Applied Animal Behaviour Science 236 (2021) 105260

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Towards on-site automatic detection of noxious events in dairy cows

Yael Salzer[a,*], Hen H. Honig[b,1], Roy Shaked[b], Elad Abeles[c], Alona Kleinjan-Elazary[b], Karin Berger[d], Shamay Jacoby[c], Barak Fishbain[d], Shai Kendler[d,2]

[a] Institute of Agricultural Engineering, Agricultural Research Organization – Volcani Center, Hamakabim Rd. 68, Rishon-Le'Zion, 7528809, Israel
[b] Animal Science, Agricultural Research Organization – Volcani Center, Hamakabim Rd. 68, Rishon-Le'Zion, 7528809, Israel
[c] Koret School of Veterinary Medicine, Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem, Yehoshua Hankin St 21, Rehovot, 76100, Israel
[d] Dept. of Environmental, Water and Agricultural Engineering, Faculty of Civil & Environmental Engineering, Technion – Israel Institute of Technology 616, Rabin Hall, Haifa, 32000, Israel

ARTICLE INFO

Keywords:
Pain assessment
Dairy cow
Machine learning
Welfare

ABSTRACT

Successful detection of pain in cows could circumscribe the therapeutic window for treatment before the cow's condition deteriorates further. While severe clinical cases that are characterized as painful have clear behavioral and physiological manifestations, mild pain may go unnoticed in cows due to their stoic nature. This work presents the first step in developing a warning system that will enable the identification of mild pain in dairy cows. In a set of three experiments, a topical application of 10 % capsaicin cream was used to elicit a noxious sensation. Experiment 1 was aimed at establishing capsaicin as a noxious model for bovines (n = 11). Each cow was treated with neutral cream on day one and the noxious cream on day two. Since the duration of capsaicin effects on bovine skin is unknown, Experiment 2 was designed to evaluate capsaicin's impact on bovines 30 min after application (n = 17). Physiological signs were collected in response to the capsaicin cream application and were compared to the application of the neutral cream. In Experiment 3 physiological signs and continuous behavioral data were recorded (n = 22, four cows participated in Experiment 1). Each cow was treated with neutral cream on day one and the noxious cream on days two and three, i.e., repeated exposure to the noxious stimulus. Heart and breathing rates were elevated soon after the noxious treatment but not for the neutral cream. Blood oxygen saturation was inconclusive. Changes in daily activity patterns consecutive to the noxious challenge included a decrease in rumination time and an increase in lying bouts. These results are in line with what would be expected for physiological and behavioral effects of pain in cows. Additional data are required to rule out habituation or sensitization to the procedure. The resulting database was then used to develop a machine-learning algorithm to detect noxious sensations by applying random forest classifiers trained with two different approaches. The learning herd approach, in which a specialist labels a set of observations and uses them to derive a classifier for new observations from the same herd, achieved 82 % ±9% accuracy. The unlearning herd approach, in which a single database is used to train a classifier that can be applied to members of other herds, resulted in an accuracy of 86 %±18 %. The data discussed in this study meet the requirements of an automatic on-site noxious detection system; real-time on-farm measurements, informative of negative high-arousal states.

1. Introduction

Timely assessment of pain is an important aspect of dairy cow welfare and a major concern for livestock veterinarians (Mench, 2018). A cow in pain may manifest a negative affective state, which in itself is a key component of animal welfare (Ede et al., 2019). Modern intensive methods of dairy cattle rearing have contributed to the increase in the incidence of morbidity (i.e., lameness, mastitis) and economic losses, including decreased milk production, lower reproduction rates, and culling (von Keyserlingk et al., 2009; Warnick et al., 2010). Because of their stoic nature, detecting pain events in cattle is difficult, and measuring their threshold is more challenging still. Thus, cows are not

* Corresponding author.
E-mail address: salzer@agri.gov.il (Y. Salzer).
[1] Animal Welfare, Veterinary Services, Ministry of Agriculture and Rural Development, Hamakabim Rd. 62, Rishon-Le'Zion, Israel.
[2] Environmental Physics Department, Israel Institute for Biological Research, 24 Lerer St. Ness Ziona 74100 Israel.

https://doi.org/10.1016/j.applanim.2021.105260
Received 24 August 2020; Received in revised form 4 February 2021; Accepted 8 February 2021
Available online 13 February 2021
0168-1591/

likely to exhibit visible signs of pain until their condition has escalated (Hudson et al., 2008), so that overall, cows are undertreated for pain (Adcock and Tucker, 2018).

There is no gold standard for pain assessment in animals. An animal cannot report its subjective experience; hence, what remains are proxy indices of cow pain or distress, based on inferential reasoning (Carstens and Moberg, 2000). The main obstacle to utilizing physiological and behavioral signs in practical farming is their lack of specificity, which involves differentiating different situations with similar manifestations. Traditional on-farm approaches for assessing pain in cattle rely on the expertise and awareness of caregivers. Veterinarians or farmers, who are highly familiar with their animals, will notice changes in their cows' normal behavior when they express pain. Accurate identification and management of pain are essential to ethically competent dairy farm care (Giovannini et al., 2017); in particular behavioral assessment facilitates pain recognition (as reviewed by Landa, 2012). Alterations in activity patterns, postural behavior, and feed intake are typical pain markers (Adcock and Tucker, 2018; Piñeiro et al., 2018). In order to improve farmers' ability to correctly identify the clinical status of a cow in a short time, novel approaches for automatic detection have been presented. These methods rely on data collected via designated on-farm, non-invasive sensors. These factors include food and water intake (Heinrich et al., 2010; Sepulveda-Varas et al., 2016), rumination time and body weight gain (Fitzpatrick et al., 2013; Steensels et al., 2016), standing and lying time, lying bouts (Van Hertem et al., 2016), weight distribution between legs (Dyer et al., 2007), and back arching (Hansen et al., 2018; Thomsen et al., 2008), as well as eye temperature (Stewart et al., 2008). Other studies have explored the potential of milk cortisol (Giovannini et al., 2017) or the plasma concentration of cortisol, haptoglobin, norepinephrine, beta-endorphin, and substance P (Bustamante et al., 2015) as biomarkers of pain.

Pain is generally associated with increased sympathetic activity affecting physiological indices such as tissue oxygen saturation, as well as heart and breathing rates (Landa, 2012; Morton and Griffiths, 1985). However, pain assessment based on these physiological factors is commonly considered inapplicable on the farm (Gleerup et al., 2015a). One of the major limitations of pain assessment research has to do with problems differentiating between the expression of the pain stimulus and the symptoms of its cause, i.e., inflammation, injury, or disease. For adequate exploration of pain per se, the stimulation should not initiate a systemic inflammatory reaction or significant damage to the flesh.

Here, in a series of three experiments, short-lasting noxious topical stimuli were used to explore the indirect symptoms of mild somatic pain. In humans, capsaicin elicits a burning sensation; it selectively activates sensory neurons that convey information about noxious stimuli to the central nervous system (Piperine et al., 1997). Gleerup and colleagues, and Di Giminiani and colleagues (Di Giminiani et al., 2014; Gleerup et al., 2015b) applied topical 10 % capsaicin to horses' and pigs' skin, respectively. Similarly, in this study, a topical application of diluted capsaicin cream, the principal pungent ingredient in hot chili peppers, was used as the stimulus (Di Giminiani et al., 2014; Gleerup et al., 2015b).

This work deals with the challenge of detecting physiological and behavioral changes associated with mild pain, which generally go unnoticed due to the stoic nature of the cow. Machine learning methods, a branch of computer analysis techniques, are designed to classify large quantities of observation-based patterns from the input data. In supervised learning, the program is trained on a set of observations from known classes, i.e., the training set. The result of this training process is a model that is then used for the classification of new observations. Two approaches to detect noxious events in cows are assessed here. The *learning herd* (LH) approach examines data from a herd of cows that are monitored and diagnosed by an expert. This generates a training set that contains observations from the entire herd. A new unlabeled observation, which was not a part of the training set, is then classified using this classifier. The *unlearning herd* (UNLH) approach is based on deriving a classification model from a set of labeled observations from a herd of cows. This classifier is then used to classify noxious events in cows from herds that were not involved in the training process. The main difference between LH and UNLH is that in LH both the training and test sets of observations are taken from the same herd. Although the LH method is expected to be accurate, since the variation between the training set and the test set is relatively small, collecting a set of labeled observations for every herd is time consuming and labor-intensive. Note that the validation set of observations is not used to derive the classifier. The UNLH approach is more streamlined since the classifier is computed once on a dedicated dataset and then is deployed to many herds. However, variations between the training set and the test set may be large and unexpected. Thus, each method has its advantages and shortcomings.

The study's protocol induced short-term discomfort, which dissipated in a short time, and was designed to temporarily impact the animal's welfare as reflected in transient, reversible behavioral changes. Overall, it was hypothesized that changes in sympathetic activity would occur in response to the noxious sensation inflicted by the application of capsaicin cream on dairy cows' skin. This would be expressed in a rise in heart and breathing rates and a drop in the oxygen saturation level. The overarching goal was the development of a technique for the early detection of cow pain, which can lead to timely care. The experiments reported here contribute to establishing a methodology that defines reliable measurement, their time window, and a machine learning algorithm that analyses these signs to assess the noxious event. This method must be robust to the naturally occurring variations in farm animals' physiological manifestations.

2. Materials and methods

This study was conducted at the Volcani Center's Experimental Dairy Farm in Beit Dagan, Israel. The Volcani Center Animal Care Committee (approval numbers IL 774 /18 and IL 820/19) approved the protocols. Three experiments were conducted. Experiment 1 was aimed at establishing capsaicin as a noxious model for bovines. It was hypothesized that topical capsaicin would elicit a noxious sensation on dairy cow skin, which would be reflected in immediate increased sympathetic activity; i.e., a rise in heart and breathing rates, and a drop in oxygen saturation level, among other physiological responses. Since there are no studies on the duration of capsaicin effects on bovine skin, Experiment 2 was designed to evaluate capsaicin's impact on bovines 30 min after application. Physiological signs were collected in response to the application of the capsaicin vs. a neutral cream. Experiment 3 consisted of recordings of the immediate physiological responses and continuous behavioral data (number of steps, lying time, etc.) to the first as well as to repeated exposure to the noxious stimulus. The dataset acquired in these experiments was then used to provide a proof-of-concept that the automatic algorithm using these data could distinguish a healthy cow experiencing the noxious event from the neutral control condition. Throughout the experiments, blood samples were drawn. The blood analyses are beyond the scope of this article.

2.1. Experiment 1

2.1.1. Animals, housing and management

The experiment was conducted in June 2018. Eleven healthy multiparous Israeli Holstein dairy cows that were not candidates for insemination with no clinical signs of any sort 14 days prior to the experiment were tested. The cows were housed together in a covered loose-housing pen with an adjacent continuous fenced outdoor yard. The cows could move freely between the two at all times. The feeding stations are aligned along the side of the open barn. Each cow was familiar with her individual station, which was controlled by an automatic RF recognition receiver. The cows are accustomed to being secured to their stations for up to an hour after milking during feeding. The cows' routine was maintained throughout the experiments. Milking took place three times a day (at 05:00 h, 13:00 h, and 20:00 h). DIM (days in milk) was not controlled for but averaged 204 DIM (range 112–302) for 671 kg of BW (body weight), with standard deviations of 67 and 85, respectively. The cows were fed a commercial TMR (total mix ratio) composed of 1.78 mega calories of NEL (net energy for lactation), 16.5 % CP (crude protein), and 31.7 % crude NDF (neutral detergent fiber), and ad libitum access to water.

2.1.2. Preparation

A day before each individual cow's test, a patch on the cow's left side of the rump was tonsured. Core temperature was monitored to exclude the progression of inflammation or background disease. A bottom-type temperature logger (sampling rate 0.1 min-1, SL52T-A, Signatrol Ltd) attached to a plastic seeder holder with no hormones was inserted into place (Burdick et al., 2012; de Oliveira et al., 2019). The vaginal temperature was measured continuously until the device was removed, a day after the final noxious treatment was administered.

2.1.3. Stimuli

A noxious stimulus (i.e., noxious treatment) was induced with 10 % capsaicin (Affix Scientific©) mixed with a universally used ambiphilic base cream from Deutscher Arzneimittel-Codex, DAC (Vetmarket©, Israel). A measured ~5 g teaspoon was applied on the tonsured patch of the cow's left rump. For control, the base cream was applied in the same manner (i.e., neutral treatment).

2.1.4. Design

Treatment was considered a fixed within-subject factor (noxious, control), and cow as a random factor. The cows were randomly divided into three time groups: after the morning (n = 4), noon (n = 4), and evening (n = 3) milking. On the experimental days (noxious treatment day and control treatment day), the cow's daily routine was interrupted for an hour after milking. The time of treatment was kept constant for each individual cow.

Since the research team's presence during the experimental days was interpreted as a threat to the cows, regardless of the treatment, it was expected that the cows' state of arousal and stress level would be affected. Specifically, when a threat is encountered, the state of arousal increases, and the cow's focus is oriented away from the stimulation, thus suppressing the experience of pain (Adcock and Tucker, 2018). On the other hand, a painful event may reverberate, leading to increased arousal and a negative affective state (Ede et al., 2019). To overcome this problem, the order of treatments was kept fixed, and each individual cow served as its own control. Specifically, the cows were treated with neutral cream on day one and the noxious cream on day two. This design allowed the cows to become familiarized with the research team and the procedure during the neutral treatment before forming conditional episodic memory linked with the noxious stimulus.

2.1.5. Experimental procedure

After milking, all the cows were led back to the cowshed. The cows were then guided to their individual feeding stations. The participating cows were kept secured in their stations and thus stood in one place throughout the experimental procedure for up to an hour. The experimental procedure (see Fig. 1) began with base-level measures that included five minutes of continuous measurements of heart and breathing rates, and blood oxygen saturation (SpO_2) (i.e., physiological metrics). The pulse rate and SpO2 were monitored with a pulse oximeter device attached to the lips of the cow's vulva (1 Hz sample rate, s500 Handheld Pulse Oximeter, SINNOR Instruments, Inc.) (Grubb and Anderson, 2017; Peter and Peter, 2002). Breathing rates were measured by counting flank movements for 15 s, every minute, for five minutes (Stewart et al., 2008). A blood sample was then collected by venipuncture from the tail vessels. Immediately afterward, the treatment was applied. A measured teaspoon containing ~5 g of cream was evenly spread within the boundaries of a 10 × 10 cm stencil on the tonsured patch (see Fig. 1). Immediately after the treatment, for the next five minutes, the physiological metrics were measured for the second time (Treister et al., 2012). A second blood sample was retaken 30 min after treatment. Then, without delay, a vegetable-based oil was applied to the tonsured patch to remove the cream (Gleerup et al., 2015b).

2.2. Experiment 2

Since there are no data on the duration of the effect of capsaicin on bovine skin, Experiment 2 followed a similar protocol as Experiment 1 with one difference. Specifically, a second physiological measurement, i.e., after the treatment, was obtained 30-minutes after the cream was

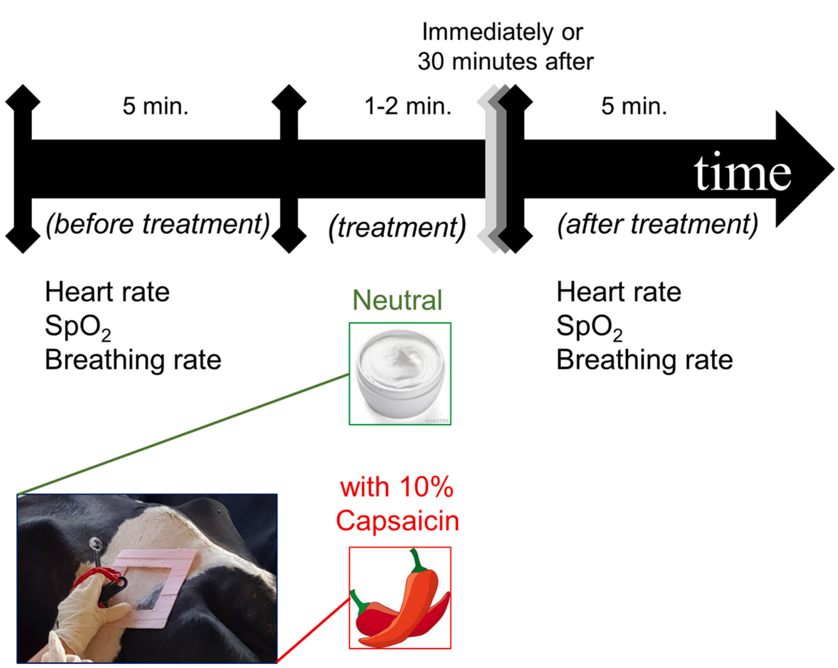

Fig. 1. Timeline of a single experimental session. Cream was applied on a tonsured patch on the left or right side of the cow's rump, using a 10 cm × 10 cm stencil.

applied.

2.2.1. Animals, housing and management

This experiment was conducted in June 2018. Seventeen healthy multiparous Israeli Holstein dairy cows that did not participate in Experiment 1, were not candidates for insemination and had no clinical signs of any sort 14 days prior to the experiment, were included in the study. They averaged 185 DIM (range 104–285) for 654 kg of BW, with standard deviations of 60 and 68 respectively.

2.2.2. Preparation, stimuli, and design

The preparation, stimuli, and design were identical to Experiment 1. The cows were randomly divided into three time groups, after the morning (n = 7), noon (n = 4), and evening (n = 4) milking. Two other cows were treated with the neutral treatment after the noon milking and noxious treatment after the evening milking, on the same day (n = 2).

2.2.3. Experimental procedure

The experimental protocol was the same as in Experiment 1 with one modification. The second time physiological metrics were measured (see Fig. 1), i.e., after the treatment, was 30 min after the cream was applied.

2.3. Experiment 3

The assessment of the effects of a noxious event on general physiological measurements as conducted in Experiments 1 and 2 was a prerequisite to the development of a machine-learning-based system. A real-time on-farm pain assessment tool should be able to detect the first as well as repeating occurrences of mild pain. However, it remained unclear whether the physiological response to a recurring noxious event would be the same as the first event. To this end, Experiment 3 examined the physiological response to the first as well as to repeated exposure to the noxious stimulus in an expanded dataset. Each cow was exposed once to the neutral treatment and twice to the noxious treatment, which was applied once on the left and once on the right side of the cow's rump. The three exposures took place on three different days.

The cows' daily activity patterns, lying time, lying bouts, number of steps, and rumination are indicators of their comfort and welfare (Piñeiro et al., 2018). As such, they can reinforce the physiological findings. Although the topical acute noxious stimulus was transient, the impact on the animal's welfare might be reflected in behavioral changes hours after the challenge was completed, when the animal was free to move around between the loose-housing pen and outdoor yards. Thus, behavioral data were also collected from the farm's domestic system database.

2.3.1. Animals, housing and management

The experiment took place in March 2019. Four cows that participated in Experiment 1 and 18 other healthy multiparous Israeli Holstein dairy cows that were not candidates for insemination and had no clinical signs of any sort 14 days prior to the experiment were included in the study. They averaged 83 DIM (range 61–103) for 652 kg of BW, with standard deviations of 13 and 59, respectively.

2.3.2. Preparation and stimuli

The preparation and stimuli were the same as in Experiment 1. Because the noxious treatment was applied to both the left and the right sides of the cow's rump, patches were tonsured on both sites on the rump.

2.3.3. Design

Each individual cow served as its own control. All the cows received the neutral treatment on day one, the noxious treatment on day two, and the second noxious treatment on day three, which was applied contralaterally to the treatment's location on the previous day. Experiment 3 took place only in the mornings, after the 05:00 a.m. milking.

2.3.4. Experimental procedure

The experimental procedure was identical to Experiment 1. Behavioral data were collected from the farm's domestic systems. Leg sensor tags (Pedometer Plus; S.A.E. Afikim) placed on the cows' metatarsus accumulated the step count, lying bouts (number of recurring switches in position from lying to standing), and lying time (time in mins. spent lying down) every 15 min. The data were then entered into the farm's management software (AfiFarm; S.A.E. Afikim). Rumination time (minutes) was collected every two hours (HR-Tags, SCR Engineers Ltd., Hadarim, Netanya, Israel; Bar and Soloman, 2010). For the analysis, the behavioral counts were summed from the time the cow was released from the station, i.e., at the end of the experimental procedure, and for the next four hours.

2.4. Statistical analyses

For each cow, in each treatment condition (neutral, noxious), the arithmetic means for heart rate, reparation rate, and SpO_2 were calculated before treatment (i.e., base-level measures) and immediately after treatment. The arithmetic means for one hour of temperature recordings before treatment and one hour after treatment were calculated. The individual difference (ID) was obtained for each cow on each treatment by subtracting the before-treatment mean from the after-treatment mean. Heart rate, reparation rate, SpO_2, and temperature IDs were then subjected to two-tailed paired sample t-tests with treatment (neutral, noxious) as a fixed within-subject factor in Experiments 1 and 2, and a one-way ANOVA with treatment (neutral, noxious, noxious repeat) as the fixed within-subject factor in Experiment 3, and the cow as a random factor. For each cow in Experiment 3, rumination, the number of steps, lying time, and lying bouts accumulated over four hours from the offset of the experimental session were subjected to a one-way ANOVA with treatment (neutral, noxious, noxious repeat) as the within-subject factor, and the cow as a random factor. Since the experimental design was within-subject, Mauchly's (1940) test for sphericity was conducted. The Greenhouse-Geisser correction was applied when sphericity assumptions were not met. See Fig. 2 and Table 1 for complete statistics.

3. Results

3.1. Experiment 1

The cows' standing movements (tail flicks, thumps, pelvic movements) led to occasional signal failures. Of the SpO_2 recordings, 9.1 % were lower than 85 %, which was not clinically viable given the cows' clinical condition, and thus were omitted from the dataset. The results indicated an immediate increase in heart rate of 16.4 beats per minute [BPM] after the noxious treatment but not after the neutral treatment (mean ID = 1.16 BPM with $t(10) = -4.28$, $p < 0.001$). Similarly, the breathing rate increased by an average of 7.38 breaths/min immediately after the noxious treatment, but not after the neutral treatment (mean ID = 1.36 breaths/min with $t(10) = -2.58$, $p < 0.02$). The cows' temperature rose after the neutral treatment by a mean ID of 0.19 C, but not after the noxious treatment ($t(10) = -2.93$, $p < 0.014$). Treatment had no significant effect on SpO_2 ($p = ns$).

3.2. Experiment 2

One cow (no. 3650) moved restlessly following the noxious treatment; breathing rate readings were unavailable. Four cows' heart rate and SpO_2 data files were lost (no. 3637, no. 3690, no. 3701, no. 3730), and vaginal loggers were misplaced for four cows (no. 3626, no. 3637, no. 3701, no. 3603). Of the SpO_2 recordings, 5 % were below 85 % and thus were omitted from the dataset. The results showed that 30 min after the noxious treatment, heart rate was still elevated by an average of 6.72 BPM but was not significantly different from the smaller rise of 1.49

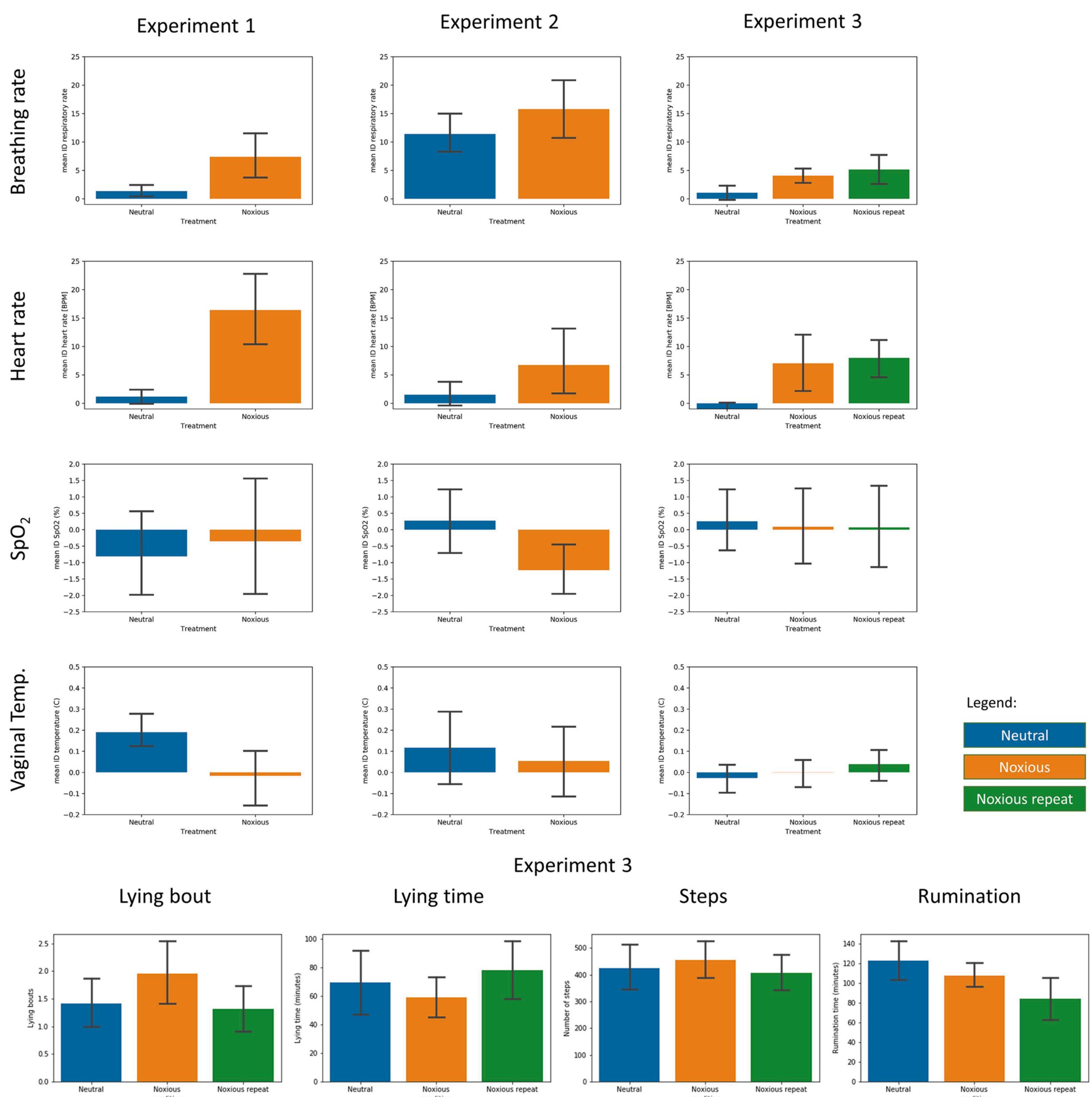

Fig. 2. Experiments 1, 2, 3: Mean individual differences (ID) in heart rate, breathing rate, SpO2, and temperature by treatment (neutral, noxious and noxious-repeat). Experiment 3: Rumination time (minutes), no. of steps, no. of lying bouts and lying time (minutes) summed four hours after treatment. Error bars denote confidences intervals (CI).

BPM after the neutral treatment ($t(12) = -2.0, p = 0.068$). A similar rise in breathing rate was observed 30 min after both treatments (neutral: mean ID = 11.37 breaths/min; noxious: mean ID = 15.79 breaths/min with $t(15) = -1.61, p = 0.12$). SpO_2 decreased 30 min after the noxious treatment (mean ID = -1.23 %) but not in the neutral treatment (mean ID = 0.27 % with $t(12) = 2.20, p < 0.04$). Treatment had no significant effect on temperature ($p = ns$).

Overall, in Experiments 1 and 2, there was an immediate rise in heart and breathing rates after the cows were exposed to the short-lasting noxious topical stimulus, whereas the decrease in SpO_2 was only observed 30 min after the challenge. The heart rate trended down to baseline 30 min after the noxious stimulus was applied. Thirty minutes after the neutral and noxious challenges, the breathing rate increased, regardless of treatment. As expected, SpO_2 decreased after the noxious treatment, but only after a lapse of 30 min.

3.3. Experiment 3

Heart rate and SpO_2 noxious (cows no. 3578, no. 3798), and a noxious-repeat treatment data file was lost (cow no. 3388). Vaginal loggers were misplaced during noxious treatment (cows no. 3683, no. 3821) and noxious-repeat (no. 3683, no. 3821, no. 3597, no. 3689, no. 3814) treatments. Of the SpO_2 recordings, 9.3 % (<85 %) were omitted from the dataset. Sphericity assumptions were not met for heart rate and breathing rate (Mauchly, 1940), for which the Greenhouse and Geisser (1959) correction was applied. Heart rate increased by 7.02 BPM after the noxious treatment, and 7.99 BPM after the noxious repeat treatment, but not after the neutral treatment (mean ID = −1.48 BPM with $F(2, 36) = 9.11, p < 0.002$). Similarly, the breathing rate increased by an average of 4.07 breaths/min after the noxious treatment, and 5.16 breaths/min after the noxious repeat, but only by 1.07 breaths/min after the neutral treatment ($F(2,42)$=5.57, $p < 0.015$). Simple comparisons revealed that the heart and breathing rate differences between the noxious and

Table 1
Experiments 1, 2, 3: Statistical analysis (t-test or ANOVA) of mean individual differences (ID) in heart rate, breathing rate, SpO_2, and temperature. Experiment 3: ANOVA of rumination time (minutes), no. of steps, no. of lying bouts and lying time (minutes) summed four hours after treatment.

	Experiment	*df*	t/F-test	*p*
Heart rate [BPM]	Exp. 1	10	t = −4.28	0.001
	Exp. 2	12	t = −2.00	0.068
	Exp. 3	2, 36	F = 9.11	0.002
Breathing rate **(per min.)**	Exp. 1	10	t = −2.58	0.02
	Exp. 2	15	t = −1.61	0.12
	Exp. 3	2, 42	F = 5.57	0.015
SpO2 (%)	Exp. 1	10	t = −0.41	0.68
	Exp. 2	12	t = −2.20	0.04
	Exp. 3	2, 39	F = 0.05	0.95
Temperature (C)	Exp. 1	10	t = −2.93	0.014
	Exp. 2	12	t = −0.50	0.62
	Exp. 3	232	F = 2.55	0.09
Rumination (min.)		2, 42	F = 5.347	**0.008**
Steps	Exp. 3	2, 42	F = 0.692	0.5
Lying time (min.)			F = 2.118	0.14
Lying bouts		2, 42	F = 3.33	**0.045**

noxious repeat treatments were non-significant (p = *n.s.* for both), thus suggesting that the noxious and noxious repeat treatments could not be differentiated based on these measurements. The effect of treatment on SpO_2 and temperature did not reach significance ($F(2, 39) < 1$, and $F(2, 32) = 2.55$, $p = 0.09$, respectively). Thus, Experiment 3 successfully replicated the results of Experiment 1, suggesting that both heart and breathing rates increased immediately after the noxious stimulation and the repetition of the noxious stimulation. Sphericity assumptions were not met for lying time, for which the Greenhouse-Geisser correction was applied. A significant difference was found for rumination time, with a mean rumination time of 123 min after the neutral treatment, 108 min after noxious treatment, and 84 min after noxious repeat treatment ($F(2, 42) = 5.347$, $p < 0.008$). More lying bouts were observed after the noxious treatment (mean = 1.95) than the neutral (mean = 1.4) or noxious repeat (mean = 1.31) treatments ($F(2,42) = 3.33$, $p = 0.045$). The effect of treatment on the step count and lying time did not reach significance ($F(2, 42) < 1$, and $F(2,42) = 2.11$, $p = 0.14$, respectively).

4. Supervised machine-learning classification of Experiments 1–3 dataset

Machine learning is a well-established branch of computer analysis techniques used to classify large numbers of observation-based patterns in the input data. Supervised machines (i.e., algorithms) are trained on a set of observations belonging to known classes, i.e., the training set. The output of this training process is a classifier. A key issue in classification is the right choice of the training set, which should include phenomena that are expected to be found in the validation set of unlabeled observations. Random forest is a supervised multi-class classifier based on a collection of decision trees (Breiman, 2001; Breiman et al., 1984). The random forest classifier (RFC) uses voting between an ensemble of decision trees (hence "random forest"). Ensemble machine learning techniques consist of a combination of classifiers {*f*}. Once the set {*f*} is trained, each new data point is assigned a label with respect to all the outputs obtained from the set {*f*} by taking a vote of their predictions. Potentially, given their diversity, ensembles may be more accurate than an individual classifier (Hansen and Salamon, 1990) with a lower likelihood of overfitting (Belgiu and Drăgu, 2016; Guan et al., 2013).

During the RFC's training stage, a random subset of the training set observations is chosen and is left unused (out-of-bag observations). The out-of-bag-observations are then employed during the derivation of the classifier to assess its performance and the effect of parameters, such as the number of trees, on the classification error. Similarly, each feature's contribution (e.g., its importance) to classification accuracy can be evaluated (Friedman, 2001; Richter et al., 2020). This can simplify the measurement system if it is set to solely measure the system's features deemed important. The features that were found to be potentially important can shed light on the phenomena under investigation (Huynh-Thu et al., 2012).

4.1. Prior assumptions

The results of Experiments 1, 2, and 3 suggested that the effects of the topical noxious stimulation on the sympathetic system were short-lasting. This assumption was tested by omitting observations from Experiment 2 during the classification and comparing the obtained classification accuracy to the case where the results from all three experiments were used. The first and repeated capsaicin applications resulted in an indistinguishable short-lived effect; hence, the noxious and repeated noxious events were pooled together in the classification process.

4.2. Data pre-processing

Similar to the experimental procedures described above, the observations, O, acquired in the experiments, were used to develop the classification process. Let $o_c{}^j(t) \in O$ be a specific physical measurement, j, acquired from cow c at time t. Then, $o_c{}^j(t=0) \in O$ is the measurement taken at t = 0; i.e., before treatment, and $o_c{}^j(t>0) \in O$ is the value of that feature after treatment. To focus on the changes putatively caused by the treatment, $o_c{}^j(t) \in O$ was modified, as follows (1):

$$\widehat{o_c{}^j}(t) = \frac{o_c{}^j(t) - o_c{}^j(t=0)}{o_c{}^j(t=0)} \tag{1}$$

The resulting normalized observation set, $\widehat{o_c{}^j}$, was then used for the classification. Each feature, j, in the reduced dataset corresponded to a change from the baseline value of that feature normalized to the baseline value. The results reported here were obtained using the data available in all the experiments. These data were the median value and standard deviation of the breathing rate, the median and standard deviation of the pulse rate, and the median of SpO_2; correlations were equal to or less than zero and were not correlated. The resulting database contained 86 observations for 29 individual cows; 11 cows participated in Experiment 1, four of which, along with additional 18 other cows, participated in Experiment 3.

4.3. Classification accuracy assessment

Distinguishing between a healthy cow experiencing a noxious event from the neutral control condition is defined as classification, where the two classes are neutral and noxious. The experimental protocol was designed to inflict uniform noxious sensation. The classification was performed using an algorithm that labeled an event according to its class. Several metrics were used to assess classification performance. A convenient way to examine the classifier's performance is to examine a confusion matrix in which the diagonal elements are the number (or percentage) of accurate classifications (TP and TN), and the off-diagonal elements are FP and FN. The following numerical criteria can be deduced from the confusion matrix. The first is the positive predictive value (PPV), which is the ratio of the number of observations that were correctly classified as positive (true positive, TP) to all the positively classified observations, the sum of the TP and the false positives (FP), as follows (2):

$$PPV = \frac{TP}{TP + FP} \tag{2}$$

Similarly, the negative predictive value (NPV) is expressed as (3):

$$NPV = \frac{TN}{TN + FN} \tag{3}$$

where the TN is the number of true negative classifications, and FN is the number of false-negative classifications. The total accuracy is given in (4):

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN} \tag{4}$$

The variability of the cows' response to the noxious stimulus was high, making the machine learning method sensitive to selecting the training set since each selection could result in different classification performance. Hence, 4000 training and classification cycles were conducted; in each cycle, the training set was randomly selected from the set of observations, $\left\{ \widehat{o_c{}^j}(t) \right\}$. The mean classification performance, standard deviations, and the distribution of PPV, NPV, and accuracy were calculated. The classification performance of the RFC was compared to nine other classifiers, including support vector machine (SVM) classifiers with different kernel functions (Suykens and Vandewalle, 1999), decision trees (Steensels et al., 2016), linear and quadratic discriminant (Singh et al., 2004), KNN (Hindman, 2015; Islam et al., 2008) and Naïve Bayes (Islam et al., 2008).

4.4. Learning herd and unlearning herd classification approaches

The random forest classification tested two different methodologies: the learning herd and the unlearning herd. In the *learning herd* (LH) classification approach, observations from a specific herd were collected and labeled by an expert. The resulting dataset was used to train a RFC. A new, unlabeled observation collected from a cow that was a member of the same herd was then classified using the obtained RFC.

The *unlearning herd* (UNLH) classification model was trained on a set of labeled observations from one or more herds. The classifier was then used to classify noxious events from cows that were members of another herd that was not involved in the training process. In the current stage of this project, it was impractical to perform the experiments on several farms; therefore, this approach was simulated by using all the observations from one specific cow as the test set. The remainder of the observations were used as the training set that excluded any observations of the test cow. The UNLH method is very advantageous from a commercial point of view; once a database of observations obtained from a specific herd has been acquired, the RFC can be calculated in advance and then deployed as an operating system in any number of herds or farms. However, variations in the cows' response to noxious events may affect accuracy. Here, the effects of specific responses were explored by repeating the UNLH process several times. Each time, all of the observations of one cow were used as the test set, and the rest of the observations were used as the training set. Comparing the accuracies obtained from different cows was considered an indication of the extent of variation in the cows' responses to noxious events.

4.5. Classification results

4.5.1. The learning herd classification approach

Classification performance was explored by repeating the computation more than 4000 times. In each repetition, a new training set was randomly selected, and the mean classification performance (NPV, PPV, and accuracy) was computed using the database from Experiments 1 and 3 that employed an identical protocol. Fig. 3 presents the distribution of the performance, the means, and the standard deviations for the classification performance, with $NPV = 0.84 \pm 0.14$, $PPV = 0.79 \pm 0.16$ and $accuracy = 0.82 \pm 0.09$. Next, the performance was calculated again,

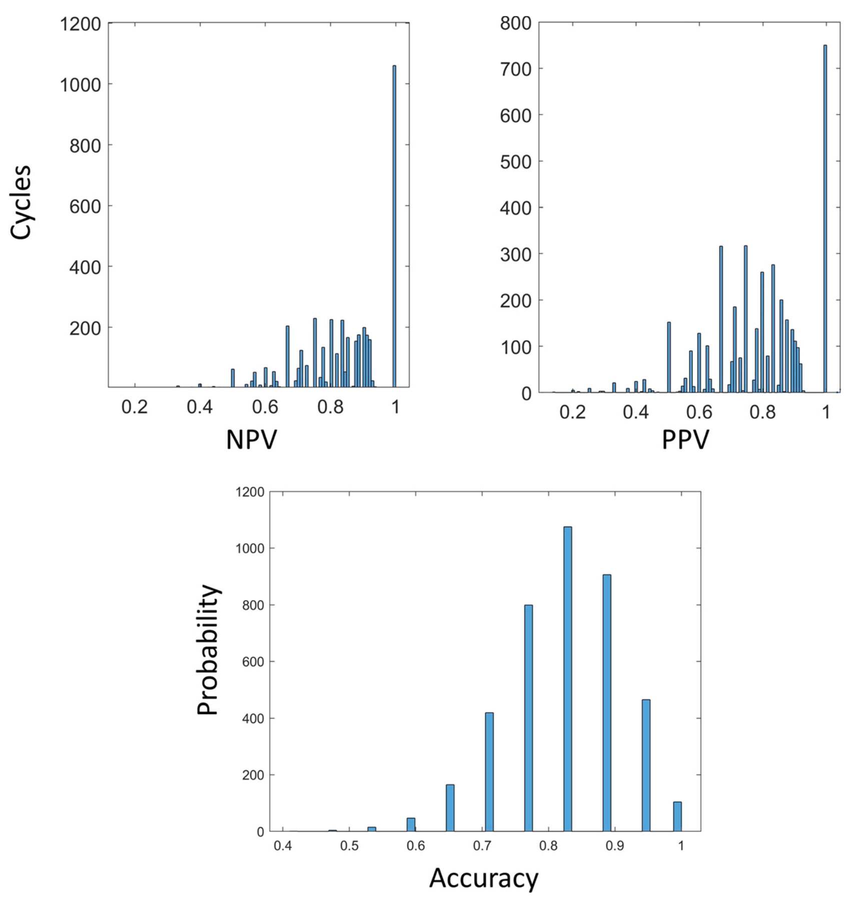

Fig. 3. Distribution of classification performance metrics for the LH approach. Top: NPV, PPV. Bottom: Accuracy. Four thousand cycles were run using the database obtained in Experiments 1 and 3. 150 trees were used to derive the classifier.

including the dataset from Experiment 2, in which post-treatment measurements were delayed by 30 min, resulting in $NPV = 0.79 \pm 0.12$, $PPV = 0.75 \pm 0.13$ and $accuracy = 0.77 \pm 0.08$. The drop in performance obtained when including Experiment 2 can be attributed to the smaller observed changes in the cows' physiological markers in Experiment 2 as compared to Experiments 1 and 3 (see Fig. 2 and Table 1). These findings suggest that the noxious effects elicited by capsaicin cream are short-lived.

The classification performance of the RFC was compared to support vector machine (SVM) classifiers with different kernel functions, decision trees, linear and quadratic discriminant, K-nearest neighbor (KNN), and Naïve Bayes. The findings, which indicated that the RFC was 8 %–20 % more accurate than all the other classifiers (see Table 1 in the supplementary material), are consistent with Fodeh et al. (2019), who successfully implemented RFC for pain classification in humans.

4.5.2. The unlearning herd classification approach

The unlearning herd (UNLH) method was tested using $\left\{\widehat{o_c{}^{j}}\right\}$, that contained observations from 29 cows. Accordingly, the method was run 29 times; in each run, all of the observations from a specific cow, $\mathbb{C}$, (i.e., the test cow) were pulled out of $\left\{\widehat{o_c{}^{j}}\right\}$, resulting in $\left\{\left\{\widehat{o_c{}^{j}}\right\}\Big\backslash\left\{\widehat{o^{j}_{c=\mathbb{C}}}\right\}\right\}$. The RFC, which was computed with the set $\left\{\left\{\widehat{o_c{}^{j}}\right\}\Big\backslash\left\{\widehat{o_{c=\mathbb{C}}{}^{j}}\right\}\right\}$, was applied to classify the observations of the test cow, $\left\{\widehat{o_{c=\mathbb{C}}{}^{j}}\right\}$. To verify that the result was unaffected by the random processes involved in the derivation of the RFC (for example the random selection of the out-of-bag observations), 100 classification models were calculated for each $\left\{\widehat{o_{c=\mathbb{C}}{}^{j}}\right\}$.

The total unlearning herd accuracy was 86 %±18 %. Fig. 4 presents the resulting confusion matrices obtained using this method for each cow. It shows that the classification accuracy varied across cows. For example, the classification accuracy of cows 3743 and 3798 was 1, whereas the classifications for cows 3808 and 3821 were inaccurate. This suggests that each cow may have had a specific sensitivity to capsaicin, which resulted in different manifestations of physiological signs after being treated with the cream. The specific response to capsaicin may potentially indicate that, in general, cows may react in an idiosyncratic fashion to noxious stimuli. These variations may be challenging for classification, especially in real cases in which classifier training is based on different herds in different conditions. The implications are discussed below.

5. Discussion

Valid pain assessment is a prerequisite for ensuring the proper health and welfare of dairy cows. To contribute to this end, this study had two main goals. The first was to confirm that physiological parameters are informative of a noxious event and can be measured on-site on the farm. The second was to provide proof-of-concept that an automatic algorithm can differentiate healthy cows suffering from mild somatic pain from those that are not. For this purpose, a short-lasting noxious or a neutral topical stimulus was applied on consecutive days to explore the physiological and indirect behavioral indices of mild somatic pain. Heart and breathing rates were elevated soon after the noxious treatment but not for the neutral cream. SpO_2 and vaginal temperature were inconclusive. The behavioral metrics demonstrated changes in daily activity patterns consecutive to the noxious challenge, including a decrease in rumination time and an increase in lying bouts, suggesting that the cows experienced stress or discomfort beyond the duration of the noxious manipulation (Leslie and Petersson-Wolfe, 2012; Piñeiro et al., 2018; Siivonen et al., 2011).

Although the precise distinction between pain and distress, which undermine the animal's welfare, is difficult to establish, an effort

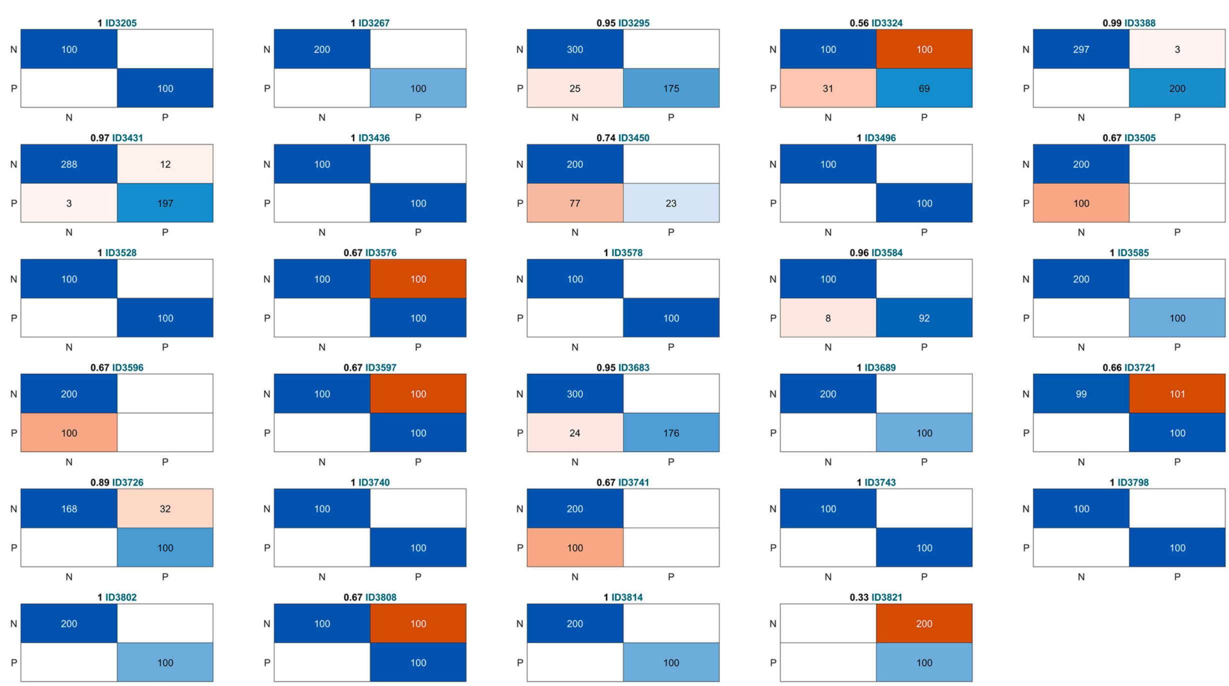

Fig. 4. The confusion matrices of 29 cows using the UNLH approach, f = 0.8, 100 cycles for each cow. Each confusion matrix is associated with an individual cow (ID, in teal). For each cow, the confusion matrix accuracy (in black) were obtained by excluding all the observations of an individual cow from $\widehat{o_c^j}$. The classifier was then computed using the remaining observations, excluding the individual cow. Classifier performance was tested on the observations for the individual cow. Y-axis denotes the true labels, and the X-axis indicates the predicted labels (N = neutral treatment, P = noxious treatment) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

towards their detection, integrated with multiple inputs, may expedite disease detection (Maltz, 2020). If correctly identified, the pain and distress associated with a range of clinical states could facilitate clinical treatment and improve animal welfare in dairy cows. Pain assessment in animals is indirect in that it relies on the expert interpretation of behavioral, physiological, and clinical responses to pain. Signs of mild pain are easily masked, making them more difficult to identify than acute management events.

Morton, Griffiths, and others (Leslie and Petersson-Wolfe, 2012; Morton and Griffiths, 1985) have argued that clinical metrics such as heart rate, pulse quality, peripheral circulation, and temperature should be included in efforts to improve early animal pain assessment. Lay and colleagues (Lay et al., 1992) reported a significant increase in heart rate after a freeze and hot iron banding. Increased heart rate and breathing rate were measured up to 48 h and 6 h, respectively, in response to induced oligofructose induced-lameness in dairy heifers (Bustamante et al., 2015). Kemp and colleagues (Kemp et al., 2008) associated the severity of mastitis in dairy cows with an increase in rectal temperatures and heart and breathing rates. In this work, there was a rise in heart and breathing rates soon after the cows were exposed to the short-lasting noxious topical stimulus. Heart rate slowed down 30 min after the noxious stimulus was applied. The breathing rate further increased 30 min after treatment in both the noxious and neutral conditions (Experiment 2). The delayed rise in the neutral breathing rate may be due to the cows' cumulative restlessness, since they were kept back at their feeding stations, and were waiting to be released to their group. In similar studies, Di Giminiani and colleagues applied capsaicin on a 4 cm^2 area on the pig's flank region, whereas Gleerup and colleagues applied capsaicin on a 100 cm^2 area on the horse's hind limb or shoulder. Gleerup and colleagues reported no differences in heart and breathing rates before and 20 min after the noxious stimulus was applied. However, they noted that soon after the noxious stimulus initiation, the heart rate increased for a few minutes before returning to baseline. The findings here are consistent with Gleerup and colleagues' results and underscore the importance of heart and breathing rates as timely clinical metrics of transient topical pain. Moreover, significant behavioral changes were observed in horses but not in pigs. Perception of cutaneous sensation depends greatly on stimulus size, also known as spatial summation, which defines the topographical variation of the receptors in the area (Green and Zaharchuk, 2001). It is likely that the summation area in the Gleerup's study, which was 25 times larger than Di Giminaiani's, played a crucial role in the robustness of the stimuli.

Pulse oximetry is an essential monitoring tool in human and animal medicine and anesthesia; however, it is rarely implemented in farm animals (Kanz et al., 2018). It was hypothesized that, similar to humans (Worley et al., 2012), oxygen saturation would decrease after noxious stimulation. However, the findings failed to support this hypothesis. Carrying biotelemetry sensors can have negative impacts on the individual animal wearing the tag (Paci et al., 2019). It is possible that the application of the pulse-oximeter device, which was attached to the lips of the cow's vulva, produced a concurrent noxious stimulation, which diminished the perceived intensity of the topical capsaicin applied to the cow's rump (Adcock and Tucker, 2018). One way of overcoming these limitations would be to use a dedicated animal-center designed medical device, in which factors such as shape, materials, and position are considered to minimize the distress and discomfort of the animal (Paci et al., 2019).

The metrics for the cardiovascular and respiratory systems have limitations because they may be indicative not only of pain but of another negative (e.g., fear-related) or positive (e.g., sexual stimulation) state of high arousal (Ede et al., 2019). Ede and colleagues criticized the use of physiological and behavioral signs and argued they might be valuable for estimating whether the cow's arousal is high or low but are of less value in concluding as to the event's valence. Here, this possible shortcoming was avoided by ensuring that the only difference between treatments was the ointment content.

The Volcani Center's Dairy Farm is an experimental farm. Cows are familiar with their feeding stations and are accustomed to being secured with headgate for up to an hour after milking during feeding. However, two aspects of the study protocol should be noted. First, the time the cows were secured may have prevented them from moving freely during the treatment. Therefore, the behavioral measures were only obtained after the cows were released back with their group. Second, although the cows are accustomed to the headgate, it is reasonable that the whole procedure was stressful and directly affected the physiological measurements. These conditions induced a similar baseline level of discomfort in both the neutral and noxious conditions. Hence, the significant differences in the physiological parameters are likely to be the capsaicin challenge's outcome.

The treatment sequence was fixed, such that the neutral treatment was administered on day one and the noxious cream on day two. The treatment sequence allowed the cows to become familiarized with the research team and procedure during the neutral treatment before forming conditional episodic memory linked with the noxious stimulus. Similar reasoning prompted Gleerup and colleagues (Gleerup et al., 2015b) to subject all their horses to the same treatment sequence, i.e., the first two days constituted the control treatment, which was followed by four days of noxious treatment. However, this design's weakness is that the order effect could not be completely ruled out in accounting for the observed differences. However, since the results here were in line with the expected effects of noxious treatment, it is less likely that the cows habituated to the testing procedure itself.

Changes in the animal's behavior may indicate an attempt on the part of the animal to avoid an unpleasant experience, protect parts or all of its body, minimize pain, and favor healing (Aubert, 1999; Molony and Kent, 1997). For instance, cows infected with clinical metritis and lameness tended to be less active and spend more time lying than non-infected cows (Piñeiro et al., 2018). Conversely, mastitis infected cows stood longer (Cyples et al., 2012; Siivonen et al., 2011), exhibited increased postural changes to minimize the discomfort experienced when lying down (de Boyer des Roches et al., 2017) and decreased food intake (Sepulveda-Varas et al., 2016). A decrease in ruminal cycles was also recorded following oligofructose induced-lameness (Bustamante et al., 2015). These behavioral and digestive patterns are pain indicators in the study of visceral and deep pain. In the present study, to confirm that the somatic noxious stimulus was transient, the cows' daily activity pattern after the experimental treatment was monitored and examined. In line with expectations, the noxious topical stimuli did not appear to affect lying time or the number of steps; however, lying bouts increased after the first exposure to the noxious stimulus, but not the second.

Interestingly, rumination time decreased after the first noxious treatment and even more after the second treatment. It is possible that the acute phase of the topical treatment diminished in minutes but that the general discomfort lasted beyond the experimental session. Future work should attempt to improve further the sensitivity and selectivity of dairy cow pain's indirect metrics. The magnitude of changes in the animal's behavioral activity patterns and physiological signs, rate of appearance, and decline may be indicative of the cow's condition and are probably more specific than measurement at a single point in time.

This work aimed at providing proof-of-concept that an automatic algorithm can distinguish between healthy cows suffering from mild somatic pain and those that are not. A learning herd (LH) and an unlearning herd (UNLH) approach were tested. The two methods differed in the population that was used to derive the classifier. In the LH approach, a classifier was derived from labeled observations of the herd. The LH classifier was used to classify new unlabeled observations of cows from the same herd it was trained on. This approach was expected to be highly accurate since each cow's characteristics were accounted for when the classifier was run. As anticipated, a mean classification accuracy of 82 % was achieved using this technique. It is reasonable to assume that performance could be further improved with a broader dataset. However, the LH technique requires having an expert label the training set observations each time the system is implemented on a new farm or a new herd. This problem is avoided when applying the UNLH approach.

The UNLH approach is broadly applicable since the classifier is trained once and can be deployed on numerous farms. The UNLH approach was simulated in this work by excluding the observations of a single cow, as though it belonged to a different herd. The UNLH achieved an average accuracy of 86 %±18 %. Noxious detection of about 25 % of the cows results in a classification accuracy below 70 %. This may be a result of the cows' specific response to the noxious stimulus or pain. The cows' current age, as well as their neonatal pain and experiences of stress, may account for some of the inter-individual variability (Adcock and Tucker, 2018). Real-life applications of the UNLH method might involve greater variation between the training set and the test sets (Shalev-Shwartz and Ben-David, 2013), resulting in lower accuracy than seen in the simulated case reported here. A straightforward step to improve accuracy would be to generate a larger training set from several farms. Another alternative would be to use the transfer learning technique (Daumé III, 2010; Segev et al., 2017; Tan et al., 2018), in which a classifier is computed from a large set of observations. The resulting classifier is expected to be accurate in the source domain but less accurate in the target domain. The accuracy is then improved by refining the classifier using a relatively small set of labeled observations obtained from the target domain. This type of fine-tuning might be cost-effective.

6. Conclusion

Acute pain related to management procedures (e.g., castration and disbudding) or controlled and spontaneous pathogenic infections are manifested clearly on both the behavioral and physiological levels. The present work contributes to a better understanding of the signs of mild pain by suggesting an algorithm that can detect covert noxious events that may be less noticeable due to the cow's stoic nature. In this work, the mild pain model involved the topical application of capsaicin cream that elicited a short-term noxious event. Cardiovascular and respiratory changes were observed immediately after the stimulus was applied, along with changes in the cows' daily activity patterns. However, since none of these changes by themselves was event-specific, a systematic collection of on-farm event-related data is imperative for improved specificity. This work provides a proof-of-concept that an automatic algorithm has the potential to distinguish cows suffering from a noxious event that might go unnoticed from those who are not. Future work should involve finding improved techniques that will allow for continuous measurements over long durations, adding the time domain to the diagnostic process. This will enable researchers to accurately monitor cows' responses to temporally defined events and explore deviations from the norm.

Funding

This work was supported by the Office of the Chief Scientist, Ministry of Agriculture and Rural Development, Israel, Grant ID 20-07-0025.

Declaration of Competing Interest

The authors report no declarations of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.applanim.2021.105260.

References

Adcock, S.J.J., Tucker, C.B., 2018. Painful procedures: when and what should we be measuring in cattle? Advances in Cattle Welfare. Elsevier Ltd, pp. 157–198. https://doi.org/10.1016/b978-0-08-100938-3.00008-5.

Aubert, A., 1999. Sickness and behaviour in animals: a motivational perspective. Neurosci. Biobehav. Rev. 23, 1029–1036. https://doi.org/10.1016/S0149-7634(99)00034-2.

Bar, D., Soloman, R., 2010. Rumination collars: what can they tell us. First North Am. Conf. Precis. Dairy Manag. 2010, 2.

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification And Regression Trees. Routledge. https://doi.org/10.1201/9781315139470.

Burdick, N.C., Carroll, J.A., Dailey, J.W., Randel, R.D., Falkenberg, S.M., Schmidt, T.B., 2012. Development of a self-contained, indwelling vaginal temperature probe for use in cattle research. J. Therm. Biol. 37, 339–343. https://doi.org/10.1016/j.jtherbio.2011.10.007.

Bustamante, H.A., Rodríguez, A.R., Herzberg, D.E., Werner, M.P., 2015. Stress and pain response after oligofructose induced-lameness in dairy heifers. J. Vet. Sci. https://doi.org/10.4142/jvs.2015.16.4.405.

Carstens, E., Moberg, G.P., 2000. Recognizing pain and distress in laboratory animals. ILAR J. 41, 62–71. https://doi.org/10.1093/ilar.41.2.62.

Cyples, J.A., Fitzpatrick, C.E., Leslie, K.E., DeVries, T.J., Haley, D.B., Chapinal, N., 2012. Short communication: the effects of experimentally induced Escherichia coli clinical mastitis on lying behavior of dairy cows. J. Dairy Sci. 95, 2571–2575. https://doi.org/10.3168/jds.2011-5135.

Daumé III, H., 2010. Frustratingly easy semi-supervised domain adaptation. ACL Work. Domain Adapt. NLPomain Adapt. NLP, 53–59.

de Boyer, des., Roches, A., Faure, M., Lussert, A., Herry, V., Rainard, P., Durand, D., Foucras, G., 2017. Behavioral and patho-physiological response as possible signs of pain in dairy cows during Escherichia coli mastitis: a pilot study. J. Dairy Sci. 100, 8385–8397. https://doi.org/10.3168/jds.2017-12796.

de Oliveira, C.C., Alves, F.V., de Almeida Martins, P.G.M., Junior, N.K., Alves, G.F., de Almeida, R.G., Mastelaro, A.P., da Costa e Silva, E.V., 2019. Vaginal temperature as indicative of thermoregulatory response in Nellore heifers under different microclimatic conditions. PLoS One 14, 1–13. https://doi.org/10.1371/journal.pone.0223190.

Di Giminiani, P., Petersen, L.J., Herskin, M.S., 2014. Capsaicin-induced neurogenic inflammation in pig skin: a behavioural study. Res. Vet. Sci. 96, 447–453. https://doi.org/10.1016/j.rvsc.2014.03.023.

Dyer, R.M., Neerchal, N.K., Tasch, U., Wu, Y., Dyer, P., Rajkondawar, P.G., 2007. Objective determination of claw pain and its relationship to limb locomotion score in dairy cattle. J. Dairy Sci. 90, 4592–4602. https://doi.org/10.3168/jds.2007-0006.

Ede, T., Lecorps, B., von Keyserlingk, M.A.G., Weary, D.M., 2019. Symposium review: scientific assessment of affective states in dairy cattle. J. Dairy Sci. 102, 10677–10694. https://doi.org/10.3168/jds.2019-16325.

Fitzpatrick, C.E., Chapinal, N., Petersson-Wolfe, C.S., DeVries, T.J., Kelton, D.F., Duffield, T.F., Leslie, K.E., 2013. The effect of meloxicam on pain sensitivity, rumination time, and clinical signs in dairy cows with endotoxin-induced clinical mastitis. J. Dairy Sci. 96, 2847–2856. https://doi.org/10.3168/jds.2012-5855.

Fodeh, S.J., Haven, N., Finch, D., Veterans, J.A.H., Bouayad, L., Veterans, J.A.H., Luther, S.L., Veterans, J.A.H., Ling, H., Haven, N., Kerns, R.D., Haven, W., Brandt, C., Haven, N., 2019. Classifying clinical notes with pain assessment using machine learning. Med. Biol. Eng. Comput. 56, 1285–1292. https://doi.org/10.1007/s11517-017-1772-1.Classifying.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232. https://doi.org/10.1214/aos/1013203451.

Giovannini, A.E.J., van den Borne, B.H.P., Wall, S.K., Wellnitz, O., Bruckmaier, R.M., Spadavecchia, C., 2017. Experimentally induced subclinical mastitis: are lipopolysaccharide and lipoteichoic acid eliciting similar pain responses? Acta Vet. Scand. 59, 1–13. https://doi.org/10.1186/s13028-017-0306-z.

Gleerup, K.B., Andersen, P.H., Munksgaard, L., Forkman, B., 2015a. Pain evaluation in dairy cattle. Appl. Anim. Behav. Sci. 171, 25–32. https://doi.org/10.1016/j.applanim.2015.08.023.

Gleerup, K.B., Forkman, B., Lindegaard, C., Andersen, P.H., 2015b. An equine pain face. Vet. Anaesth. Analg. 42, 103–114. https://doi.org/10.1111/vaa.12212.

Green, B.G., Zaharchuk, R., 2001. Spatial variation in sensitivity as a factor in measurements of spatial summation of warmth and cold. Somatosens. Mot. Res. 18, 181–190.

Greenhouse, S.W., Geisser, S., 1959. On Methods in the Analysis of Profile Data. Springer.

Grubb, T.L., Anderson, D.E., 2017. Assessment of clinical application of pulse oximetry probes in llamas and alpacas. Vet. Med. Sci. 3, 169–175. https://doi.org/10.1002/vms3.68.

Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z., Yang, X., 2013. Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests. Int. J. Remote Sens. 34, 5166–5186. https://doi.org/10.1080/01431161.2013.788261.

Hansen, L.K., Salamon, P., 1990. Neural network ensembles. IEEE Trans. Pattern Anal. Mach. Intell. 12, 993–1001. https://doi.org/10.1109/34.58871.

Hansen, M.F., Smith, M.L., Smith, L.N., Jabbar, K.A., Forbes, D., 2018. Computers in Industry Automated monitoring of dairy cow body condition, mobility and weight

using a single 3D video capture device. Comput. Ind. 98, 14–22. https://doi.org/10.1016/j.compind.2018.02.011.
Heinrich, A., Duffield, T.F., Lissemore, K.D., Millman, S.T., 2010. The effect of meloxicam on behavior and pain sensitivity of dairy calves following cautery dehorning with a local anesthetic. J. Dairy Sci. 93, 2450–2457. https://doi.org/10.3168/jds.2009-2813.
Hindman, M., 2015. Building better models: prediction, replication, and machine learning in the social sciences. Ann. Am. Acad. Pol. Soc. Sci. 659, 48–62. https://doi.org/10.1177/0002716215570279.
Hudson, C., Whay, H., Huxley, J., 2008. Recognition and management of pain in cattle. In Pract. 30, 126–134. https://doi.org/10.1136/inpract.30.3.126.
Huynh-Thu, V.A., Saeys, Y., Wehenkel, L., Geurts, P., 2012. Statistical interpretation of machine learning-based feature importance scores for biomarker discovery. Bioinformatics 28, 1766–1774. https://doi.org/10.1093/bioinformatics/bts238.
Islam, M.J., Wu, Q.M.J., Ahmadi, M., Sid-Ahmed, M.A., 2008. Investigating the performance of naive- bayes classifiers and K- nearest neighbor classifiers. Inst. Electr. Electron. Eng. (IEEE) 1541–1546. https://doi.org/10.1109/iccit.2007.148.
Kanz, P., Krieger, S., Drillich, M., Iwersen, M., 2018. Technical note: evaluation of a wireless pulse oximeter for measuring arterial oxygen saturation and pulse rate in newborn Holstein Friesian calves. J. Dairy Sci. 101, 6437–6442. https://doi.org/10.3168/jds.2017-14266.
Kemp, M.H., Nolan, A.M., Cripps, P.J., Fitzpatrick, J.L., 2008. Animal-based measurements of the severity of mastitis in dairy cows. Vet. Rec. 163, 175–179. https://doi.org/10.1136/vr.163.6.175.
Landa, L., 2012. Pain in domestic animals and how to assess it: a review. Vet. Med. 57, 185–192. https://doi.org/10.17221/5915-VETMED.
Lay, D.C., Friend, T.H., Bowers, C.L., Grissom, K.K., Jenkins, O.C., 1992. A comparative physiological and behavioral study of freeze and hot-iron branding using dairy cows. J. Anim. Sci. 70, 1121–1125. https://doi.org/10.2527/1992.7041121x.
Leslie, K.E., Petersson-Wolfe, C.S., 2012. Assessment and management of pain in dairy cows with clinical mastitis. Vet. Clin. North Am. Food Anim. Pract. 28, 289–305. https://doi.org/10.1016/j.cvfa.2012.04.002.
Maltz, E., 2020. Individual dairy cow management: achievements, obstacles and prospects. J. Dairy Res. 87, 145–157. https://doi.org/10.1017/S0022029920000382.
Mauchly, J.W., 1940. Significance test for sphericity of a normal n-Variate distribution. Ann. Math. Stat. 11, 204–209. https://doi.org/10.1214/aoms/1177731915.
Mench, J.A., 2018. Advances in agricultural animal welfare. Advances in Agricultural Animal Welfare. https://doi.org/10.1016/c2015-0-05920-2.
Molony, V., Kent, J.E., 1997. Assessment of acute pain in farm animals using behavioral and physiological measurements. J. Anim. Sci. 75, 266–272. https://doi.org/10.2527/1997.751266x.
Morton, D.B., Griffiths, P.H.M., 1985. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. Vet. Rec. 116, 431–436. https://doi.org/10.1136/vr.116.16.431.
Paci, P., Mancini, C., Price, B.A., 2019. Designing for Wearability, pp. 1–12. https://doi.org/10.1145/3371049.3371051.
Peter, J., Peter, C., 2002. Clinical Examination of Farm Animals. Blackwell Publishing Ltd.
Piñeiro, J.M., Rajala-Schultz, P.J., Schuenemann, G.M., Bas, S., Barragan, A.A., Sanders, D.E., Lakritz, J., 2018. Assessment of daily activity patterns and biomarkers of pain, inflammation, and stress in lactating dairy cows diagnosed with clinical metritis. J. Dairy Sci. 101, 8248–8258. https://doi.org/10.3168/jds.2018-14510.
Piperine, O.I.P.O., Caterina, N., M.j Schumacher, M.A., Tominaga, M., Rosen, T.A., Levine, J.D., Julius, D., 1997. The capsaicin receptor: a heat-activated ion channel in the pain pathway. Nature 389, 816–824.
Richter, T., Fishbain, B., Markus, A., Richter-Levin, G., Okon-Singer, H., 2020. Using machine learning-based analysis for behavioral differentiation between anxiety and depression. Sci. Rep. 10, 16381.
Segev, N., Harel, M., Mannor, S., Crammer, K., El-Yaniv, R., 2017. Learn on source, refine on target: a model transfer learning framework with random forests. IEEE Trans. Pattern Anal. Mach. Intell. 39, 1811–1824. https://doi.org/10.1109/TPAMI.2016.2618118.
Sepulveda-Varas, P., Proudfoot, K.L., Weary, D.M., von Keyserlingk, M.A.G., 2016. Changes in behaviour of dairy cows with clinical mastitis. Appl. Anim. Behav. Sci. 175, 8–13. https://doi.org/10.1016/j.applanim.2014.09.022.
Shalev-Shwartz, S., Ben-David, S., 2013. Understanding machine learning: From theory to algorithms, Understanding Machine Learning: From Theory to Algorithms. https://doi.org/10.1017/CBO9781107298019.
Siivonen, J., Taponen, S., Hovinen, M., Pastell, M., Lensink, B.J., Pyörälä, S., Hänninen, L., 2011. Impact of acute clinical mastitis on cow behaviour. Appl. Anim. Behav. Sci. 132, 101–106. https://doi.org/10.1016/j.applanim.2011.04.005.
Singh, K.P., Malik, A., Mohan, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) - a case study. Water Res. 38, 3980–3992. https://doi.org/10.1016/j.watres.2004.06.011.
Steensels, M., Antler, A., Bahr, C., Berckmans, D., Maltz, E., Halachmi, I., 2016. A decision-tree model to detect post-calving diseases based on rumination, activity, milk yield, BW and voluntary visits to the milking robot. Animal 10, 1493–1500. https://doi.org/10.1017/S1751731116000744.
Stewart, M., Stafford, K.J., Dowling, S.K., Schaefer, A.L., Webster, J.R., 2008. Eye temperature and heart rate variability of calves disbudded with or without local anaesthetic. Physiol. Behav. 93, 789–797. https://doi.org/10.1016/j.physbeh.2007.11.044.
Suykens, J.A.J., Vandewalle, J., 1999. Least squares support vector machine classifiers. Neural Process. Lett. 9, 293–300.
Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C., 2018. A survey on deep transfer learning. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 11141, 270–279. https://doi.org/10.1007/978-3-030-01424-7_27. LNCS.
Thomsen, P.T., Munksgaard, L., Tøgersen, F.A., 2008. Evaluation of a lameness scoring system for dairy cows. J. Dairy Sci. 91, 119–126. https://doi.org/10.3168/jds.2007-0496.
Treister, R., Kliger, M., Zuckerman, G., Aryeh, I.G., Eisenberg, E., 2012. Differentiating between heat pain intensities: the combined effect of multiple autonomic parameters. Pain 153, 1807–1814. https://doi.org/10.1016/j.pain.2012.04.008.
Van Hertem, T., Bahr, C., Tello, A.S., Viazzi, S., Steensels, M., Romanini, C.E.B., Lokhorst, C., Maltz, E., Halachmi, I., Berckmans, D., 2016. Lameness detection in dairy cattle: single predictor v. Multivariate analysis of image-based posture processing and behaviour and performance sensing. Animal 10, 1525–1532. https://doi.org/10.1017/S1751731115001457.
von Keyserlingk, M.A.G., Rushen, J., de Passillé, A.M., Weary, D.M., 2009. Invited review: the welfare of dairy cattle—key concepts and the role of science. J. Dairy Sci. 92, 4101–4111. https://doi.org/10.3168/jds.2009-2326.
Warnick, L.D., Janssen, D., Guard, C.L., Gröhn, Y.T., 2010. The effect of lameness on milk production in dairy cows. J. Dairy Sci. 84, 1988–1997. https://doi.org/10.3168/jds.s0022-0302(01)74642-5.
Worley, A., Fabrizi, L., Boyd, S., Slater, R., 2012. Multi-modal pain measurements in infants. J. Neurosci. Methods 205, 252–257. https://doi.org/10.1016/j.jneumeth.2012.01.009.