Lameness Detection as a Service: Application of Machine Learning to an Internet of Cattle

John Byabazaire*†, Cristian Olariu‡, Mohit Taneja*†, Alan Davy*†
*Telecommunications Software and Systems Group, Waterford Institute Of Technology, Waterford, Ireland
†CONNECT- Centre for Future Networks and Communications, Ireland
‡Innovation Exchange, IBM Ireland, Dublin
{jbyabazaire, mtaneja, adavy}@tssg.org, cristian.olariu@ie.ibm.com

***Abstract*—Lameness is a big problem in the dairy industry, farmers are not yet able to adequately solve it because of the high initial setup costs and complex equipment in currently available solutions, and as a result, we propose an end-to-end IoT application that leverages advanced machine learning and data analytics techniques to identify lame dairy cattle.**

As part of a real world trial in Waterford, Ireland, 150 dairy cows were each fitted with a long range pedometer. The mobility data from the sensors attached to the front leg of each cow is aggregated at the fog node to form time series of behavioral activities (e.g. step count, lying time and swaps per hour). These are analyzed in the cloud and lameness anomalies are sent to farmer's mobile device using push notifications. The application and model automatically measure and can gather data continuously such that cows can be monitored daily. This means there is no need for herding the cows, furthermore the clustering technique employed proposes a new approach of having a different model for subsets of animals with similar activity levels as opposed to a one size fits all approach. It also ensures that the custom models dynamically adjust as weather and farm condition change as the application scales. The initial results indicate that we can predict lameness 3 days before it can be visually captured by the farmer with an overall accuracy of 87%. This means that the animal can either be isolated or treated immediately to avoid any further effects of lameness.

***Index Terms*—Lameness, Internet of Things (IoT), Data Analytics, Smart Agriculture, Machine Learning, Micro services, Fog Computing.**

I. INTRODUCTION

Lameness is one of the major problems in dairy cattle [1]. It is one of the factors for reduced performance on many dairy farms, at least through reduced reproductive efficiency, milk production and increased culling [2]. Lameness is the third disease of economic importance in breeding with an average of 11% of cows and a high variability inter-breeding [2]. An all-encompassing definition of lameness includes any abnormality which causes a cow to change the way that she walks, and can be caused by a range of foot and leg conditions, themselves caused by disease, management or environmental factors [3]. Prevention, early detection and treatment of lameness is therefore important to reduce these negative effects of lameness on dairy cows [4], [5].

Traditional approaches for lameness detection are based on locomotion scoring (1-5 scale) that requires observing a cow while walking preferably at the exit of the milking parlor [5]. Such approaches are subjective, time consuming and can only be implemented on very small farms. As the size of the farm increases it becomes impossible to monitor all the animals and it's going to necessitate extra labour which in turn will increase farm expenditure. Observation of postural abnormalities predictive of lameness while cows are locked at stanchions is also used as an alternative detection method [6].

To overcome the challenges in the above approaches, in recent studies, new approaches have been put forth. Some automated lameness assessment techniques have been developed which overcome many problems associated with gait scoring technique. These techniques today are becoming popular on many commercial dairy farms to detect lame cows.

However, it's important to note that whilst interactions among cattle in the same pasture are often inevitable, authors in [4] conclude that under some situations, movement patterns of sub-set of individual tracked cows may have levels of independency that are sufficient for analysis as individual experiments. Also the need to incorporate individual farm setting and geographical context is still lacking in most solutions. For example, in Ireland, animals will stay in the field during summer when the weather is good and they are kept sheltered during the winter. Notably, the activity of the animal will change for example increased lying during the summer may be indicative of lameness which may not be true during the winter. Considering such individual differences while analysing and building models based on cow activity within the same herd may help improve the accuracy of such a model and hence reduce on false alarms.

In this paper, we present an end-to-end IoT application that leverages threshold based clustering and machine learning classification to predict lameness in dairy cattle. The application automatically measures and gathers activity data (Lying time, step count and swap per hour) continuously, such that cows can be monitored daily. This means there is no need for herding the cows. Furthermore, the clustering technique employed ensures that the models dynamically adjusts depending on farm and weather conditions and automatically selects a custom learning model for that cluster.

The rest of the paper is structured as follows: In section II, we present the related work and state of the art, In section III, we present the system architecture and work flow. Here we explain how data moves from the sensors to the fog node and then to the IBM cloud and the general architecture of the IoT end-to-end application of the smart dairy farm setup as a part of our real world testbed deployment, In section IV, we present materials and methods. Here we explain data collection, analysis and the learning model, In section V, we present a discussion of the results, and finally in section VI, we present the concluding remarks and future work.

II. Related work

1) Pressure Plate / Load cell: In these solutions, the main aim is to investigate how the weight is distributed across the legs of the animal as it walks through a marked area. Neveux et al. [10] studied the use of a platform outside the automatic milking system to measure the weight distribution of cows while standing on different surfaces. Chapinal et al. [10], [11] and Pastell et al. [12] later adjusted the experimental setup to measure lameness and hoof lesions. The drawbacks of such solutions may not be only the costs of new and complex equipment but also other technical concerns. For example, Pastell et al. [12] suggested that a cow may suffer pain when walking, which is not as obvious when the cow is standing still. Pressure plate / Load cell require the cow to be in a certain position.

2) Image processing techniques: This category studies the use of image processing techniques to analyse the posture of the animal as it walks through a milking parlour. Poursaberi et al. [13] proposed a method based on detecting the arc of back posture and fitting a circle through selected points on the spine line of a cow as it walks. Viazzi et al. [14] further studied the idea and an algorithm based on Body Movement Pattern was tested under farm conditions. Further study on this method shows that it still has challenges on real farm conditions. Some of these challenges were explored by Poursaberi et al. [15], Van Hertem et al. [16] and Viazzi et al. [17]; (1) changing lighting conditions causing noise and shadows in the images that impede extraction of the back posture and (2) continuous background changes that interfere with cow segmentation from the images.

3) Using Accelerometers: Here, techniques use both 2D and 3D accelerometers to record movement patterns of the animal. This data is then used to build the daily activities of the cow say; walking, lying down. Munksgaard et al. [18] proposed the use of sensors that measure acceleration in different dimensions to automatically monitor activity (standing and lying behaviour) of cows. Their results indicate excellent accuracies between the sensor data attached to the legs of the cows and observations for lying and standing (0.99), activity (0.89) and for number of steps (0.84). Since then, a vast number of studies have used accelerometers to measure dairy cow activity and behaviour.

It is important to note that in all the above solutions the equipment or device must be placed in a controlled position and the cows must either be coerced in or they must go through a controlled procedure. Because cows have a stoic nature, guiding them will bias the measurements, because they will try to hide their weakness and pain compared to measurements during normal routine without the presence of a human or predator. Therefore, there is still need for a more automated solution that monitors the animals everywhere they are, either in the fields grazing, during milking or lying down in the shade. Although there are other sensor based systems, the system presented in this paper differs from these by offering the following advantages;

- Sensor agnostic: The model is built to take in activity data from any kind of sensor used to monitor activity of the animal. This among other thing will reduce the initial installation costs if a farm already has a system in place.
- Avoids vendor lock-in: Design, creation and development of services following a microservices based application design principles to tackle the problem of vendor lock-in and to support multi-vendor interoperability.
- Multiple end-users: Because it is designed as a service, this makes it easy to integrate with the exiting systems. It could be a farmer with an existing system or even an agri-tech service provider who wants to provide more services to his clients.

III. Experimental setup

A. Architecture and Data flow

As shown in Figure 1, after the Receiver receives data from the sensors and transceiver, it then sends the data to the communication unit (RS485 to USB) through wired connection, which in turn sends it to the local PC (which acts as controller and fog node, and is configured[1] as - Intel Core 3rd Generation i7-3540M CPU @ 3.00GHz, 16.0 GB RAM, 500 GB Local Storage) through wired connection via USB interface. The fog node consists of a local database which stores all the data from the sensors before it is preprocessed. The total size of the daily data collection at the fog node is about 10.1MB of unprocessed data. This is then preprocessed and aggregated to form behavioural activities. For this study, three of these are used for the analysis. Also on board the fog node is dashboard which the farmer can interact with [22].

For communication between the fog node and cloud node, Message Queue Telemetry Transport (MQTT) [23] was used. This is made up of two functional components namely; MQTT clients (such as publishers and subscribers) and MQTT broker (for mediating messages between publishers and subscribers). In this study these components are as follows:

- MQTT Publisher: Script running on fog node
- MQTT Broker: IBM Watson IoT Platform (Cloud node)
- MQTT Subscriber: Application designed and hosted on IBM Cloud

After the critical analysis, data preprocessing and aggregation at the fog node, the processed data is sent to the cloud for historical storage and analysis via the IBM Watson IoT Platform. The cloud is also the site for fusion of the data from other sources, such as weather data.This data was also used to investigate the effects of weather on lameness.

[1]The minimum suggested configuration for the given setup is a Dual Core processor @ 2.3GHz, 4.0 GB RAM, 100 GB local storage.

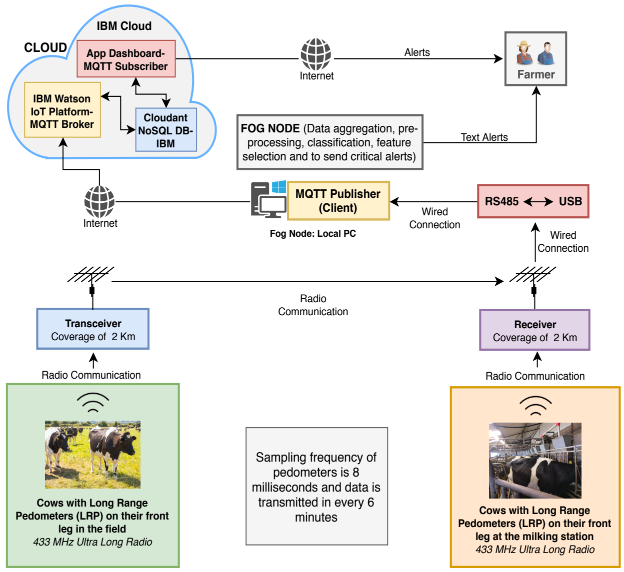

Fig. 1. Overall architecture of the test-bed

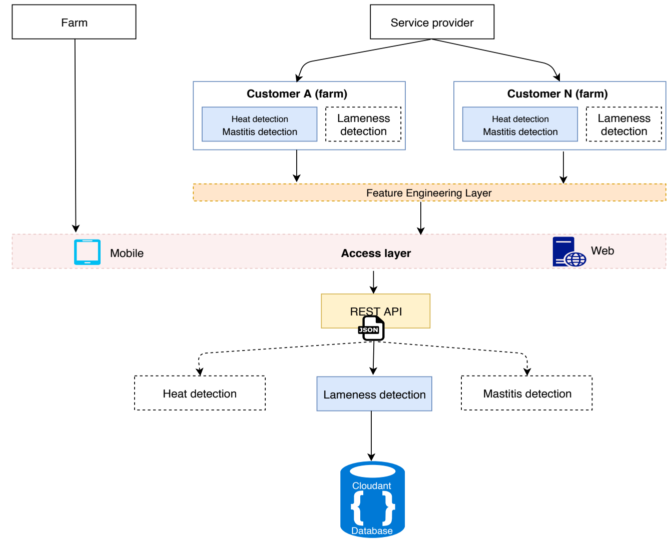

Fig. 2. Microservices based flow of the LDA

B. *Lameness detection as a service*

Unlike all the available implementations that are based on a monolith design approach, the applications designed in this paper follows microservices approach. Because the Lameness Detection Algorithm (LDA) expects a certain number of features, an implementation of a feature engineering layer is added for existing systems or service providers with exiting systems, for example a service can be an agri-tech company providing any other solution like heat detection who wants to integrate our LDA in their system. This ensures that data is transformed to output only the required features and also reject those that can not be engineered to form the required features. Such operations could include feature mapping, for example the LDA expects lying time, step count and swaps but a service provider might have activity counter instead of step count and (Standup+Liedown) instead of swaps. It is important to note that this layer will be different for each service provider since the underlying sensor technology might be different. This is then passed via the access layer which includes both mobile and web via a REST API which in turn calls the LDA. Figure 2 shows the design of the proposed system.

IV. Materials and Methods

A. *Track a cow Long-Range Pedometer*

As part of the experiment[2], a local dairy farm with 150 cows in Waterford, Ireland was used. Commercially available Track a cow Long-Range Pedometer (LRP, ENGS Systems©®, Israel) specifically designed for use in dairy cattle were attached to the front leg of each cow as shown in figure 3. These have an approximate net weight of 124 g, sampling frequency of 8 milliseconds and transmit data every 6 minutes either via the transceiver placed in the field or a receiver placed near the milking parlour each with a coverage of up to 2km. The LRP collects acceleration data which is then converted into activity counters like step count at the fog node. It also does some preprocessing and has on-board memory with a retention capacity of up to 12 hours. Therefore the cows are continuously monitored and data transmitted whether they are in the field during good weather conditions during the summer or adverse winter conditions when they are kept in house.

[2]The ethical approval for the experimentation was taken from Research Ethics Committee of Waterford Institute of Technology, Ireland prior to the deployment in July, 2017.

Fig. 3. LRP attached as a part of the experiment to the front leg of the cows.

B. *Data*

The data from the sensors is sent to the fog node from the receiver where it is preprocessed and aggregated into three behavioural activities; (1) Step count, this is the number of steps an animal makes per hour, (2) Lying time, the number of hours an animal spends lying down and (3) Swaps, this is the number of times an animal moves from lying down to standing up. The choice of the 3 features is guided by literature study that they are among the best predictor of a lame cow or one transitioning to lameness. The data is then summed to form daily time series. Out of 150 cows used in the trial, only 146 cows were used in the analysis. Only data from July to December 2017 was included in this analysis.

During this period, 32 animals were confirmed as lame (cows were confirmed as lame by either an agricultural scientist or by the farmer). Because the number of none lame animals was small, the split also made sure to have at least 75% of the lame animals in the training and the rest in the testing fold. But because this is a live experiment, we hope to re-train the models after sometime. The initial performance on both the training and testing are reported in a later section.

C. Data analysis and machine learning

1) Cow profiles: In order to build robust profiles that are distinguishable by the learning model, one needs to understand how each test profile (lame and non-lame) relates to the rest of the herd. The most common approach would be to compare the activity level of lame and non-lame animals and investigate how these deviate from the mean of the entire herd. However as it is known, the mean can be affected by a single value being too high or low compared to the rest of the sample. This is why a median or quantiles are sometimes taken as a better measure. To that effect, we investigated the relationship between the herd mean and the herd median. The results of this as shown in figure 4 shows that these almost trace out each other for all the three activities: Lying time, Step count and Swaps per hour. This is one of the features of a normal distribution and therefore it would not matter whether the mean or median is used. Authors [19] in argued that animals grazing within the

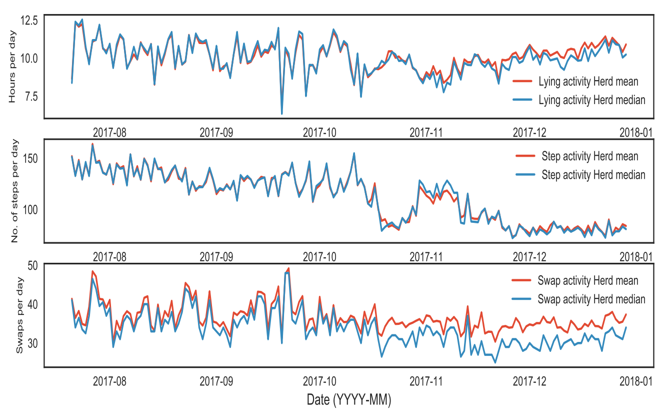

Fig. 4. Comparing the Mean and Median of the various Animal Activities

same pasture can influence the movement, grazing locations, and activities of other animals randomly, with attraction, or with avoidance, therefore most of the animals will have their activity levels equal to the herd mean. For this reason and the one discussed above, the herd mean was used as the baseline and any deviation from such behaviour due to lameness will be classified as an anomaly. We also think that this will help eliminate the effects of external factors like weather and location of the farm as these will be affecting the whole herd and only leave the individual effects of lameness on the cow.

To form a profile for each animal to characterize normal behaviour, we use a window of certain number of days using clustering-based techniques, this helps use to define Lameness Activity Region (LAR, period during which the animal is confirmed as lame) and Normal Activity Region (NAR, period during which the animal is confirmed as non-lame) ground truth which will act as an input for a classification model for predicting lameness.

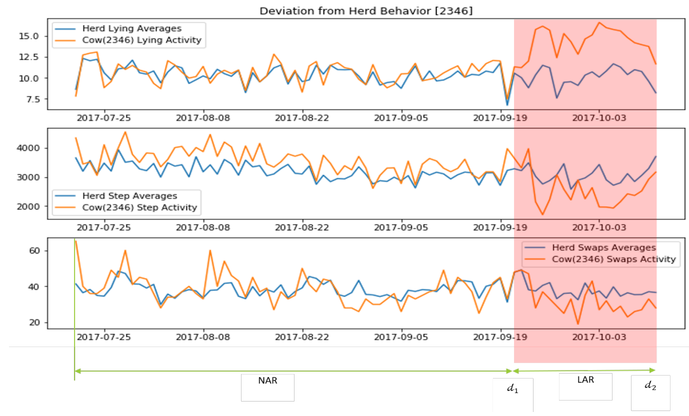

Fig. 5. Relationship between herd mean and cow activity for 2346

2) Clustering: From the above, we discovered that not all animals behaved the same way. For example some animal had their activity levels (step count lying time and swaps) tracing out the herd mean, others with activity levels always higher than the herd mean and the other category always lower than the herd mean. It's also important to note that even when they became lame they had different activity levels depending on which category they belonged to. Therefore the clustering model is based on this. We set thresholds and based on this we form three clusters. If any two of the activity levels are below a certain threshold, then that animal is assigned into one of the below clusters:

Active: These are animals in the herd that have activity levels always higher than the herd mean. These have the mean deviation of any two of the activities is greater than threshold h

Normal: These are animals in the herd that have activity levels always tracing out the herd mean. These have the mean deviation of any two of the activities is less than h but great or equal to zero.

Dormant: These are animals in the herd that have activity levels always lower than the herd mean. These have the mean deviation of any two of the activities is less than zero. This threshold was carefully chosen by a repetitive evaluation process. It is also important to note that these clusters are dynamic, that is the animals keep switching between clusters they belong to. The optimal time to re-cluster was about 2 weeks. This can be caused by many factors like age and weather. So it is the role of the clustering model to keep regrouping the animals before selecting the appropriate classification model for that cluster. Table I shows the distribution of the clusters at the time of analysis. The total number used to build clusters was 146 as three of the animals were eliminated due other health related issues and one animal lost the tag during the experiment.

TABLE I
DISTRIBUTION OF THE CLUSTERS

Active	Normal	Dormant
25	109	12

3) Classification: Classification algorithms are a family of machine learning algorithms that output a discrete value. The output variables are sometimes called labels or categories. These kind of problems always require the examples be classified into two or more classes. Classification problems with two labels are called binary classification problems while those with more than two are called multi-class. We formulated our problem as a binary class problem with Lame as being the positive class and Non-lame as the negative class. For model training, three months data (July 2017 to September 2017) was used and the rest was used for testing (August 2017 to December 2017). This split also made it possible to have 80% of the lameness incidences in the training and 20% in the testing.

V. DISCUSSION OF RESULTS

A. Rationale of the Clustering model

In a study about association patterns of visually observed cattle, Stephenson et al [21] concluded that herds with 40 or less cows did not exhibit preferential or avoidance associations. This means that they lived together as a single group. In contrast, larger herd sizes (53-240 cows) tended to to form associations with other cows stronger than what you would expect by chance. Therefore, the clustering step is only relevant to large herd sizes. Needless to mention, automated lameness solutions are meant for large herd sizes as it is assumed that for small ones, the farmer can visually inspect the cows easily. We compared the results of a one-size-fits-all model and a cluster specific models. Overall cluster specific models reduced the classification error by 8% as compared to a one-size-fits-all model without clustering. For example figure 6 shows an animal that was confirmed as lame from 03/12/2017 to 15/12/2017. Well as the normal cluster model could correctly identify all the days the animal was lame, the one-size-fits-all model could only pick up some days as show by the red points.

B. Fog-Cloud data reduction

Some of the downsides of all the current approaches are that they are either fully cloud based, that is all the data is sent to the cloud for processing of fully on premises, that is all the processing is done on the farm. The disadvantages of the former are many but the scope of this work focused on the reduction of data exchanged between the cloud and the fog node. The disadvantage of the latter are the opportunities

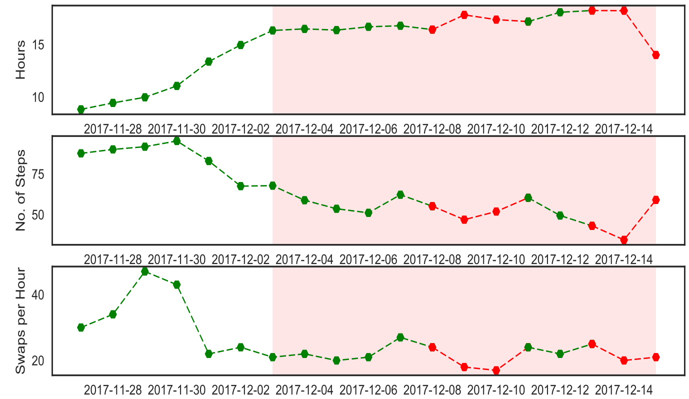

Fig. 6. Animal confirmed as lame between 03/12/2017 to 15/12/2017 but could not be correctly identified by a one-size-fits-all model

missed that come with fusion of data from different sources, for example weather data. This work applied the fog architecture to the problem and was able to reduce the amount of data exchanged between the fog and cloud node from 10.1MB to 1.61MB daily. On daily basis this reduction seems negligible but in the long run it becomes significant. The system was also able to benefit from processing and notification at the fog node. Figure 7 show the daily data reduction between the fog and cloud node.

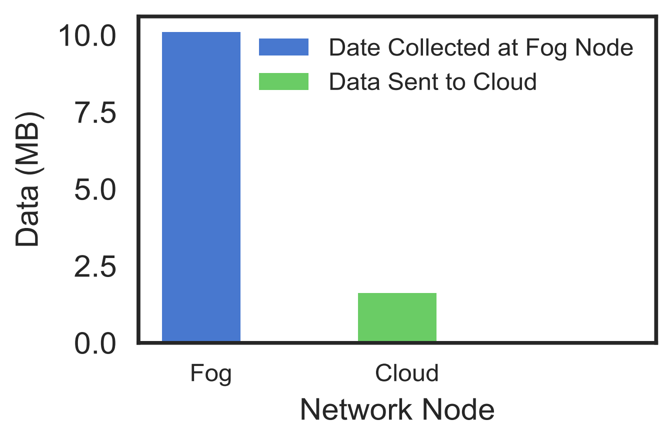

Fig. 7. Daily data reduction between the fog node and the cloud node

C. Lameness

We experimented on a number of sklearn [20] classification algorithms raging from Support vector machine, Random forest, K-Neighbors and Decision trees while noting the accuracy. The best two performing was Random forest (RF) and K-Nearest Neighbors(K-NN) with an accuracy of 91% and 1 day before visual signs and 87% with 3 days before visual signs respectively. It is important to note that although three classification models were trained (one for each cluster), the performance and accuracy reported in this paper are only for normal cluster. This is because the other two were imbalanced for a proper evaluation.

K-NN: This has a number of parameters that should be fine tuned in order to achieve the desired results. Among these, we evaluated different K-values (2-5), which is the number of neighbours to consider while assigning the nearest class. We set the distance metric to euclidean. The highest accuracy was obtained with $k = 2$ although this was over fitting the data. Optimal results were obtained at $k = 4$ which gave an accuracy of 87% with 3 days before the visual signs could be seen. In all, the normal cluster model had a sensitivity of 89.7% and specificity of 72.5% Figure 8 shows some of the correct predictions. One particular cow was confirmed as lame between 16/10/2017 and 25/10/2017 and the model could correctly classify all the days as shown by the red points.

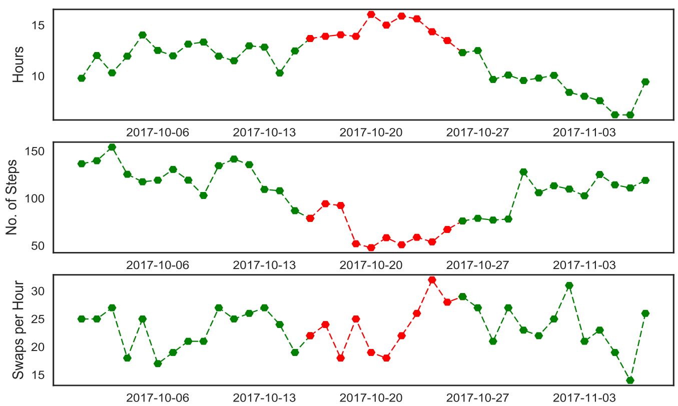

Fig. 8. Red points indicating lameness anomalies identified by the normal cluster model for cow ID 1988

VI. Conclusion and Future Work

Although current approaches have been able to mitigate lameness in the dairy industry and even a few of them made it to commercial markets, these are still expensive because of the complex equipment, neglect individual animal behaviour and also changes in environment and weather conditions.

Our results showed that with a custom model for a small group of animals, we were able to reduce the classification error of the LDA by 8% as opposed to a one-size fits all approach. The solution is also environment and weather agnostic. In our future work we intend to investigate a more robust clustering technique as the current one is only based on threshold. Also evaluate the other cluster models

Acknowledgment

This work has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grant Number 13/RC/2077.

References

[1] O Callaghan, K. A., et al.: Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle, ANIMAL WELFARE-POTTERS BAR THEN WHEATHAMPSTEAD- 12.4 (2003): 605-610.

[2] Chapinal, N., et al. Using gait score, walking speed, and lying behavior to detect hoof lesions in dairy cows. Journal of dairy science 92.9 (2009): 4365-4374.

[3] Dairy Cow Mobility and Lameness - AHDB Dairy. Accessed on 1st Nov 2016 from https://dairy.ahdb.org.uk/technical-information/animal-health-welfare/lameness

[4] Poursaberi, A.; Bahr, C.; Pluk, A.; Berckmans, D.; Veermae, I.; Kokin, E.; Pokalainen, V Online lameness detection in dairy cattle using Body Movement Pattern (BMP), Intelligent Systems Design and Applications (ISDA), 2011 11th International Conference on, Issue Date: 22-24 Nov. 2011

[5] M. Alsaaod , J.J. Niederhauser, G. Beer, N. Zehner, G. Schuepbach-Regula, A. Steiner Development and validation of a novel pedometer algorithm to quantify extended characteristics of the locomotor behavior of dairy cows, Journal of Dairy Science Volume 98, Issue 9, September 2015, Pages 62366242

[6] Garca-Muoz, A., et al. "Evaluation of two methodologies for lameness detection in dairy cows based on postural and gait abnormalities observed during milking and while restrained at headlock stanchions." Preventive veterinary medicine 128 (2016): 33-40

[7] Rajkondawar, P. G., et al. A system for identifying lameness in dairy cattle. Applied En-gineering in Agriculture 18.1 (2002): 87. 8.

[8] N. Chapinal, A. de Passill,M. Pastell, L. Hnninen, L.Munksgaard, and J. Rushen, Measurement of acceleration while walking as an automated method for gait assessment in dairy cattle, Journal of Dairy Science, vol. 94, no. 6, pp. 28952901, 2011.

[9] Xiangyu Song, Toon Leroy, Erik Vranken, Willem Maertens, Bart Sonck, Daniel Berckmansa Automatic detection of lameness in dairy cattleVision-based trackway analysis in cows locomotion, Computers and Electronics in Agriculture Volume 64, Issue 1, November 2008, Pages 3944

[10] Neveux, S., et al. Hoof discomfort changes how dairy cattle distribute their body weight. Journal of dairy science 89.7 (2006): 2503-2509

[11] N. Chapinal, , A.M. de Passill, J. Rushen, S. Wagner Automated methods for detecting lameness and measuring analgesia in dairy cattle, Journal of Dairy Science Volume 93, Issue 5, May 2010, Pages 20072013

[12] Pastell, Matti, et al. Measures of weight distribution of dairy cows to detect lameness and the presence of hoof lesions. Journal of dairy science 93.3 (2010): 954-960.

[13] A. Poursaberi, C. Bahr, A. Pluk, A. Van Nuffel, 1, D. Berckmans Real-time automatic lameness detection based on back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques, Computers and Electronics in Agriculture Volume 74, Issue 1, October 2010, Pages 110119

[14] Viazzi, Stefano, et al. Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. Journal of Dairy Science 96.1 (2013): 257-266.

[15] Poursaberi, Ahmad, et al. Image based separation of dairy cows for automatic lameness detection with a real time vision system. 2009 Reno, Nevada, June 21-June 24, 2009. American Society of Agricultural and Biological Engineers, 2009.

[16] Van Hertem, Tom, et al. Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity. Journal of dairy science 96.7 (2013): 4286-4298.

[17] Viazzi, Stefano, et al. "Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows." Computers and Elec-tronics in Agriculture 100 (2014): 139-147

[18] Munksgaard L., Reenen C.G., Boyce R. Automatic monitoring of lying, standing and walking behavior in dairy cattle. J. Anim. Sci. 2006;84(Suppl.):304.

[19] Stephenson, M.B.; Bailey, D.W. Do movement patterns of GPS-tracked cattle on extensive rangelands suggest independence among individuals? Agriculture 2017.

[20] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

[21] Stephenson, M.B.; Bailey, D.W.; Jensen, D. Associations patterns of visually-observed cattle on Montana, USA foothill rangelands. Appl. Anim. Behav. Sci. 2016, 178, 715.

[22] M. Taneja, J. Byabazaire, A. Davy and C. Olariu, "Fog assisted application support for animal behaviour analysis and health monitoring in dairy farming," 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), pp. 819-824, Feb 2018.

[23] MQTT, http://mqtt.org/, (Last Accessed on Feb 03,2018)