IEEE Transactions on Industrial Informatics, Available online 18 April 2024
DOI: 10.1109/TII.2024.3384594

Delineating mastitis cases in dairy cows: Development of an IoT-enabled intelligent decision support system for dairy farms

Mohammad Farhan Khan, Vivi Mørkøre Thorup, and Zhenhua Luo

***Abstract*— Mastitis, an intramammary bacterial infection, is not only known to adversely affect the health of a dairy cow but also to cause significant economic loss to the dairy industry. The severity and spread of mastitis can be restrained by identifying the early signs of infection in the cows through an intelligent decision support system. Early intervention and control of infection largely depend on the availability of on-site high throughput machinery, which can analyse milk samples regularly. However, due to limited resources, marginal and small farms usually can not afford such high-end machinery, hence the financial loss in such farms due to mastitis may become significant. To overcome such limitations, this paper proposes a low-complexity yet affordable automated system for accurate prediction of early signs of clinical mastitis infection in dairy cows. In this work, behavioural data collected through IoT-enabled wearable sensors for cows is utilised to develop a support vector machine (SVM) model for the daily prediction of mastitis cases in a dairy farm. The data set from the research herd utilises the information of 415 cows collected in the span of 4.75 years in which 75 cows had mastitis. In addition to relevant behavioural features, other statistically significant features such as daily milk yield, lactation period, and age are also utilised as features. Our study indicates that the SVM model comprising a subset of behavioural and non-behavioural features can deliver a mastitis prediction accuracy of 89.2%.**

***Index Terms*— Animal health informatics, automated detection, clinical mastitis, decision support system, IoT wearable sensor.**

I. Introduction

Mastitis is a production disease in dairy cattle, which triggers an inflammatory condition in the mammary gland [1]. Some of the major mastitis-causing pathogens are contagious and easily spread from one cow to another, hence adversely impacting the dairy industry [2]. The UK economy annually suffers a loss of approximately £170 million due to mastitis in cattle [3]. To reduce the financial losses resulting from to proliferation of the disease, it is desirable to predict the onset of mastitis in the earliest phase and thereby enable intervention.

Diagnosing mastitis in cows can help control the disease, which in turn can assist animal welfare by hindering the severe effect of mastitis otherwise due to prediction delays [4]. The intervention of mastitis requires a decision support system that can extract vital information related to cows and utilise the core information to forecast the possible signs of disease. Several studies have been conducted in the domain of decision support systems that can predict the signs of disease using wearable sensors [5].

The study conducted by Nogami *et al.* [6] used a wearable sensor system to measure the body temperature of a calf by mounting the sensor on the calf's tail. In another study, Miura *et al.* [7] used a wireless sensor system for detecting estrus in cattle by monitoring their ventral tail base surface temperature. On the other hand, Benaissa *et al.* [8] has developed automatic wearable monitoring systems to quantify the behaviour of cows with the help of leg- and neck-mounted accelerators. Similarly, Norton *et al.* [9] has investigated the role of wearable sensing systems in accurately predicting stress in horses in real-time.

Integrating machine learning algorithms with diverse data sets can identify complex patterns associated with the disease and can provide timely insights by contributing to improved animal welfare, enhanced milk quality, and increased farm productivity. Support vector machine (SVM), decision trees, neural networks, and logistic regression are some of the well-known machine learning algorithms that are widely utilised in the literature for classifying mastitis disease [10]. Out of various machine learning algorithms, SVM is considered one of the most efficient and frequently opted algorithms [11]. SVM can differentiate between the classes of infected and healthy cows by creating an optimal classification boundary [12]. It is worthwhile noticing that the performance of machine learning algorithms largely depends on the quality of the feature space, which might be compromised due to noise and bias within the data set [13]–[15].

Usually, a real-world data set is comprised of noisy data values, which severely decrease the quality of the feature space by developing erroneous conjectures. Another major factor that reduces the performance of machine learning algorithms is a bias among the classes, causing a tendency to miss the adequate information of the minority class and favour the

"This study was carried out with the support of Innovate UK under Grant no. 104989."

M. F. Khan is with the Sir David Bell Building, Digby Stuart College, University of Roehampton, London SW15 5PU, UK (e-mail: mohammad.f.khan@roehampton.ac.uk).

V. M. Thorup was with IceRobotics Ltd, now: Peacock Technology, Unit 13 Alpha Centre, Stirling University Innovation Park, Scotland FK 4NF, UK. She is now with the Department of Animal Health, Aarhus University, Tjele - DK 8830, Denmark (e-mail: vivim.thorup@anis.au.dk).

Z. Luo is with the School of Water, Energy, Environment and Agrifood, Cranfield University, Cranfield - MK43 0AL, UK (e-mail: z.luo@cranfield.ac.uk).

© 2024 IEEE. This is the Author Accepted Manuscript issued with: Creative Commons Attribution License (CC:BY 4.0).
The final published version (version of record) is available online at DOI:10.1109/TII.2024.3384594. Please refer to any applicable publisher terms of use.

majority class during the learning process.

II. Related work and motivation

Some recent studies have focused on predicting the onset of clinical mastitis using machine learning models. The study conducted by Hyde *et al.* [3] has developed a random forest model that can predict contagious and environmental mastitis cases in dairy cows. Ghafoor and Sitkowska [16] have developed a highly accurate random forest classifier that utilises the up-sampling method for reducing data bias while detecting clinical mastitis. Similarly, Fadul-Pacheco *et al.* [17] developed a random forest model on various features including somatic cell score and milk production. The prediction model can forecast the onset of clinical mastitis with an accuracy of 72% during first lactation and the same 72% on a daily basis. On the other hand, the model proposed by Fan et al. [18] can attain a sensitivity of 95% for predicting clinical mastitis cases by up-sampling the feature space which includes electrical conductivity (EC). Similarly, the study conducted by Luo *et al.* [19] can predict clinical mastitis cases with an accuracy of 97% by including EC as one of the features.

According to most of the other recent studies reported in Table I [20]–[30], the two crucial factors that highly contribute in differentiating between healthy cows and mastitis infected cows are somatic cell count (SCC) and milk EC. However, to evaluate crucial milk parameters such as SCC and milk EC of diseased cows requires a high throughput on-farm machinery, which can analyse milk samples regularly [31]. However, the resources that are available to marginal and small dairy farmers are limited and prevent them from affording such machinery [32].

Under such scenarios, the classic approach would be monitoring the behavioural pattern of a cow, which tends to change after contracting the disease [33]. A wearable internet of things (IoT) sensor is capable of continuous collection of the movement and behavioural data of individual cows, and it may raise a simple alert for a diseased cow at early phase of infection. The anomaly in the behaviour of a cow can be closely monitored by observing the core behavioural pattern of each cow through lying time, standing time, step count, etc. [35]–[37]. However, usually the relationship of behavioural factors with animal welfare is highly complex, hence making the disease almost unobservable at its earliest phase [37]. Using machine learning to identify and exact the hidden patterns in the behavioural data, would provide a mean to handle such complex relationship and effectively differentiate diseased cow from healthy ones [38], [39].

In this study, both of the aforementioned obstacles are dealt with while developing machine learning algorithms for mastitis detection, so that the proposed model can predict the change in the behaviour of the cow. To the best of our knowledge, to date, no study has considered the role of bias within the behavioural data set while developing a machine learning model for predicting the early phase of clinical mastitis in dairy cows. Fig 1 illustrates the procedure adopted in this study for diagnosing clinical mastitis in Holstein Friesian dairy cows, which includes behavioural data collected through IceQube sensor, synchronisation with the farm study along with other non-behavioural data sets, under-sampling of data to remove bias between healthy and mastitis classes, and different feature elimination procedures for developing machine learning models. It should be noted that for this study, a healthy cow has not had mastitis identified during the previous 14 days. A cow was allowed to have other types of disease during those 14 days, to allow real-world situations.

The reminder of this paper is organized as follows: An overview of an IoT enabled sensor, data acquisition, pre-processing, and statistical analysis of the feature space and its role in the development of an unbiased SVM model is discussed in Section III. Results are reported and their meaning is interpreted in detail in Section IV, followed by concluding remarks in Section V.

III. Materials and methods

A. IoT sensor

This study used data collected through an IoT-enabled wearable sensor [40], namely a 3-dimensional accelerometer (IceQube, IceRobotics Ltd, South Queensferry, UK). Fig. 2 illustrates the attachment of a sensor that records data from individual cows. The IceQube sensor captures various behavioural features of a cow, such as lying time, standing time, lying bout, standing bout, step count, IceScore, and motion index. Lying time is recorded when the hind leg is in a horizontal position, while standing time is recorded when the hind leg is not in a horizontal position, that is, lying and standing postures are mutually exclusive. A lying bout is recorded when a sensor changes from vertical to horizontal. Similarly, a standing bout is recorded when a sensor changes from a horizontal to a vertical position. The step count measures the number of times a cow lifts its leg and places it back down again. The motion index is a summarisation of the acceleration of the leg in all three dimensions, thus expressing overall leg activity [41]. The IceScore is a proprietary feature. The latter two features have no unit. The features are summarised in steps of 15 minutes, and at each milking data were transferred from the IceQube sensor through the local farm server to a cloud-based server, CowAlert [40].

B. Data acquisition and pre-processing

The complete data set utilised in this study was from a period of approximately 4.75 years, from October 2013 to June 2018. It was captured from the Langhill herd at Crichton Royal Farm, Dumfries, UK, which belongs to Scotlands Rural College, where the temperature rarely exceeds 20°C and the temperature-humidity index stays below 68. The Crichton Royal farm runs two distinct genetic groups of Holstein Friesian dairy cows; the Select group is a genetic line with increased milk fat and crude protein yield, while the Control group is another genetic line that remains close to average merit for milk fat plus crude protein yield [42]. Cows remained in the herd for three lactations, or until there was a suitable replacement unless the animal was culled for welfare reasons. Calving took place all year round. Lactating cows were loose-housed in four pens of 24 cubicles with rubber mattresses,

TABLE I
MACHINE LEARNING ALGORITHMS USING SCC AND MILK EC AS FEATURES TO DIFFERENTIATE BETWEEN HEALTHY COWS AND MASTITIS INFECTED COWS.

Reference	Features used	Data balancing	Best performing algorithm	Best performing metric
[3]	SCC and other parameters	-	Random forest	Accuracy = 98.2%
[16]	Udder inhale/exhale limits and body temperature	Random up-sampling	Random forest	Accuracy = 99.1%
[17]	SCC and other parameters	SMOTE up-sampling	Random forest	AUC = 72%
[18]	EC and other parameters	Random under-sampling	Decision tree based ensemble	Sensitivity = 95%
[19]	EC and other parameters	SMOTE up-sampling	Decision tree	Accuracy = 98%
[20]	SCC	-	Linear discriminant analysis	Accuracy = 79%
[21]	SCC and other parameters	-	Gradient boosted tree	Accuracy = 84.9%
[22]	SCC, EC and other parameters	-	SVM	Sensitivity = 89%
[23]	EC and other parameters	-	SVM	Sensitivity = 84.6%
[24]	SCC, EC and other parameters	-	Random forest	Accuracy = 90%
[25]	SCC, EC and other parameters	-	Extra trees classifier	AUC = 79%
[26]	SCC and other parameters	-	Linear regression	AUC = 71%
[27]	SCC, EC and other parameters	-	Conventional neural network	Accuracy = 76.2%
[28]	SCC and mid-infrared spectroscopy	-	Partial least squares regression	AUC = 85%
[29]	SCC and other parameters	-	SVM	Accuracy = 79.6%
[30]	SCC, EC, behavioural and other parameters	-	Recurrent neural network	Accuracy = 84.3%

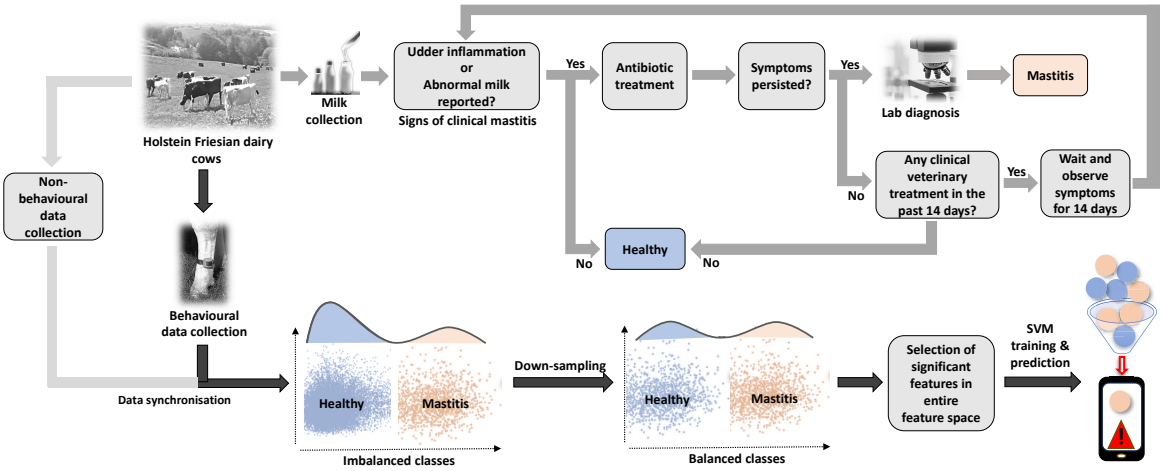

Fig. 1. Illustration of the procedure in the farm adopted by farm staff for diagnosing clinical mastitis in Holstein Friesian dairy cows. The behavioural data collected through IceQube sensor is synchronised with the farm study along with other non-behavioural data sets and sampled together to create balanced classes of healthy and mastitis cases for developing the feature defined SVM models.

Fig. 2. IceQube sensor attached to the hind leg of a cow. Photos: Courtesy of IceRobotics Ltd.

top-dressed with sawdust, and grooved concrete floors in the alleys with automatic scrapers. Cows were milked three times daily in a conventional milking parlour, the milking periods ranged from 0730 to 1000 h; 1430 to 1630 h; and 2130 to 2330 h. Milk yield was measured at each milking. The cows left the pen as a group for about 40 min/milking.

Mastitis was diagnosed by farm staff during the milking procedure by checking for the signs of mastitis at each milking instance. A cow was declared as having mastitis if signs of clinical mastitis were noted, including udder changes (inflammation/swelling) or abnormal milk (clots/flakes), with or without signs of systemic illness. Following diagnosis, cows were treated for three days with antibiotics and analgesics. If symptoms persisted, treatment was repeated. An aseptically collected sample was taken from the affected quarter/quarters and sent for bacteriological culture to identify the cause of each clinical mastitis case. In the present study, a cow was defined as healthy when she had no clinical veterinary treatment in the past 14 days. All treatments made were entered into the research database, and codes referring to different treatments have been established as routine practice. In the present study, these treatment codes were grouped into two categories: (i) clinical mastitis, and (ii) all other disorders, including reproductive, metabolic, and leg disorders etc.

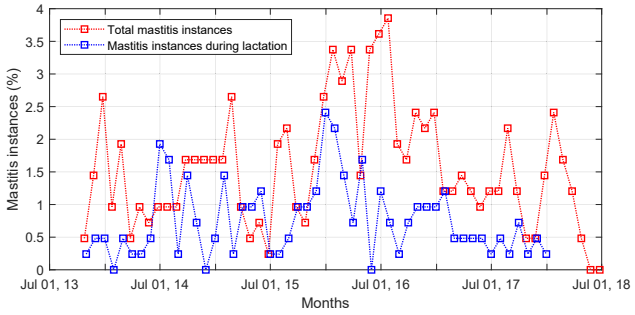

Fig. 3. Frequency of mastitis instances with respect to healthy instances in each month over the observation period. To estimate the prevalence, the total number of affected cows has been divided by total number of cows at each month.

To predict the daily mastitis cases, data from the different sources were integrated, meaning all the data recorded with small step sizes or at each milking, such as the behavioural features and the milk yield, were averaged or summarised to 24-hour values.

The overall data set consisted of 415 cows, out of which 75 cows had a mastitis diagnosis. Each cow was continuously recorded around the clock in steps of 15 minutes. There were 175395 valid log entries in the data set, where each entry comprised a detailed profile of a cow. A complete profile of each cow includes cow ID, unique ear tag, date and time of data logging, icescore which calculates the difference between motion index and steps, steps taken by a cow, standing time, lying time, lying bouts which counts the number of times a cow gets down from standing position, standing bouts which counts the number of times a cow gets up from lying position, total daily milk yield (milking three times a day), calving date, cow weight at calving, cow weight at drying off instance, date of birth.

The frequency of mastitis instances in the herd over the observation period along with the frequency of infection across the lactation period is illustrated in Fig. 3. The IceQube sensor records behavioural data that sometimes fails to capture the data and records values as zero, which is then completely removed from the data set due to its rare occurrence. In the current data set, there are approximately 430 additional bad log entries, which implies around 0.24% of missing log entries in the data set.

Note that, other factors than disease do affect cow behaviour; management procedures regarding feeding, milking, cleaning the stall, season, and physiological stage (when did the cow have her last calf, is the cow in estrus, is it a high or low yielding cow). The management procedures we assumed in this study were identical for all cows, and the physiological stage is accounted for in the model by including age, days since calving, and milk yield.

C. Support vector machine

The SVM is one of the most highly recognised machine learning algorithms, which can predict the class of a new data value by utilising a selection criterion developed with the help of training inputs. The SVM can differentiate between the healthy and the diseased class by formulating a mean classification hyperplane [12].

Consider a set of training data values (x_i, y_i), $i = 1, 2, \cdots, n$ where input feature $x_i \in \mathbb{R}^n$ and output $y_i \in \{-1, 1\}$. The SVM maximises the gap between two hyperplanes H_1 and H_2 (associated with the diseased and healthy class) by maximising the modified dual Lagrange objective function:

$$\max \sum_{i=1}^{n} \beta_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_i \beta_j y_i y_j (x_i \cdot x_j) \qquad (1)$$

$$\text{s.t. } \hat{C} \geq \beta_i \geq 0 \ \forall i, \quad \sum_{i=1}^{n} \beta_i y_i = 0$$

where, β is Lagrange constraint, and $\hat{C}$ is a soft control variable (box constraint), which determines a trade-off between margin width and classification mistakes. To handle a complex non-linear classification problem, the expression $(x_i \cdot x_j)$ in the objective function can be mapped to a non-linear kernel function $K(x_i, x_j)$, where K can be a quadratic function, cubic function, radial basis function (RBF), etc.

IV. Results and Discussion

Table II reports the descriptive statistics of the aforementioned data recorded during approximately 4 years and 7 months. From Table II, observing the mean and standard deviation of all the features along with respective skewness and kurtosis values, it can be observed that the larger portion of the data points of all the features (except daily milk yield) of mastitis infected cows are conflated into the healthy data points, hence might require a non-linear kernel function for formulating an accurate SVM hyperplane.

Further, the number of data points for mastitis and healthy class are highly skewed, which is likely to create a positive bias in predicting healthy cases [15]. The bias in a SVM model can be eliminated in two ways, that is either by quantitatively down-sampling the healthy data points or up-sampling the mastitis data points by introducing synthetic data points. However, the current database is comprised of approximately 1511 times higher quantitative skewness, therefore opting for up-sampling criteria will introduce a large number of synthetic data points that might resemble unrealistic mastitis infection. Therefore, in this study, a down-sampling criteria is opted to quantitatively balance the skewness between the classes.

The down-sampling of the database can be performed by either randomly choosing a subset of 116 data points from the healthy class or utilising a day prior healthy data of all the cows that have contracted mastitis. Given that the objective of the study is to predict the earliest phase of clinical mastitis for early intervention, then among both the down-sampling possibilities, the latter scenario is comparatively more feasible, because it helps analysing the imperceptible change in the behaviour of the cows after contracting the disease.

Thus, in this study, 24-hour prior recorded data was utilised to analyse the change in the feature values compared to the one when the cow had contracted the infection. The reason behind opting for this period is to allow the farmers to test

TABLE II
DESCRIPTIVE STATISTICS OF FEATURES USED FOR CLASSIFICATION OF MASTITIS CASES AND HEALTHY CASES.

Feature ID	Features	Mastitis ($N_1 = 116$)				Healthy ($N_2 = 175279$)			
		Mean	SD	Skewness	Kurtosis	Mean	SD	Skewness	Kurtosis
1	IceScore (no unit)	1.31	0.50	0.382	2.721	1.37	0.51	0.856	5.09
2	Standing bout (number/day)	10.0	3.56	0.581	3.532	9.71	3.38	1.351	7.40
3	Step count (number/day)	1105	1791	2.460	9.100	1453	1083	4.027	72.27
4	Motion index (no unit)	3901	2930	2.599	10.288	5204	5064	36.030	3523.30
5	Standing time (sec/day)	48194	9236	0.423	3.258	46653	8705	0.033	3.48
6	Lying time (sec/day)	38206	9236	-0.423	3.258	39746	8705	-0.034	3.48
7	Lying bout (number/day)	10.0	3.68	0.491	3.571	9.84	3.43	1.264	7.06
8	Milk yield (kg/day)	12.3	9.11	0.877	4.085	28.3	8.80	0.433	3.43
9	Days since calving (days)	142	94.7	0.679	2.762	189	112.6	0.673	3.93
10	Calving weight (kg)	622	60.0	0.411	3.047	587	69.9	0.266	2.93
11	Dry-off weight (kg)	710	83.7	0.1394	3.339	692	108	7.357	112.50
12	Age (days)	1442	363	0.764	3.816	1319	402	0.690	3.00

the milk for bacterial growth and start the treatme which may increase the likelihood of curing and de excessive usage of antibiotics. Fig. 4 illustrates the di of all the twelve feature values of 75 cows that mastitis during the period of data collection. Note of 75 infected cows, 41 cows were reinfected in 4 years and 7 months, and hence has been consid new mastitis case for developing machine learnin Fig. 4(h), the daily yield of mastitis cows is visu lower than the yield of healthy cows. Although mastitis are still being milked, their milk is disc therefore not recorded as soon as a laboratory inv confirms contraction of the infection. Hence, for infe the remaining milking slots (in a day) have been re zero which creates a bias in a data logging process. in this work, a daily milk yield feature is to be used t the impact of the bias of a feature while training model.

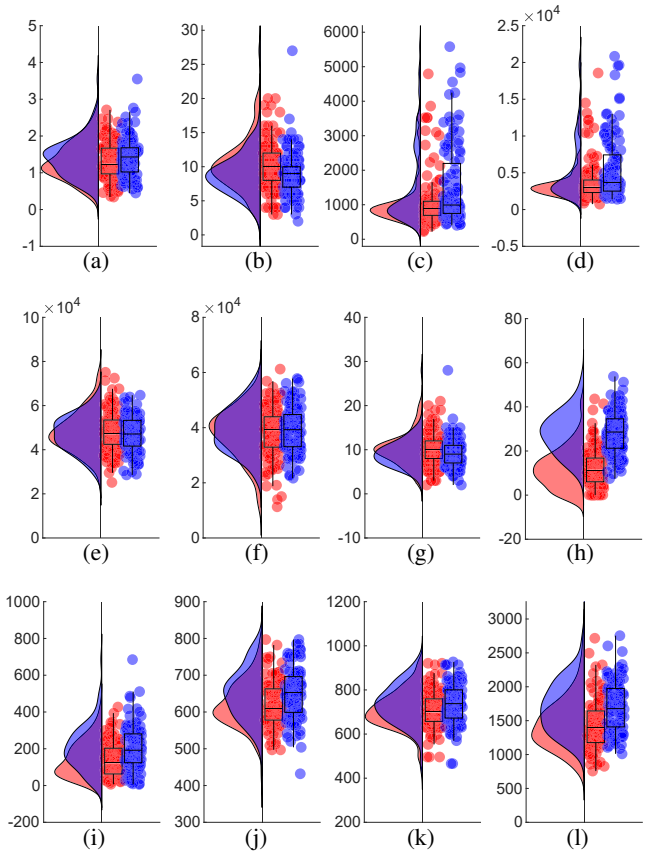

Fig. 4. A raincloud representation of feature space of cows with mastitis (shown in red colour) and with healthy cows (shown in blue colour): (a) IceScore, (b) standing bout, (c) step count, (d) motion index, (e) standing time, (f) lying time, (g) lying bout, (h) milk yield, (i) days since calving, (j) calving weight, (k) dry-off weight, and (l) age.

A. Analysis of variance and χ^2 test

Before developing the SVM model for predicting mastitis cases on a daily basis, it is vital to statistically analyse the significance of the feature space using analysis of variance and χ^2 test [43]. Table III demonstrates the variation between cows with mastitis with respect to the 24-hour prior data of the healthy cows using a sum of squares, F-ratio, and p-value of each feature, respectively. Following are the null and alternate hypothesis statements for the features considered in Table III:

Null hypothesis (H_0): There is no significant difference between the specific features of mastitis-infected cows and healthy cows.

Alternate hypothesis (H_1): There is a significant difference between the specific features of mastitis-infected cows and healthy cows.

In Table III, the sum of squares represents the deviation of all the data points from the mean value, while F-ratio signifies the relative variability of group mean values. Note that, the magnitude of a sum of squares largely depends on the average value of that specific feature. Observing Table III along with Fig. 4, it can be asserted that IceScore, lying time, calving weight, and dry-off weight are not very good classification features, because their null hypothesis is accepted, meaning there is no significant difference between the mastitis infected cows and healthy cows for these features.

The low F-ratio of IceScore, lying time, calving weight, and dry-off weight suggest that the group means are very close, hence the variability is low between the groups relative to within the group. On the other hand, the higher F-ratio of the remaining features signifies greater variability of group means as compared to the within-group variability. Hence such factors are likely to contribute to a greater extent in accurately classifying the cows as diseased or healthy.

TABLE III
ESTIMATION OF FEATURE RELATION OF MASTITIS INFECTED COWS WITH RESPECT TO 24 HOUR PRIOR HEALTH DATA USING ANOVA TEST.

Features	Sum of squares	F-ratio	p-value	Null hypothesis
IceScore	0.15	0.51	0.47	Accept
Standing bout	124.57	9.54	<0.01	Reject
Step count	7.81×10^6	7.80	<0.01	Reject
Motion index	1.09×10^8	7.89	<0.01	Reject
Standing time	6.04×10^8	5.01	0.02	Reject
Lying time	1.42×10^7	0.13	0.71	Accept
Lying bout	89.38	6.44	0.01	Reject
Milk yield	1.21×10^4	119.63	<0.01	Reject
Days since calving	1.63×10^5	13.28	<0.01	Reject
Calving weight	33.37	0.00	0.95	Accept
Dry-off weight	347.65	0.02	0.89	Accept
Age	1.88×10^6	9.94	<0.01	Reject

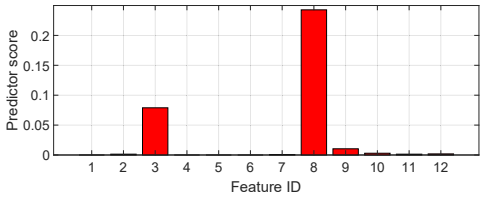

Fig. 5. Confidence score of individual feature estimated using MRMR algorithm.

Note that, there are other factors than the disease which affect cow behaviour; examples include management procedures regarding feeding, milking, cleaning the stall, season, and physiological stage. In the data set, the management procedures were assumed identical for all cows, and the accounted physiological stage is considered in the model by including age, days since calving, and milk yield.

B. Minimum redundancy maximum relevance

Another algorithm that has been adopted in this work for identifying the significance of features is minimum redundancy maximum relevance (MRMR), which ranks the features of the entire feature space using a forward addition scheme [44]. The MRMR algorithm chooses an optimum set of maximally and mutually diverse features to adequately represent the response variable. Fig. 5 illustrates the estimated confidence score of the features using the MRMR algorithm. A high score indicates maximum confidence in the feature in differentiating infected cows from healthy ones. Observing Fig. 5, it can be asserted that milk yield, step count, and days since calving are the three most important features that have maximum relevance compared to other features for predicting infection cases.

C. Multicollinearity

Multicollinearity is one of the most important factors that need to be investigated in the data set prior to developing machine learning algorithms [45]. In the realm of machine learning, the presence of multicollinearity demands meticulous scrutiny before model development due to its potential to significantly distort the results and undermine the reliability

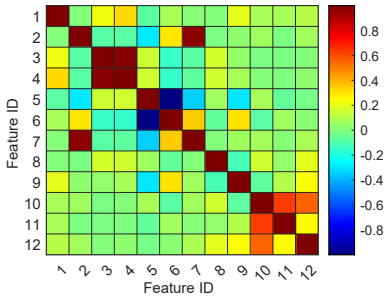

Fig. 6. PCC of all the features with respect to other features in feature space.

of the predictive model. The linearly correlated variables may exhibit redundant information, making it challenging to identify which predictors contribute unique insights to the model. Consequently, the model may include unnecessary variables, increase complexity without enhancing predictive performance, or overlook important features, diminishing the model's explanatory power [46], [47]. Hence, in this current study, if any two features have a linear Pearson correlation coefficient (PCC) > 0.5 are considered moderately correlated, hence one feature is dropped from the feature space while retaining the other feature. Fig. 6 illustrates the PCC of variables concerning each other, and it has been found that eight features namely icescore, standing bout, step count, standing time, lying time, milk yield, days since calving, and calving weight are not strongly correlated, hence considered for disease classification.

D. Performance of SVM classification model

To develop the SVM model for predicting the daily mastitis cases, a 5-fold cross-validation approach was adopted along with four kernel variants. In the 5-fold cross-validation approach, 1-fold is reserved for testing purposes while the remaining four are used for training purposes. To implement the 5-fold cross-validation approach, the data set is randomly shuffled and split into five closely equal sub-samples. Then five iterations are conducted, and in each iteration, a unique 1-fold of the sub-sample is reserved for testing while the remaining 4-folds are used for training. In the end, the average value of performance metrics namely accuracy, area under curve (AUC), sensitivity, specificity, and Matthews correlation coefficient (MCC) are recorded. Note that, for evaluating the performance of the SVM models, a mastitis-infected cow was deemed as a positive instance, while a healthy cow was considered as a negative instance.

Fig. 7 illustrates the variation in the accuracy of the SVM model for predicting daily mastitis infection cases in cattle for different kernel functions $K(x_i, x_j)$ explored through systematic grid-search. For the RBF kernel, the systematic grid search has been performed by varying a combination of box constraint and kernel scale variables in the range of $[10^{-3}, 10^2]$ with a step size of 0.25. For the remaining kernel functions only box constraint was varied in the aforementioned range with a step size of 0.25.

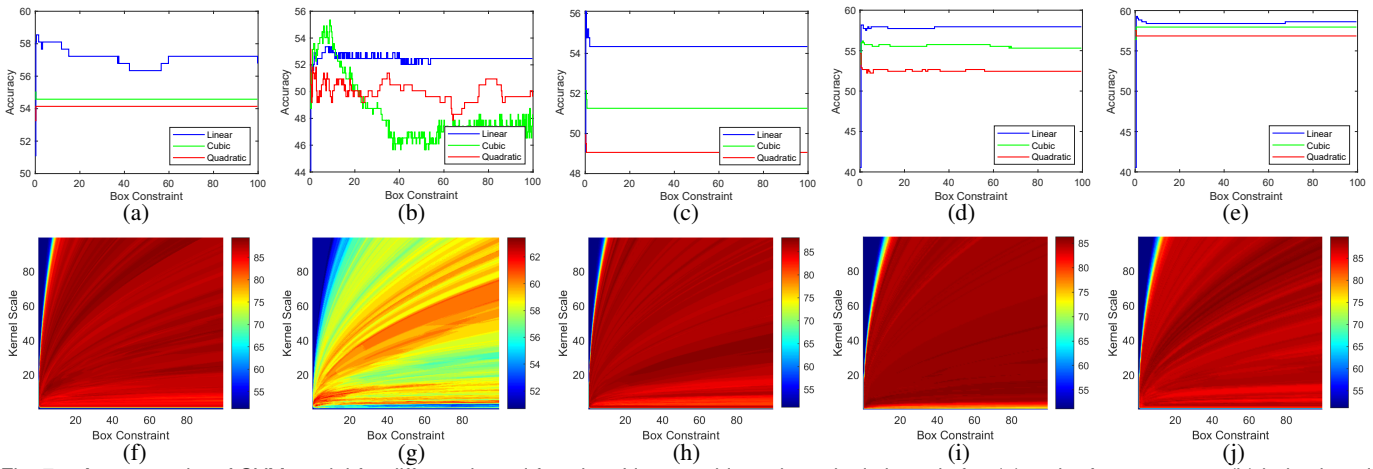

Fig. 7. Accuracy plot of SVM model for different kernel function: Linear, cubic and quadratic kernels for, (a) entire feature space, (b) behavioural features, (c) statistically significant features, (d) MRMR features, (e) PCC features; RBF kernel for, (f) entire feature space, (g) behavioural features, (h) statistically significant features, (i) MRMR features, (j) PCC features.

From Fig. 7(a) it can be stated that for the entire feature space, the linear kernel can achieve the highest accuracy at $\hat{C}$ = 0.50. However, the highest achievable accuracy of 58.5% by linear kernel states that feature space is not linearly separable, hence it is necessary to explore non-linear kernel functions such as cubic, quadratic, and RBF. Observing the performance of cubic and quadratic kernel functions in Fig. 7(a), it is found that both functions performed worse than the linear kernel, that is, cubic and quadratic kernels are able to achieve the highest accuracy of 55.0%, and 54.1%, respectively. The box constraint value that gives giving best performance for both cubic and quadratic kernels is 0.25.

From the AUC value of all the three kernels from Table IV, it can be asserted that the SVM models developed using the aforementioned kernels hold substandard predictive power and have a higher percentage of error by classifying an excessive number of false positive cases.

Again observing Fig. 7(f) it can be stated that by considering the entire feature space, the systematic grid-search can find a maximum accuracy of 89.4% for RBF kernel scale = 13.00 and $\hat{C}$ = 5.75. Observing other metrics of the RBF kernel in Table IV shows that the RBF kernel can achieve an AUC of 0.95 with sensitivity and specificity of 88.5% and 90.4%, respectively. The higher value of sensitivity, specificity, and AUC suggest that the SVM model with an RBF kernel is capable of predicting mastitis cases more accurately with a lesser risk of assuming healthy instances as false positive and mastitis instances as false negative. To further strengthen the preposition, MCC is reported in Table IV. MCC gives more importance to the true prediction of both positive and negative instances. The high MCC values support the findings of the SVM RBF kernel to be the highest-performing model.

A similar pattern is also followed by the SVM model developed by considering a subset of seven behavioural features recorded through the IceQube sensors. Fig. 7(b) and (g) illustrate the highest possible performance of different kernel functions by excluding all non-behavioural features of the cow. The highest achievable accuracy of 63.4% is attained by the RBF kernel for kernel scale = 2.00 and $\hat{C}$ = 0.75.

On the other hand, compared to the accuracy of the RBF kernel, the remaining three kernels fell short by 8 to 10 percentage points. The linear, cubic, and quadratic kernels can achieve the highest accuracy at $\hat{C}$= 6.75, 8.75, and 0.001 respectively. Further, the lower value of MCC indicates poor performance by non-RBF kernels in creating a balance for accurately predicting positive and negative instances.

In contrast, Figs. 7(c) and (h) illustrate the importance of statistically significant features in attaining closer behaviour to that of the entire feature space. The subset of significant features reported in Table III through ANOVA and χ^2 test are utilised to estimate a drop in classification performance. The objective behind opting for statistically significant features is to analyse the usefulness of less sparse data in reducing the computational complexity of the low-end system. Statistical analysis has indicated that opting for significant feature space for optimised RBF kernel (with kernel scale = 2.25 and $\hat{C}$ = 0.50) produced the highest AUC value, which signifies that opting for significant features tended to slightly suppress the occurrence of false positive instances. Note that, the features flagged in statistical analysis can be used individually but the accuracy will be severely affected in those predictions. As mentioned in Table IV, using all the statistically significant features with the linear kernel gives only 55.4% accuracy, which implies that the relation of the feature distribution is not straightforward and there is a need to explore more complex relations which can utilise all the statistically significant features together.

Similar to statistically significant features, MRMR and PCC algorithms are used to extract the important features that are likely to contribute towards the classification of diseased states with respect to healthy ones. From Figs. 7(d)-(e) and (i)-(j), it can be stated that 8 features shortlisted by PCC analysis have attained the accuracy of 89.2% for kernel scale = 18.5 and $\hat{C}$ = 2.75. While only 3 features selected by MRMR analysis can achieve an accuracy of 85.3% (for kernel scale = 3.00 and $\hat{C}$ = 2.00), which is relatively close to the highest achievable accuracy of 89.4%. Hence, it is evident that MRMR features are the most optimal feature sub-spaces

TABLE IV
PERFORMANCE OF FOUR VARIANTS OF THE SVM MODEL TRAINED OVER FIVE DIFFERENT SETS OF FEATURES.

Feature set	No. of features	SVM kernel	Performance metrics				
			Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)	MCC
All features	12	Linear	58.5	0.51	49.2	51.1	0.15
		Cubic	55.0	0.51	49.1	52.0	0.14
		Quadratic	54.1	0.51	49.9	51.7	0.13
		RBF	**89.4**	**0.95**	**88.5**	**90.4**	**0.79**
Behavioural features	7	Linear	53.4	0.53	57.4	58.8	0.16
		Cubic	55.5	0.52	56.0	57.6	0.14
		Quadratic	53.1	0.51	49.1	51.1	0.14
		RBF	63.4	0.70	58.2	59.8	0.18
Statistically significant features	8	Linear	55.4	0.55	62.6	63.3	0.53
		Cubic	54.1	0.51	50.0	53.4	0.20
		Quadratic	54.1	0.52	51.5	55.8	0.20
		RBF	**88.1**	**0.97**	**87.5**	**88.7**	**0.76**
MRMR features	3	Linear	58.1	0.52	52.6	53.3	0.19
		Cubic	55.8	0.51	51.2	52.2	0.19
		Quadratic	55.1	0.51	51.8	53.1	0.20
		RBF	85.3	0.91	84.7	86.8	0.69
PCC features	8	Linear	59.1	0.52	52.6	53.4	0.16
		Cubic	58.2	0.51	51.2	51.9	0.14
		Quadratic	58.1	0.51	51.1	51.5	0.14
		RBF	**89.2**	**0.96**	**87.8**	**90.1**	**0.78**

that can achieve comparatively high accuracy by dropping a large number of irrelevant features, which is very helpful in reducing the computational complexity of portable devices. Note that, the objective of the paper is to propose a prediction model for a specific farm operating under specific conditions by collecting behavioural and non-behavioural data, hence the proposed model can't be deemed as generalised one.

Comparing the proposed machine learning algorithms with the most relevant techniques reported by Hyde *et al.* [3], Fadul-Pacheco *et al.* [17], Fan *et al.* [18], and Luo *et al.* [19], it can be stated that all of the techniques rely on either of the two crucial factors namely SCC and EC to predict the mastitis cases in dairy farms. By utilising SCC or EC, the aforementioned techniques are able to achieve from reasonable to a higher level of sensitivity, however evaluating the SCC and EC might not be possible for marginal and small dairy farmers due to limited resources. It is worth noting that, the PCC analysis conducted on the data collected from Scottish Holstein Friesian dairy cows came out to be 0.5722, which shows that a practically weak relationship exists between mastitis and SCC for Holstein Friesian dairy cows in that specific farm. On the other hand, Ghafoor and Sitkowska [16] opted for the up-sampling criteria for removing the bias in the data set, which has introduced a large number of synthetic data points resembling unrealistic mastitis infection.

V. CONCLUSION

In this paper, a solution for detecting clinical mastitis in the early stage of infection for marginal dairy farms has been introduced with the help of an IoT-enabled wearable sensor and machine learning. The complete data set consisted of seven behavioural features plus eight simplistic non-behavioural features, which have been collected in a span of approximately 5 years and aggregated to a step size of 24 hours. To analyse the importance of each feature and curtail excessive computational complexity for low-end systems, PCC, MRMR, and ANOVA tests have been employed. The statistical analyses show that out of all the features, three features namely milk yield, step count, and days since calving are the most important features that significantly capture the difference between healthy cows and infected cows. The proposed SVM model developed using the eight PCC features is able to predict the daily mastitis cases in dairy cows with the near highest achievable accuracy, sensitivity, and specificity of 89.2%, 87.8%, and 90.1%, respectively. It is worth noting that, if other diseases elicit the same behavioural response and decline the milk yield like clinical mastitis then the model will not be able to distinguish the other disease from mastitis. In addition, for all the RBF kernels, the best performing $\hat{C}$ value was lying at a low magnitude value, which depicts that the RBF kernel-based SVM model is giving more freedom to misclassification of data points that are leading to a smoother separation. Hence it can be concluded that, for predicting clinical mastitis cases, strict separation of classes is going to deteriorate the performance of the SVM classification model and should be avoided.

REFERENCES

[1] G. Keefe, "Update on control of Staphylococcus aureus and Streptococcus agalactiae for management of mastitis", *Vet. Clin. North Am. Food Anim. Pract.*, vol. 28, no. 2, pp. 203-216, 2012.

[2] F. Gomes, M. J. Saavedra, and M. Henriques, "Bovine mastitis disease/pathogenicity: Evidence of the potential role of microbial biofilms", *Pathog. Dis.*, vol 74, no. 3, Art. no. ftw006, 2016.

[3] R. M. Hyde, P. M. Down, A. J. Bradley, J. E. Breen, C. Hudson, K. A. Leach, and M. J. Green, "Automated prediction of mastitis infection patterns in dairy herds using machine learning", *Sci. Rep.*, vol. 10, no. 1, Art. no. 4289, 2020.

[4] K. Sah, P. Karki, R. D. Shrestha, A. Sigdel, A. T. Adesogan, and G. E. Dahl, "MILK Symposium review: Improving control of mastitis in dairy animals in Nepal", *J. Dairy Sci.*, vol. 103, no. 11, pp. 9740-9747, 2020.

[5] Y. Feng, H. Niu, F. Wang, S. Ivey, J. Wu, H. Qi, R. Almeida, S. Eda, and Q. Cao, "SocialCattle: IoT-based mastitis detection and control through social cattle behavior sensing in smart farms", *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10130-10138, 2022.

[6] H. Nogami, H. Okada, T. Miyamoto, R. Maeda, and T. Itoh, "Wearable wireless temperature sensor nodes appressed to base of a calf's tail", *Sens. Mater.*, vol. 26, no. 8, pp. 539-545, 2014.

[7] R. Miura, K. Yoshioka, T. Miyamoto, H. Nogami, H. Okada, and T. Itoh, "Estrous detection by monitoring ventral tail base surface temperature using a wearable wireless sensor in cattle", *Anim. Reprod. Sci.*, vol. 180, pp. 50-57, 2017.

[8] S. Benaissa, F.A. Tuyttens, D. Plets, T. De Pessemier, J. Trogh, E. Tanghe, L. Martens, L. Vandaele, A. Van Nuffel, W. Joseph, and B. Sonck, "On the use of on-cow accelerometers for the classification of behaviours in dairy barns", *Res. Vet. Sci.*, vol. 125, pp. 425-433, 2019.

[9] T. Norton, D. Piette, V. Exadaktylos, and D. Berckmans, "Automated real-time stress monitoring of police horses using wearable technology", *Appl. Anim. Behav. Sci.*, vol. 198, pp. 67-74, 2018.

[10] N. Esener, A. M. Guerra, K. Giebel, D. Lea, M. J. Green, A. J. Bradley, and T. Dottorini, "Mass spectrometry and machine learning for the accurate diagnosis of benzylpenicillin and multidrug resistance of Staphylococcus aureus in bovine mastitis", *PLoS Comput. Biol.*, vol. 17, no. 6, Art. no. e1009108, 2021.

[11] Z. Wen, J. Shi, Q. Li, B. He, and J. Chen, "ThunderSVM: A fast SVM library on GPUs and CPUs", *J. Mach. Learn. Res.*, vol. 19, no. 1, pp. 797-801, 2018.

[12] T. Joachims, "Making large-scale SVM learning practical," in *Advances in Kernel Methods: Support Vector Learning, 1*st ed. London, UK: MIT Press, 1998, ch. 11, sec. 2, pp. *169–184.*

[13] J. Feng, M. Huang, L. Zhao, Y. Yang, and X. Zhu, "Reinforcement learning for relation classification from noisy data", in *Proc. AAAI*, New Orleans, LA, USA, vol. 32, no. 1, 2018, pp. 5779-5786.

[14] J. Hoffmann, Y. Bar-Sinai, L. M. Lee, J. Andrejevic, S. Mishra, S. M. Rubinstein, and C. H. Rycroft, "Machine learning in a data-limited regime: Augmenting experiments with synthetic data uncovers order in crumpled sheets", *Sci. Adv.*, vol. 5, no. 4, Art. no. eaau6792, 2019.

[15] M. F. Khan, R. K. Gazara, M. M. Nofal, S. Chakrabarty, E. M. A. Dannoun, R. Al-Hmouz, and M. Mursaleen, "Reinforcing synthetic data for meticulous survival prediction of patients suffering from left ventricular systolic dysfunction." *IEEE Access*, vol. 9, pp. 72661-72669, 2021.

[16] N. A. Ghafoor, and B. Sitkowska, "MasPA: A machine learning application to predict risk of mastitis in cattle from AMS sensor data", *AgriEng.*, vol. 3, no. 3, pp. 575-583, 2021.

[17] L. F. Pacheco, H. Delgado, and V. E. Cabrera, "Exploring machine learning algorithms for early prediction of clinical mastitis", *Int. Dairy J.*, vol. 119, Art. no. 105051, 2021.

[18] X. Fan, R. D. Watters, D. V. Nydam, P. D. Virkler, M. Wieland, and K. F. Reed, "Multivariable time series classification for clinical mastitis detection and prediction in automated milking systems", *J. Dairy Sci.*, vol. 106, no. 5, pp. 3448-3464, 2023.

[19] W. Luo, Q. Dong, and Y. Feng, "Risk prediction model of clinical mastitis in lactating dairy cows based on machine learning algorithms", *Prev. Vet. Med.*, Art. no. 106059, 2023.

[20] T. Bobbo, S. Biffani, C. Taccioli, M. Penasa, and M. Cassandro, "Comparison of machine learning methods to predict udder health status based on somatic cell counts in dairy cows", *Sci. Rep.*, vol. 11, no. 1, Art. no. 13642, 2021.

[21] M. Ebrahimi, M. M. Dehcheshmeh, E. Ebrahimie, and K. R. Petrovski, "Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep Learning and gradient-boosted trees outperform other models", *Comput. Biol. Med.*, vol. 114, Art. no. 103456, 2019.

[22] N. Mammadova, and I. Keskin, "Application of the support vector machine to predict subclinical mastitis in dairy cattle", *Sci. World J.*, vol. 2013, Art. no. 603897, 2013.

[23] B. Miekley, I. Traulsen, and J. Krieter, "Mastitis detection in dairy cows: the application of support vector machines", *J. Agric. Sci.* vol. 151, no. 6, pp. 889-897, 2013.

[24] E. Ebrahimie, F. Ebrahimi, M. Ebrahimi, S. Tomlinson, and K. R. Petrovski, "Hierarchical pattern recognition in milking parameters predicts mastitis prevalence", *Comput. Electron. Agric.*, vol. 147, pp. 6-11, 2018.

[25] C. Post, C. Rietz, W. Bscher, and U. Müller, "Using sensor data to detect lameness and mastitis treatment events in dairy cows: A comparison of classification models", *Sensors*, vol. 20, no. 14, Art.no. 3863, 2020.

[26] J. V. Bolanos, C. C. C. Serrano, D. V. Mejía, J. C. R. Rojas, C. E. Giraldo, J. U. Carmona, and A. C. Márquez, "Application of udder surface temperature by infrared thermography for diagnosis of subclinical mastitis in Holstein cows located in tropical highlands", *J. Dairy Sci.*, vol. 104, no. 9, pp. 10310-10323, 2021.

[27] T. Bobbo, R. Matera, G. Pedota, A. Manunza, A. Cotticelli, G. Neglia, and S. Biffani, "Exploiting machine learning methods with monthly routine milk recording data and climatic information to predict subclinical mastitis in Italian Mediterranean buffaloes", *J. Dairy Sci.*, vol. 106, no. 3, pp. 1942-1952, 2023.

[28] L. Rienesl, N. Khayatzdadeh, As. Köck, C. E. Danner, N. Gengler, C. Grelet, L. Monica Dale, A. Werner,F. J. Auer,J. Leblois, and J. Sölkner, "Prediction of acute and chronic mastitis in dairy cows based on somatic cell score and mid-infrared spectroscopy of milk", *Animals*, vol. 12, no. 14, Art no. 1830, 2022.

[29] A. Satoa, and K. Satoa, "Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking and super-learner ensembles versus single machine learning models", *J. Dairy Sci.*, 2024. DOI: 10.3168/jds.2023-24243.

[30] S. A. Naqvi, M. T. M. King, R. D. Matson, T. J. DeVries, Rob Deardon, and Herman W. Barkema. "Mastitis detection with recurrent neural networks in farms using automated milking systems", *Comput. Electron. Agric.*, vol. 192, Art. no. 106618, 2022.

[31] T. Halasa, and C. Kirkeby, "Differential somatic cell count: Value for udder health management", *Front. Vet. Sci.*, vol. 7, Art. no. 1153, 2020.

[32] K. Van Herck, and J. Swinnen, "Small farmers, standards, value chains, and structural change: Panel evidence from Bulgaria", *Br. Food J.*, vol. 117, no. 10, pp. 2435-2464, 2015.

[33] G. Grodkowski, T. Szwaczkowski, K. Koszela, W. Mueller, K. Tomaszyk, T. Baars, and T. Sakowski, "Early detection of mastitis in cows using the system based on 3D motions detectors", *Sci. Rep.*, vol. 12, no. 1, Art. no. 21215, 2022.

[34] K. A. Leach, D. A. Tisdall, N. J. Bell, D. C. J. Main, and L. E. Green, "The effects of early treatment for hindlimb lameness in dairy cows on four commercial UK farms", *Vet. J.*, vol. 193, no. 3, pp. 626-632, 2012.

[35] K. J. Ji, R. E. Booth, and N. Blackie, "A retrospective case study into the effect of hoof lesions on the lying behaviour of Holstein-Friesian in a loose-housed system", *Animals*, vol. 11, no. 4, Art. no. 1120, 2021.

[36] D. T. Byrne, H. Esmonde, D. P. Berry, F. McGovern, P. Creighton, and N. McHugh, "Sheep lameness detection from individual hoof load", *Comput. Electron. Agric.*, vol. 158, pp. 241-248, 2019.

[37] H. F. M. Sobte, and S. Buijs, "Impact of paper bedding on lying behaviour and welfare in lactating dairy cows", *Appl Anim Behav Sci.*, vol. 239, Art. no. 105321, 2021.

[38] S. Neethirajan, "The role of sensors, big data and machine learning in modern animal farming", *Sens. Bio-Sens. Res.*, vol. 29, Art. no. 100367, 2020.

[39] N. W. O'Leary, D. T. Byrne, A. H. O'Connor, and L. Shalloo, "Invited review: Cattle lameness detection with accelerometers", *J. Dairy Sci.*, vol. 103, no. 5, pp. 3895-3911, 2020.

[40] IceRobotics Ltd. [Online]. Available: https://www.icerobotics.com. Accessed on: Dec 20, 2021.

[41] J. Maselyne, M. Pastell, P. T. Thomsen, V. M. Thorup, L. Hänninen, J. Vangeyte, A. Van Nuffel, and L. Munksgaard, "Daily lying time, motion index and step frequency in dairy cows change throughout lactation", *Res. Vet. Sci.*, vol. 110, pp. 1-3, 2017.

[42] G. E. Pollott, and M. P. Coffey, "The effect of genetic merit and production system on dairy cow fertility, measured using progesterone profiles and on-farm recording", *J. Dairy Sci.*, vol. 91, no. 9 pp. 3649-3660, 2008.

[43] D. J. Sheskin, "Inferential statistical tests employed with two independent samples (and related measures of association/correlation)," in *Handbook of Parametric and Nonparametric Statistical Procedures, 3*rd ed., London, UK: CRC Press, 2003, ch. $Test-16$, pp. *493-572*.

[44] H. Peng, and C. Ding, "Minimum redundancy and maximum relevance feature selection and recent advances in cancer classification", in *Proc. SDM,* Newport Beach, CA, USA, 2005, pp. 52–59.

[45] R. A. Fisher, "Statistical methods for research workers," in *Breakthroughs in Statistics: Methodology and Distribution, 1*st ed., New York, USA: Springer, 1970, ch. 11, pp. *66-70*.

[46] N. Shrestha, "Detecting multicollinearity in regression analysis", *Am. J. Appl. Math. Stat.*, vol. 8, no. 2, pp. 39-42, 2020.

[47] G. J. Chen, "A simple way to deal with multicollinearity", *J. Appl. Stat.*, vol. 39, no. 9, pp. 1893-1909, 2012.

Mohammad F. Khan received his B.Tech. and M.Tech. degrees in electronics engineering from the Z. H. College of Engineering and Technology, Aligarh Muslim University, India, in 2010 and 2012, respectively; and Ph.D. degree in electronic engineering from the School of Engineering and Digital Arts, University of Kent, U.K., in 2017. From 2017 to 2022, he held different research/academic positions that were associated with the number of universities such as Cranfield University (UK), University of Central Lancashire (UK), University of Warwick (UK), and Indian Institute of Technology Roorkee (India). Since 2022, he has been associated with the University of Roehampton (UK) as a Lecturer in Computing. His research interests include data processing, industrial data science, feature engineering, explainable AI, computer vision, applied soft computing, control theory application, and computational biology.

Vivi M. Thorup has a MSc in biology. She earned her PhD in biomechanical gait analysis in pigs from Copenhagen University in 2007. Currently, she holds a position as Assistant Professor at Aarhus University, Institute of Animal and Veterinary Sciences. Her scientific interest lies in the intersection of innovation, precision livestock farming, animal welfare, and modelling. Her current activities revolve around management strategies that promote animal health and welfare while reducing climate impact. She has over 20 years of experience with research and development in academia as well as in industry.

Zhenhua Luo received his BSc in nanotechnology in 2006, and PhD degrees in materials science and engineering in 2011, at the University of New South Wales, Australia.

From 2013 to 2016, he was a Research Fellow at the University of Southampton, UK. Since 2016, he has been an academic at Cranfield University, UK, and is currently a Senior Lecturer in Energy Storage and Harvesting. He has published over 50 research articles and 3 patents. His current research interests include renewable energy and energy harvesting, focusing on piezoelectric and thermoelectric, and solar-to-hydrogen generation. He is currently a steering board member of the UK Energy Harvesting Network.

Dr Luo was a recipient of the Innovate UK ICURe award in 2019, and the UK universities Award for Excellence in Hydrogen Research and Innovation in 2024. He has also been nominated for the Earthshot Prize in 2024.

https://dspace.org/

Delineating mastitis cases in dairy cows: development of an IoT-enabled intelligent decision support system for dairy farms

Khan, Mohammad Farhan; Thorup, Vivi Mørkøre; Luo, Zhenhua

2024-07-01
IEEE
https://doi.org/10.1109/TII.2024.3384594
Khan MF, Thorup VM, Luo Z. (2024) Delineating mastitis cases in dairy cows: development of an IoT-enabled intelligent decision support system for dairy farms. IEEE Transactions on Industrial Informatics, Volume 20, Issue 7, July 2024, pp. 9508-9517

Downloaded from a DSpace repository. https://dspace.org/