Çizelge 4.4 'te yöntemler arasındaki farklılıklar kullanılan veriye bağlı olarak verilmiştir. Yöntemler arasındaki en önemli fark yöntemlerin stokastik ya da deterministik olmasıdır. Çünkü stokastik yöntemlerde tahmin edilen hasar rezerv tutarı için hata değeri elde edilirken, deterministik yöntemlerde bu değer elde edilemez.

Mack Zincir Merdiven Yöntemi, kümülatif hasar tutarları göz önünde bulundurularak rezerv tutarının tahmin edildiği yaygın kullanılan bir yöntemdir. Oluşan ve ödenmiş hasar tutarı bilgisi bulunduğu Münih Zincir Merdiven

¹ Barnett & Zehnwirth (2000) Cape Cod Yöntemi'ni regresyon modeli kullanarak, stokastik olarak uygulamıştır.

Yöntemi ile daha uygun sonuçlar elde edilebilir. Çünkü Mack Zincir Merdiven Yöntemi 'nde tek bir değişken göz önünde bulundurulurken, Münih Zincir Merdiven Yöntemi 'nde hem oluşan hem de ödenmiş hasar tutar bilgisi göz önünde bulundurularak hasar rezerv tutarı tahmin edilir.

Bornhuetter-Ferguson Yöntemi beklenen hasar tutarına bağlı bir yöntemdir. Beklenen hasar tutarına ilişkin güvenilir bilgi olmadığı durumlarda Bornhuetter-Ferguson Yöntemi rezerv tutarı tahminine ilişkin güvenilir sonuçlar vermeyebilir. Ayrıca hasar tutarları için aralarında çok az salınım varsa gelişim örüntü değerleri büyük olabilir ve buna bağlı olarak rezerv tutarı çok yüksek tahmin edilebilir (Friedland, 2009).

Cape Cod Yöntemi ise son hasar ve prim tutarına bağlı bir yöntemdir. Cape Cod Yöntemi 'nde genellikle bildirilmiş hasar tutarları kullanılır. Ayrıca ödenmiş hasar tutarları da bu yöntemde kullanılır (Friedland, 2009).

Bornhuetter-Ferguson Yöntemi ve Cape Cod Yöntemi genellikle reasürans uygulamalarında kullanılmaktadır (Casualty Loss Reserve Seminar, 1992; Friedland, 2009).

Hasar Gelişim Yöntemi 'nde beklenen hasar tutarına ilişkin herhangi bir önsel kestirim değerine ihtiyaç yoktur. Hasar Gelişim Yöntemi ile Bornhuetter-Ferguson Yöntemi ile tahmin edilen hasar rezerv tutarları yakın değerlerdir (Schmidt & Zocher, 2008).

Toplamsal Yöntem 'de genellikle aşamalı hasar tutarı kullanılarak rezerv miktarı tahmin edilebilir. Ancak birikimli hasar tutarı da hasar rezervini tahmin etmekte kullanılır. Toplamsal Yöntem ile Cape Cod Yöntemi ile tahmin edilen hasar rezerv tutarları yakın değerlerdir (Schmidt & Zocher, 2008).

Hasar rezervi tahmininde kullanılacak yöntemin seçilmesinde aktüerin önemli bir rolü vardır. Stokastik modellerde kullanılacak dağılımların parametreleri en çok olasılık yöntemi ile tahmin edilebilir. Ancak burada uygun dağılımın belirlenmesi önemlidir.

Güvenilir önsel kestirim bilgilerine göre verilen yöntemler arasında seçim yapılır.

Eğer portföy birden fazla portföyden (subportfolio) oluşuyor ise Zincir Merdiven Yöntemi ve Toplamsal Yöntem, bu çeşit portföylere ilişkin hasar rezervi tutarının tahmininde kullanılabilir (Schmidt, 2006).

BEŞİNCİ BÖLÜM

5. UYGULAMA

Çalışmanın bu bölümünde Trafik Sigortası (Karayolları Motorlu Araçlar Zorunlu Mali Sorumluluk Sigortası) için risk sınıflandırması yapılmış ve sonuçları karşılaştırılmıştır. Risk sınıflandırması CHAID ve CRT algoritmaları ile yapılmıştır. Ödenmiş hasar tutarı sürekli bağımlı değişken olduğundan ve ödenmiş hasar tutarına göre sınıflandırma yapıldığından bu karar ağaçları sınıflandırmada kullanılmıştır. Ayrıca CHAID algoritması bir düğümü ikiden daha çok dala ayırırken, CRT algoritması bir düğümü iki dala ayırmaktadır. Çalışmada karar ağaçları arasındaki bu farkı da ortaya koymak için CHAID ve CRT algoritmaları seçilmiştir. Risk grupları belirlenerek, gruplar arasındaki ödenmiş hasar tutarı farklılıklarına değinilmiştir. Ödenmiş hasar tutarına göre yıllar itibariyle risk grupları belirlendikten sonra hasar rezervi tahmin edilmiştir. Ödenmiş hasar tutarı bilgisi olduğundan, hasar rezervi tahmininde Mack Zincir Merdiven Yöntemi kullanılmıştır. Sınıflandırmadan önce ve sınıflandırmadan sonraki tahmin edilen hasar rezervleri karşılaştırılmıştır.

5.1. Trafik Sigortası (Karayolları Motorlu Araçlar Zorunlu Mali Sorumluluk Sigortası)

Trafik Sigortası (Karayolları Motorlu Araçlar Zorunlu Mali Sorumluluk Sigortası) ülkemizde prim üretimi açısından önemli bir büyüklüğe ve geniş bir uygulama alanına sahip olan bir sigorta türüdür.

Trafik Sigortası 2918 sayılı Karayolları Trafik Kanunu gereğince mecburi olup, Trafik Sigortası olmayan araçlar trafiğe çıkamazlar.

Trafik Sigortası prim ve teminatları, Türkiye Sigorta ve Reasürans Şirketler Birliği ve T.C. Başbakanlık Hazine Müsteşarlığı tarafından müştereken belirlenir.

Hazine Müsteşarlığı çalışmalarının en önemlilerinden birisi, sigorta şirketlerinin Trafik Sigortasına ilişkin tüm bilgilerini bir araya getiren merkezi bir bilişim alt yapısının oluşturulmasıdır. 16.12.2003 tarih ve 25318 sayılı Resmi Gazetede yayımlanan Trafik Sigortası Bilgi Merkezi Yönetmeliği ile kısaca TRAMER olarak

adlandırılan Trafik Sigortası Bilgi Merkezi kurulmasıyla Trafik Sigortası daha sağlıklı bir alt yapıya kavuşturulmuştur.

TRAMER'in kurulması ile temel olarak; Trafik Sigortasında denetim etkinliğinin sağlanması, sigorta kaçaklarının önlenmesi, sahtekarlığa karşı önlem alınması, toplumda sigorta sistemine olan güvenin artırılması, daha sağlıklı fiyatlandırma yapılması, haksız rekabetin önlenmesi gibi hususlar amaçlanmaktadır.

5.2. Veri Kümesinin Oluşturulması

Çalışmamızda ihtiyaç duyulan Trafik Sigortası poliçelerine ilişkin sürücü bilgileri (sürücünün cinsiyeti, sürücünün yaşı vb.), araç bilgileri (aracın model yılı, araç türü, yakıt türü vb.) ve hasar tutarına ilişkin veri elde edilemediğinden, TRAMER tarafından açıklanan istatistikler temel alınmış ve hasar yapan poliçe adetlerine bağlı olarak hasar verisi ile hasar üreten poliçelerin araç türü, il plaka kodu, yakıt türü gibi değişkenler türetilerek çalışmaya esas olan veri kümesi elde edilmiştir.

Veri kümesinin oluşturulmasında TRAMER 'de bulunan poliçe adedi ve toplam ödenmiş hasar tutarı bilgilerinden yararlanılmıştır. Çizelge 5.1 'de 2005-2008 yılları arasında TRAMER 'de bulunan poliçe adedi ve toplam ödenmiş hasar tutarı bilgileri verilmiştir. Ortalama hasar tutarına ilişkin bilgi bulunmadığından, tüm poliçelerin en az bir hasara sahip olduğu varsayımı ile ortalama hasar tutarı, toplam ödenmiş hasar tutarının poliçe adedine bölünmesiyle elde edilmiştir. Bu varsayım kullanılarak TRAMER 'de bulunan poliçe sayısının on binde biri oranında veri türetilmiştir.

TRAMER 'de illere ilişkin poliçe adedi ve toplam ödenmiş hasar tutarı bilgisi olduğundan hasar tutarının türetilmesinde il gruplarından yararlanılmıştır. Çalışmada il grupları İstanbul, Ankara, İzmir, Antalya, Bursa ve diğer şehirler olarak seçilmiştir ve hasar tutarları türetilirken illere göre hasar bilgileri dikkate alınmıştır. TRAMER 'de araç ve sürücüye ait bilgi olmadığından, araç türü, aracın model yılı, sürücünün yaşı ve cinsiyeti gibi değişkenler türetilirken TÜİK ve Emniyet Müdürlüğü verilerinden yararlanılmıştır.

Çizelge 5.1. Yıllara göre TRAMER’den elde edilen toplam poliçe adedi ve ödenmiş hasar tutarı

YILLAR	POLİÇE ADEDİ	TOPLAM ÖDENMİŞ HASAR TUTARI (TL)	ORTALAMA HASAR TUTARI (TL)
2005	8.998.046	823.629.000	91,53
2006	9.846.075	980.692.000	99,60
2007	10.430.316	1.158.017.000	111,02
2008	10.940.481	1.358.014.000	124,13
TOPLAM	40.214.918	4.320.352.000	107,43

Çizelge 5.1.’de verilen poliçe adedinin on binde biri olan toplam 40.214 tane poliçe verisi analizde kullanılmak üzere aşağıda açıklanan varsayımlar yardımıyla türetilmiştir.

Çalışmada her yıldaki poliçenin diğer yılda olmadığı (bir yıllık poliçe) ve her poliçede en az bir hasar olduğu ve ödenmiş hasar tutarının lognormal dağılıma sahip olduğu varsayımları yapılmıştır. Hasar tutarı değerleri genellikle sağa çarpık ve uzun kuyruklu dağılım yapısı göstermektedir. Hasar tutarı modellenmesinde genellikle lognormal, pareto, gamma gibi sağa çarpık uzun kuyruklu dağılımlar kullanılmaktadır. Çalışmada hasar tutarı modellenmesinde lognormal dağılım kullanılmıştır.

Lognormal dağılım, logaritması normal dağılım gösteren rastlantı değişkenlerin, tek kuyruklu ve sağa çarpık olasılık dağılımıdır. Sigortacılıkta hasar tutarının modellenmesinde yaygın olarak kullanılmaktadır.

Lognormal dağılımın parametreleri $\mu(-\infty < \mu < \infty)$ ve $\sigma(\sigma > 0)$ olmak üzere olasılık yoğunluk fonksiyonu ;

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right] \quad (5.1)$$

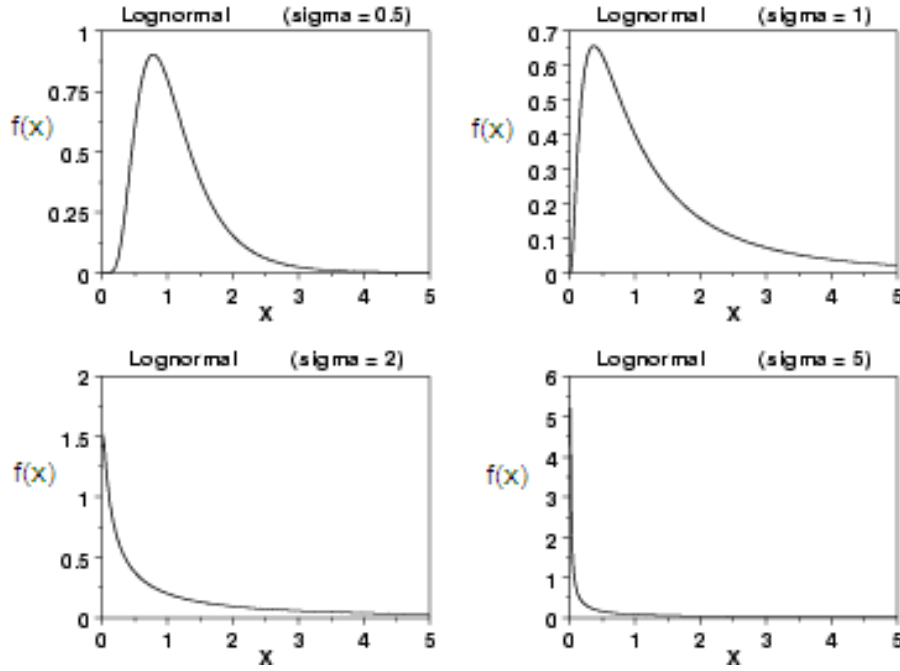
şeklindedir. Dağılımın beklenen değeri ve varyansı

$$E(x) = e^{\mu + \frac{\sigma^2}{2}} \quad (5.2)$$

$$V(x) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2} \quad (5.3)$$

şeklinde elde edilir.

Lognormal dağılımın farklı σ parametreleri için olasılık yoğunluk fonksiyonu Şekil 5.1'de verilmiştir.



Şekil 5.1. Lognormal dağılımın olasılık yoğunluk fonksiyonu grafikleri

Lognormal dağılımın parametrelerinden biri olan μ TRAMER 'de bulunan her il grubunun ortalama hasar tutarına göre belirlenmiştir. Yıllar itibariyle her yılın ortalama hasar tutarının logaritmsı alınarak μ parametresi belirlenmiştir. Şekil 5.1 'de verilen farklı σ değerleri için lognormal dağılım incelenerek, σ parametresi kuyruk yapısına uygun olarak 0,5 olarak varsayılmıştır. σ parametresinin 0,5 olarak varsayılmasının nedeni, σ parametresi bire yakın olduğunda ödenmiş hasar tutarı değerleri sıfıra çok yakındır, çalışmada tüm poliçelerin en az bir hasara sahip olduğu varsayımı yapıldığından dolayı σ değerinin birden küçük olarak varsayılmıştır. Ancak σ değeri sıfıra yaklaştıkça dağılımın yayılımı azalmaktadır. Veri kümesinde hem çok küçük hem de çok büyük ödenmiş hasar tutarları olması için σ parametresi 0,5 olarak belirlenmiştir. İllere göre belirlenen

parametre deęerleri yıllar için toplam poliçe sayısı ve toplam hasar tutarı göz önünde bulundurularak hesaplanmıştır. İllere göre parametre deęerleri 2005 yılı için Çizelge 5.2 'de, 2006 yılı için Çizelge 5.3 'te, 2007 yılı için Çizelge 5.4 'te ve 2008 yılı için Çizelge 5.5 'te verilmektedir.

Çizelge 5.2. 2005 yılı için illere göre lognormal dağılımın parametre deęerleri

	2005 yılı				
	ÖDENMİŞ HASAR TUTARI	POLİÇE ADEDİ	ORTALAMA HASAR TUTARI	$\mu=\ln(\text{ORTALAMA HASAR TUTARI})$	σ
İSTANBUL	300.854.000	2.160.950	139,22303	4,93608	0,5
ANKARA	97.486.000	948.248	102,80644	4,63285	0,5
İZMİR	56.811.000	607.053	93,58491	4,53887	0,5
ANTALYA	25.617.000	372.830	68,70960	4,22989	0,5
BURSA	33.415.000	336.550	99,28688	4,59801	0,5
DİĞER	309.446.000	4.572.415	67,67671	4,21474	0,5
TOPLAM	823.629.000	8.998.046			

(Kaynak: TRAMER (Ödenmiş hasar tutarı ve poliçe adedi için))

Çizelge 5.3. 2006 yılı için illere göre lognormal dağılımın parametre deęerleri

	2006 yılı				
	ÖDENMİŞ HASAR TUTARI	POLİÇE ADEDİ	ORTALAMA HASAR TUTARI	$\mu=\ln(\text{ORTALAMA HASAR TUTARI})$	σ
İSTANBUL	354.944.000	2.302.036	154,18699	5,03817	0,5
ANKARA	112.428.000	1.013.324	110,94971	4,70908	0,5
İZMİR	68.527.000	665.624	102,95152	4,63426	0,5
ANTALYA	30.587.000	412.155	74,21237	4,30693	0,5
BURSA	41.324.000	370.600	111,50567	4,71408	0,5
DİĞER	372.882.000	5.082.336	73,36823	4,29549	0,5
TOPLAM	980.692.000	9.846.075			

(Kaynak:TRAMER (Ödenmiş hasar tutarı ve poliçe adedi için))

Çizelge 5.4. 2007 yılı için illere göre lognormal dağılımın parametre değerleri

	2007 yılı				
	ÖDENMİŞ HASAR TUTARI	POLİÇE ADEDİ	ORTALAMA HASAR TUTARI	$\mu=\ln(\text{ORTALAMA HASAR TUTARI})$	σ
İSTANBUL	425.854.000	2.420.284	175,95208	5,17021	0,5
ANKARA	134.981.000	1.070.542	126,08660	4,83697	0,5
İZMİR	78.372.000	696.642	112,49968	4,72295	0,5
ANTALYA	36.645.000	442.082	82,89186	4,41754	0,5
BURSA	45.584.000	393.231	115,92168	4,75291	0,5
DİĞER	436.581.000	5.407.535	80,73568	4,39118	0,5
TOPLAM	1.158.017.000	10.430.316			

(Kaynak:TRAMER(Ödenmiş hasar tutarı ve poliçe adedi için))

Çizelge 5.5. 2008 yılı için illere göre lognormal dağılımın parametre değerleri

	2008 yılı				
	ÖDENMİŞ HASAR TUTARI	POLİÇE ADEDİ	ORTALAMA HASAR TUTARI	$\mu=\ln(\text{ORTALAMA HASAR TUTARI})$	σ
İSTANBUL	500.606.000	2.474.703	202,28933	5,30970	0,5
ANKARA	161.263.000	1.111.331	145,10798	4,97748	0,5
İZMİR	93.589.000	720.123	129,96252	4,86725	0,5
ANTALYA	43.225.000	472.590	91,46406	4,51595	0,5
BURSA	55.310.000	431.499	128,18106	4,85344	0,5
DİĞER	504.021.000	5.730.235	87,95817	4,47686	0,5
TOPLAM	1.358.014.000	10.940.481			

(Kaynak:TRAMER(Ödenmiş hasar tutarı ve poliçe adedi için))

Çizelge 5.2, Çizelge 5.3, Çizelge 5.4 ve Çizelge 5.5 'te il gruplarına göre verilen μ ve σ parametreleri kullanılarak, yıllar itibariyle poliçe adedi ve toplam ödenmiş hasar tutarı bilgisi türetilmiştir ve elde edilen sonuçlar Çizelge 5.6 'da özetlenmiştir.

Çizelge 5.6. Yıllara göre türetilen toplam poliçe adedi ve ödenmiş hasar tutarı

YILLAR	POLİÇE ADEDİ	TOPLAM ÖDENMİŞ HASAR TUTARI (TL)	ORTALAMA HASAR TUTARI (TL)
2005	8.998	941.152	104,60
2006	9.846	1.129.027	114,67
2007	10.430	1.311.366	125,73
2008	10.940	1.561.553	142,74
TOPLAM	40.214	4.943.098	122,92

Poliçelere ilişkin hasar tutarları türetildikten sonra, bağımsız değişkenler olan araç türü, araç modeli, sürücünün cinsiyeti, sürücünün yaşı ve yakıt türü, TÜİK ve Emniyet Genel Müdürlüğü tarafından açıklanan istatistikler göz önünde bulundurularak türetilmiştir. Bağımlı değişken ve bağımsız değişkenler Çizelge 5.7 'de verilmiştir. Çizelge 5.7 'de verilen değişkenlere ilişkin bilgilere ulaşılabildiğinden, bu değişkenler risk sınıflandırmasında kullanılmıştır.

Çizelge 5.7. Bağımlı değişken ve bağımsız değişkenler

Bağımsız Değişkenler		Bağımlı Değişken
İl plaka kodu	Yakıt türü	Ödenmiş hasar tutarı(TL)
Araç türü	Sürücünün cinsiyeti	
Araç modeli	Sürücünün yaşı	

Ödenmiş Hasar Tutarı (TL): Ödenmiş hasar tutarının her il sınıfı için lognormal dağılıma sahip olduğu varsayımı ile. Çizelge 5.3, Çizelge 5.4, Çizelge 5.5 ve Çizelge 5.6'da verilen μ ve σ parametrelerine göre türetilmiştir. Risk sınıflandırması ödenmiş hasar tutarına göre yapılmıştır.

İl Plaka Kodu: Bağımsız değişkenlerden biri olan il plaka kodu, ödenmiş hasar tutarlarının hangi ilde olduğuna dair bilgi vermektedir. Burada İl Plaka Kodu altı sınıftan oluşmaktadır. TRAMER 'de verilen hasar tutarları ve poliçe adetleri göz önünde bulundurularak en çok hasar getiren ilk beş il ele alınmıştır. Geride kalan diğer iller ise tek bir sınıf olarak değerlendirilmiştir. Çünkü ilk beş il yıllar itibarıyla

hasar tutarının ve poliçe adetlerinin büyük bir kısmına sahiptir. İl Plaka Kodu grupları; İstanbul, Ankara, İzmir, Antalya, Bursa ve diğer iller olarak belirlenmiştir. İllere ait poliçe adetleri ve oranları Çizelge 5.8 'de verilmiştir. Çizelge 5.8 'de verilen oranlar, yıllar itibariyle illere ait poliçe adetlerinin toplam poliçe adetlerine oranı ile bulunmuştur. Çizelge 5.8 'de verilen oranlara göre yıllar itibariyle poliçelere il plaka kodu değişkeni atanmıştır.

Çizelge 5.8. İl plaka kodu oranları

İL PLAKA KODU	2005 yılı		2006 yılı		2007 yılı		2008 yılı	
	Poliçe Adedi	Oran	Poliçe Adedi	Oran	Poliçe Adedi	Oran	Poliçe Adedi	Oran
İSTANBUL	2.160.950	0,24016	2.302.036	0,23380	2.420.284	0,23204	2.474.703	0,22620
ANKARA	948.248	0,10538	1.013.324	0,10292	1.070.542	0,10264	1.111.331	0,10158
İZMİR	607.053	0,06746	665.624	0,06760	696.642	0,06679	720.123	0,06582
ANTALYA	372.830	0,04143	412.155	0,04186	442.082	0,04238	472.590	0,04320
BURSA	336.550	0,03740	370.600	0,03764	393.231	0,03770	431.499	0,03944
DİĞER	4.572.415	0,50816	5.082.336	0,51618	5.407.535	0,51844	5.730.235	0,52376
TOPLAM	8.998.046	1	9.846.075	1	10.430.316	1	10.940.481	1

(Kaynak:TRAMER)

Araç Türü: Araç türü analizde kullanılan bir bağımsız değişkendir. Araç Türü altı sınıftan oluşmaktadır. Araç Türü'ne ilişkin sınıflar; otomobil, minibüs (8-14 koltuk), otobüs (15 koltuk ve üstü), kamyonet, kamyon ve diğer araçlar şeklinde belirlenmiştir. Araç türü değişkenine ilişkin istatistikler TRAMER 'den alınmıştır. Yıllar itibariyle araç türü değişkeni oranları Çizelge 5.9 'da verilmiştir. Çizelge 5.9 'da verilen oranlara göre poliçelere araç türü değişkeni atanmıştır.

Çizelge 5.9. Araç türü oranları

ARAÇ TÜRÜ	2005 yılı		2006 yılı		2007 yılı		2008 yılı	
	Police Adedi	Oran	Police Adedi	Oran	Police Adedi	Oran	Police Adedi	Oran
OTOMOBİL	5.749.531	0,63898	6.058.641	0,61534	6.440.953	0,61752	6.707.712	0,61311
MINİBÜS (8-14 KOLTUK)	313.414	0,03483	325.699	0,03308	337.915	0,03240	335.551	0,03067
OTOBÜS (15 KOLTUK VE ÜSTÜ)	147.170	0,01636	154.487	0,01569	162.667	0,01560	169.160	0,01546
KAMYONET	1.382.310	0,15362	1.585.011	0,16098	1.773.564	0,17004	1.910.349	0,17461
KAMYON	494.360	0,05494	496.681	0,05044	486.416	0,04663	468.772	0,04285
DİĞER	911.261	0,10127	1.225.556	0,12447	1.228.801	0,11781	1.348.937	0,12330
TOPLAM	8.998.046	1	9.846.075	1	10.430.316	1	10.940.481	1

(Kaynak:TRAMER)

Sürücünün Cinsiyeti: Sürücülere ilişkin bilgiler TRAMER tarafından açıklanan istatistiklerde yer almadığından, sürücünün cinsiyeti için Emniyet Genel Müdürlüğü'nün yıllar itibariyle sürücülerin cinsiyetlerine ilişkin istatistiklerinden yararlanılmıştır. Sürücünün cinsiyeti, erkek ve kadın olmak üzere iki sınıftan oluşmaktadır. Yıllar itibariyle sürücü cinsiyet oranları Çizelge 5.10'da verilmektedir.

Çizelge 5.10. Sürücünün cinsiyeti oranları

CİNSİYET	2005	2006	2007	2008
ERKEK	0,84260	0,83987	0,83724	0,76688
KADIN	0,15740	0,16013	0,16276	0,23312
TOPLAM	1	1	1	1

(Kaynak: Emniyet Genel Müdürlüğü)

Sürücünün Yaşı: Sürücünün yaşına ilişkin olarak da gerek TÜİK kayıtlarında gerekse Emniyet Genel Müdürlüğü kayıtlarında yeterli bilgi bulunamamıştır. Bu nedenle sürücünün yaş bilgisi, 20 ile 70 yaşları arasında tekdüze (uniform) dağılıma göre türetilmiştir.

Araç Modeli: TÜİK tarafından açıklanan motorlu taşıtlara ilişkin istatistiklerde yer alan kayıtlı araçların model yıllarına ilişkin istatistiklere göre araç model yılı değişkeni türetilmiştir. TÜİK tarafından açıklanan ve araç modeli değişkenine esas olan istatistikler Ek-1, Ek-2, Ek-3 ve Ek-4 'te verilmiştir.

Yakıt Türü: Araçları sınıflandırmada önemli değişkenlerden biri olan Yakıt Türü bağımsız değişkeni, benzin, dizel ve otopaz(LPG) olmak üzere üç sınıftan oluşmaktadır. Ülkemizde trafiğe kayıtlı araçların yakıt türlerine ilişkin olarak TÜİK tarafından yıllar itibariyle açıklanan istatistiklere göre belirlenen yakıt türü için oranlar Çizelge 5.11 'de verilmiştir.

Çizelge 5.11. Yakıt türü oranları

YAKIT TÜRÜ	2005	2006	2007	2008
BENZİN	0,52193	0,49974	0,47049	0,43924
DİZEL	0,35716	0,36809	0,38161	0,39280
OTOGAZ (LPG)	0,12092	0,13218	0,14790	0,16796

(KAYNAK:TÜİK)

5.3. Trafik Sigortası İçin Risk Sınıflandırması

Ödenmiş hasar tutarlarına göre risk sınıflandırması yapılırken, ödenmiş hasar tutarı değişkeni illere göre türetildiğinden sınıflandırmanın yansız olması için il değişkeni analizin dışında tutulmuştur. SPSS Clementine 11.1 programında yıllar itibariyle Karar Ağaçları'ndan CHAID ve CRT algoritmaları ile risk sınıflandırması yapılmıştır. SPSS Clementine 11.1'de karar ağacı derinliği seçimi kullanıcıya bağlıdır. Derinlik sayısı ile sınıf sayısı doğru orantılıdır. Derinlik sayısı ne kadar artarsa, sınıf sayısı da derinlik sayısına paralel olarak artmaktadır. Ancak derinlik sayısına karar vermedeki en önemli ölçüt yapraklarda yeteri kadar verinin bulunmasıdır., Yapraklarda yeteri kadar veri bulunduğu ve karar ağaçlarının veri kümesiyle ilişkisi yüksek (0,50 'den fazla) olduğundan (Çizelge 5.12), derinlik sayısı beş olarak belirlenmiştir. CHAID ve CRT algoritmalarından hangisinin veri seti için daha uyumlu olduğu, SPSS Clementine 11.1 'de analiz (analysis) ile

yapılmaktadır. Veri setine uygun karar ağaçlarının analizinin sonuçları Çizelge 5.12'de verilmiştir.

Çizelge 5.12. Yıllar itibariyle CHAID ve CRT algoritmalarının veri kümesi ile doğrusal ilişki sonuçları

YILLAR	CHAID	CRT
2005	0,612	0,613
2006	0,597	0,600
2007	0,554	0,558
2008	0,537	0,540

Çizelge 5.12 yıllara göre karar ağaçlarının veri kümesi ile olan doğrusal ilişkisini göstermektedir. Bu değerlerin bire yakın olması yapılan sınıflandırmanın veri kümesine daha iyi uyum sağladığını göstermektedir. Çizelge 5.12 'ye göre dört yıl için CRT algoritmasının ilişki değeri risk sınıflandırmasında daha iyi sonuç verdiği görülmektedir. CRT algoritmasının yaptığı sınıflandırma veri kümesine daha uygun olduğundan dolayı risk sınıflandırmasında, CRT algoritmasından elde edilen sınıflar kullanılmıştır.

2005-2008 yılları için CHAID ve CRT algoritmalarına göre yapılan sınıflandırma sonucunda (EK 5 - EK 12) hasar tutarına etki eden en önemli değişkenin araç türü olduğu görülmüştür. Yıllara göre risk sınıflandırmasında karar ağaçlarının ilk derinliğinde bulunan düğümlere göre risk sınıfları belirlenerek, bu risk sınıflarına göre hasar rezervi tahmin edilmiştir. Ancak karar ağaçları incelendiğinde yıllar itibariyle oluşan risk grupları birbirlerinden farklılık göstermektedir. Örnek olarak, Ek 6'da 2005 yılı için verilen CRT algoritmasında ilk derinlikte iki risk grubu bulunmaktadır. Birinci risk grubunda kamyon, minibüs ve otobüs, ikinci risk grubunda kamyonet, otomobil ve diğer araçlar bulunmaktadır. Birinci risk grubu için karar ağacının ikinci derinliği incelendiğinde, birinci risk grubu araç türü bakımından tekrar iki risk grubuna ayrılmaktadır (Node 2 ve Node 3). Minibüs bir risk grubu, kamyon ve otobüs ise diğer bir risk grubudur. Bu şekilde her bir düğüm derinlik arttıkça iç düğümlere ayrılmaktadır. CRT algoritması incelendiğinde ilgilenilen yıllar için bulunan gruplar özellikleri bakımından farklılık göstermektedir. Rezerv hesaplamasında her yıl için ayrı grupla çalışılması mümkün olmadığı ve

gruplar arası farkın gösterilebilmesi için karar ağaçlarından elde edilen risk gruplarının tümü yerine, karar ağaçlarının ilk derinliğinde bulunan risk grupları analizde kullanılmıştır. 2008 yılı için yapılan risk sınıflandırmasında ilk derinlikte bulunan risk grupları diğer yıllarda bulunandan farklılık göstermekle birlikte, bu farklılık tek bir araç türünden kaynaklandığı için çalışmada rezerv tahminlerinde bu farklılık göz ardı edilmiştir.

Çizelge 5.13. 2005, 2006 ve 2007 yılları için CRT algoritması ile yapılan risk sınıflandırmasında ilk derinlikteki sınıflar

1. RİSK SINIFI	2. RİSK SINIFI
Kamyon	Diğer
Minibüs	Kamyonet
Otobüs	Otomobil

Çizelge 5.13, 2005 yılı için Ek-6, 2006 yılı için Ek-8 ve 2007 yılı için Ek-10 'da verilen CRT algoritmasının ilk derinliğinde bulunan sınıfları göstermektedir.

Çizelge 5.14. 2008 yılı için CRT algoritması ile yapılan risk sınıflandırmasında ilk derinlikteki sınıflar

1. RİSK SINIFI	2. RİSK SINIFI
Kamyon	Diğer
Kamyonet	
Minibüs	Otomobil
Otobüs	

Çizelge 5.14, 2008 yılı için Ek-12 'de verilen CRT algoritmasının ilk derinliğinde bulunan sınıfları göstermektedir.

Çizelge 5.13 ve Çizelge 5.14 incelendiğinde portföydeki toplam sigortalılar ödenmiş hasar türüne göre iki risk grubuna ayrılmaktadır. Risk sınıflarının rezerve etkisinin incelenebilmesi için, çalışmada tüm yıllar için risk sınıfları birinci risk

grubu kamyon, minibüs ve otobüs, ikinci risk grubu diğer araçlar, kamyonet ve otomobil olarak belirlenmiş ve yıllar itibariyle hasar rezervi tahmin edilmiştir.

CHAID ve CRT algoritmalarındaki her düğümde, düğüm numarası, düğümde bulunan poliçe sayısı, düğümde bulunan poliçe yüzdesi ve düğümün ortalama hasar tutarı tahminlerini vermektedir (Kumar, 2004).

Çizelge 5.15 'de kök düğümün ve ilk derinlikteki düğüm bilgileri verilmektedir. Burada "TOPLAM" olarak belirtilen kök düğümü ifade etmektedir ve sınıflandırma yapılmadan tüm poliçelere ilişkin bilgileri vermektedir. Risk gruplarına göre verilen değerler ise ilk derinlikte bulunan bilgilerdir. Bu bilgiler incelendiğinde yıllar itibariyle portföyde daha çok kamyonet, otomobil ve diğer araçlar bulunmaktadır ve bu araçların ortalama hasar tutarları ikinci risk grubunda bulunan araçların ortalama hasar tutarlarından daha düşüktür.

Çizelge 5.15. Yıllara göre CRT algoritmasının kök düğüm ve ilk derinlikteki düğüm bilgileri

Ek-No	Yıllar	Düğüm	Düğümdeki Poliçe Sayısı (n)	%	Tahmini Ortalama Hasar Tutarı (TL)
Ek-6	2005	TOPLAM	8.998	100,00	104,60
		1. Risk Grubu	994	11,05	255,00
		2. Risk Grubu	8.004	88,95	85,92
Ek-8	2006	TOPLAM	9.846	100,00	114,67
		1. Risk Grubu	1.038	10,54	271,92
		2. Risk Grubu	8.808	89,46	96,14
Ek-10	2007	TOPLAM	10.430	100,00	125,73
		1. Risk Grubu	992	9,51	294,92
		2. Risk Grubu	9.438	90,49	107,95
Ek-12	2008	TOPLAM	10.940	100,00	142,74
		1. Risk Grubu	2.882	26,34	244,75
		2. Risk Grubu	8.058	73,66	106,25

Yıllar İtibariyle yapılan sınıflandırmaya göre, kamyon, minibüs ve otobüsten oluşan birinci risk grubu portföyün daha az bir kısmını, kamyonet, otomobil ve diğer araçlardan oluşan ikinci risk grubu ise portföyün büyük kısmını kapsamaktadır.

İkinci risk grubu için araç başına ortalama hasar tutarı, birinci risk grubunun ortalama hasar tutarından daha düşüktür.

Örnek olarak Çizelge 5.15 'te verilen 2005 yılına ait bilgiler aşağıda yorumlanmaktadır:

2005 yılı için ortalama hasar tutarı 104,60 TL 'dir.

Birinci risk grubu portföyün %11,05 'ine sahiptir ve ortalama hasar tutarı 255 TL 'dir.

İkinci risk grubu ise portföyün %88,95 'ine sahiptir ve ortalama hasar tutarı 85,92 TL 'dir.

Portföyün büyük bir kısmını kapsayan ikinci risk grubu daha düşük hasar tutarına, portföyün küçük bir kısmı olan birinci risk grubu daha yüksek hasar tutarına sahiptir. Sigorta şirketleri hasar tutarına göre, müşteri profillerini belirleyebilir ve portföylerini risk sınıflarına göre değerlendirebilirler.

5.4. Hasar Rezervi Tahmini

Çalışmanın başında rezerv hesaplamasına ilişkin çeşitli yöntemler verilmişti. Uygulamada türetilebilen veride sadece ödenmiş hasar tutarları olduğundan Mack Zincir Merdiven Yöntemi ile hasar rezervi tahmin edilmiştir. Burada dört yıllık veri kümesi kullanıldığı için, hasar gelişim yılı da dört dönem olarak belirlenmiştir. 2005 yılındaki poliçeler 2005 kaza yılına ait ödenmiş hasar tutarlarının belirlenmesinde, 2006 yılındaki poliçeler 2006 kaza yılına ait ödenmiş hasar tutarlarının belirlenmesinde, 2007 yılındaki poliçeler 2007 kaza yılına ait ödenmiş hasar tutarlarının belirlenmesinde ve 2008 yılındaki poliçeler 2008 kaza yılına ait ödenmiş hasar tutarlarının belirlenmesinde kullanılmıştır. Poliçelere hasar gelişim yıllarını atamak için TRAMER tarafından açıklanan hasar gelişim üçgenlerindeki gelişim yıllarına ilişkin tutarlar dikkate alınarak, veri kümesine hasar gelişim yılları da değişken olarak eklenmiş, ancak risk sınıflandırmasında analiz dışında tutulmuştur.

Çizelge 5.16, Çizelge 5.17 ve Çizelge 5.18 yıllar itibariyle aşamalı ödenmiş hasar tutarlarını vermektedir.

Çizelge 5.16. Risk sınıflandırması yapılmadan oluşturulan hasar gelişim üçgeni
(aşamalı hasar tutarlarına göre) (TL)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	234.208,73	568.947,20	103.553,30	34.442,96
2006	269.787,52	726.073,11	133.165,94	
2007	384.768,73	926.597,41		
2008	1.561.552,71			

Çizelge 5.16 risk sınıflandırması yapılmadan aşamalı hasar tutarı bilgisini vermektedir. Çizelge 5.16 'ya göre ;

- 2005 yılında meydana gelmiş ve aynı yıl ödenen hasar tutarı 234.208,73 TL'dir.
- 2005 yılında meydana gelmiş ve 2006 yılında ödenen hasar tutarı 568.947,20 TL'dir.
- 2005 yılında meydana gelmiş ve 2007 yılında ödenen hasar tutarı 103.553,30 TL'dir.
- 2005 yılında meydana gelmiş ve 2008 yılında ödenen hasar tutarı 34.442,96 TL'dir.

Diğer kaza yıllarına ait ödenmiş hasar tutarı değerleri de aynı şekilde yorumlanır.

Çizelge 5.17. Birinci risk grubu (kamyon, minibüs, otobüs) için hasar gelişim üçgeni (aşamalı hasar tutarlarına göre) (TL)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	70.009,15	145.517,41	30.414,50	7.525,57
2006	67.983,02	181.575,76	32.690,19	
2007	95.015,50	197.540,12		
2008	316.634,62			

Çizelge 5.18. İkinci risk grubu (kamyonet, otomobil, diğer araçlar) için hasar gelişim üçgeni (aşamalı hasar tutarlarına göre) (TL)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	164.199,57	423.429,79	73.138,80	26.917,39
2006	201.804,50	544.497,35	100.475,75	
2007	289.753,23	729.057,29		
2008	1.244.918,09			

Mack Zincir Merdiven Yöntemi'nin uygulanabilmesi için birikimli ödenmiş hasar tutarlarına ihtiyaç olduğundan, aşamalı hasar tutarlarına göre oluşturulan hasar gelişim üçgenleri (Çizelge 5.16 - 5.18), birikimli hasarlara göre yeniden düzenlenmiş ve Çizelge 5.19 - 5.21 'de verilmiştir.

Çizelge 5.19. Risk sınıflandırması yapılmadan hasar gelişim üçgeni (birikimli hasar tutarlarına göre)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	234.208,73	803.155,92	906.709,23	941.152,18
2006	269.787,52	995.860,63	1.129.026,57	
2007	384.768,73	1.311.366,14		
2008	1.561.552,71			

Çizelge 5.19 risk sınıflandırması yapılmadan birikimli hasar tutarı bilgisini vermektedir. Çizelge 5.19 'a göre ;

- 2005 yılında meydana gelmiş ve aynı yıl ödenen hasar tutarı 234.208,73 TL'dir.
- 2005 yılında meydana gelmiş, 2005 ve 2006 yıllarında ödenen toplam hasar tutarı 803.155,92² TL 'dir. Bu değer 2005 yılında meydana gelmiş ve 2005 yılında ödenen aşamalı hasar tutarı (234.208,73 TL) ile 2005 yılında meydana gelmiş 2006 yılında ödenmiş hasar tutarının (568.947,20 TL) toplamından oluşmaktadır.
- 2005 yılında meydana gelmiş, 2005,2006 ve 2007 yıllarında ödenen toplam hasar tutarı 906.709,23 TL 'dir (234.208,73 + 568.947,20 + 103.553,30).
- 2005 yılında meydana gelmiş, 2005, 2006, 2007 ve 2008 yıllarında ödenen toplam hasar tutarı 941.152,18TL 'dir.

biçimindedir. Diğer kaza yıllarına ait ödenmiş hasar tutarı değerleri de aynı şekilde yorumlanır.

² Yapılan hesaplamada yuvarlama olduğundan dolayı 0,01 'lik bir fark görülmektedir.

Çizelge 5.20. Birinci risk grubu (kamyon, minibüs, otobüs) için hasar gelişim üçgeni (birikimli hasar tutarlarına göre) (TL)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	70.009,15	215.526,56	245.941,06	253.466,63
2006	67.983,02	249.558,79	282.248,97	
2007	95.015,50	292.555,62		
2008	316.634,62			

Çizelge 5.21. İkinci risk grubu (kamyonet, otomobil, diğer araçlar) için hasar gelişim üçgeni (birikimli hasar tutarlarına göre) (TL)

Kaza Yılı	Hasar Gelişim Yılı			
	0	1	2	3
2005	164.199,57	587.629,36	660.768,16	687.685,55
2006	201.804,50	746.301,85	846.777,60	
2007	289.753,23	1.018.810,52		
2008	1.244.918,09			

Birikimli hasar tutarlarına göre yeniden düzenlenen hasar gelişim üçgenleri kullanılarak, Mack Zincir Merdiven Yöntemi 'ne göre tahmin edilen, rezerv tutarları Çizelge 5.22 , Çizelge 5.23 ve Çizelge 5.24 'te verilmiştir.

Çizelge 5.22. Mack zincir merdiven yöntemine göre risk sınıflandırması yapılmadan tahmin edilen rezerv miktarı (TL)

Kaza Yılı	Hasar Gelişim Yılı				Hasar Rezervi
	0	1	2	3	
2005	234.208,73	803.155,92	906.709,23	941.152,18	0,00
2006	269.787,52	995.860,63	1.129.026,57	1.171.915,00	42.888,00
2007	384.768,73	1.311.366,14	S₂₂	1.540.288,00	228.922,00
2008	1.561.552,71	S₃₁	S₃₂	6.418.915,00	4.857.362,00
Toplam					5.129.172,00

Hasar rezervi **5.129.172,00** olarak tahmin edilmiştir. Hasar rezervi birikimli hasarlarla hesaplandığından aşağıdaki biçimde de elde edilir :

$$\text{Hasar Rezervi} = (941.152,18 + 1.171.915,00 + 1.540.288,00 + 6.418.915,00) -$$

$$(941.152,18 + 1.129.026,57 + 1.311.366,14 + 1.561.552,71)$$

$$\cong 5.129.172,00$$

Mack Zincir Merdiven Yöntemi'ne göre elde edilen 5.129.172,00 TL'lik hasar rezervinin standart hatası 206.841,55'tir.

Risk sınıflandırması yapılmadan R programında hesaplanan rezerv miktarına ilişkin kodlar Ek-13 'te verilmiştir.

Sınıflandırma yapılmadan önce ve sonra tahmin edilen rezerv miktarlarının karşılaştırılması için iki risk grubunun rezerv tahminleri Çizelge 5.23 ve Çizelge 5.24 kullanılarak elde edilmiştir.

Çizelge 5.23. Mack zincir merdiven yöntemine göre birinci risk grubu (kamyon, minibüs, otobüs) için tahmin edilen rezerv miktarı (TL)

Kaza Yılı	Hasar Gelişim Yılı				Hasar Rezervi
	0	1	2	3	
2005	70.009,15	215.526,56	245.941,06	253.466,63	0,00
2006	67.983,02	249.558,79	282.248,97	290.886,00	8.637,00
2007	95.015,50	292.555,62	S ₂₂	342.417,00	49.862,00
2008	316.634,62	S ₃₁	S ₃₂	1.205.033,00	888.398,00
Toplam					946.897,00

Hasar rezervi **946.897,00** olarak tahmin edilmiştir. Hasar rezervi birikimli hasarlarla hesaplandığından aşağıdaki biçimde de elde edilir :

$$\text{Hasar Rezervi} = (253.466,63+290.886,00+342.417,00+1.205.033,00)-$$

$$(253.466,63+282.248,97+292.555,62+316.634,62)$$

$$\cong 946.897,00$$

Mack Zincir Merdiven Yöntemi 'ne göre elde edilen 946.897,00TL 'lik hasar rezervinin standart hatası 93.269,66 'dır. Birinci risk grubu için R programında hesaplanan rezerv miktarına ilişkin kodlar Ek-14 'te verilmiştir

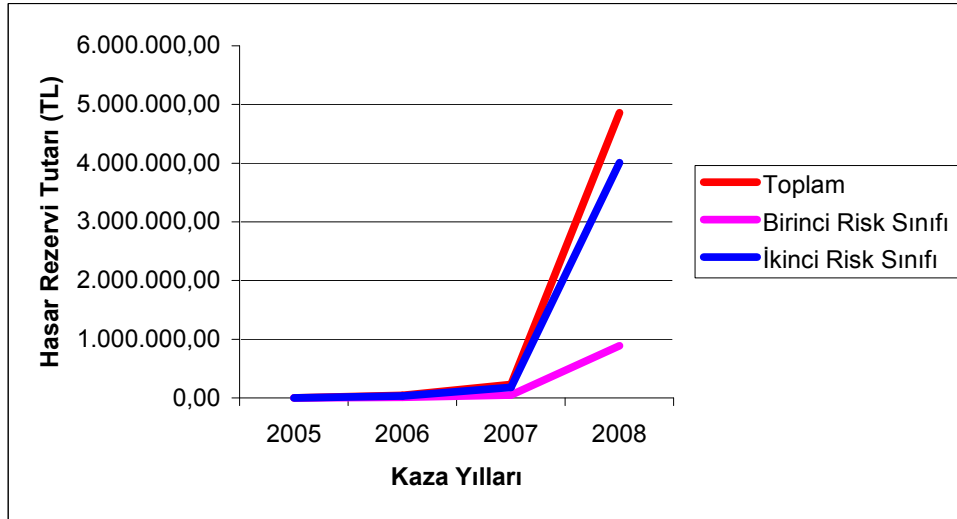
Çizelge 5.24. Mack zincir merdiven yöntemine göre ikinci risk grubu (kamyonet, otomobil, diğer araçlar) için tahmin edilen rezerv miktarı (TL)

Kaza Yılı	Hasar Gelişim Yılı				Hasar Rezervi
	0	1	2	3	
2005	164.199,57	587.629,36	660.768,16	687.685,55	0,00
2006	201.804,50	746.301,85	846.777,60	881.272,00	34.495,00
2007	289.753,23	1.018.810,52	S ₂₂	1.198.316,00	179.505,00
2008	1.244.918,09	S ₃₁	S ₃₂	5.253.512,00	4.008.594,00
Toplam					4.222.594,00

Hasar rezervi **4.222.594,00** olarak tahmin edilmiştir. Hasar rezervi birikimli hasarlarla hesaplandığından aşağıdaki biçimde de elde edilir :

$$\begin{aligned} \text{Hasar Rezervi} &= (687.685,55+881.272,00+1.198.316,00+5.253.512,00)- \\ &\quad (687.685,55+846.777,60+1.018.810,52+1.244.918,09) \\ &\cong 4.222.594,00 \end{aligned}$$

Mack Zincir Merdiven Yöntemi 'ne göre elde edilen 4.222.594,00 TL'lik hasar rezervinin standar hatası 104.695,12 'dir. İkinci risk grubu için R programında hesaplanan rezerv miktarına ilişkin kodlar Ek-15 'te verilmiştir.



Şekil 5.2. Kaza yıllarına göre hasar rezervi tutarı

Risk sınıfları göz önünde bulundurulmadan Mack yöntemine göre tahmin edilen rezerv miktarı 5.129.172,40 TL 'dir. Birinci risk sınıfı için hasar rezervi 946.896,60 TL ve ikinci risk sınıfı için hasar rezervi 4.222.594,33 TL olmak üzere toplam hasar rezervi 5.169.490,93 TL 'dir. Arada 40.318,53 TL fark olup, risk sınıflandırması yapılmadan önce hesaplanan hasar rezervi tutarının, risk sınıflandırması yaptıktan sonraki hasar rezervi tutarından daha düşük olduğu görülmektedir.

ALTINCI BÖLÜM

6. SONUÇLAR VE ÖNERİLER

Sigorta şirketlerinin portföyünde bulunan sigortalılar, risk özellikleri bakımından farklılık gösterirler. Bu nedenle sigorta şirketleri, müşteri profilini belirlerken ya da fiyatlandırma işlemlerinde risk sınıflarından yararlanır.

Sigorta şirketlerinin yükümlülüklerini karşılayabilmesi için ayırdığı rezerv miktarı önemlidir. Çünkü sigorta şirketi henüz gerçekleşmemiş ya da şirket bilgisi dahilinde olmayan hasarlar için rezerv ayırmak zorundadırlar. Hayat dışı branşlarda faaliyet gösteren sigorta şirketleri, hasar karşılığı olarak ayıracakları rezerv miktarını geçmiş hasar ödemelerine göre tahmin etmektedir. Sigorta şirketlerinin portföylerinde bulunan her risk grubu farklı risk özelliklerine sahip olduğundan, risk gruplarına ilişkin rezerv miktarı farklılık gösterebilir.

Bu bilgiler kullanılarak çalışmada Trafik Sigortası için risk sınıflandırması yapılmış ve risk sınıflandırmasının hasar rezervi üzerindeki etkisi incelenmiştir. Poliçe ayırımında sürücü, araç, hasar sayısı ve hasar tutarı bilgileri elde edilemediğinden, TRAMER, TÜİK ve Emniyet Müdürlüğü'ndeki istatistikler kullanılarak, 2005-2008 yılları için hasar bilgileri, hasarı yapan araca ait bilgiler de dahil olmak üzere veri türetilmiştir.

Yıllar itibariyle risk sınıflandırması CHAID ve CRT algoritmaları ile yapıp, hasar tutarı bağımlı değişken olduğundan, risk grupları hasar tutarına göre belirlenmiştir. Yapılan karar ağaçlarının yıllar itibariyle ilk derinliği göz önünde bulundurulduğunda, portföy iki risk grubuna ayrılmaktadır. Birinci risk grubu kamyon, minibüs ve otobüs; ikinci risk grubu kamyonet, otomobil ve diğer araçlardan oluşmaktadır. Yıllar itibariyle birinci risk grubunun portföydeki payının, ikinci risk grubunun portföydeki payından daha az olduğu görülmüştür. Sonuç olarak portföyün büyük bir kısmı kamyonet, otomobil ve diğer araçlardan oluşmaktadır. Risk gruplarının ortalama hasar tutarları karşılaştırıldığında, birinci risk grubunun ortalama hasar tutarının daha yüksek olduğu sonucuna ulaşılmıştır. Risk sınıflandırması ile, portföydeki payı az olan birinci risk grubunun ortalama hasar tutarı fazla iken, portföyde daha çok paya sahip olan ikinci risk grubunun ortalama hasar tutarı düşük olduğu görülmüştür. Bu sonuç sigorta şirketleri

açısından önemlidir. Çünkü sigorta şirketlerinin portföylerini daha az hasar getiren araçlardan oluşturma eğilimleri vardır. Risk sınıflandırması sonuçlarına göre, sigorta şirketleri portföylerini uygun şekilde değiştirebilir ve fiyatlandırmayı risk gruplarının özelliklerine göre düzenleyebilir.

Risk sınıflandırması sonucunda hasar tutarına etki eden değişkenlerin, araç türü dışında, aracın modeli ve sürücünün yaşı olduğu görülmüştür. Araç türü, aracın modeli ve sürücünün yaşı değişkenlerine göre yıllar itibariyle yapılan sınıflandırmada risk grupları birbirinden farklılık göstermektedir. Risk gruplarının rezerve etkisi ve risk grupları arasında rezerv bakımından farklılığın incelenmesi için aynı risk grupları ile çalışılması gerektiğinden, yıllar itibariyle karar ağaçlarının ilk derinliğinde bulunan risk grupları kullanılarak, risk sınıflandırmasının rezerve etkisi incelenmiştir.

Risk sınıfları elde edildikten sonra, risk gruplarının aynı zamanda hasar rezervi miktarı üzerindeki etkisi de incelenmiştir. Risk gruplarına göre Mack Zincir Merdiven Yöntemi ile tahmin edilen rezerv miktarının, sınıflandırma yapılmadan tahmin edilen rezerv miktarından daha yüksek olduğu görülmüştür.

Risk sınıflandırmasına göre bulunan rezerv miktarları, sınıflandırma yapılmadan bulunan rezerv miktarlarından yüksek ya da düşük olabilir. Sigorta şirketleri düşük miktarda rezerv ayırmak isteyebilirler. Ancak bu durum yükümlülüklerini karşılayamama riskini ortaya çıkarabilir. Bu çalışmada açıklanan yöntemlerle sigorta şirketlerinin, portföylerini risk özelliklerine göre sınıflandırarak, üstlendikleri risklere ve risk gruplarına göre rezerv tahmini yapabilir ve bu durum şirketlerin geleceğini planlamasında etkili olabilir.

Çalışma kullanılan veri kümesi TRAMER, TÜİK ve Emniyet Müdürlüğü'nden alınan istatistiklere dayandırılmasına rağmen değişkenler rassal olarak türetildiği için sonuçlar rassal sonuçlardır. Trafik Sigortası için geçmişte yürürlükte olan poliçelerin bilgilerine ulaşıldığı takdirde farklı risk sınıfları bulunabilir ve buna bağlı olarak risk sınıflarının rezerve olan etkisi daha farklı olarak elde edilebilir.

Risk sınıflandırması sigortacılıkta genellikle portföylerin ve prim tutarlarının belirlenmesinde kullanılmaktadır. Özellikle kasko sigortaları için araç (türü, modeli, yakıt tipi, marka, araç değeri vb), sürücü bilgileri (yaşı, ikametgah, cinsiyet vb),

hasar tutarı ve hasar sayısı bilgileri kullanılarak çalışma genişletilebilir. Poliçe bilgileri elde edildiđi takdirde prim tutarları, hasar tutarı ve hasar sayısına bađlı olarak, VM yöntemleriyle sınıflandırma yapılarak, elde edilen grupların risk özellikleri karşılaştırılabilir.

KAYNAKLAR

- Akpınar, H., 2000, Veri Tabanlarında Bilgi Keşfi ve Veri Madenciliği, İ.Ü. İşletme Fakültesi Dergisi, c:29, 1, 1-22.
- Antonio, K., Beirlant, J., 2008, Risk Classification in Non-Life Insurance, Encyclopedia of Quantitative Risk Assessment, Insurance/Actuarial Risks Section, John Wiley & Sons, 2176p.
- Apte, C, Grossman, E., Pednault, E.P.D., Rosen, B.K., Tipu, F.A., White, B., 1999, Probabilistic Estimation Based Data Mining for Discovering Insurance Risks, IEEE Intelligent Systems, Vol.14, Issue 6, 49-58.
- Barnett, J., Zehnwrith, B., 2000, Best Estimates for Reserves, Proceedings of Casualty Actuarial Society, Vol.LXXXVII, Part 2, No.167, 245-321.
- Barnett, J., 19-22 June 2007, Cape Cod Credibility, Actuarial Studies in Non-Life Insurance (ASTIN), Orlando, Florida, United States of America.
- Berry, M., Linoff, G., 2004, Data Mining Techniques:for marketing, sales, and customer, Second Edition, Wiley Publishing, Indianapolis, 643p.
- Carvalho, D.H., Freitas, A.A., 2004, A hybrid decision tree/genetic algorithm method for data mining, Information Sciences, Volume 163, Issues 1-3.
- Casualty Loss Reserve Seminar, 1992, Denver, Colorado, 1158p.
- Ciarapica, F., Giacchetta, G., 2009, Classification and prediction of occupational injury risk using soft computing techniques: An Italian Study, Safety Science, Vol.47, Issue 1, 36-49.
- Dahl, P., 2003, Introduction to Reserving, Lecture Notes, www.math.su.se/matstat/und/sakii/pdf/dahl2003.pdf.
- Dolgun, M.Ö., 2006, Büyük Alışveriş Merkezleri İçin Veri Madenciliği Uygulamaları, Hacettepe Üniversitesi, İstatistik Anabilim Dalı, Yüksek Lisans Tezi.
- Emniyet Genel Müdürlüğü, www.emniyet.gov.tr.
- Friedland, J., 2009, Estimating Unpaid Claims Using Basic Techniques, Casualty Actuarial Society, Version II, 412p.
- Guo, L., 2003, Applying Data Mining Techniques in Property/Casualty Insurance, Casualty Actuarial Society E-Forum.
- Han, J., Kamber, M., 2006, Data Mining: Concepts and Technique, Morgan Kaufmann Publishers, Second Edition, 743p.
- Kaas, R., Goovaerts, M., Dhane, J., Denuit, M., 2001, Modern Actuarial Risk Theory, Kluwer Academic Publishers, Boston, 309p.

- Kahane, Y., Levin, N., Meiri, R., Zahavi, J., 2007, Applying Data Mining Technology for Insurance Rate Making-an Example of Automobile Insurance, Asia-Pacific Journal of Risk and Insurance: Vol. 2: Iss. 1, Article 3.
- Kdnuggets, Data Mining Community 's resources for Data Mining and Analytics Software, <http://www.kdnuggets.com/software/classification-tree-rules.html>.
- Kubrusly, J., Dias, M., Lage, M., Lopes, H., 2009, A Hybrid Chain Ladder and Gaussian Process Regression Method for IBNR Estimation, Contributed Paper at 4th BSMIF (Brazilian Conference on Statistical Modelling in Insurance and Finance).
- Kumar, T., 2004, SPSS Clementine for Data Mining in Institutional Research, Office of Information Management & Analysis, University of Northern Iowa, Power Point Presentation.
- Larose, D., 2005, Discovering Knowledge in Data, John Wiley & Sons, New Jersey, 222p.
- Quarg, G., Mack, T., 2008, Munich Chain Ladder: A Reserving Method that Reduces Gap between IBNR Projections Based on Paid Losses and IBNR Projections Based on Incurred Losses, Variance Journal, Volume 2, Issue 2, 266-299.
- SAS, 2002, Data Mining in the Insurance Industry, A SAS White Paper.
- Safavian, S.R., Landgrebe, D., 1991, A Survey of Decision Tree Classifier Methodology, Institute of Electrical and Electronics Engineers, Vol. 21, No. 3, 660-674.
- Sahasrabuddhe, R., 2008, Principles of Chain-Ladder "Method" Selecting and Updating Claims Development Factors, Casualty Actuarial Society E-Forum.
- Schmidt, K.D., 2006, Methods and Models of Loss Reserving Based on Run-Off Triangles: A Unifying Survey, Casualty Actuarial Society E-Forum.
- Schmidt, K.D., Zocher, M., 2008, The Bornhuetter-Ferguson as a General Principle of Loss Reserving, Casualty Actuarial Society E-Forum.
- Struzzieri, P., Hussian, P., 1998, Using Best Practices to Determine a Best Reserve Estimate, Casualty Actuarial Society Forum.
- Sugumaran, V., Muralidaran, V., Ramachandran, K.I., 2007, Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing, Mechanical Systems and Signal Processing, Volume 21, Issue 2, 930-942.
- TRAMER, www.tramer.org.tr.
- TÜİK, 2005, 2006, 2007, 2008, Motorlu Kara Taşıtları İstatistikleri.

- Wang, K.K., 2003, A Data Mining Analysis of Actuarial Data, Department of Statistics, University of Auckland, Master of Science Dissertation.
- Werner, G., Modlin, C., 2010, Basic Ratemaking, Casualty Actuarial Society, 316p.
- Westphal, C., Blaxton, T., 1998, Data Mining Solutions-Methods and Tools for Solving Real Word Problems, John Wiley & Sons, 603p.
- Zhou, Z., Chen, Z., 2002, Hybrid Decision Tree, Knowledge-Based Systems, Volume 15, Issue 8, 515-528.

EKLER DİZİNİ

	<u>Sayfa</u>
EK-1. 2005 yılı için araç türüne göre araç modelleri olasılıkları	64
EK-2. 2006 yılı için araç türüne göre araç modelleri olasılıkları	65
EK-3. 2007 yılı için araç türüne göre araç modelleri olasılıkları	66
EK-4. 2008 yılı için araç türüne göre araç modelleri olasılıkları	67
EK-5. 2005 yılı için CHAID algortiması.....	68
EK-6. 2005 yılı için CRT algortiması.....	69
EK-7. 2006 yılı için CHAID algortiması.....	70
EK-8. 2006 yılı için CRT algortiması.....	71
EK-9. 2007 yılı için CHAID algortiması.....	72
EK-10. 2007 yılı için CRT algortiması.....	73
EK-11. 2008 yılı için CHAID algortiması.....	74
EK-12. 2008 yılı için CRT algortiması.....	75
EK-13. R programında Mack Zincir Merdiven Yönteminin Kodları ve çıktısı (Toplam birikimli hasar tutarına göre).....	76
EK-14. R programında Mack Zincir Merdiven Yönteminin Kodları ve çıktısı (Birinci Risk grubunun birikimli hasar tutarına göre).....	77
EK-15. R programında Mack Zincir Merdiven Yönteminin Kodları ve çıktısı (İkinci Risk grubunun birikimli hasar tutarına göre).....	78

EK-1. 2005 yılı için araç türüne göre araç modelleri olasılıkları

2005 yılı						
Model Yılı	Otomobil	Minibüs	Otobüs	Kamyonet	Kamyon	Diğer
1979*	0,06653	0,09860	0,11372	0,12225	0,21307	0,25753
1980	0,00557	0,00451	0,00769	0,00624	0,01244	0,01273
1981	0,00354	0,00492	0,00603	0,00372	0,01784	0,01738
1982	0,00600	0,00841	0,01084	0,00430	0,02270	0,02172
1983	0,00745	0,01194	0,02127	0,00622	0,02349	0,02459
1984	0,00978	0,01441	0,02031	0,00736	0,02425	0,03071
1985	0,01210	0,01602	0,01892	0,00925	0,02637	0,02600
1986	0,01594	0,01815	0,02118	0,00863	0,02472	0,02046
1987	0,02003	0,01920	0,01727	0,00756	0,01878	0,02543
1988	0,02283	0,01814	0,01799	0,00652	0,02101	0,02402
1989	0,02039	0,01481	0,01962	0,00618	0,01618	0,01874
1990	0,03670	0,02340	0,03273	0,01063	0,02420	0,03190
1991	0,03429	0,02309	0,03255	0,01208	0,02064	0,03222
1992	0,04975	0,03605	0,04357	0,01669	0,03033	0,03150
1993	0,06907	0,04432	0,05418	0,02819	0,04053	0,04257
1994	0,05968	0,03380	0,03382	0,02439	0,02795	0,02603
1995	0,03600	0,02339	0,02059	0,01173	0,02193	0,02246
1996	0,04059	0,04000	0,03164	0,02740	0,03729	0,03085
1997	0,04921	0,06481	0,05057	0,05313	0,05769	0,03720
1998	0,06217	0,08274	0,05467	0,07290	0,06548	0,03588
1999	0,04510	0,06601	0,04026	0,05031	0,02977	0,02657
2000	0,06691	0,07819	0,05486	0,05717	0,03492	0,02241
2001	0,05542	0,06153	0,04579	0,05198	0,03070	0,01543
2002	0,01565	0,01890	0,01454	0,02028	0,01165	0,00776
2003	0,02486	0,02901	0,02572	0,04101	0,01920	0,01263
2004	0,07183	0,06437	0,07242	0,12509	0,05072	0,04467
2005	0,06434	0,05521	0,07644	0,14251	0,05149	0,09009
2006	0,02828	0,02607	0,04081	0,06629	0,02467	0,01051
TOPLAM	1	1	1	1	1	1

* 1979 ve önceki modelleri kapsar.

(TÜİK MOTORLU KARA TAŞITLARI İSTATİSTİKLERİ, 2005)

EK-2. 2006 yılı için araç türüne göre araç modelleri olasılıkları

2006 yılı						
Model Yılı	Otomobil	Minibüs	Otobüs	Kamyonet	Kamyon	Diğer
1980	0,06682	0,09365	0,10870	0,11175	0,20768	0,23325
1981	0,00330	0,00444	0,00538	0,00308	0,01614	0,01497
1982	0,00558	0,00755	0,00964	0,00363	0,02086	0,01868
1983	0,00694	0,01064	0,01892	0,00540	0,02173	0,02117
1984	0,00912	0,01294	0,01805	0,00642	0,02250	0,02646
1985	0,01128	0,01431	0,01709	0,00773	0,02469	0,02241
1986	0,01487	0,01614	0,01905	0,00740	0,02233	0,01763
1987	0,01871	0,01715	0,01566	0,00642	0,01768	0,02192
1988	0,02135	0,01611	0,01640	0,00555	0,01985	0,02068
1989	0,01908	0,01336	0,01780	0,00534	0,01529	0,01614
1990	0,03437	0,02103	0,02949	0,00922	0,02289	0,02749
1991	0,03212	0,02083	0,02916	0,01050	0,01958	0,02776
1992	0,04666	0,03262	0,03916	0,01473	0,02874	0,02713
1993	0,06481	0,04103	0,04892	0,02458	0,03846	0,03670
1994	0,05598	0,03177	0,03075	0,02121	0,02651	0,02250
1995	0,03382	0,02205	0,01868	0,01017	0,02085	0,01947
1996	0,03817	0,03799	0,02878	0,02381	0,03547	0,02675
1997	0,04629	0,06197	0,04612	0,04610	0,05484	0,03224
1998	0,05843	0,07920	0,05004	0,06324	0,06232	0,03114
1999	0,04239	0,06341	0,03690	0,04359	0,02839	0,02307
2000	0,06292	0,07577	0,05028	0,04940	0,03333	0,01953
2001	0,05208	0,06025	0,04165	0,04501	0,02925	0,01346
2002	0,01473	0,01871	0,01332	0,01765	0,01111	0,00679
2003	0,02354	0,02911	0,02413	0,03573	0,01848	0,01110
2004	0,06804	0,06491	0,06772	0,10883	0,04862	0,03977
2005	0,06219	0,05602	0,07302	0,12657	0,05012	0,09231
2006	0,06934	0,05770	0,09428	0,14320	0,06343	0,12237
2007	0,01707	0,01934	0,03092	0,04375	0,01887	0,00712
TOPLAM	1	1	1	1	1	1

* 1980 ve önceki modelleri kapsar.

(TÜİK MOTORLU KARA TAŞITLARI İSTATİSTİKLERİ, 2006)

EK-3. 2007 yılı için araç türüne göre araç modelleri olasılıkları

2007 yılı						
Model Yılı	Otomobil	Minibüs	Otobüs	Kamyonet	Kamyon	Diğer
1981	0,06585	0,09088	0,10101	0,10390	0,20498	0,23091
1982	0,00526	0,00694	0,00844	0,00323	0,01911	0,01736
1983	0,00655	0,00960	0,01676	0,00482	0,02012	0,01970
1984	0,00861	0,01171	0,01591	0,00575	0,02087	0,02463
1985	0,01065	0,01296	0,01519	0,00693	0,02322	0,02085
1986	0,01403	0,01465	0,01721	0,00667	0,02202	0,01642
1987	0,01767	0,01561	0,01409	0,00582	0,01694	0,02041
1988	0,02019	0,01436	0,01476	0,00433	0,01903	0,01926
1989	0,01806	0,01207	0,01607	0,00397	0,01470	0,01499
1990	0,03253	0,01890	0,02658	0,00757	0,02197	0,02551
1991	0,03039	0,01900	0,02619	0,00954	0,01887	0,02569
1992	0,04420	0,02961	0,03508	0,01341	0,02773	0,02518
1993	0,06139	0,03776	0,04376	0,02227	0,03715	0,03409
1994	0,05304	0,02991	0,02765	0,01911	0,02565	0,02097
1995	0,03206	0,02096	0,01679	0,00917	0,02015	0,01820
1996	0,03620	0,03639	0,02616	0,02136	0,03434	0,02503
1997	0,04390	0,05973	0,04202	0,04130	0,05313	0,03015
1998	0,05544	0,07631	0,04581	0,05662	0,06042	0,02918
1999	0,04021	0,06135	0,03415	0,03896	0,02756	0,02162
2000	0,05969	0,07398	0,04592	0,04409	0,03235	0,01831
2001	0,04941	0,05910	0,03825	0,04013	0,02835	0,01265
2002	0,01399	0,01860	0,01225	0,01577	0,01079	0,00639
2003	0,02242	0,02906	0,02240	0,03201	0,01807	0,01046
2004	0,06480	0,06524	0,06376	0,09739	0,04732	0,03759
2005	0,05948	0,05631	0,06900	0,11394	0,04895	0,08861
2006	0,06682	0,05829	0,08974	0,12967	0,06260	0,13661
2007	0,04525	0,04205	0,07399	0,09828	0,04640	0,04190
2008	0,02195	0,01867	0,04107	0,04399	0,01721	0,00733
TOPLAM	1	1	1	1	1	1

* 1981 ve önceki modelleri kapsar.

(TÜİK MOTORLU KARA TAŞITLARI İSTATİSTİKLERİ, 2007)

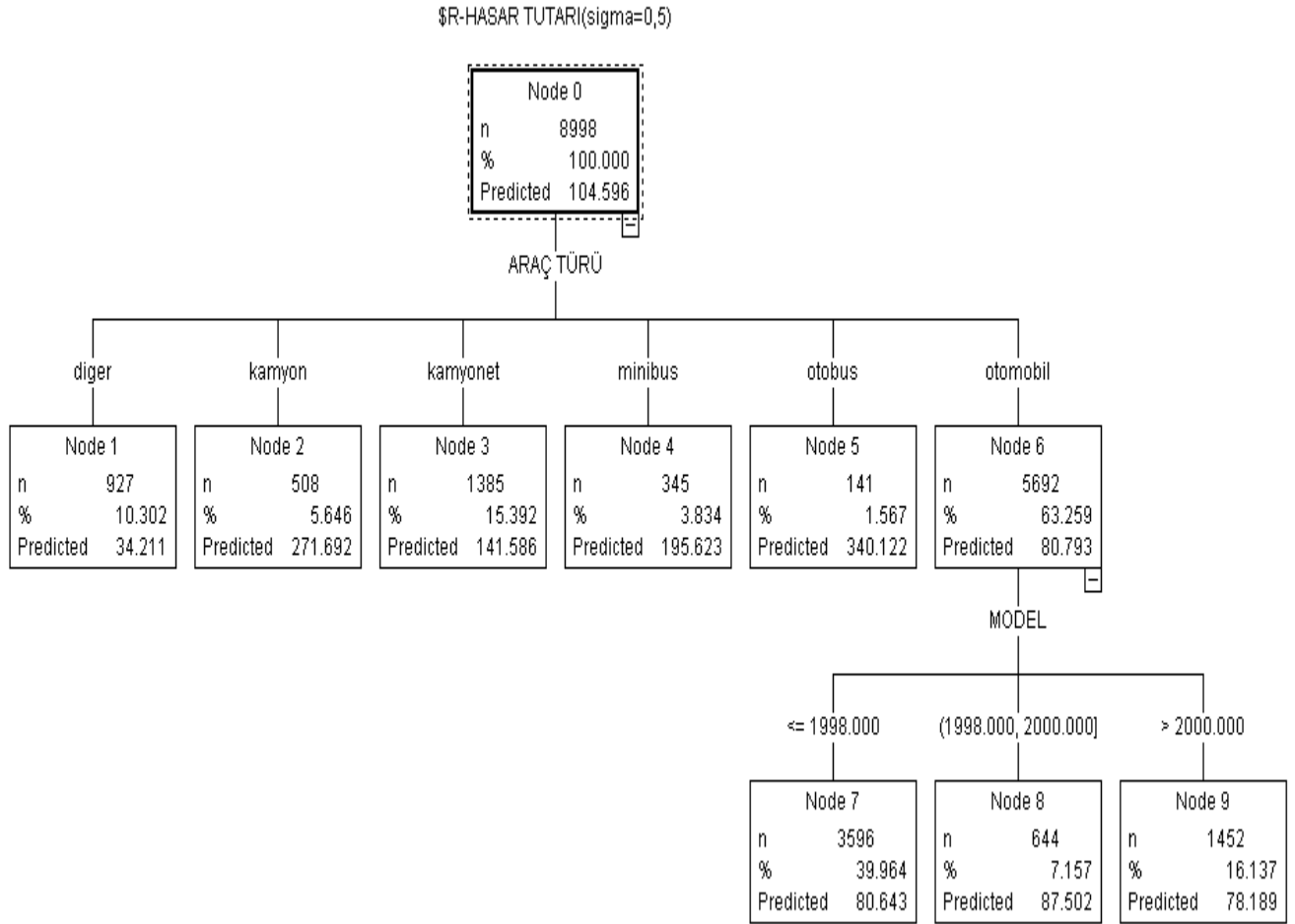
EK-4. 2008 yılı için araç türüne göre araç modelleri olasılıkları

2008 yılı						
Model Yılı	Otomobil	Minibüs	Otobüs	Kamyonet	Kamyon	Diğer
1982	0,06681	0,08972	0,09708	0,09362	0,20904	0,23285
1983	0,00619	0,00854	0,01475	0,00434	0,01891	0,01852
1984	0,00815	0,01046	0,01417	0,00518	0,01967	0,02313
1985	0,01006	0,01156	0,01337	0,00627	0,02195	0,01955
1986	0,01323	0,01314	0,01528	0,00606	0,02091	0,01542
1987	0,01672	0,01395	0,01263	0,00533	0,01621	0,01917
1988	0,01911	0,01281	0,01317	0,00466	0,01835	0,01808
1989	0,01713	0,01097	0,01425	0,00447	0,01425	0,01403
1990	0,03083	0,01715	0,02358	0,00772	0,02134	0,02388
1991	0,02881	0,01717	0,02337	0,00884	0,01826	0,02398
1992	0,04197	0,02691	0,03112	0,01249	0,02692	0,02359
1993	0,05827	0,03469	0,03949	0,02060	0,03609	0,03193
1994	0,05042	0,02827	0,02517	0,01752	0,02493	0,01967
1995	0,03049	0,01996	0,01522	0,00844	0,01963	0,01710
1996	0,03447	0,03502	0,02395	0,01962	0,03340	0,02357
1997	0,04179	0,05793	0,03866	0,03781	0,05174	0,02840
1998	0,05279	0,07451	0,04263	0,05175	0,05890	0,02750
1999	0,03832	0,06005	0,03203	0,03558	0,02693	0,02041
2000	0,05687	0,07268	0,04301	0,04018	0,03166	0,01730
2001	0,04708	0,05838	0,03561	0,03666	0,02760	0,01196
2002	0,01335	0,01853	0,01148	0,01439	0,01052	0,00604
2003	0,02149	0,02947	0,02075	0,02924	0,01757	0,00990
2004	0,06205	0,06627	0,06015	0,08905	0,04612	0,03565
2005	0,05715	0,05701	0,06557	0,10470	0,04787	0,08461
2006	0,06426	0,05901	0,08555	0,11914	0,06131	0,13794
2007	0,04364	0,04337	0,07145	0,09055	0,04558	0,05315
2008	0,05535	0,04175	0,09000	0,10029	0,04245	0,03633
2009	0,01320	0,01072	0,02650	0,02551	0,01186	0,00634
TOPLAM	1	1	1	1	1	1

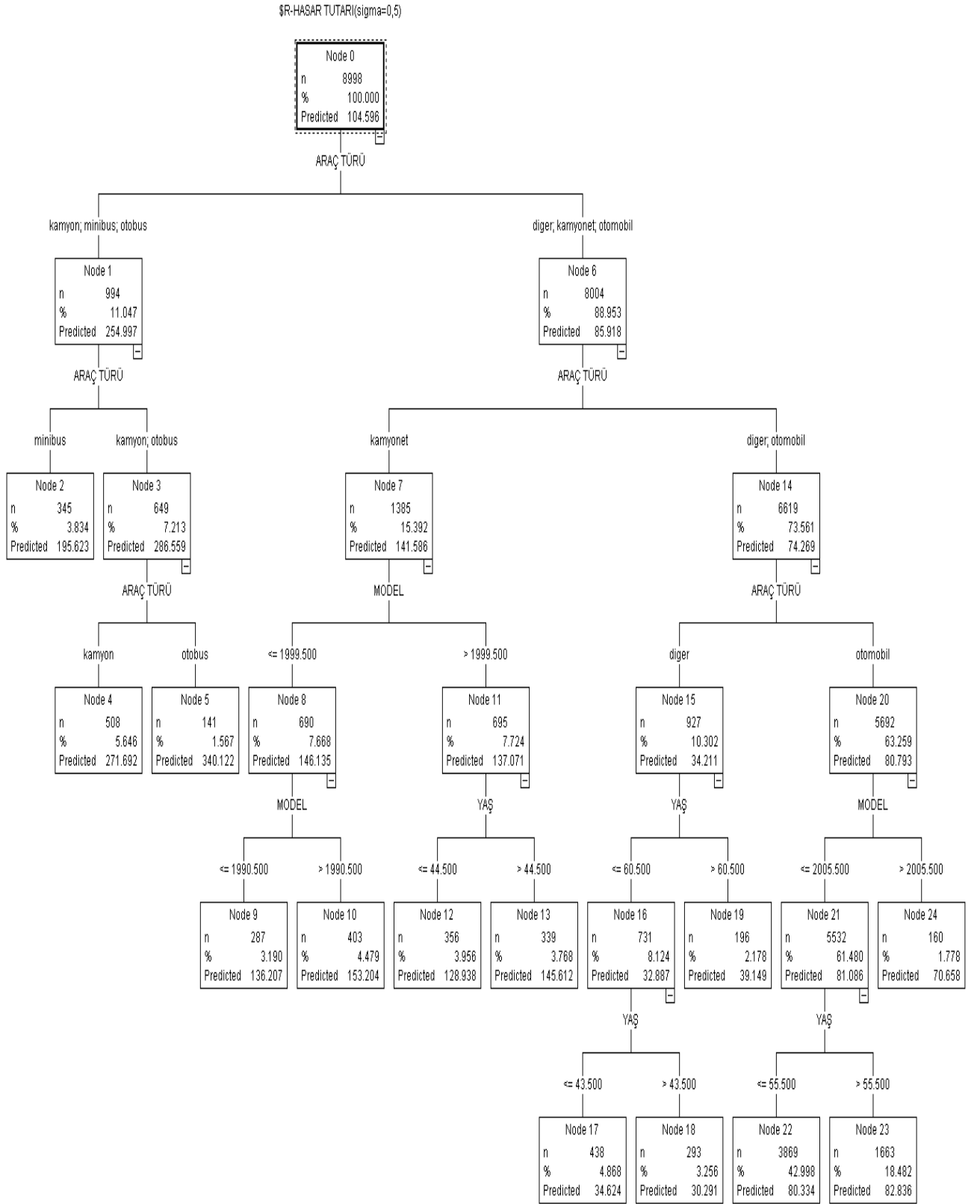
* 1982 ve önceki modelleri kapsar.

(TÜİK MOTORLU KARA TAŞITLARI İSTATİSTİKLERİ, 2008)

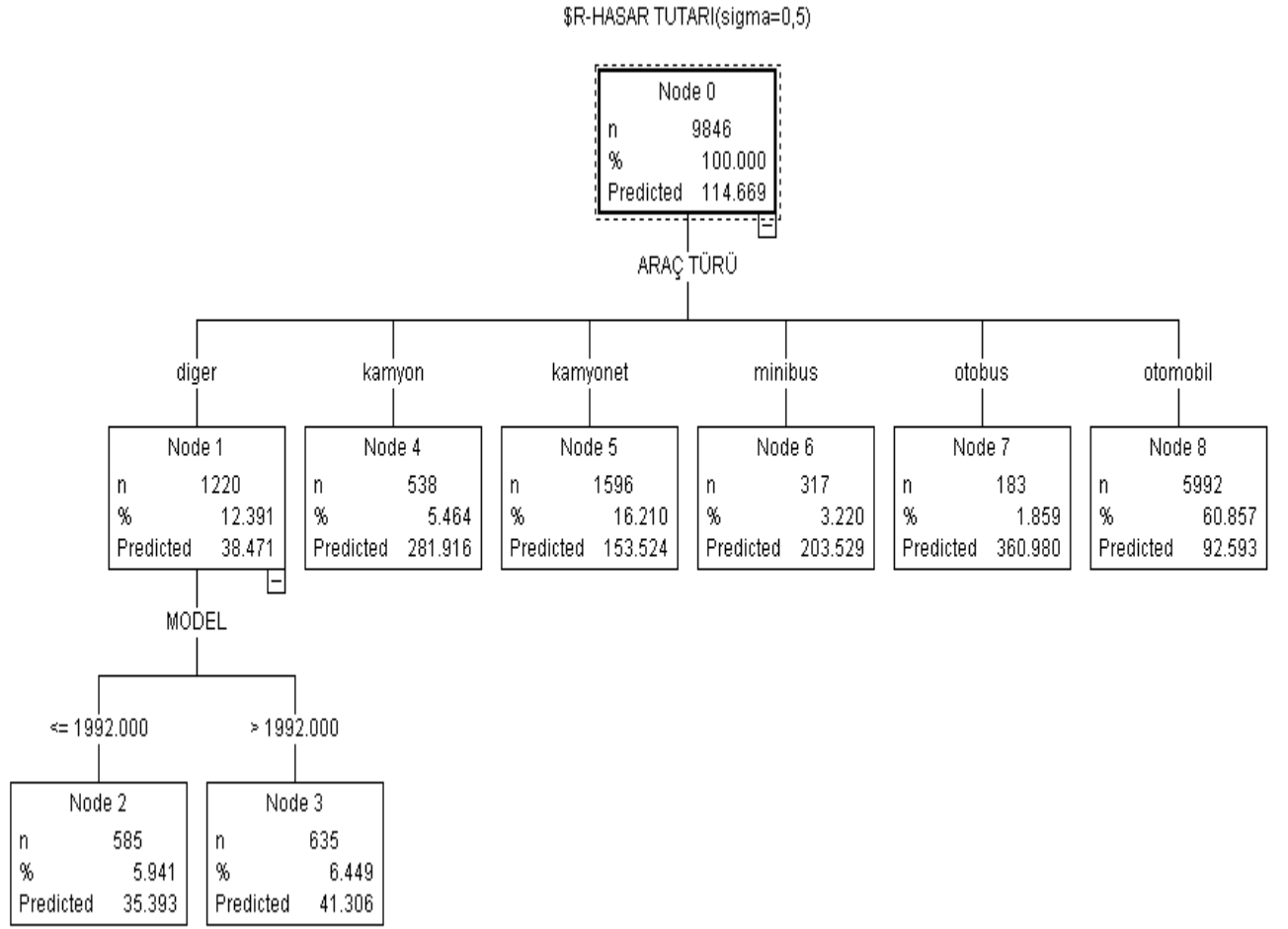
EK-5. 2005 yılı için CHAİD algoritması



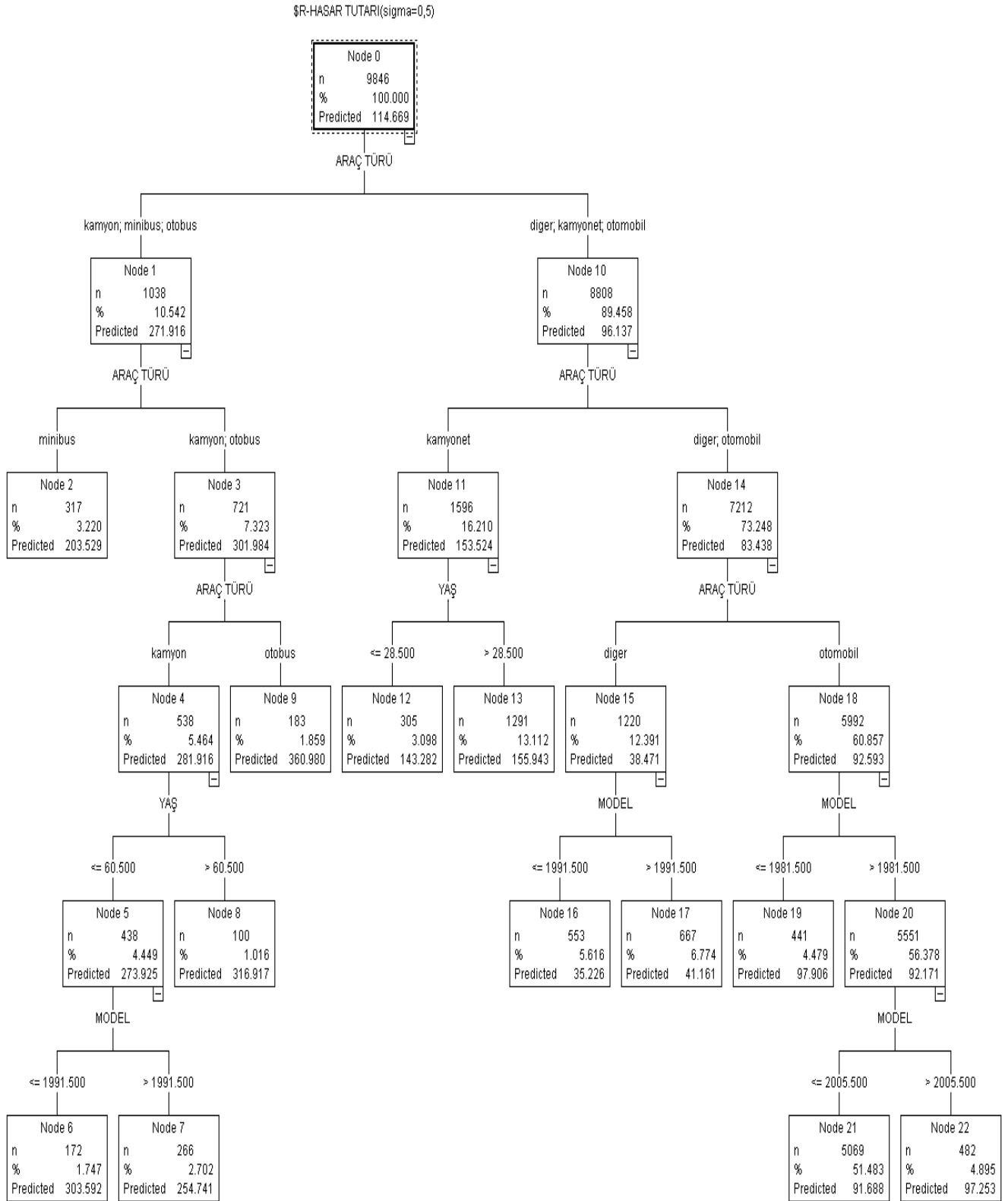
EK-6. 2005 yılı için CRT algoritması



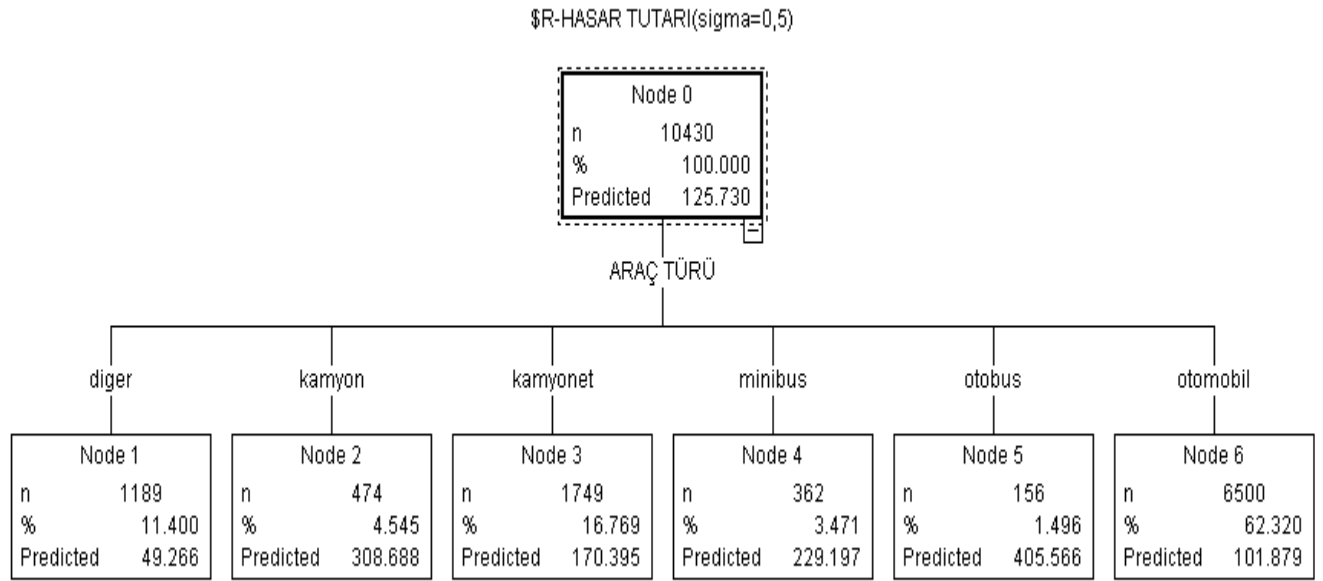
EK-7. 2006 yılı için CHAİD algoritması



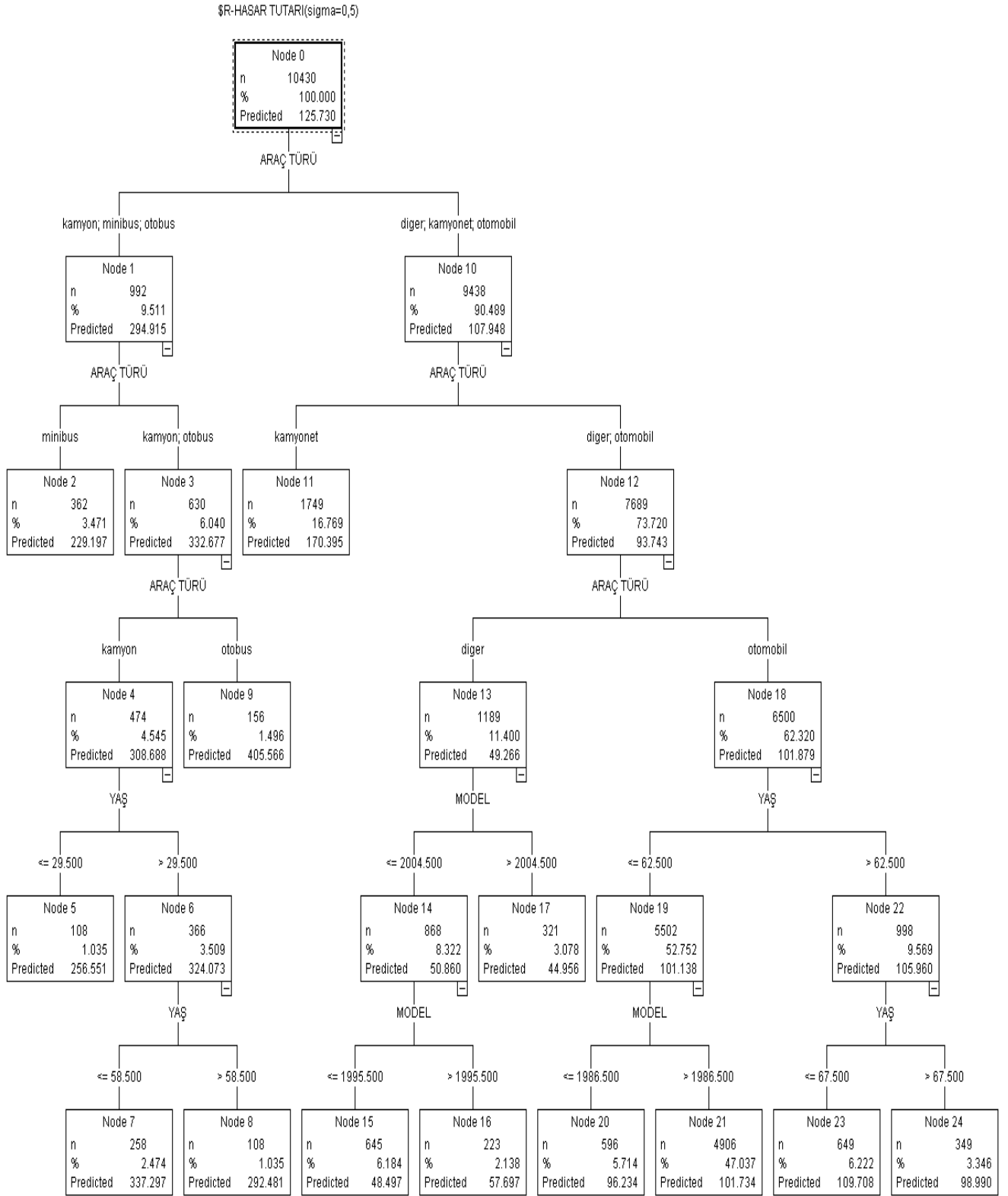
EK-8. 2006 yılı için CRT algoritması



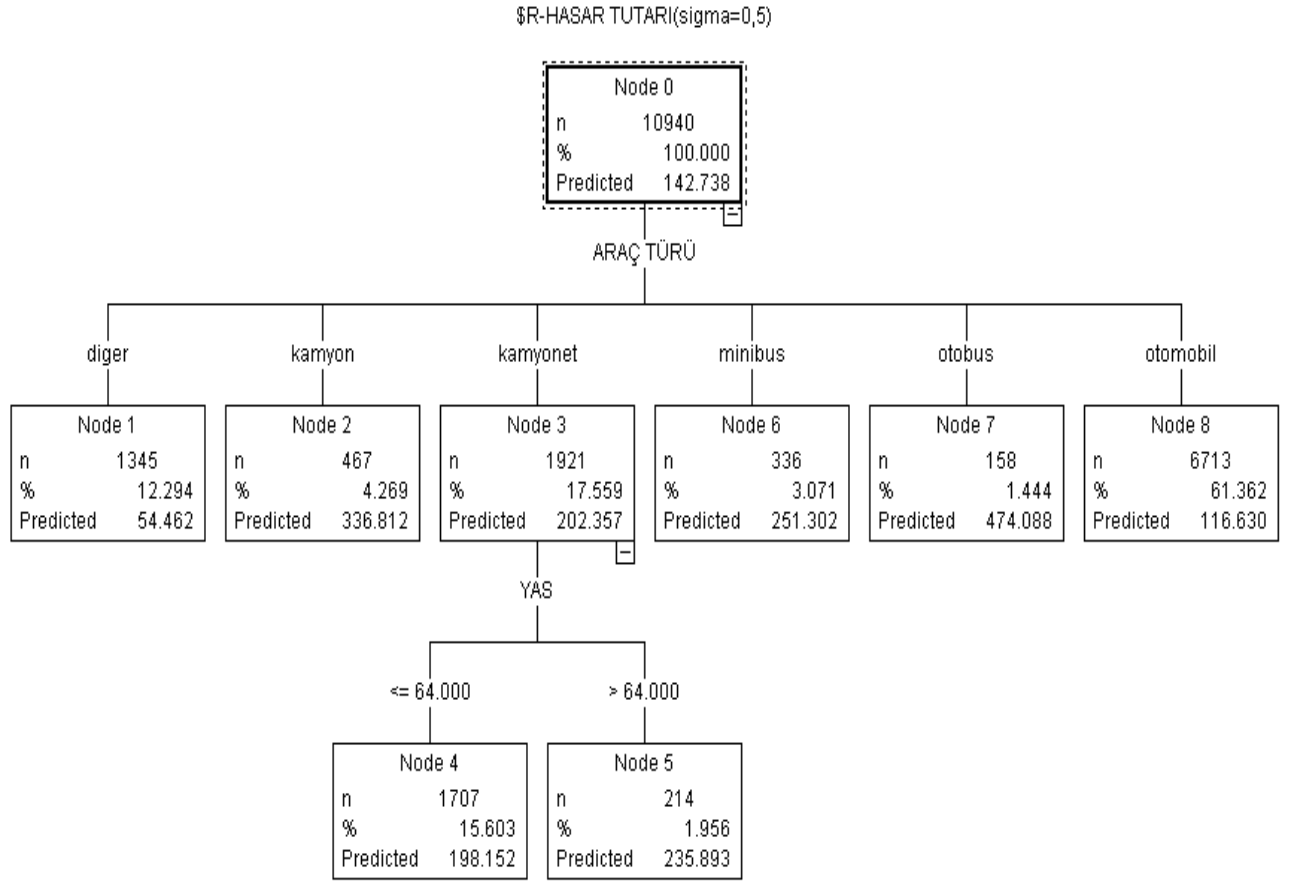
EK-9. 2007 yılı için CHAİD algoritması



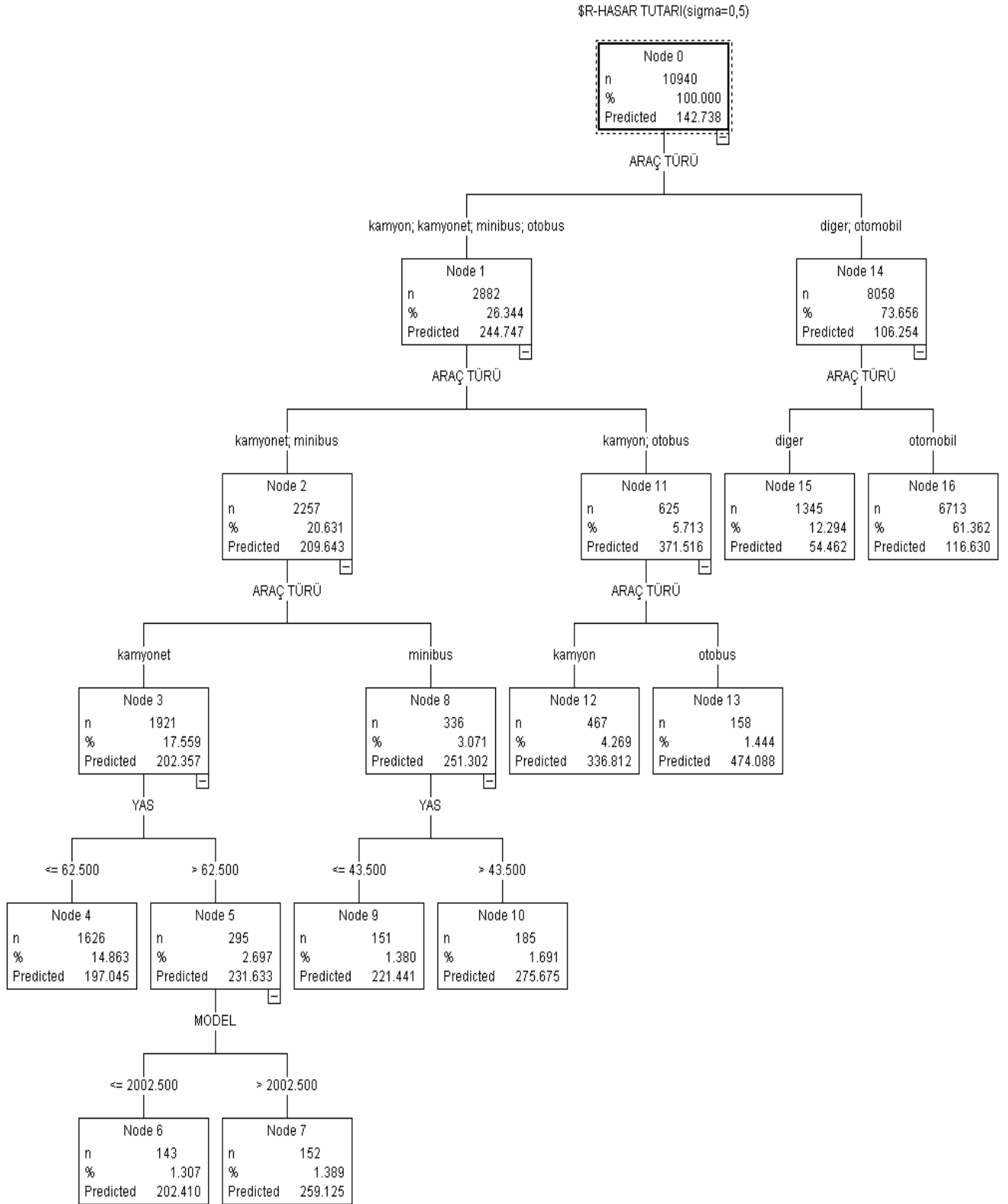
EK-10. 2007 yılı için CRT algoritması



EK-11. 2008 yılı için CHAİD algoritması



EK-12. 2008 yılı için CRT algoritması



**EK-13. R programında Mack Zincir Merdiven Yönteminin kodları ve çıktısı
(Toplam birikimli hasar tutarına göre)**

```
data=read.table("c:/toplaml.txt")
> data
      X0      X1      X2      X3
2005 234208.7 803155.9 906709.2 941152.2
2006 269787.5 995860.6 1129026.6    NA
2007 384768.7 1311366.1    NA    NA
2008 1561552.7    NA    NA    NA
> M<-MackChainLadder(data,est.sigma="Mack")
> M
MackChainLadder(Triangle = data, est.sigma = "Mack")
```

	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
2005	941,152	1.000	941,152	0	0	NaN
2006	1,129,027	0.963	1,171,915	42,888	192	0.00447
2007	1,311,366	0.851	1,540,288	228,922	4,994	0.02181
2008	1,561,553	0.243	6,418,915	4,857,362	206,568	0.04253

```

Totals
Latest      : 4,943,097.60
Dev         : 0.49
Ultimate    : 10,072,270.00
IBNR        : 5,129,172.40
Mack S.E.   : 206,841.55
CV(IBNR)    : 0.0403264950854866
```

**EK-14. R programında Mack Zincir Merdiven Yönteminin kodları ve çıktısı
(Birinci risk grubunun birikimli hasar tutarına göre)**

```
data=read.table("c:/birinci.txt")
> data
      X0      X1      X2      X3
2005 70009.15 215526.6 245941.1 253466.6
2006 67983.02 249558.8 282249.0    NA
2007 95015.50 292555.6    NA    NA
2008 316634.62    NA    NA    NA
> M<-MackChainLadder(data,est.sigma="Mack")
> M
MackChainLadder(Triangle = data, est.sigma = "Mack")

      Latest  Dev.To.Date  Ultimate  IBNR  Mack.S.E  CV(IBNR)
2005 253,467    1.000    253,467    0      0      NaN
2006 282,249    0.970    290,886    8,637    100    0.0116
2007 292,556    0.854    342,417    49,862    2,452    0.0492
2008 316,635    0.263    1,205,033    888,398    93,149    0.1049

      Totals
Latest : 1,144,905.84
Dev     : 0.55
Ultimate : 2,091,802.44
IBNR    : 946,896.60
Mack S.E. : 93,269.66
CV(IBNR) : 0.0985003613046203
```

**EK-15. R programında Mack Zincir Merdiven Yönteminin kodları ve çıktısı
(İkinci risk grubunun birikimli hasar tutarına göre)**

```
data=read.table("c:/ikinci.txt")
```

```
> data
```

```
      X0      X1      X2      X3
2005 164199.6 587629.4 660768.2 687685.6
2006 201804.5 746301.8 846777.6    NA
2007 289753.2 1018810.5    NA    NA
2008 1244918.1    NA    NA    NA
```

```
> M<-MackChainLadder(data,est.sigma="Mack")
```

```
> M
```

```
MackChainLadder(Triangle = data, est.sigma = "Mack")
```

	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
2005	687,686	1.000	687,686	0	0	NaN
2006	846,778	0.961	881,272	34,495	1,062	0.0308
2007	1,018,811	0.850	1,198,316	179,505	8,246	0.0459
2008	1,244,918	0.237	5,253,512	4,008,594	103,059	0.0257

```
Totals
Latest      : 3,798,191.76
Dev         : 0.47
Ultimate    : 8,020,786.09
IBNR        : 4,222,594.33
Mack S.E.   : 104,695.12
CV(IBNR)    : 0.0247940280256542
```

ÖZGEÇMİŞ

Adı Soyadı : Damla BARLAS

Doğum Yeri : Ankara

Doğum Yılı : 1985

Medeni Hali : Bekar

Eğitim ve Akademik Durumu:

Lise : 2000-2003 Mamak Anadolu Lisesi

Lisans : 2003-2007 Hacettepe Üniversitesi İstatistik Bölümü

Yabancı Dil : İngilizce, Almanca

İş Tecrübesi:

Aralık, 2007 - ... Hacettepe Üniversitesi Aktüerya Bilimleri Bölümü,
Araştırma Görevlisi