



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

KİMİ KÜLTÜR IRKI SIĞIRLARDA BESİ SONU CANLI AĞIRLIĞINI ETKİLEYEN ÇEŞİTLİ VÜCUT ÖLÇÜLERİNİN REGRESYON AĞACI YÖNTEMİYLE BELİRLENMESİ

Rifat AKŞAHAN

YÜKSEK LİSANS TEZİ

Zootekni Anabilim Dalını

Haziran-2015
KONYA
Her Hakkı Saklıdır

ÖZET

YÜKSEK LİSANS TEZİ

KİMİ KÜLTÜR IRKI SIĞIRLARDA BESİ SONU CANLI AĞIRLIĞINI ETKİLEYEN ÇEŞİTLİ VÜCUT ÖLÇÜLERİNİN REGRESYON AĞACI YÖNTEMİYLE BELİRLENMESİ

Rifat AKŞAHAN

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Zootekni Anabilim Dalı

Danışman: Doç. Dr. İsmail KESKİN

2015, 27 Sayfa

Jüri
Prof. Dr. Saim BOZTEPE
Doç. Dr. İsmail KESKİN
Doç. Dr. Ecevit EYDURAN

Bu çalışmanın amacı,Türkiye'nin Afyon ilinin Bolvadin ilçesinde yetiştirilen 103 baş (Siyah Alaca, Simental, Esmer, ve Melez) tosunun besi sonu canlı ağırlığını (BSCA) etkileyen vücut ölçülerini belirlemektir. Bu amaç için, CHAID algoritmasını esas alan regresyon agacı metodu kullanılmıştır. Bağımlı değişken olarak, BSCA kullanılmıştır. Bu çalışmada kullanılan bağımsız değişkenler sırasıyla, cidago yüksekliği (CY), sırt yüksekliği (SIY), sağrı yüksekliği (SAY), kuyruk sokumu yüksekliği (KSY), vücut uzunluğu (VU), arka sağrı genişliği (ASG), göğüs derinliği (GD), ve göğüs çevresi (GÇ) ve besi süresi (BS) olmuştur. Bu çalışmada en iyi model kriterlerini elde etmek amacıyla CHAID algoritması için ebeveyn ve çocuk düğümde bulunan minimum tosun sayıları 8:4 olarak ayarlanmıştır. CHAID algoritmasının etkinliğini belirlemek için model kalite kriterleri olarak belirleme katsayısı, düzeltilmiş belirleme katsayısı ve gerçek ve tahmin edilen BSCA değerleri arasındaki Pearson korelasyon katsayısı kullanılmıştır. BSCA etkileyen bağımsız değişkenler; GÇ (Düz. $P<0.01$), VU (Düz. $P<0.05$), KSY (Düz. $P<0.01$) ve BS (Düz. $P<0.01$) olmuştur.

Belirleme katsayısı (%), düzeltilmiş belirleme katsayısı (%) ve gerçek ve tahmin edilen BSCA değerleri arasındaki Pearson korelasyon katsayısı sırasıyla, 87.82 (%), 87.32(%) ve 0.937 bulunmuştur. Model kalite kriterleri, CHAID algoritması kullanımının yüksek bir tahmin doğruluğu sağladığı göstermiştir. Ortalama olarak, en ağır BSCA (545.850 kg); göğüs çevresi GÇ > 190 cm olan tosunların oluşturduğu alt gruptan elde edilmiştir. İkinci olarak en ağır BSCA (505.562 kg); 183 < GÇ $\leq$ 190 cm ve BS > 4 ay olan tosunların oluşturduğu alt gruptan elde edilmiştir.

Sonuç olarak, farklı vücut ölçülerinden BSCA tahmin etmek amacıyla CHAID algoritması kullanımı, hayvan ıslahı çalışmaları için tavsiye edilebilir.

Anahtar Kelimeler: CHAID algoritması, Regresyon Ağacı, Siyah Alaca, Simental, Esmer Sığır

ABSTRACT

MS THESIS

DETERMINATION OF BODY MEASUREMENTS INFLUENCING FINAL LIVE WEIGHT VIA REGRESSION TREE METHOD IN SOME CATTLE BREEDS

Rifat AKŞAHAN

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF SELÇUK UNIVERSITY
THE DEGREE OF MASTER SCIENCE
IN ANIMAL SCIENCE

Advisor: Assoc. Prof. Dr. İsmail KESKİN
2015, 27 Pages

Jury
Prof. Dr. Saim BOZTEPE
Assoc. Prof. Dr. İsmail KESKİN
Assoc. Prof. Dr. Ecevit EYDURAN

The aim of this study is to determine body measurements affecting final live weight (FLW) at fattening period of 103 young (Holstein, Simmental, Brown Swiss and crossbreed) bulls reared in Bolvadin district of Afyon province of Turkey. For this aim, a regression tree method on the basis of CHAID (Chi-Squared Automatic Interaction Detector) algorithm were used. FLW was considered as response variable. The independent variables included in the study were withers height (WH), back height (BH), front rump height (FRH), back rump height (BRH), body length (BL), back rump width (BRW), chest dept (CD), chest circumference (CC), and fattening period (FP). Minimum numbers of young bulls in parent and child nodes for CHAID algorithm were assigned as: 8:4, in order to obtain the best model quality criteria for the data evaluated under the study. Model quality criteria like coefficient of determination (%), adjusted coefficient of determination (%), and the Pearson correlation between actual and predicted fattening FLW values were estimated to determine the effectiveness of CHAID algorithm. The independent variables influencing FLW were CC (Adj. $P<0.01$), BL (Adj. $P<0.05$), BRH (Adj. $P<0.01$), and FP (Adj. $P<0.01$).

Coefficient of determination (%), adjusted coefficient of determination (%), and the Pearson correlation between actual and predicted fattening final weight values were found 87.82 (%), 87.32 (%), and 0.937, illustrating that use of CHAID algorithm visually provided a high predictive accuracy. The heaviest FLW (545.850 kg) at fattening period were obtained from the sub-group of the young bulls with CC > 190 cm, averagely. The secondly heaviest sub group of the young bulls with $183 < CC \leq 190$ cm and FP > 4 months provided a FLW of 505.562 kg.

Consequently, the use of CHAID algorithm for predicting FLW from various body measurements could be advised to be useful for breeding purposes.

Keywords: CHAID algorithm, Regression Tree, Holstein, Simmental, Brown Swiss

ÖNSÖZ

Lisansüstü eğitimim boyunca ve bu tezin hazırlanmasında benden desteklerini esirgemeyip bilgi ve deneyimleri ile yol gösteren Danışman Hocam Sayın Doç. Dr. İsmail KESKİN'e en içten samimiyetle teşekkürlerimi sunarım.

Tezin kaynak araştırma ve sonuçları değerlendirme aşamasında yardımcı olan Doç. Dr. Ecevit EYDURAN'a teşekkür ederim.

Tez verilerinin toplanma aşamasında yardımcı olan meslektaşım ve aynı zamanda arkadaşım Şükrü GÜNER'e teşekkür ederim.

Bu günlere gelmemde büyük pay sahibi olan aileme ve dostlarıma teşekkürlerimi sunarım.

Rifat AKŞAHAN
KONYA-2015

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

SİMGELER VE KISALTMALAR ..... **viii**

1. GİRİŞ **1**

2. KAYNAK ARAŞTIRMASI **3**

2.1. Hayvancılık Alanında Yapılan Çalışmalar 3
2.2. Diğer Alanlarda Yapılan Çalışmalar 7

3. MATERYAL VE YÖNTEM **10**

3.1. Materyal 10
3.1.1. Sığırlardan alınan parametreler 10
3.2. Yöntem 14
3.2.1. Sınıflama ve regresyon ağacı (cassification and regression tree) 14
3.2.2. İstatistiksel analizler 17
3.2.3. Model değerlendirme kriterleri 17

4. ARAŞTIRMA BULGULARI **19**

5. SONUÇ VE ÖNERİLER **23**

KAYNAKLAR **24**

ÖZGEÇMİŞ **27**

SİMGELER VE KISALTMALAR

Kısaltmalar

ASG	Arka Sağrı Genişliği
BS	Besi Süresi
BSCA	Besi Sonu Canlı Ağırlık
CY	Cidago Yüksekliği
CV	Varyasyon Katsayısı
GÇ	Göğüs Çevresi
GD	Göğüs Derinliği
KSY	Kuyruk Sokumu Yüksekliği
SAY	Sağrı Yüksekliği
SIY	Sırt Yüksekliği
SRA	Sınıflama ve Regresyon Ağacı
VU	Vücut Uzunluğu

1. GİRİŞ

Bağımsız değişkenin (ya da değişkenlerin) kendi ölçü birimi cinsinden bir birim değişmesine karşılık bağımlı değişkenin kendi ölçü birimi cinsinden ortalama olarak ne kadar değişeceğini gösteren katsayıya regresyon katsayısı denmektedir. Bağımsız değişken veya değişkenler ile bağımlı değişken arasındaki ilişkiyi inceleyen regresyon analizi; basit veya çoklu olabildiği gibi, doğrusallık, normallik, homojenlik, toplanabilirlik gibi bir takım varsayımların yerine getirilmesinden sonra uygulanabilmektedir. Bu varsayımların yerine gelmemesi durumunda veri setindeki orijinal değerler ya logaritmik dönüşüm, ya da karekök dönüşüm yöntemleri gibi bir takım dönüştürme işlemlerine tabi tutularak uygun hale getirilmeye çalışılmaktadır (Efe ve ark., 2000). Parametrik yöntemlerde her ne kadar dönüştürme metotları kullanılarak ön koşul varsayımlar yerine getirilmeye çalışılıyorsa da, yapılan analiz neticesinde veri setine ilişkin yanlı sonuçlar elde edilebilmektedir. Bu nedenle alternatifin olmadığı durumlarda verileri dönüştürme yoluna gidilmesi istatistik açıdan daha doğrudur (Kayri ve Boysan, 2008).

Karar ağaçları, bir problemi oluşturan veri setlerinin yapısına göre bir ağaç yapısı şeklinde sınıflandırma ve regresyon modelleri oluşturmaktadır. Söz konusu ağaç yapılarının oluşturulmasında kullanılan karar kurallarının anlaşılabilir olması yöntemin kullanımını yaygın hale getirmiştir. Karar ağaçları, sınıflama ve regresyon probleminin çözümünde çok aşamalı ve ardışık bir yaklaşım ile karmaşık yapıdaki verileri aşamalı bir hale dönüştürerek basit bir karar verme işlemini gerçekleştirmektedir (Safavian ve ark., 1991). Ağaç modellerinde bağımlı değişken kategorik yapıda olduğu durumlarda sınıflama ağacı, bağımlı değişken sürekli bir değişken olduğu durumlarda ise regresyon ağacı modeli kullanılmaktadır (Sümbüloğlu ve Akdağ, 2007; Nefeslioğlu ve ark., 2010).

Sınıflama ve Regresyon ağaçları (SRA) bağımsız değişkene ait hiçbir ön koşul öne sürmeden kesikli ya da sürekli bağımlı değişkenin sınıf üyeliğini tahmin etmeye yarayan ters ağaç şeklindeki modellerdir. Kategorik ya da sürekli, bir ya da birden fazla bağımsız değişkenin kombinasyonları kullanılarak, tekrarlamalı ikili homojen bölünmelerle, bağımlı değişkendeki değişimi ortaya çıkarmaya ve bağımlı değişkenin değerlerini tahmin etmeye yarayan ve görsel olarak ters ağaç şeklindeki modellere ağaç modelleri denir. İstatistik verilerin görsel olarak sunulması, aralarındaki etkileşimin belirlenmesi için karar ağaç modelleri sıklıkla kullanılmaktadır. Ağaç modellerinin işleyiş yapısı, bağımsız değişkene ait temel basit sorulardan alınan cevapların oluşturduğu yolları (ağaç dalları) takip etmektedir. Bu yollar ise (ağaç dalları) bağımlı

değişkeni hangi bağımsız değişken ya da değişkenlerin etkilediğini gösterir (Temel, 2004).

Bu çalışmada, besi sonu canlı ağırlığını (BSCA) etkileyen çeşitli vücut ölçülerinin ((cidago yüksekliği (CY), sırt yüksekliği (SIY), besi süresi (BS), sağrı yüksekliği (SAY), kuyruk sokumu yüksekliği (KSY), vücut uzunlu (VU), arka sağrı genişliği (ASG), göğüs derinliği (GD) ve göğüs çevresi (GÇ)) regresyon ağacı yöntemiyle incelenmesi amaçlanmıştır.

2. KAYNAK ARAŞTIRMASI

2.1. Hayvancılık Alanında Yapılan Çalışmalar

Doğan (2003), Bala tarım işletmesinde yetiştirilen 440 baş Holştayn ırkı ineğe ait 305 günlük süt verimini bağımlı, laktasyon süresi, kuru süre, ilk damızlıkta kullanma yaşı, buzağılama aralığı, gebelik başına tohumlama sayısı, servis periyodu, buzağılama mevsimi, gebelik süresi özelliklerini de bağımsız değişken olarak ele alarak CHAID analiz yöntemini uygulamıştır. CHAID yöntemi, bir popülasyonu; bağımlı değişkendeki varyasyonu gruplar içi minimum ve gruplar arası maksimum olacak şekilde farklı alt gruplara veya bölümlere tekrarlı olarak ayıran ve değişkenler arasındaki etkileşim veya kombinasyonları bulan bir yöntemdir. Çalışmada, süt verimini en iyi açıklayan ilk alt kategori olarak 55-74 günlük kuruda kalma süresi tahmin edilmiştir. 55-74 gün olarak belirlenen kuruda kalma süresi ile en fazla etkileşim içerisinde bulunan değişken ilk damızlıkta kullanma yaşı olarak tespit edilmiştir. İlk damızlıkta kullanma yaşı kendi içinde 10-16 ay ve 16-32 ay olarak iki alt gruba ayrılmış bu alt gruplar arasında da 10-16 ay ilk damızlıkta kullanma yaşının süt veriminde en iyi süre olduğu belirtilmiştir.

Eyduran ve ark. (2008), Akkaraman ırkının varyeteleri olan Norduz (103 baş) ve Karakaş (101 baş) kuzuları üzerinde yaptıkları çalışmada, doğum ağırlığı üzerinde etkili olan faktörleri (doğum tipi, doğum yılı, cinsiyet ve ana yaşı) belirlemek amacıyla regresyon ağacı metodunu kullanmışlardır. Çalışmada, doğum ağırlığını birinci dereceden doğum tipinin, ikinci dereceden ırk ve cinsiyetin, üçüncü dereceden ise ana yaşının etkilediği belirtilmiştir. Doğum tipi tekiz olan kuzuların bulunduğu grubu temsil eden Düğüm 1'deki tekiz kuzuların doğum ağırlığını sadece cinsiyetin etkilediğini ifade etmişlerdir. Tekiz erkek hayvanların (düğüm 3) ortalama doğum ağırlığı 5.065 kg, tekiz ve dişi kuzularda (düğüm 4) ortalama doğum ağırlığı 4.609 kg olarak tahmin edilmiştir. Düğüm 3 ve Düğüm 4'ü oluşturan tekiz-erkek ve tekiz-dişi kuzu gruplarında yeterli homojenite sağlandığından bölünme işlemi durmuştur. İkiz doğan Norduz kuzularının (Düğüm 6) doğum ağırlığı üzerine ana yaşı faktörünün etkili olduğu görülmüştür (Düz.P<0,05). Düğüm 6'da (ikiz ve Norduz kuzuları) ana yaşı bakımından iki çocuk düğüme bölünmüştür (Düğüm 7 ve Düğüm 8). Düğüm 7; 4 ve 5 yaşlı analarda doğan Norduz ikiz kuzularının oluşturduğu grubu temsil etmekle birlikte ortalama doğum ağırlığı 4.475 kg olarak bulunmuştur. Düğüm 8'de ise 2-3 ve 6 yaşlı analardan doğan Norduz ırkı ikiz kuzuların oluşturduğu gruptur. Dört 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(Düz.P<0,05).

Bakır ve ark. (2009), Muş devlet üretme çiftliğinde yetiştirilen 770 baş Brown Swiss ırkı sığırlarda ergin çağ süt verimi ile ineğin yaşı, ilk tohumlama yaşı, laktasyon sırası, laktasyon süresi, kuruda kalma süresi, ilk buzağılama yaşı, buzağılama mevsimi, doğum tipi ve cinsiyet gibi değişkenler arasındaki ilişkiyi CHAID analizi kullanarak değerlendirmişlerdir. CHAID analizi sonuçları incelendiğinde, ergin çağ süt verimi üzerinde laktasyon süresinin 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 belirtilmiştir.

Mendeş ve Akkartal, (2009), üç farklı mevsimde yetiştirilen (ilkbahar = 66, yaz = 66 ve kış = 92) 224 adet Ross-308 etçi piliçlerin kesim ağırlıklarının tahmin edilmesinde ve kesim ağırlığını etkileyen değişkenlerin belirlenmesinde, regresyon ağacı analiz metodunu kullanmışlardır. Araştırıcılar, kesim ağırlığını tahmin etmek için 2. hafta canlı ağırlığı, bacak uzunluğu, bacak genişliği, göğüs kemiği uzunluğu, göğüs genişliği, göğüs çevresin ve vücut uzunluğu özeliklerini kullanmışlardır. Analiz sonucunda, incelenen bu bağımsız değişkenden sadece canlı ağırlık, göğüs kemiği uzunluğu, bacak genişliği ve göğüs çevresinin kesim ağırlığı üzerinde etkili olduğu belirtilmiştir. Çalışma sonucunda, yüksek kesim ağırlıklarının; 2. hafta canlı ağırlığı 295.95 g'dan daha ağır, göğüs kemiği uzunluğu 55.82 mm ve göğüs çevresi 14.18 cm olan piliçlerden ya da 2. hafta canlı ağırlığı 295.95 g'dan hafif ve göğüs kemiği uzunluğu 60.26 mm'den daha uzun ve bacak genişliği 8.32 cm'den daha geniş olan piliçlerden elde edildiği tespit edilmiştir.

Bakır ve ark. (2010), farklı çiftliklerde yetiştirilen Holstein-Friesian ırkı sığırlarda 305 günlük süt verimi üzerinde etkili olan; kuruda kalma süresi, laktasyon sırası, işletme, buzağılama mevsimi ve buzağılama yaşı gibi çevre faktörlerini belirlemek amacıyla regresyon ağacı metodunu kullanmışlardır. Çalışmada, 305 günlük süt verimi üzerinde kuruda kalma süresi ve laktasyon sırasının birinci derecede, işletme ve buzağılama mevsimi faktörlerinin ikinci derecede ve buzağılama yaşı faktörü ise üçüncü derecede etkili faktörler olduğu tespit edilmiştir.

Topal ve ark. (2010), Gümüşhane ili Kelkit ilçesinde 206 adet İsveç kırmızısı sığırlarda doğum ağırlıkları ve gerçek süt verimini etkileyen faktörlerin etkilerini belirlemek amacıyla regresyon ağacı analiz tekniğini uygulamışlardır. Çalışmada, doğum ağırlığını birinci dereceden etkileyen faktörün doğum tipi olduğunu ve bunu

doğum mevsimi, cinsiyet faktörlerinin izlediğini saptamışlardır. Doğum ağırlığı üzerinde etkili olan bir başka faktör de anaların doğum sırasındaki vücut kondüsyon puanı olduğunu belirlemişlerdir. Ayrıca, laktasyon süresi ve pik süt veriminin ise gerçek süt verimi üzerinde etkili olduğunu belirtilmiştir.

Oruçoğlu (2011), Holstein ırkı ineklerin 305 günlük süt verimini etkileyen çevre faktörlerini regresyon ağacı analizi yöntemi ile incelemiştir. Laktasyonlara göre; ilk buzağılama yaşı (ay), buzağılama yaşı (ay), buzağılama aralığı (gün) ve 305 günlük süt verimi ortalamaları arasında önemli düzeyde farklılık bulunmuştur. Doğum ile ilk kontrol arası süre (gün) ortalamaları ve kuruda kalma süresi (gün) ortalamaları arasında önemli düzeyde farklılık bulunmamıştır. Regresyon ağacı analizinde, 305 günlük süt veriminde etkili olabilecek bütün bağımsız değişkenler modele dâhil edilerek analiz edilmiştir. Analiz sonucu elde edilen tüm regresyon ağaçlarında risk oranları yüksek bulunmuştur. Risk oranlarının yüksek bulunması, 305 günlük süt verimindeki varyasyonun fazla ve kullanılan laktasyon sayısının az olmasından kaynaklanabileceği belirtilmiştir. Bağımlı değişkendeki varyasyonun yüksek ve örnek büyüklüğünün az olduğunda regresyon ağacının kullanılması halinde testin gücünün yeterli olamayacağı ifade edilmiştir.

Mohammad ve ark. (2012), Pakistan'da yaşları 11-13 ay arasında değişen 239 baş farklı ırklarda erkek koyunlar (Mengali-48, Balochi-48, Harnai-48, Beverigh-47, Rakshani-48) üzerinde yürüttükleri bir çalışmada regresyon ağacı yöntemiyle; cidago yüksekliği, vücut uzunluğu ve göğüs çevresi gibi vücut ölçülerini kullanarak canlı ağırlık tahmini yapmışlardır. Regresyon ağacı sonuçlarına göre, canlı ağırlıktaki varyasyonun % 72'si cidago yüksekliği, vücut uzunluğu, göğüs çevresi özellikleri ve ırk faktörü tarafından açıklanmıştır. Bütün özellikler ve ırk faktörü önemli bulunmuştur. Tüm koyunlar içerisinde, göğüs çevresi 89 cm'den geniş olan koyunların en yüksek canlı ağırlığa (48.6 kg) sahip olduğu görülmüştür. Regresyon ağacı diyagramı incelendiğinde, canlı ağırlık üzerinde en etkili değişkenin, göğüs çevresi olduğu tespit edilmiştir. Çalışmada, göğüs çevresinin artmasıyla birlikte canlı ağırlıkta da artış görüldüğü bildirilmiştir.

Çak ve ark. (2013), İsviçre Esmeri sığırlarda laktasyon süt verimini etkileyen faktörlerin (laktasyon süresi, laktasyon sırası ve buzağılama mevsimi) etkilerini regresyon ağacı metodu kullanarak belirlemeye çalışmışlardır. Çalışmada, laktasyon süt verimi üzerinde laktasyon süresinin birinci derecede etkili olduğu, laktasyon sırasının ve buzağılama mevsiminin ise ikinci derece etkili olduğu bildirilmiştir.

Eyduran ve ark. (2013a), 46 baş Türk Saanen keçisinde laktasyon süt verimi, laktasyon uzunluğu, somatik hücre sayısı ve meme özellikleri (meme üst yüksekliği, meme genişliği, meme uzunluğu, meme çapı, meme hacmi) arasındaki ilişkiyi, farklı istatistiksel metotlarla (çoklu doğrusal regresyon, stepwise regresyon, çoklu regresyon analizinde faktör analizi ve regresyon ağacı) incelemişlerdir. Regresyon ağacı metoduna ait sonuçlar incelendiğinde, en yüksek süt verimi (297.357 *l*), meme açısı 50^0 dereceden geniş keçilerden elde etmişlerdir. Meme açısı 50^0 dereceden dar olan keçilerin süt veriminin laktasyon süresinden etkilendiği tespit edilmiştir ve laktasyon süresinin de 191 günden uzun olanların süt verimi (243.66 *l*), laktasyon süresinin 191 ve 191 günden kısa olanlardan yüksek olmuştur (192.25 *l*). Ayrıca meme açısı 50^0 dereceden daha dar ve laktasyon süresi 191 ve 191 günden küçük olan keçilerin süt verimi meme hacmi özelliğinden etkilenmiştir (Düz.P<0,05).

Eyduran ve ark. (2013b), regresyon ağacı analiz metodundan yararlanarak Brown Swiss sığırlarda 305 günlük süt verimi ile kuruda kalma süresi, laktasyon sırası, buzağılama mevsimi, yılı ve aralığı arasındaki ilişkileri irdelemişlerdir. Çalışmada, 645 Brown Swiss sığıra ait 1884 laktasyon kaydı kullanılmıştır. Regresyon ağacı diyagramından, süt verimini en çok etkileyen faktörün buzağılama yılı olduğu (P<0.01), ikinci dereceden etkili olan faktörlerin ise laktasyon sırası (P<0.01) ve buzağılama aralığı (P<0.01) olduğu belirlenmiştir. 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 olarak 2002 ve 2003 yıllarına ait laktasyon kayıtlarına sahip olan ineklerin süt verimi, buzağılama aralığı tarafından etkilenmiştir.

Yılmaz ve ark. (2013), İsviçre Esmeri ırkı sığırların 1984 - 2010 yıllarına ait kayıtlar üzerinde yaptıkları çalışmada, buzağılama mevsimi, buzağılama yılı ve buzağı cinsiyeti faktörlerinin doğum ağırlığı üzerindeki etkilerini regresyon ağacı metoduyla incelemişlerdir. Çalışmada, doğum ağırlığı üzerinde birinci derecede doğum yılının etkili olduğu, cinsiyetin ve buzağılama mevsiminin ise ikinci derece etkili faktörler olduğu belirlenmiştir. Yapılan regresyon ağacı analizi sonucunda, en yüksek doğum ağırlığı ortalaması 1991-2004 yılları arasında doğan buzağılardan elde edilmiştir. 1991-1995 yılları arasında doğan buzağılardan oluşan alt grubun doğum ağırlığı üzerine cinsiyetin etkisi önemli bulunmuştur (Düz.P<0,05). Buzağılama mevsimi, 2002 ile 2004 yılları arasında doğan buzağıların oluşturduğu grubun doğum ağırlığı üzerinde etkili olmuştur.

2.2. Diğer Alanlarda Yapılan Çalışmalar

Oğuzlar (2004), hane halkı işgücü anketi sonuçlarını SRA analizi yöntemi metodunu kullanarak bağımlı değişken (iş arayan ve aramayan grup) ile cinsiyet, bitirilen yaş, hane halkı reisine yakınlık, en son bitirilen okul, medeni durum gibi bağımsız değişkenler arasındaki ilişkiyi değerlendirmiştir. Analiz sonucunda, erkeklerin daha fazla iş aradığını, hane halkı reisinin çocuğu, gelini, torunu veya diğer akrabası olan kadınlar daha fazla iş aradığını, daha yüksek tahsil görmüş grupta yer alan kişilerin daha düşük tahsil görenlere nazaran daha çok iş aradıkları belirlenmiştir. Kadınlar için iş arama değişkeni üzerinde etkili ikinci değişken hane halkı reisine olan yakınlıkları iken, erkekler için ikinci önemli değişkenin medeni durum olduğu elde etmiştir.

Temel (2004), Nöroloji bölümünde 206 denek üzerinde yaptığı anket sonuçlarını kullanmış ve deneklerin RLS (Restless Legs Symptoms) hastası olup olmama durumuna etki eden değişkenleri sınıflama ağacı ile tespit etmiştir. RLS hastalığında gebelik ve menopoz önemli birer risk faktörü olduğundan, kadınlara özel ve genel olmak üzere iki ayrı analiz gerçekleştirmiştir. Birinci analiz; birinci derece akrabalarında RLS hastalığı olup olmama durumu, kilosu, günde kaç saat uyuduğu, ayda kaç gece uykudan uyandığı ayda kaç gündüz saatlerinde uyukladığı, ayda kaç gün moralinin bozuk olduğu, genel olarak sağlık durumu, mesleği ve Parkinson ilacı kullanıp kullanmaması soruları ve ikinci analizde de birinci analizden farklı olarak deneklere; gebelik geçirip geçirmediği, depresyona girip girmediği, yaşı ve boyu sorulan sorular ile RLS grubunu ve RLS hastası olmayan grubunu birbirinden ayıran önemli bağımsız değişkenler olarak tespit etmiştir. Ayrıca araştırıcı diğer sınıflama ve tahmin problemlerinde sorun oluşturan bağımsız değişkenler arasındaki etkileşimi, SRA için bir sorun teşkil etmediğini aksine modeli daha güvenilir kıldığını bildirmektedir.

Emel ve Taşkın (2005), yapmış oldukları çalışmada, ele alınan bir perakendeci işletme için işletmenin müşterilere göre kişiselleştirilmiş satış hareketlerini içeren veri tabanından yararlanarak ayrıntılı ve göreceli ölçüm sonuçlarını içeren bir satış analizini ele almışlardır. Oluşturulacak sınıflama tipi satış tahmin modeli için SRA karar ağacı tekniğinden yararlanmışlardır. Yapılan SRA karar ağacı tekniği uygulaması sonucunda harcama tutarına göre müşterileri sınıflara ayırmışlardır. Müşterilerin tanımlanması ve hedef müşteri grubunun belirlenmesi, hatta belirlenen hedef pazarların adlandırılabilmeleri, özel teklif sunmaya değecek müşteri veya adayların seçilmesine, müşterilerin işletmeye bağlılığının güçlendirilmesine ve tekrar satın almalarının

sağlanmasına veya müşteri portföyünün değiştirilmesine yönelik etkin pazarlama stratejilerinin oluşumuna önemli katkıda bulunduğunu belirtmişlerdir. Ayrıca, elde edilen sağlıklı bilgilere dayalı pazarlama stratejileri envanter, dağıtım, müşteriye hızlı cevap verebilme gibi işletmenin bir çok politikasını olumlu yönde etkilediğini söylemişlerdir. Böylece maliyetlerin düşürülmesi sureti ile kar artışı olacağını belirtmişlerdir.

Kayri ve Boysan (2008), Yüzüncü Yıl Üniversitesi'nden 437 öğrenciye (313 erkek, 124 bayan) uygulanan sınırlılık şemaları envanteri, Genel öz yeterlilik ölçeği ve beck depresyon envanterine verilen cevaplardan oluşturdukları veri setini SRA tekniği ile analiz etmişlerdir. çalışmada bağımlı ve bağımsız değişkenler arasındaki ilişki bir karar ağacı yapısı şeklinde modellenmiştir. SRA yöntemi, bu çalışmada sınırlılık algısının, depresyon için bir bilişsel yatkınlık faktörü olduğunu tespit etmiştir.

Kıran (2010), yapmış olduğu çalışmada sınıflama ve regresyon ağaçları (SRA) yöntemiyle lojistik regresyonun sınıflama özelliklerinin karşılaştırılmasını amaçlamıştır. Bu kapsamda 2007–2009 yılları arası Sosyal Güvenlik Kurumu ilaç provizyon sisteminden alınan solunum sistemi hastalıklarında reçeteye yazılan antibiyotikler içerisinden, penisilin gurubu antibiyotik kullanan hastaların profilini belirlemek amacıyla bir uygulama yapılmış ve çalışmaya alınan veri seti SRA analizinin lojistik regresyon analizine göre daha iyi bir doğru sınıflandırma oranına sahip olduğunu göstermiştir.

Kavzoğlu ve ark. (2012), Trabzon iline ait litoloji, eğim, arazi örtüsü, bakı, yükseklik ve yol ağı katmanlarından faydalanılarak regresyon ağaçları yönteminin heyelan risk potansiyelinin analizindeki etkinliğini araştırmışlardır. Yapmış oldukları çalışmada elde edilen sonuçlar, regresyon ağaçları ile üretilen duyarlılık haritasının doğruluğunun lojistik regresyon yöntemiyle üretilene göre %5 daha yüksek olduğunu göstermişlerdir. Sonuç olarak, Trabzon ili için heyelan riski yüksek bölgelerin şehrin Kuzey-Batı ve Kuzey-Doğu kesimlerinde yoğunlaştığını şehrin Güney ve Güney-Doğu kesimlerinde ise heyelan riski düşük alanlar belirtmişlerdir. Heyelan duyarlılığının belirlenmesinde dikkate alınan risk faktörleri göz önüne alınarak tematik haritalar değerlendirildiğinde bölgede eğimin yüksek, yol ağının sık ve yüksekliğin düşük olduğu bölgelerde heyelan riskinin daha yüksek olduğu, tematik haritalarda yüksekliğin arttığı, ormanlık ve kayalık arazi örtüsüne sahip düşük eğimli bölgelerde heyelan riskinin azaldığı kanısına varmışlardır. Araştırmacılar, regresyon ağaçları yönteminin çok

değişkenli konumsal verilerin modellenmesinde etkin bir yaklaşım olduğu belirtmişlerdir.

Özkan (2012), orman ekosistemlerinde hedef türlerin yetişme ortamı özelliklerine göre modellenmesi ile ilgili parametrik olmayan yöntem olan SRA metodu kullanmıştır. Bu yöntem kullanılarak hem kategorik (sınıflandırma ağacı) hem de sürekli (regresyon ağacı) bağımlı değişkenler modellemiştir. Örnek göstermek amacıyla Yukarıgökdere yöresinde Toros sedirine ve tür çeşitliliğine ilişkin sınıflandırma ağacı yöntemi uygulamıştır. Toros sedirinin dağılımında başta yükselti (ELVN) olmak üzere sıcaklık indisi (HI), yıllık toplam yağış (BIO12), eğim (SLOP), arazi şekli (LFC) ve yamaç konumu indisi (TPI) etkili olduğunu saptamıştır. Regresyon ağacının uyguladığı tür çeşitliliği dağılımında da yükselti (ELVN) en önemli yetişme ortamı faktörü olduğunu ve mevsimsel sıcaklık sapması (BIO4), radyasyon indeksi (RI), eğim (SLOP), arazi şekli (LFC), yıllık toplam yağış (BIO12), en sıcak üç aylık dönemin yağışı (BIO18), en sıcak üç aylık dönemin ortalama sıcaklığı (BIO10) ve anakaya (ROCK) tür çeşitliliği ağacı modelinin oluşmasında katkıda bulunan diğer değişkenler olduğunu tespit etmiştir. Sonuç olarak orman ekoloji alanında ağaçlandırma, restorasyon ve koruma çalışmalarında günümüz ve geleceğe yönelik doğru kararların verilebilmesi için modellerin iklim değişim senaryolarının dijital verilerine uyarlanmasında SRA yönteminin başarılı bir şekilde kullanılabileceğini bildirmiştir.

Avcı ve Altay (2014), 1990-2010 yılları arasında Türkiye, Arjantin, Meksika, Malezya ve Tayland'da yaşanan finansal krizlerin öngörülebilirliğini regresyon ağaçları modeli ile analiz etmişlerdir. Modellerde kriz dönemlerini belirlemek amacıyla finansal baskı endeksi hesaplanmış ve bu endeksi açıklamak üzere literatürden on beş farklı gösterge seçilmişlerdir. Finansal krizlerin öngörüsünde en başarılı bulunan göstergeler para piyasası baskı endeksi, M2'nin rezervlere oranı ve yurtiçi kredilerin endüstriyel üretime oranı olarak bulmuşlardır. 1994 ve 2001 Türkiye, 2002 ve 2009 Arjantin, 1994 ve 2009 Meksika, 1997 ve 2009 Malezya ve 1997 Tayland krizleri regresyon ağaçları modeli ile farklı göstergeler tarafından başarılı sonuç elde etmişlerdir.

3. MATERYAL VE YÖNTEM

3.1. Materyal

Araştırmanın hayvan materyalini Afyonkarahisar İli Bolvadin İlçesinde bulunan Avşar Tarım ve Hayvancılık İşletmesinde yetiştirilen 38 baş Siyah Alaca, 27 baş Simental, 23 baş Montofon ve 15 baş Melez erkek sığırlar oluşturmuştur. Bu hayvanlardan çeşitli vücut ölçüleri (cidago yüksekliği (CY), sırt yüksekliği (SIY), besi süresi (BS), sağrı yüksekliği (SAY), kuyruk sokumu yüksekliği (KSY), vücut uzunluğu (VU), arka sağrı genişliği (ASG), göğüs derinliği (GD), ve göğüs çevresi (GÇ)) ve canlı ağırlık değerleri (besi sonu canlı ağırlık (BSCA)) elde edilmiştir.

3.1.1. Sığırlardan alınan parametreler

Ele alınan parametreleri ölçmek için Şekil 3.1.'deki gibi 0.5 cm hassasiyetinde ölçü bastonu ve Şekil 3.2.'deki gibi 0.5 cm hassasiyetinde ölçü şeridi (mezura), tartmak için ise Şekil 3.3.'deki gibi 1 kg hassasiyetinde canlı hayvan kantarı (baskül) kullanılmıştır. Besi süreleri ile ilgili bilgiler işletmede bulunan kayıt defterinden alınmıştır.



Şekil 3.1. Ölçü bastonu



Şekil 3.2. Ölçü şeridi (mezura)



Şekil 3.3. Canlı hayvan kantarı (baskül)

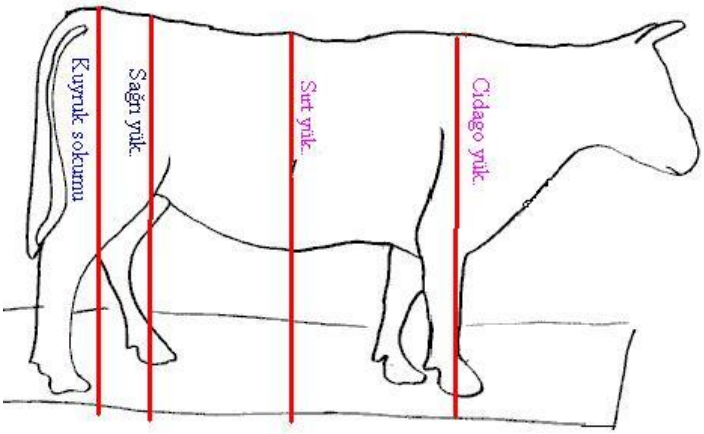
Besi sonu canlı ağırlık (BSCA): Hayvanlar basküle çıkartılıp sabit durmaları sağlanılarak tartılmış ve besi sonu canlı ağırlıkları (kg) belirlenmiştir.

Cidago yüksekliği (CY): Şekil 3.4.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar dik pozisyonda ölçü bastonu yardımı ile cidagonun en yüksek noktasından yere kadar olan dikey mesafesi ölçülerek (cm) belirlenmiştir.

Sırt yüksekliği (SIY): Şekil 3.4.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar dik pozisyonda ölçü bastonu yardımı ile son sırt omuru diken çıkıntısından yere kadar olan dikey mesafe ölçülerek (cm) belirlenmiştir.

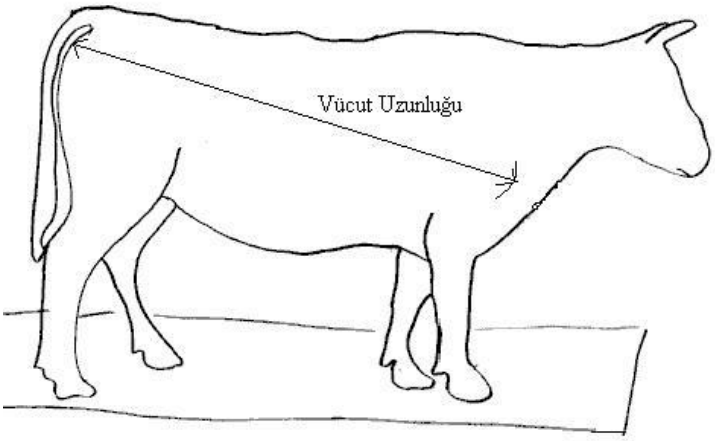
Sağrı yüksekliği (SAY): Şekil 3.4.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar dik pozisyonda ölçü bastonu yardımı ile sağrı kemiğinin (acez omurları) en yüksek noktasından yere kadar olan dikey mesafe ölçülerek (cm) belirlenmiştir.

Kuyruk sokumu (KSY): Şekil 3.4.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar dik pozisyonda ölçü bastonu yardımı ile kuyruğun başladığı en yüksek noktadan yere kadar olan düz mesafe ölçülerek (cm) belirlenmiştir.



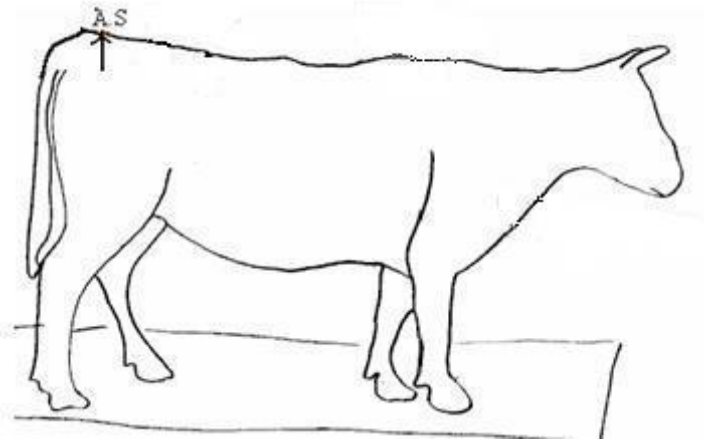
Şekil 3.4. Yükseklik ölçüleri

Vücut uzunluğu: Şekil 3.5.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar dik pozisyonda ölçü bastonu yardımı ile omuz ucundan (*Atriculus huneri*) oturak yumrusu ucuna (*Tuber ichii*) kadar olan meyilli bir hattı ölçülerek (cm) belirlenmiştir.



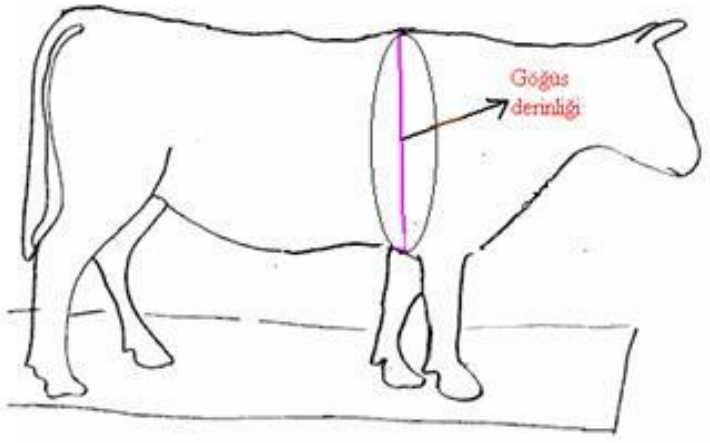
Şekil 3.5. Vücut uzunluğu

Arka sağrı genişliği (*Tuber ichii*) : Şekil 3.6.'da gösterildiği gibi düz bir zemin üzerinde hayvanlar normal duruş pozisyonunda ölçü bastonu yardımı ile Tuber ichii'lerin dış kenarları arası kalan kısım ölçülerek (cm) belirlenmiştir.



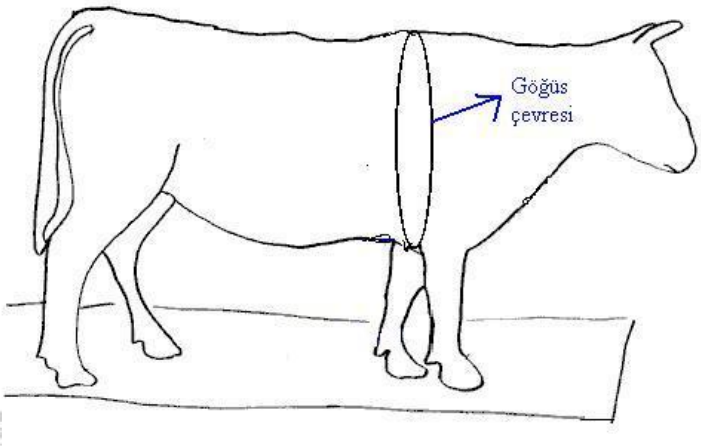
Şekil 3.6. Arka sağrı Genişliği

Göğüs derinliği: Şekil 3.7.'de gösterildiği gibi düz bir zemin üzerinde hayvanlar normal duruş pozisyonunda ölçü bastonu yardımı ile omuz (kürek kemiği) arkasında cidagonun en yüksek noktasından göğüs kemiğine kadar olan mesafe ölçülerek (cm) belirlenmiştir.



Şekil 3.7. Göğüs derinliği

Göğüs çevresi: Şekil 3.8.'de gösterildiği gibi normal duruş pozisyonunda bulunan hayvanlar kürekler arkasından ölçü şeridini (mezura) gövde etrafında dolaştırmak suretiyle ile ölçülüp (cm) belirlenmiştir.



Şekil 3.8. Göğüs çevresi

3.2. Yöntem

3.2.1. Sınıflama ve regresyon ağacı (cassification and regression tree)

Sınıflama ve regresyon ağacı, bütün bağımsız değişkenleri kullanıp verileri alt gruplara ayırarak oluşturulan bir ağaçtır. Sınıflama ve regresyon ağaçlarının en başında, herhangi bir parçalanma içermeyen ve bağımlı değişkenin yer aldığı kök düğümü bulunur. İlk olarak bu kök düğümü iki parçaya ayrılır. Bu iki parçaya, ebeveyn dalı adı verilir. Regresyon ağacının oluşturulmasında temel prensip, cevap değişkeninde maksimum homojenliği sağlayacak şekilde yinelemeli olarak iki yavru düğüme ayrılmasıdır. Ayrılma ve bölünme sonucu oluşan düğümler, alt küme olarak da adlandırılır (Mendeş ve ark., 2009). Ağacın oluşturulmasındaki temel hedef, cevap değişkenlerinde yinelemeli olarak oluşturulan herhangi bir yavru düğümde homojenlik mümkün olduğunca sağlanmışsa, bu düğümlerde artık parçalanma süreci sone erer ve bu düğüm terminal olarak adlandırılır (Oruçoğlu, 2011). Bu süreçte; program, modele alınan bütün açıklayıcı değişkenleri test ederek, sonuçta yeni oluşacak olan düğümde, en yüksek homojenliği sağlayacak şekilde açıklayıcı değişkenin kesim değerini (eğer açıklayıcı değişken kategorik ise kategoriyi) belirler (Mendeş ve ark., 2009).

İlk defa Morgan ve Sonquist (1963) tarafından ortaya atılan (Temel, 2004), Breiman ve ark., (1984) tarafından yazılan "Classification ve Regression Trees" isimli eser ile gelişme sağlayan sınıflandırma ve regresyon ağaçları, veri madenciliğinin sınıflandırma ile ilgili konuları arasında yer almaktadır. Sınıflandırma ağaçlarında kullanılabilecek birçok safsızlık ölçüsü (Gini, Twoing, Ordered Twoing ve Least

Squared Deviation) bulunmaktadır (Kurt ve ark., 2008). Burada amaç hedef değişkeni açısından mümkün olduğunca homojen veri kümeleri üretmektir.

Sınıflama ve regresyon ağaçlarını oluşturabilmek için ağaca önce büyüme daha sonra budama işlemi uygulanır (Küçükoğlu, 2010). Regresyon ağacı analizinin budama sürecinde yüksek hatalı sınıflandırmaya neden olan diğer bir ifade ile modelin tahmin edici değerine az katkıda bulunan düğümler ya da dallar atılmaktadır. Budama işlemi, cost-complexity parametresine göre yapılmakta, bu parametre, değer ya da maliyet fonksiyonu, hatalı sınıflandırma ve ağaç büyüklüğü (regresyon ağacındaki düğüm sayısı) değerleri dikkate alınarak hesaplanmaktadır (Mendeş ve ark., 2009). Budama işlemi, ağaçta oluşmuş fakat sonucu etkilemeyen ve sınıflanmaya herhangi bir katkısı olmayan dalların ağaçtan alınmasıyla bir bakıma ağacın gereksiz ayrıntılardan arındırılmasıdır. Budama uygulanmasının amacı önceki aşamalarda modele dahil edilmiş değişkenlerin ileriki aşamalarda tekrar modele dahil olmasını engellemektir. Budama süreci, en az katkı sağlayan düğümden başlayarak, önemli katkı sağlayan düğümler kalıncaya kadar devam ettirilir. Budanmış ağaç diğer yarısına gelene kadar tekrar budanır. Bu budama işlemi ağacın boyu artık değişmeyene kadar devam eder (Küçükoğlu, 2010).

Regresyon ağacı yöntemi ile sınıflama ağacı için hatalı sınıflandırma oranı, regresyon ağacı içinde model tarafından açıklanamayan varyans yani hata varyansının hesaplanabileceğini, dolayısıyla, regresyon analizindeki R^2'ye benzer ölçü ile modelin belirleme katsayısı hesaplanabilmektedir (Mendeş ve ark., 2009).

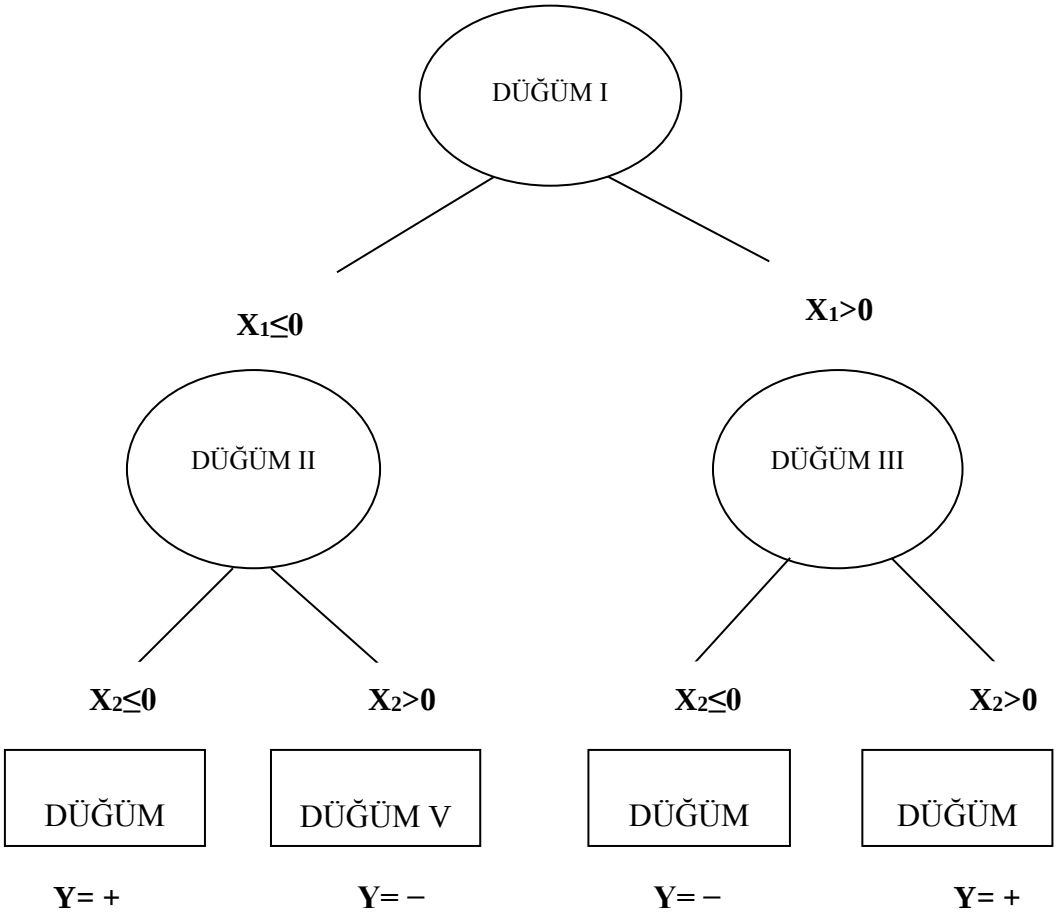
Regresyon ağaçları, sınıflandırmaya dayalı en çok kullanılan model olmasındaki en büyük etken, kullanımı ve sonuçlarının anlaşılır olmasından dolayıdır. Temel olarak regresyon ağaçlarından sonuç elde etmenin 2 aşaması bulunmaktadır. Birincisi ağacın kurulmasıdır. Regresyon ağacının kurulmasında, daha hızlı ve güvenli bir sonuç alınması için verileri en büyük parçaya bölecek sorular kök düğümden itibaren sorulmaya başlanmalıdır. Regresyon ağacı yapısı oluşturulduktan sonra, ikinci aşamaya geçilip eldeki veri, ağacın uygun dalına yerleştirilir.

Bir ağaç modelinde Y bağımlı, X_1 ve X_2 ise bağımsız değişken olarak ele alınırsa, X_1 ve X_2; $[-1; +1]$ tanım aralığında değişen düzgün (uniform) olasılık dağılımından tesadüfi olarak seçilen n_1 ve n_2 büyüklükteki örneklerin içerdiği değişken değerlerine çarpma kuralı uygulanarak aşağıdaki sonuçlar elde edilir. Bu sonuçlara göre;

$X_1.X_2 \geq 0$ ise Y' nin değeri pozitif,

$X_1.X_2 \leq 0$ ise Y' nin değeri negatiftir.

Bu şekilde bağımlı değişkenin pozitif ve negatif olmak üzere iki seviyesi oluşur (Temel, 2004).



Şekil 3.9. Örnek Ağaç Modeli (Temel, 2004)

Ağaç modellerinde karar verme noktalarına düğüm denmektedir. Şekil 3. 9'daki ağaç modelinde başlangıç düğümü, gözlem değerlerinin tümünü ihtiva eden ve en karmaşık düğüm olan Düğüm I'dir. Bu düğüme kök düğümü ya da aile düğümü de denir. Kök düğümü iki alt düğüme (çocuk düğümüne) bölünür. Şekil 3. 9'daki ağaç modelindeki çocuk düğümleri Düğüm II ve Düğüm III'tür. Çocuk düğümlerinde henüz karar verme gerçekleşmemiştir. Aile düğümünden her çocuk düğümüne bölünme gerçekleştiği için çocuk düğümü aile düğümüne göre daha homojendir. Daha sonra çocuk düğümleri ayırma kriterleri dikkate alınarak karar noktalarına yani terminal düğümlere ulaşılır. Terminal düğümlerde ele alınan özelliklerin sınıf üyelikleri

tanımlanır. Terminal düğümler ağaçtaki en homojen düğümler olduğu için, daha sonra bölünme gerçekleşmez (Temel, 2004).

Ağaç modellerinde genel olarak daire ile gösterilenler çocuk düğümleri, kare ile gösterilenler ise terminal düğümlerdir. Şekil 3. 9.'daki ağaç modelinde Düğüm IV, Düğüm V, Düğüm VI ve Düğüm VII terminal düğümleri göstermektedir (Temel, 2004).

Ağaç modellerinde, başlangıç düğümünden başlayarak ikili tekrarlı ayırmalarla daha homojen alt gruplara ulaşılıp karar noktalarında bağımlı değişkenin durumu tanımlanmaktadır. Bu şekilde regresyon ağaçlarındaki düğüm noktalarında yer alan gözlemler sahip oldukları bağımsız değişkenin değerlerine göre iki çocuk düğümden uygun olana atanırlar. Düğüm I'deki gözlemler eğer sıfırdan küçük veya sıfıra eşit iseler Düğüm II'ye, sıfırdan büyük iseler Düğüm III'e atanırlar. Buna benzer şekilde Düğüm II'deki gözlemler bağımsız değişkenin aldığı değere göre Düğüm IV ve Düğüm V'e, Düğüm III'deki değerler ise Düğüm VI ve Düğüm VII'ye atanırlar (Temel, 2004).

3.2.2. İstatistiksel analizler

Bu çalışmada, Türkiye'nin Afyon ilinin Bolvadin ilçesinde yetiştirilen 103 baş (Siyah Alaca, Simental, Esmer ve Melez) tosunun besi sonu canlı ağırlığını (BSCA) etkileyen vücut ölçüleri belirlenmesi amaçlanmıştır. Bu amaca ulaşmak için, veri madenciliği algoritmalarından CHAID algoritması kullanılmıştır. Regresyon ağacı metodu olarak CHAID algoritması ile regresyon ağacı oluşturulmuştur. Bu çalışmada en iyi model değerlendirme kriterlerini elde etmek amacıyla CHAID algoritması için ebeveyn ve çocuk düğümde bulunan minimum tosun sayıları 8:4 olarak ayarlanmıştır.

Bağımlı değişken olarak, BSCA kullanılmıştır. Bu çalışmada kullanılan bağımsız değişkenler sırasıyla, cidago yüksekliği (CY), sırt yüksekliği (SIY), sağrı yüksekliği (SAY), kuyruk sokumu yüksekliği (KSY), vücut uzunluğu (VU), arka sağrı genişliği (ASG), göğüs derinliği (GD), ve göğüs çevresi (GÇ) ve besi süresi (BS) olmuştur.

3.2.3. Model değerlendirme kriterleri

Bu çalışmada, CHAID algoritması için ebeveyn ve çocuk düğümlerinde minimum tosun sayıları 8:4 olarak ayarlanmıştır. CHAID algoritmasının etkinliğini belirlemek amacıyla model değerlendirme kriterleri olarak belirleme katsayısı,

düzeltilmiş belirleme katsayısı, ve gerçek ve tahmin edilen BSCA değerleri arasındaki Pearson korelasyon katsayısı kullanılmıştır (Grzesiak ve Zaborski, 2012). Bu kriterlere ilişkin formüller Çizelge 3.1'de verilmiştir.

Çizelge 3.1 Model Değerlendirme Kriterleri

Model Kalite Kriteri	Formül
Belirleme Katsayısı (%)	$R^2(\%) = \left[1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}\right] * 100$
Düzeltilmiş Belirleme Katsayısı (%)	$Düz\text{-}R^2(\%) = \left[1 - \frac{\frac{1}{n-k-1}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}\right] * 100$
Gerçek ve Tahmin BSCA değerleri arasındaki Korelasyon katsayısı	$r = \frac{\sum d_{Y_i} d_{\hat{Y}_i}}{\sqrt{\sum d_{Y_i}^2 \sum d_{\hat{Y}_i}^2}}$

Çizelge 3. 1'de;

Yi: i. tosunun BSCA değerini, $\hat{Y}_i$: *i.* tosunun BSCA tahmin değerini, $\bar{Y}$: Gerçek BSCA değerlerinin ortalamasını, ε_i : *i.* tosunun hata terim değeri, $\bar{\varepsilon}$: Hata terim değerlerinin ortalamasını, *n:* toplam tosun sayısını ve *k*: modele giren önemli bağımsız değişkenlerin sayısını ifade etmektedir.

Model değerlendirme kriterlerinin mümkün olduğunca yüksek olması, kullanılan CHAID algoritması ile özelleştirilen modelin daha iyi olması anlamına gelmektedir. Elde edilen veriler SPSS 22 programında analiz edilmiştir.

4. ARAŞTIRMA BULGULARI

Çalışmada kullanılan 38 baş Siyah Alaca, 15 baş Melez, 23 baş Esmer ve 27 baş Simental ırkı erkek sığırlara ait çeşitli vücut ölçüleri (cidago yüksekliği (CY), sırt yüksekliği (SIY), besi süresi (BS), sağrı yüksekliği (SAY), kuyruk sokumu yüksekliği (KSY), vücut uzunluğu (VU), arka sağrı genişliği (ASG), göğüs derinliği (GD), ve göğüs çevresi (GÇ)) ve canlı ağırlık değerlerine (besi sonu canlı ağırlık (BSCA)) ait tanıtıcı istatistikler Çizelge 4. 1.'de verilmiştir.

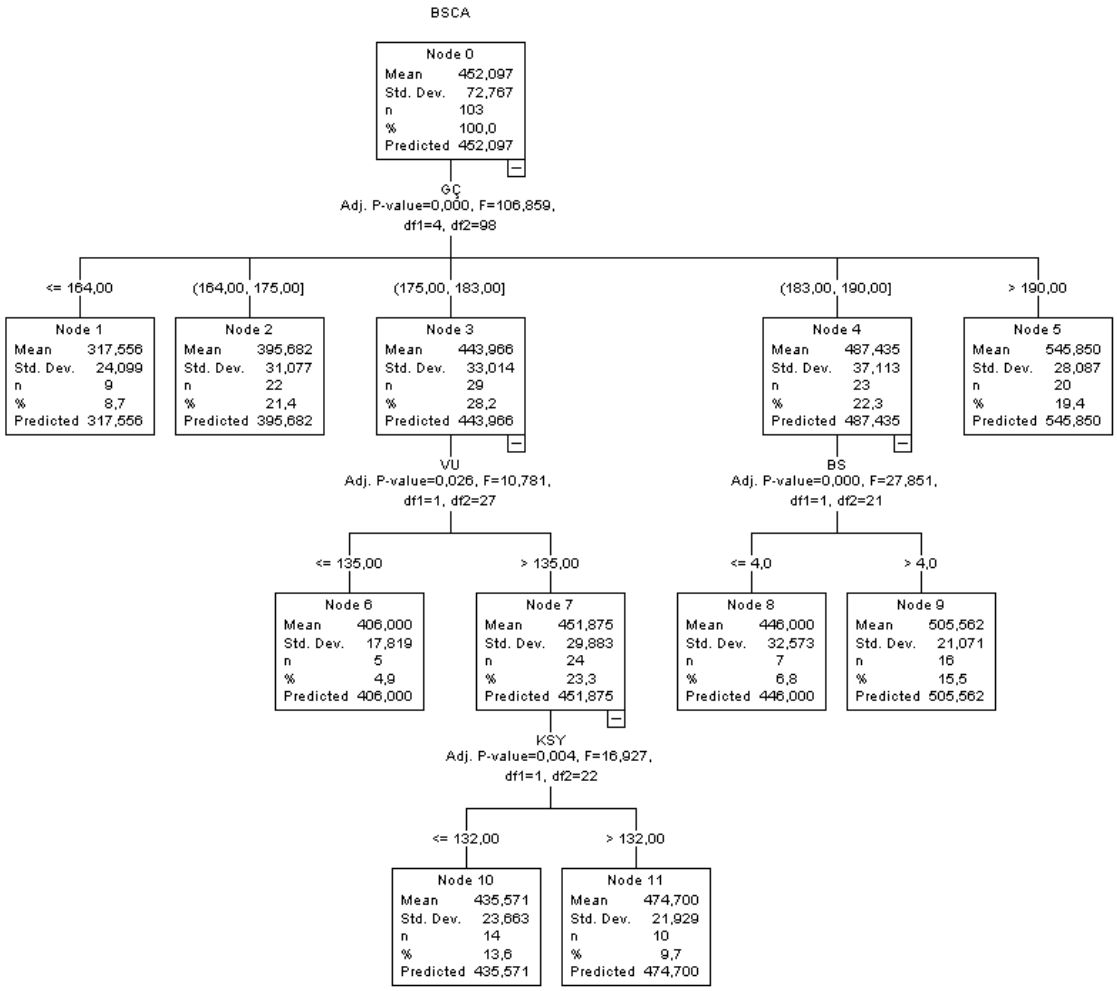
Çizelge 4. 1. Çeşitli vücut ölçüleri ve ırklara göre tanıtıcı istatistikleri

Değişkenler	Irk	N	$\overline{X} \pm S_{\overline{X}}$	Minimum	Maximum	CV
BSCA	1	38	454.1±12.4	276	571	16.80
	2	23	455.9±17.6	285	588	18.47
	3	27	458.7±13.4	313	587	15.20
	4	15	429.2±12.5	335	521	11.29
BS	1	38	4.763±0.122	4	6	15.77
	2	23	4.957±0.194	4	6	18.73
	3	27	4.815±0.142	4	6	15.28
	4	15	4.467±0.215	4	6	18.67
CY	1	38	129.11 ± 1.11^{a}	111	144	5.29
	2	23	125.80 ± 1.21^{ab}	110	134.5	4.61
	3	27	125.89 ± 0.955^{ab}	119	135	3.94
	4	15	123.63 ± 1.10^{b}	116	133	3.43
SIY	1	38	131.76 ± 0.984^{A}	114.5	146	4.60
	2	23	128.96 ± 1.19^{AB}	118	138	4.41
	3	27	128.11 ± 1.03^{AB}	119	140	4.18
	4	15	125.37 ± 1.27^{B}	117.5	134	3.91
SAY	1	38	134.84 ± 1.06^{A}	116	149	4.83
	2	23	131.52 ± 1.15^{AB}	118	143	4.19
	3	27	130.39 ± 0.992^{AB}	123	141	3.95
	4	15	128.67 ± 1.16^{B}	118	135	3.49
KSY	1	38	134.96 ± 0.936^{A}	117	147	4.27
	2	23	131.37 ± 0.934^{AB}	120	140	3.41
	3	27	130.09 ± 0.774^{B}	124	141	3.09
	4	15	129 ± 0.975^{B}	121	135	2.93
VU	1	38	143.24±1.34	124	160	5.79
	2	23	141.39±2.01	122	163	6.83
	3	27	141.70±1.31	128	155	4.80
	4	15	141.40±1.52	129	153	4.16
ASG	1	38	48.105±0.615	40	59	7.88
	2	23	49.04±1	38	58	9.81
	3	27	49.519±0.746	41	55	7.83
	4	15	48.8±0.947	44	57	7.52
GD	1	38	62.842±0.581	53	69	5.70
	2	23	61.96±1.07	50	71	8.31
	3	27	61.815±0.710	55	69	5.97
	4	15	61.933±0.808	56	67	5.05
GÇ	1	38	181.12±1.79	155	197	6.11
	2	23	180.22±2.85	154	203	7.58
	3	27	181.44±1.66	162	197	4.75
	4	15	179.27±2.02	164	191	4.37

$^{a, b}$: $P<0.05$; $^{A, B}$: $P<0.01$

Çizelge 4. 1.'den de görülebileceği gibi her bir ırk için besi sonu canlı ağırlık değerleri 454.1, 455.9, 458.7 ve 429.2 kg arasında değişmiştir.

Yapılan varyans analizi neticesinde Siyah Alaca, Esmer, Simental ve Melez ırkı erkek sığırlara ait BSCA, BS, VU, ASG, GD ve GÇ değerleri arasındaki farklar istatistik olarak önemsiz bulunmuştur. Irklar arasında CY bakımından % 5'e, SIY, SAY ve KSY için ise % 1'e göre önemli farklılıklar tespit edilmiştir.



Şekil 4. 1. Besi sonu canlı ağırlık tahmini için regresyon ağacı diyagramı

Ekonomik öneme sahip BSCA özelliğini etkileyen değişkenleri belirlemek amacıyla CHAID algoritması ile oluşturulan regresyon ağacı diyagramı Şekil 4.1'de sunulmuştur. Regresyon ağacına ait hesaplanan model kalite değerleri sırasıyla, % 87.82 R^2, % 87.32 Düz. R^2 ve gerçek ve tahmin edilen BSCA değerleri arasındaki korelasyon 0.937 olarak hesaplanmıştır. Bu değerler, BSCA özelliğinde meydana gelen toplam

varyasyonun yaklaşık % 90'ının GÇ, VU, BS ve KSY bağımsız değişkenleri tarafından açıklandığını ifade etmektedir. Şekil 4.1 incelendiğinde, oluşturulan regresyon ağacının en tepesinde bulunan ve kök düğüm olarak adlandırılan düğüm, analize dahil edilen bütün hayvanların bulunduğu grubu temsil etmektedir. Regresyon ağacı diyagramı incelendiğinde, BSCA üzerinde birinci derecede etkili bağımsız değişkenin GÇ (Düz. P=0.000), ikinci derecede etkili bağımsız değişkenlerin VU (Düz. P=0.026) ve BS (Düz. P=0.000), üçüncü derecede etkili bağımsız değişkenin KSY (Düz. P=0.004) olduğu belirlenmiştir.

Üzerinde çalışma yapılan tüm sığırlar (Düğüm 0), GÇ değişkeni bakımından 5 alt gruba (Düğüm 1, Düğüm 2, Düğüm 3, Düğüm 4 ve Düğüm 5) ayrılmıştır. Düğüm 1; göğüs çevresi GÇ $\leq$ 164 cm olan tosunların oluşturduğu alt grubu, Düğüm 2; göğüs çevresi 164 < GÇ $\leq$ 175 cm arasında olan tosunların oluşturduğu alt grubu, Düğüm 3; göğüs çevresi 175 < GÇ $\leq$ 183 cm arasında olan tosunların oluşturduğu alt grubu, Düğüm 4; göğüs çevresi 183 < GÇ $\leq$ 190 arasında olan tosunların oluşturduğu alt grubu ve Düğüm 5; göğüs çevresi GÇ $\geq$ 190 cm tosunların oluşturduğu alt grubu temsil etmektedir (Şekil 4.1). Düğüm 1 (göğüs çevresi GÇ $\leq$ 164 cm olan tosunların) BSCA ortalaması 317.556 (S=24.099) kg, Düğüm 2 (göğüs çevresi 164 < GÇ $\leq$ 175 cm arasında olan tosunların) BSCA ortalaması 395.682 (S=31.077) kg, Düğüm 3 (göğüs çevresi 175 < GÇ $\leq$ 183 cm arasında olan tosunların) BSCA ortalaması 443.966 (S=33.014) kg, Düğüm 4 (göğüs çevresi 183 < GÇ $\leq$ 190 cm arasında olan tosunların) BSCA ortalaması 487.435 (S=37.113) ve Düğüm 5 (göğüs çevresi GÇ $\geq$ 190 cm tosunların) BSCA ortalaması BSCA 545.850 (S=28.087) olarak bulunmuştur (Şekil 4.1).

Oluşturulan alt gruplardan yeterince homojen yapıya sahip olan Düğüm 1, Düğüm 2 ve Düğüm 5'e terminal düğüm adı verilir. Düğüm 1'den Düğüm 5'e doğru gidildikçe yani GÇ arttıkça BSCA ortalama değeri 317.556 kg' dan 545.850 kg'a artmıştır.

Düğüm 3 (göğüs çevresi 175 < GÇ $\leq$ 183 cm arasında olan tosunların) BSCA, VU tarafından etkilenmiştir (Düz P=0.026). Düğüm 3 (göğüs çevresi 175 < GÇ $\leq$ 183 cm arasında olan tosunların oluşturduğu alt grup) VU bakımından iki yeni alt gruba (Düğüm 6 ve Düğüm 7) ayrılmıştır. Düğüm 6; göğüs çevresi 175 < GÇ $\leq$ 183 cm ve vücut uzunluğu VU $\leq$ 135 cm olan tosunların oluşturduğu alt grubu, Düğüm 7; göğüs çevresi 175 < GÇ $\leq$ 183 cm ve vücut uzunluğu VU > 135 cm olan tosunların oluşturduğu alt grubu ifade etmektedir. Düğüm 6 ve Düğüm 7 için BSCA ortalamaları

sırasıyla, 406.000 (S=17.819) kg ve 451.875 (S=29.883) kg olarak tahmin edilmiştir. Düğüm 7 (göğüs çevresi 175 < GÇ ≤ 183 cm ve vücut uzunluğu VU > 135 cm olan tosunlar) BSCA sadece KSY tarafından etkilendiği saptanmıştır (Düz. P= 0.004).

Düğüm 7 (göğüs çevresi 175 < GÇ ≤ 183 cm, vücut uzunluğu VU > 135 cm olan tosunların oluşturduğu alt grup) KSY bakımından iki yeni alt gruba (Düğüm 10 ve Düğüm 11) ayrılmıştır. Düğüm 10; göğüs çevresi 175 < GÇ ≤ 183 cm, vücut uzunluğu VU > 135 cm ve kuyruk sokumu yüksekliği KSY ≤ 132 cm olan tosunların oluşturduğu alt grubu, Düğüm 11; göğüs çevresi 175 < GÇ ≤ 183 cm, vücut uzunluğu VU > 135 cm ve kuyruk sokumu yüksekliği KSY > 132 cm olan tosunların oluşturduğu alt grubu ifade etmektedir.

Düğüm 4 (göğüs çevresi 183 < GÇ ≤ 190 cm arasında olan tosunların) BSCA özelliği üzerinde analize dahil edilen bağımsız değişkenlerden sadece BS etkili bulunmuştur (Düz. P=0.000). Düğüm 4; BS bakımından iki yeni alt gruba (Düğüm 8 ve Düğüm 9) ayrılmıştır. Düğüm 8; göğüs çevresi 183 < GÇ ≤ 190 cm ve besi süresi BS ≤ 4 olan tosunların oluşturduğu alt grubu, Düğüm 9; göğüs çevresi 183 < GÇ ≤ 190 cm ve besi süresi BS > 4 olan tosunların oluşturduğu alt grubu ifade etmektedir. Düğüm 8 ve Düğüm 9 için BSCA ortalamaları sırasıyla, 446.000 (S=32.573) kg ve 505.562 (S=21.071) kg olarak hesaplanmıştır. Göğüs çevresi 183 < GÇ ≤ 190 cm tosunların 4 aydan daha fazla sürede beside tutulması BSCA ortalama olarak arttırdığı söylenebilir.

5. SONUÇ VE ÖNERİLER

Klasik sınıflama modellerine alternatif olarak ortaya çıkan regresyon ağacı (CHAID algoritması), modelde yer alan değişkenler üzerinde hiçbir varsayım gerektirmemesi nedeniyle, eldeki verilerin fazla sayıda ve kompleks bir yapıda olduğu durumlarda sınıflama analizlerinde kolaylıkla uygulanabilen güçlü bir tekniktir. CHAID algoritması ile oluşturulan regresyon ağaç diyagramı incelendiğinde, bağımlı değişkenleri hangi bağımsız değişkenlerin etkilediği kolaylıkla görülebilmektedir. Bu çalışmada CHAID algoritmasına ilişkin elde edilen sonuçlar aşağıda maddeler halinde özetlenmiştir.

1. Regresyon ağacına ait hesaplanan model değerlendirme kriterleri sırasıyla, % 87.82 R^2, % 87.32 Düz. R^2 ve gerçek ve tahmin edilen BSCA değerleri arasındaki korelasyon 0.937 olarak bulunmuştur.
2. BSCA üzerinde etkili bağımsız değişkenlerin sırasıyla, GÇ (Düz. P=0.000), VU (Düz. P=0.026), BS (Düz. P=0.000) ve KSY (Düz. P=0.004) olduğu belirlenmiştir.
3. Göğüs çevresi 175 < GÇ $\leq$ 183 cm arasında olan tosunların BSCA, VU tarafından etkilenmiştir (Düz. P=0.026).
4. Göğüs çevresi 175 < GÇ $\leq$ 183 cm ve vücut uzunluğu VU > 135 cm olan tosunların BSCA'nın sadece KSY tarafından etkilendiği saptanmıştır (Düz. P= 0.004).
5. Düğüm 4 (göğüs çevresi 183 < GÇ $\leq$ 190 cm arasında olan tosunların) BSCA özelliği üzerinde analize dahil edilen bağımsız değişkenlerden sadece BS etkili bulunmuştur (Düz. P=0.000).
6. Ortalama olarak, en ağır BSCA (545.850 kg); göğüs çevresi GÇ > 190 cm olan tosunların oluşturduğu alt gruptan elde edilmiştir. İkinci olarak en ağır BSCA (505.562 kg); 183 < GÇ $\leq$ 190 cm ve BS > 4 ay olan tosunların oluşturduğu alt gruptan elde edilmiştir.

Sonuç olarak, farklı vücut ölçülerinden BSCA tahmin etmek amacıyla CHAID algoritması kullanımı, hayvan ıslahı çalışmaları için tavsiye edilebilir.

KAYNAKLAR

Avcı, M.A., Altay, N.O., 2014, Finansal Krizlerin Öngörüşünde Regresyon Ağaçları Modeli: Gelişmekte Olan Ülkelere Yönelik Bir Analizi, *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 12; 191-212.

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, *J. Anim. Vet. Adv.,* 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, *J. Anim. Vet. Adv.,* 9 (1): 55-59.

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., 1984, Classification And Regression Trees, New York: *Chapman and Hall.*

Çak, B., Keskin, S., Yılmaz, O., 2013, Regression tree analysis for determining of affecting factors to lactation milk yield in Brown Swiss Cattle, *Asian J. Anim. Vet. Adv.,* 8 (4); 677-682.

Doğan, İ., 2003, Holştayn Irkı İneklerde Süt Verimine Etki Eden Faktörlerin CHAID Analizi İle İncelenmesi. *Ankara Üniv. Vet. Fak. Derg.*, 50: 65-70.

Efe, E., Bek, Y., Şahin, M., 2000, SPSS'te Çözümleri ile İstatistik Yöntemler II. Kahramanmaraş: *Kahramanmaraş Sütçü İmam Üniv. Rek.,* Yayın No:10.

Emel G. G., Taşkın Ç., 2005, Veri Madenciliğinde Karar Ağaçları ve Bir Satış Analizi Uygulaması. *Eskişehir Osmangazi Üniv. Sos. Bil. Derg., 6 (2); 221-239*

Eyduran, E., Karakus, K., Keskin, S., Cengiz, F., 2008, Determination of Factors Influencing Birth Weight Using Regression Tree (RT) Method, *J. Appl. Anim. Res.* 34: 109-112.

Eyduran, E., Yilmaz, I., Kaygisiz, A., Aktaş, Z. M., 2013a, 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 of Animal & Plant Sciences,* 23(4); 956-963, 1018-7081.

Eyduran, E., Yilmaz, I., Kaygisiz, A., Tariq, M. M., 2013b, Estimation of 305-D Milk Yield Using Regression Tree Method in Brown Swiss Cattle, *The Journal of Animal & Plant Sciences,* 23(3); 731-735.

Grzesiak, W., Zaborski, D., 2012, Example of the Use of Data Mining Methods in Animal Breeding. ISBN: 978-953-51-0720-0.

Kavzoğlu, T., Şahin, E. K., Çölkesen, İ., 2012, Heyelan Duyarlılığının İncelenmesinde Regresyon Ağaçlarının Kullanımı: Trabzon Örneği, *Harita Derg.*,147; 21-33

Kayri, M., Boysan, M., 2008, Bilişsel Yatkınlık İle Depresyon Düzeyleri İlişkisinin Sınıflandırma ve Regresyon Ağacı ile İncelenmesi, *Hacettepe Üniv. Eğitim Bil. Derg.*, 34: 168-177.

Kıran Z. B., 2010, Lojistik Regresyon ve Cart Analizi Teknikleriyle Sosyal Güvenlik Kurumu İlaç Provizyon Sistemi Verileri Üzerinde Bir Uygulama, Yüksek Lisans Tezi, *Gazi Üniv. Fen Bil. Enst. İstatistik Bölümü*, Ankara.

Kurt, I., Ture, M., Kurum, A. T., 2008, Comparing Performances of Logistic Regression, Classification and Regression Tree, and Neural Networks for Predicting Coronary Artery Disease, *Expert Systems with Applications*, 34; 366-374.

Küçükoğlu, O., 2010, Veri madenciliği yöntemlerinin hayvancılıkta kullanımı, Yüksek Lisans Tezi *Çanakkale Onsekiz Mart Üniv. Fen Bil. Enst.*, Çanakkale.

Mendeş, M., Akkartal, E., 2009, Regression tree model in predicting slaughter weight of broiler, *Italian Journal of Animal Science*, 8; 615-624.

Mohammad, M. T., 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 J. Anim. Plant. Sci.*, 22 (1); 20-24.

Morgan, J. N. and Sonquist, J. A., 1963, Problems in the analysis of survey data, and a proposal, *Journal of the American Statistical Association* 58; 415-434.

Nefeslioglu, H.A., Sezer, E., Gokceoglu, C., Bozkir, A.S., Duman, T.Y., 2010, Assessment of Landslide Susceptibility by Decision Trees in the Metropolitan Area of Istanbul, Turkey, *Mathematical Problems in Engineering*, doi:10.1155/2010/901095.

Oğuzlar, A., 2004, CART Analizi ile Hane halkı İşgücü Anketi Sonuçlarının Özetlenmesi, *İktisadi ve İdari Bilimler Dergisi*, 18 (3-4);79-90.

Oruçoğlu, O., 2011, Holstein Irkı İneklerin 305 Günlük Süt Verimini Etkileyen Çevre Faktörlerinin Regresyon Ağacı ile Belirlenmesi, Yüksek Lisans Tezi, *Süleyman Demirel Üniv. Fen Bil. Enst.*, Isparta.

Özkan, K., 2012, Sınıflandırma ve Regresyon Ağacı Tekniği (SRAT) İle Ekolojik Verinin Modellenmesi, *SDÜ Orman Fak. Derg.*, 13: 1-4

Safavian, S.R., Landgrebe, D., 1991, A Survey of Decision Tree Classifier Methodology, *IEEE Transactions on Systems Man and Cybernetics*, 21; 660-674.

Sümbüloğlu, K., Akdağ, B., 2007, Regresyon Yöntemleri ve Korelasyon Analizi, *Hatiboğlu Yayınevi*. Ankara.

Temel, Ö.G., 2004, Sınıflama ve Regresyon Ağaçları, Yüksek Lisans Tezi, *Mersin Üniv. Sağlık Bil. Enst. Biyoistatistik Anabilim Dalı*, Mersin.

Topal, M., Aksakal, V., Bayram, B., Yaganoglu, M., 2010, An analysis of the factor affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis, *The J. Anim & Plant Sci.*, 20; 63-69.

Yılmaz, I., Eyduran E., Kaygisiz, A., 2013, Determination of Non-Genetic Factors Influencing Birth Weight Using Regression Tree Method in Brown-Swiss Cattle Canadian, *Journal of Applied Sci.*. 1 (3); 382-387.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Rifat AKŞAHAN
Uyruğu : T.C.
Doğum Yeri ve Tarihi : Bolvadin 06.02.1986
Telefon : 0274 371 20 01
Faks : 0274 371 20 01
e-mail : raksahan@hotmail.com

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Lise	:	Bolvadin Anadolu Lisesi, Bolvadin, Afyonkarahisar	2004
Üniversite	:	Ege Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü	2010
Yüksek Lisans	:	Selçuk Ünivrsitesi, Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı	2015
Doktora	:		

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2011-Halen	Gıda, Tarım ve Hayvancılık Bakanlığı	Ziraat Mühendisi