

Siyah Alaca Sığırlarda Farklı Büyüme ve Gelişme Dönemlerindeki Vücut Ölçülerinden Canlı Ağırlık Tahmininde Kullanılan Veri Madenciliği Algoritmalarının Karşılaştırılması

Comparison of Data Mining Algorithms used in Predictive of Live Weight from Body Measurements in Holstein Cattle at Different Growth and Development Periods

ÖZET

Bu çalışmanın amacı 0-12 aylık yaşta farklı büyüme ve gelişme dönemindeki sığırların bazı vücut ölçümlerinden canlı ağırlık tahmininde kullanılan veri madenciliği algoritmalarının karşılaştırılmasıdır. Çalışmada 24 baş dişi ve 18 baş erkek olmak üzere toplamda 42 baş sığıra ait kimi vücut ölçülerinden göğüs çevresi (GÇ), göğüs derinliği (GD), vücut uzunluğu (VU), cidago yüksekliği (CY), sağrı yüksekliği (SY) ile cinsiyet ve yaş özelliği bağımsız değişken, canlı ağırlık ise bağımlı değişken olarak ele alınmıştır. Vücut ölçülerinden canlı ağırlığının tahmin edilmesinde ise veri madenciliği algoritmalarından Çoklu Doğrusal Regresyon (MLR), Rastgele Orman (RF), Karar Ağacı (DT) ve En Yakın Komşu (kNN) algoritmaları çapraz doğrulama (cross-validation) 5 alınarak kullanılmıştır. Vücut ölçüleri ile canlı ağırlık (CA) arasında pozitif bir korelasyon olduğu tespit edilmiştir ($P < 0.01$). CA'nın en yüksek korelasyonu GÇ ($r=0.972$) ve GD ($r=0.972$) olarak bulunurken, en düşük korelasyon ise SY ($r=0.887$)'de bulunmuştur. MLR, RF, DT ve kNN algoritmalarında belirleme katsayıları (R^2) sırasıyla % 93.90, % 91.20, % 85.70 ve % 85.60 olarak tespit edilmiştir. RF ve MLR'ye ait R^2 değerleri; kNN ve DT modellerine nazaran daha yüksek bulunurken, MSE ve MAE tahmin performans ölçütleri ise bu modellerden daha düşük olarak tespit edilmiştir. Veri madenciliği algoritmalarından MLR modelin R^2 değeri 0.939 ile en yüksek iken, en düşük değer ise kNN (0.856) modelinde tespit edilmiştir.

Sorumlu Yazar

Gizem COŞKUN

coskun.gizem96@gmail.com

ID 0000-0003-2519-7885

Yazar

Özcan ŞAHİN

osahin006@gmail.com

ID 0000-0003-2170-2055

Yazar

İlker Ali ÖZKAN

ilkerozkan@selcuk.edu.tr

ID 0000-0002-5715-1040

Yazar

İbrahim AYTEKİN

aytekin@selcuk.edu.tr

ID 0000-0001-7769-0685

Gönderilme Tarihi :

24 Mart 2022

Kabul Tarihi :

09 Ağustos 2022

Sonuç olarak; sığırların vücut ölçüleri kullanılarak canlı ağırlığı tahmin etmedeki doğruluğun yüksek olduğu ve sadece GÇ ve GD'nin ele alınması ile de doğru ve güvenilir sonuçlar elde edileceği söylenebilir.

Anahtar Kelimeler: Algoritma, Canlı Ağırlık Tahmini, Sığır, Veri madenciliği, Vücut Ölçüleri.

ABSTRACT

The aim of this study was to compare the live weight predictive performance of different data mining algorithms from some body measurements of cattle during different growth and development ages ranging from birth to 12 months. In the study, some body measurements of 42 cattle, 24 females and 18 males, were determined by chest circumference (CC), chest depth (CD), body length (BL), withers height (WH), rump height (RH), sex and age were considered as the independent variable and live weight as the dependent variable. In the estimation of body weight from body measurements, Multiple Linear Regression (MLR), Random Forest (RF), Decision Tree (DT) and Nearest Neighbor (kNN) algorithms from data mining algorithms were utilized by using cross validation 5. A positive correlation was found between body measurements and live weight (LW) ($P < 0.01$). While the highest correlation of LW was found to be CC ($r = 0.972$) and CD ($r = 0.972$), the lowest correlation was found for RH ($r = 0.887$). Determination coefficients (R^2) in MLR, RF, DT, and kNN algorithms were determined as 93.90 %, 91.20 %, 85.70 %, and 85.60 %, respectively. While R^2 values of RF and MLR were found higher than kNN and DT models, MSE and MAE estimation performance criteria were found to be lower than these models. Among the data mining algorithms, the R^2 value of the MLR model was the highest with 0.939, while the lowest value was found in the kNN (0.856) model. As a result, it can be said that the accuracy in estimating the live weight of cattle using body measurements is high, and accurate and reliable results can be obtained by considering only a CC and CD.

Keywords: Algorithm, Live Weight Prediction, Cattle, Data mining, Body Measurements.

GİRİŞ

Sürdürülebilir bir hayvancılık için sürü yönetiminde hayvan refahı, genetik ve çevresel iyileşme, hayvan sağlığı ve pazarlama gibi uygulamaları daha iyi entegre eden stratejilerin geliştirilmesi gerekmektedir. Fizyolojik ölçümler; ırk standartlarının bilinmesi, sürü yönetimindeki bakım ve besleme olanaklarındaki eksikliklerinin ortaya çıkarılması ile büyüme ve gelişme yeteneklerinin tespiti açısından önemlidir (Riva, Rizzi, Marelli, ve Cavalchini, 2004; Zülkadir, Şahin, Aytekin, ve Boztepe, 2008) Hayvanlardan alınan kimi vücut ölçüleri sürü yönetimindeki fiziksel durumları hakkında bilgiler vermekte olup elde edilen veriler ile canlı ağırlıkları arasında yüksek bir korelasyon olduğu da bilinmektedir (Çankaya, Altop, Kul, ve Erener, 2009). Bu nedenle hayvancılık işletmesindeki sürülerin canlı ağırlıklarının takibi işletme açısından önem arz etmektedir. Hayvansal üretimin temel amacı, her bir hayvanın karlılığını mümkün olduğu kadar artırmaktır. Bu bağlamda istenilen karlılık seviyesinin elde edilmesi çevre faktörlerinin optimize edilmesiyle sağlanabilir (Gül, Keskin, Biçer, Gündüz, ve Behrem, 2020). Çiftlik hayvanlarında ekonomik öneme sahip olan canlı ağırlığın bilinmesi hayvanların besin madde ihtiyacı, damızlıkta kullanma yaşının belirlenmesi, büyümesinin izlenmesi gibi sürü yönetim uygulamaları açısından gereklidir (Tüzemen, Yanar, Akbulut, Uğur, ve Aydın, 1995; Tariq, Younas, Khan, ve Schlecht, 2013).

Hayvanların canlı ağırlıkları işletmelerde genellikle kantarlar vasıtasıyla doğru bir şekilde belirlenebilmektedir. Tartıma ek olarak hayvanların farklı bölgelerinden ölçü bastonu veya ölçü şeridi ile yapılan kolay ve basit ölçümler canlı ağırlık tahmininde alternatif bir yol olarak kullanılabilir (Koç ve Akman, 2007; Tariq vd., 2013). Hayvanların büyüme ve gelişme takibinin yapılması ve vücut yapılarını ifade edebilmek için belirli aralıklarla vücut ölçülerinin alınması ile hayvanlar arası mukayeseler yapılabileceği gibi sürü yönetimindeki bakım ve besleme programlarının düzenlenmesinde de fikir edinilebileceğini ifade etmişlerdir (Zülkadir vd., 2008). Bu bağlamda vücut ölçüleri, cinsiyet, verim türü, yaş ve ırk gibi önemli faktörlere göre değişiklik göstermektedir (Pesmen ve Yardimci, 2008). Günümüze kadar canlı ağırlık ve vücut ölçüleri arasındaki ilişkiyi yorumlamak amacıyla; kanonik korelasyon (Kara-

bacak, Altay, ve Aytekin, 2019) yapay sinir ağları ve çoklu doğrusal regresyon yöntemleri (Akkol, Akilli, ve Cemal, 2017), Stepwise-regresyon yöntemi (Koç ve Akman, 2007), path analizi (Keskin, Dağ, ve Şahin, 2005) bulanık mantık yöntemi (Taşdemir, Ürkmez, ve İnal, 2011; Ameen ve Mikail, 2018), Çok Değişkenli Uyarlanabilir Regresyon Uzantıları (MARS) algoritmaları (Aytekin, Eyduran, Karadas, Aksahan, ve Keskin, 2018), CART, CHAID, Exhaustive CHAID algoritmaları (Altay, 2022) ve Bayesian Regularized Neural Network, Random Forest Regression, Support Vector Regression algoritmaları (Tırınk, 2022) gibi birçok metot kullanılmıştır.

Veri madenciliği algoritmaları, deneysel gözlemlerden veya oluşturulan veri setlerinden belirli bir işlevi öğrenme yeteneği anlamına gelen yapay zekanın bir alt dalı olup, elde edilen veri setlerinden anlaşılır bilgiler edinmek için bilgisayar algoritmalarının dizayn ve geliştirme süreçlerini amaç edinen, geniş uygulama yelpazesine sahip ve RF, kNN, MLR ve DT gibi birçok güçlü algoritması mevcut olan bir bilim dalıdır (Peng, 2015; Mahesh, 2020). Bu veri madenciliği algoritmalarından; MLR genellikle bir veya birden fazla bağımsız değişken ile tahmini değer arasındaki ilişkiyi belirlemek için (Gültepe, 2019), basit bir yapıya sahip olan RF sınıflandırma algoritması sınıflandırma algoritması ise binlerce değişkeni silmeden veya doğruluğu bozulmadan işlemek için (Breiman, 2004), Karar ağacı modeli, her düğümün bir özelliği (niteliği) temsil ettiği, her bağlantının (dal) bir kararı temsil ettiği ve her bir yapının bir sonucu olduğu bir ağacın bilinmesi için (Gültepe, 2019) ve kNN ise bir özelliğin kendisine en yakın olan özellikler arasındaki yakınlığı göre (Kılınç, Borandağ, Yücalar, Tunalı, Şimşek, ve Özçift, 2016) hem sınıflandırma hem de regresyon problemlerini çözebilmek için (Mahesh, 2020) kullanılan algoritmalarından bazılarıdır. Verileri anlamlandırma ve bir kanıtı dayandırma bilimi olan istatistik, boyutu büyük olan veri setlerinde araştırmacıları sınırlandırabilmektedir. Birçok alanda klasik istatistik yöntemleri kullanılsada teknolojinin gelişmesi ile çeşitli ve büyük boyutlardaki verinin analiz edilmesinde yetersiz kalabilmektedir. Bu bağlamda veri madenciliği algoritmaları ön plana çıkmaktadır. İstatistik yöntemlerle analiz edilemeyecek büyüklükteki veri setleri veri madenciliği algoritmaları ile kolaylıkla analiz edilebilmektedir. Veri madenciliği algoritmalarının en önemli avantajı ise herhangi bir ön koşula tabii olmamaları ve ör-

nekleme yöntemini kullanarak çapraz doğrulama yoluyla sınıflandırmayı kontrol edebilmeleridir (Tüzüntürk, 2010; Emre ve Erol, 2017)

Aynı zamanda algoritmaların pratik olması, doğru, başarılı ve güvenilir sonuçlar elde edilmesi ile birçok hayvancılık alanlarında da kullanılmaktadır (Teke, Orhan, Küçüksille, Bilginturan, ve Teke, 2013; Shahinfar, Page, Guenther, Cabrera, Fricke, ve Weigel, 2014; Cihan, Kalıpsız, ve Gökçe, 2020)

Mevcut çalışmada doğumdan 12 aylık yaşa kadar farklı büyüme ve gelişme dönemlerindeki sığırların bazı vücut ölçülerinden canlı ağırlık tahmininin belirlenmesinde yaygın olarak kullanılan MLR, RF, DT ve kNN veri madenciliği algoritmalarının performanslarının değerlendirilmesi amaçlanmıştır.

MATERYAL VE YÖNTEM

Bu çalışmanın hayvan materyalini Selçuk Üniversitesi Ziraat Fakültesi Prof. Dr. Orhan DÜZGÜNEŞ Araştırma ve Uygulama Tesislerinde yetiştirilen doğumdan 12 aylık yaşa kadar toplam 42 baş (dişi=24 ve erkek=18 baş) Siyah Alaca sığır oluşturmaktadır. Çalışmadaki hayvanlar; doğumdan süttan kesime kadar (0-2 aylık) % 18 HP ve 2671 kkal/kg ME buzağı başlangıç yemi, süttan kesimden 6 aylığa kadar (3-6 aylık) % 17.50 HP ve 2653 kkal/kg ME buzağı büyütme yemi ve 7. aydan 12 aylığa kadar % 15.53 HP ve 1553 kkal/kg ME'li kesif yem ile beslenmişlerdir. Kesif yeme ilaveten saman, kuru yonca otu ve mısır silajı kaba yem olarak sabah ve akşam iki öğünde tam rasyon ile *ad-libitum* olarak beslenmiştir.

Yöntem

Hayvanların canlı ağırlığını saptamak için 50 g hassasiyetinde kantar, vücut ölçülerinden cidago yüksekliği (CY), sağrı yüksekliği (SY) vücut uzunluğu (VU), göğüs derinliği (GD) ölçü bastonu ile göğüs çevresi (GÇ) ölçü şeridi ile belirlenmiştir (Hauptner, Germany). Çalışmada vücut ölçüleri Göncü (2021)'nin bildirdiğine göre alınmıştır.

GD: Cidagonun en yüksek noktasından göğüs kemiği arasındaki dikey mesafe, GÇ: Kürekler arkasındaki göğüs çevresi, CY: Cidagonun en yüksek noktasından yere kadar olan dikey mesafe, SY: Sağrı kemiğinin en yüksek noktasından yere kadar olan dikey mesafe ve VU: Omuz ucundan oturarak yumrusuna kadar olan meyilli mesafedir.

İstatistikî analiz

Canlı ağırlık bağımlı değişken, cidago yüksekliği, sağrı yüksekliği, vücut uzunluğu, göğüs derinliği, göğüs çevresi, yaş ve cinsiyet ise bağımsız değişken olarak tanımlanmıştır. Çalışmadaki tüm analizler çapraz doğrulama (cross-validation) 5 alınarak ORANGE programı versiyon 3.31 yardımıyla gerçekleştirilmiştir (Demšar, Curk, Erjavec, Gorup, Hočevar, Milutinovič, Mozina, Polajnar, Toplak, Staric, Stajdohar, Umek, Zagar, Zbontar, Zitnik, ve Zupan, 2013). ORANGE programında 4 farklı veri madenciliği algoritması LR, RF, DT ve kNN modelleri kullanılarak analizler yapılmıştır. Tanıtıcı istatistikler ve Pearson korelasyon katsayıları ise Minitab 16.1.1 istatistik paket programında yapılmıştır (Minitab, 2010).

Canlı ağırlık tahminde kullanılan veri madenciliği algoritmalarının tahmin performanları olarak hata kareler

ortalamasının karekökü (RMSE), ortalama mutlak hata (MAE), ortalama karesel hata (MSE) ve belirtme katsayısı (R^2) kullanılmış olup, aşağıda eşitlik 1, 2, 3 ve 4'te verilmiştir (Zaborski, Ali, Eyduran, Grzesiak, Tariq, Abbas, Waheed, ve Tirink, 2019; Altay, Boztepe, Eyduran, Keskin, Tariq, Bukhari, ve Ali, 2021; Topuz, 2021; Hodson, 2022).

$$RMSE = \sqrt{\frac{1}{2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{2} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (2)$$

$$MSE = \frac{1}{2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2} \quad (4)$$

Yukarıdaki denklemlerde; n: Gözlenen sayısını, Y_i : gerçek canlı ağırlık değerini, \hat{Y}_i : tahmin edilen canlı ağırlık değerini ifade etmektedir.

BULGULAR VE TARTIŞMA

Mevcut çalışmada 0-12 aylık yaşta 42 baş sığıra ait bazı vücut ölçülerinin canlı ağırlık ve yaş özelliğine ait tanıtıcı istatistikleri Çizelge 1'de verilmiştir.

Çizelge 1. Bazı vücut ölçüleri, canlı ağırlık ve yaş özelliğine ait tanıtıcı istatistikler

Özellikler	N	\bar{X}	s_x	Min	Maks
CA (kg)	42	140.8	68.5	29	318
GÇ (cm)	42	113.54	23.10	69	162
CY (cm)	42	87.98	13.18	63	119
VU (cm)	42	90.58	15.89	53	127
GD (cm)	42	40.37	7.93	25	62
SY (cm)	42	93.26	13.60	67	132
Yaş (gün)	42	173.4	108.1	7	363

Vücut ölçülerinden GÇ (cm), CY (cm), VU (cm), GD (cm) ve SY (cm)' nin minimum ve maksimum değerleri sırasıyla; 69-162, 63-119, 53-127, 25-62 ve 67-132 aralığında, yaş (gün) ve CA (kg)'nin minimum ve

maksimum değerleri ise aynı sıra ile 7-363 ve 29-318 aralığında tespit edilmiştir. Bazı vücut ölçüleri, yaş ve canlı ağırlık arasındaki Pearson korelasyon katsayıları Çizelge 2'de verilmiştir.

Çizelge 2. Farklı büyüme ve gelişme dönemlerine ait bazı vücut ölçüleri, yaş ve canlı ağırlık arasındaki korelasyon katsayıları

	CA	GÇ	CY	VU	GD	SY
GÇ	0.972**					
CY	0.902**	0.901**				
VU	0.951**	0.957**	0.919**			
GD	0.972**	0.976**	0.917**	0.959**		
SY	0.887**	0.892**	0.972**	0.923**	0.912**	
Yaş	0.916**	0.944**	0.851**	0.892**	0.931**	0.836**

**; P<0.01

Çizelge 2'ye göre vücut ölçüleri ile CA arasında pozitif bir korelasyon olduğu belirlenmiştir (P<0.01). CA'nın en yüksek korelasyonu GÇ (r=0.972) ve GD (r=0.972) olarak bulunurken, en düşük korelasyon ise SY (r=0.887)'de bulunmuştur.

Bazı vücut ölçüleri, cinsiyet ve yaş özellikleri kullanılarak canlı ağırlık tahmininde veri madenciliği algoritmalarının sonuç karşılaştırılması ise Çizelge 3'de verilmiştir.

Çizelge 3. GÇ, CY, VU, GD, SY, yaş ve cinsiyet özellikleri ile canlı ağırlık tahmininde kullanılan algoritma sonuçlarının karşılaştırılması

Algoritmalar	MSE	RMSE	MAE	R ²
kNN	658.733	25.666	16.990	0.856
DT	653.811	25.570	19.300	0.857
RF	404.080	20.102	14.718	0.912
MLR	277.544	16.660	13.197	0.939

Çizelge 3'de sonuçlarının karşılaştırılması yapılan algoritmalarından kNN, DT, RF ve MLR'ye ait belirleme katsayıları (R²) sırasıyla, 0.856, 0.857, 0.912 ve 0.939 olarak saptanmıştır. RF ve MLR'ye ait R² değerleri; kNN ve DT algoritmalarına nazaran daha yüksek bulunurken, MSE ve MAE değerleri ise bu algoritmalarından daha düşük olarak saptanmıştır. Veri madenciliği algoritmalarından MLR modelin R² değeri (0.939) ile en yüksek iken, en düşük değer ise kNN (0.856) modelinde saptanmıştır. Veri madencili-

ğinde veri setlerinin performansı üzerine faydalı ve önemli özellikleri tespit etmek, veri setlerinden gereksiz değişkenlerin çıkartmak ve algoritmaların yüksek performansı için özellik seçimi yapılmaktadır (Budak, 2018). Bu bağlamda çok fazla değişken bazı algoritmalarda modelin performansını düşürebileceğinden CA ile en yüksek korelasyon gösteren bağımsız değişkenler (GÇ ve GD) kullanılarak veri madenciliği algoritmalarının sonuç karşılaştırmaları yapılmış olup Çizelge 4'de verilmiştir.

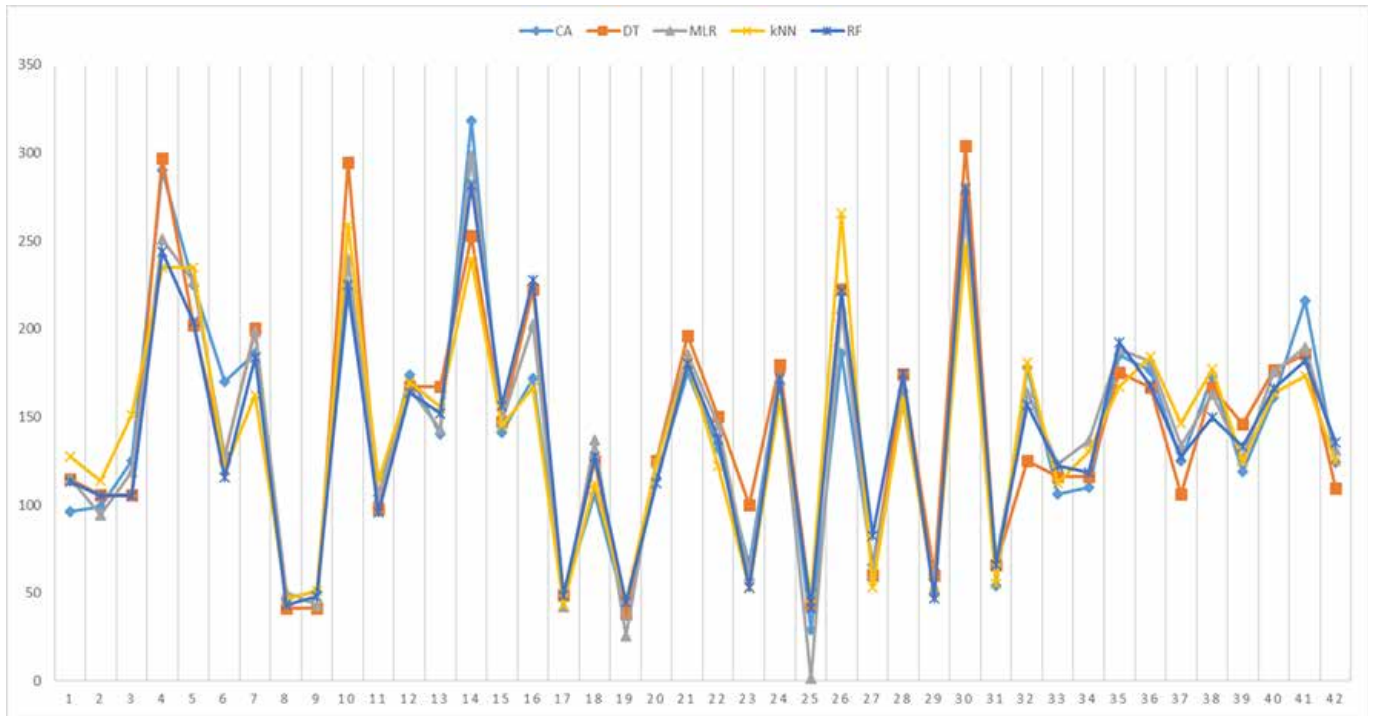
Çizelge 4. GÇ, GD, yaş ve cinsiyet özellikleri kullanılarak canlı ağırlık tahmininde algoritma sonuçlarının karşılaştırılması

Algoritmalar	MSE	RMSE	MAE	R ²
kNN	708.710	26.622	18.362	0.845
DT	565.685	23.784	17.655	0.877
RF	420.330	20.502	14.804	0.908
MLR	263.189	16.223	12.636	0.943

Veri madenciliğinde kullanılan modellerin bazı vücut ölçüleri ile canlı ağırlığın tahmin edilmesinin yanısıra Çizelge 4'de sadece GÇ ve GD vücut ölçülerinin kullanılması ile yine algoritmalarda yüksek R² değerleri elde edilmiştir. GÇ ve GD ait verilerin kullanılması sonucu en iyi uyum performansı 0.943 R² değeri ile MLR modelinde tespit

edilmiştir.

Araştırmada kullanılan gözlemlenen canlı ağırlıklar ile veri madenciliği algoritmalarından tahmin edilen canlı ağırlıklar Şekil 1'de grafiksel olarak verilmiştir. Şekil 1'e göre kullanılan veri madenciliği algoritmaları ile tahmin edilen CA değerleri birbiri ile benzer eğilimde olduğu görülebilir.



Şekil 1. Gözlenen canlı ağırlıklar ile veri madenciliği algoritmalarından tahmin edilen canlı ağırlıkların karşılaştırılması

Koç ve Akman'da (2007) farklı dönemlerdeki 18 baş Siyah Alaca tosunlarında yaptıkları bir çalışmada vücut ölçülerinden canlı ağırlığının tahmin etmek için stepwise-regresyon yöntemi vasıtasıyla 3 farklı regresyon eşitliği

elde edilmiş olup, regresyon eşitliklerinde R² > 084 olarak tespit edilmiştir. Araştırmacılar CA ile en yüksek korelasyonun GÇ (r= 0.921) arasında olduğunu, GÇ'nin tek başına CA tahmin etmek için yeterli olacağını bunlara ilaveten BÇ

ve SY'nin kullanılması ile tıracağına ifade etmişlerdir. tahmindeki doğruluğu art-

Vanvanhossou, Diogo, ve Dossa (2018), Shorthorn sığır ırkında vücut ölçümleri ve kondüsyon skorlarından canlı ağırlığın tahmin edilmesi amaçlanmıştır. Çalışmada canlı ağırlık ve ele alınan bütün vücut ölçüleri arasında pozitif bir korelasyon ($r=0.90-0.97$) bulunurken, vücut ölçüleri istatistik olarak önemli bulunmuştur ($P<0.001$). En yüksek korelasyon canlı ağırlık ve göğüs çevresi arasında, en düşük korelasyon ise canlı ağırlık ile sağrı yüksekliği arasında bulunmuştur. Vücut kondüsyon skoru ile canlı ağırlık arasındaki korelasyon ise oldukça düşük bulunmuştur ($r=0.30$). Sonuç olarak araştırmacılar canlı ağırlığı tahmin etmede fizyolojik ölçümlerin kullanılmasının yeterli olacağını ifade etmişlerdir.

Özkaya ve Bozkurt (2009), Siyah Alaca, Esmer ve melez ırklarından fizyolojik ölçümleri kullanılarak canlı ağırlığın tahminlenmesi amaçlanmıştır. Çalışmada tüm vücut ölçülerinin kullandığı regresyon denkleminin R^2 değerlerini; Siyah Alaca, Esmer ve melez ırklarında sırasıyla %92.2, %95.0 ve %68.2 olarak bulunmuştur. GÇ'si tüm ırklar için istatistiki olarak önemli bulunmuştur ($P<0.05$). Siyah Alaca, Esmer ve melez ırklarında en yüksek korelasyon canlı ağırlık ile göğüs çevresinde aynı sıra ile 0.78, 0.95 ve 0.94 olarak bulmuşlardır. Araştırmacılar Esmer ve melez ırkların vücut ölçüleri kullanılarak canlı ağırlığı tahmin etmedeki doğruluğunun Siyah Alaca ırkına göre daha yüksek olduğunu ve tüm vücut ölçümleri arasında GÇ'sinin canlı ağırlığı tahmin etmede en iyi özellik olduğunu ifade etmişlerdir.

Ulutaş, Saatçi, ve Özlütürk (2001), yaptıkları bir çalışmada Doğu Anadolu Kırmızısı buzağılarında vücut ölçüleri kullanılarak canlı ağırlığın tahmin edilmesi amaçlamışlardır. Bütün vücut ölçülerinin kullanılmasıyla elde edilen $R^2=0.97$, vücut ölçülerinden sadece göğüs çevresi kullanılması sonucunda ise $R^2=0.95$ olarak bulunmuştur. Sonuç olarak sadece GÇ'nin kullanılmasının; bütün vücut ölçümlerinin kullanılmasıyla elde edilen R^2 kadar güvenilir olduğunu ve hayvanlarda canlı ağırlık tahmini için iyi bir tahmin edici olarak rahatlıkla kullanılabileceğini bildirmiştir.

Huma ve Iqbal (2019), 2-36 aylık 131 baş koyunlar-

da canlı ağırlığını tahmin etmek için bazı makine öğrenmesi algoritmaları kullanılmıştır. Çalışmada linear model, regression trees, random forest ve support vector machine algoritmaları kullanılmıştır. Veri seti test ve eğitim seti olarak 2'ye ayrılmıştır. Araştırmacılar çalışmada en iyi sonucu random forest modelinin verdiğini ve random forest modelinin eğitim ve test veri setleri için R^2 lerinin sırasıyla 0.988 ve 0.916 olduğunu ifade etmişlerdir.

54 dişi 29 erkek olmak üzere toplamda 83 keçide yürütülen diğer bir çalışmada regresyon ağacı diyagramından faydalanılarak vücut ölçülerinden canlı ağırlığın tahminlenmesi amaçlanmıştır. Hem erkek hem de dişi keçilerde canlı ağırlık ile bütün vücut ölçüleri arasında pozitif bir korelasyon olduğunu ($P<0.01$), dişilerde en yüksek korelasyonun CA ile CY arasında ($r=0.82$), erkeklerde ise CA ile VU arasında ($r=0.83$) olduğunu belirtmişlerdir. Araştırmacılar sınıflandırma ve regresyon ağacı (CART) algoritmasının yüksek belirleme katsayısı sebebiyle keçilerin canlı ağırlığını tahmin etmede kullanılabileceğini ve canlı ağırlığı tahmin etmede VU'nun kayda değer bir role sahip olduğunu ifade etmişlerdir (Louis-Tyasi, Tshogofatso-Mkhonto, Cyril-Mathapo, ve Madikadike-Molabe, 2021).

Literatürde vücut ölçülerinden canlı ağırlığın tahmininde bazı veri madenciliği modelleri kullanılmış olup, canlı ağırlığının tahmin edilmesinde kullanılan bazı vücut ölçülerini istatistiki olarak ya önemli ($P<0.05$) yada çok önemli ($P<0.01$) bulunmuştur. Mevcut çalışma ile karşılaştırıldığında bulunan sonuçların uyumlu olduğu söylenebilir. Çalışmalarda ele alınan kimi vücut ölçülerinin CA ile yüksek korelasyon ve yüksek regresyon katsayısına sahip olmaları nedeniyle tahminlerde rahatlıkla kullanılabileceği, bir veya birden fazla bağımsız değişken ile tahmini değer arasındaki ilişkiyi, binlerce değişkeni silmeden veya doğruluğu bozulmadan işlenmesi ve hem sınıflandırma hem de regresyon problemlerini çözebilmeleri nedeniyle pratik, basit, güvenilir ve kalite uyum kriterlerinden dolayı kullanılabileceği sonucuna varılabilir.

SONUÇ VE ÖNERİLER

Canlı ağırlık, çiftlik hayvanların yaşamlarının her döneminde bilinmesine ihtiyaç duyulan önemli bir özelliktir. Mevcut çalışmadan elde edilen sonuçlar canlı ağırlık ile bazı

vücut ölçüleri arasında önemli bir ilişki olduğunu göstermektedir. GÇ, CY, VU, GD, SY kullanılarak canlı ağırlığı tahmin etmede MLR veri madenciliği algoritmasının belirleme katsayısı (R^2) % 93.90 gibi yüksek bir orana sahiptir. GÇ, CY, VU, GD, SY kullanılarak canlı ağırlık tahmin edildiği gibi sadece GÇ ve GD'nin kullanılmasının da yeterli olacağı tahmin performansları incelendiğinde anlaşılmaktadır. Ayrıca sığırların canlı ağırlığının veri madenciliği algoritmaları ile tahmininde vücut özelliklerinden güvenle kullanılabilceği sonucuna varılmıştır.

Açıklama

Tüm yazarlar, bu çalışmanın geçerli olduğuna, çalışma hakkındaki görüşlerini temsil ettiğine ve bu çalışmanın yayınlanmasına izin vermişlerdir. Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır. Veri toplama ve hayvancılık uygulamaları 5996 sayılı kanunun 9. maddesinde belirtilen hayvan refahı kurallarına göre gerçekleştirilmiştir.

KAYNAKLAR

- Akkol, S., Akilli, A., ve Cemal, I., 2017, Comparison of artificial neural network and multiple linear regression for prediction of live weight in hair goats. *Yuzuncu Yıl University Journal of Agricultural Sciences*, 27(1), 21-29.
- Altay, Y., Boztepe, S., Eydurun, E., Keskin, İ., Tariq, M. M., Bukhari, F. A., ve Ali, I., 2021, Description of Factors Affecting Wool Fineness in Karacabey Merino Sheep using Chaid and Mars Algorithms. *Pakistan Journal of Zoology*, 53(2), 1-7.
- Altay, Y., 2022, Phenotypic Characterization of Hair and Honamli Goats by Using Classification Trees Algorithms and Multivariate Adaptive Regression Splines (MARS), *Journal of the Faculty of Veterinary Medicine, Kafkas University* 28(3), 401-410.
- Ameen, A. A., ve Mikail, N., 2018, Live body weight prediction in hair goats by application of fuzzy logic. *Applied Ecology and Environmental Research*, 16(6), 7563-7574.
- Aytekin, İ., Eydurun, E., Karadas, K., Aksahan, R., ve Keskin, İ., 2018, Prediction of fattening final live weight from some body measurements and fattening period in young bulls of crossbred and exotic breeds using MARS data mining algorithm. *Pakistan Journal of Zoology*, 50(1), 189-195.
- Breiman, L., 2004. Consistency for a simple model of random forests. Technical Report 670, Berkeley
- Budak, H., 2018, Özellik seçim yöntemleri ve yeni bir yaklaşım. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22 (özel sayı), 21-31.
- Cihan, P., Kalıpsız, O., ve Gökçe, E., 2020, Yenidoğan kuzularda bilgisayar destekli tanı. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 26(2), 385-391.
- Çankaya, S., Altop, A., Kul, E., ve Erener, G., 2009, Faktör analiz skorları kullanılarak karayaka kuzularında canlı ağırlık tahmini. *Anadolu Tarım Bilimleri Dergisi*, 24(2), 98-102.
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., ve Zupan, B., 2013, Orange: data mining toolbox in Python. *The Journal of Machine Learning Research*, 14(1), 2349-2353.
- Emre, İ. E., ve Erol, Ç. S., 2017, Veri Analizinde İstatistik mi Veri Madenciliği mi?. *Bilişim Teknolojileri Dergisi*, 10(2), 161-167.
- Göncü, S., 2021, *Sığırcılık*, Akademisyen Kitapevi, Ankara.
- Gül, S., Keskin, M., Biçer, O., Gündüz, Z., ve Behrem, S., 2020, Effects of different lambing season on some reproductive characteristics of ewes and growth performance of lambs in Awassi sheep. *Lalahan Hayvancılık Araştırma Enstitüsü Dergisi*, 60(1), 32-36.
- Gültepe, Y., 2019, Makine öğrenmesi algoritmaları ile hava kirliliği tahmini üzerine karşılaştırmalı bir değerlendirme. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 8-15.
- Hodson, T. O., 2022, Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development Discussions*, 1-10.
- Huma, Z. E., ve Iqbal, F., 2019, Predicting the body weight of Balochi sheep using a machine learning

- approach. Turkish Journal of Veterinary and Animal Sciences, 43(4), 500-506.
- Karabacak, A., Altay, Y., ve Aytekin, İ. 2019, Relationship between some body measurements and ultrasound measurements at the end of fattening of Akkaraman and İvesi lambs. Bahri Dağdaş Hayvancılık Araştırma Dergisi, 8(2), 56-64.
- Keskin, İ., Dağ, B., ve Şahin, Ö., 2005, Anadolu Merinosu erkek kuzularında besi başı vücut ölçüleri ile sıcak karkas ağırlığı arasındaki ilişkilerin path analizi ile araştırılması, Hayvancılık Araştırma Dergisi, 15(2), 6-10.
- Kılınç, D., Borandağ, E., Yücalar, F., Tunalı, V., Şimşek, M., ve Özçift, A., 2016, kNN algoritması ve R dili ile metin madenciliği kullanılarak bilimsel makale tasnifi. Marmara Fen Bilimleri Dergisi, 28(3), 89-94.
- Koç, A., ve Akman, N., 2007, Siyah-Alaca tosunların değişik dönemlerdeki vücut ölçüleri ve vücut ölçülerinden canlı ağırlığın tahmini. Adnan Menderes Üniversitesi Ziraat Fakültesi Dergisi, 4(1/2), 21-25.
- Louis-Tyasi, T., Tshogofatso-Mkhonto, A., Cyril-Mathapo, M., ve Madikadike-Molabe, K., 2021, Regression tree analysis to predict body weight of South African non-descript goats raised at Syferkuil farm, Capricorn district of South Africa. Biotechnology in Animal Husbandry, 37(4), 293-304.
- Mahesh, B., 2020, Machine learning algorithms-a review. International Journal of Science and Research, 9, 381-386.
- Minitab, 2010, Minitab 16.1.1 for Windows. State College, PA, USA: Minitab Inc.
- Özkaya, S., ve Bozkurt, Y., 2009, The accuracy of prediction of body weight from body measurements in beef cattle. Archives Animal Breeding, 52(4), 371-377.
- Peng, H., 2015, Air quality prediction by machine learning methods. Doctoral dissertation, University of British Columbia. 1-86.
- Pesmen, G., ve Yardimci, M., 2008, Estimating the live weight using some body measurements in Saanen goats. Archiva Zootechnica, 11(4), 30-40.
- Riva, J., Rizzi, R., Marelli, S., ve Cavalchini, L. G., 2004, Body measurements in Bergamasca sheep. Small Ruminant Research, 55(1-3), 221-227.
- Shahinfar, S., Page, D., Guenther, J., Cabrera, V., Fricke, P., ve Weigel, K., 2014, Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. Journal of Dairy Science, 97(2), 731-742.
- Tariq, M., Younas, M., Khan, A. B., ve Schlecht, E., 2013, Body Measurements and Body Condition Scoring as Basis for Estimation of Live Weight in Nili-Ravi Buffaloes. Pakistan Veterinary Journal, 33(3), 325-329.
- Taşdemir, Ş., Ürkmez, A., ve İnal, Ş., 2011, A fuzzy rule-based system for predicting the live weight of Holstein cows whose body dimensions were determined by image analysis. Turkish Journal of Electrical Engineering & Computer Sciences, 19(4), 689-703.
- Teke, E.Ç., Orhan, H., Küçüksille, E.U., Bilginturan, S., ve Teke, H., 2013, Veri madenciliği süreci ile siyah alaca sığırlarda canlı ağırlık tahmini. 8. Ulusal Zootekni Bilim Kongresi, 5-7 Eylül, Çanakkale, 402-403.
- Tırınk, C., 2022, Comparison of Bayesian Regularized Neural Network, Random Forest Regression, Support Vector Regression and Multivariate Adaptive Regression Splines Algorithms to Predict Body Weight from Biometrical Measurements in thalli sheep. Journal of the Faculty of Veterinary Medicine, Kafkas University, 28(3), 411-419.
- Topuz, S., 2021, Eğitsel Verilerde Weka ve Orange Veri Madenciliği Yazılımlarından Elde Edilen Analiz Sonuçlarının Karşılaştırılması, Yüksek Lisans Tezi, Eğitim Bilimleri Enstitüsü, Ankara, 1-68.
- Tüzemen, N., Yanar, M., Akbulut, Ö., Uğur, F ve Aydın, R., 1995, Prediction Of Body Weights From Body Measurements İn Holstein-Friesian Calves. Atatürk Üniversitesi Ziraat Fakültesi Dergisi, 26(2), 245-252.
- Tüzüntürk, S., 2010, Veri madenciliği ve istatistik. Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 29(1), 65-90.
- Ulutaş, Z., Saatçi, M., ve Özlütürk, A., 2001, Prediction of body weights from body measurements in East Anatolian Red calves. Atatürk Üniversitesi Ziraat

- Fakültesi Dergisi, 32(1), 61-65.
- Vanvanhossou, S. F. U., Diogo, R. V. C., ve Dossa, L. H., 2018, Estimation of live body weight from linear body measurements and body condition score in the West African Savannah Shorthorn Cattle in North-West Benin. *Cogent Food and Agriculture*, 4(1), 1549767.
- Zaborski, D., Ali, M., Eyduran, E., Grzesiak, W., Tariq, M. M., Abbas, F., Waheed, A., ve Tirink, C., 2019, Prediction of selected reproductive traits of indigenous Harnai sheep under the farm management system via various data mining algorithms. *Pakistan Journal of Zoology*, 51(2), 421-431.
- Zülkadir, U., Şahin, Ö., Aytekin, İ., ve Boztepe, S., 2008, Malya kuzularda canlı ağırlık ve bazı vücut ölçülerinin tekrarlanma dereceleri. *Selçuk Üniversitesi Ziraat Fakültesi Dergisi*, 22 (45), 89-93.