



**HAYVANCILIKTA ELDE EDİLEN FARKLI VERİLERE
REGRESYON AĞACI METODUNUN UYGULANMASI**

Yusuf KOÇ

Yüksek Lisans Tezi

ZOOTEKNİ ANABİLİM DALI

1.Danışman: Doç. Dr. Ecevit EYDURAN

2.Danışman: Prof. Dr. Ömer AKBULUT

2016

Her hakkı saklıdır

FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**HAYVANCILIKTA ELDE EDİLEN FARKLI VERİLERE REGRESYON
AĞACI METODUNUN UYGULANMASI**

Yusuf KOÇ

ZOOTEKNİ ANABİLİM DALI

İĞDIR

2016

Her hakkı saklıdır

Doç. Dr. Ecevit EYDURAN ve Prof. Dr. Ömer AKBULUT danışmanlığında Yusuf KOÇ tarafından hazırlanan bu çalışma 18 Nisan 2016 tarihinde aşağıdaki jüri üyeleri tarafından Zootekni Anabilim Dalı'nda Yüksek lisans tezi olarak kabul edilmiştir.

Başkan: Doç. Dr. Ecevit EYDURAN

İmza:

Üye: Yrd. Doç. Dr. Şenol ÇELİK

İmza:

Üye: Yrd. Doç. Dr. Mehmet Kazım KARA

İmza:

Üye: Yrd. Doç. Dr. İsa YILMAZ

İmza:

Üye: Yrd. Doç. Dr. Yakup Erdal ERTÜRK

İmza:

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun / /2016 tarih ve sayılı kararı ile onaylanmıştır.

(İmza)

.....

Prof. Dr. Bünyamin YILDIRIM

Enstitü Müdürü

ÖZET

HAYVANCILIKTA ELDE EDİLEN FARKLI VERİLERE REGRESYON AĞACI METODUNUN UYGULANMASI

KOÇ, Yusuf

Yüksek Lisans Tezi, Zootekni Anabilim Dalı

1. Tez Danışmanı: Doç. Dr. Ecevit EYDURAN

2. Tez Danışmanı: Prof. Dr. Ömer AKBULUT

Nisan 2016, 58 Sayfa

Bu çalışmanın amacı, hayvancılıkla ilgili bir veri setinde, farklı ebeveyn-yavru düğüm kombinasyonları için, CHAID, Exhaustive CHAID ve CART veri madenciliği algoritmalarının tahmin performanslarını değerlendirmektir. Bu amaca ulaşmak için, cinsiyet (erkek ve dişi), doğum tipi (tekiz ve ikiz), doğum yılı (2005, 2006, 2007, 2008 ve 2009) işletme (Araştırma istasyonu, Mastung, Quetta, ve Noshki), ana yaşı (20-78 ay), ve ana ağırlığından (25-48 kg) sütten kesim ağırlığını tahmin etmek için 1884 baş Mengali koyunu kullanılmıştır. En iyi karar ağacı algoritmasını seçmek için belirleme katsayısı (R^2 %), düzeltilmiş belirleme katsayısı (VK%), varyasyon katsayısı R^2 -düz. (%), standart sapma oranı, nispi yaklaşık hata (NYH) regresyon, hata kareler ortalamasının karekökü (HKOK) ve gerçek ve tahmin edilen sütten kesim ağırlıkları arasındaki Pearson korelasyon katsayısı hesaplanmıştır. Exhaustive CHAID ve CART veri madenciliği algoritmaları ile karşılaştırıldığında CHAID algoritmasının ve biyolojik konularda daha uygun ağaç yapısı oluşturduğu belirlenmiştir. Sonuç olarak, model kalite ölçütlerinin tahmin edilmesi ile birlikte oluşturulan ağaç yapısının biyolojik uygunluğunun göz önüne alınması tavsiye edilmektedir.

Anahtar Kelimeler: CART, CHAID, Exhaustive CHAID, Regresyon ağacı, Sütten kesim ağırlığı.

ABSTRACT

APPLICATION OF REGRESSION TREE METHOD FOR DIFFERENT DATA FROM ANIMAL SCIENCE

KOÇ, Yusuf

Master Thesis, Department of Animal Science

1st Thesis Advisor: Assoc. Prof. Dr. Ecevit EYDURAN

2nd Thesis Advisor: Prof. Dr. Omer AKBULUT

April, 2016, 58 pages

The aim of this study was to evaluate predictive performances of CHAID, Exhaustive CHAID, and CART regression tree methods for different combinations of parent node: child node in the data set regarding animal science. To achieve the aim, 1884 Mengali lambs were provided for predicting weaning weight from sex (male and female), birth type (single and twins), birth year (2005, 2006, 2007, 2008 and 2009), farm (Research station, Mastung, Quetta, and Noshki), birth weight, dam age, and dam weight. To choose the best regression tree method, regression model quality criteria such as coefficient of determination ($R^2\%$), adjusted coefficient of determination (Adj- $R^2\%$), coefficient of variation (%), SD ratio, relative approximation error (RAE), Root Mean Square Error (RMSE), Pearson correlation between actual and predicted weaning weights were estimated for each combination. It was determined that CHAID algorithm constructed more suitable tree structures, biologically in comparison to Exhaustive CHAID and CART data mining algorithms. Consequently, it is recommended that the biological suitability of the constructed tree structure should be taken account together with estimating model quality criteria.

Key words: CART, CHAID, Exhaustive CHAID, Regression tree, Weaning weight

ÖNSÖZ ve TEŞEKKÜR

Veri madenciliği diğer sınıflama yöntemleri ile karşılaştırıldığında çok sayıda avantaja sahip olduğu görülmektedir. Veri madenciliği analizinde kullanılacak değişkenler çok değişik sayısal değişkenler olabilir. Bununla birlikte, sınıflayıcı veya sıralayıcı yapıya sahip kategorik değişkenler de olabilir. Diğer çok değişkenli modelleme yöntemleri araştırmacılara daha fazla data gerektirirken, veri madenciliği yöntemi araştırmacılara daha az girdi gerektirmektedir. Ayrıca, veri madenciliği ile elde edilen sonuçların yorumlanmasının daha kolay olduğunu söylemek mümkündür. Bu avantajlar sayesinde veri madenciliği ile hayvancılık alanından elde edilen verilerin regresyon ağacı yöntemiyle değerlendirilmesi ve analiz sonuçlarının yorumlanması metotlarının kullanılması, hayvansal verilerin değerlendirilmesine farklı bir boyut kazandırılacaktır. Bununla birlikte bu tarz çalışmalarda diğer çalışmalardan farklı olarak farklı algoritmalar karşılaştırılmış ve uygun algoritmalar belirlenmiştir.

Öncelikle verilerini kullanmamıza izin veren Prof.Dr. Mohammad Masood TARIQ ve Dr. Abdul WAHEED'e çok teşekkür ederim. Bu tezin yazım ve hazırlama aşamasında bana yol gösteren, tecrübe ve önerileri ile beni destekleyen, maddi ve manevi daha doğruya ulaşmama yardımcı olan saygıdeğer danışman hocalarım Doç. Dr. Ecevit EYDURAN ve Prof. Dr. Ömer AKBULUT'a, Zootekni bölümü öğretim üyelerinden Doç. Dr. Yaşar ŞEKER'e ve Yrd. Doç. Dr. İsa YILMAZ' a sonsuz teşekkürlerimi sunarım.

Nisan – 2016

İÇİNDEKİLER

	Sayfa No
ÖZET.....	i
ABSTRACT	ii
ÖNSÖZ ve TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
SİMGELER ve KISALTMALAR DİZİNİ	vi
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	viii
1. GİRİŞ	1
2. KAYNAK ÖZETLERİ	3
3. MATERYAL ve METOT	17
3.1. Materyal	17
3.2. Metot	17
3.2.1. Veri madenciliği algoritmaları	20
3.2.1. 1. CART algoritması	20
3.2.1. 2. CHAID ve Geniş (Exhaustive) CHAID algoritmaları.....	22
3.2.1.2.1.CHAID algoritması oluşum aşamaları.....	23
3.2.2. Model kalite kriterleri.....	25
4. BULGULAR ve TARTIŞMA	27
4. 1. Sütten Kesim Ağırlığına İlişkin Hesaplanan Uyum İyiliği Ölçütlerine Ait Sonuçlar.....	27
4.1.1.CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait bulgular.....	27

4.1.2. Geniş (Exhaustive) CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait bulgular.....	29
4.1.3.CART algoritmasına ilişkin bulgular.....	31
4.2 . CHAID Algoritmasına İlişkin Bulgular.....	34
5. SONUÇ ve ÖNERİLER.....	49
KAYNAKLAR.....	54
ÖZGEÇMİŞ.....	59



SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

%	Yüzde
cm	Santimetre
g	Gram
kg.....	Kilogram
mm.....	milimetre

Kısaltmalar

CA.....	Canlı ağırlık
CHAID.....	Otomatik Ki-Kare Etkileşim Belirleme Analizi
DA.....	Doğum ağırlığı
HKOK.....	Hata kareler ortalaması karekökü
NHY.....	Nisbi yaklaşık hata
R^2	Belirleme katsayısı
R^2 -düz. (%).....	Düzeltilmiş belirleme katsayısı
r	Korelasyon katsayısı
RA.....	Regresyon ağaçları
ROH.....	Relative ortalama hata
RT.....	Risk tahmini
RTD.....	Risk tahmin değeri
SA.....	Sınıflandırma Ağacı
SRA.....	Sınıflandırma ve Regresyon Ağaçları
SS-ORAN.....	Standart sapma oranı
VK (%).....	Varyasyon katsayısı
SKA	Sütten kesim ağırlığı

ŞEKİLLER DİZİNİ

	Sayfa No
Şekil 4.1. 500: 250, 480:240, 460:230, 440:220, 420:210:400:200 ve 380:190 Minimum ebeveyn-yavru düğüm oranları için CHAID algoritmasına ait regresyon ağacı diyagramı.....	35
Şekil 4.2. 300: 150 Minimum ebeveyn-yavru düğüm oranları için CHAID algoritmasına ait regresyon ağacı diyagramı	37
Şekil 4.3. 200: 100 Minimum ebeveyn-yavru düğüm oranı için CHAID algoritmasına ait regresyon ağacı diyagramı	41
Şekil 4.4. 100: 50 Minimum ebeveyn-yavru düğüm oranı için CHAID algoritmasına ait regresyon ağacı diyagramı	46

ÇİZELGELER DİZİNİ

	Sayfa No
Çizelge 4.1. CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar.....	28
Çizelge 4.2. Geniş (Exhaustive) CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar.....	30
Çizelge 4.3. CART algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar.....	31
Çizelge 4.4. Sütten kesim ağırlığı için CHAID algoritması sonuçları.....	32
Çizelge 4.5. Sütten kesim ağırlığı için Geniş (Exhaustive) CHAID algoritması sonuçları.....	33
Çizelge 4.6. Sütten kesim ağırlığı için Pruning CART algoritması sonuçları	33

1.GİRİŞ

Hayvan ıslahında et, süt ve yumurta gibi ekonomik öneme sahip olan verim özellikleri ile vücut özellikleri arasındaki ilişkiler önem arz etmektedir. Öte yandan, hayvancılıkta üzerinde durulan verim özelliğini etkileyen çevre faktörlerinin saptanması ve verim özelliği ile diğer kantitatif özellikler arasındaki sebep-sonuç ilişkilerinin belirlenmesi önemlidir. Farklı vücut ve testis ölçülerinden canlı ağırlığının tahmin edilmesi ve farklı meme özelliklerinden yola çıkarak süt veriminin tahmin edilmesi, sebep sonuç ilişkilerine örnek olarak verilebilir.

Hayvancılıkta sebep-sonuç ilişkileri; basit doğrusal regresyon analizi (*simple linear regression analysis*), çoklu doğrusal regresyon analizi (*multiple linear regression analysis*), Ridge regresyon analizi, faktör analiz skorları ile çoklu regresyon analizi (*factor analysis scores in multiple regression analysis*), temel bileşen analiz skorları ile çoklu regresyon analizi (*principal component analysis scores in multiple regression analysis*), Path Analizi (*Path Analysis*) ve regresyon ağacı analiz (*Regression Tree Analysis*) yöntemleri ile belirlenmektedir (Khan *et al.*, 2014). Ancak, ekonomik öneme sahip kantitatif verim özelliklerini etkileyen faktörlerin etki düzeylerinin belirlenmesinde ise Genel Doğrusal Modeller (*General Linear Models*) oldukça yaygın olarak kullanılmaktadır (Eyduran *et al.*, 2008).

Hayvancılıkla ilgili veriler için kullanılan analiz yöntemlerinden regresyon ağacı analizi yöntemi, yukarıda bahsedilen diğer analiz yöntemlerine alternatif olarak kullanılabilir (Eyduran *et al.*, 2008). Regresyon ağacı yöntemi, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi ve etkileşimi tespit eden, popülasyonu homojen alt sınıflara ayıran non-parametrik bir analiz yöntemidir (Kayrı ve Boysan, 2008). Regresyon ağacı yöntemi, bağımsız değişkenlerin dağılımına ilişkin herhangi bir varsayım gerektirmemesi, çoklu bağlantılılık (*Multicollinearity*), sapan değerler (*outliers*) ve kayıp gözlemlerden (*missing observation*) etkilenmemesi gibi avantajlardan dolayı tercih edilmektedir (Mendes ve Akkartal, 2009).

Regresyon ağacı oluşturmada yaygın olarak kullanılan CART, CHAID ve Geniş (Exhaustive) CHAID algoritmaları, nominal, ordinal ve sürekli değişkenlerin analizi için kullanılan non-parametrik metotlardır (Ali *et al.*, 2005). Bağımlı değişkenin sürekli olmasıyla oluşturulan ağaca regresyon ağacı denir. Yukarıda bahsedilen algoritmaların uygulandığı farklı kaynaklarda yapılan analiz, regresyon ağacı metodu ya da regresyon ağacı analizi olarak kullanılmıştır (Eyduran *et al.*, 2008).

Regresyon ağacı yöntemi; tıp, mühendislik ve endüstri alanlarında çok kullanılan bir non-parametrik yöntemdir. Ancak, hayvancılık alanında regresyon ağacı uygulamalarına ilişkin çalışmaların sayısı az olmakla birlikte gün geçtikçe artmaktadır (Eyduran *et al.*, 2008; Mendes ve Akkartal, 2009; Bakır *et al.*, 2010; Çak *et al.*, 2013; Eyduran *et al.*, 2013a; Eyduran *et al.*, 2013b; Yılmaz *et al.*, 2013). Bununla birlikte, hayvancılık ile ilgili veri setleri için tahmin performansı bakımından regresyon ağacı oluşturmak için kullanılan CART (Yakubu, 2009), CHAID ve Exhaustive CHAID (Khan *et al.*, 2009) gibi farklı veri madenciliği algoritmalarının karşılaştırılmasıyla ilgili çalışmaların sayısı son derece azdır (Ali *et al.*, 2015). Söz konusu bu algoritmalarla oluşturulan regresyon ağacı sonuçları ileride yapılacak seleksiyon çalışmalarında önemli rol oynayacaktır (Ali *et al.*, 2015).

Bu çalışmada, hayvancılık alanında elde edilen bir veri seti üzerinde regresyon ağacının analiz sonuçlarının nasıl yorumlandığından bahsedilecektir. Hayvancılık ile ilgili veri setleri için algoritmalar (CART, CHAID ve Geniş (Exhaustive) CHAID) karşılaştırılarak uygun algoritma belirlenmiştir. Bununla birlikte atılan farklı minimum ebeveyn-yavru düğüm oranlarını incelenmiş uygun ebeveyn düğüm: yavru düğüm oranı belirlenmeye çalışılmıştır. Ayrıca, hayvancılık alanında regresyon ağacı yönteminin uygulandığı çalışmalara değinilecektir. Bu çalışma ile birbiriyle benzerlik gösteren bireylerden oluşan alt grupların tespit edilmesi amaçlanmıştır.

2. KAYNAK ÖZETLERİ

Doğan (2003), üzerinde durulan özelliği etkileyen faktörlerin hangi seviyesinde etkinin yüksek olduğunu belirlemek amacıyla kullanılan CHAID algoritmasının, bir popülasyonu, bağımlı değişkendeki varyasyonu gruplar içi minimum ve gruplar arası maksimum olacak şekilde farklı alt gruplara veya bölümlere tekrarlı olarak ayıran bir teknik olduğundan bahsetmiştir. Bu çalışma ile 1985-1992 yılları arasında Bala Tarım İşletmesi'nde yetiştirilen Holstein ırkı ineğe ait 2xEÇx305 (açıklama) 440 baş örnekten elde edilen; Ggünlük süt verimi bağımlı değişken; Laktasyon süresi, kuru süre, ilk sıfat yaşı, iki buzağılama arası geçen süre, servis sayısı, servis periyodu, buzağılama mevsimi, gebelik süresi özellikleri de bağımsız değişken olarak alınmıştır. Maksimum süt verimini veren bağımsız değişkenlere ait alt grup kombinasyonları CHAID veri madenciliği algoritması ile belirlenmiştir. Süt verimini en iyi açıklayan ilk alt kategorinin 55-74 günlük kuruda kalma süresi olduğu, 55-74 gün olarak belirlenen kuruda kalma süresi ile en fazla etkileşim içerisinde bulunan alt kategorinin ise 10-16 aylık ilk sıfat yaşının olduğu saptanmıştır.

Grzesiak *et al.* (2003), kısmi laktasyon kayıtlarını kullanarak 305 günlük süt verimini tahmin etmeye çalışmışlardır. Araştırmacılar süt verimini tahmin etmek amacıyla yapay sinir ağları ve çoklu regresyon metotlarını kullanmışlardır. Modellerin karşılaştırılmasında, hata kareler ortalamasının karekökü (HKOK), standart sapma oranı (SS oran), tahmine ilişkin relative ortalama hata (ROH) ve düzeltilmiş belirleme katsayısı (R^2 -düz.(%)) gibi model kalite ölçütleri kullanılmıştır. Çoklu regresyon modeline nazaran sinir ağları modelinin 305 günlük süt verimi gibi kantitatif özelliklerin tahmin edilmesinde alternatif bir metot olabileceği sonucuna varılmıştır.

Caraviello *at al.*, (2006), makine öğrenme algoritmalarını kullanarak süt sığırcılığı yapan büyük işletmelerde yetiştirilen laktasyon dönemindeki Holstein ırkı ineklerin üreme özelliklerini etkileyen faktörleri tanımlamaya çalışmışlar, çoklu bağlantılılık, kayıp gözlem ve değişkenler arasındaki kompleks interaksiyonların olması gibi durumunda söz konusu makine öğrenme algoritmalarının büyük kolaylık sağladığından bahsetmişlerdir.

Araştırmacılar, 103 çiftçi üzerinde bakım ve idare, iş gücü, beslenme, barınak olanakları, üreme, genetik seleksiyon, iklim ve süt verimi ile ilgili bir anket çalışması uygulamışlardır. Bu çalışmada, ilkinde doğurma oranı için 31076 laktasyon kaydı, 14804 inek ve 317 bağımsız değişken kullanılırken laktasyonun 150. günündeki genetik durumu için 17587 laktasyon kaydı, 9516 inek ve 341 bağımsız değişken kullanılmıştır. İlkine doğurma oranı için alternatif bir karar ağacı algoritmasının kayıtları %75.6 oranında doğru sınıfladığı tespit edilmiştir. Bu algoritma ayak tırnağı kesme sıklığı, altlık tipi, inek kısıtlama sistemi, istekli durma süresi gibi bağımsız değişkenlerin önemi belirlenmiştir. Laktasyonun 150. günündeki gebelik durumu için kullanılan alternatif karar ağacı algoritmasının kayıtların %71.4'ünü doğru sınıflandırdığı belirlenmiştir. Gebelik durumu için etkili bağımsız değişkenlerin ise inek başına altlık boşluğu, spermanın erime sıcaklığı, düşük vücut kondüsyon skoruna sahip ineklerin oranı, doğum bölmesinde bulunan ineklerin sayısı, doğumdan sonraki ilk süt verimi ve kullanılan aşım stratejisinin olduğu saptanmıştır.

Grzesiak *et al.* (2006), 305 günlük süt verimini tahmin etmek amacıyla yapay sinir ağları ve Wood metotlarını kullanmışlardır. Metotların uyum iyiliği kalitesini karşılaştırmak için; nispi yaklaşık hata (NYH) ve hata kareler ortalamasının karekökü (HKOK) ve belirleme katsayısını (R^2 (%)) kullanılmıştır. Çalışmanın sonucunda, Wood metodu yerine yapay sinir ağları kullanımının daha doğru olabileceği sonucuna varılmıştır.

Eyduran *et al.* (2008), Akkaraman ırkının varyeteleri olan Norduz (103 baş) ve Karakaş (101 baş) kuzuları üzerinde yaptıkları çalışmada, doğum ağırlığı üzerinde etkili olan faktörleri belirlemek amacıyla CHAID algoritmasına dayalı regresyon ağacı metodunu kullanmışlardır. Regresyon ağacı diyagramı incelendiğinde, bağımlı değişken olan doğum ağırlığını birinci dereceden doğum tipinin etkilediği görülmektedir (Düz.P<0.01). Bununla birlikte doğum ağırlığını ikinci dereceden ırk (Düz.P<0.01) ve cinsiyet (Düz.P<0.01) etkilerken, üçüncü dereceden ana yaşının etkilediği saptanmıştır (Düz.P<0.01). Çalışmada, doğum tipi tekiz olan kuzuların bulunduğu grubu temsil eden Dügüm-1'deki tekiz kuzuların doğum ağırlığı üzerine sadece cinsiyet faktörünün etkili olduğu belirlenmiştir. Tekiz erkek hayvanların (dügüm-3) ortalama doğum ağırlığı 5.065 kg, tekiz ve dişi kuzularda (dügüm-4)

ortalama doğum ağırlığı 4.609 kg olarak tahmin edilmiştir. Dügüm-3 ve Dügüm-4'ü oluşturan tekiz-erkek ve tekiz-dişi kuzu gruplarında yeterli homojeniteye ulaşıldığından dolayı bölünme işlemi durmuştur. İkiz doğan Norduz kuzularının (Dügüm-6) doğum ağırlığı üzerine ana yaşı faktörünün etkili olduğu görülmüştür (Düz.P<0.01). Dügüm-6 (ikiz ve Norduz kuzuları) ana yaşı bakımından iki yavru düğüme bölünmüştür (Dügüm-7 ve Dügüm-8). Dügüm-7; 4 ve 5 yaşlı analarda doğan Norduz ikiz kuzularının oluşturduğu grubu temsil etmekle birlikte ortalama doğum ağırlığı 4.475 kg bulunmuştur. Sonuç olarak 4 ve 5 yaşlı analardan doğan Norduz ikiz kuzuların doğum ağırlığı, 2, 3 ve 6 yaşlı analardan doğan Norduz ikiz kuzulardan daha ağır olduğu tespit edilmiştir.

Bakır *et al.* (2009), Brown Swiss ırkı sığırlarda ergin çağ süt verimi ile 9 bağımsız değişken (ineğin yaşı, ilk tohumlama yaşı, laktasyon sırası, laktasyon süresi, kuruda kalma süresi, ilk buzağılama yaşı, buzağılama mevsimi, doğum tipi ve cinsiyet) arasındaki ilişkiyi CHAID algoritması kullanarak değerlendirmişlerdir. Araştırmacılar, Muş devlet üretme çiftliğinde yetiştirilen Brown-Swiss ırkı hayvanlara ait 770 kayıt kullanmışlardır. CHAID sonuçları incelendiğinde laktasyon süresinin ergin çağ süt verimi üzerinde birinci derecede etkili olduğu, cinsiyet ve ineğin yaşının ikinci derecede etkili olduğu ve ilk tohumlama yaşı ve kuruda kalma süresinin ise ergin süt verimini etkileyen önemli değişkenlerden olduğu belirlenmiştir (P<0.001).

Mendeş *et al.* (2009), 224 adet Ross-308 etçi piliçlerin kesim ağırlığını etkileyen değişkenlerin belirlenmesinde, regresyon ağacı analiz metodunu kullanmışlardır. Araştırmacılar, kesim ağırlığını tahmin etmek için ikinci hafta canlı ağırlığı, bacak uzunluğu, bacak genişliği, göğüs kemiği uzunluğu, göğüs genişliği, göğüs çevresi ve vücut uzunluğu özelliklerini kullanmışlardır. Analiz sonucunda incelenen bu sekiz bağımsız değişkenden sadece dördünün (canlı ağırlık, göğüs kemiği uzunluğu, bacak genişliği ve göğüs çevresi) kesim ağırlığı üzerinde etkili olduğu belirtilmiştir. Sonuçlara göre yüksek kesim ağırlıklarının; 2. hafta canlı ağırlığı 295.95 g'dan daha ağır, göğüs kemiği uzunluğu 55.82 mm'den daha uzun ve göğüs çevresi 14.18 cm'den daha geniş olan Ross-308 etçi piliçlerden ya da 2. hafta canlı ağırlığı 295.95 g'a eşit ve 295.95 g'dan daha hafif ve göğüs kemiği

uzunluđu 60.26 mm'den daha uzun ve bacak geniřliđi 8.32 cm'den daha geniř olan piliçlerden elde edildiđi saptanmıřtır.

Nisbet *et al.* (2009), SPSS, STATISTICA ve SAS paket programlarını kullanarak CART, CHAID, Boosted Classifiers ve Regression, MARS, Regresyon ve Sınıflama için Random Forest, Makine Öğrenme (Bayesyan, Destek Vektörleri, En Yakın Komşuluk, Bağımsız Bileşenler gibi ileri veri madenciliđi algoritmalarını detaylı olarak incelemiřlerdir.

Piwczynski (2009), Polonya'nın Pomarze ve Kujavay bölgesinde bulunan on sürüden yařları 2-8 arasında deđiřen 6586 bař Polonya Merinosunu üreme performans indeksi bakımından (reproductive performance index) deđerlendirmiřtir. Bu çalıřmada, çiftleřtirilen koyunlardan elde edilen kuzu sayısına iliřkin meydana gelen varyasyondan sorumlu deđiřkenleri tanımlanmak amacıyla CART sınıflama ađacı tekniđi kullanılmıř ve en önemli deđiřkenlerin ana yařı, sürü, dođum tipi ve koyunların 16. ay canlı ađırlıđı olduđu belirlenmiřtir. Ayrıca, çiftleřtirilen her koyun başına elde edilen kuzu sayısı üzerinde modele alınan faktörlerin (sürü, ana yařı, dođum tipi ve vücut ađırlıđı) etkili olduđu bildirilmiřtir.

Bakır *et al.* (2010), farklı çiftliklerde yetiřtirilen Holstayn ırkı sığırlar üzerinde yaptıkları çalıřmada; kuruda kalma süresi, laktasyon sırası, iřletme, buzađılama mevsimi ve yařı gibi çevre faktörlerinin 305 günlük süt verimi üzerindeki etkilerini CHAID algoritması ile oluřturulan regresyon ađacını kullanarak tespit etmiřlerdir. Çalıřmada kuruda kalma süresi ($P<0.01$) ve laktasyon sırasının ($P<0.01$), 305 günlük süt verimi üzerinde birinci derecede etkili faktörler olduđu iřletme ($P<0.01$) ve buzađılama mevsimi ($P<0.01$) faktörlerinin ikinci derecede etkili faktörler olduđu ve buzađılama yařı faktörünün ($P<0.01$) ise üçüncü derecede etkili olduđu tespit edilmiřtir.

Karabađ *ve ark.* (2010), kınalı keklikler üzerinde yürüttükleri bir arařtırmada, CART algoritması temelinde sınıflandırma ađacı yöntemi kullanılarak çıkıř gücüne etki eden yumurta özelliklerini belirlemiřlerdir. Arařtırmada, dıř yumurta özelliklerinden yumurta ađırlıđı, yumurta hacmi, yumurta uzunluđu ve yumurta geniřliđinin çıkıř üzerinde önemli

etkiye sahip olduđu % 75.6 isabetle tahmin edilmiştir. Çıkış gücü 18.1 g'dan hafif yumurtalar için % 56.0 olarak belirlenirken, 18.1 g'dan ağır yumurtalarda ise % 80 olarak gerçekleşmiştir. Yumurta hacmi ve yumurta genişliği sırasıyla 27.2 ve 3.14 cm'den büyük olduğunda çıkış gücü yaklaşık % 82.1 olmuştur. Buna karşın, yumurta hacmi 27.2'den büyük ve yumurta genişliği 3.14 cm'den küçük olduğunda istenilen düzeyde oluşum meydana gelmediği bildirilmiştir. Yumurta ağırlığı 18.1'den büyük olduğunda çıkış gücü, yumurta hacmi, yumurta uzunluğu ve genişliği tarafından önemli derecede etkilenmiştir.

Tariq *et al.* (2012), Pakistan'da yetiştirilen bazı yerli koyun ırkları (Mengali-48, Balochi-48, Harnai-48, Beverigh-47, Rakshani-48) üzerinde yürüttükleri bir çalışmada CHAID algoritmasına dayalı regresyon ağacı yöntemiyle, cidago yüksekliği, vücut uzunluğu ve göğüs çevresi gibi vücut ölçülerini kullanarak canlı ağırlık tahmini yapmışlardır. Araştırmanın materyali, 11-13 ay arasında değişen 239 baş erkek koyundan oluşmuştur. Regresyon ağacı sonuçlarına göre canlı ağırlıktaki varyasyonun % 72'si cidago yüksekliği ($P<0.001$), vücut uzunluğu ($P<0.001$), göğüs çevresi ($P<0.001$) özellikleri ve ırk faktörü tarafından açıklanmıştır ($P<0.001$). Tüm koyunlar içerisinde göğüs çevresi 89 cm'den büyük olan koyunların en yüksek canlı ağırlığa (48.6 kg) sahip olmuştur. Regresyon ağacı diyagramı incelendiğinde canlı ağırlık üzerinde en etkili değişkenin, göğüs çevresi olduğu belirlenmiştir. Ayrıca yapılan bu çalışmada, göğüs çevresinin artmasıyla birlikte canlı ağırlığında arttığı belirtilmiştir.

Topal *et al.* (2010), İsveç kırmızısı sığırlarda doğum ağırlıkları ve gerçek süt verimini etkileyen faktörlerin etkilerini belirlemek amacıyla CART algoritmasına dayalı regresyon ağacı analiz tekniğini uygulamışlardır. Araştırmacılar, doğum ağırlığını etkileyen en önemli faktörün doğum tipi olduğunu ve bunu doğum mevsimi, cinsiyet faktörlerinin izlediğini ve doğum sırasında analar için hesaplanan vücut kondüsyon puanının da doğum ağırlığı üzerinde etkili olduğunu belirlemişlerdir. Ayrıca araştırmacılar, laktasyon süresi ve pik süt veriminin ise gerçek süt verimini etkileyen önemli değişkenler olduğu belirtmişlerdir.

Grzesiak *et al.* (2011), süt ineklerinde döl tutmayı etkileyen faktörlerin etkisini belirlemek amacıyla Naive Bayes (NBC) ve (CART) algoritmalarını kullanmışlardır.

Araştırmacılar, laktasyon sayısı, suni tohumlama mevsimi, ineğin tohumlama yaşı, HF gen oranı, gebelik oranı, gebelik süresi, süt protein ve yağ verimi ve bir önceki buzağılamadaki cinsiyeti bağımsız değişken olarak kullanmışlardır. Bağımlı değişken, iyi (bir ya da iki suni tohumlamadan sonra bir ineğin gebe kalma durumu) ve kötü (ikiden fazla suni tohumlamadan sonra bir ineğin gebe kalma durumu) döl tutma olmak üzere iki seviyeli (binary) olarak değerlendirilmiştir. Doğru sınıflama oranı (accuracy rate) % 83 bulunmuştur. NBC'ye nazaran CART algoritması zayıf döl tutma sınıfının duyarlık t oranı (sensitivity) tanımlanmasında daha etkili bulunmuştur ($P<0.01$). Özgüllük oranı (Specificity) her iki metot için benzer bulunmuştur. CART algoritmasında döl tutma düzeyini belirleyen değişkenler arasında, buzağılamadan döl tutmaya kadar geçen süre, buzağılama aralığı, ortalama vücut kondüsyonu ile suni tohumlamadaki kondüsyon arasındaki fark değişkenleri çok önemli bulunmuştur. Bu çalışma ile, özellikle CART algoritmasının kullanımının, ineklerin uygun yapay tohumlama zamanının tespitinde yararlı olabileceği sonucuna varılmıştır.

Oruçoğlu (2011), Siyah-Alaca ırkı ineklerin 305 günlük süt verimini etkileyen çevre faktörlerini belirlemek amacıyla CHAID algoritmasını kullanmıştır. Araştırmacı, 2001- 2009 yılları arasında Aydın ili Balta köy sınırları içerisinde Arif Gürdal Tarım İşletmesinde yetiştiriciliği yapılan 198 sağmal ineğe ait toplam 424 adet laktasyon süt verim kaydı kullanılmıştır. Çalışmada CHAID algoritması ile karar ağacı oluşturmak için bağımsız değişken olarak laktasyon sayısı, buzağılama mevsimi, buzağılama yılı ve gebelik başına tohumlama sayısı, ilkine buzağılama yaşı, buzağılama aralığı, kuruda kalma süresi, buzağılama yaşı ve doğum ile ilk kontrol süt tarihine (DİM) kadar geçen süre kullanılmıştır. Minimum ebeveyn: yavru Düğüm oranı 80:40 ya da 60:30 olduğunda, 305 gün süt verimini etkileyen en önemli bağımsız değişkenin DIM olduğu belirlenmiştir ($P<0.001$). Diğer önemli değişkenlerin ise ilkine buzağılama yaşı ($P<0.004$) ve buzağılama aralığı ($P<0.006$) olduğu bildirilmiştir. DIM değeri 26 ve 26'dan daha kısa olan ineklerin süt verimi ilkine buzağılama yaşı tarafından etkilenmiştir. Ayrıca, 26 ve 26 aydan daha kısa olan inekler arasında ilkine buzağılama yaşı 28.19 ay ile 30.90 ay olanların süt verimi üzerinde etkili

olan deęişken sadece buzaęılama aralıęı olmuştur (Düz.P=0.006). En yüksek süt verimi ise (8489.906 kg) DIM deęeri 26 aydan daha uzun ineklerden elde edilmiştir. Ebeveyn: yavru düęümleri minimum veri adedi oranı 50:25, 40:20, 30:15, 20:10 ve 10:5 olduęunda 305 gün süt verimini etkileyen en önemli baęımsız deęişkenin DIM olduęu belirlenmiştir (Düz.P<0.001). Dięer önemli deęişkenlerin ise ilkine buzaęılama yaşı (Düz.P<0.004) ve bunu buzaęı cinsiyeti (Düz.P<0.007), buzaęılama mevsimi (Düz. P<0.002), buzaęılama aralıęı (Düz. P<0.006) olduęu bildirilmiştir. DIM deęeri 26 ve 26 aydan daha kısa olan ineklerin süt verimini ilkine buzaęılama yaşı etkilemiştir. Aynı zamanda, 26 ve 26 ay dan kısa olan inekler arasında ilkine buzaęılama yaşı 26,13 ay dan daha küçük olanların süt verimi üzerinde etkili olan deęişken buzaęı cinsiyeti olmuştur (Düz. P<0.007).Daha sonra 28,19 ay ile 30,90 ay olanların süt verimi üzerinde etkili olan buzaęılama aralıęı olmuştur (Düz. P<0.006). En yüksek süt verimi ise (8489.906 kg) DIM deęeri 26 aydan daha uzun olan ineklerden elde edilmiştir.

Sut ve Simsek (2011), insanlarda kafa yaralanması sonucu oluřan ölüm oranını tahmin etmek amacıyla altı farklı karar aęacı algoritmasını (CART, CHAID, Exhaustive-CHAID, QUEST, Random Forest RC, and Boosted Tree Classifiers and Regression (BTCR)) birbiriyle karşılařtırmıştır. Bu çalışmanın veri setini, kafa yaralanmasına maruz kalan 1603 kiři oluřturmuştur. İncelenen algoritmaların performansları, hassasiyet oranı (sensitivity), özgülük oranı (specificity), pozitif/negatif tahmin oranı (positive/negative predictive) and isabet oranı (accuracy rate) ölçütleri kullanılarak karşılaştırılmıştır. Ayrıca, her bir algoritmaya iliřkin ROC eęrisi altında kalan alanlar hesaplanmış, 0.801 ile 0.954 arasında deęiřtięi bildirilmiştir (P<0.001). Kafa yaralanması sonucu oluřan ölüm oranının tahmin edilmesinde, ROC eęrisi altında kalan en büyük alana sahip algoritmanın BTCR olduęu (0.954) ve bu algoritmaya iliřkin isabet oranının %93 olduęu saptanmıştır. Buna karşı, ROC eęrisinin altına kalan en küçük alana sahip algoritmanın CART (0.801) olduęu ve bu algoritma için isabet oranının % 91.1 olduęu belirtilmiştir. Bu çalışmada, kafa yaralanmaları sonucu oluřan ölüm oranına iliřkin doęru tahminler yapılması bakımından, BTCR algoritmasının kullanılmasının yararlı olabileceęi sonucuna varılmıştır.

Grzesiak ve Zborski (2012), hayvancılık alanında veri madenciliği metotlarının kullanımı ile ilgili teorik bilgiler vermiştir. Çalışmada, CART, CHAID, MARS (Multivariate Adaptive and Regression Splines), ANN (Artificial Neural Network), ve diğer makine öğrenme metotları (Naive Bayes Classifier (NBC), Destek Vektör Makinaları (Support Vector Machines =SVM), ve k-en yakın komşuluklar (k-nearest neighbors = k-NN) gibi veri madenciliği metotlarının teorik alt yapısından bahsedilmiştir.

Piwczynski ve Sitkowska (2012a), 455 baş Polonya Holstein ırkı inekte somatik hücre sayısını etkileyen laktasyonla ilgili faktörlerin (laktasyon sırası, sürü büyüklüğü, verim düzeyi, buzağılama yılı, buzağılama mevsimi, test günü mevsimi, laktasyon aşaması ve test sağımında elde edilen süt miktarı) etkilerini belirlemek amacıyla sınıflama ağacı ve lojistik regresyon metotlarını karşılaştırmışlardır. Sınıflama ağacının oluşmasında Entropy fonksiyonu ve Gini katsayısı gibi ayırma ölçütleri temel alınmıştır. Modellerin uyum kalitesi, karesel hata (squared error), yanlış sınıflama oranı (misclassification rate), cumulative lift, kolmogorov-smirnov istatistiği ve ROC eğrisi altında kalan alan gibi ölçütlere göre değerlendirilmiştir. Somatik hücre sayısı bakımından en iyi modelleme, Entropy fonksiyonunu temel alan sınıflama ağacı tekniği ile elde edilmiştir. Çalışmanın sonunda, somatik hücre üzerinde etkili olan faktörlerin önem sıralaması, sürü verim düzeyi ($P<0.001$), laktasyon sırası ($P<0.001$), buzağılama yılı ($P<0.001$), test sağım günü ($P<0.001$), sürü büyüklüğü ($P<0.001$), ve süt örneğinin alındığı ay ($P<0.001$) şeklinde bulunmuştur. Mevcut sonuçlar değerlendirildiğinde, yüksek süt verimli ineklerin bulunduğu sürülerde daha özel bakım koşullarının sağlanmasının gerekli olduğu sonucuna varılmıştır.

Piwczynski *et al.* (2012b), 20044 baş Polonya Merinos kuzularının ölüm oranından sorumlu olan faktörleri belirlemek amacıyla logistik regresyon ve sınıflama ağacı metotlarını kullanmışlardır. Bu kuzuların doğumdan 100. güne kadar olan yaşama gücü binomiyal skalada (canlı ve ölü) ifade edilmiştir. Araştırmacılar, Gini indeks ve Entropy fonksiyonu gibi ayırma ölçütlerine göre iki farklı model geliştirmiştir. Karar ağacı modelleri ve çoklu regresyonlar; hata kareler (squared error), yanlış sınıflandırma oranı (misclassification rate), cumulative lift, kolmogorov-smirnov istatistiği (kolmogorov-

smirnov statistic) ve ROC eğrisi altında kalan alan gibi ölçütlere göre mukayese edilmiştir. Entropy fonksiyonu ve Gini indeksini esas alan sınıflama ağaçlarının, yaşama gücüyle ilgili özelliklere (doğumdan 100. güne kadar olan yaşama gücü) ilişkin varyasyonu açıklamada daha isabetli sonuçlar verdiği bildirilmiştir. Bununla birlikte, en iyi sınıflama modeli Gini indeksi ile elde edilmiş, geliştirilen bu modelde farklılaşmayı sağlayan en önemli faktörlerin sürü, kuzuların doğum yılı ve tipi olduğu bildirilmiştir.

Yakubu *et al.* (2012), yaptıkları çalışmada, Uda koyunlarının canlı ağırlığı tahmininde CART algoritmasını uygulamışlardır. Araştırmacılar, kuzey merkez Nijerya'nın Nasarawa eyaletindeki farklı sürülerden rastgele seçilen 499 Uda koçu üzerinde çalışmışlar, canlı ağırlık ve bazı morfometrik özellikler arasında orta ve yüksek düzeyli ($r=0.43-0.76$) korelasyonlar bulmuşlardır ($P<0.01$). Canlı ağırlığın tahmin edilmesinde en önemli bağımsız değişkenlerin; göğüs çevresi, omuz genişliği, sağrı genişliği, vücut ve yüz uzunluğu olduğu belirlenmiştir ($P<0.001$). Böylece modele dahil edilen bu beş değişken ideal regresyon ağacının elde edilmesini sağlamışlardır. CART algoritması ile oluşturulan karar ağacı incelendiğinde, göğüs çevresi 94.05 cm' den geniş ve yüz uzunluğu 20.85 cm'den daha uzun olan hayvanların daha ağır olabileceği beklenmektedir. Sonuç olarak, elde edilen bu bilgilerin, Uda koyunlarının bakım-idare, seleksiyon ve ıslah çalışmaları açısından faydalı olabileceği kanaatine varılmıştır.

Çak *et al.* (2013), Brown Swiss sığırlarda laktasyon süt verimini etkileyen faktörlerin etkilerini regresyon ağacı (CHAID algoritması) metodu kullanarak belirlemeye çalışmışlardır. Araştırmacılar, modele dahil edilen değişkenlerden laktasyon süresinin (Düz. $P<0.01$) laktasyon süt verimi üzerinde birinci derecede etkili olduğunu, laktasyon sırası (Düz. $P<0.01$) ve buzağılama mevsim (Düz. $P<0.01$) değişkenlerinin ise ikinci derece etkili olduğunu bildirmişlerdir. Sonuç olarak, araştırmacılar en uygun buzağılama mevsiminin kış ve uzun laktasyon süresi ve laktasyon sırası 1'den büyük olan ineklerde süt verimi daha yüksek olabileceği belirtmişlerdir.

Eyduran *et al.* (2013a) regresyon ağacı algoritması olarak CHAID kullanarak Brown Swiss sığırlarda 305 günlük süt verimi ile kuruda kalma süresi, laktasyon sırası, buzağılama

mevsimi, yılı ve aralığı arasındaki ilişkileri incelemişlerdir. Bunun için 645 baş Brown Swiss sığıra ait 1884 laktasyon kaydı değerlendirilmiştir. Regresyon ağacı diyagramı incelendiğinde, süt verimini en çok etkileyen faktörün buzağılama yılı olduğu (Düz.P<0.01), ikinci dereceden etkili olan faktörlerin ise laktasyon sırası (Düz.P<0.01) ve buzağılama aralığı (Düz.P<0.01) olduğu ortaya çıkmıştır. Laktasyon sırası faktörü sadece 1996 ve daha önce, 1997-1998 ve 2002-2003 buzağılama yılları arasında laktasyon kaydı olan ineklerin süt verimini etkilemektedir (Düz.P<0.01). Buzağılama yılı 2002 ve 2003 olan ineklerin süt verimi, buzağılama aralığı tarafından etkilenmiştir.

Eyduran *et al.* (2013b), 46 baş Türk Saanen keçilerinde laktasyon süt verimi, somatik hücre sayısı ve meme özellikleri arasındaki ilişkiyi, çoklu doğrusal regresyon, stepwise regresyon, çoklu regresyon analizinde faktör analizi ve regresyon ağacı gibi farklı istatistiksel metotlarla incelemişlerdir. Bu çalışmada kullanılan kantitatif özellikler, laktasyon süt verimi, meme üst yüksekliği, meme genişliği, meme uzunluğu, meme çapı, meme hacmi, somatik hücre sayısı, laktasyon uzunluğu ve laktasyon süt veriminden oluşmaktadır. Exhaustive CHAID algoritması ile oluşturulan regresyon ağacı ilişkin sonuçlar incelendiğinde en yüksek süt veriminin (297.357 lt), meme açısı 50⁰ dereceden geniş keçilerden elde edildiği belirlenmiştir. Meme açısı 50⁰ dereceden daha dar olan keçiler içerisinde laktasyon süresi 191 günden büyük olanların süt verimi (243,66 lt), laktasyon uzunluğu 191 ve 191 günden daha kısa olanlardan (192.25 lt) yüksek olmuştur. Meme açısı 50⁰ dereceden dar olan keçilerin süt verimi, laktasyon süresinden etkilenmektedir (Düz.P<0.05). Ayrıca meme açısı 50⁰ dereceden daha dar ve laktasyon süresi 191 ve 191 günden daha kısa olan keçilerin süt verimi meme hacmi özelliğinden etkilenmiştir (Düz.P<0.05).

Eyduran *et al.* (2013c), Mengali koyunlarının doğum ağırlığı üzerinde etkili olan çevre faktörlerini belirlemek için CHAID algoritmasını kullanmışlardır. Çalışmada en etkili değişkenin doğum tipi olduğu belirlenmiştir. Ayrıca doğum yılı ve ana yaşı değişkenlerinin de doğum ağırlığı üzerinde etkili olduğu bildirilmiştir. Çalışmada, 59 aydan daha yaşlı analardan doğan kuzuların en yüksek doğum ağırlığına sahip olduğu tespit edilmiştir. Ana

yaşı faktörü, tekiz ve erkek kuzuların doğum ağırlığı üzerinde etkili olmuştur. Ancak, tekiz ve dişi kuzuların doğum ağırlığını etkileyen faktörün doğum yılı olduğu bildirilmiştir.

Maia *et al.* (2013), atların termal konforu ile yüzey ısısı arasındaki ilişkiyi incelemek amacıyla karar ağacına dayalı model kullanmışlardır. Bağımsız değişken olarak; koltuk altı, sağrı, göğüs ve kasık bölgelerindeki yüzey sıcaklıklarını kullanmışlardır. Bağımlı değişken, atlarda konfor durumunu temsil eden iki seviyeli (konforlu ve konforsuz) bir değişken olmuştur. Çalışmada uygulanan karar ağacı modeli ile % 74 oranında bir isabet oranı sağlanmış, termal konforun tahmin edilmesinde en etkili bağımsız değişkenlerin koltuk altı ve göğüs yüzey sıcaklıklarının olduğu belirlenmiştir. Sonuç olarak, veri madenciliği tekniklerinin, atların termal konfor ile ilişkili yeni değişkenlerin keşfedilmesinde faydalı olabileceği bildirilmiştir.

Piwczynski *et al.* (2013), 1257 baş Holstein ırkı ineğin olduğu bir populasyonda, ölü doğum (stillbirths) ve buzağılama kolaylığını (calving ease) etkileyen çevre faktörlerinin etkilerini belirlemek amacıyla CART ve QUEST sınıflama algoritmalarını kullanmışlardır. Çalışmada CART ve QUEST algoritmaları için elde edilen sınıflama ağaçları, üç ayırma ölçütüne (Pearson ki-kare, Entropi fonksiyonu, ve gini indeksi) ve beş uyum iyiliği ölçütüne (karesel hata (squared error), yanlış sınıflama oranı (misclassification rate), kümülatif yükselme (cumulative lift), kolmogorov-smirnov istatistiği ve ROC eğrisi altında kalan alan) göre değerlendirilmiştir. Ki - kare istatistiği ve Entropi fonksiyonunu esas alan sınıflama ağaçları, buzağılama kolaylığına ilişkin varyasyonu daha doğru tanımlamışlardır. Ancak, ayırma ölçütlerine bakılmaksızın, ölü doğum özelliğine ilişkin elde edilen sınıflama ağaç modelleri birbirine benzer bulunmuştur. Araştırmada, buzağılama kolaylığı üzerinde etkili olan değişkenlerin önem sıralamasına bakıldığında, en etkili değişkenin buzağıların canlı ağırlığı olduğu ve bunu laktasyon sırası, yetiştirme sistemi, gebelik süresi uzunluğu ve buzağı cinsiyeti değişkenlerinin izlediği belirtilmiştir. Ölü doğum üzerinde sadece buzağı doğum ağırlığı değişkeninin etkili olduğu saptanmıştır. Bu çalışma sonunda, grafiksel bir model olan sınıflama ağacı modelinin buzağılamadan sorumlu olan faktörlerin belirlenmesinde yararlı olabileceği bildirilmiştir.

Schlegel *et al.* (2013), bazı karar ağacı algoritmalarına ilişkin (CHAID, Exhaustive CHAID ve CART) yapılan uygulamalarda 1780 baş dişi domuz kullanmış, toplam doğan yavru domuz sayısına ilişkin (bağımlı değişken) toplam varyasyonun açıklanmasında bağımsız değişken olarak canlı doğan yavru domuz sayısı, sütten kesim ağırlığına ulaşan yavru domuz sayısı, bir batında doğan yavru sayısı, baba sayısı ve yapay tohumlama tekniği gibi değişkenleri modele dahil etmişlerdir. Araştırmacılar, yavru sayısının önemli bir değişken olmadığını, bir batında doğan yavru domuzlardan sütten kesim ağırlığına gelenlerin sayısının fazla bilgi sunmadığını, ancak tohumlama tekniğinin karar ağaçlarında yeni dallanmalar oluşturarak etkili olduğunu bildirmişlerdir. Çalışmanın sonucunda, dişi domuzların farklı değişkenlere göre gruplandırılmasında sınıflama ağacı metotları kullanımının yararlı olabileceği bildirilmiştir.

Yılmaz *et al.* (2013), Brown Swiss ırkı sığırlar üzerinde yaptıkları çalışmada, buzağılama mevsimi, buzağılama yılı ve buzağı cinsiyeti faktörlerinin doğum ağırlığı üzerindeki etkilerini Geniş (Exhaustive) CHAID regresyon ağacı algoritmasıyla incelemişlerdir. Araştırmacılar, doğum yılının (Düz.P<0.01) doğum ağırlığı üzerinde birinci derecede etkili bir faktör olduğunu ve bunu ikinci derece etkili cinsiyet (Düz.P<0.01) ve buzağılama mevsimi (Düz.P<0.01) faktörlerinin izlediğini ifade etmişlerdir. Oluşturulan regresyon ağacı sonucunda, en yüksek doğum ağırlığı ortalamasının 1991-2004 yılları arasında doğan buzağılardan elde edildiği belirtilmiştir. Çalışmada 2002-2004 yılları arasında yaz ve sonbahar mevsiminde doğan buzağıların oluşturduğu alt grubun doğum ağırlığı ortalaması, ilkbahar ve yaz mevsiminde doğan buzağıların oluşturduğu alt grubun ortalamasından daha hafif olduğu tespit edilmiştir. 1991-1995 yılları arasında doğan buzağılardan oluşan alt grubun doğum ağırlığı üzerine cinsiyetin etkisi önemli bulunmuştur (Düz.P<0.01). Buzağılama mevsimi, 2002 ile 2004 yılları arasında doğan buzağıların oluşturduğu grubun doğum ağırlığı üzerinde etkili olmuştur (Düz.P<0.01). Sonuç olarak araştırmacılar, çoklu bağlantılılık ve sapan gözlemler (outliers) ve kayıp gözlemlerden (missing observations) etkilenmediğinden regresyon ağacı yönteminin kullanılmasını tavsiye etmişlerdir.

Khan *et al.* (2014), Harnai koyunlarında çoklu bağlantılılık (multicollinearity) problemini elemine etmek suretiyle bazı vücut ölçülerinden (vücut uzunluğu, cidago yüksekliği, göğüs çevresi, karın çevresi, yüz uzunluğu, kulaklar arası uzunluk, kuyruk genişliği ve uzunluğu) canlı ağırlığı tahmin etmişlerdir. Araştırmacılar, bağımsız değişkenlere yani vücut ölçüleri için faktör ve temel bileşen analizleri uygulamış, birbirinden bağımsız (uncorrelated) olan yeni bağımsız değişkenler (faktör ve temel bileşen skor değişkenleri) elde ederek çoklu bağlantılılığın olumsuz etkilerini tamamen kaldırmışlardır. Ayrıca, çoklu bağlantılılıktan etkilenmeyen regresyon ağacı modeli kullanarak canlı ağırlığı tahmin etmişlerdir. Regresyon ağacı (Geniş (Exhaustive) CHAID) sonuçları incelendiğinde, canlı ağırlığı birinci derecede etkileyen en önemli bağımsız değişkenin yüz uzunluğu olduğu (Düz. $P < 0.001$) ve bunu ikinci derece cidago yüksekliği (Düz. $P < 0.001$) ve göğüs çevresi (Düz. $P < 0.001$), üçüncü derecede göğüs çevresi (Düz. $P < 0.001$) ve vücut uzunluğu (Düz. $P < 0.001$) değişkenlerinin izlediği görülmüştür. Yüzü uzun olan koyunların daha ağır olduğu, yüz uzunluğu 13.750-15.650 cm arasında olan hayvanların canlı ağırlıklarının farklılaşmasını sağlayan özelliğin cidago yüksekliği olduğu, yüz uzunluğu 13.750-15,650 cm, cidago yüksekliği > 15.5 cm, ve göğüs çevresi > 48.45 olan koyunların 32.598 kg ortalama ağırlığa sahip olduğu, yüz uzunluğu 18.65 -23 cm olan koyunların canlı ağırlığı üzerinde göğüs çevresinin etkili olduğu, yüz uzunluğu 18.65 -23 cm ve göğüs çevresi > 46.75 cm olan hayvanların canlı ağırlığının vücut uzunluğu özelliği tarafından etkilendiği saptanmıştır. Sonuç olarak, oluşturulan regresyon ağacına göre canlı ağırlığa ilişkin toplam varyasyonun % 84.4'ünün açıklandığı belirtilmiştir.

Bayram *et.al.* (2015), tarafından Gümüşhane ilinde yetiştiricilik yapan özel bir süt sığırı işletmesinde 2004 ile 2006 yılları arasında doğum yapan 613 baş Siyah Alaca ineğin 947 buzağılama kaydını kullanarak yaptıkları çalışmada, bazı çevre faktörlerinin Siyah Alaca sığırlarda güç ve ölü doğuma etkilerini CHAID algoritması ile değerlendirmişlerdir. Buzağılama güçlüğü üzerinde etkili olan en önemli değişkenlerin, buzağı doğum ağırlığı, buzağılama tipi ve mevsimi olduğunu ve buzağılara ait doğum ağırlığının artıkça, buzağılama güçlüğünün önemli oranda (Düz. $P < 0.01$) olduğunu tespit etmişlerdir. Doğumu

ikiz olan buzağılarda doğum riski (% 18.9), tek doğanlardan (% 5.0) oldukça yüksektir (Düz.P<0.01). Ölü doğumu etkileyen çevresel etkiler sırasıyla; doğum sırası, buzağının doğum ağırlığı, buzağı cinsiyeti, buzağılama mevsimi ve buzağılama güçlüğü olmuştur. Ölü doğum üzerine önemli etkili değişken doğum sırası olup, ilkin doğum yapan ineklerde ölü doğum oranı (% 18.7), çoklu doğum yapanlardan (% 5.4) yaklaşık 3.5 kat daha yüksektir. Buzağıya ait doğum ağırlığı arttıkça, ölü doğum oranı önemli oranda artmıştır (Düz.P<0.01). Kış mevsiminde gerçekleşen doğumlarda ölü doğum (% 19.7), diğer mevsimlerden (% 3.6) yüksek bulunmuştur (Düz.P<0.01). Güç doğan buzağılarda ölü doğum (%10.0), kolay doğanlardan (% 1.9) daha yüksek olduğu belirtilmiştir (Düz.P<0.01). Sonuç olarak, Siyah Alaca sığırlarda güç doğuma etkili değişkenler buzağı doğum ağırlığı, doğum tipi ve buzağılama mevsimi olurken ölü doğum üzerine doğum sırası, buzağının doğum ağırlığı, buzağı cinsiyeti, buzağılama mevsimi ve buzağılama güçlüğü etkili değişkenler olmuştur.

3. MATERYAL ve METOT

3.1. Materyal

CHAID, Geniş (Exhaustive) CHAID ve CART veri madenciliği algoritmalarının hayvancılıkla ilgili çalışmalarda kullanımlarını göstermek amacıyla yürütülen bu tez çalışmasının hayvan materyalini, Pakistan'ın farklı işletmelerinde yetiştirilen 1884 Mengali koyunlarına ilişkin veriler oluşturmuştur. Bu çalışmada, hayvancılıkla ilgili bir veri setinde, farklı ebeveyn çocuk kombinasyonu (500:250'den 10:5'e kadar) kullanılmıştır. Sütten kesim ağırlığını tahmin etmek için, cinsiyet (erkek ve dişi), doğum tipi (tekiz ve ikiz), doğum yılı (2005, 2006, 2007, 2008 ve 2009) çiftlik (Araştırma istasyonu, Mastung, Quetta ve Noshki), ana yaşı (20-78 ay) ve ana ağırlığı (25-48 kg) gibi bağımsız değişkenler kullanılmıştır.

3.2 Metot

CHAID, Geniş (Exhaustive) CHAID ve CART veri madenciliği algoritmalarını esas alan regresyon ağacı yöntemi, canlı ağırlık, süt verimi ve yapağı verimi gibi kantitatif olan bir bağımlı değişken ile birden fazla nicel, sıralı ve sürekli bağımsız değişken arasındaki ilişki yapısı hakkında, araştırmacılara bilgi veren kullanışlı bir metottur. Regresyon ağacı yönteminde bağımlı ve bağımsız değişkenlerin yapısı; sürekli, sınıflayıcı ve sıralı olabilmektedir (Mendeş ve Akkartal, 2009). Ancak, sınırlı değer alan bağımlı değişkenler için kurulan ağaç modeller sınıflama ağaçları (Classification Trees) olarak isimlendirilirken, sürekli bağımlı değişkenler için oluşturulan ağaç modeller ise Regresyon ağaçları (Regression Trees) olarak isimlendirilirler.

Yukarıda bahsedilen ağaca dayalı algoritmaları esas alan regresyon ağacı yöntemi; büyük ve karmaşık veri setlerinde, bağımlı değişken üzerinde etkili olan bağımsız değişkenlerin önemlilik durumlarını kolay anlaşılabilir ve görsel bir karar ağacı yapısı şeklinde görülmesini sağlar. Regresyon ağacı, önemli bağımsız değişkenlerden yola çıkarak veri setini mümkün olduğunca homojen alt gruplara ayrılması sonucu oluşan bir karar

ağacıdır. Oluşturulan karar ağacı en tepede bulunan kök düğüm ile başlar ve tekrarlamalı olarak birbirini takip eden alt gruplar (yavru düğümleri) şeklinde oluşmaya devam etmektedir (Çamdeviren *et al.*, 2005). Ağaç yapısında, yeterince homojenlik sağlanan alt gruplarda artık bölünme yani yeni alt gruplar oluşmamaktadır. Diğer bir ifadeyle, sonraki aşamalarda bölünmeyen alt gruplara ya da düğümlere terminal düğüm adı verilir. Veri madenciliği algoritmaları ile oluşturulan regresyon ağaçlarında amaç, heterojen bir veri setinden yola çıkarak mümkün olduğunca homojen alt gruplar yani terminal düğümler elde ederek düğümler arası varyansın payını arttırmaktır.

Regresyon ağaçlarını oluşturma adımları Küçükoğlu (2010) tarafından belirtildiği gibi oluşturulmuştur.

Bu adımlar;

- a) Regresyon analizine başlarken regresyon ağacında kök düğüm ile başlanır.
- b) m_c herhangi bir "c" dalına ilişkin tahmin değerini, V_c ; c dalına ayırma varyansını ve n_c 'de bu daldaki gözlem sayısını belirtmek üzere ayırma işlemi $S = \sum \sum (Y_i - m_c)^2$ biçiminde ifade edilen S değeri minimum oluncaya kadar devam ettirilir. Bu ifade $S = \sum n_c V_c$ biçiminde de gösterilebilir, m_c ise $\sum c Y_i / n_c$ dir.
- c) Düğüm içerisinde bütün bireyler dikkate alınan tüm bağımsız değişkenler için aynı değeri alıyorsa ağacın bölünme işlemi durdurulur. Aksi takdirde S minimize edilebilecek en son noktaya kadar düşürülür. Eğer S'deki en büyük azalma belirlenen δ gibi bir eşik değerinden küçük ise bölünme sonlanır.
- d) Her bir yeni düğüm için 1. Adıma geri dönülür.

İdeal bir regresyon ağacını uygun bir biçimde elde etmek için uygun yöntemlerden birisi çapraz geçerlilik (crossvalidation) yönteminin kullanılmasıdır. Herhangi k düğümü için $R(k)$ nodun varyansını ya da k düğümünün risk değerini gösterir.

$$R(k) \text{ hesaplanırken; } R(k) = \frac{1}{N(k)} \sum_{i \in k} [Y_i - Y(k)]^2$$

$N(k)$: düğümün k içerisindeki gözlem sayısını gösterir.

Y_i : Hedef deęişken deęerini gösterir.

$Y(k)$: k düęümün ortalamasıdır.

Risk deęeri baęımlı deęişkenin oranına baęlıdır.

Hata varyansı ya da açıklanmayan varyasyon S_e^2 şöyle hesaplanır.

$$S_e^2 = \frac{Risk}{S_y^2}$$

Bu bilgilere baęlı olarak baęımlı deęişkendeki varyasyonun oluşturulan ağaç tarafından açıklanabilen kısmı ise,

$$S_x^2 = 1 - S_e^2 \text{ şeklindedir (Çamdeviren et al., 2005; Mendeş ve Akkartal, 2009).}$$

Ağacın oluşturulmasındaki temel amaç, cevap deęişkeni bakımından olabildiğince homojen olan yeni yavru düęümler ya da yavru dallar (child branches) oluturmaktır. Bununla birlikte yavru düęümlerin oluşturulması, tekrarlamalı olduğundan; oluşturulan herhangi bir yavru düęümde homojenlik mümkün olduğunca sağlanmışsa, bu düęümde artık parçalanma süreci durdurulur ve bu düęüm terminal yada üç düęüm olarak isimlendirilir. CART analizinde, yukarıda belirtilen işlem yapılırken genel olarak veri seti iki alt gruba ayrılır. Bu alt gruplardan birisi bilgi ya da çalışma grubu (learning, training) dięeri ise test ya da geçerlilik (testing, validation) grubudur. Çalışma grubu *Learning sample* olarak tanımlanır ve veri seti L ile gösterilir (Oruçoęlu 2011).

$$X = \{x_1, x_2, \dots, x_n, \dots, x_N\}$$

$$J = \{j_1, j_2, \dots, j_n, \dots, j_N\}$$

$$L = \{(x_1, j_1), (x_2, j_2), \dots, (x_n, j_n), \dots, (x_N, j_N)\}$$

$$x_n \in X ; j_n \in \{1, 2, \dots, C\} \text{ ve } n = 1, 2, \dots, N \text{ dir.}$$

Burada;

L : başlangıç veri seti (baęımsız deęişkenlere ait ölçüm vektörü bilgilerinin tümünü içine alır).

X : baęımsız deęişkenlere ait ölçüm vektörü,

J : deney ünitelerinin ait olduğu sınıflar (kategoriler) vektörü,

x_n : n. birey ya da objeye ait bağımsız değişkenin değeri,
 j_n : n. birey ya da objenin ait olduğu sınıf (kategori),
 C : deney ünitelerinin ait olduğu sınıf (kategori) sayısı ve
 N : toplam deney ünite sayısıdır.

3.2.1. Veri madenciliği algoritmaları

3.2.1.1. CART algoritması

İlk olarak Breiman *et al.* (1984) tarafından geliştirilen CART veri madenciliği algoritması, veri setindeki (learning sample) bütün bireyleri içine alan kök düğümden başlayarak, tekrarlamalı olarak homojen düğümler elde edilene kadar her aşamada bir düğümden sadece iki yavru düğüm şeklinde bölünen ikili karar ağacı algoritmasıdır. CART non-parametrik bir teknik olmakla birlikte bir veri setinin dağılımı ile ilgili özelleştirme gerektirmez ve önemli olan bağımsız değişkenleri otomatik olarak ağaç yapısına dahil eder (Nisbet *et al.*, 2009).

CART için her bölünme sadece bir bağımsız değişkenin değerine bağlıdır. 1 adet kategorisi olan sıralı değişken için 2I-1-1 adet bölünme söz konusudur. Bununla birlikte, CART algoritması ile elde edilecek bir karar ağacında yorumlama kolaylığının sağlanması açısından tahminleme performansını düşüren gereksiz dalların ya da çocuk düğümlerin ağaç yapısından budamak için SPSS programında budama (pruning) işlemi aktif hale getirilmelidir. Ancak söz konusu budama işlemini, CHAID ve Geniş (Exhaustive) CHAID veri madenciliği algoritmaları için SPSS programı otomatik olarak yapmaktadır. Diğer iki algoritmadan farklı olarak CART algoritması ile oluşturulan bir karar ağacında aynı değişkenin bir dalda birkaç kez bölünme (surrogate splits) olması söz konusudur. Diğer bir anlatımla, CART algoritmasının daha fazla dallanma potansiyeli olan bir algoritma olduğu söylenebilir. Bu gibi durumları önlemek için çalışılan veri setini öğrenme (Learning, Training) ve test (testing, validation) setlerine ayrılması, ebeveyn-yavru düğümde bulunması gereken birey sayısının ve ağaç derinliği (tree depth) sayısının dikkatli bir şekilde

ayarlanması önerilebilir. Genelde öğrenme ve test setleri için 80:20 oranının kullanımı tavsiye edilir. İyi bir ağaç yapısı için bu iki set için hesaplanan model kalite ölçütlerinin birbirine yakın olması gerekmektedir. Diğer bir seçenek olarak iki set oluşturmak yerine çapraz geçerlik (validation) 10 alınıp mevcut veri seti için ortalama bir hata değeri hesaplanabilir.

Bir bireye ait bağımlı değişkene ait değer ya da tüm bağımsız değişkenlere ait değerler kayıpsa, CART algoritmasında bu birey analize dahil edilmez.

Y Bağımlı değişken sıralı, isimsel ya da sürekli olabilir.

$X_m, m = 1, \dots, M$ Bütün bağımsız değişkenlerin olduğu set

$\tilde{n} = \{x_n, y_n\}_{n=1}^N$ Öğrenme örneğini içine alan veri setinin tamamı

$\tilde{n}(t)$ t düğüme atanan öğrenme örnekleri

w_n n . bireye ilişkin ağırlık ile ilgili durum ağırlığı,

f_n n . bireye ilişkin frekans ağırlığı,

Bağımlı değişken SKA gibi sürekli olduğu zaman Gini, Twoing ve Sıralı Twoing ayırma kriterleri yerine ayırma kriteri (splitting criterion) $\Delta(\delta.(t)) = i(t) - P_L.i(t_L) - P_R.i(t_R)$

$$\Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R)$$

En Küçük Kareler Sapma (LSD) safsızlık ölçümü ile birlikte kullanılır.

$$i(t) = \frac{\sum_{n \in \tilde{n}(t)} w_n f_n (y_n - \bar{y}(t))^2}{\sum_{n \in \tilde{n}(t)} w_n f_n}$$

$$p_L = N_w(t_L)/N_w(t), p_R = N_w(t_R)/N_w(t), N_w(t) = \sum_{n \in \tilde{n}(t)} w_n f_n$$

Burada;

P_L : t. düğümde sol yavru düğüme atanan bireylerin (kuzuların) oranı,

P_R : t. düğümde sağ çocuk düğüme atanan bireylerin (kuzuların) oranı,

$i(t)$: t düğüme ait safsızlık ölçüsü,

$i(t_L)$: sol yavru düğüme ait safsızlık ölçüsü ve

$i(t_R)$: sağ yavru düğüme ait safsızlık ölçüsünü temsil etmektedir (Oruçoğlu, 2011).

$$\bar{y}(t) = \frac{\sum_{n \in \tilde{h}(t)} w_n f_n y_n}{N_w(t)}$$

3.2.1.1.2. Variable importance (Değişkenlik önemi)

CART algoritmasında elde edilen son ağaç (T) ile ilgili olarak X bağımsız bir değişkeninin önem derecesi, söz konusu bu X değişkeninin katkı (improvement) değerlerinin bütün düğümlerdeki ağırlıklı toplamı olarak tanımlanır.

$$M(X) = \sum_{t \in T} \Delta(\tilde{s}_X, t)$$

X bağımsız değişkenine ait önem derecesi VI (X), en yüksek önem dereceli bağımsız değişkene oranla X değişkeninin regresyon ağacına sağladığı katkırı gösteren normalleştirilmiş miktar olarak ifade edilir. Önem derecesi 0 ile 100 arasında değişir. En yüksek önem derecesine sahip değişken 100 değerini alır.

$$VI(X) = \frac{M(X)}{\max_X M(X)} \times 100$$

3.2.1. 2. CHAID ve Geniş (Exhaustive) CHAID algoritmaları

CART ile karşılaştırıldığında, bir düğümün çok sayıda bölünmesine (multiple splitting) izin veren CHAID algoritması Kass (1980) tarafından geliştirilmiştir. Bu bakımından CHAID'e benzeyen Geniş (Exhaustive) CHAID algoritması ise Biggs *et al.*,

(1991) tarafından geliştirilmiştir. Birleştirme, bölünme ve durma aşamalarına sahip olan CHAID algoritmaları kök düğümden başlayarak tekrarlamalı olarak homojen düğümler oluşturur ve böylece düğümler arası ve düğümler içi varyans artırılır veya azaltılır (Nisbet *et al.* 2009). Düğümler arası varyans, toplam birey sayısına, analizde kullanılan bağımsız değişkenlerin etkili olup olmamasına, ebeveyn:yavru düğümlerde bulunması planlanan minimum birey sayısına ve ağacın derinliğine göre değişebilir (Ali *et al.*, 2015). SPSS programında, ağaç derinliği her iki CHAID algoritması için 3 olmasına karşın, CART için bu derinlik 5 olarak ayarlanmıştır. Ancak, ağaç derinliği değiştirilebilir.

3.2.1.2.1.CHAID algoritması oluşum aşamaları

a. Birleşme (Merging) aşaması; Bağımsız değişkene ilişkin önemsiz kategoriler birleştirilir. Bu aşamada da düzeltilmiş P (Düz-P) değeri hesaplanır. X düğümü ayrılabilirse her bir belirleyici X değeri için önemsiz kategoride birleşir X in her bir sonuç kategorisi çocuk düğümün birinde sonuçlanacak. Ayrıca hesaplama adımları birleşerek p-değeri ayarlanır bölünme adımlarında kullanılır.

- 1- Eğer X bağımsız değişkeninin sadece 1 kategorisi varsa, düzeltilmiş P değeri 1 olarak ayarlanır
- 2- Eğer X bağımsız değişkeninin 2 kategorisi varsa, 8. Aşamaya gidilir.
- 3- Aksi halde, X bağımsız değişkenine ait kabul edilebilir kategori çiftlerini bulur. Sıralı değişken için birbirine benzeyen ardışık kategori çiftleri bulurken isimsel değişken için herhangi iki kategoriyi bulur.
- 4- P değeri, kullanıcılar tarafından özelleştirilen $\alpha_{\text{Birleştirme}}$ değerinden büyükse kategori çifti tek bir kategori olarak birleştirilir. Aksi halde, 7. Aşama gidilir.
- 5- (Tercihe bağlı) eğer yeni oluşturulan bileşik kategori 3 ya da daha fazla orijinal kategoriden oluşursa,
- 6- 2. adıma gitmek için

7-Birkaç gözleme sahip olan herhangi bir kategori (özel bir kullanıcı ile karşılaştırılarak minimum kısımlı boyuta ayrılır). En büyük düz.- P değerine sahip olan ve ölçülmüş diğer kategoriye en benzer kategori ile birleştirilir.

8- Düzeltilmiş P (Düz.-P), sonraki tartışmalarda kullanmak için Bonferroni uygulayarak kategoriler birleştirilmiştir.

b. Splitting (Ayrırma)

Her tahmin edici için, en iyi ayırma birleşme adımlarında bulunur. Tahmin edici, ayırma adımını seçmek için, en iyi ayırma düğümünü tahmin eder. Seçim, düzenlenmiş p-değeri, her bir tahmin edici ile ilişkili olarak karşılaştırılır. Düzenlenmiş p-değeri birleştirme adımında elde edilir.

1- Tahmin edici, en küçük düzenlenmiş p-değerine sahip olanı seçer.

2- Eğer, düzenlenmiş p-değeri, özel kullanıcı alfa (α) seviyesine eşit ya da daha az ise, bölünen düğümü kullanarak tahmin edilir. Aksi takdirde düğüm olmaz çünkü düğüm yeterli homojeniteye ulaştığından bölünme olmaz, bu düğüm terminal düğüm olarak ortaya çıkar.

c. Stopping (Durdurma)

Ağaç oluşturma sürecindeki durdurma adımları, aşağıdaki kurallara göre durdurulmalı;

-Eğer bir düğüm homojen ise yani bütün durumlarda düğüm bağımlı değişkenin aynı değerine sahip ise düğüm bölünmez.

-Eğer bir düğümdeki bütün durumlar her belirleyici için aynı değere sahip ise düğüm bölünmez.

-Eğer geçerli ağaç derinliği kullanıcıda belirtilen maksimum ağaç derinliğine ve sınır değerine ulaşırsa ağacın ilerleme süreci durur.

-Eğer bir düğümün boyutu, kullanıcının minimum düğüm boyutu değerinden az ise düğüm bölünmez.

-Eğer bölünen bir düğümün çocuk düğümdeki sonuçları kullanıcının düğüm boyutundan az ise çocuk düğümünün boyut değeri minimumdur. Çocuk düğümler P-değerinin en büyüğü tarafından ölçülmüş olan en fazla benzer durumdaki ile birleştirilir (minimumla karşılaştırılır). Fakat çocuk düğümünün sonuç sayısı 1 olursa bu durumda düğüm bölünmez.

3.2.2. Model kalite kriterleri

Kriterler	Formüller
Belirleme Katsayısı (Coefficient of Determination)	$R^2 = \left[1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \right] * 100$
Düzeltilmiş Belirleme Katsayısı (Adjusted Coefficient of Determination)	$\bar{R}^2 = \left[1 - \frac{\frac{1}{n-k-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2} \right] * 100$
Risk Tahmini (Risk Estimate)	$RT = \frac{1}{n-k-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$
Varyasyon Katsayısı (%) (Coefficient of Variation %)	$\sqrt{\frac{\frac{1}{n-1} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2}{\bar{Y}}} * 100$
Standart Sapma Oranı (Standart Deviation Ratio)	$\sqrt{\frac{\frac{1}{n-1} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2}{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2}}$

Relatif Yaklaşık Hata (Relative Approximation Error)	$\sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n Y_i^2}}$
Hata Kareler Ortalamasının Karekökü (Root Mean Square Error)	$\sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$
Modifiye Relatif Yaklaşık Hata (Modified Relative Approximation Error)	$\sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n \sum_{i=1}^n Y_i^2}}$

Burada;

$$\varepsilon_i = Y_i - \hat{Y}_i \text{ (Gerçek SKA – Tahmin edilen SKA)}$$

SKA=sütten kesim ağırlığı

n= Toplam hayvan sayısı (n)

i 1,2,...,n(muamele)

\bar{Y} genel ortalama

\hat{Y}_i i . kuzuya ait tahmin değeri

ε_i i . kuzuya ait şansa bağlı hata değeri

$\bar{\varepsilon}$ şansa bağlı hata değerlerine ait ortalama hata değeri

Y_i i . kuzuya ait SKA değerini ifade etmektedir.

4. BULGULAR ve TARTIŞMA

Bu çalışmada, süttten kesim ağırlığı üzerine bazı çevre faktörlerinin etkilerini incelemek amacıyla CHAID, Geniş (Exhaustive) CHAID ve CART algoritmalarının performansları karşılaştırılmış, en uygun algoritma ve ebeveyn yavru düğümü oranı belirlenmeye çalışılmış, elde edilen sonuçlar sırasıyla Çizelge 4.1-4.6'da verilmiştir. Genel olarak bütün algoritmalar için değerlendirmeye alınan ebeveyn ve yavru düğümlerinde, 500:250 oranından 10:5 oranına doğru bir azalma gerçekleştiğinde uyum ölçüt değerlerinin daha iyi olduğu anlaşılmıştır. Buna karşın, gerek Geniş (Exhaustive) CHAID gerekse CART algoritmalarından üretilen regresyon ağaçlarının biyolojik olarak uygun sonuçlar üretmediği anlaşılmış, söz konusu ağaçları oluşturan dallarda aynı değişkenlerin tekrarlandığı gözlenmiştir. Literatürde ilk olarak süttten kesim ağırlığı için CHAID, Geniş (Exhaustive) CHAID ve CART algoritmalarına dayalı oluşturulan regresyon ağacı modellerinin performanslarına ilişkin bilgiler sunulmuştur.

4. 1. Süttten Kesim Ağırlığına İlişkin Hesaplanan Uyum İyiliği Ölçütlerine Ait Sonuçlar

4.1.1.CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait bulgular

Minimum ebeveyn düğüm (Çizelge 4.1): çocuk düğüm oranı 500:250'den 10:5'e doğru gidildikçe CHAID algoritmasına ilişkin risk tahmin değerleri 0.992'den 0.510'a, NYH değerleri 0.249'dan 0.178'e, HKOK değerleri 0.996'dan 0.506'ya ve SS-oranının ise 0.706'dan 0.506'a azaldığı görülmüştür. Bununla birlikte, belirleme katsayısı değerleri % 50.176'dan %74.385'e yükseldiği belirlenmiş ve süttten kesim ağırlığının gerçek ve tahmin değerleri arasındaki Pearson korelasyonlarının 0.708'den 0.863'e doğru yükseldiği saptanmıştır ($P<0.01$). Tüm sonuçlar değerlendirildiğinde, CHAID algoritmasının kullanılması sonucu düğümler içi (hata) varyansın da azaldığı dikkat çekmiştir. Çizelge 4.1'e bakıldığında, ebeveyn-yavru oranlarının 380:190 ile 500:250 arasında CHAID algoritmasına ilişkin uyum iyiliği ölçütlerinin değişmediği anlaşılmaktadır. Uyum iyiliği ölçütleri bakımından Mengali kuzularının süttten kesim ağırlığı için CART algoritması kullanımının diğer algoritmalara nazaran biraz daha avantajlı olduğu tespit edilmiştir. CART

algoritmasından elde edilen sonuçlara ilişkin yorumların biyolojik olarak daha uygun olduğu ortaya konulmuştur.

Çizelge 4.1. CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar

Ebeveyn düğüm	Çocuk düğüm	RT	R^2 (%)	NYH	HKOK	SS_oran	r	R^2 -düz.(%)	VK(%)
500	250	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
480	240	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
460	230	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
440	220	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
420	210	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
400	200	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
380	190	0.992	550.176	0.249	0.996	0.706	0.708	50.128	6.222
360	180	0.951	552.235	0.249	0.996	0.706	0.708	50.102	6.222
340	170	0.976	550.979	0.244	0.975	0.691	0.723	52.184	6.092
320	160	0.966	51.482	0.246	0.983	0.697	0.717	51.429	6.140
300	150	0.955	552.034	0.244	0.977	0.693	0.721	51.981	6.105
280	140	0.955	552.034	0.244	0.977	0.693	0.721	51.955	6.105
260	130	0.953	552.135	0.244	0.976	0.692	0.722	52.071	6.098
240	120	0.926	553.491	0.240	0.962	0.682	0.731	53.368	6.011
220	110	0.926	553.491	0.240	0.962	0.682	0.731	53.401	6.011
200	100	0.902	554.696	0.237	0.950	0.673	0.739	54.575	5.933
180	90	0.814	559.116	0.225	0.902	0.640	0.769	59.015	5.636
160	80	0.801	559.769	0.224	0.895	0.635	0.773	59.631	5.592
140	70	0.791	60.271	0.222	0.889	0.630	0.776	60.156	5.555
120	60	0.721	663.787	0.212	0.849	0.602	0.799	63.707	5.303
100	50	0.761	661.778	0.218	0.872	0.618	0.786	61.696	5.448

Çizelge 4.1. (devam)

Ebeveyn düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS-oran	r	R^2 -düz.(%)	VK(%)
80	40	0.687	665.495	0.207	0.829	0.588	0.809	65.397	5.178
60	30	0.663	666.700	0.203	0.814	0.577	0.817	66.614	5.085
40	20	0.571	771.321	0.189	0.756	0.536	0.844	71.222	4.721
20	10	0.532	773.280	0.182	0.730	0.517	0.856	73.173	4.558
10	5	0.510	774.385	0.178	0.714	0.506	0.863	74.320	4.460

4.1.2. Geniş (Exhaustive) CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait bulgular

Ebeveyn düğüm: çocuk düğüm oranı 500:250'den 10:5'e doğru gidildikçe Geniş CHAID algoritmasına ilişkin risk tahmin değerleri 0.992'den 0.511'e, NYH değerleri 0.249'dan 0.179'a, HKOK değerleri 0.996'dan 0.715'e ve SS-oranının ise 0.706'dan 0.507'ye azaldığı görülmüştür (Çizelge 4.2). Bununla beraber, belirleme katsayısı değerleri % 50.176'dan %74.335'e yükseldiği ve süttten kesim ağırlığına ait gerçek ve tahmin değerleri arasındaki Pearson korelasyonlarının 0.708'den 0.862'e doğru yükseldiği saptanmıştır ($P<0.01$).

Çizelge 4.1 ve 4.2'ye incelendiğinde 500:250 ile 380:190 arasında bulunan ebeveyn-yavru düğüm oranları için Geniş (Exhaustive) CHAID algoritmasının, CHAID algoritmasına benzer sonuçlar verdiği gözlenmiştir. Ancak, 300:150, 240:120 ve 80:40 ebeveyn-yavru düğümü oranları için Geniş (Exhaustive) CHAID algoritmasından türetilen karar ağaçlarının bazı dallarında doğum ağırlığı kovaryeti ardı ardına yer almıştır. Bahsedilen bu üç oran için CART algoritmasının kullanılması önerilebilir. Geniş (Exhaustive) CHAID algoritmasının, CHAID algoritmasına nazaran daha uzun sürede işlem yaptığı bildirilmiştir (Tariq et al, 2012.)

Çizelge 4.2. Geniş (Exhaustive) CHAID algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar

Ebeveyn düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS -Oranı	r	$R^2_{düz.}$ (%)	VK(%)
500	250	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
480	240	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
460	230	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
440	220	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
420	210	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
400	200	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
380	190	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
360	180	0.951	52.235	0.244	0.975	0.691	0.723	52.158	6.092
340	170	0.976	50.979	0.247	0.988	0.701	0.714	50.903	6.173
320	160	0.966	51.482	0.246	0.983	0.697	0.717	51.429	6.140
300	150	0.955	52.034	0.244	0.977	0.693	0.721	51.955	6.105
280	140	0.955	52.034	0.244	0.977	0.693	0.721	51.955	6.105
260	130	0.953	52.135	0.244	0.976	0.692	0.722	52.071	6.098
240	120	0.926	53.491	0.240	0.962	0.682	0.731	53.368	6.011
220	110	0.922	53.692	0.240	0.960	0.681	0.733	53.575	5.998
200	100	0.877	55.952	0.234	0.937	0.664	0.748	55.812	5.852
180	90	0.808	59.417	0.225	0.899	0.637	0.771	59.341	5.615
160	80	0.795	60.070	0.223	0.892	0.632	0.775	59.979	5.570
140	70	0.792	60.221	0.222	0.890	0.631	0.776	60.092	5.561
120	60	0.746	62.531	0.216	0.864	0.612	0.791	62.442	5.395
100	50	0.740	62.833	0.215	0.860	0.610	0.793	62.707	5.374
80	40	0.706	64.540	0.210	0.840	0.596	0.803	64.442	5.248
60	30	0.662	66.750	0.203	0.813	0.577	0.817	66.656	5.082
40	20	0.565	71.622	0.188	0.752	0.533	0.846	71.535	4.695
20	10	0.534	73.179	0.183	0.731	0.518	0.855	73.106	4.564
10	5	0.511	74.335	0.179	0.715	0.507	0.862	74.273	4.464

4.1.3. CART algoritmasına ilişkin bulgular

Farklı ebeveyn-çocuk düğüm oranlarının bulunduğu Çizelge 4.3 incelendiğinde, ebeveyn düğüm: yavru düğüm oranı 500:250'den 10:5'e doğru gidildikçe CART algoritmasına ilişkin risk tahmin değerlerinde (1.022'den 0.565'e) azalma kaydedilmiştir. CART algoritması için NYH değerleri ise 0.253'den 0.189'a, HKOK değerleri 1.011'dan 0.756'ya ve SS-oranının ise 0.719'dan 0.536'ya azaldığı görülmüştür. Bununla birlikte, belirleme katsayısı R^2 (%) değerleri % 48.667'den %71.621'e yükseldiği belirlenmiş ve sütten kesim ağırlığının gerçek ve tahmin değerleri arasındaki Pearson korelasyonların 0.697'den 0.844'e doğru yükseldiği saptanmıştır (P<0.01). Tüm sonuçlar değerlendirildiğinde, CART algoritması ile tahminlenen düğümler içi (hata) varyansın da azaldığı dikkat çekmiştir.

Çizelge 4.3. CART algoritması için hesaplanan uyum iyiliği ölçütlerine ait sonuçlar

Ebeveyn	Yavru	RT	R^2 (%)	NYH	HKOK	SS-Oranı	r	R^2 -düz. (%)	VK(%)
500	250	1.022	48.667	0.253	1.011	0.719	0.697	48.593	4.463
480	240	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
460	230	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
440	220	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
420	210	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
400	200	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
380	190	1.022	48.667	0.253	1.011	0.719	0.697	48.593	6.332
360	180	0.920	53.790	0.253	1.011	0.719	0.697	48.538	6.332
340	170	0.920	53.790	0.240	0.959	0.680	0.733	53.701	5.991
320	160	0.863	56.653	0.240	0.959	0.680	0.733	53.701	5.991
300	150	0.863	56.653	0.232	0.929	0.659	0.753	56.553	5.803
280	140	0.863	56.653	0.232	0.929	0.659	0.753	56.553	5.803
260	130	0.863	56.653	0.232	0.929	0.659	0.753	56.553	5.803
240	120	0.829	58.361	0.232	0.929	0.659	0.753	56.553	5.803
220	110	0.829	58.361	0.227	0.910	0.645	0.764	58.292	5.686

Çizelge 4.3. (devam)

Ebeveyn Düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS-Oranı	r	R^2 _{-düz.} (%)	VK(%)
200	100	0.802	59.717	0.227	0.910	0.645	0.764	58.292	5.686
180	90	0.802	59.717	0.224	0.895	0.635	0.773	59.622	5.593
160	80	0.743	62.681	0.224	0.895	0.635	0.773	59.622	5.593
140	70	0.732	63.233	0.215	0.862	0.611	0.792	62.575	5.385
120	60	0.712	64.238	0.214	0.855	0.606	0.795	63.150	5.343
100	50	0.693	65.192	0.211	0.844	0.598	0.802	64.154	5.270
80	40	0.672	66.247	0.208	0.832	0.590	0.807	65.099	5.200
60	30	0.637	68.005	0.205	0.820	0.581	0.814	66.166	5.120
40	20	0.582	70.767	0.199	0.798	0.566	0.825	67.919	4.983
20	10	0.572	71.270	0.191	0.763	0.541	0.841	70.676	4.764
10	5	0.565	71.621	0.189	0.756	0.536	0.844	71.169	4.724

Çizelge 4.4. Sütten kesim ağırlığı için CHAID algoritma sonuçları

Ebeveyn düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS-Oranı	r	R^2 _{-düz.} (%)	VK(%)
500	250	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
400	200	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
300	150	0.955	52.034	0.244	0.977	0.693	0.721	51.981	6.105
200	100	0.902	54.696	0.237	0.950	0.673	0.739	54.575	5.933
100	50	0.761	61.778	0.218	0.872	0.618	0.786	61.696	5.448

Çizelge 4.5. Sütten kesim ağırlığı için Geniş (Exhaustive) CHAID algoritma sonuçları

Ebeveyn* düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS-Oranı	r	R^2 _{-düz.} (%)	VK(%)
500	250	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
400	200	0.992	50.176	0.249	0.996	0.706	0.708	50.128	6.222
300	150	0.955	52.034	0.244	0.977	0.693	0.721	51.955	6.105
200	100	0.877	55.952	0.234	0.937	0.664	0.748	55.812	5.852
100	50	0.740	62.833	0.215	0.860	0.610	0.793	62.707	5.374

Çizelge 4.6. Sütten kesim ağırlığı için Pruning CART algoritma sonuçları

Ebeveyn* düğüm	Yavru düğüm	RT	R^2 (%)	NYH	HKOK	SS-Oranı	r	R^2 _{-düz.} (%)	VK(%)
500	250	1.022	48.669	0.253	1.011	0.717	0.697	48.595	6.316
400	200	1.022	48.669	0.253	1.011	0.717	0.697	48.595	6.316
300	150	0.863	56.655	0.232	0.929	0.659	0.753	56.555	5.804
200	100	0.829	58.363	0.227	0.910	0.645	0.764	58.294	5.686
100	50	0.704	64.641	0.209	0.839	0.595	0.804	64.527	5.243

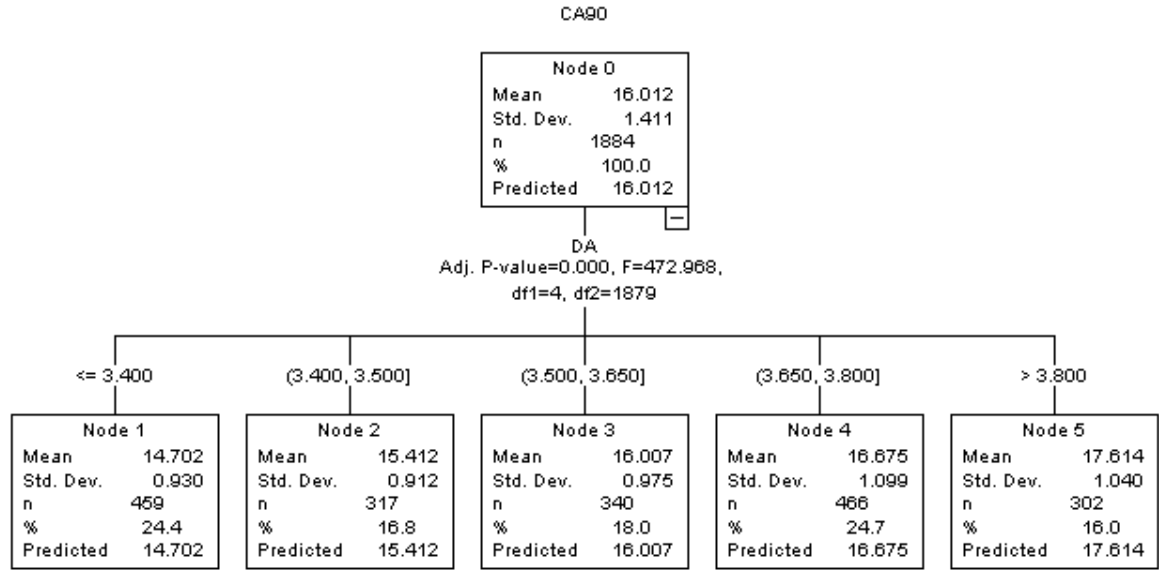
Atanılan farklı minimum ebeveyn-yavru düğüm oranlarının bulunduğu Çizelge 4.6 incelendiğinde, ebeveyn düğüm: yavru düğüm oranı 500:250'den 100:50'e doğru gidildikçe pruning CART algoritmasına ilişkin risk tahmin değerleri 1.022'den 0.704'e doğru azalma kaydedilmiştir. Pruning CART algoritması için NYH değerleri ise 0.253'den 0.209'a, RMS değerleri 1.011'dan 0.839'a azaldığı görülmüştür. Bununla birlikte, belirleme katsayısı değerleri % 48.669'den % 64.641'e yükseldiği belirlenmiş ve sütten kesim ağırlığının gerçek ve tahmin değerleri arasındaki Pearson korelasyonların 0.697'den 0.804'e doğru artma saptanmıştır ($P < 0.01$). Tüm sonuçlar değerlendirildiğinde, CART algoritmasından tahminlenen düğümler içi (hata) varyansların da azaldığı dikkat çekmiştir. Ayrıca pruning CART algoritması (Şekil 10) CART algoritmasıyla (Şekil 9) karşılaştırıldığında, aynı değişkenle ve aynı verilere göre değerlendirildiğinde çıkan analiz sonucunda önemli derecede bir sapma kaydedilmemiştir.

4. 2. CHAID Algoritmasına İlişkin Bulgular

Farklı minimum ebeveyn-çocuk düğüm oranları (500:250, 480:240, 460:230, 440:220, 420:210, 400:200 ve 380:190) için CHAID algoritması ile oluşturulan regresyon ağacı Şekil 4.1’de gösterilmiştir. Regresyon ağacının en üstünde bulunan düğüm kök düğüm (root node) olup, çalışmaya konu olan bütün kuzuları (n=1884) içine almaktadır. Sütten kesim ağırlığının genel ortalamasının 16.012 kg, standart sapmasının ise 1.411 kg olduğu belirlenmiştir. Regresyon ağacının en üstünde bulunan düğüm doğum ağırlığı bakımından 5 yavru düğüme (Düğüm 1, 2, 3, 4 ve 5) ayrılmıştır. Düğüm 1, kg cinsinden doğum ağırlığı $DA \leq 3.4$ kg’dan olan kuzuların oluşturduğu alt-grubu; Düğüm 2 doğum ağırlığı $3.4 < DA \leq 3.5$ olan kuzuların oluşturduğu alt-grubu; Düğüm 3, doğum ağırlığı $3.5 < DA \leq 3.650$ olan kuzuların oluşturduğu alt-grubu; Düğüm 4, doğum ağırlığı $3.650 < DA \leq 3.80$ olan kuzuların oluşturduğu alt-grubu ve Düğüm 5 ise $DA > 3.5$ olan kuzuların oluşturduğu alt-grubu temsil etmektedir. Regresyon ağacında Düğüm 1’den 5’e doğru gidildikçe sütten kesim ağırlığının yükseldiği (14.702-17.614kg) görülmektedir. Bu durum, doğum ağırlığı yüksek olan kuzuların sütten kesim ağırlığının yüksek olacağı anlamına gelmektedir. Mengali koyunlarının sütten kesim ağırlığı ortalama (standart sapma) değerleri sırasıyla, Düğüm 1 için 14.702 (S=0.930) kg, Düğüm 2 için 15.412 (S=0.912) kg, Düğüm 3 için 1.6007 (S=0.975) kg, Düğüm 4 için 16.675 (S=1.099) kg ve Düğüm 5 için 17.614 (S=1.040) kg olarak kaydedilmiştir. Toplamda 1884 koyunun, %24.4’ü Düğüm 1, %16.8’i Düğüm 2, %18’i Düğüm 3, %24.7’sini Düğüm 4 ve %16’sını Düğüm 5 oluşturmaktadır.

Sütten kesim ağırlığı bakımından gerçek ve tahmin değerleri arasındaki Pearson korelasyon katsayısı 0.708 bulunmuştur ($P < 0.01$). Her bir kuzu için hesaplanan hata değerlerine ait standart sapmanın, gerçek değerlere ait standart sapmaya oranı (SS oran) 0.706 bulunmuştur. Ancak kurulan bir modelin mükemmel bir uyuma sahip olabilmesi için SS oran değerinin 0-0.10 arasında olması gerekmektedir (Grezesiak ve Zaborski, 2012). Şekil 4.1 incelendiğinde, modele eklenen değişkenlerden sadece doğum ağırlığı kovaryetinin sütten kesim ağırlığı üzerinde etkili olduğu belirlenmiştir (düz- $P < 0.0000$, $F = 472.968$, $sd = 4$, $sd_2 = 1879$). Sütten kesim ağırlığına ait toplam varyasyonun %50.176’sı doğum ağırlığı

tarafından açıklanmaktadır. CHAID algoritması için üzerinde durulan özelliğe yani sütten kesim ağırlığına ait açıklanan varyasyon oranının artırılması, modele doğum ağırlığı dışında farklı değişkenlerin dahil edilebilmesine bağlıdır. Ancak bu durum, ebeveyn-yavru düğüm oranlarının azaltılması ile mümkündür (Çizelge 4.1 ve Şekil 4.1).



Şekil 4.1. 500:250, 480:240, 460:230, 440:220, 420:210:400:200 ve 380:190 Minimum ebeveyn-yavru düğüm oranları için CHAID algoritmasına ait regresyon ağacı diyagramı

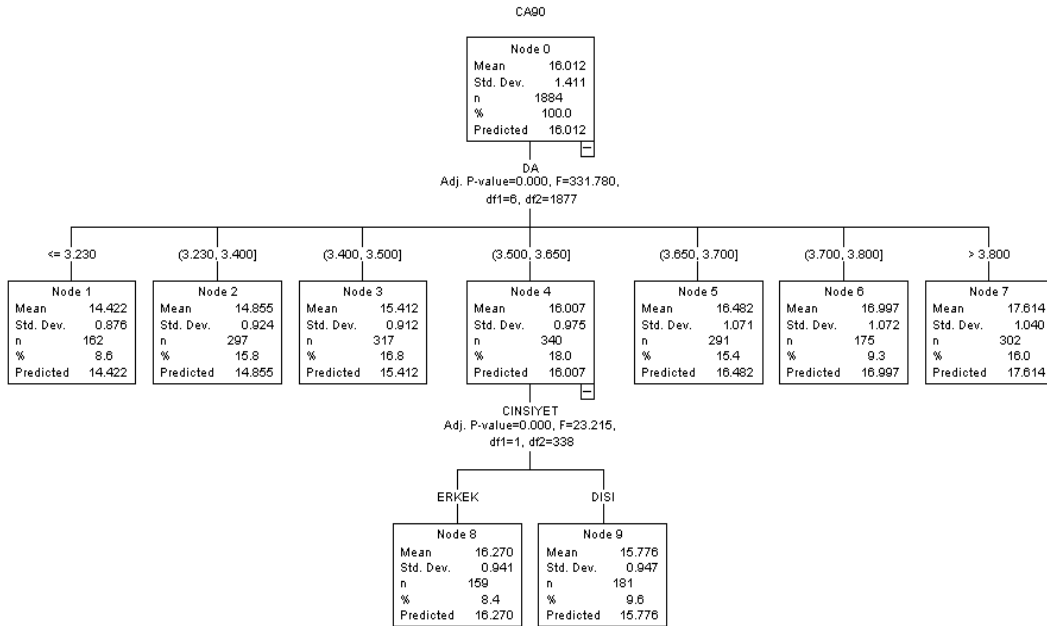
Minimum ebeveyn-yavru düğüm oranı 300:150 için CHAID algoritmasına ait regresyon ağacı Şekil 2’de gösterilmiştir. Regresyon ağacının en üstünde bulunan kök düğüm (root node) , çalışmaya konu olan bütün kuzuları (n=1884) içine almaktadır. Kök düğümde sütten kesim ağırlığının genel ortalamasının 16.012 kg, standart sapmasının ise 1.411 kg olduğu belirlenmiştir. Regresyon ağacının en üstünde bulunan düğüm doğum ağırlığı bakımından 7 yavru düğüme (Düğüm 1, 2, 3, 4, 5, 6 ve 7) ayrılmıştır. Düğüm 1, kg cinsinden doğum ağırlığı $DA \leq 3.230$ kuzuların oluşturduğu alt-grubu; Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların oluşturduğu alt-grubu; Düğüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların oluşturduğu alt-grubu;

Düğüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların oluşturduğu alt-grubu; Düğüm 5, doğum ağırlığı $3.650 < DA < 3.700$ olan kuzuların oluşturduğu alt-grubu; Düğüm 6, doğum ağırlığı $3.700 < DA < 3.800$ olan kuzuların oluşturduğu alt grubu ve Düğüm 7 ise, $DA > 3.800$ kg olan koyunların oluşturduğu alt grubu temsil etmektedir. Ayrıca Düğüm 4 ($3.500 < DA \leq 3.650$) cinsiyet faktörü bakımından iki yavru düğüme (Düğüm 8 ve Düğüm 9) ayrılmıştır. Regresyon ağacında Düğüm 1'den 7'ye doğru gidildikçe sütten kesim ağırlığının yükseldiği (14.422-17.614kg) görülmektedir. Bu durum, doğum ağırlığı yüksek olan kuzuların sütten kesim ağırlığının yüksek olacağı anlamına gelmektedir. Mengali koyunlarının sütten kesim ağırlığı ortalama (standart sapma) değerleri sırasıyla, Düğüm 1 için 14.422 (S=0.876) kg, Düğüm 2 için 14.855 (S=0.924)kg, Düğüm 3 için 15.412 (S=0.912) kg, Düğüm 4 için 16.007(S=0.975) kg, Düğüm 5 için 16.482 (S=1.071) kg, Düğüm 6 için 16.997 (S=1.072) kg ve Düğüm 7 için 17.614 (S=1.040) kg olarak kaydedilmiştir. Toplamda 1884 koyunun, % 8.6'sı Düğüm 1, % 15.8'i Düğüm 2, % 16.8'i Düğüm 3, % 18'sini Düğüm 4, % 15.4'ünü Düğüm 5, % 9.3'ünü Düğüm 6 ve %16'sını Düğüm 7 oluşturmaktadır.

İncelenen 1884 kuzu içerisinde doğum ağırlığı $3.500 < DA < 3.650$ olan kuzuların (Düğüm 4) sütten kesim ağırlığı cinsiyet faktöründen etkilenmektedir (düz-P=0000, F=23.215, sd=1, sd2=338). Düğüm 4'ü oluşturan kuzular cinsiyet faktörü tarafından düğüm 8 ve 9 olmak üzere iki yeni yavru düğüme ayrılmıştır. Buna ek olarak, doğum ağırlığı $3.500 < DA < 3.650$ olan erkek kuzuların (düğüm 8) sütten kesim ağırlığı ortalaması, $3.500 < DA < 3.650$ olan dişi kuzuların sütten kesim ağırlığı ortalamasından yüksek bulunmuştur. Şekil 4.2 incelendiğinde, 1, 2, 3, 5, 6, 7, 8 ve 9 nolu düğümlerin terminal düğüm olduğu belirlenmiştir. CHAID algoritmasından türetilen regresyon ağacında terminal düğümlerde yeterince homojenliğe ulaşılmış, sonraki aşamalarda başka dallanmalar olmamıştır.

CHAID algoritmasının kullanılması ile her bir kuzunun sütten kesim ağırlığı için tahmin değeri hesaplanmış, sütten kesim ağırlığı bakımından gerçek ve tahmin değerleri arasındaki Pearson korelasyon katsayısı 0.721 bulunmuştur (P < 0.01). Her bir kuzu için hesaplanan hata değerlerine ait standart sapmanın, gerçek değerlere ait standart sapmaya

oranı (SS oran) 0.693 bulunmuştur. Veri madenciliği çalışmalarında, bir modelin mükemmel bir uyuma sahip olabilmesi için SS oran değerinin 0-0.10 arasında olması gerekmektedir (Grezesiak ve Zaborski, 2012). Şekil 4.2 incelendiğinde, modele eklenen değişkenlerden sadece doğum ağırlığı kovaryetinin (düz-P=0000, F=331.780, sd1=4, sd2=1877) ve cinsiyet faktörünün (Adj-P<0000, F=23.215, sd1=1, sd2=338) sütten kesim ağırlığı üzerinde etkili olduğu belirlenmiştir. Sütten kesim ağırlığına ait toplam varyasyonun %52.034'sı doğum ağırlığı ve cinsiyet tarafından açıklanmaktadır. CHAID algoritması için üzerinde durulan özelliğe yani sütten kesim ağırlığına ait açıklanan varyasyon oranının artırılması, modele doğum ağırlığı ve cinsiyet dışında farklı değişkenlerin modelde olması ile ilişkilidir. Bilindiği üzere açıklanan varyasyon payının artırılması, minimum ebeveyn-yavru düğüm oranlarının azaltılması ile mümkün olacaktır (Çizelge 4.1 ve Şekil 4.2).



Şekil 4.2. 300:150 Minimum ebeveyn-yavru düğüm oranları için CHAID algoritmasına ait regresyon ağacı diyagramı

Sütten kesim ağırlığını tahmin etmek için ayarlanan minimum ebeveyn-yavru düğüm oranı 200:100 için CHAID algoritmasına ait regresyon ağacı Şekil 4.3’de gösterilmiştir. Kök düğümde sütten kesim ağırlığının genel ortalamasının 16.012 kg, standart sapmasının ise 1.411 kg olduğu belirlenmiştir. Regresyon ağacının en üstünde bulunan düğüm doğum ağırlığı bakımından 8 yavru düğüme (Düğüm 1, 2, 3, 4, 5, 6, 7 ve 8) ayrılmıştır. Düğüm 1, kg cinsinden doğum ağırlığı $DA \leq 3.230$ kuzuların oluşturduğu alt-grubu; Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların oluşturduğu alt-grubu; Düğüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların oluşturduğu alt-grubu; Düğüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların oluşturduğu alt-grubu; Düğüm 5, doğum ağırlığı $3.650 < DA < 3.700$ olan kuzuların oluşturduğu alt-grubu; Düğüm 6, doğum ağırlığı $3.700 < DA < 3.800$ olan kuzuların oluşturduğu alt grubu; Düğüm 7, doğum ağırlığı $3.800 < DA < 3.900$ olan kuzuların oluşturduğu alt grubu ve Düğüm 8 doğum ağırlığı ise $DA > 3.900$ olan kuzuların oluşturduğu grubu temsil etmektedir. Ayrıca, Düğüm 2 ($3.230 < DA < 3.400$) ana yaşı faktörü bakımından iki yavru düğüme (Düğüm 9, Düğüm 10), Düğüm 3 ve 4 cinsiyet faktörü bakımından iki yavru düğüme (Düğüm 11, Düğüm 12) ve düğüm 5 ise yıl faktöründen etkilenerek iki yavru düğüme (Düğüm 15, Düğüm 16) ayrılmışlardır. Regresyon ağacında Düğüm 1’den 9’a doğru gidildikçe sütten kesim ağırlığının yükseldiği (14.422-17.718kg) görülmektedir. Bu durum, daha ağır doğum ağırlığı sahip olan kuzuların sütten kesim ağırlığı bakımından daha ağır olacağı anlamına gelmektedir. Mengali koyunlarının sütten kesim ağırlığı ortalama (standart sapma) değerleri sırasıyla, Düğüm 1 için 14.422 (S=0.876) kg, Düğüm 2 için 14.855 (S=0.924)kg, Düğüm 3 için 15.412 (S=0.912) kg, Düğüm 4 için 16.007(S=0.975) kg, Düğüm 5 için 16.482 (S=1.071) kg, Düğüm 6 için 16.997 (S=1.072) kg, Düğüm 7 için 17.476 (S=1.020) kg ve Düğüm 8 için ise 17.718 (S=1.046) olarak kaydedilmiştir. Toplamda 1884 koyunun, %8.6’sı Düğüm 1, %15.8’i Düğüm 2, %16.8’i Düğüm 3, %18.0’ini Düğüm 4,%’ünü 15.4 Düğüm 5, % 9.3’ünü Düğüm 6, % 6.9’unu Düğüm 7 ve %9.1’ni ise Düğüm 8 oluşturmaktadır.

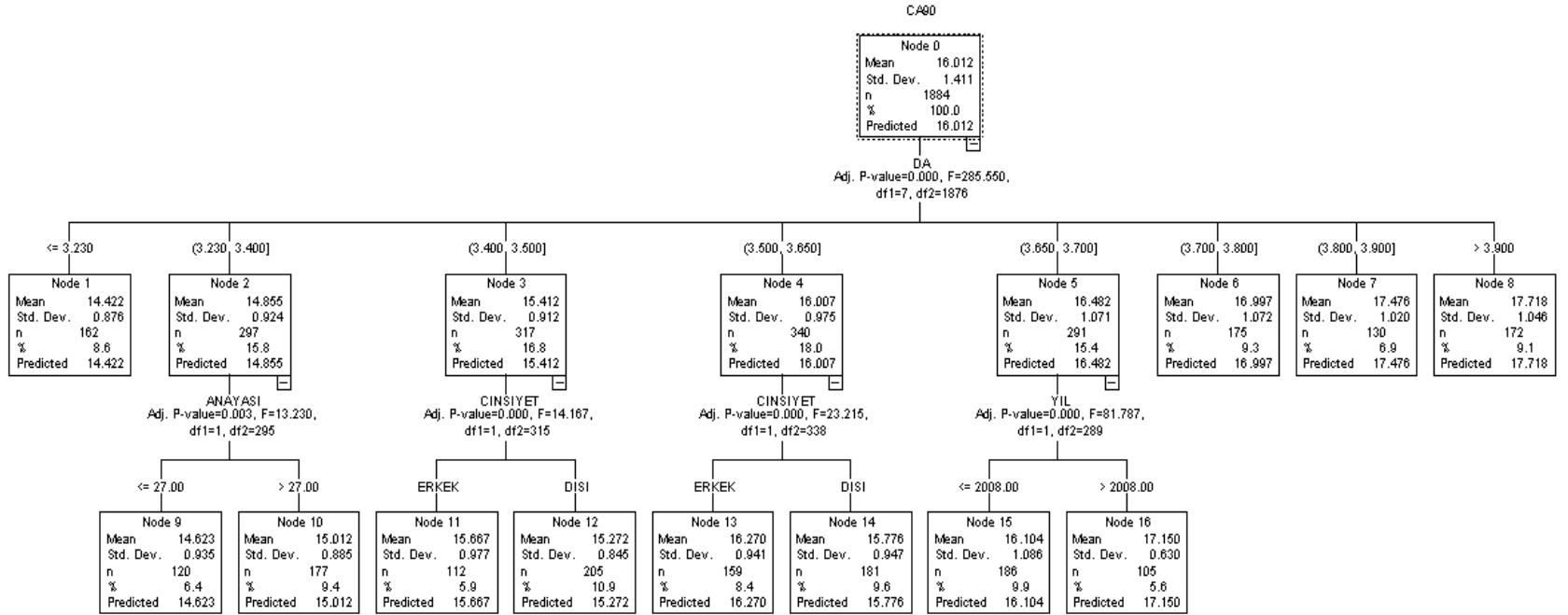
Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların sütten kesim ağırlığını ana yaşı kovaryeti etkilemiştir. (düz-P<0.003, F=13.230, sd1=1, sd2=295). Ana yaşı $AY \leq 27$

olan analardan doğan ve doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların ortalama sütten kesim ağırlığı 14.623 (S=0.935) kg, ana yaşı $AY > 27$ olan analardan doğan ve doğum ağırlığı $3.230 < DA \leq 3.400$ arasında olan kuzuların ortalama sütten kesim ağırlığı ise 15.012 (S=0.885) kg olarak tahmin edilmiştir. Doğum ağırlığı, $3.400 < DA < 3.500$ kg olan kuzuların sütten kesim ağırlığını cinsiyet faktörü etkilemiş, ve Dügüm 3 iki alt yavru düğüme (Dügüm 11, Dügüm 12) ayrılmıştır (düz-P=0.000, F=14.167, sd1=1, sd2=315). Cinsiyeti erkek olan kuzuların ortalama sütten kesim ağırlığı 15.667 (S=0.977) kg, cinsiyeti dişi olan kuzuların ortalama sütten kesim ağırlığı 15.272 (S=0.846) kg olarak tahmin edilmiştir (Dügüm 11 ve Dügüm 12) Dügüm 4, $3.500 < \text{doğum ağırlığı} \leq 3.650$ olan kuzuların sütten kesim ağırlığını cinsiyet faktörün etkilediği belirlenmiştir (düz-P<0.000, F=23.215, sd1=1, sd2=338). Dügüm 4, cinsiyet faktörü bakımından iki yeni yavru düğüme (dügüm 13 ve 14) ayrılmış, doğum ağırlığı $3.500 < DA \leq 3.650$ olan erkek kuzuların ortalama sütten kesim ağırlığının 16.270 (S=0.941) kg, cinsiyeti dişi olan kuzulardan 15.776 (S=0.947) kg daha ağır olduğu tahmin edilmiştir. Ayrıca Dügüm 5, doğum yılı faktörü bakımından iki yeni yavru düğüme (Dügüm 15 ve Dügüm 16) ayrılmıştır. Doğum ağırlığı $3.650 < DA \leq 3.700$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkilemiştir. (Düz-P<0.000, F=81.787, sd1=1, sd2=289). Doğum yılı $DOĞ YIL. \leq 2008$ olan kuzuların ortalama sütten kesim ağırlığı 14.104 (S=1.086) kg, doğum yılı $DOĞ YIL > 2008$ olan kuzuların ortalama sütten kesim ağırlığı 17.150 (S=0.630) kg olarak tahmin edilmiştir. Bu verilere göre doğum yılı 2008 yılından büyük olan kuzuların doğum ağırlığı daha fazla olduğu söylenebilir.

Şekil 4.3 incelendiğinde, 1, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 ve 16. düğümler terminal düğümlerdir. CHAID algoritmasından türetilen regresyon ağacında terminal düğümlerde yeterince homojenliğe ulaşıldığından dolayı, sonraki aşamalarda başka dallanmalar olmamıştır.

CHAID algoritmasının kullanılması ile her bir kuzunun sütten kesim ağırlığı için tahmin değeri hesaplanmış, sütten kesim ağırlığı bakımından gerçek ve tahmin değerleri arasındaki Pearson korelasyon katsayısı 0.739 bulunmuştur (P<0.01). Her bir kuzu için hesaplanan hata değerlerine ait standart sapmanın, gerçek değerlere ait standart sapmaya

oranı (SS oran) 0.673 bulunmuştur. Veri madenciliği çalışmalarında, özelleştirilen bir modelin mükemmel bir uyuma sahip olabilmesi için SS oran değerinin 0-0.10 arasında olması gerekmektedir (Grezesiak ve Zaborski, 2012). Şekil 4.3'e bakıldığında, modele dahil edilen değişkenlerden doğum ağırlığı kovaryetinin (Düz-P<0000, F=295.550, sd1=7, sd2=1876), ana yaşı faktörü (Düz-P<0.003, F=13.230, sd1=1, sd2=295) cinsiyet faktörü (Düz-P<0000, F=14.167, sd1=1, sd2=315) ve (Düz-P<0000, F=23.215, sd1=1, sd2=338) ve yıl faktörünün (Düz-P<0000, F=81.787, sd1=1, sd2=289) sütten kesim ağırlığı üzerinde etkili olduğu belirlenmiştir. Sütten kesim ağırlığına ait toplam varyasyonun %54.695'i doğum ağırlığı, ana yaşı, cinsiyet ve yıl faktörleri tarafından açıklanmaktadır. CHAID algoritması için üzerinde durulan özelliğe yani sütten kesim ağırlığına ait açıklanan varyasyon oranının artırılması, modele doğum ağırlığı, ana yaşı, cinsiyet ve yıl dışında farklı değişkenlerin modelde olması ile ilişkilidir (Çizelge 4.1 ve Şekil 4.3).



Şekil 4.3. 200:100 Minimum ebeveyn-yavru düğüm oranı için CHAID algoritmasına ait regresyon ağacı diyagramı

Sütten kesim ağırlığını tahmin etmek için ayarlanan minimum ebeveyn-yavru düğüm oranı 100:50 için CHAID algoritması ile oluşturulan regresyon ağacı Şekil 4.4’de gösterilmiştir. Kök düğümde sütten kesim ağırlığına ait genel ortalamanın 16.012 kg, standart sapmanın ise 1.411 kg olduğu belirlenmiştir. Regresyon ağacının en üstünde bulunan düğüm doğum ağırlığı bakımından 8 yavru düğümüne (Düğüm 1, 2, 3, 4, 5, 6, 7 ve 8) ayrılmıştır. Düğüm 1, kg cinsinden doğum ağırlığı $DA \leq 3.230$ kuzuların oluşturduğu alt-grubu; Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların oluşturduğu alt-grubu; Düğüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların oluşturduğu alt-grubu; Düğüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların oluşturduğu alt-grubu; Düğüm 5, doğum ağırlığı $3.650 < DA \leq 3.700$ olan kuzuların oluşturduğu alt-grubu; Düğüm 6, doğum ağırlığı $3.700 < DA \leq 3.800$ olan kuzuların oluşturduğu alt grubu; Düğüm 7, doğum ağırlığı $3.800 < DA \leq 3.900$ olan kuzuların oluşturduğu alt grubu ve Düğüm 8 doğum ağırlığı $DA > 3.900$ olan kuzuların oluşturduğu grubu temsil etmektedir. Ayrıca Düğüm 1, 2, 3, 4 ve 5 yıl faktörü bakımından ikiyeşer yavru düğümüne (Düğüm 9-22) ayrılmıştır. Düğüm 6 ve 8 cinsiyet faktörü bakımından iki yavru düğümüne (Düğüm 23, Düğüm 24, Düğüm 25, Düğüm 26) ayrılmışlardır. Regresyon ağacının ilk derinliğinde Düğüm 1’den 8’e doğru gidildikçe sütten kesim ağırlığının yükseldiği (14.422–17.718 kg) kaydedilmektedir. Bununla birlikte regresyon ağacında ikinci derinlikte de Düğüm 9’dan 26’ya gidildikçe sütten kesim ağırlığının yükseldiği (14.148-17.254 kg) gözlemlenmektedir. Bu durum, doğum ağırlığı yüksek olan kuzuların sütten kesim ağırlığının yüksek olacağı anlamına gelmektedir. Mengali koyunlarının sütten kesim ağırlığı ortalama (standart sapma) değerleri sırasıyla, Düğüm 1 için 14.422 (S=0,876) kg, Düğüm 2 için 14.855 (S=0.924) kg, Düğüm 3 için 15.412 (S=0.912) kg, Düğüm 4 için 16.007 (S=0.975) kg, Düğüm 5 için 16.482 (S=1.071) kg, Düğüm 6 için 16.997 (S=1.072) kg, Düğüm 7 için 17.476 (S=1.020) kg ve Düğüm 8 için ise 17.718 (S=1.046) olarak kaydedilmiştir. İncelenen 1884 kuzunun, %8.6’sı Düğüm 1, %15.8’i Düğüm 2, %16.8’i Düğüm 3, %18.0’ini Düğüm 4,% 15.4 Düğüm 5, % 9.3’ünü Düğüm 6, % 6.9’unu Düğüm 7 ve %9.1’ni Düğüm 8 oluşturmaktadır.

Düğüm 1, $DA \leq 3.230$ doğum ağırlığı olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek iki yavru düğüme (Düğüm 9, Düğüm 10) ayrılmıştır (Düz-P<0.014, F=8.756, sd1=1, sd2=160). Doğum yılı 2007 ve 2007 yılından küçük olan kuzuların (Düğüm 9) sütten kesim ağırlığı 14.148 (S=0.876) kg olarak tahmin edilmiştir. Doğum yılı 2007 yılından büyük olan kuzuların (Düğüm 10) sütten kesim ağırlığı 14.556 (S=0.963) kg olarak kaydedilmiştir. Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek üç yavru düğüme (Düğüm 11, Düğüm 12, Düğüm 13) ayrılmıştır. (Düz-P<0.000, F=46.692, sd1=2, sd2=294). Doğum yılı $DOĞ\ YIL \leq 2005$ olan kuzuların (Düğüm 11) sütten kesim ağırlığı 14.129 (S=0.719) kg olarak tahmin edilmiştir. Doğum yılı 2005 ve 2008 yılları arasında olan kuzuların (Düğüm 12) sütten kesim ağırlığı 14.775 (S=0.951) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ\ YIL > 2008$ olan kuzuların (Düğüm 13) sütten kesim ağırlığı 15.420 (S=0.586) kg olarak tahmin edilmiştir. Düğüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek üç yavru düğüme (Düğüm 14, Düğüm 15, Düğüm 16) ayrılmıştır (Düz-P<0.000, F=60.744, sd1=2, sd2=314). Doğum yılı $DOĞ\ YIL \leq 2005$ olan kuzuların (Düğüm 14) sütten kesim ağırlığı 14.782 (S=0.587) kg olarak tahmin edilmiştir. Doğum yılı 2005 ve 2008 yılları arasında olan kuzuların (Düğüm 15) sütten kesim ağırlığı 15.218 (S=0.937) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ\ YIL \geq 2008$ olan kuzuların (Düğüm 16) sütten kesim ağırlığı 16.107 (S=0.506) kg olarak tahmin edilmiştir. Düğüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek üç yavru düğüme (Düğüm 17, Düğüm 18, Düğüm 19) ayrılmıştır. (Düz-P<0.000, F=42.240, sd1=2, sd2=337). Doğum yılı $DOĞ\ YIL \leq 2006$ olan kuzuların (Düğüm 17) sütten kesim ağırlığı 15.791 (S=0.838) kg olarak tahmin edilmiştir. Doğum yılı 2006 ve 2008 yılları arasında olan kuzuların (Düğüm 18) sütten kesim ağırlığı 15.707 (S=1.055) kg olarak tahmin edilmiştir. Doğum yılı 2008 ve 2008 yılından büyük olan kuzuların (Düğüm 19) sütten kesim ağırlığı 16.710 (S=0.501) kg olarak tahmin edilmiştir. Düğüm 5, doğum ağırlığı $3.650 < DA \leq 3.700$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek üç yavru düğüme (Düğüm 20, Düğüm 21, Düğüm 22) ayrılmıştır. (Düz-P<0.000, F=42.172, sd1=2, sd2=288). Doğum yılı $DOĞ\ YIL \leq 2007$ olan kuzuların (Düğüm 20) sütten kesim ağırlığı 16.024 (S=1.103) kg

olarak tahmin edilmiştir. Doğum yılı 2007 ve 2008 yılları arasında doğan kuzuların (Düğüm 21) sütten kesim ağırlığı 16.237 (S=1.051) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ YIL \leq 2008$ olan kuzuların (Düğüm 22) sütten kesim ağırlığı 17.150 (S=0.630) kg olarak tahmin edilmiştir. Düğüm 6, doğum ağırlığı $3.700 < DA \leq 3.300$ cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm23, Düğüm 24) ayrılmıştır. (Düz-P<0.000, F=54.478, sd1=1, sd2=173). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.385 (S=0.782) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 16.290(S=1.171) kg olarak tahmin edilmiştir. Burada erkekler dişilere oranla daha fazla ağırlığa sahip olduğu saptanmıştır. Ayrıca Düğüm 8 ise, doğum ağırlığı $DA > 3.900$ cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm25ve Düğüm 26) ayrılmıştır. (Düz-P<0.000, F=23,143 sd1=1, sd2=170). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.999 (S=0.840) kg olarak tahmin edilmiş, Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 17.254(S=1.186) kg olarak tahmin edilmiştir. Burada da erkekler dişilere oranla daha fazla ağırlığa sahip olduğu belirlenmiştir. İkinci derecede derinlikte bulunan Düğüm 15 işletme faktörü tarafından etkilenecek iki yavru düğüme (Düğüm27 ve Düğüm 28) ve Düğüm 18 ise cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 29 ve Düğüm 30) ayrılmışlardır. 2006 ve 2008 yılları arasında doğan ve doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların sütten kesim ağırlığı işletme faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 27, Düğüm 28) ayrıldığı gözlenmiştir. (Düz-P<0.000, F=27.623, sd1=1, sd2=167). işletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 15.725 (S=0.812) kg olarak tahmin edilmiştir. İşletmesi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.973 (S=0.898) kg olarak tahmin edilmiştir. Mastung işletmesi Noshki işletmesine göre kuzuların canlı ağırlığı daha fazla olduğu tahmin edilmiştir. 2007 ve 2008 yılları arasında doğan ve doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların sütten kesim ağırlığı cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 29 ve Düğüm 30) ayrıldığı gözlenmiştir. (Düz-P<0.002, F=10.208, sd1=1, sd2=152). Cinsiyeti erkek olan kuzuların doğum ağırlığı 16.020 (S=1.009) kg ve cinsiyeti dişi olan kuzuların doğum ağırlığı 15.485 (S=1.035) kg olarak kaydedilmiştir. Erkekler doğum ağırlığı bakımından dişilerle karşılaştırıldığında daha fazla doğum ağırlığına sahip olduğu gözlenmiştir.

Şekil 4.4 incelendiğinde, 7, 9, 10, 11, 12, 13, 14, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ve 30'uncu düğümler terminal düğümlerdir. CHAID algoritmasından türetilen regresyon ağacında terminal düğümlerde yeterince homojenliğe ulaşılmış, sonraki aşamalarda başka dallanmalar olmamıştır.

CHAID algoritmasının kullanılması ile her bir kuzunun sütten kesim ağırlığı için tahmin değeri hesaplanmış, sütten kesim ağırlığı bakımından gerçek ve tahmin değerleri arasındaki Pearson korelasyon katsayısı 0.786 bulunmuştur ($P < 0.01$). Her bir kuzu için hesaplanan hata değerlerine ait standart sapmanın, gerçek değerlere ait standart sapmaya oranı (SS oran) 0.618 bulunmuştur. Veri madenciliği çalışmalarında, bir modelin mükemmel bir uyuma sahip olabilmesi için SS oran değerinin 0-0.10 arasında olması gerekmektedir (Grezesiak ve Zaborski, 2012). Sütten kesim ağırlığına ait toplam varyasyonun % 61.778'i doğum ağırlığı, yıl, cinsiyet ve işletme gibi değişkenler tarafından açıklanmaktadır. CHAID algoritması için üzerinde durulan özelliğe yani sütten kesim ağırlığına ait açıklanan varyasyon oranının artırılması, modele doğum ağırlığı, yıl, cinsiyet ve işletme dışında farklı değişkenlerin modelde olması ile ilişkilidir. Açıklanan varyasyon payının artırılması, ebeveyn-yavru düğüm oranlarının azaltılması ile mümkün olabilmektedir (Çizelge 4.1 ve Şekil 4.4).

Sütten kesim ağırlığını tahmin etmek için ayarlanan ebeveyn-yavru düğüm oranı 10:5 için CHAID algoritmasına ait regresyon ağacı analizi yapılmıştır(şekil çok büyük olduğundan gösterilmemiştir). Kök düğümde Sütten kesim ağırlığının genel ortalamasının 16.012 kg, standart sapmasının ise 1.411 kg olduğu belirlenmiştir. Regresyon ağacının en üstünde bulunan düğüm doğum ağırlığı bakımından 8 yavru düğüme (Düğüm 1, 2, 3, 4, 5, 6, 7 ve 8) ayrılmıştır. Düğüm 1, kg cinsinden doğum ağırlığı, $DA \leq 3.230$ olan kuzuların oluşturduğu alt-grubu; Düğüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların oluşturduğu alt-grubu; Düğüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların oluşturduğu alt-grubu; Düğüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların oluşturduğu alt-grubu; Düğüm 5, doğum ağırlığı $3.650 < DA \leq 3.700$ olan kuzuların oluşturduğu alt-grubu; Düğüm 6, doğum ağırlığı $3.700 < DA \leq 403.800$ olan kuzuların oluşturduğu alt grubu; Düğüm 7, doğum ağırlığı $3.800 < DA \leq 3.900$ olan kuzuların oluşturduğu alt grubu ve Düğüm 8 de ise $DA > 3.900$ den büyük olan kuzuların oluşturduğu grubu temsil etmektedir. Ayrıca Düğüm 1, işletme faktörü bakımında iki yavru düğüme (Düğüm 9, Düğüm 10), düğüm 2, dört yavru düğüme (Düğüm 11, Düğüm 12, Düğüm 13, Düğüm 14), düğüm 3, üç yavru düğüme (Düğüm 15, Düğüm 16, Düğüm 17), düğüm 4, dört yavru düğüme (Düğüm 18, Düğüm 19, Düğüm 20, Düğüm 21), düğüm 5, beş yavru düğüme (Düğüm 22, Düğüm 23, Düğüm 24, Düğüm 25, Düğüm 26), düğüm 6, 7 ve 8 ise ikişer yavru düğüme ayrılmıştır(Düğüm 27, Düğüm 28, Düğüm 29, Düğüm 30, Düğüm 31, Düğüm 32). Regresyon ağacının ilk derinliğinde Düğüm 1'den 8'e doğru gidildikçe sütten kesim ağırlığının yükseldiği (14.422-17.718 kg) kaydedilmektedir. Bununla birlikte regresyon ağacında ikinci derinlikte de Düğüm 9'dan 32'ye gidildikçe sütten kesim ağırlığının yükseldiği (14.148-17.254 kg) gözlemlenmektedir. Bu durum, doğum ağırlığı yüksek olan kuzuların sütten kesim ağırlığının yüksek olacağı anlamına gelmektedir. Mengali kuzularının sütten kesim ağırlığı ortalama (standart sapma) değerleri sırasıyla, Düğüm 1 için 14.422 (S=0.876) kg, Düğüm 2 için 14.855 (S=0.924)kg, Düğüm 3 için 14.412 (S=0.912) kg, Düğüm 4 için 16.007(S=0.975) kg, Düğüm 5 için 16.482 (S=1.071) kg, Düğüm 6 için 16.997 (S=1.072) kg, Düğüm 7 için 17.476 (S=1.020) kg ve Düğüm 8 için ise 17.710 (S=1.048) olarak kaydedilmiştir. Toplamda 1884 koyunun, %8.6'sı Düğüm 1, %15.8'i Düğüm 2, %16.8'i Düğüm 3, %18.0'ini Düğüm 4,% 15.4 Düğüm 5, % 9.3'ünü Düğüm 6, % 6.9'unu Düğüm 7 ve

%9.1'ni Dügüm 8 oluşturmaktadır. Dügüm 1, $DA \leq 3.230$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek iki yavru düğüme (Dügüm 9, Dügüm 10) ayrılmıştır. (Düz- $P < 0.014$, $F=8.756$, $sd1=1$, $sd2=160$). Doğum yılı, $DOĞ YIL \leq 2007$ olan kuzuların (Dügüm 9) sütten kesim ağırlığı 14.148 ($S=0.599$) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ YIL > 2007$ olan kuzuların (Dügüm 10) sütten kesim ağırlığı 14.556 ($S=0.963$) kg olarak kaydedilmiştir. Dügüm 2, doğum ağırlığı $3.230 < DA \leq 3.400$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek dört yavru düğüme (Dügüm 11, Dügüm 12, Dügüm 13, Dügüm 14) ayrılmıştır. (Düz- $P < 0.000$, $F=38.622$, $sd1=3$, $sd2=293$). Doğum yılı $DOĞ YIL \leq$

2005 olan kuzuların (Dügüm 11) sütten kesim ağırlığı 14.129 ($S=0.719$) kg olarak tahmin edilmiştir. Doğum yılı 2005 ve 2006 yılı dahil olan kuzuların (Dügüm 12) sütten kesim ağırlığı $15,162$ ($S=0,841$) kg olarak tahmin edilmiştir. Doğum yılı 2006 ve 2008 yılı dahil kuzuların (Dügüm 13) sütten kesim ağırlığı 14.585 ($S=0.948$) kg olarak tahmin edilmiştir, doğum yılı $DOĞ YIL > 2008$ olan kuzuların (Dügüm 14) sütten kesim ağırlığı 15.420 ($S=0.586$) kg olarak kaydedilmiştir. Dügüm 3, doğum ağırlığı $3.400 < DA \leq 3.500$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek üç yavru düğüme (Dügüm 15, Dügüm 16, Dügüm 17) ayrılmıştır. (Düz- $P < 0.000$, $F=60.744$, $sd1=2$, $sd2=314$). Doğum yılı $DOĞ YIL \leq 2005$ olan kuzuların (Dügüm 15) sütten kesim ağırlığı 14.782 ($S=0.587$) kg olarak tahmin edilmiştir. Doğum yılı $2005 < DOĞ YIL < 2008$ olan kuzuların (Dügüm 16) sütten kesim ağırlığı 15.218 ($S=0.937$) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ YIL \geq 2008$ kuzuların (Dügüm 17) sütten kesim ağırlığı 16.107 ($S=0.506$) kg olarak tahmin edilmiştir. Dügüm 4, doğum ağırlığı $3.500 < DA \leq 3.650$ olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek dört yavru düğüme (Dügüm 18, Dügüm 19, Dügüm 20) ayrılmıştır. (Düz- $P < 0.000$, $F=34.727$, $sd1=3$, $sd2=336$). Doğum yılı $DOĞ YIL \leq 2005$ olan kuzuların (Dügüm 18) sütten kesim ağırlığı 15.394 ($S=0.716$) kg olarak tahmin edilmiştir. Doğum yılı $2005 < DOĞ YIL < 2006$ olan kuzuların (Dügüm 19) sütten kesim ağırlığı 16.111 ($S=0.796$) kg olarak tahmin edilmiştir. Doğum yılı $2006 < DOĞ YIL \leq 2008$ olan kuzuların (Dügüm 20) sütten kesim ağırlığı 15.707 ($S=1.055$) kg olarak tahmin edilmiştir. Doğum yılı $DOĞ YIL > 2008$ olan kuzuların (Dügüm 21) sütten kesim ağırlığı 16.710 ($S=0.501$) kg olarak kaydedilmiştir. Dügüm 5, doğum ağırlığı

3.650 < DA ≤ 3.700 olan kuzuların sütten kesim ağırlığını yıl faktörü etkileyerek beş yavru düğüme (Düğüm 22, Düğüm 23, Düğüm 24, Düğüm 25, Düğüm 26) ayrılmıştır. (Düz-P < 0.000, F=27.329, sd1=2, sd2=286). Doğum yılı DOĞ YIL ≤ 2005 olan kuzuların (Düğüm 22) sütten kesim ağırlığı 15.715 (S=0.852) kg olarak tahmin edilmiştir. Doğum yılı 2005 < DOĞ YIL < 2006 yılları arasında doğan kuzuların (Düğüm 23) sütten kesim ağırlığı 16.633 (S=0.797) kg olarak tahmin edilmiştir. Doğum yılı 2006 < DOĞ YIL < 2007 yılları arasında olan kuzuların (Düğüm 24) sütten kesim ağırlığı 15.829 (S=1.230) kg olarak tahmin edilmiştir. Doğum yılı 2007 < DOĞ YIL < 2008 yılları arasında olan kuzuların (Düğüm 25) sütten kesim ağırlığı 16.237 (S=1.051) kg olarak tahmin edilmiştir. Doğum yılı DOĞ YIL ≥ 2008 olan kuzuların (Düğüm 26) sütten kesim ağırlığı 17.150 (S=0.630) kg olarak tahmin edilmiştir. Düğüm 6, doğum ağırlığı 3.700 < DA ≤ 3.800 cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 27, Düğüm 28) ayrılmıştır. (Düz-P < 0.000, F=54.478, sd1=1, sd2=173). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.385 (S=0.782) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 16.290 (S=1.171) kg olarak tahmin edilmiştir. Burada erkekler dişilere oranla daha fazla ağırlığa sahip olduğu saptanmıştır. Düğüm 7, doğum ağırlığı 3.800 < DA ≤ 3.900 cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 29, Düğüm 30) ayrılmıştır. (Düz-P < 0.000, F=38.223, sd1=1, sd2=128). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.824 (S=0.805) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 16.795 (S=1.059) kg olarak tahmin edilmiştir. Burada erkekler dişilere oranla daha fazla ağırlığa sahip olduğu saptanmıştır. Ayrıca Düğüm 8, doğum ağırlığı DA > 3.900 olan kuzuları cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 31, Düğüm 32) ayrılmıştır. (Düz-P < 0.000, F=23,143 sd1=1, sd2=170). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.999 (S=0.840) kg olarak tahmin edilmiş, Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 17.254 (S=1.186) kg olarak tahmin edilmiştir. Burada da erkekler dişilere oranla daha fazla ağırlığa sahip olduğu belirlenmiştir. İkinci derecede derinlikte bulunan Düğüm 9 işletme faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 33, Düğüm 34) ayrılmıştır. (Düz-P < 0.004, F=13.454, sd1=1, sd2=54). işletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.518 (S=0.590) kg olarak tahmin edilmiştir. işletmesi Quetta olan işletmedeki kuzuların doğum ağırlığı 13.958 (S=0.514) kg olarak tahmin edilmiştir. Noshki işletmesi Quetta işletmesine göre

kuzuların canlı ağırlığı daha fazla olduğu tahmin edilmiştir. Dügüm 10 işletme faktörü tarafından etkilenecek iki yavru düğüme (Dügüm35, Dügüm 36) ayrılmıştır. (Düz-P<0.010, F=10.807, sd1=1, sd2=104). işletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 14.896 (S=1.269) kg olarak tahmin edilmiştir. İşletmesi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.304 (S=0.496) kg olarak tahmin edilmiştir. Dügüm 11 işletme faktörü tarafından etkilenecek iki yavru düğüme (Dügüm 37 ve Dügüm 38) ayrılmıştır. (Düz-P=0.000, F=20.998, sd1=1, sd2=55). İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 14.729 (S=0.349) kg olarak tahmin edilmiştir. İşletmesi Noshki olan işletmedeki kuzuların doğum ağırlığı 13.895 (S=0.692) kg olarak tahmin edilmiştir. Dügüm 13 ana yaşı faktörü tarafından etkilenecek altı yavru düğüme (Dügüm39, Dügüm 40, Dügüm41, Dügüm 42, Dügüm43, Dügüm 44) ayrılmıştır. (Düz-P=0.000, F=10.763, sd1=5, sd2=92). Ana yaşı $A \text{ YAŞ} \leq 27$ kuzuların doğum ağırlığı 14.197 (S=0.668) kg olarak belirlenmiştir. Dügüm 40, ana yaşı $27 < A \text{ YAŞ} \leq 36$ ay olan kuzuların doğum ağırlığı 14.908 (S=1.003) kg olarak saptanmıştır. Dügüm 41, ana yaşı $36 < A \text{ YAŞ} \leq 41$ ay olan kuzuların doğum ağırlığı 14.230 (S=0.536) kg olarak saptanmıştır. Dügüm 42, ana yaşı $41 < A \text{ YAŞ} \leq 48$ ay olan kuzuların doğum ağırlığı 15.066 (S=0.312) kg olarak saptanmıştır. Dügüm 43, ana yaşı $48 < A \text{ YAŞ} \leq 57$ ay olan kuzuların doğum ağırlığı 13.662 (S=0.789) kg olarak saptanmıştır. Dügüm 44, ana yaşı 57 aydan büyük olan kuzuların doğum ağırlığı 15.994 (S=0.926) kg olarak saptanmıştır. Dügüm 14 işletme faktörü tarafından etkilenecek iki yavru düğüme (Dügüm45, Dügüm 46) ayrılmıştır. (Düz-P<0.000, F=104.007, sd1=1, sd2=92). İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 15.725 (S=0.359) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 14.829 (S=0.480) kg olarak tahmin edilmiştir. Dügüm 15 işletme faktörü tarafından etkilenecek iki yavru düğüme (Dügüm47, Dügüm 48) ayrılmıştır. (Düz-P<0.000, F=28.050, sd1=1, sd2=51). İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 15.264 (S=0.395) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.534 (S=0.512) kg olarak kaydedilmiştir. Dügüm 16 işletme faktörü tarafından etkilenecek üç yavru düğüme (Dügüm49, Dügüm 50, Dügüm 51) ayrılmıştır. (Düz-P<0.000, F=18.258, sd1=2, sd2=166). Araştırma istasyonunda olan işletmedeki kuzuların doğum ağırlığı 16.133 (S=0.613) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki

kuzuların doğum ağırlığı 15.472 (S=0.824) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.973 (S=0.898) kg olarak kaydedilmiştir. Dügüm 17 işletme faktörü tarafından etkilenecek iki yavru düğüme (Dügüm52, Dügüm 53) ayrılmıştır. (Düz-P=0.000, F=205.858, sd1=1, sd2=93). İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 15.477 (S=0.325) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 16.385 (S=0.264) kg olarak kaydedilmiştir. Dügüm 18 işletme faktörü tarafından etkilenecek üç yavru düğüme (Dügüm54, Dügüm 55, Dügüm 56) ayrılmıştır. (Düz-P<0.000, F=30.734, sd1=2, sd2=38). İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 15.941 (S=0.484) kg olarak tahmin edilmiştir. İşletme ismi Quetta olan işletmedeki kuzuların doğum ağırlığı 15.179 (S=0.824) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.551 (S=0.349) kg olarak kaydedilmiştir. Dügüm 19 işletme faktörü tarafından etkilenecek üç yavru düğüme (Dügüm 57, Dügüm 58, Dügüm 59) ayrılmıştır. (Düz-P<0.000, F=23.634, sd1=2, sd2=48). Araştırma istasyonunda olan işletmedeki kuzuların doğum ağırlığı 16.920 (S=0.397) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 16.588 (S=0.302) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 15.576 (S=0.736) kg olarak gözlenmiştir. Dügüm 20 ana yaşı faktörü tarafından etkilenecek beş yavru düğüme (Dügüm60, Dügüm 61, Dügüm62, Dügüm 63, Dügüm64) ayrılmıştır. (Düz-P<0.000, F=12.908, sd1=4, sd2=149). Ana yaşı 23 ve 23 aydan küçük kuzuların doğum ağırlığı 16.531 (S=0.994) kg olarak belirlenmiştir. Dügüm 61, ana yaşı $23 < A \text{ YAŞ} \leq 41$ ay olan kuzuların doğum ağırlığı 15.490 (S=1.010) kg olarak saptanmıştır. Dügüm 62, ana yaşı $41 < A \text{ YAŞ} \leq 48$ ay olan kuzuların doğum ağırlığı 16.263 (S=0.415) kg olarak saptanmıştır. Dügüm 63, ana yaşı $48 < A \text{ YAŞ} \leq 57$ ay olan kuzuların doğum ağırlığı 14.699 (S=0.986) kg olarak saptanmıştır. Dügüm 64, ana yaşı $A \text{ YAŞ} > 57$ aydan kuzuların doğum ağırlığı 15.863 (S=1.114) kg olarak belirlenmiştir. Dügüm 21, cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Dügüm 65, Dügüm 66) ayrılmıştır. (Düz-P<0.000, F=45.704, sd1=1, sd2=92). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.036 (S=0.411) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 16.457(S=0.412) kg olarak tahmin edilmiştir. Burada erkekler dişilere oranla daha fazla ağırlığa sahip olduğu gözlenmiştir. Dügüm 22, cinsiyet

faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 67, Düğüm 68) ayrılmıştır. (Düz-P<0.000, F=27.929, sd1=1, sd2=25). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 16.109 (S=0.594) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 14.779(S=0.605) kg olarak tahmin edilmiştir. Düğüm 23 işletme faktörü tarafından etkilenecek iki yavru düğüme (Düğüm69, Düğüm 70) ayrılmıştır. (Düz-P<0.000, F=27.133, sd1=1, sd2=30). İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 16.893 (S=0.632) kg olarak belirlenmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 15.508 (S=0.260) kg olarak kaydedilmiştir. Düğüm 24 işletme faktörü tarafından etkilenecek üç yavru düğüme (Düğüm71, Düğüm 72, Düğüm 73) ayrılmıştır. (Düz-P<0.000, F=143.577, sd1=2, sd2=54). Araştırma istasyonundaki ismi Quetta olan işletmedeki kuzuların doğum ağırlığı 16.939 (S=0.651) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 16.482 (S=0.449) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 14.485 (S=0.367) kg olarak kaydedilmiştir. Düğüm 25 işletme faktörü tarafından etkilenecek üç yavru düğüme (Düğüm74, Düğüm 75, Düğüm 76) ayrılmıştır. (Düz-P<0.000, F=74.861, sd1=2, sd2=67). Araştırma istasyonundaki ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 17.216 (S=0.549) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 15.946 (S=0.902) kg olarak belirlenmiştir. İşletme ismi Quetta olan işletmedeki kuzuların doğum ağırlığı 15.374 (S=0.526) kg olarak saptanmıştır. Düğüm 26, cinsiyet faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 77, Düğüm 78) ayrılmıştır. (Düz-P<0.000, F=28.627, sd1=1, sd2=103). Cinsiyeti erkek olan kuzuların sütten kesim ağırlığı 17.389 (S=0.594) kg olarak tahmin edilmiş ve Cinsiyeti dişi olan kuzuların sütten kesim ağırlığı 16.792(S=0.466) kg olarak belirlenmiştir. Düğüm 27 işletme faktörü tarafından etkilenecek üç yavru düğüme (Düğüm79, Düğüm 80, Düğüm 81) ayrılmıştır. (Düz-P<0.000, F=28.961, sd1=2, sd2=110). Araştırma istasyonunda olan işletmedeki kuzuların doğum ağırlığı 18.143 (S=0.358) kg olarak tahmin edilmiştir. İşletme ismi Quetta olan işletmedeki kuzuların doğum ağırlığı 17.357 (S=0.552) kg olarak tahmin edilmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 16.456 (S=1.199) kg olarak saptanmıştır. Düğüm 28, sütten kesim ağırlığını yıl faktörü etkileyerek iki yavru düğüme (Düğüm 82, Düğüm 83) ayrılmıştır. (Düz-P<0.000,

F=18.806, sd1=1, sd2=60). Doğum yılı 2008 ve 2008 yılından küçük olan kuzuların (Düğüm 82) sütten kesim ağırlığı 15.942 (S=1.151) kg olarak tahmin edilmiştir. Doğum yılı 2008 yılından büyük olan kuzuların (Düğüm 83) sütten kesim ağırlığı 17.214 (S=0.583) kg olarak belirlenmiştir. Düğüm 29 işletme faktörü tarafından etkilenecek iki yavru düğüme (Düğüm 84, Düğüm 85) ayrılmıştır. (Düz-P<0.000, F=17.750, sd1=1, sd2=84). Araştırma istasyonundaki kuzuların doğum ağırlığı 18.654 (S=0.151) kg olarak tahmin edilmiştir. İşletme ismi İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 17.690 (S=0.787) kg olarak kaydedilmiştir. Düğüm 31 işletme faktörü tarafından etkilenecek üç yavru düğüme (Düğüm86, Düğüm 87, Düğüm 88) ayrılmıştır. (Düz-P<0.000, F=17.784, sd1=2, sd2=104). Araştırma istasyonundaki kuzuların doğum ağırlığı 18.816 (S=0.451) kg olarak tahmin edilmiştir. İşletme ismi Mastung olan işletmedeki kuzuların doğum ağırlığı 18.167 (S=0.522) kg olarak belirlenmiştir. İşletme ismi Noshki olan işletmedeki kuzuların doğum ağırlığı 17.399 (S=1.096) kg olarak saptanmıştır. Düğüm 32, sütten kesim ağırlığını yıl faktörü etkileyerek iki yavru düğüme (Düğüm 89, Düğüm 90) ayrılmıştır. (Düz-P<0.023, F=8.170, sd1=1, sd2=63). Doğum yılı 2008 ve 2008 yılından küçük olan kuzuların (Düğüm 89) sütten kesim ağırlığı 16.939 (S=1.302) kg olarak tahmin edilmiştir. Doğum yılı 2008 yılından büyük olan kuzuların (Düğüm 90) sütten kesim ağırlığı 17.759 (S=0.751) kg olarak belirlenmiştir.

Şekil incelendiğinde, 12, 30, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89 ve 90'nıncı düğümler terminal düğümlerdir. CHAID algoritmasından türetilen regresyon ağacında terminal düğümlerde yeterince homojenliğe ulaşılmış, sonraki aşamalarda başka dallanmalar olmamıştır.

CHAID algoritmasının kullanılması ile her bir kuzunun sütten kesim ağırlığı için tahmin değeri hesaplanmış, sütten kesim ağırlığı bakımından gerçek ve tahmin değerleri arasındaki Pearson korelasyon katsayısı 0.863 bulunmuştur (P<0.01). Her bir kuzu için hesaplanan hata değerlerine ait standart sapmanın, gerçek değerlere ait standart sapmaya oranı (Soran) 0.506 bulunmuştur. Sütten kesim ağırlığına ait toplam varyasyonun % 74.385'i doğum ağırlığı, yıl, cinsiyet ve işletme gibi değişkenler tarafından

açıklanmaktadır. CHAID algoritması için üzerinde durulan özelliğe yani sütten kesim ağırlığına ait açıklanan varyasyon oranının artırılması, modele doğum ağırlığı, yıl, cinsiyet ve işletme dışında farklı değişkenlerin modelde olması ile ilişkilidir. Bilindiği üzere açıklanan varyasyon payının artırılması, ebeveyn-yavru düğüm oranlarının azaltılması ile mümkündür (Çizelge 1) ve analiz yapıldığında Şekilde görülebilir).



5. SONUÇ ve ÖNERİLER

Bu çalışmada, farklı ebeveyn-yavru düğüm oranları için üç algoritmanın (CART, CHAID, Geniş (Exhaustive) CHAID) tahmin performansı karşılaştırılmıştır. Ayrıca, hangi algoritma ile daha uygun ağaç elde edildiği de gözlemlenmiştir. Çalışmada farklı ebeveyn- yavru düğüm oranları karşılaştırılmıştır. Genel olarak, ebeveyn-yavru düğüm oranları 500:250'den 10:5'e doğru azaltıldığında uyum kriterlerinin daha iyi sonuç verdiği görülmüştür. Çizelge 4.4, 4.5 ve 4.6'daki uyum ölçütleri bakımından yapılan iki yönlü varyans analizi sonucunda ebeveyn:yavru düğüm oranları arasında fark olduğu (500:250, 400:200, 300:150, 200:100 ve 100:50), ancak algoritmalar arasında fark olmadığı tespit edilmiştir. CART algoritmasının, CHAID algoritmalarına nazaran aşırı dallanmalar oluşturduğundan dolayı CART algoritmasına ilişkin elde edilen karar ağacının zor yorumlanacağı anlaşılmıştır.

Bu çalışmada elde edilen verilere göre farklı ebeveyn-yavru düğüm oranları incelenmiş ve algoritmalarda elde edilen uyum iyiliği ölçütlerine ilişkin sonuçlar aşağıda özetlenmiştir.

CHAID algoritması için hesaplanan uyum iyiliği ölçütleri dikkate alındığında;

1. Risk tahmin değeri (RT) ebeveyn- yavru düğüm oranları 500:250 ile 380:190 arasında sabit ($RT=0.922$) kaldığı, ebeveyn- yavru düğüm oranları 380:190 ile 100:50 arasında olduğunda risk tahmin değerinin ($RT=0.922-0.761$) düştüğü görülmektedir.
2. Belirleme katsayısı ($R^2(\%)$), 500:250 ile 380:190 arasındaki ebeveyn- yavru düğüm oranları için sabit ($R^2(\%)=50.176$) kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50'ye doğru azaldığında belirleme katsayısının ($R^2(\%)=50.176-61.778$) yükseldiği görülmektedir.
3. Nisbi yaklaşık hata (NYH) değeri 500:250 ile 360:180 arasındaki ebeveyn-yavru düğüm oranları için (NYH =0.249) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'dan 100:50'ye doğru azaldığında nisbi yaklaşık hata (NYH=0.249-0.218) değerinin düştüğü kaydedilmiştir.

4. Hata kareler ortalamasının karekökü (HKOK) değeri 500:250 ile 380:190 arasında tanımlanan ebeveyn- yavru düğüm oranları için (HKOK =0.996) sabit kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50'ye doğru azaldığında hata kareler ortalamasının karekökü (HKOK =0.996-0.872) düştüğü görülmüştür.
5. Standart sapma oranı (SS-oran) değerlerinin 500:250 ile 360:180 arasında tanımlanan ebeveyn- yavru düğüm oranları için (SS-oran =0.706) sabit kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50' ye doğru azaldığında standart sapma oranı (SS-oran=0.691-0.618) düştüğü görülmüştür.
6. Pearson korelasyon katsayısı (r) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında (r =0.708) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'den 100:50'ye doğru azaldığında pearson korelasyon katsayısı değeri (r =0.708-0.786) arttığı kaydedilmiştir.
7. Düzeltmiş belirleme katsayısı ($R^2_{\text{düz.}}(\%)$) değerlerinin 500:250 ile 380:190 arasında tanımlanan ebeveyn- yavru düğüm oranları için ($R^2_{\text{düz.}}(\%) =50.128$) sabit kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50' ye doğru azaldığında düzeltilmiş belirleme katsayısı değeri ($R^2_{\text{düz.}}(\%) =50.128-61.696$) yükselmiştir.
8. Varyasyon katsayısı (VK(%)) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında (VK(%))=6.222) sabit kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50'ye doğru azaldığında varyasyon katsayısı değerlerinin (VK(%))=6.222-5.448) düştüğü görülmüştür.

Geniş (Exhaustive) CHAID algoritması için hesaplanan uyum iyiliği ölçütleri dikkate alındığında;

1. Risk tahmin değerlerinin (RT) tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında sabit (RT=0.992) kaldığı, ebeveyn- yavru düğüm oranları 360:180 ile 100:50'e doğru azaltıldığında risk tahmin değerinin (RT=0.951-0.740) düştüğü görülmektedir.

2. Belirleme katsayısı ($R^2(\%)$) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında sabit ($R^2(\%)=50.176$) kaldığı, ebeveyn- yavru düğüm oranları 360:180'dan 100:50'ye doğru azaldığında belirleme katsayısı değerinin ($R^2(\%)=52.235-62.835$) yükseldiği görülmektedir.
3. Nisbi yaklaşık hata (NYH) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 360:180 arasında sabit (NYH =0.249) kaldığı, ebeveyn - yavru düğüm oranları 340:170'dan 100:50'ye doğru azaldığında nisbi yaklaşık hata (NYH=0.244-0.215) değeri düştüğü kaydedilmiştir.
4. Hata kareler ortalamasının karekökü (HKOK) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında (HKOK =0.996) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'den 100:50'ye doğru azaldığında hata kareler ortalamasının karekök değerlerinin (HKOK =0.975-0.860) düştüğü görülmüştür.
5. 500:250 ile 360:180 arasında tanımlanan ebeveyn- yavru düğüm oranları için standart sapma oranı (SS-oran) değerlerinin (SS-oran =0.706) sabit kaldığı, ebeveyn- yavru düğüm oranları 380:190'dan 100:50'ye doğru azaldığında ise standart sapma oranı değerlerinin (SS-oran =0.691-0.610) düştüğü görülmüştür.
6. Pearson korelasyon katsayısı (r) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında (r =0.708) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'den 100:50'ye doğru azaldığında pearson korelasyon katsayısı değerinin (r =0.723-0.793) arttığı kaydedilmiştir.
7. 500:250 ile 380:190 arasında tanımlanan ebeveyn- yavru düğüm oranları için düzeltilmiş belirleme katsayısı ($R^2_{-düz.}(\%)$) değerlerinin ($R^2_{-düz.}(\%) =50.128$) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'dan 100:50'ye doğru azaldığında düzeltilmiş belirleme katsayısı değerinin ($R^2_{-düz.}(\%)=52.128-62.707$) yükselmiştir.

8. Varyasyon katsayısı (VK(%)) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında (VK(%)=6.222) sabit kaldığı, ebeveyn- yavru düğüm oranları 360:180'dan 100:50'ye doğru azaldığında varyasyon katsayısı değerinin (VK(%)=6.092-5.374) düştüğü görülmüştür.

CART algoritması için hesaplanan uyum iyiliği ölçütleri dikkate alındığında;

1. Risk tahmin değerlerinin (RT) tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında sabit (RT=1.022) kaldığı, ebeveyn- yavru düğüm oranlarının 360:180 ile 100:50'e doğru azaltıldığında risk tahmin değerinin (RT=0.920-0.693) düştüğü görülmektedir.
2. Belirleme katsayısı değerlerinin (R^2 (%)) tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında sabit (R^2 (%)=48.667) kaldığı, ebeveyn- yavru düğüm oranlarının 360:180'dan 100:50'ye doğru azaldığında belirleme katsayısı (R^2 (%)=53.790-65.192) yükseldiği görülmektedir.
3. 500:250 ile 360:180 arasında tanımlanan ebeveyn- yavru düğüm oranları için Nisbi yaklaşık hata (NYH) değerlerinin (NYH =0.253) sabit kaldığı, ebeveyn- yavru düğüm oranlarının 340:170'dan 100:50'ye doğru azaldığında Nisbi yaklaşık hata (NYH=0.240-0.211) değeri düştüğü kaydedilmiştir.
4. Hata kareler ortalamasının karekökü (HKOK) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 360:180 arasında (HKOK =1.011) sabit kaldığı, ebeveyn- yavru düğüm oranlarının 340:170'den 100:50'ye doğru azaldığında hata kareler ortalamasının karekökü (HKOK =0.959-0.844) düştüğü görülmüştür.
5. 500:250 ile 360:180 arasında tanımlanan ebeveyn- yavru düğüm oranları için Standart sapma oranı (SS-oran) değerlerinin (SS-oran =0.719) sabit kaldığı, ebeveyn- yavru düğüm oranlarının 340:170'dan 100:50'ye doğru azaldığında standart sapma oranı (SS-oran =0.680-0.598) düştüğü görülmüştür.
6. Pearson korelasyon katsayısı (r) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 360:180 arasında (r =0.697) sabit kaldığı,

ebeveyn- yavru düğüm oranlarının 340:170'den 100:50'ye doğru azaldığında pearson korelasyon katsayısı değeri ($r=0.733-0.802$) arttığı kaydedilmiştir.

7. Düzeltilmiş belirleme katsayısı ($R^2_{\text{düz.}}(\%)$) değerlerinin tanımlanan ebeveyn- yavru düğüm oranları için 500:250 ile 380:190 arasında ($R^2_{\text{düz.}}(\%)=48.593$) sabit kaldığı, ebeveyn- yavru düğüm oranlarının 360:180'dan 100:50'ye doğru azaldığında düzeltilmiş belirleme katsayısı değeri ($R^2_{\text{düz.}}(\%)=48.538-64.154$) yükselmiştir.

8. 500:250 ile 360:180 arasında tanımlanan ebeveyn- yavru düğüm oranları için Varyasyon katsayısı (VK(%)) değerlerinin (VK(%)=6.332) sabit kaldığı, ebeveyn- yavru düğüm oranlarının 360:180'dan 100:50'ye doğru azaldığında varyasyon katsayısı (VK(%)=5.991-5.270) düştüğü görülmüştür.

CART, pruning CART ile karşılaştırıldığında arasında önemli bir fark olmadığı görülmüştür.

Sonuç olarak incelenen veri seti için CHAID algoritması ile oluşturulan karar ağacının daha anlaşılabilir ve kolay olduğu sonucuna varılmıştır. Literatürde, canlı ağırlık tahmininde kullanılan vücut ölçülerinin canlı ağırlık ile yüksek genetik korelasyonundan hareketle hayvan ıslahı bakımından daha efektif sonuçlar elde edilmesi muhtemeldir (Khan *et al.*, 2014; Ali *et al.*, 2015). Sonuç olarak, genetik korelasyona sahip özellikler ve akrabalı yetiştirme katsayılarının CART, CHAID ve Geniş (Exhaustive) CHAID algoritmaları ile birlikte kullanılması sonucu oluşturulan regresyon karar ağacının ileride yapılacak hayvan ıslahı çalışmalarında kullanılabileceği ümit edilmektedir. Diğer bir anlatımla, canlı ağırlık, süt verimi ve yapağı verimi gibi incelenen kantitatif verim özellikler bakımından bireysel akrabalı yetiştirme katsayılarının regresyon ağacı analiz yöntemlerinde kullanılması, akrabalı yetiştirme dejenerasyonuna ilişkin sınır değerlerinin tespit edilmesine olanak sağlayabilir.

KAYNAKLAR

- Ali, M., Eyduran, E., Tariq, M.M., Tirink, C., Abbas, F., Bajwa, M.A., Baloch, M.H., Nizamani, A.H., Waheed, A., Awan, M.A., Shah, S.H., Ahmad, Z. Jan, S. 2015. Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in harnai sheep. *Pakistan Journal of Zoology*, vol. 47(6), 1579-1585.
- Bakır, G., Keskin, S., Mirtagioğlu, H. 2009. Evaluating the relationship between mature age milk yield and several traits using CHAID analysis in Brown Swiss Cows. *Journal of Animal and Veterinary Advances* , 8 (3) : 587-589.
- Bakır, G., Keskin, S., Mirtagioğlu, H. 2010. Determination of the effective factors for 305 days milk yield by regression tree (RT) method. *Journal of Animal and Veterinary Advances* , 9 (1) : 55-59.
- Bayram, B., Topal, M., Aksakal, V. Önk, K. 2015. Investigate the effects of non-genetic factors on calving difficulty and stillbirth rate in holstein friesian cattle using the CHAID analysis. *Kafkas Universitesi Veteriner Fakültesi Dergisi*. 21 (5): 645-652.
- Biggs, D., B. de Ville, Suen, E. 1991. A method Of Choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics*, 18, 49-62.
- Breiman, L., J. H. Friedman, R. A. Olshen, Stone. C. J. 1984. *Classification and Regression Trees. New York: Chapman & Hall/CRC*.
- Cak, B., Keskin, S., Yılmaz, O. 2013. Regression tree analysis for determining of affecting factors to lactation milk yield in Brown Swiss Cattle. *Asian Journal of Animal Veterinary Advances*, 8 (4) : 677-682.
- Caraviello, D.Z., Weigel, K.A., Craven, M., Gianola, D., Cook, N.B., Nordlund, K.V., Fricke, P.M., Wiltbank, M.C. 2006. Analysis of reproductive performance of lactating cows on large dairy farms using machine learning algorithms. *Journal of Dairy Science*. 89:4703-4722.

- Çamdeviren H., Mendeş M., Ozkan M.M., Toros F., Şaşmaz T., Oner S. 2005. Determination of depression risk factors in children and adolescents by regression tree methodology. *Acta Medica Okayama*, 59(1):19-26.
- Doğan, I. 2003. Holştayn ırkı ineklerde Süt verimine etki eden faktörlerin CHAID analizi ile incelenmesi. *Ankara Üniversitesi Veteriner Fakültesi Dergisi*, 50:65-70.
- Eyduran, E., Tatlıyer, A., Tariq, M., Waheed, A. 2013a. Application of classification and regression tree methods in agriculture. *Ulusal Tarım Kongresi*, 26-29 Ekim, Antalya.
- Eyduran, E., Yılmaz, I., Kaygısız, A., Aktaş, Z. M. 2013b. An Investigation on relationship between lactation milk yield, somatic cell count and udder traits in first lactation Turkish Saanen goat using different statistical techniques. *The Journal Animal Plant Science*, 23(4): 956-963.
- Eyduran, E., Yılmaz, I., Tariq, M. M., Kaygısız, A. 2013c. Estimation of 305-d milk yield using regression tree method in brown Swiss cattle. *The Journal Animal Plant Science*, 23(3): 731-735
- Goodman, L.A. 1979. Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74, 537-552.
- Grzesiak, W., Lacroix, R., Wójcik, J., Błaszczuk, P. 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Canadian Journal of Animal Science*, 83(2): 307-310.
- Grzesiak, W., Błaszczuk, P., Lacroix, R. 2006. Methods of predicting milk yield in dairy cows predictive capabilities of wood's lactation curve and artificial neural networks (ANNs). *Computers and Electronic in Agriculture*, 54: 69-83.
- Grzesiak, W., Zaborski, D. 2012. Examples of the use of data mining methods in animal breeding. *Additional information is available at the end of the chapter.* <http://dx.doi.org/10.5772/50893>

- Grzesiak, W., Zaborski, D., Sablik, P., Pilarczyk, R. 2011. Detection of difficult conceptions in dairy cows using selected data mining methods. *Animal Science Papers and reports I*. 29(4), 293-302.
- Karabağ, K., Alkan. S., Mendeş, M. 2010. Kınalı keklik (*Alectoris chukar*) yumurtalarında çıkış gücüne etki eden faktörlerin sınıflandırma ağacı yöntemi ile belirlenmesi. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi* 16(5): 723-727.
- Kass, G. 1980. An exploratory technique for investigating large quantities of categorical Data. *Applied Statistics*, 29:2, 119-127.
- Kayri, M., Boysan, M. 2008. Assesment of relation between cognitive vulnerability and depression's level by using classification and regression tree analysis. *Hacettepe Üniversitesi Egitim Fakültesi Dergisi*, 34, 168-177.
- Khan, MA., Tariq, M. M., Eyduran, E., Tathyer, A., Rafeeq, M., Abbas, F., Rashid, N., Awan, M.A., Javed, K. 2014. Estimating body weight from several body measurements in Harnai sheep without multicollinearity problem. *The Journal of Animal Plant Science*. 24(1) p.120-126.
- Küçüköğlü, O. 2010. *Veri madenciliği yöntemlerinin hayvancılıkta kullanımı*. Çanakkale Onsekiz Mart Üniversitesi Fen Bilimleri Enstitüsü (Yüksek Lisans Tezi), Çanakkale.
- Maia, A. P., Oliveiraa, S. R. M., Moura, D. J., Sarubbi, J., Vercellino, R. A., Medeiros, B. B. L., Griskaa, P. R. 2013. A decision tree-based model for evaluating the thermal comfort of horses. *Scientia Agricola Science agriculture* 70 (6).
- Mendeş, M., Akkartal, E. 2009. Regression tree analysis for predicting slaughter weight in broilers. *Italian Journal of Animal Science*.,8: 615-624.
- Oruçoğlu, O. (2011). *Holstein ırkı ineklerin 305 günlük süt verimini etkileyen çevre faktörlerinin regresyon ağacı ile belirlenmesi*. Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü (Yüksek lisans tezi).

- Piwczynski, D. 2009. Using classification trees in statistical analysis of discrete sheep reproduction traits. *Journal Central European Agriculture*. 10(3) 303-310
- Piwczynski, D., Sitkowska, B. 2012a. Statistical modelling of somatic cell counts using the classification tree technique. *Aarchiv Tierzucht* 55(4);332- 345
- Piwczynski, D., Nogalski, Z., Sitkowska, B. 2013. Statistical modeling of calving ease and stillbirths in dairy cattle using the classification tree technique. *Livestock Science*. 154(1-3) 19-27
- Piwczynski, D., Sitkowska, B., Wisniewska, E. 2012b. Application of classification trees and logistic regression to determine factors responsible for lamb mortality. *Small ruminant research*,103; 225-231.
- Schlegel , H., Fischer, K., Wahner, M. 2013. Practical aspects for classification of performance parameters of pigs by decision tree. *Research in pig breeding*, 7, (2).
- Sut, N., Simsek, O. 2011. Comparison of regression tree data mining methods for prediction of mortality in head injury. *Expert Systems with Applications*, 38(12);15534-15539.
- Tariq, M.M., Rafeeq, M.,Bajwa, M.A., Abbas, F., Waheed, A., Bukhari, F. A., Akhtar, P. 2012. Prediction of body weight from body measurements using regression tree (RT) method for indigenous sheep breeds in Balochistan, Pakistan. *The Journal of Animal Plant Science*. 22(1): 20-24.
- Topal, M., Aksakal, V., Bayram, B., Yağanoğlu, A. M. 2010. An analysis of factors affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. *The Journal of Animal Plant Science*, 20(2): 63-69.
- Yakubu, A. 2012. Application of regression tree methodology in predicting the body weight of udder sheep. Scientific papers. *Animal Science and Biotechnologies*, 45 (2).
- Yılmaz, İ., Eydurhan, E., Kaygısız, A. 2013. Determination of non-genetic factors influencing birth weight using regression tree method in Brown-Swiss cattle. *Canadian Journal of Applied Science*. 1(3): 382-387.

ÖZGEÇMİŞ

1986 yılında Muş İli İnardı köyünde doğdu. İlköğretim ve Ortaöğretimi Muş ilinde tamamladı. 2008 yılında Iğdır Üniversitesi, Ziraat Fakültesinde lisans öğrenimine başladı ve 2012 yılında mezun oldu. Aynı yıl Iğdır Üniversitesi Fen Bilimleri Enstitüsü Zootekni Anabilim Dalında Lisansüstü öğrenimine başladı.

