UNIVERSIDADE FEDERAL DE LAVRAS

LEONARDO AUGUSTO COELHO RIBEIRO

APPROACHES OF MACHINE LEARNING AND VALIDATION STRATEGIES TO PREDICT GRAZING BEHAVIOR IN BEEF CATTLE USING SENSORS

LAVRAS – MG

2020

LEONARDO AUGUSTO COELHO RIBEIRO

APPROACHES OF MACHINE LEARNING AND VALIDATION STRATEGIES TO PREDICT GRAZING BEHAVIOR IN BEEF CATTLE USING SENSORS

Thesis project submitted to the Federal University of Lavras as part of requirements of Graduate Program in Animal Science, area of the concentration in Ruminants Production and Nutrition, to obtain Master degree.

Prof (a). Marina de Arruda Camargo Danés
Advisor
Prof. Daniel Rume Casagrande
Co-advisor

LAVRAS – MG
2020

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Ribeiro, Leonardo Augusto Coelho.
Approaches of machine learning and validation strategies to predict grazing behavior in beef cattle using sensors / Leonardo Augusto Coelho Ribeiro. - 2020.
65 p. : il.

Orientador(a): Marina de Arruda Camargo Danés.
Coorientador(a): Daniel Rume Casagrande.
Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2020.
Bibliografia.

1. Use of sensor to predict grazing. 2. Use of machine learning to predict animal behavior. 3. Cross-validation strategies. I. Danés, Marina de Arruda Camargo. II. Casagrande, Daniel Rume. III. Título.

LEONARDO AUGUSTO COLEHO RIBEIRO

APPROACHES OF MACHINE LEARNING AND VALIDATION STRATEGIES TO PREDICT GRAZING BEHAVIOR IN BEEF CATTLE USING SENSORS

Thesis presented to the Federal University of Lavras, as part of the requirements of the Graduate Program in Animal Science, area of concentration in Ruminant Production and Nutrition, to obtain the Master degree.

Approved: March 11th, 2020

Marina de Arruda Camargo Danés, PhD. UFLA
Thiago Fernandes Bernardes, D. Sc. UFLA
Luiz Gustavo Ribeiro Pereira, D. Sc. EMBRAPA
Marcio André Stefanelli Lara, D. Sc. UFLA

Prof. Marina de Arruda Camargo Danés, PhD.
Advisor
Prof. Daniel Rume Casagrande, D. Sc.
Co-advisor

LAVRAS – MG

2020

Às pessoas que me ajudaram a construir essa obra, em especial a minha mãe Suely e ao meu pai Joaquim, o alicerce de toda essa jornada.

Dedico

ACKNOWLEDGEMENTS

Nesta página de documento que me é reservada, gostaria de agradecer as pessoas que de certa maneira, dentro das suas possibilidades, contribuíram para a construção deste degrau de minha vida.

Iniciando os agradecimentos pelos meus pais e minha família, núcleo de fraternidade, diversidade. Pela credibilidade, carinho e amor depositado em minha pessoa. Aos familiares *in memoriam* que muito me ajudaram e me viram progredir no meio acadêmico, mas que por força maior não estão presentes para ver essa obra.

Aos amigos de Lavras, aos amigos do Setor de Forragicultura e do NEFOR que me acolherem e inseriram-me, este ser aberrante, nos estudos em forragicultura. Aos professores do departamento de Zootecnia desta universidade, agradeço pela paciência e compreensão na construção de meu conhecimento. Em especial a Profª. Marina Danés, com o coração maior que o peito me escolheu como seu mestrando e me proporcionou uma das maiores experiências de vida nesses últimos dois anos. Sem esquecer, é claro, de todos os Minions que me ajudaram nos experimentos, nas discussões dos *papers* e a grande experiência que me passaram na área de gado de leite. Tenham certeza aprendi muito sobre gado de leite com vocês!

Sem mencionar nomes, gostaria de ressaltar meus agradecimentos aquelas pessoas mais humildes que por advém da vida não tiveram a oportunidade de continuar seus estudos, mas que com os estudos adquiridos na "escola da vida" souberam me orientar em momentos de complicadas decisões, em momentos em que estava diante de mim uma prova de superação.

I did not forget to acknowledge my friends who lived together in Madison, WI – United States. For the University of Wisconsin and the Animal Sciences Department to accept me in these 6 months, I will miss the traditional coffee time on 10 o'clock in 4th floor. The friends in 9th floor, thanks for the patience in understand me and my "Medieval English". The Dorea's Lab guys for teaching me about my master's project and collaborated with my professional development and career. In special I would like to appreciate Professor João Dórea who introduced me to a new country and helped me in the adaptation process. You are extremely important in my life; without you I guess I would not have learned about US culture, rules, costumes and condition. Thank you so much!

This research was made possible by the financial contributions part by the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) – Finance Code 001 and for Research Foundation of Minas Gerais State (FAPEMIG).

*“I do not know anyone who has got to the top without hard work.
This is the recipe. It will not always get you to the top, but should get you pretty near”*

Margaret Thatcher

---

*“Eu não conheço ninguém que tenha chegado ao topo sem muito trabalho.
Essa é a receita. Nem sempre você irá chegar ao topo, mas vai chegar bem perto.”*

Margaret Thatcher

RESUMO

O uso de técnicas de *machine learning* tem sido importante para enfrentar os desafios atuais na pecuária de precisão, pois apresenta novas ferramentas para análise com dados preditivos em larga escala em diversos campos, incluindo a de tecnologia de sensores. Nesse contexto, o objetivo do estudo foi avaliar abordagens de *machine leaning* e estratégias de validação para predição de tempo de pastejo com base em dados gerados por sensores do tipo acelerômetro e giroscópio em bovinos de corte. As abordagens de *machine learning* avaliadas foram *generalizer linear regression* (GLR), *random forest* (RF) e *artificial neural network* (ANN), e as estratégias de validação foram: remover 20% dos dados aleatoriamente para validação (*holdout*), remover todos os dados de um animal por vez para validação (LOAO) e remover os dados de cada um dos últimos 5 dias da avaliação comportamental para validação (LODO). Seis bovinos Nelore de 345 ± 21 kg peso corporal, foram mantidos em pastagem de *B. brizantha* cv. Marandu com sensores acelerômetro e giroscópio acoplados. O comportamento dos animais foi registrado visualmente em um período de 10 horas durante 15 dias. Os dados obtidos pelos giroscópios não foram utilizados, devido a intervalos muito longos de registro dos sensores resultando em um banco de dados incompleto. Os valores de acurácia dos modelos GLR, RF e ANN foram, respectivamente: 57,1%, 76,9% e 74,2% para validação holdout, 53,1% 58,7% e 72% para validação LOAO, e 47,4%, 58,5% e 59,7% para a validação LODO. O modelo de predição linear GLR não foi adequado para predição do comportamento animal a partir de dados de sensores. As ferramentas de machine learning RF e ANN são mais adequadas para processarem dados complexos como esses. Claramente, a estratégia de validação interfere na acurácia do modelo preditivo e isso deve ser levado em consideração na interpretação de dados da literatura. Os baixos valores de acurácia na validação LODO mostram que modelos preditivos não funcionam adequadamente em condições de pastejo diferentes das utilizadas no desenvolvimento do modelo. O modelo validado por LOAO e desenvolvido com ANN atingiu valor de acurácia promissor, o que sugere que, com a correta ferramenta de machine learning, é possível predizer comportamento de pastejo de novos animais, que não foram utilizados no desenvolvimento do modelo. A validação holdout, utilizada na maioria dos estudos com sensores, apresenta valores inflados em decorrência de condições de ambiente (e.g. animal ou condições de pastejo) que influenciam da mesma forma os dados do conjunto de treinamento e de validação do modelo.

Palavras chave: acelerômetro, comportamento animal, crossvalidação, pecuária de precisão, pastejo

ABSTRACT

Machine learning approaches have been crucial for addressing current challenges in precision livestock, as it presents new tools for developing large scale predictive analytics in many fields including the area of sensor technology. In this context, the objectives of our study were to evaluate the following strategies of cross-validation used to predict grazing and not-grazing activities in grazing cattle. The machine learning approaches were generalizer linear regression (GLR), random forest (RF) and artificial neural network (ANN) as well as the cross-validation strategies evaluated were: 20% of the dataset randomly exclude to build the validation dataset (holdout), leave-one-animal-out (LOAO), and leave-one-day-out (LODO). Six Nellore bulls, 345 ± 21 kg body weight, were kept on pasture of Marandu Palisadegrass and had accelerometer and gyroscope sensor attached on neck. Animal behavior was registered through visual observation within a period of 10 hours for 15 days. The gyroscope record data were not used because a larger gap in a datapoint was observed. The overall accuracy of GLR, RF, and ANN were respectively 57.1%, 76.9%, and 74.2% in holdout validation, 53.1%, 58.7% and 72% in LOAO and 47.4%, 58.8% and 59.7% in LODO. GLR was not adequate model to predict animal behavior using our dataset. RF and ANN are more efficient to process complex dataset as these. Clearly, the validation strategy inferring in accuracy results and this is an important point in data analysis. Low values validation accuracy results in LODO shown us that predictive models are not adequate to use in different conditions of pasture. LOAO with ANN was the best validation strategy and it could predict animal behavior of different animals without used in predict model. Holdout validation, widely used in several similar studies, present an inflate accuracy values due to environmental conditions (e.g. animal or grazing conditions) that influence in dataset using in training and the validation dataset of the model.

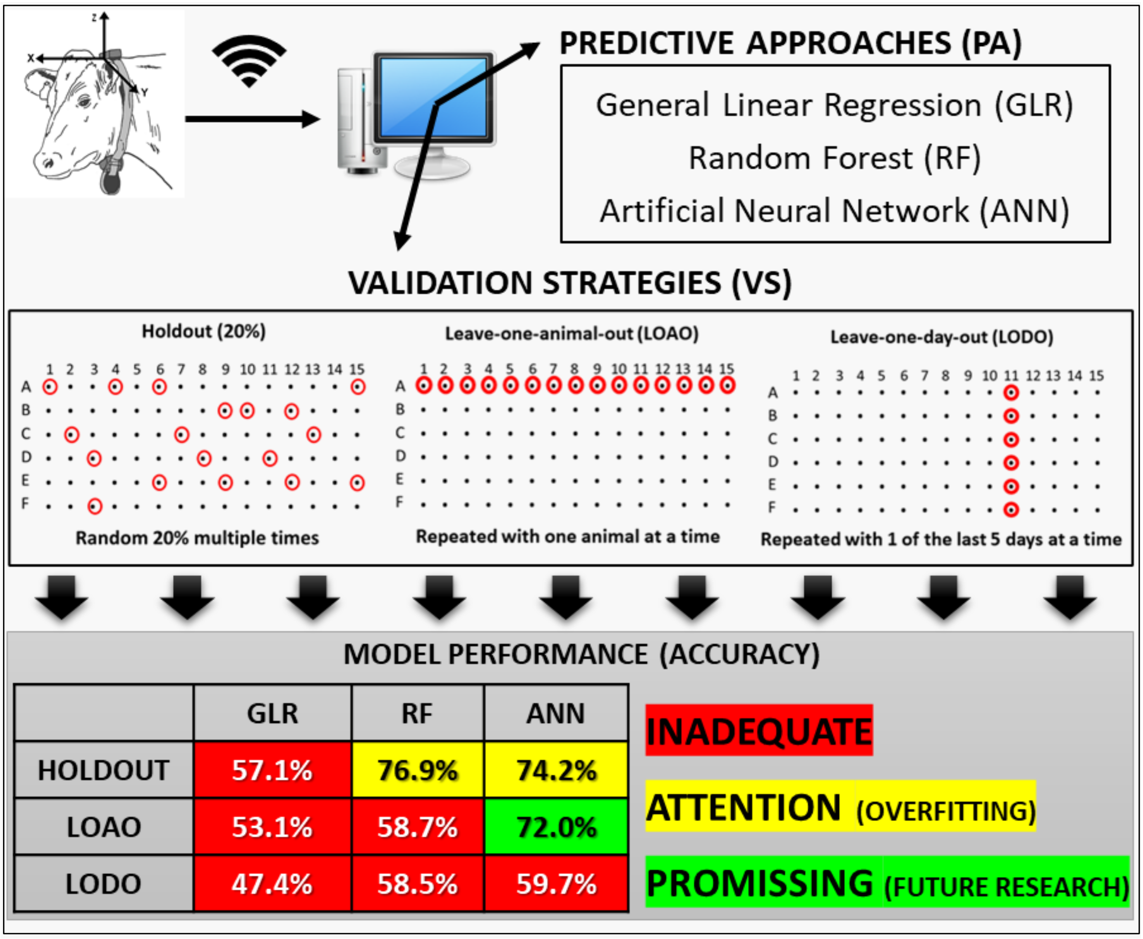
Keywords: accelerometer, animal behavior, cross-validation, grazing, precision-livestock

INTERPRETIVE SUMMARY

Elaborated by Leonardo and advised by Marina

A adoção de tecnologias de automação e processamento de dados na agricultura e na pecuária possibilita a tomada de decisão ágil, com base em informações precisas e atualizadas em tempo real. No entanto, entre a geração de dados e as tomadas de decisão existe um caminho complexo que depende do conhecimento biológico e matemático para saber quais perguntas fazer para o banco de dados e com que ferramentas processá-lo. Neste estudo, testamos três técnicas de modelagem preditiva e três estratégias de validação de modelos para predizer tempo de pastejo de bovinos a partir de dados de acelerômetro. A técnica de machine-learning random-forest é a mais utilizada para processar dados de sensores em comportamento animal e apresentou desempenho adequado em nosso estudo. No entanto, nossa comparação de estratégias de validação mostrou que a estratégia mais utilizada nos trabalhos publicados (holdout) apresenta valores de acurácia inflados por interdependência biológica dos dados nos conjuntos de treinamento e validação. Quando os modelos são aplicados em novos animais ou novas condições de pastejo, a acurácia diminui. A técnica de machine-learning redes neurais artificiais foi capaz de produzir um modelo com acurácia promissora para utilização mesmo em animais novos, que não estavam no treinamento do modelo. No entanto, nenhum modelo foi capaz de predizer tempo de pastejo com confiabilidade quando usado em nova condição de pastejo, enfatizando a importância de incluir uma grande diversidade de condições de pastejo no treinamento do modelo.

Resumo gráfico



Master's dissertation in Animal Science at Federal University of Lavas

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACC	Accelerometer
ANN	Artificial Neural Network
AUC	Area Under the Curve
BT	Bagged Trees
BW	Body Weight
CH	Canopy Height
CV	Cross-validation
DT	Decision Tree
FM	Forage Mass
GLM	Generalized Linear Models
GLR	Generalized Linear Regression
GPS	Global Position System Hectare
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LOAO	Leave-one-animal-out
LODO	Leave-one-day-out
LOOCV	Leave-one-out cross-validation
LR	Logistic Regression
ML	Machine Learning
RF	Random Forest
SVM	Support Vector Machine

TABLE OF CONTENTS

1 INTRODUCTION

The rapid advance in technology development experienced in agricultural systems has created a strong potential for growth of automation and adoption of digital technologies (DREWRY et al., 2019). Among the variety of digital technologies, the use of wearable sensors has been adopted as an alternative for real-time monitoring of animal performance, health and welfare. As such, those monitoring systems usually uses animal behavioral traits, characterized by activities as lying, eating and drinking as unique indicators of health issues, housing conditions, and productivity performance (PENG et al., 2019). To accomplish that, the raw information captured by the sensor technology is processed by predictive models in order to generate the phenotype or alert of interest (WANG, 2019). In this context, the use of machine learning algorithms has been widely employed as way to achieve greater predictive ability.

The emerging field of machine learning is core to pattern recognition and information extraction. Machine learning has been crucial for addressing current challenges in precision agriculture, as it presents new tools for developing large scale predictive analytics in many fields including the area of sensor technology (MOROTA et al., 2018). However, although machine learning techniques can improve prediction quality when applied to sensor data (VALLETTA et al., 2017), very little attention is paid to confounding effects that may inflate prediction quality, more specifically when cross-validation strategies are used.

The k-fold cross-validation is the most commonly used strategy to validate models in animal behavior (HAMILTON et al., 2019 and RIABOFF et al., 2019). The form of using cross-validation are variable, can be use with or without replication data, as leave-one-out and leave-pair-out cross-validation (SMITH et al., 2014). Regardless the type of cross validation adopted in predictive modelling problems, the independence of train and validation dataset are assumed but not always is ensured. Several authors have suggested that biological knowledge needs to be considered when choosing the strategy of cross-validation as way to create the most independent train and validation datasets (LAHART et al., 2019; DÓREA et al., 2018; SHETTY et al., 2017). Those authors demonstrated that when cows from the same group were randomly assigned in training and validations sets, prediction accuracy was inflated when compared with validation set built from an exclusion of all cows of a specific group (herd, farm, or trial, for example). The reason for such results was attributed the carryover effects of diet, management practices, season, remaining in both training and validation datasets.

In grazing systems, which sensor can be used to predict feeding behavior validation may not only be needed under different animal-level condition, but also under new grazing management conditions in order to generalize the predictive ability of trained machine learning algorithms.

In this context, the objectives of our study was to evaluate the following strategies of cross-validation used to predict grazing and not-grazing activities in grazing cattle. The cross-validation strategies evaluated were: 20% of the dataset randomly exclude to build the validation dataset (holdout), leave-one-animal-out, and leave-one-day-out. Additionally, our study aimed to evaluate the predictive performance of machine learning methods within and across cross-validation strategies.

2 LITERATURE REVIEW

Data analytics, big data and precision technologies have potential to make a revolution in pasture based ruminant production system. The likely benefits include increased efficiency, improved product quality and animal health, reduced cost, reduced environmental impacts (STEENEVELD et al., 2015).

The profitability in this system depends on the relationship between plant and animal. Variations in intensity and frequency of grazing affect the development of animals and plants. The sward structure effect the grazing process by ruminants. Time spent searching, energy spent walking and browsing forage are some variables affected by sward structure (e.g. low pasture offer, leaf stem proportion; CASAGRANDE et al., 2011).

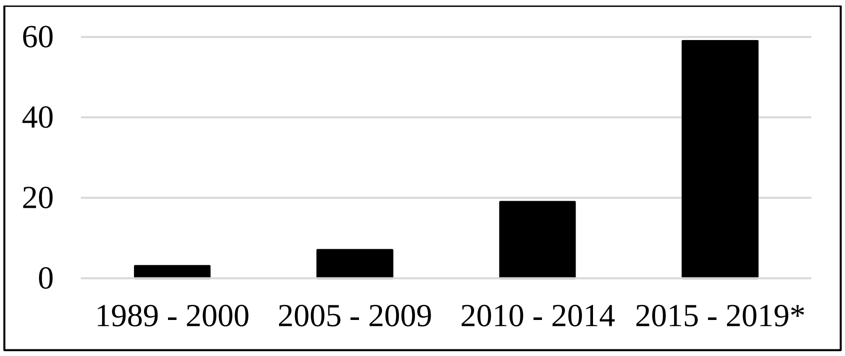
Enormous efforts have been done to develop sensor-based tools to register animal behavior. Several studies, using different types of sensors, were conducted with the most diverse objectives, such as detecting lameness (BARWICK et al., 2018a), estimating energy expenditure on grazing (BEKER et al., 2010; BROSH et al., 2006), and detecting and monitoring the spatial distribution of urine patches (BETTERIDGE et al., 2010).

A review of the literature was performed using the keywords “sensor” and “behavior”, in conjunction with “grazing”, “animal”, “accelerometer”, “machine learning", "prediction", "monitoring", "analyzer" and "observer". The criteria used for the article selection were: (i) written in English, (ii) type of sensor used mentioned (e.g. GPS, accelerometer), (iii) sensor attached to the animal (wearable sensor), (iv) the aim was monitoring animal behavior. Table A.1 in the appendix presents details about title, author, year of publication, location (country where the experiment occurred), major phenotype registered (e.g. location and behavior), type of sensor (e.g. accelerometer or GPS), sensor location in the animal (e.g. neck, ear, jaw, leg), animal species (e.g. dairy cows, beef cattle, goats, sheep), number of animals evaluated, and analytical procedure used (e.g. Decision Tree, Support Vector Machine, K-Nearest Neighbors, Artificial Neural Network).

The data basis consulted were Web of Science, Elsevier and the journals *Journal of Dairy Science*, *Journal of Animal Science and Computer* and *Electronic in Agriculture*. All database consulted returned 4,026 documents, but only 94 articles met the selection criteria and were included in our analysis.

Even though the first study was published in 1989, the topic really received heavy attention in the last decade (Figure 1), following improvements in the technology and mathematical tools. Global Position System (GPS) and accelerometer (ACC) are the most used devices to measure animal behavior based in three main parameters: the location, the body posture and the movements of animal. The earliest article used a medical device attached to the animal to measure jaw movement, head position, and walking posture (MATSUI, 1989). The dissemination of GPS technology made possible to track the location of animals in pasture (HUIRCÁN et al., 2010) which, associated with a reduction in manufacture costs of devices (ANDRIAMANDROSO et al., 2016), contributed to the increased number of studies with animals.

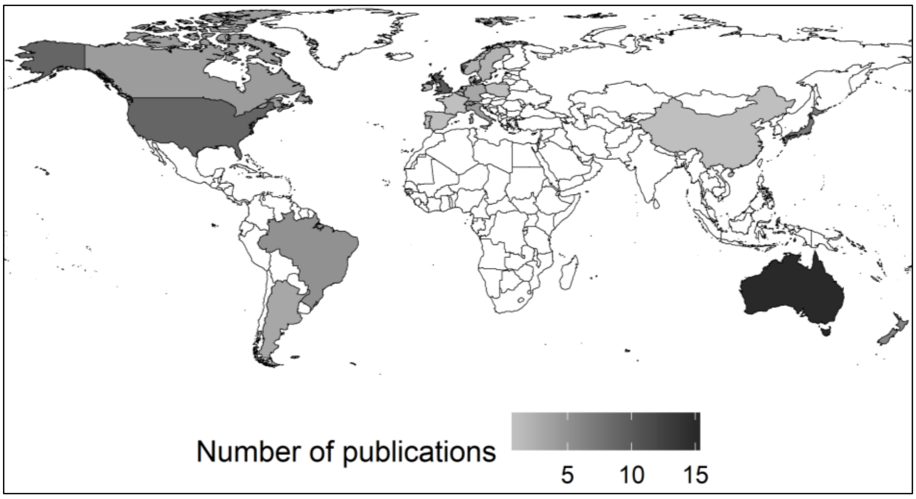
Figure 1 – Evolution of articles published per year using sensors in literature review.



Legend: * publications until November 2019

Initially, the studies were concentrated in America, Japan and Europe, with a later expansion to different regions in last five years. Nineteen publications are accumulated between 2010 to 2014, this number increase to 59 in November 2019. Figure 2 illustrates the location of experiments occurs using sensors (e.g. accelerometer, GPS) from 1989 to 2019. Australia has the greatest number of publications, with studies predominantly in pasture system. On the other hand, in the United States, most of the experiments used dairy cows in tie-stall systems. Finally, the United Kingdom, another country with many publications, mainly used sensors on sheep.

Figure 2 – Distribution of experiments across countries using sensor for monitoring grazing behavior.



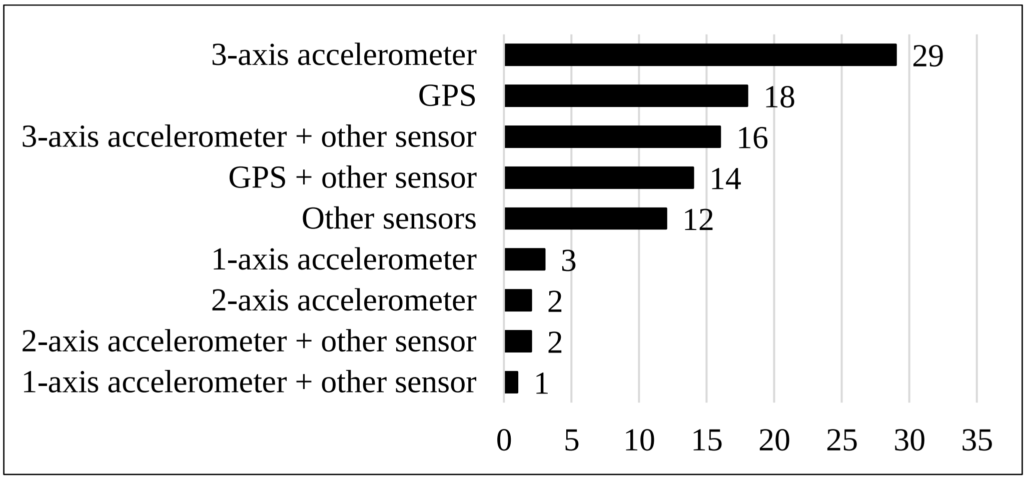
Legend: Argentina (3); Australia (15); Belgium (2); Brazil (5); Canada (4); China (1); Denmark (2); France (1); Germany (3); Ireland (2); Israel (5); Italy (4); Japan (7); Netherlands (2); New Zealand (6); Norway (2); Poland (1); Portugal (1); Scotland (1); Spain (1); Sweden (1); Switzerland (2); United Kingdom (10); United States (9).

Most of the articles (66%) were focused in pasture system production, in which animals in paddocks were observed to predict grazing. Articles classified as other production systems (44%) included animals in free-stalls and tie-stalls, for example. Dairy cows were the most common animal utilized in the publications (45%), followed by beef cattle (26%), sheep (15%), goats (5%) and calves (2%). In 7% of the articles the authors used more than one species.

The most used device was the ACC. However, there is more than one type of ACC. Some devices can record data in three axes (X, Y and Z), others record data in two axes or just in one axis. Moreover, some studies used ACC simultaneously with other types of sensors. The category “other sensors” includes studies that did not use neither ACC or

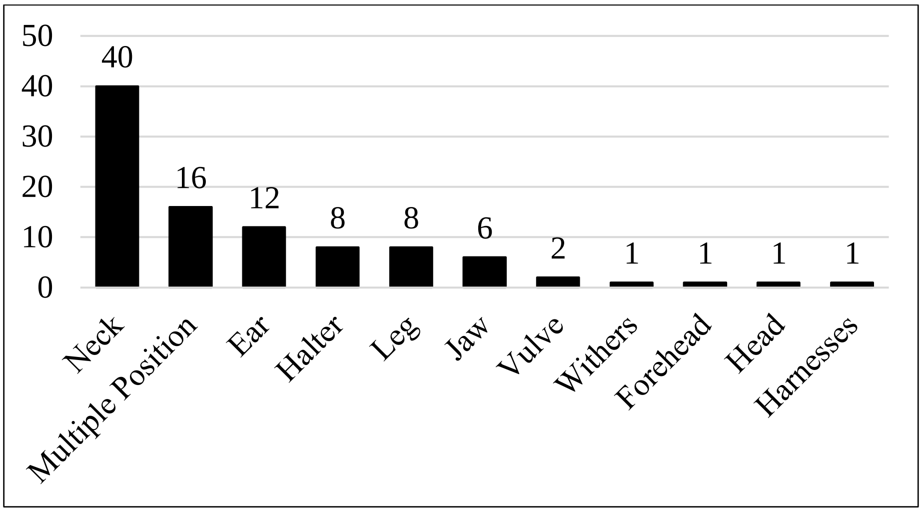
GPS, but rather sensors such as data loggers, magnetometers, gyroscopes and pedometers. In major of articles return in review the sensors used were primarily designed for the respective studies. Even thought some commercial sensors such as the RumiWatch® ACC (Itin + Hoch GmbH, Liestal, Switzerland) are already used for behavior classification in farms situations, some studies that used them did not described the sensors.

Figure 3 – Number of articles by type of sensor utilized.



The position in which the sensor is attached to the animal can affect the accuracy of models developed using machine learning (BARWICK et al., 2018b). The results of our review indicate that the region where sensors are most attached is the neck, followed by the ear of the animals (Figure 4).

Figure 4 – Number of articles by the location of the sensor in the animal



2.1 Sensors used

An accelerometer sensor (ACC) operates basically in signal output transformed from physical acceleration, motion or gravity (ANDRIAMANDROSO et al., 2016). They measure acceleration forces in G-forces. Having common using in cows, sheep and beef cattle to measure features behaviors, ruminations, grazing or feeding and other activities, for example (SHALLOO et al., 2018). Several other activities reported in the literature are walking (BARWICK et al., 2018b), drinking, grooming (SMITH et al., 2016) and resting (DUNNE et al., 2017).

It is the most common sensor. Although, this type of sensor differenced in how many axes worked. Exist ACC sensors that record on a single-axis (1-axis) to detected magnitude and direction of acceleration (YASHITOSHI et al., 2013). Others research used a 2-axis accelerometer (SPEDENER et al., 2019), but are more common used 3-axis accelerometer in researches (SAKAI et al., 2019). It appears also use ACC data from GPS device (GUO et al., 2018).

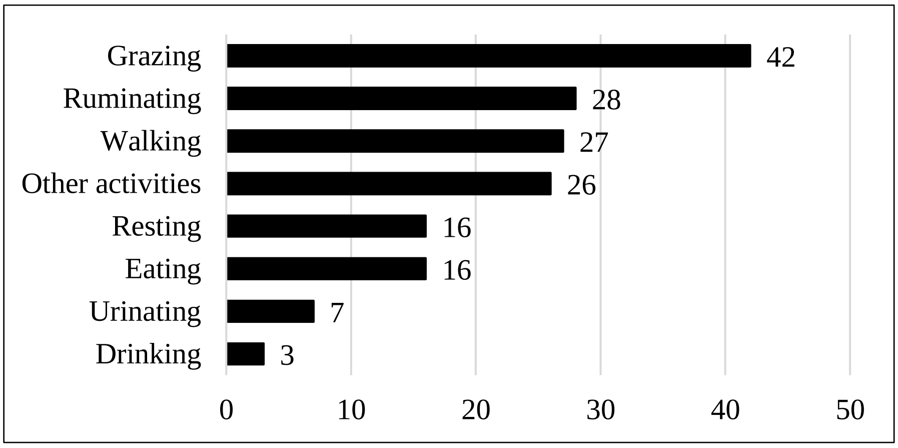
A gyroscope sensor can be used to detect the tilt. Usually, tilt is measured in two axes of reference plane (pitch and roll). Differently of ACC, tilt doesn't measure motion, but ACC can be used as a tilt sensor, but the inverse is possible. The initials tilt sensor used for animal behavior was a mercury tilt sensor (CHAMPION; RUTTER; PENNING, 1997). Currently, tilt sensor are integrated with GPS device (UMSTATTER, WATERHOUSE, HOLLAND, 2008). We used one independent tilt sensor in all steps that recorded a sample at 800 Hz rate (800 samples/second).

In all experimental steps, we used a 3-axis (X, Y, Z) accelerometer sensor recorded a sample at 800 Hz rate (800 samples/second). This sensor was charged with a coin battery cell-attached in hardware.

2.2 Use of sensors to predict grazing

The main objective of use sensor in animals is to record behavior data. Location was a feature used in initial studies when the GPS device was developed. Many behavior variables are common among the studies (Figure 5); grazing being the most used one. In figure 5, other activities are the ones that did not fit into any of the other classifications such as social licking or licking salt.

Figure 5 – Number of articles by behavior feature

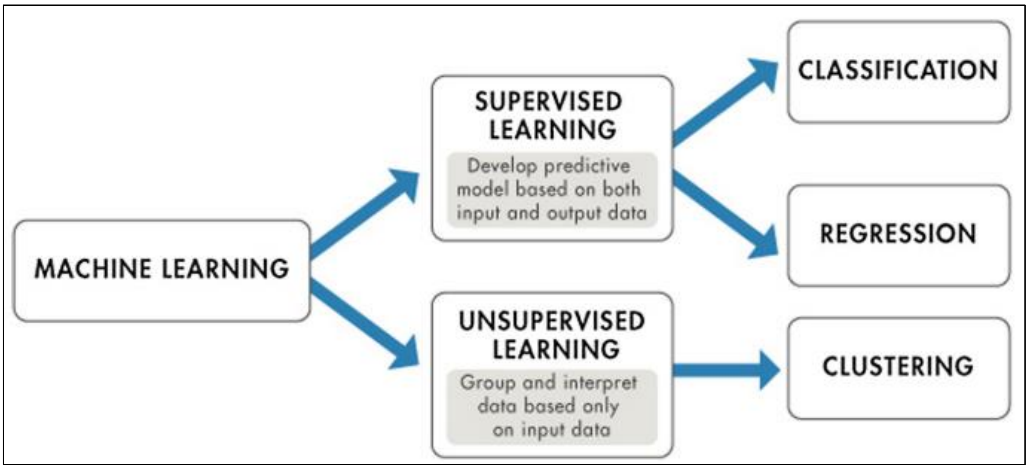


2.3 Use of machine learning to predict animal behavior

Machine learning (ML) is a technique for regression and/or classification of nonlinear systems. The main focus is the prediction without prior knowledge of the underlying data, different from statistical models that use observed data (LARY et al., 2016). As mentioned before, the use of sensor to record animal behavior data increased in the last five years. The use of ML techniques increased as well, and the commonly used algorithms to predict animal behavior are Artificial Neural Networks (ANN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Linear Discriminant Analysis (LDA), K-Nearest Neighbors (KNN) and Generalized Linear Models (GLM). A total of 32% of articles found on this review used ML techniques and RF is the most used algorithm to classifier animal behavior (LUSH et al., 2018). However, other 68% of articles used statistic correlation or other statistic model, or the sensor default prediction without a preview data treatment and in some cases generated by the sensor software used without described if the software were used machine leaning techniques

The algorithms of machine learning can be grouped in two forms: unsupervised learning and supervised learning (Figure 6). Valletta et al., 2017 describe unsupervised learning methods as able uncover structure un unlabeled data. Through reduced the dimensionality (Dimensionality Reduction), identify groups of observation sharing similar attributes (Clustering), and determining the distribution of the data (Density estimation). Already, supervised learning can identify the relationship between an outcome and a set of explanatory variables from a dataset as a starting point to training models.

Figure 6 – Diagram of machine learning process, supervised learning against unsupervised learning.



Source: BUNKER; THABTAH, 2019

In this currently study was used supervised machine learning techniques to classifier raw data. In general, classification techniques seek to categorize samples into groups based on the predictor characteristics. Some techniques take a mathematical path (e.g. Linear Discriminant Analysis), and others take an algorithmic path (e.g. K-Nearest Neighbors) (KUHM, 2013).

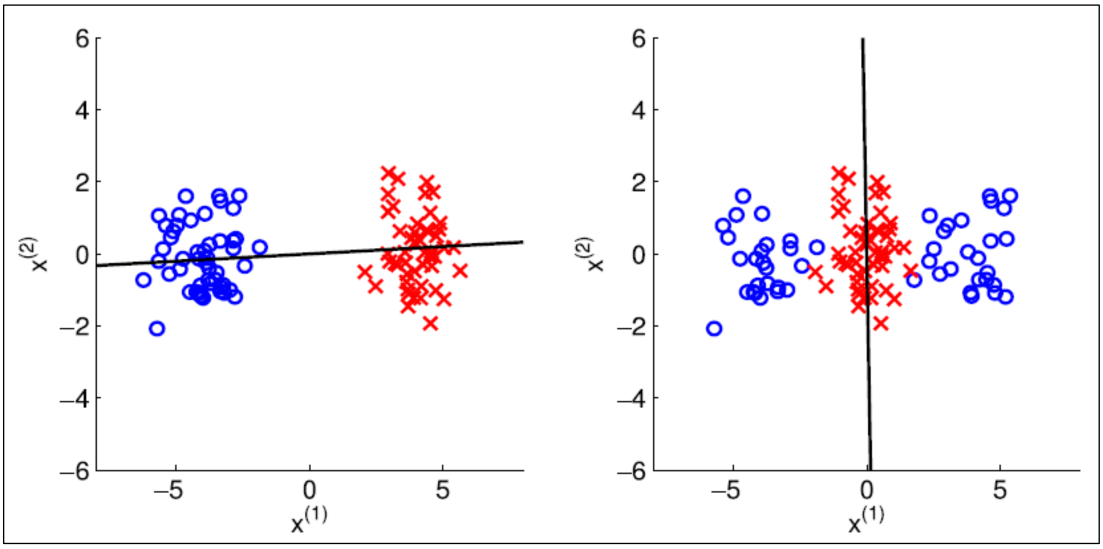
2.3.1 Linear Classification Models

2.3.1.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a technique developed in the year 1936, originally called Fisher's Discriminant Analysis (SCHOLKOPF, 1999). In the first time this technique was used to describe 2-class problem. The multi-class version was latter generalized by C. R. Rao as Multiple Discriminant Analysis.

These method projects a dataset onto a lower-dimensional space with good class-separability to avoid overfitting. The result combination a linear classifier, or in same cases, dimensionality reducer before subsequent classification (BISHOP, 2006). The basic idea of LDA is to find samples pairs in the same class get closer to each other and sample pair in different classes are far apart (Figure 7) (SUGIYAMA, 2015).

Figure 7 – Example of classifier data by Linear Discriminant Analysis (LDA) technique.



Legend: The solid lines denote the found subspaces to which training samples are projected.

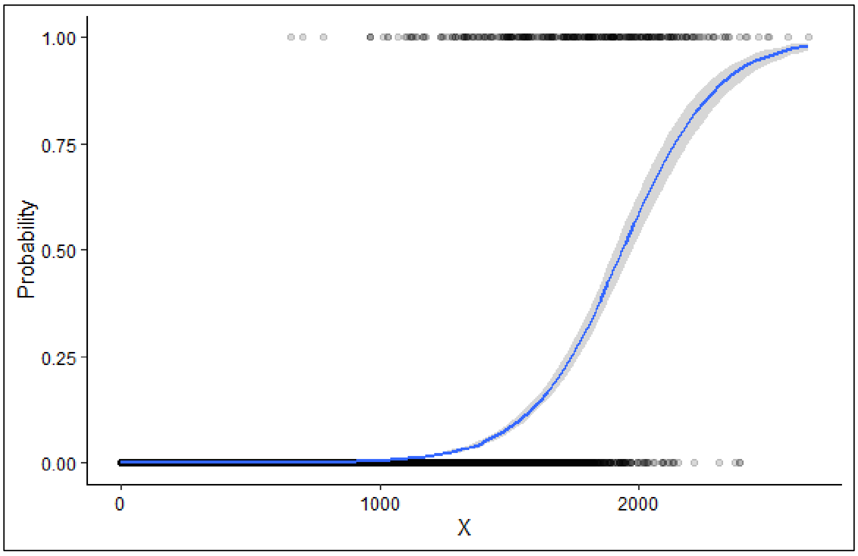
Source: SUGIYAMA, 2015

2.3.1.2 Generalized Linear Models

Generalized linear models (GLM) is a generic approach to a broad range of response modeling problems. Normal, Poisson, and binomial responses are the most commonly used, but other distributions can be used as well (FARAWAY, 2010). The GLM can be fitted using a common procedure and mechanism for hypothesis testing available. Use of a GLM is by no means sufficient as there are aspects of analysis of all the different GLMs which are specific to that particular response type. For example, while a logistic regression is a GLM the user still needs to understand the particular interpretations of odds in this type of model (GALLAGHER, 2007).

Logistic Regression (LR) was developed by David Cox in 1958, is the one of the larger class of techniques called Generalized Liner Models (GLMs) that encompass many different probability distributions (KUHN, 2013). Gudivada, (2016) consider LR an essentially classification algorithm. The word “regression” become from the linear regression, because the goal for the algorithms is to find the decision boundaries among the classes. The logistic function is a sigmoid function, which takes any real value between zero and one and allow us to estimate the probability (Figure 8).

Figure 8 – Illustrate example of classifier data by Logistic Regression.



Legend: Generic data generated in R for the propose illustrate example.

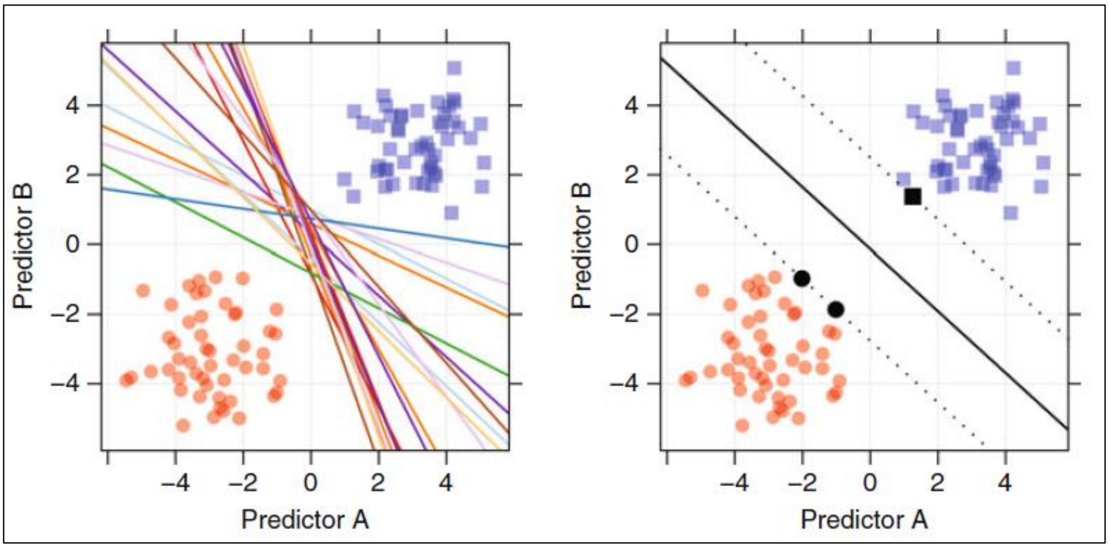
Source: elaborated by the author

2.3.2 Nonlinear Classification Models

2.3.2.1 Support Vector Machine

Support Vector Machine (SVM) was developed by Vladimir Vapnik in mid-1960s, are considered a class of statistical models (KUHN, 2013). According Vapnik, (2010), SVM is the most flexible and effective machine learning tool by produces significant accuracies. The objective to use SVM algorithm is to find a hyperplane in an *n* numbers of features (*n*-dimensional space) that distinctly classifies the data points (Figure 9).

Figure 9 – Illustrate example of Support vector machine classifier data.



Legend: In left are possible hyperplanes. In right are linear maximum margin classifier. The solid black point indicates the support vectors.

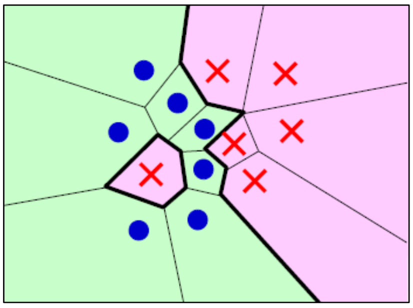
Source: KUHN, (2013)

To classifier data showed in figure 9 in two classes of data points, there are many possible hyperplanes that could be chosen (Figure 9 – left), but if choose one hyperplanes randomly, probability is not separate data points in equal distance of hyperplane. Using the support vectors, is possible to choose a hyperplane that represent the maximum distance between data points of both classes (Figure 9 – right) (SUGIYAMA, 2015).

2.3.2.2 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is one of many supervised learning algorithms used in data mining and machine learning, it is a good algorithm classifier. The algorithm works how similar is a data from other (DUTTA, 2015). But it is not robust against outliers, see figure 10 (SUGIYAMA, 2015).

Figure 10 – Illustrate example of K-Nearest Neighbors classifier data.



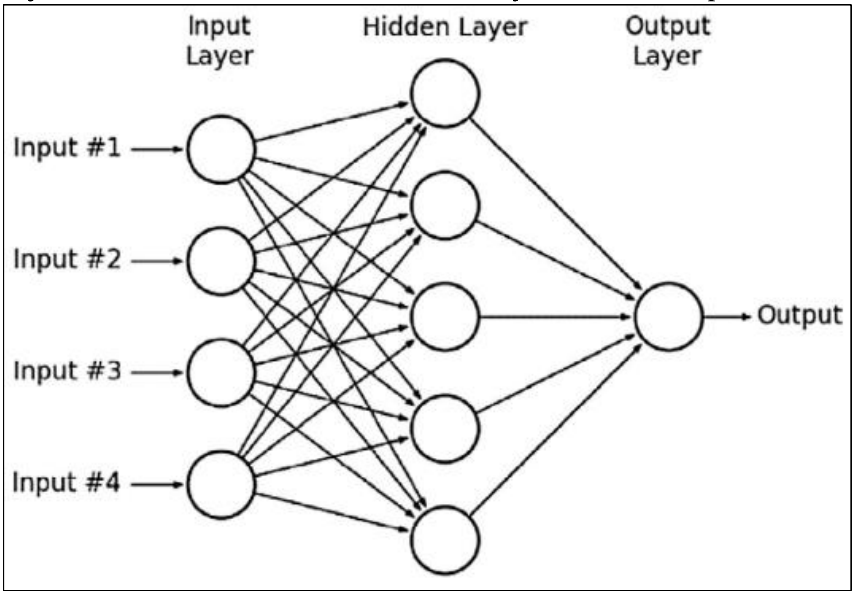
Source: SUGIYAMA, (2015)

Kuhn, (2013) describe the metric of algorithm classifier are based using datapoints geographic neighborhood to predict the classification. If the dataset contain outlier, for example, and this outlier is similar than one classes label probability the algorithm classifier and isolate the datapoint.

2.3.2.3 Artificial Neural Network

Artificial Neural Networks (ANN) works based on the human nervous system (DÓREA, 2018). Similar than biological neurons in which the neuron are responsible for receiving sensory input from an external stimulus via dendrites process the information a gives the output through axons, in artificial neural network the mechanism are not different. Independent variables or inputs are multiplied by a connection weight synapses, all inputs are analyzer and an activation function are apply. After this the results are be produce (Figure 11). The process of perceptron leaning is the adaptation of weight values until an acceptable relation between input and output obtain (SIQUEIRA-BATISTA et al., 2014).

Figure 11 – Illustrate example of structure of an ANN with four inputs nodes in the input layer, five hidden nodes in the hidden layer and one output in the output layer.



Source: BUNKER; THABTAH, 2019.

2.3.3 Classification Trees

2.3.3.1 Decision Trees

Decision Tree (DT) are a model used for classification and regression (GARETH, 2013). In this topic we are disserting about classifier form. There is a simple and intuitive predictive model similar then “if this than that” for choice decision (HUTCHINSON; GIGERENZER, 2005). DT are used to classifier animal behavior asking a series of simple yes or no questions (e.g. the animal grazing or not-grazing). Followed Valletta et al., (2017) a decision tree is constructed by three steps:

(1) Find the yes or no rule that best splits the data with respect to one of the features.
(2) The best split is the one that produces the most homogeneous groups.
(3) Repeat steps one and two until all data are correctly classified or some stopping rule is reached.

2.3.3.2 Random Forest

Random Forest (RF) was development to reduce overfitting problem in decision tree (VALLETTA et al., 2017). But, in RF instead of a single tree as DT, multiple tree is calculated. It is simply a collection of decision tree whose results are aggregated into one final result (BREIMAN, 2001).

2.3.4 Cross-validation strategies

To evaluate the performance of many machine learning models are needed to test it on some unsee data. Based the model performance on unsee data can shown if the model has fit, underfitting or overfitting. Cross-validation (CV) is one of the techniques used to test effectiveness of machine learning approach models. The k-fold cross-validation is the most commonly used strategy to validate models in animal behavior (HAMILTON et al., 2019; RIABOFF et al., 2019). The form of using CV are variable, can be use with or without replication data, as leave-one-out and leave-pair-out cross-validation (SMITH et al., 2014).

2.3.4.1 Holdout

In this strategy the sample are randomly split the complete data into training (performance) and test sets. A common split when using the holdout strategy is using 80% of data for training and the remaining 20% of the data for testing. In some cases, using 70% of data for training and 30% of data for testing (BUNKER; THABTAH, 2019).

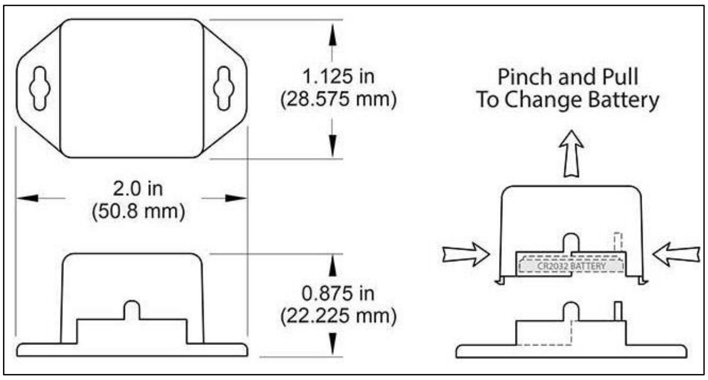
2.3.4.2 Leave-one out cross-validation

Leave-one-out cross-validation (LOOCV) is one form of cross-validation. In these processes, the number of folds is equal the number of instances in the dataset. Therefore, the algorithm is applied once time for each instance (WEBB et al., 2011). In LOOCV a single observation are isolated from the all data set. The algorithm are predicting model based in dataset without the part isolated. This part isolated are will used to validation process (SMITH et al., 2014).

3 MATERIAL AND METHODS

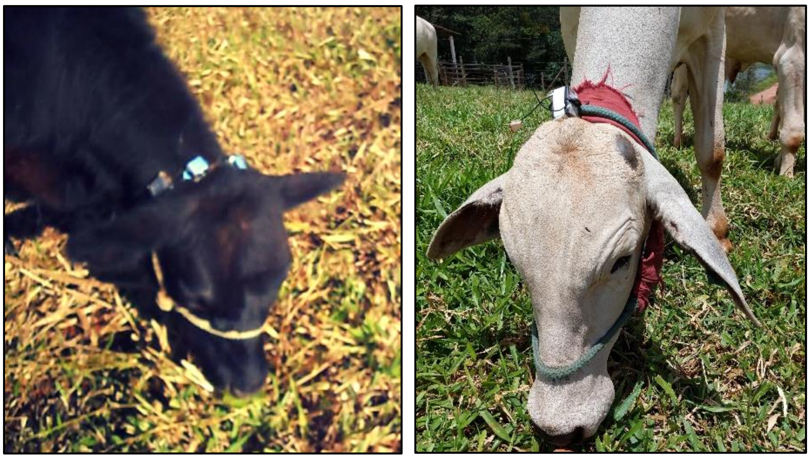
All experiments were conducted in the experimental farm of the Department of Animal Sciences at the Federal University of Lavras, located in Brazil (21°13'51.53" S, 44°58'10.52" W; 918 m above sea level). Two experiments were conducted, a preliminary study and the main study. The procedures of this project were authorized by the Ethics and Animal Welfare Committee of the university (protocol number 016/2018). In both studies, two types of sensors were used, a 3-axis ACC and a gyroscope (MONNIT, Salt Lake City, Utah, United States), both recording at 800 Hz rate (800 datapoints/second) and charged with a coin battery within the hardware case (Figure 12). Sensors were manufactured by MONNIT®, as was the software (MONNIT, Salt Lake City, Utah, United States) used to receive and export the raw datasets. The cases were attached to a halter that was fitted in the animals every collection day (Figure 13).

Figure 12 – Specifications of accelerometer and gyroscope sensors used in all experiments.



Source: MONNIT® (2018).

Figure 13 – Animals used on experiments



Legend: On left are Angus heifer used in preliminary study. On right are Nellore steer used in the main trial. Both with sensors attached on the halter.

3.3 Preliminary Experiment

The preliminary study was a pilot aimed at getting familiar with the sensors and data processing. We measured battery life with different sampling frequencies, the coverage area relative to the gateway, means to attach the sensors, the quality of the data base, and finally the processing of this amount of data.

We used two Angus heifers (Figure 14), with 224 ± 88,2 kg of body weight, grazing a 0.8 ha paddock of Marandu Palisadegrass with free access to fresh water.

Figure 14 – Angus heifers used in the pilot trial.

During collection periods, a smaller area, of 0.25 ha, was delimitated using an electric wire to keep the animals within the gateway best coverage area (Figure 15). Moreover, we positioned two range extensors to improve signal coverage.

Figure 15 – Preliminary experimental area map.

Legend: The experimental area used in the pilot trial (image by Google Earth Pro v7.3.2.5491. Scale 1:90. Picture obtained in 9/22/2018)
Source: elaborated by the author

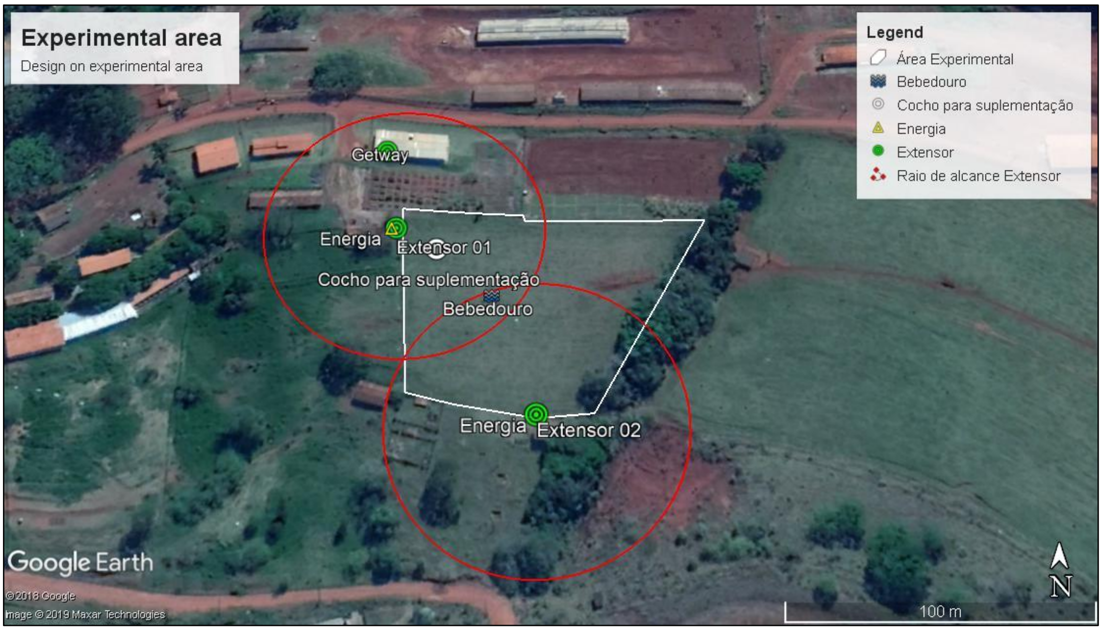
A total of 11 collection days were conducted during the pilot experiment. Each observation period lasted twelve hours (from approximately sunrise to sunset) on a day and only one animal received the sensors at a time. To register animal behavior, the Excel® macro VBA (Visual Basic Application) was set to record the exact time (minutes and seconds) that an activity changed. Prior to the beginning of the observations, the clock of both computer and sensor were synchronized. The behavior features observed were grazing, standing ruminating, lying ruminating, standing idleness, lying idleness, and drinking water. The animals were used to close human presence and did not change behavior when people were observing them in the paddock.

The sensors were set to record data in 10 Hz (10 samples/second). Data was constantly sent to the gateway plugged in a computer and the software stored it in a cloud server for posterior analysis. To analyze the data collected, we created three sets of behavior variables. Set 1 had six classes of behaviors: grazing, standing ruminating, lying ruminating, standing idle, lying idle and drinking water. Set 2 had four classes of behaviors: grazing, ruminating (both standing and lying down), idle (both standing and lying down) and drinking water. Set 3 had three classes behaviors: grazing, ruminating and idle (drinking water was compiled with idle). The sets of variables were analyzed with four machine leaning approaches. Bagged Trees (BT), Support Vector Machine (SVM), K-Nearest Neighbors (K-NN) and Linear Discriminant Analysis (LDA) were employed on the raw data extracted from the 3-axis accelerometers and 2-axis gyroscope. To assess prediction quality, a 5 k-fold cross-validation was performed and overall accuracy, true positives, false negatives and area under the curve (AUC) of the receiver operating characteristics were calculated for each machine learning approach and each set of behavior parameters.

3.4 Main Experiment

The study was conducted in the experimental farm at the Federal University of Lavras, Brazil (21°13'51.53" S, 44°58'10.52" W; 0,8ha; 918m above sea level). Before introducing the animals, the paddock was fertilized with 50 kg of urea/hectare. After 30 days from the application of fertilizing all animals were placed in the paddock (Figure 16) and canopy height means were measured daily using a sward stick (BARTHRAM, 1985). Forage allowance was measured every day. In each evaluation, 5 representative points (with heights close to the average of the paddock) were selected, and in each point, the material within a 1.0 x 0.5 m frame was cut to ground level. Total forage mass was weighed, and a representative subsample of 0.5 kg was taken and separated into leaf blades, stems (including leaves sheaths), reproductive stem and dead material (as indicated by more than 50% of the tissue area being senescent) to determine the sward morphological composition. After separation, forage samples were oven-dried at 66°C for 72 hours to a constant weight.

Figure 16 – Experimental area map used in second step.



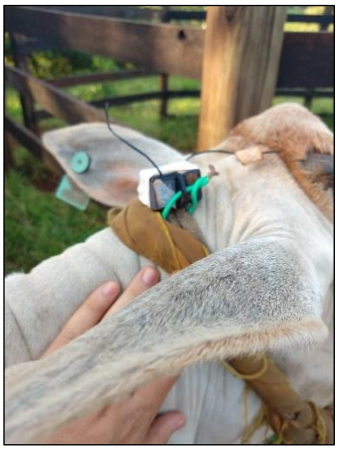
Legend: Experimental area used in step two image by Google Earth Pro v7.3.2.5776.
Scale 1:100. Picture obtained in 11/18/2019.
Source: elaborated by author

Six Nellore bulls, (345.3 ± 21.4 kg body weight), previously adapted to human handling were used. All animals were kept on pasture of Marandu Palisadegrass with water and mineral salt ad libitum. Animal behavior was collected through visual observation within a period of 10 hours (0800 to 1800) during 15 days in total. Animals had access to the paddock on day 1, when the average sward height was 31.5 cm, forage allowance was 1,826 kg of DM.ha^{-1} and forage composition was 58.1% of leaves, 19.9% of stem, and 22.0% of dead material. On day 16, average sward height was 19.5 cm, the forage allowance was 1,230 kg DM.ha^{-1} and the composition was 29.7% of leaves, 34.2% of stem, 36.1% of dead material. Table 1.

Animal behavior was classified into two classes: grazing and not-grazing. Grazing activity was considered when the animal was searching and grazing, and every other activity including ruminating, idleness, walking, and drinking were considered not-grazing activity. To facilitate the visual observation of behavioral activities, each animal was painted with a nontoxic ink on thoracic region and using halters of different colors.

A 3-axis (X, Y and Z) wireless accelerometer sensor (MONNIT, 2018) range ± 2 g; was attached to the halters on back of the neck of each animal (Figure 17). The X, Y and Z axes indicate longitudinal (front-to-back), horizontal (side-to-side) and vertical (upto-down) head movements, respectively (SHEPARD et al., 2010). Devices were setup to send simultaneously data from each animal at the frequency of 1 data point (G-force for X, Y, and Z) per second (Figure 22). Energy supply for the sensors was powered with a coin cell battery, model CR2032 of 3.0 voltage. To register visual animal behavior, the Excel® macro VBA (Visual Basic for Applications) was set to record the exact time (minutes and seconds) that an activity changed. Prior to the beginning of the observations, the clock of both computer and sensor were synchronized.

Figure 17 – Sensors utilized in experiment on attached on halter Nellore bull.



3.4.4 Predictive Approaches

Three machine learning approaches (Generalized Linear Model, Random Forest, and Artificial Neural Network) were employed on the raw data extracted from the 3-axis accelerometers to predict two classes of animal behavior: grazing and not-grazing.

Generalized linear model (GLM) is a method to provide a linear modeling predictor function of exploratory variables and dealing within non-normal error structures (BOURNE et al., 2007). To fit a generalize linear model two hyperparameters were required lambda and alpha (both defined between 0 and 1). The lambda parameter controls the amount of the regularization applied to the model, where larger lambda shrinkage the coefficients toward to zero, and alpha parameter controls the distribution between LASSO ($\ell 1$) and ridge regression ($\ell 2$) penalties. The best values were chosen by performing a random discrete grid search using a combination of maximum runtime per 360 seconds and/or max number of 100 models as early stopping criterion, and misclassification as stopping metric.

Random forest (RF), is one of the most used machine learning technique in sensor data for classification of animal behavior (LUSH et al., 2018; ALVARENGA et al., 2016). This method was developed to solve problem of overfitting in regular decision tree (BREIMAN, 2001). A random discrete grid search was performed using the following hyperparameters: number of trees (10, 20, 40, 80, 100, 200, and 300), minimum number of observations per leaf (1, 2, 10, 20, and 30), number of variables in the subset (2, 3, 4, 5, and 6), and maximum tree depth (1, 10, 20, 40, and 80). The combination of maximum runtime per 860 seconds and/or max number of 100 models were used as early stopping criterion and misclassification as stopping metric.

Finally, artificial neural network (ANN) is based on the human nervous system (STAUDENMAYER et al., 2009), which can deal with complex relationship between input data and the response variable (BREWSTER et al., 2018). A random grid search to determine the best ANN architecture was carried out using six activation functions (Rectifier, RectifierWithDropout, Tanh, TanhWithDropout, Maxout, and MaxoutWithDropout), different number of hidden layers (1, 2, 3, and 4) and number of neurons (20, 30, 50, 80, and 100) by hidden layers, except 100 neurons that was not used for 3 and 4 hidden, dropout ration (0.10 and 0.20), and the regularization hyperparameters $\ell1$ and $\ell2$ (both in a range from 0 to 0.0001). The maximum number of 100 models was used as early stopping criterion and misclassification as stopping metric.

The tuning of all hyperparameters (GLR, RF, and ANN) was perform using 5-fold cross-validation using the H2O package (ERIN LEDELL et al., 2020) implemented in R (R CORE TEAM, 2019). The parameter tuning was performed only using the training set. The validation set was previously excluded to compose the three “Cross-Validation Strategies” evaluated:

1) Holdout: 20% of the training set was randomly excluded and used in validation set and 80% was used in the training set; 2) Leave-one-animal-out (LOAO): all sensor data from a specific animal was excluded per time and used as validation set; and 3) Leave-one-day-out (LODO): the last five days of observations were excluded and used in the validation, while the first 10 days were used in training process. For LOAO strategy, our goal was to evaluate the predictive performance of machine learning methods when information of new animals is utilized. The LODO strategy was implemented to evaluate the predictive performance of a machine learning methods when a new grazing management (or grazing condition) is imposed. Sward characteristics are greatly affected by the process of grazing, thus every consecutive day mimics a new grazing condition.

As mentioned before, the 5-fold cross-validation was used during the training process for all validation strategies and machine learning methods. To evaluate the performance of the prediction models, accuracy, error-rate (ERR), sensitivity, specificity, positive predicted value (PPV), and negative predicted value (NPV) were calculated using the following equations: Accuracy = (TP + TN) / (TP + TN + FP + FN), ERR = 1 – Accuracy, Sensitivity = TP / (TP + FN), Specificity = TN / (TN + FP), where TP, FP, FN and TN are true positive, false positive, false negative and true negative respectively.

4 RESULTS AND DISCUSSIONS

4.1 Preliminary Experiment

The overall accuracy of all machine learning approaches was better for Set 3 (Table 1). In Set 3, when BT was implemented, the true positives rates for grazing and ruminating behavior were 72% and 59%, respectively, (Figure 18). Idle behavior was poorly predicted (26%) by this method. The AUC curves were 0.77; 0.71; and 0.55 for BT, SVM, K-NN and LDA, respectively.

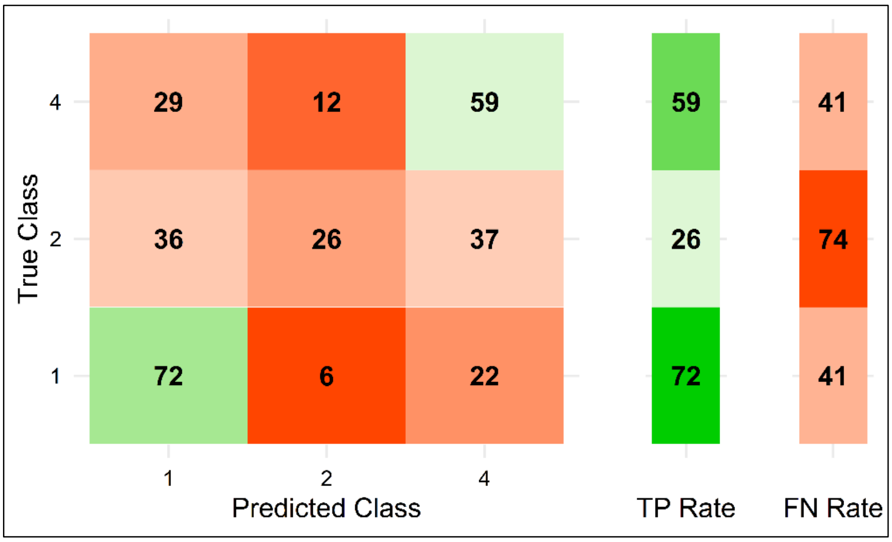
Table 1 – Overall accuracies validation of all machine learning approaches used to classifier different sets of Angus behavior.

Machine Learning Approaches	**Overall Accuracy (%)**		
	Set 1 (6 classes)	Set 2 (4 classes)	Set 3 (3 classes)
BT	51.3	57.9	58.8
SVM	48.1	53.7	54.7
K-NN	46.5	53.2	54.4
LDA	41.4	46.9	45.9

Legend: Each set behavior contain respectively behavior features. Set1: grazing, ruminating standing, ruminating lying down, idle standing, idle lying down and drinking. Set2: grazing, ruminating (both standing and lying down), idle (both standing and lying down) and drinking. Set3: grazing, ruminating and idle.

The use of Set 3 obtained higher accuracy values to predict grazing time using BT, as well as SVM and K-NN. These results encourage us to group the behavior traits of the main trial in only two classes (grazing and not-grazing), as well as to use a type of BT approach (RF) as a reference algorithm in the main trial.

Figure 18 – Confusion matrix of classification accuracy of Angus heifer cattle behavior using decision-tree algorithm to Set3.

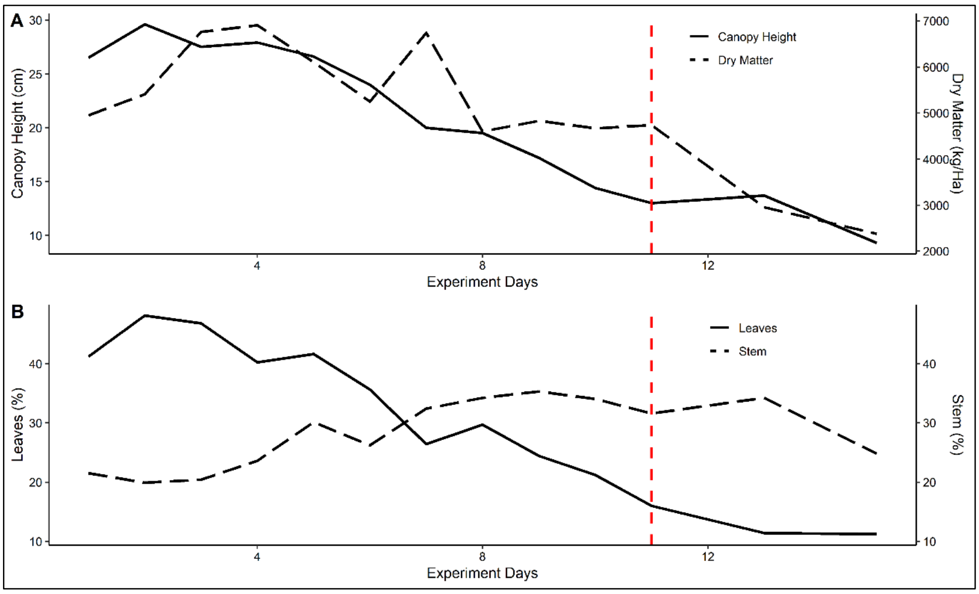


Legend: (1) grazing, (2) idle and (4) ruminating behaviors, respectively. TP Rate, corresponding True Positive Rate. FN Rate, corresponding False Negative Rate. The classification accuracy for each behavior can be read on the diagonal. The deeper green color highlights results with higher precision.

4.2 Main Experiment

The description of the sward structure is presented in Figure 19. As the animals grazed down the pasture, forage availability, canopy height and proportion of leaves in the sward decreased, while the proportion of stem increased. Changes in sward structure, especially forage mass and the ratio leaf:stem directly affect variables of grazing behavior, such as grazing time and bite rate (DA SILVA; CARVALHO, 2005).

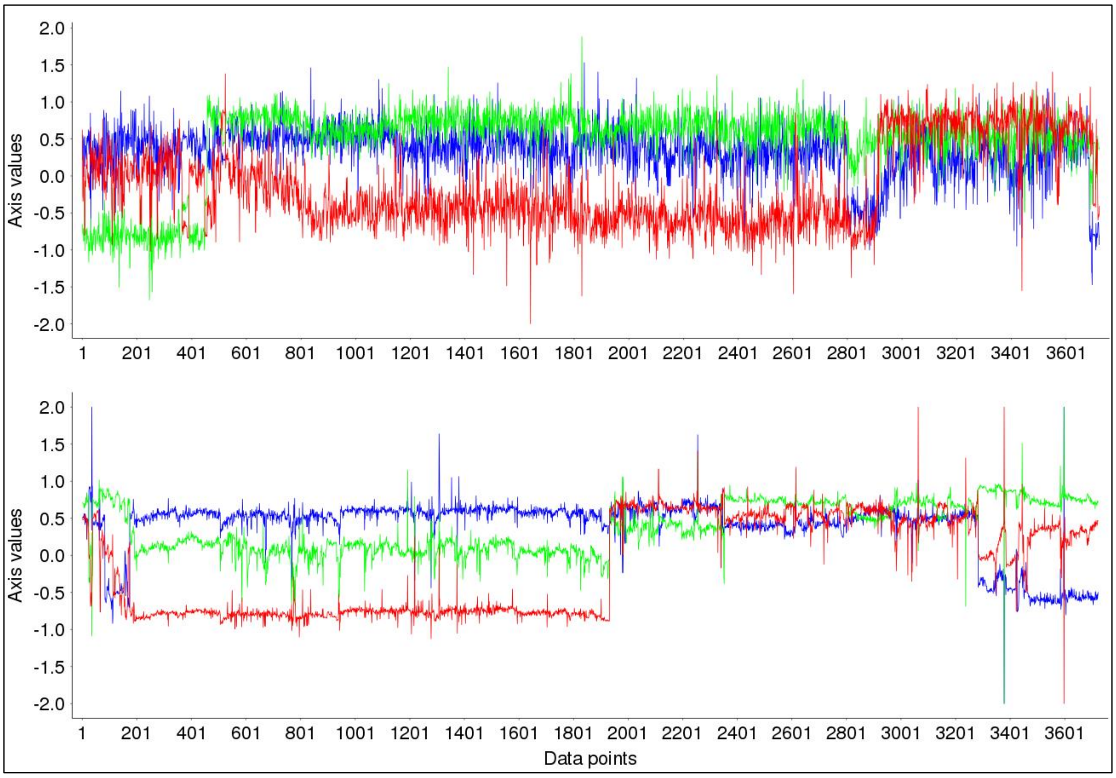
Figure 19 – (A) Evaluation of canopy height and forage offer during of experiment and (B) percentage of leaves and stem during of experiment.



Legend: The vertical red line indicates the division data using for training model and validation model (last five days).

Therefore, it is of outmost importance to evaluate whether predictions developed with one sward structure work well in a different structure. Raw ACC data plotted in graphic (Figure 20) showing a clear difference between grazing and not-grazing. The total number of observations for each activity (grazing and not-grazing), each dataset (training and validation) and each validation strategy are presented in Table 2.

Figure 20 – Raw data distribution sampled from one experimental point day.



Legend: for grazing (top) or not-grazing (bottom) behavior categories visually observed. The X, Y, and Z accelerometer axis values (G-forces) are represented in blue, green and red colors respectively.

Table 2 – Sample size throughout different validation strategies by individual behavior categories.

Behavior	Total	Development			Validation		
		LOAO[1]	LODO[1]	Holdout	LOAO	LODO	Holdout
Grazing	51,095	43,023	28,608	40,972	8,072	22,487	10,123
Not-grazing[1]	58,143	48,895	45,599	46,391	9,248	12,544	11,752

[1]Leave-one-animal-out (LOAO), Leave-one-day-out (LODO), and testing model in a random 20 % of the data set (holdout). Not-grazing category included animal ruminating, idle, walking, and drinking water.

The performance of each ML approach for model training and validation is presented in Table 3 and 4, respectively. The accuracy of all predictive approaches was greater for the training set than for the validation set. Such result is especially important when machine learning techniques are employed. Several machine learning methods, including the ones used in this current study (RF and ANN), have many parameters that are tuned during training process. During that process, the high accuracies commonly observed are due to overfitting and the main reason of having an independent dataset is to avoid this type of issue. Machine learning algorithms are more susceptible to overfitting compared to GLR, for example. This fact can be noted by comparing the changes in accuracy for each predictive approach on training (Table 3) and validation datasets (Table 4). The accuracy for GLR averaged 56.0% for the training set and 52.5% for the validation set, while the RF and ANN dropped from 77.1% and 74.7% to 64.7% and 68.6%, respectively.

Table 3 – Performance of machine learning approaches in model development to predict grazing or not-grazing behavior categories visually observed in Nellore cattle using different validations strategies.

	Accuracy	Error Rate	Sensitivity	Specificity	PPV[1]	NPV[1]
Leave-one-animal-out						
GLR[1]	55.9	44.1	9.6	96.9	73.4	54.7
RF[1]	78.0	22.0	76.7	79.1	76.5	79.3
ANN[1]	74.1	25.9	69.5	78.0	73.6	74.4
Leave-one-day-out						
GLR	56.0	44.0	66.7	45.1	55.3	57.0
RF	77.4	22.2	60.3	88.1	76.1	78.0
ANN	75.8	24.2	63.5	83.6	70.8	78.5
Holdout (20%)						
GLR	56.1	43.9	17.9	89.7	60.3	55.5
RF	75.9	24.1	72.8	78.6	75.0	76.7
ANN	74.2	25.8	71.5	76.6	73.0	75.3

Generalized Linear Regression (GLR), Random Forest (RF), Artificial Neural Network (ANN), Positive Predicted Values (PPV) and Negative Predicted Values (NPV).

Table 4 – Validation of the machine learning approaches to predict grazing or not-grazing behavior categories visually observed in Nellore cattle using different validation strategies.

	Accuracy	Error Rate	Sensitivity	Specificity	PPV[1]	NPV[1]
Leave-one-animal-out						
GLR[1]	53.1	46.9	16.2	84.5	47.2	54.2
RF[1]	58.7	41.3	56.0	61.0	55.0	61.9
ANN[1]	72.0	28.0	65.3	77.8	72.0	72.0
Leave-one-day-out						
GLR	47.4	52.6	27.0	67.7	45.4	48.2
RF	58.5	41.5	51.1	71.8	76.5	45.0
ANN	59.7	40.3	54.0	70.0	76.3	45.9
Holdout (20%)						
GLR	57.1	42.8	24.7	85.9	60.7	56.3
RF	76.9	23.1	74.0	79.4	76.1	77.5
ANN	74.2	25.8	70.7	77.3	72.8	75.4

Generalized Linear Regression (GLR), Random Forest (RF), Artificial Neural Network (ANN), Positive Predicted Values (PPV) and Negative Predicted Values (NPV).

Regardless of the validation strategy, RF and ANN had a more accurate prediction of grazing activity than GLR and ANN. The linear approach (GLR) achieved an expressively worse accuracy (averaging 56.0%) compared to the machine learning techniques (averaging 64.7% and 68.6% for RF and ANN, respectively). The RF technique had showed superior accuracy in classifying data from sensors for animal behavior in similar studies (SMITH et al., 2016), being the most used tool to process this type of data. The overall accuracy found in our study for machine learning methods was similar to studies predicting grazing behavior through sensor technology (PENG et al., 2019; BARWICK et al., 2018b; RAHMAN et al., 2018; ALVARENGA et al., 2016; GONZÁLEZ et al., 2015). Such results indicate the potential of coupling machine learning techniques and sensor technology to predict complex behaviors such as grazing activity. Predicting grazing behavior would greatly improve farm management decisions related to grazing management (HOMBURGER et al., 2014), animal health (BARKER et al., 2018), and welfare (KUŹNICKA; GBURZYŃSKI, 2017).

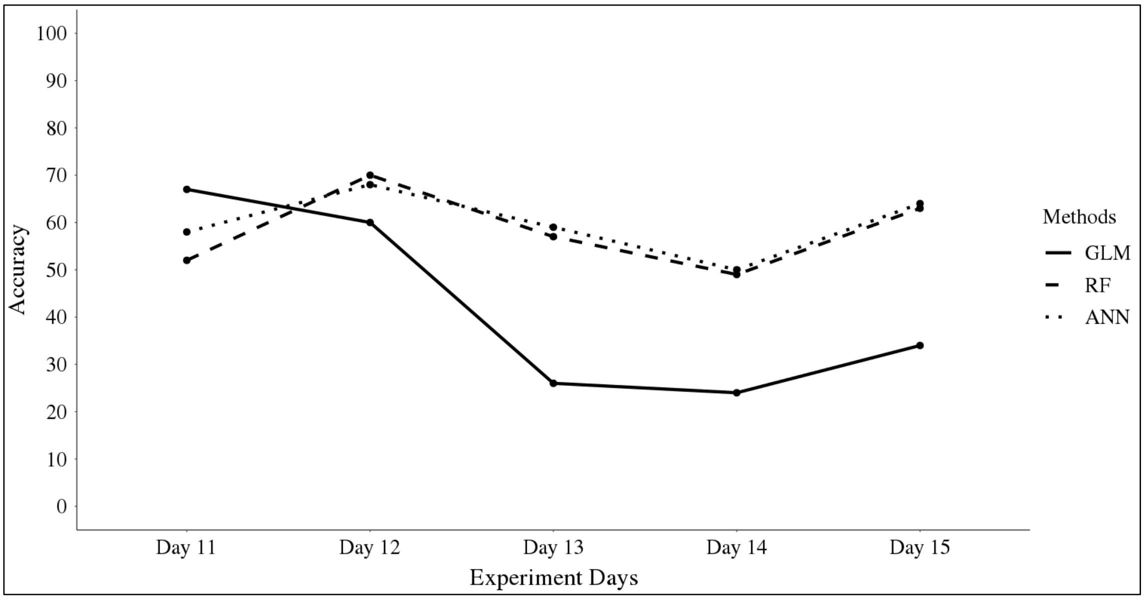
The capacity of detecting grazing activity is expressed by the sensitivity of the model, while the capacity of detecting not-grazing activity is shown by the specificity. Even though GLR had the highest specificity value in the LOAO and holdout validation strategies, its sensitivity was very poor (Table 4). On the other hand, for the LODO strategy, GLR had the highest sensitivity value but poor specificity. Additionally, GLR had the greatest error rate among the techniques for all validation strategy, confirming that it is not an adequate technique to predict grazing in these conditions (*i.e.* with a varying sward structure).

The first thing worth noticing is that the accuracy values are smaller than the ones in Table 3. This is expected because during the process of modeling developing, the training uses data within the dataset (5k-fold), while the validation process creates independent sets to test the model. Since the final goal is to predict traits in new animals and conditions that were not included in the training set, it is important to achieve high accuracy in the validation analysis.

The holdout validation strategy yielded the highest accuracy values for all three machine learning approaches. This strategy randomly selects 20% of the database to use as an independent set for validation. However, this independence might not be completely true. Even though the datapoints in the validation set were not used for training the model, the remaining 80% still contain other datapoints from the same animal and same pasture conditions. This can bring carryover effects that inflate prediction accuracy due to overfitting problems (WANG; BOVENHUIS, 2019; DÓREA et al., 2018) when compared with a truly independent validation set, in which all datapoints from the same animal or from the same pasture conditions are excluded. Indeed, the LOAO and LODO validation strategies yielded lower accuracy values, confirming the overfitting effect of holdout.

Interestingly, the overall performance of the models was worse for LODO than for LOAO. With the LODO strategy, the training set (first 11 days of grazing period) had very different sward conditions from the validation set (last 5 days of grazing period), as shown in Table 1 and Figure 1. None of the algorithms was able to account for this change in pasture condition and yielded inadequate accuracy (all smaller than 60%) and sensitivity (all smaller than 54%) values. However, the more the sward condition changed, as days passed, the worse was the accuracy for GLR (Figure 21). This did not happen for RF and ANN, that kept their accuracy value throughout the validation period.

Figure 21 – Accuracy values of last five days as a validation to grazing for LODO.



The poor performance of the models with LODO validation is an interesting finding since the goal of this kind of research is to develop a tool that can be used to accurately predict grazing behavior in order to assist with pasture management under different environmental conditions (e.g. different grass and animal types). Sward structure changes drastically depending on management strategy and climate conditions and the algorithm needs to be able to perform regardless. One possible explanation for the poorest performance of LODO compared to LOAO was that the signal created by the ACC as a result of a behavioral activity would be more similar among animals than grazing managements. Animals will tend to behave and move similarly during grazing as long as they are under the same management. Thus, although animal to animal variation might exist, it is probably smaller compared to how a new (unknown by the algorithm) grazing management (e.g. sward height) will affect the behavioral pattern within and across animals. In this context, the development predictive analytics to be coupled with sensor technology using datasets from limited grazing management scenarios may limit the efficacy of technology implementation because models will not perform as expected. Such lack of model generalization was not observed when one animal was excluded at a time (LOAO), which means that although a one animal is not used for algorithm training, the data acquired by other animals may better represent the behavioral patterns of new animals (unknown by the algorithm).

When all data from a single animal were excluded from the training set and used for validation (LOAO), the accuracy was still poor for GLR (53.1%) and RF (58.7%) but increased with ANN (72%). This validation strategy aimed to test if the algorithm would still be able to predict grazing behavior when the sensor is placed in a new animal, that was not used in the training. This is of outmost importance for the application of the tool in commercial farms, where the visual observation of animal behavior is infeasible. We were able to develop a model, using ANN, that yielded adequate prediction capacity. With a larger dataset (*i.e.* more animals and grazing management scenarios), it is expected that this accuracy could be increased, considering an appropriate strategy of validation which is a key factor to create robust and accurate sensor technology and minimize frustration of users in real-life applications.

Therefore, our data suggests that in order to build a robust prediction model, it is important to combine statistical knowledge with biological information to avoid misleading recommendations. Most of the studies that validate the use of sensors to predict animal behavior used the k-fold strategy, excluding random sets of data and ignoring the interdependence of training and validation sets. This overfitting inflates the accuracy and may incur in frustration when such tool is applied in a commercial setting, with new animals and varying pasture conditions.

5 CONCLUSIONS

Our results demonstrated that the validation strategy does interfere with the accuracy of predicting models and that random choice of datapoints (such as in holdout validation) inflate accuracy values. Removing all data from one animal (LOAO) or one pasture condition (LODO) decreased accuracy, suggesting carry over effects not accounted for with the holdout validation strategy. Therefore, the performance of an algorithm developed within a specific scenario and validated with holdout strategies might perform poorly in a different scenario, limiting the practical application of the tool.

Another conclusion was that the linear approach was not adequate to predict such as complex behaviors such as grazing activity, regardless of the validation strategy. The RF method had the best performance with the holdout validation but decreased to inadequate values when LOAO or LODO were used. Finally, when compared to the holdout validation, ANN also lost accuracy with LODO but not with LOAO, achieving an adequate accuracy with this validation strategy. This is a promising result that strengthens the potential for such technologies to become decision-making tools in the farm.

REFERENCES

ALVARENGA, F. A. P. et al. **Using a three-axis accelerometer to identify and classify sheep behaviour at pasture.** Applied Animal Behaviour Science, v. 181, p. 91–99, 1 ago. 2016.

ANDRIAMANDROSO, A. L. H. et al. **A review on the use of sensors to monitor cattle jaw movements and behavior when grazing.** Biotechnology, Agronomy and Society and Environment, v. 20, n. S1, p. 273–286, 2016.

B. SCHOLKOPFT; K.-R. MULLERT. **Fisher discriminant analysis with kernels.** Neural networks for signal processing IX, 1999.

BARKER, Z. E. et al. **Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle.** Journal of Dairy Science, v. 101, n. 7, p. 6310–6321, 1 jul. 2018.

BARTHRAM, G. T. **Experimental techniques: the HFRO sward stick.** The Hill Farming Research Organization Biennial Report 1984/1985, v. Biennial R, n. HFRO, p. 29–30, 1985.

BARWICK, J. et al. **Categorising sheep activity using a tri-axial accelerometer.** Computers and Electronics in Agriculture, v. 145, p. 289–297, 1 fev. 2018b.

BARWICK, J. et al. **Predicting lameness in sheep activity using tri-axial acceleration signals.** Animals, v. 8, n. 1, p. 1–16, 11 jan. 2018a.

BEKER, A. et al. **Energy expenditure and activity of different types of small ruminants grazing varying pastures in the summer.** Journal of Applied Animal Research, v. 37, n. 1, p. 1–14, mar. 2010.

BETTERIDGE, K. et al. **Sensors for detecting and logging spatial distribution of urine patches of grazing female sheep and cattle.** Computers and Electronics in Agriculture, v. 73, n. 1, p. 66–73, 2010.

BISHOP, CM. **Pattern recognition and machine learning.** New York (NY, USA): Springer; 2006.

BOURNE, C. M. et al. **Generalized Linear Model Analysis in Ecology.** Memorial University of Newfoundland and Labrador, p. 1–50, 2007.

BREIMAN, L. **Random forests.** Machine Learning, v. 45, n. 1, p. 5–32, out. 2001.

BREWSTER, L. R. et al. **Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data.** Marine Biology, v. 165, n. 4, p. 62, 8 abr. 2018.

BROSH, A. et al. **Energy cost of cows' grazing activity: Use of the heart rate method and the Global Positioning System for direct field estimation.** Journal of Animal Science, v. 84, n. 7, p. 1951–1967, jul. 2006.

BUNKER, R. P.; THABTAH, F. **A machine learning framework for sport result prediction.** Applied Computing and Informatics, v. 15, n. 1, p. 27–33, 2019. Available in: <https://doi.org/10.1016/j.aci.2017.09.005>.

C. M. BISHOP. **Pattern Recognition and Machine Learning.** Secaucus, NJ, USA: Springer, 2006.

CARVALHO, P. et al. **Importância da estrutura da pastagem na ingestão e seleção de dietas pelo animal em pastejo.** Anais da XXXVIII Reunião Anual da Sociedade Brasileira de Zootecnia, v. 1, n. 2000, p. 853–871, 2001.

CASAGRANDE, D. R. et al. **Canopy characteristics and behavior of nellore heifers in brachiaria brizantha pastures under different grazing heights at a continuous stocking rate.** Revista Brasileira de Zootecnia, v. 40, n. 11, p. 2294–2301, 2011.

CHAMPION, R. A.; RUTTER, S. M.; PENNING, P. D. **An automatic system to monitor lying, standing and walking behaviour of grazing animals.** Applied Animal Behaviour Science, v. 54, n. 4, p. 291–305, nov. 1997.

CHELOTTI, J. O. et al. **A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle.** Computers and Electronics in Agriculture, v. 127, p. 64–75, 1 set. 2016.

DA SILVA, S. C.; CARVALHO, P. C. de F. **Foraging behaviour and herbage intake in the favourable tropics/sub-tropics.** Grassland: A Global Resource, p. 81–95, 2005.

DÓREA, J. R. R. et al. **Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows.** Journal of Dairy Science, v. 101, n. 7, p. 5878–5889, 1 jul. 2018.

DREWRY, J. L. et al. **Assessment of digital technology adoption and access barriers among crop, dairy and livestock producers in Wisconsin.** Computers and Electronics in Agriculture, v. 165, n. August, p. 104960, 2019.

DUNNE, R. et al. **Behavior classification of dairy cows fitted with GPS collars.** In: KANG, U AND LIM, EP AND YU, JX AND MOON, Y. (Ed.). Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Lecture Notes in Artificial Intelligence. Australia: Springer Verlag, 2017. 10526 LNAIp. 15–25.

DUTTA, R. et al. **Dynamic cattle behavioural classification using supervised ensemble classifiers.** Computers and Electronics in Agriculture, v. 111, p. 18–28, 1 fev. 2015.

ERIN LEDELL ET AL. **h2o: R Interface for the "H2O" Scalable Machine Learning Platform**, 2020.

FARAWAY, J. J. **In International Encyclopedia of Education** (Third Edition), 2010.

GALLAGHER, C. **Extending the Linear Model With R: Generalized Linear, Mixed Effects and Nonparametric Regression Models.** Journal of the American Statistical Association, 2007.102:480, 1477, DOI: 10.1198/jasa.2007.s238

GARETH JAMES, DANIELA WITTEN, TREVOR HASTIE, ROBERT TIBSHIRANI. **An Introduction to Statistical Learning: with Applications in R.** New York: Springer, 2013.

GONZÁLEZ, L. A. et al. **Behavioral classification of data from collars containing motion sensors in grazing cattle.** Computers and Electronics in Agriculture, v. 110, p. 91–102, 1 jan. 2015.

GUDIVADA, V. N. et al. **Cognitive Analytics: Going Beyond Big Data Analytics and Machine Learning.** In: Handbook of Statistics. Elsevier B.V., 2016. v. 35p. 169–205.

GUO, L. et al. **Comparison of grazing behaviour of sheep on pasture with different sward surface heights using an inertial measurement unit sensor.** Computers and Electronics in Agriculture, v. 150, p. 394–401, jul. 2018.

HAMILTON, A. W. et al. **Identification of the rumination in cattle using support vector machines with motion-sensitive bolus sensors.** Sensors (Switzerland), v. 19, n. 5, mar. 2019.

HOMBURGER, H. et al. **Inferring behavioral states of grazing livestock from high-frequency position data alone.** PLoS ONE, v. 9, n. 12, p. e114522, 4 dez. 2014.

HUIRCÁN, J. I. et al. **ZigBee-based wireless sensor network localization for cattle monitoring in grazing fields.** Computers and Electronics in Agriculture, v. 74, n. 2, p. 258–264, nov. 2010.

HUTCHINSON, J. M. C.; GIGERENZER, G. **Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet.** Behavioural Processes, v. 69, n. 2, p. 97–124, 2005.

KUHN, MAX; KJELL. JOHNSON. 2013. **Applied Predictive Modeling**. New York: Springer.

KUŹNICKA, E.; GBURZYŃSKI, P. **Automatic detection of suckling events in lamb through accelerometer data classification.** Computers and Electronics in Agriculture, v. 138, p. 137 – 147, jun. 2017.

LAHART, B. et al. **Predicting the dry matter intake of grazing dairy cows using infrared reflectance spectroscopy analysis.** Journal of Dairy Science, v. 102, n. 10, p. 8907–8918, 2019.

LARY, D. J. et al. **Machine learning in geosciences and remote sensing.** Geoscience Frontiers, v. 7, n. 1, p. 3–10, 2016.

LUSH, L. et al. **Classification of sheep urination events using accelerometers to aid improved measurements of livestock contributions to nitrous oxide emissions.** Computers and Electronics in Agriculture, v. 150, n. April, p. 170–177, 1 jul. 2018.

MASASHI SUGIYAMA. 2015. **Introduction to Statistical Machine Learning.** Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

MATSUI, K.; OKUBO, T. **A 24-hour Automatic Recording of Grazing and Ruminating Behaviour in Cattle on Pasture by a Data Logger Method.** Nihon Chikusan Gakkaiho, v. 60, n. 10, p. 940–945, 1989.

MOROTA, G. et al. **Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture.** Journal of Animal Science, v. 96, n. 4, p. 1540–1550, 2018.

PENG, Y. et al. **Classification of multiple cattle behavior patterns using a recurrent neural network with long short-term memory and inertial measurement units.** Computers and Electronics in Agriculture, v. 157, p. 247–253, 1 fev. 2019.

R CORE TEAM. R: **A language and environment for statistical computing R** Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/., 2019. Disponível em: <http://www.r-project.org>.

RAHMAN, A. et al. **Cattle behaviour classification from collar, halter, and ear tag sensors.** Information Processing in Agriculture, v. 5, n. 1, p. 124–133, 1 mar. 2018.

RIABOFF, L. et al. **Evaluation of pre-processing methods for the prediction of cattle behaviour from accelerometer data.** Computers and Electronics in Agriculture, v. 165, p. 104961, out. 2019.

SAKAI, K. et al. **Behavior classification of goats using 9-axis multi sensors: The effect of imbalanced datasets on classification performance.** Computers and Electronics in Agriculture, v. 166, p. 105027, nov. 2019.

SCHOLKOPF, B.; BURGES, A.; SMOLA, A. **Advances in kernel methods: support vector learning.** Cambridge, Mass., MIT Press. p. 147-167, ISBN 0-262-19416-3. 1999.

SHALLOO, L. et al. **Review: Grass-based dairy systems, data and precision technologies.** Animal, v. 12, n. s2, p. S262–S271, 22 dez. 2018. Disponível em: <https://www.cambridge.org/core/product/identifier/S175173111800246X/type/journal_article>.

SHETTY, N. et al. **Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk.** Journal of Dairy Science, v. 100, n. 1, p. 253–264, 2017.

SIQUEIRA-BATISTA, R. et al. **Artificial Neural Networks and medical Education as redes Neurais artificiais e o Ensino da medicina.** Andréia Patrícia Gomes III Alcione de Paiva Oliveira III Ricardo dos Santos Ferreira III Vanderson Esperidião-Antonio III. Revista Brasileira de educação Médica, v. 38, n. 4, p. 557–565, 2014.

SMITH, D. et al. **Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems.** Computers and Electronics in Agriculture, v. 131, p. 40–50, 2016.

SMITH, G. C. S. et al. **Correcting for Optimistic Prediction in Small Data Sets.** American Journal of Epidemiology, v. 180, n. 3, p. 318–324, 1 ago. 2014.

SPEDENER, M. et al. **Microhabitat selection of free-ranging beef cattle in south-boreal forest.** Applied Animal Behaviour Science, v. 213, p. 33–39, abr. 2019.

STAUDENMAYER, J. et al. **An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer.** Journal of Applied Physiology, v. 107, n. 4, p. 1300–1307, out. 2009.

STEENEVELD, W.; HOGEVEEN, H.; OUDE LANSINK, A. G. J. M. **Economic consequences of investing in sensor systems on dairy farms.** Computers and Electronics in Agriculture, v. 119, p. 33–39, 2015.

SUGIYAMA, M. **Introduction to Statistical Machine Learning.** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2015.

TRUDELL, J.; WHITE, R. G. **The Effect of Forage Structure and Availability on Food Intake, Biting Rate, Bite Size and Daily Eating Time of Reindeer**. The Journal of Applied Ecology.

VALLETTA, J. J. et al. **Applications of machine learning in animal behaviour studies.** Animal Behaviour, v. 124, p. 203–220, 2017.

VAPNIK, V. **The Nature of Statistical Learning Theory.** Springer, 2010.

WANG, G. **Machine learning for inferring animal behavior from location and movement data.** Ecological Informatics, v. 49, n. December 2018, p. 69–76, 2019.

WANG, Q.; BOVENHUIS, H. **Validation strategy can result in an overoptimistic view of the ability of milk infrared spectra to predict methane emission of dairy cattle.** Journal of Dairy Science, v. 102, n. 7, p. 6288–6295, 1 jul. 2019.

WEBB, G. I. et al. **Leave-One-Out Cross-Validation.** In: SAMMUT, C.; WEBB, G. I. (Ed.). Encyclopedia of Machine Learning. Boston, MA: Springer US, 2011. p. 600–601.

YOSHITOSHI, R. et al. **Distinguishing cattle foraging activities using an accelerometry-based activity monitor.** Rangeland Ecology and Management, v. 66, n. 3, p. 382–386, 2013.

INDEX

APPENDIX

Table A.1 – Summary of the reviewed articles.

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Adiamandroso et al., 2017	Belgium	Dairy cows	19	Neck	3-axis accelerometer	Behavior	Sensor default prediction
Ahmed et al., 2018	Australia	Dairy cows	22	Around the vulva	GPS	Behavior	Sensor default prediction
Alsaaod et al., 2015	Switzerland	Dairy cows	20	Leg	3-axis accelerometer	Behavior	Not shown
Alsaaod et al., 2017	Switzerland	Dairy cows	17	Ear	3-axis accelerometer	Behavior	Sensor default prediction
Alvarenga et al., 2016	Brazil	Sheep	10	Jaw	3-axis accelerometer	Behavior	Decision Tree
Augustine et al., 2013	United States	Beef Cattle	9	Neck	GPS	Behavior	Sensor default prediction
Barker et al., 2018	United Kingdom	Dairy cows	19	Neck	GPS	Location	Decision Tree
Barwick et al., 2018	Australia	Sheep	5	Ear / Neck / Leg	3-axis accelerometer	Behavior	Quadratic Discriminant Analysis
Barwick et al., 2018	Australia	Sheep	1	Ear / Neck / Leg	GPS	Behavior	Quadratic Discriminant Analysis
Becciolini et al., 2018	Italy	Beef Cattle	12	Neck	GPS	Behavior	Linear Discriminant Analysis

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Betteridge et al., 2010	New Zealand	Beef Cattle / Sheep	20 / 20	Neck	GPS	Location	Sensor default prediction
Betteridge et al., 2010	New Zealand	Beef Cattle / Sheep	12 / 20	Around Vulva	Other sensors	Behavior	Sensor default prediction
Betteridge et al., 2010	New Zealand	Sheep / Beef Cattle	20 / 12	Around Vulva	GPS	Location	Sensor default prediction
Betteridge et al., 2013	New Zealand	Dairy cows	9	Around the vulva	GPS	Location	Lying Threshold Model
Bikker et al., 2016	Netherlands	Dairy cows	15	Ear	3-axis accelerometer	Behavior	Sensor default prediction
Borchers et al., 2016	United Kingdom	Dairy cows	48	Leg	3-axis accelerometer	Behavior	Sensor default prediction
Borchers et al., 2017	United States	Dairy cows	53	Leg / Neck	3-axis accelerometer	Behavior / Calving	Random Forest / Linear Discriminant Analysis / Artificial Neural Network
Champion et al., 1997	United Kingdom	Sheep / Beef Cattle	8 / 12	Leg	Mercury tilt	Behavior	Sensor default prediction
Chelotti et al., 2016	Argentina	Dairy cows	1	Halter	GPS	Behavior	Hidden Markov Models
Chelotti et al., 2018	Argentina	Dairy cows	2	Forehead	Sound record	Machine learning	Support Vector Machine /

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
							Decision Tree / Random Forest
de Campos et al., 2016	Brazil	Goats	3	Masseter muscule	Fiber Bragg Gratings	Behavior	Sensor default prediction
de Passillé et al., 2010	Canada	Dairy calves	7	Leg	3-axis accelerometer	Behavior	Spearman correlations
Debauche et al., 2013	Belgium	Beef Cattle	1	Halter	GPS	Behavior	Sensor default prediction
Decandia et al., 2016	Italy	Sheep	1	Jaw	3-axis accelerometer	Behavior	Discriminant Analysis
Decandia et al., 2018	Italy	Sheep	8	Jaw	3-axis accelerometer	Behavior	Discriminant analysis
di Virgilio et al., 2018	Argentina	Sheep	3	Neck	GPS	Location	Sensor default prediction
Draganova et al., 2010	New Zealand	Dairy cows	17	Neck / Leg	GPS	Location	Sensor default prediction
Dunne et al., 2017	Australia	Dairy cows	24	Neck	GPS	Behavior	Sensor default prediction
Dutta et al., 2015	Australia	Dairy cows	24	Neck	3-axis accelerometer	Behavior	Discriminant Analysis / K-Nearest Neighbors / Naive Bayes / Linear Discriminant Analysis
González et al., 2015	Australia	Beef Cattle	42	Neck	GPS	Behavior	Decision Tree
Greenwood et al., 2014	Australia	Beef Cattle	10	Ear / Jaw	3-axis accelerometer	Location	Sensor default prediction

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Grinter et al., 2019	United States	Sheep	19	Collar	1-axis accelerometer	Behavior	Linear Regression
Guo et al., 2009	Australia	Dairy cows	6	Neck	GPS	Location	Hidden Markov Model
Guo et al., 2018	Australia	Sheep	3	Collar	3-axis accelerometer	Behavior	Linear Discriminant Analysis
Hendriks et al., 2019	New Zealand	Dairy cows	309	Leg	3-axis accelerometer	Behavior	Sensor default prediction
Henkin et al., 2007	Israel	Beef Cattle	300	Collar	GPS	Location	Sensor default prediction
Homburger et al., 2014	Switzerland	Dairy cows	120	Collar	GPS	Behavior	Random Forest / Linear Discriminant Analysis / Support Vector Machine
Karam et al., 2015	Brazil	Beef Cattle	1	Jaw	Fiber Bragg Gratings	Behavior	Decision Tree
Klefot et al, 2016	United States	Dairy cows	4	Neck	3-axis accelerometer	Behavior	Linear Discriminant Analysis
Kuźnicka et al., 2017	Poland	Sheep	1	Neck	3-axis accelerometer	Behavior	Support Vector Machine
Lush et al., 2018	United Kingdom	Sheep	30	Ear	3-axis accelerometer	Behavior	Random Forest
Mansbridge et al., 2018	United Kingdom	Sheep	6	Ear / Collar	3-axis accelerometer	Behavior	Random Forest / Support Vector Machine / K-Nearest Neighbors

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Maroto-Molina et al., 2019	Spain	Sheep / Beef Cattle	50 / 25	Neck / Ear	GPS	Location	Sensor default prediction
Mason et al., 2013	United Kingdom	Sheep	4	Neck	3-axis accelerometer	Behavior	Sensor default prediction
Matsui et al., 1989	Japan	Beef Cattle	5	Halter	3-axis accelerometer	Behavior	Not shown
Misselbrook et al., 2016	United Kingdom	Dairy cows	12	Around the vulva	GPS	Location	Sensor default prediction
Moreau et al., 2009	Germany	Goats	2	Neck	3-axis accelerometer	Behavior	Sensor default prediction
Mulvenna et al., 2018	United Kingdom	Dairy cows / Pygmy goats/ Sheep	2	Halter	Magnet sensor	Behavior	Generalized linear mixed model
Nadimi et al., 2012	Denmark	Sheep	11	Neck	2-axis accelerometer	Behavior	Artificial neural networks
Nielsen et al., 2018	Sweden	Dairy cows	30	Leg	3-axis accelerometer	Behavior	Not shown
Nielson et al., 2013	Sweden	Dairy cows	20	Halter / Leg	2-axis accelerometer	Behavior	Sensor default prediction
Oudshoorn et al., 2012	Denmark	Beef Cattle	10	Neck	2-axis accelerometer	Behavior	Sensor default prediction
Pegorini et al., 2015	Brazil	Calf	1	Jaw	Optical fiber bragg	Intake	Decision Tree
Peng et al., 2019	Japan	Beef Cattle	6	Neck	3-axis accelerometer	Behavior	Convolutional Neural Network /

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
							Recurrent Neural Network2019
Poulopoulou et al., 2019	Italy	Beef Cattle	8	Neck	3-axis accelerometer	Behavior	Sensor default prediction
Quellet et al., 2016	Canada	Dairy cows	12	Ear	3-axis accelerometer	Behavior	Sensor default prediction
Rehman et al., 2018	Australia	Beef Cattle	1	Ear / Neck / Halter	3-axis accelerometer	Behavior	Random Forest
Reiter et al., 2018	Australia	Dairy cows	10	Ear	1-axis accelerometer	Behavior	Sensor default prediction
Reynolds et al., 2019	United States	Dairy cows	49	Ear	3-axis accelerometer	Behavior	Sensor default prediction
Riaboff et al., 2019	France	Dairy cows	10	Neck	3-axis accelerometer	Behavior	Decision Tree
Robert et al., 2009	United States	Beef Calves	15	Leg	3-axis accelerometer	Behavior	Linear mixed models
Roland et al., 2018	Germany	Dairy cows	15	Ear	3-axis accelerometer	Behavior	Hidden Markov Model
Rombach et al., 2018	Switzerland	Dairy cows	18	Halter	3-axis accelerometer	Behavior	Sensor default prediction
Rutten et al., 2017	Netherlands	Dairy cows	400	Ear	3-axis accelerometer	Behavior	Logistic Regression
Sakai et al., 2019	Japan	Goats	3	Withers	3-axis accelerometer	Behavior	K-Nearest Neighbors / Decision Trees
Schoenbaum et al., 2017	Israel	Beef Cattle	94	Neck / Leg	GPS	Behavior	Sensor default prediction

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Serrano et al., 2018	Portugal	Sheep	6	Harnesses	GPS	Location	Regression model and correlation
Shahriar et al., 2016	Australia	Dairy cows	32	Collar	3-axis accelerometer	Heat events	K-means clustering
Sheibe et al., 2006	Germany	Dairy cows	4	Collar	3-axis accelerometer	Behavior	Sensor default prediction
Smith et al., 2014	Australia	Beef Cattle	7	Halter	3-axis accelerometer	Behavior	Decision Tree
Smith et al., 2016	Australia	Dairy cows	24	Neck	GPS	Behavior	Support Vector Machine / K-Nearest Neighbors / Logistic Regression / Random Forest
Spender et al., 2019	Norway	Beef Cattle	16	Neck	2-axis accelerometer	Location	Sensor default prediction
Stephenson et al., 2017	United States	Dairy cows	11	Neck / Leg	GPS	Behavior	Sensor default prediction
Tamura et al., 2019	Japan	Dairy cows	38	Neck	3-axis accelerometer	Behavior	Decision Tree
Tani et al., 2013	Japan	Beef Cattle	4	Jaw	1-axis accelerometer	Behavior	Sensor default prediction
Thomas et al., 2011	Australia	Beef Cattle	6	Neck	GPS	Location	Sensor default prediction
Tofastrud et al., 2018	Norway	Beef Cattle	18	Neck	GPS	Behavior	Sensor default prediction
Turner et al., 2000	United States	Beef Cattle	8	Collar	GPS	Location	Sensor default prediction

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Umstatter et al., 2008	Scotland	Behavior	1	Neck	GPS	Behavior	Linear Discriminant Analysis
Ungar et al., 2005	Israel	Beef Cattle	5	Neck	GPS	Behavior	Multiple regression
Ungar et al., 2011	Israel	Beef Cattle	9	Neck / Leg	GPS	Behavior	Sensor default prediction
Ungar et al., 2018	Israel	Beef Cattle	800	Leg	3-axis accelerometer	Behavior	Sensor default prediction
Valente et al., 2013	Brazil	Beef Cattle	12	Neck	GPS	Behavior	Sensor default prediction
Vázques Diosdado et al., 2015	United Kingdom	Dairy cows	6	Neck	3-axis accelerometer	Behavior	Decision Trees
Werner et al., 2018	Ireland	Dairy cows	15	Halter	3-axis accelerometer	Behavior	Sensor default prediction
Werner et al., 2019	Ireland	Dairy cows	12	Neck	3-axis accelerometer	Behavior	Sensor default prediction
Williams et al., 2016	United Kingdom	Dairy cows	40	Neck	GPS	Behavior	Decision Tree / Naïve Bayes / Random Forest
Wolfger et al., 2017	United States	Dairy cows	15	Ear	Other sensor	Location	Not shown
Yayota et al., 2017	Japan	Goats	16	Halter	Wearable camera	Behavior	Sensor default prediction
Yoshitosh et al., 2013	Japan	Beef Cattle	4	Neck	1-axis accelerometer	Behavior	Linear Discriminant Analysis / Logistic Regression

PUBLICATION	COUNTRY	SPECIES	n	SENSOR POSITION	SENSOR UTILIZED	MAJOR PHENOTYPE	ANALYTICAL METHOD USED
Zambelis et al., 2019	Canada	Dairy cows	10	Ear	3-axis accelerometer	Behavior	Multivariate mixed models
Zehner et al., 2017	Switzerland	Dairy cows	60	Halter	Noseband pressure sensor	Behavior	Sensor default prediction
Zhang et al., 2018	China	Dairy cows	15	Neck	Sensor Network Transceiver	Location	Area coverage
Zobel et al., 2015	Canada	Goats	6	Leg	3-axis accelerometer	Behavior	Pearson correlation