Application of Machine Learning and Analytics for Cattle Behaviour Classification within an Internet of Things Deployment

By

JOHN BYABAZAIRE (NO. 20073944)

Waterford Institute *of* Technology
INSTITIÚID TEICNEOLAÍOCHTA PHORT LÁIRGE

Department of Computing and Mathematics
School of Science and Computing

Thesis submitted in partial fulfillment of the requirement for
the award of a degree of Masters by Research

Supervisor

DR. ALAN DAVY

MARCH 2020

DECLARATION

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Masters by Research is entirely my own work and has not been taken from the work of others save to extent that such work has been cited and acknowledged within the text of my work.

SIGNED: .... STUDENT ID: 20073944

DEDICATION

I dedicate this work to my mother Mary Basigirenda Ateenyi. Thanks for being an inspiration.

Acknowledgment

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Alan Davy for believing in me and his tireless support towards my masters study and the related research. I particularly thank him for challenging me to be an independent researcher by allowing me to work on my own, but steered me in the right the direction whenever he thought I needed it.

I would also like to thank my supervisor in IBM Dr. Cristian Olariu for his insightful comments and encouragement, but also for the hard question which made me to widen my research from various perspectives. I would also like to thank the data analytics team; Dr. Haytham Assem, Dr. Bora Caglayan and Dr. Teodora Sandra Buda and the entire team of Innovation Exchange, IBM Labs Ireland.

I would also like to thank my project mate Mohit Taneja for continuosly challenging me to push my boundaries. I also acknowledge all the students in TSSG, you guys made my study memorable. I also extend my gratitude to the entire TSSG community for their insightful advise during my study.

In a special way, I would like to thank my family: my parents, my brothers and sister for supporting me throughout my study and writing this thesis. I also extend my gratitude to all my friends, especially Sarah Kemigisa Patience.

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.

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, vendor incompatibility and complex equipment in currently available solutions, and as a result, this work presents a hybrid model and an end-to-end Internet of Things (IoT) application that leverages machine learning and data analytics techniques to predict lameness in dairy cattle.

As part of a real world trial in Waterford, Ireland, 150 cows were each fitted with a long range pedometer. The mobility data from sensors attached to the front leg (left leg for 50% of the cows and right leg for the other 50%) of each cow is aggregated to form time series of behavioral activities (Step count, lying time and swaps per hour). These are analyzed in the cloud and alerts of predicted lame animals are sent to the 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 as this would bias the results because cows are stoic in nature. 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 is extended to other farms. The initial results indicate that the application can predict lameness 3 days before it can be visually seen by the farmer with an overall accuracy of 87%. This means that the animal can either be isolated or treated (usually by administering antibiotics) immediately to avoid any further effects of lameness.

The application designed in this study is based on a fog-to-cloud architecture. In this architecture, some of the cloud services and applications are run closer to the physical IoT devices at the network edge. The application also implements a microservices based design approach. The solution can therefore be decoupled as a single service which can be accessed via an Application Programming Interface (API) either by the farmer seeking such a service or an agri-tech service provider who wants to provide such a service to his exiting customers. This also aids data preprocessing and aggregating between the fog node and the cloud. The result of this show an overall data reduction from 10.1MB to 1.62MB exchanged between the fog node and cloud node daily. This is the first time such an approach is implemented for lameness detection and generally for welfare monitoring for dairy cattle.

Table of Contents

List of Tables

List of Figures

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Lameness in Dairy cattle

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, reduced milk production due to milk withdraw and increased culling [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 very important to reduce its effects on dairy cows [4, 5].

Traditional approaches for lameness detection are based on gait scoring that requires observation of cows walking, preferably at the exit of the milking parlor [5]. This is a time-consuming task that can only be implemented on small size dairy farms [4]. As the size of the farm increases, it requires more scoring experts which will in turn increase the operation costs. Since this approach is also based on a person's observations of cows walking to the milking parlour or standing at the milking stall, this makes it also subjective. 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. In this study these have been categorised as; Pressure plate / Load cell for example Force Plate System [7], Image processing techniques, for example Vision Based Trackway Analysis [8], Shape Analysis [4] and those that are based on Accelerometers or pedometers (activity based techniques) [9].

Although these techniques today are becoming popular on commercial dairy farms [10, 11], they still have a lot of challenges. In pressure plate/Load cells, 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. The other drawbacks of such solutions are the high costs of new and complex equipment and other technical concerns. For example, in other setups, the animals are required to be standing in a certain position [12]. Image processing techniques have shown good results in experiments but still have challenges on real farm conditions[13, 14]. Changing lighting conditions causing noise and shadows in the images affects the accuracy of the systems. This category is also affected by high setup costs.

Based on these techniques especially activity based techniques, many general welfare related solutions have been developed. Although most agri-tech companies claim to do lameness detection based on this kind of technology, no such details have been made publically or commercially available. The other challenge with the available solutions is that they are made and fine tuned for specific farm conditions and environments. For example, in Ireland animals stay in sheds during winter and in the fields during the summer. Reduced activity levels during the summer may be inductive of lameness yet the same will not be true during the winter. Some animals naturally have reduced activity level than others and therefore may not get lame the same way. All the solutions above assume a one-size-fits all approach. So, such solutions cannot scale or if they are introduced to new environments they will have a lot of errors. Therefore, there is still need for an AI-based solution that monitors the animals everywhere they are, either in the fields grazing, during milking or lying down in the shade and consider individual animal behaviours (population agnostic) and changes in weather and environment (environmental agnostic).

As part of this study, a real-world trial was conducted 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 (50% left leg and 50% right leg) is aggregated 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 study introduces a novel system and method for predicting lameness in livestock using an AI-based approach by forming a profile for each animal (clustering) to characterize normal behaviour employing a window of certain number of days using a clustering-based technique, followed by a method for defining the formulation of the Lameness Activity Region (LAR) and Normal Activity Region (NAR) ground truth which will act as an input for a classification model for predicting the presence of the lameness. 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 as guiding them would bias the results as compared to measurements taken during a normal routine[15], 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.

1.2 Motivation and Contribution

The high costs associated with lameness are still a major challenge in the dairy industry. Due to the advances in technology, several solutions have been proposed and some of them made it to the market [10, 11], but due to the high initial costs involved and the complex setup process, most farmers still do not find such solutions cost effective[16]. While conducting a survey on the farmers' preferences for automatic lameness detection systems in dairy cattle, the authors in [16] reported that farmers who already had an estrus detection system were more willing to have an add-on. Also sensors attached to the cows were preferred. In a review by authors in [17] about automatic lameness detection, some suggestions were made. One of these was the need for automatic and continuous measurement of the parameters. This is because most solutions available require the animals to be guided one way or another. The other suggestion was the need for custom solutions, systems that need less space or those that can be included in the existing farm infrastructure [17].

The work presented herein is envisioned as the solution to the problems mentioned above, thus bridging the gap, and providing an innovative way that integrates Internet Of Things(IoT), fog computing and machine learning to predict lameness in dairy cattle. The application designed in this study also follows a microservices [18] based approach for design, creation and deployment. Therefore, the study introduces the following core novelty:

1. To investigate the existence of clusters in herds and the application of cluster specific models to predict lameness as opposed to one-size-fits all approach.
2. To investigate the effects of external factors like weather, local farm conditions on lameness prediction.
3. To investigate the application of modern AI techniques to predict rather than detect lameness in livestock.

1.2.1 Publications

From the work reported in this thesis, two conference papers have been published and one journal article. The first paper is included in the appendix and it summarizes the work in this thesis. The machine learning, data analytics and clustering reported in this paper are all extracts of this thesis. The data analysis and data flow reported in the second paper are also extracts of this thesis. The detailed machine learning and model flow reported in this thesis are presented in the journal article.

1. **J. Byabazaire**, C. Olariu, M. Taneja, A. Davy, Lameness detection as a service: Application of machine learning to an internet of cattle, in: 2019 16th IEEE Annual Consumer Communications Networking Conference (CCNC), 2019, pp. 1–6. doi:10.1109/CCNC.2019.8651681
2. M. Taneja, **J. Byabazaire**, A. Davy, C. Olariu, Fog assisted application support for animal behaviour analysis and health monitoring in dairy farming, in: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), 2018, pp. 819–824. doi: 10.1109/WF-IoT.2018.8355141.

1.3 Thesis Structure

This thesis presents an end-to-end IoT application based on machine learning to predict lameness in dairy cattle. It discusses a novel approach to using the herd mean as a baseline, forming clusters based on activity levels of the animals and having a different model for each cluster as opposed to a one-size-fits all approach. An overall accuracy of 87% was achieved and three days prediction before visual signs are visible by the farmer. This thesis is organized as follows.

Chapter 1 gives a summary of the thesis, discusses the problem domain and also highlights challenges in the current systems. It also outlines the motivation and novel contributions of this research.

Chapter 2 discusses the background and current state-of-the-art. This is divided into 3 sections; section 2.2.1, the various technologies used in lameness detection, section 2.2.2, discusses those solutions that applied machine learning and finally section 2.2.3 gives an account of the underlying architectures used in all the previous solutions. The chapter concludes with a summary of the drawbacks of the current approaches which forms the basis of this research.

Chapter 3 contrasts 2 major software design approaches namely Monolith and Microservices. It gives an account of why microservices was used in the design approach of the application. It also discusses the data collection process and highlights the proposed system design.

Chapter 4 details the implementation process of the proposed system. It starts with the architectural design. Here, details of how various portions fit in together are discussed. Form how data on the sensors is transfered to the fog node, how it is processed at the fog node, transfered to the IBM IoT platform and finally to the cloudant database in the cloud. It the discusses the implementation of the clustering and machine learning models and then concludes with the design of the mobile application to facilitate feedback to the learning model.

Finally chapter 5 gives the results from the research and chapter 6 concludes and highlights directions for future research.

CHAPTER 2

BACKGROUND AND STATE OF THE ART

This chapter highlights the background information and the state of the art. Section 2.1 gives the background information on the three domain area which form the core of this thesis; lameness, fog computing and machine learning. In section 2.2, it then presents the state of the art solutions considered to relate to this problem. Since this is a multidisciplinary problem, these will be presented in three different subsections. Subsection 2.2.1 will present general lameness detection approaches. These will be further sub-divided depending on the technology used. In 2.2.3 a discussion of the under laying architectural designs used in various solutions from the networking point of view will be presented. This will mainly compare generic cloud based solutions and to the proposed fog-cloud based approach. Finally in subsection 2.2.2 will describe solutions that apply machine learning to solve the problem.

2.1 Background

2.1.1 Lameness

Lameness is a big problem in the dairy industry which farmers are not yet able to adequately solve. 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 considered 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]. The prevalence of lameness has been reported to differ from farm to farm and from one region to the other. Also different categories of lameness have been reported differently. Ger et al. [19] reported that on an average Irish farm, 20 in every 100 cows will be affected be lameness in a given year. In the United States authors in [20] and [21] reported a mean lameness prevalence of 25% , whereas in California and the northeastern United States, overall lameness prevalence was estimated to be 34% and 63%, respectively [22]. British and German studies reported a lameness prevalence of 37% and 48% [23, 24], whereas a prevalence of 16% was reported in the Netherlands [25]. In our experimental setup (Ireland), since its start in July 2017, up to 26 cases (out of 150 cows) of lameness have been recorded. Lameness can be classified into three main categories: solar ulcers, digital disease (white line abscess, foreign bodies in the sole and pricked or punctured sole), and inter digital disease (lesions of the skin between claws and heel including foul in the foot, inter digital fibroma and dermatitis). More than 65% of cases of lameness are said to be caused by diseases [26]. Other causes include injuries to the upper skeleton or major muscles, septic joints and injection site lesions.

Lameness has many negative effects, among these; reduction in feed intake, reduction in milk production (mainly due to withdrawal because of antibiotics) and weight loss. Lameness therefore has a drastic effect on the performance of a dairy farm. Lameness is mostly detected at advanced stage and thus requiring immediate and often costly treatment. Once an animal becomes lame, it can take several weeks to recover. Lameness thus represents a significant cost to dairy farmers in terms of time, financial expenditure for veterinary calls, medication and treatment, and also for loss in production. Prevention, early detection and treatment of lameness is therefore important to reduce these negative effects of lameness on dairy cows [4, 5]. Table 2.1 [19] summaries the costs estimates for each type of lameness.

Table 2.1: Costs associated with each type of lameness

(Ger et al, Economic cost of lameness in Irish dairy herds, Forage Nutrition, 2012)

Type of lameness	Digital	Inter digital	Solar ulcer	Average case
Prevalence (%)	45	35	20	
	Cost (€)	**Cost (€)**	**Cost (€)**	**Cost (€)**
Total cost of a single case	282.85	136.12	504.58	275.26

Traditional approaches for lameness detection are based on locomotion scoring (1-4 scale, there are other variations of 1-5) that requires observation of cows walking, preferably at the exit of the milking parlour [27]. This is not only a time-consuming task that can only be implemented on small size dairy farms [4] but also subjective. Therefore, these methodologies for lameness detection are based on a person's observations of cows walking to the milking parlour or standing at the milking stall. Observation of postural abnormalities predictive of lameness while cows are locked at stanchions is also used as an alternative detection method [6]. Figure 2.1[28] illustrates this process.

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 techniques like; Force Plate System [7], Accelerometer [9], Vision Based Trackway Analysis [8] and Shape Analysis which uses Image Processing Technique [4]. These techniques today are becoming popular in many commercial dairy farms to help in detection of lameness in cows. However, the data obtained from the experimental setup must still be translated into understandable gait variables that are functionally relevant for cattle gait (and therefore related to gait attributes used in locomotion scoring systems) [17].

The Introduction of Internet of Things (IoT) changed the way we relate with the physical world. By embedding sensors and actuators, physical things like animals, cars and buildings are able to communicate over the Internet. In the dairy industry, this gave raise to terms like Precision Dairy

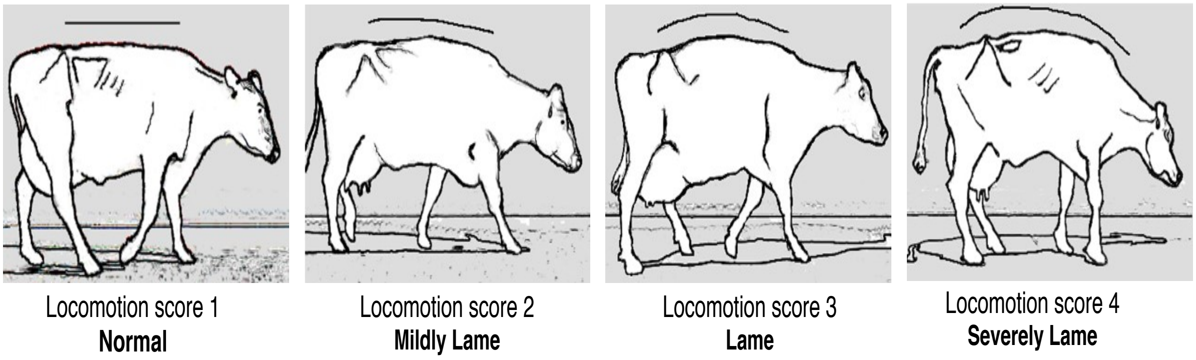


Figure 2.1: Locomotion scoring based on 1-4 scale
(Amstel et al, Manual for Treatment and Control of Lameness in Cattle, 2006)

Farming. Precision Dairy is the use of communications technologies to measure physiological, behavioural, and production indicators on individual animals to improve management strategies and farm performance. Using techniques like these, several management and welfare solutions have been developed. Recently there has been wide usage of activity sensors to predict estrous cycles in dairy cattle. The success of these and decreasing prices of sensors, prompted several scholars to investigate the applicability of the same in lameness detection.

However, in all these solutions, data aggregation and processing is either done on a local server or in the cloud. It is important to note that these sensors produce unnecessary data which for the case of a local server it would not be important to store all of it. On the other hand, for cloud solutions the disadvantages are many ranging from high bandwidth costs to transmit every byte of data produced per time interval, energy costs, security and of course delayed decision making in case one must be made.

2.1.2 Machine Learning

Machine learning is a subset of Artificial Intelligence that leverages data analytics techniques and statistical methods to give a computer system the ability to learn from data with minimal human interactions. Well as traditional computer systems require the programmer to explicitly specify and define all the necessary conditions, machine learning systems implicitly learn from data. The advantages and possibilities this could bring to smart agriculture are many. For example; we are be able to predict the trajectory of pests to create a quarantine zone, forecasting the demand and supply of agricultural commodities or even predicting diseases in livestock. Although there are many categories of machine learning algorithms, only two main categories will be discussed in this study; supervised and unsupervised learning.

Supervised learning

This form of machine learning uses training examples (input and out put) to predict the outcomes. In this instance, the user must provide the system with input and output pairs to train the system. Over time, the system can automatically construct outputs or targets for new data sets. There are two major types of supervised machine learning problems; classification and regression. Classification algorithms are a family of machine learning algorithms that output a discrete value. The output variables are sometimes called labels or categories. These kinds 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. For example predicting whether a cow is lame or not. On the other hand, the goal of regression algorithms is to predict a continuous number, or a floating-point. For example predicting house prices given the standard of living. Examples of algorithms in this family of machine learning include; Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbors (KNN), Decision tress.

Unsupervised learning

Unsupervised learning is a family of machine learning algorithms where there is no known output, no training for the algorithm. In unsupervised learning, the learning algorithm is just provided with the input data and then tasked to extract patterns from this data. It's not a matter of mapping the input to an output, but detecting more obscure trends in the input data. There is also a sub-set category known as semi-supervised, which combines unlabeled data and human-based training. There is no clear distinction between the sub-categories in this family of algorithms. For this work, we will discuss two sub-categories; Transformation algorithms and Clustering algorithms. Unsupervised transformation is used to create a new transforms of the data which might be easier for other machine learning algorithms to understand compared to the original formate of the data. Common examples of theses are dimensionality reduction algorithms like; Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA). On the other hand, clustering algorithms slit data into small groups with related features. Mostly used example in this case is K-mean.

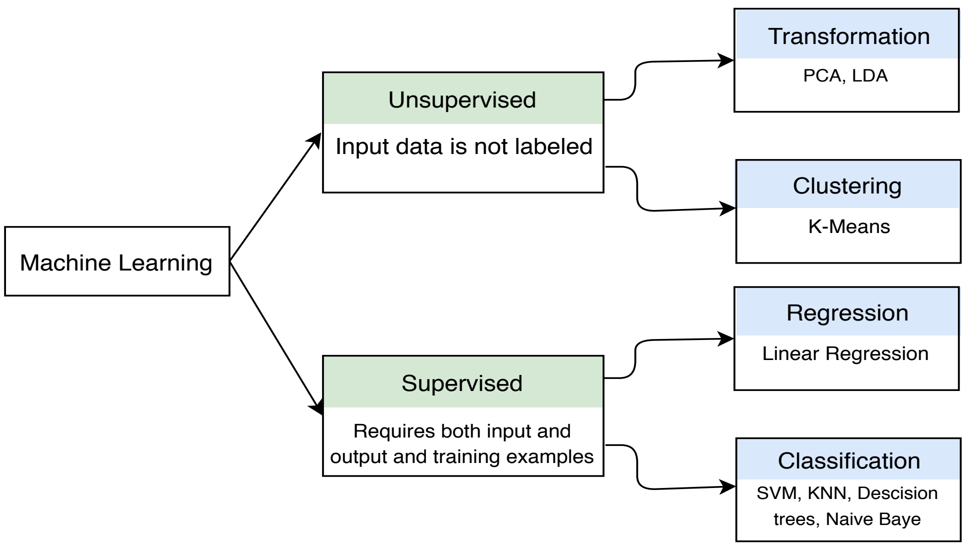


Figure 2.2: Hierarchy of machine learning

The choice of a learning algorithms is not at will but rather dependent on a number of factors. The size of the dataset, as some algorithms will work well on large datasets, of course the desired output, that is a discrete or continuous values, some will work well in highly dimensional data while others will fail, some might require the data to be normalized. The problem in this study required a discrete output and we experimented on a number of learning algorithms.

2.1.3 Fog Computing

Fog computing (or sometimes Edge computing) is a term that was introduced by Cisco. Because of the unprecedented volume and variety of data generated by IoT devices, there was need to act on the data closer to it's source. For time critical applications like in the transport sector, by the time the data makes its way to the cloud for analysis,the opportunity to act on it might be gone. Fog Computing extends the Cloud Computing paradigm to the edge of the network, thus enabling a new breed of applications and services [29]. Among the many advantages of fog computing promises, for the scope of this work two were explored;

- Providing intelligence and decision support faster and closer to source of data
- Only sends preprocessed data to the cloud for historical analysis and longer term storage

Authors in [35] reported the suitability of fog computing in context of IoT. While the primary and fundamental insights into data can be used for early event detections, the result of fog analytics can be further sent to the cloud for detailed analysis, and enhancing the learning patterns. The result of the cloud-based historical analysis can then be used to fine-tune and improve the analytics model for fog analytics [29].

2.2 State of the art

2.2.1 Technological approach

Traditional approaches for lameness detection are based on locomotion scoring (1-5 scale), this requires observation of cows walking, preferably at the exit of the milking parlour [5]. This is a time-consuming task that can only be implemented on small size dairy farms [4]. Therefore, these methodologies for lameness detection are based on a person's observations of cows walking to the milking parlour or standing at the milking stall. 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 locomotion scoring. This section summarizes these in three categories depending on the technology used; Pressure plates, Image/vision processing and Activity based techniques that use accelerometers and(or) pedometers.

2.2.1.1 Pressure Plate / Load cell

The main aim of these solutions is to investigate how the weight is distributed across the legs of the animal as it walks through or stands in a marked area. Pastell et al[30] used four strain gauge balances installed into a milking robot after careful inspection of the positions of the legs of the cows. The balances were connected to a four channel amplifier and the data transfered to a personal computer using a dedicated computer program. Preliminary analysis of the data gives evidence that limb and hoof disorders can be detected by the system. Because this was limited to a milking robot, the authors in [30] later changed the setup. A system consisting of a mat made of electromechanical film Emfit, which can detect only dynamic forces was used and this could be set up in any corridor along which the cows walk.

Neveux et al.[12] studied the use of a platform outside the automatic milking system to measure the weight distribution of cows while standing on different surfaces. In three different experiments, Holstein dairy cows were trained to stand on a platform that measured the weight placed on each limb. In later studies, different versions of the set-up was later used by Chapinal et al. [27, 31] and Pastell et al. [32] to measure lameness and hoof lesions. The asymmetry in weight bearing between left and right limbs (Leg Weight Ratio or LWR) was found to be a sensitive measure for detecting severely lame cows (AUC = 0.87). Pastell et al. [32] suggested that a cow may suffer pain when walking, which is not as obvious when the cow is standing still.

Authors in [33] developed a pressure-sensitive position mat which provides spatio-temporal and relative force information of two complete gait cycles of cows walking over a marked zone. The Gaitwise system which automatically measures the gait of the cows after milking without human interference, transforms raw data from the sensor into 20 basic variables that describe the general gait of cows walking over the marked zone together with an additional 10 more specific gait variables that are closely related to gait characteristics used in locomotion scoring systems. A postliminary study by Van Nuffel et al. [34] revealed that variables of speed and asymmetry measured by the Gaitwise system were closely related to locomotion scores given by observers based on videos of cows walking over the Gaitwise system and matched the Gaitwise variables. Initial results from these experiments show that they could correctly classify 84% of the cows as non-lame, mildly lame and severely lame cows with a sensitivity of 85%, 76% and 90% and a specificity of 86%, 89% and 100% respectively [33].

Rajkondawar et al. [7] developed a Reaction Force Detection (RFD) system. The system consists of two parallel and a floor plate that are each supported by four single-axis load cells. When a cow walks through the system, the load cell reaction forces are recorded as electric signals that change over time. They concluded that the system could recognize lame animals and identify the affected limbs. Because the duration of the experiment was short, they could only identify severely lame cows. They also developed a mathematical scoring system for lameness based on their RFD system. Their conclusive remarks indicate that with more data the system would have the ability to detect lameness in individual limbs.

Despite the many challenges solutions in this category face, a study [7] concluded that their system could recognize lame animals and identify the exact affected limbs. This would particularly be important in fully automated scenarios where a lame cow is isolated and the affect limb is automatically tagged. However in all the solutions in the above category, 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[17]. Because cows have a stoic nature, guiding them will bias the measurements, as they will try to hide their weakness and pain compared to measurements during normal routine without the presence of a human or predator. The other drawbacks of such solutions may not be only the costs of new and complex equipment but also other technical concerns. For example, in other setups, the animals are required to be standing in a certain postion. Pastell et al. [30] 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. Neveux et al.[12] also concluded that in cases were lameness affects more than one limb, it might not be apparent in the change in weight distribution.

2.2.1.2 Image/Vision processing techniques

This category studies the use of image processing techniques to analyse either the back posture or the legs of the animal either as it walks through a milking parlour or standing still. Zhao et al.[35] proposed a technique that uses leg swing analysis and computer vision to develop an automatic and continuous system for scoring the locomotion of cows to detect and predict lameness. Side-view videos were recorded after the cows were milked. The motion curve was plotted by extracting the position of the moving leg by image processing, and the motion curve was analyzed to generate six features referring to the gait asymmetry, speed, tracking up, stance time, stride length, and tenderness. Poursaberi et al. [36] studied how to automatically calculate the back arch of a cow as it walks by fitting a circle through selected points on the side view contour line of the back spine. The average inverse radius of four frames displaying the hind feet in contact with the ground was calculated for each cow. This approach was further explored and a Body Movement Pattern technique was used to describe the movement of the back and head of the cows using a side view [32]. An algorithm based on this Body Movement Pattern was tested under farm conditions by Viazzi et al. [13], who reported a correct classification of 81% and 91% using both thresholds, at the population level and at the individual level, respectively. .

Another study using the same techniques was applied to monitor animal behaviors. Yangyang et al. [37] developed a new deep learning method (i.e., an integration of background-subtraction and inter-frame difference) for automatically recognizing dairy calf scene-interactive behaviors (e.g., entering or leaving the resting area, and stationary and turning behaviors in the inlet and outlet area of the resting area) based on computer vision-based technology. Results show that the recognition success rates for the calf's science-interactive behaviors of pen entering, pen leaving, staying (standing or laying static behavior), and turning were 94.38%, 92.86%, 96.85%, and 93.51%, respectively. Similar techniques have also been applied in pigs with a high level of success [38, 39]

Compared to pressure plate/load cell, image processing techniques have a simpler setup and reduced setup cost. In cases were an animal might be lame on more than one limb, this does not affect the results of solutions in this category. Further study on this method however, shows that it still has challenges on real farm conditions. Some of these challenges were explored by Poursaberi et al. [14], Van Hertem et al. [40] and Viazzi et al. [13]; (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.

2.2.1.3 Activity based techniques

Here, techniques use accelerometers (2D and 3D) and pedometers 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. [41] 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). Chapinal et al. [9] used five 3D accelerometers on cows, one on each limb and concluded a single device attached to one of the legs appeared to be sufficient to measure the walking speed of cows, which was associated with locomotion scores. In other studies accelerometers were mounted on a hind leg on 348 cows in 401 lactations on four commercial farms [42]. Since then, a vast number of studies have used accelerometers to measure dairy cow activity and behaviour [43–46].

From this, many general welfare related and oestrus solutions have been developed [47, 48]. Although most agri-tech companies claim to do lameness detection based on this kind of technology, no such details have been made publicly or commercially available. The other challenge with the available solutions in this category is that they made and fine tuned for specific farm conditions and environments. For example, in Ireland animals stay in sheds during winter and in the fields during the summer. Reduced activity levels during the summer may be inductive of lameness yet the same will not be true during the winter. Some animals naturally have reduced activity level than others.

So, such solutions cannot scale or if they are introduced to new environments they will have a lot of errors. Therefore, they are not weather and environment agnostic.

2.2.2 Applied Machine Learning

Koen C et al. [49] used gaitwise measurements of spatiotemporal kinematic and force variables based on a pressure sensitive walkway. In their approach, the extracted parameters are fed to an artificial neural network in order to perform a lameness classification. This then performs an automatic classification of the degree of lameness in dairy cows. Ravi et al developed a technique that is based on body weight and artificial neural network. They concluded that to predict the heath status of cows RBF neural architect with Leven-burg Marquardt (LM) optimization algorithm, gave the best performance in comparison to MLP with highest classification accuracy rate (83.19%) at 80:20 data partition strategy. Dihua et al. [50] proposed a method based on YOLOv3 deep learning algorithm and relative step size characteristic vector to classify lame and non-lame cows. Videos were decomposed into sequence frames, and leg targets of cows in each frame were detected by YOLOv3 algorithm. A total of 210 videos were selected for verification using LSTM, support vector machine (SVM), K-Nearest Neighbour (KNN) and decision tree classifier (DTC) algorithms. Their showed that accuracy of lameness detection based on LSTM performed better.

2.2.3 Architectural view

From a network architectural view, there are three scenarios that can be considered(see figure 2.3 below). In figure 2.3a, On premises processing, here all data, algorithms are done at the farm. Figure 2.3b show a cloud based approach. Here the collected data is all sent to the cloud for processing. And finally figure 2.3c, fog-cloud based approach. Here some processing is done on the farm and the heavy processing is done in the cloud. It is important to note that each one of them has its own advantages and disadvantages. For example is case one, you might not need an Internet connect but miss the advantages that come with connecting to that cloud such fusion of farm data with say weather data which could improve the accuracy of the model. In scenario two, one might argue that there is a reduced setup cost because you do not need extra equipment at the farm. But you might miss the many advantages that fog computing promises. For example improved efficiency and reduced amount of data that needs to be transported to the cloud for processing, analysis and storage. Until now, all of the available solutions are based on scenarios one and two.

Fog computing itself is proposed for very time sensitive applications for example in transport and manufacturing were critical decisions have to be made in a second. In a farm environment we investigate the applicability of fog computing to solve two particular problems; reduced amount of data that needs to be transported to the cloud for processing, analysis and storage and we also show that it is relevant in network constrained environments where some processing might need to be done at the farm and notifications generated. In the agri-tech space, Bhargava et. al. [51] represents a fog computing technique for real-time activity recognition and localization on-board wearable IoT) devices. Other scholars have also applied the technique to precision agriculture [52–54].

2.3 Summary

Although most farms are equipped with some kind of oestrus detection system [55] which are based on accelerometers, lameness detection systems based on the same have not been successful. This is because of vendor lock-in. Each of the system would require its own hardware. Worth mentioning

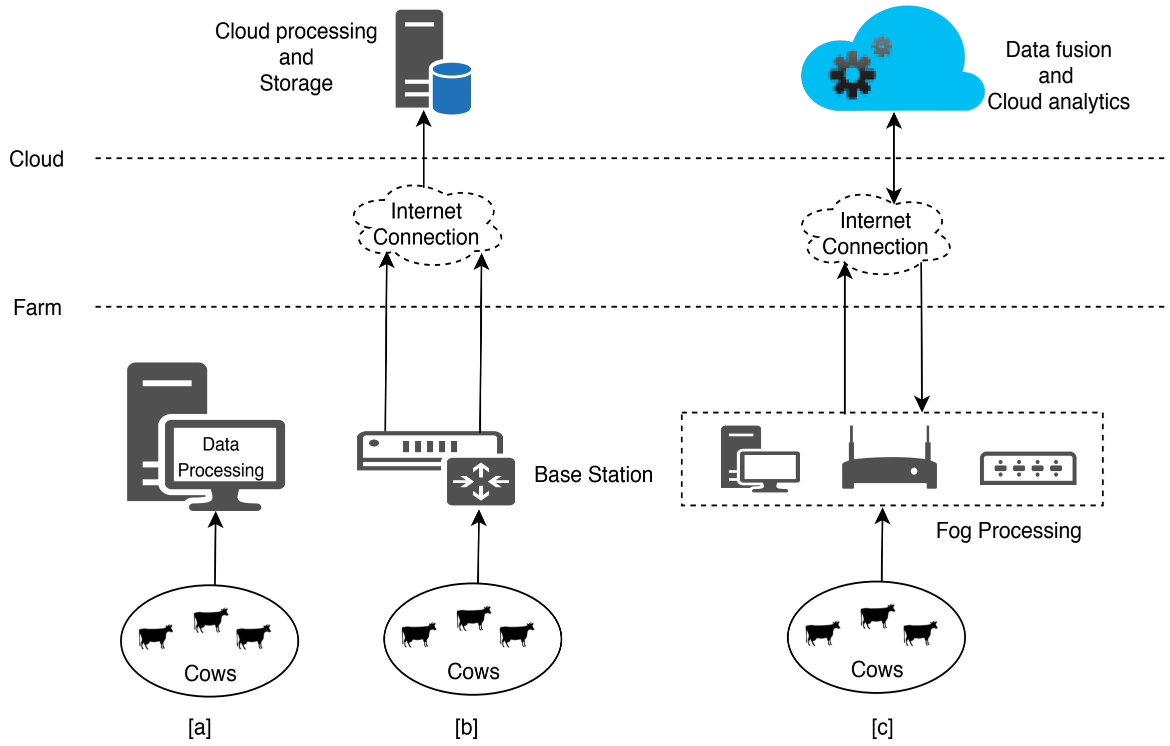


Figure 2.3: Comparison between three approaches.

again is the insight in the review by [16] that farmers who already had an oestrus detection system were willing to have an add-on for lameness detection. All the current systems luck this kind of integration. Another assumption made by all the current solutions is that all the animals will get lame the way. To put this into context, consider a farm in Ireland; There are basically two main seasons, the summer where animals stay in the field and graze freely and winter where animals are kept in house. In both cases the activity levels of the animals are different. Interestingly, in the second case they will relate to a lame animal in case one. Therefore a learning model should be able to consider such external factors. It is from the above argument that three research questions listed below were formulated;

1. Does considering individual cows or small subsets of the herd per particular model impact the overall accuracy of a learning model as opposed to a one size fits all.
2. What impact would fog computing bring to smart agriculture in terms of data exchange between the farm and the cloud.
3. What impact would fusion of external data sources like weather have on a lameness prediction and detection model.

CHAPTER 3

DESIGN METHODOLOGY

The pipeline of the end-to-end application designed in this study spans between the fog and cloud environments. The choices of the processes to run on either are made to ensure scalability and application performance. Each of the environments consists of three processes. The side has the following processes;

- Sensors: These are attached to the animals and collect data. They can also store data incase they are out of reach of the transmitter.
- Pre-processing: The sensors collect data inform of signals. Pre-processing helps convert the signal data to activities step count, which are later used for analysis.
- Clustering: This takes in activity data from the previous step and categories the animals into three clusters. This is essential as it determines the machine learning model selected in the cloud and the model configuration setting.

On the other hand, the cloud environment runs on the IBM cloud. This communicates with the fog environment using MQTT. This handles all the machine learning processing. Once a cluster is determined, an appropriate model is built. This is then evaluated with a human in the loop. If it does not meet the required accuracy, more training data is added, otherwise, it is deployed for prediction. Fig 3.1 summaries this process.

3.1 System Design

Some of the downsides of the current approaches are vendor lock-in, hardware specific implementation and all these hider collaboration and innovativeness. Therefore there was need to design and develop a system that is sensor agnostic, avoids vendor lock-in and can give access to multiple end users. To this end, two design approaches were compared.

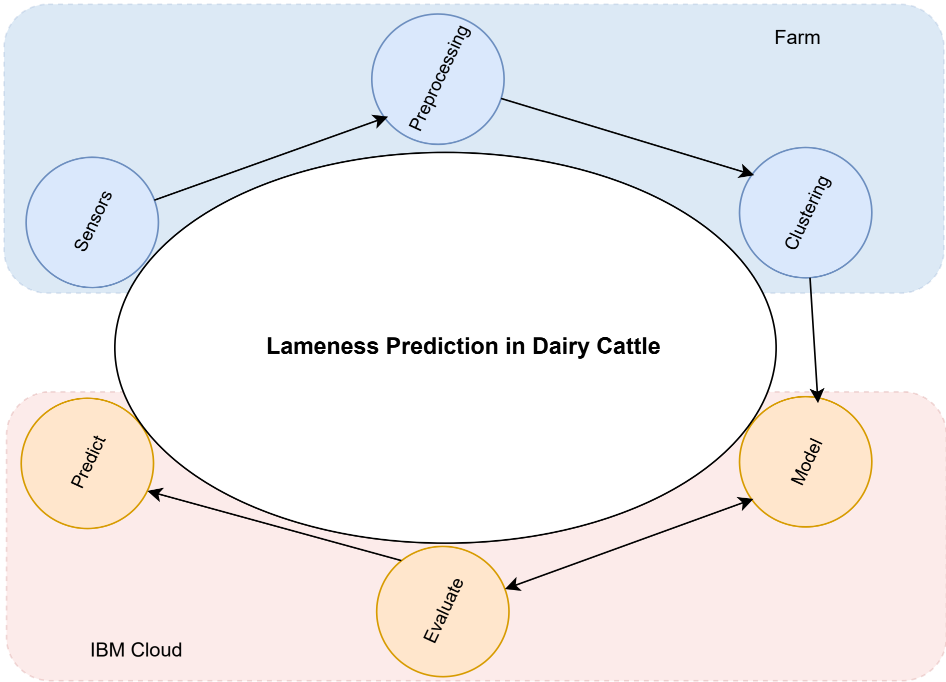


Figure 3.1: Theoretical flow of process

3.1.1 Monolith

A monolithic design approach, the application is structured as one unified model. Monolithic, in this context, means composed all in one piece. Monolithic software is designed to be self-contained; components of the program are interconnected and interdependent rather than loosely coupled as is the case with modular software programs. In a tightly-coupled architecture, each component and its associated components must be present in order for code to be executed or compiled. These are usually easy to debug and test, because they are interconnected, there is no need for API calls which improves performance. On the other hand, this has disadvantages underscored bellow;

- Monolithic applications are difficult to scale when different modules have conflicting configuration requirements
- Monolithic applications are a barrier to adopting new technologies. Since changes in frameworks or languages will affect an entire application, it is extremely expensive in both time and cost. This was important especially because this is still a growing area with dynamic key players.

3.1.2 Microservices

Microservices are a trend in software architecture, which emphasises the design and development of highly maintainable and scalable software. Microservices manage growing complexity by functionally decomposing large systems into a set of independent services [56]. Each microservice is a small self-contained application that has its own architecture consisting of implementation logic. Ideally, these microservices expose a REST and these also consume APIs provided by other services. This

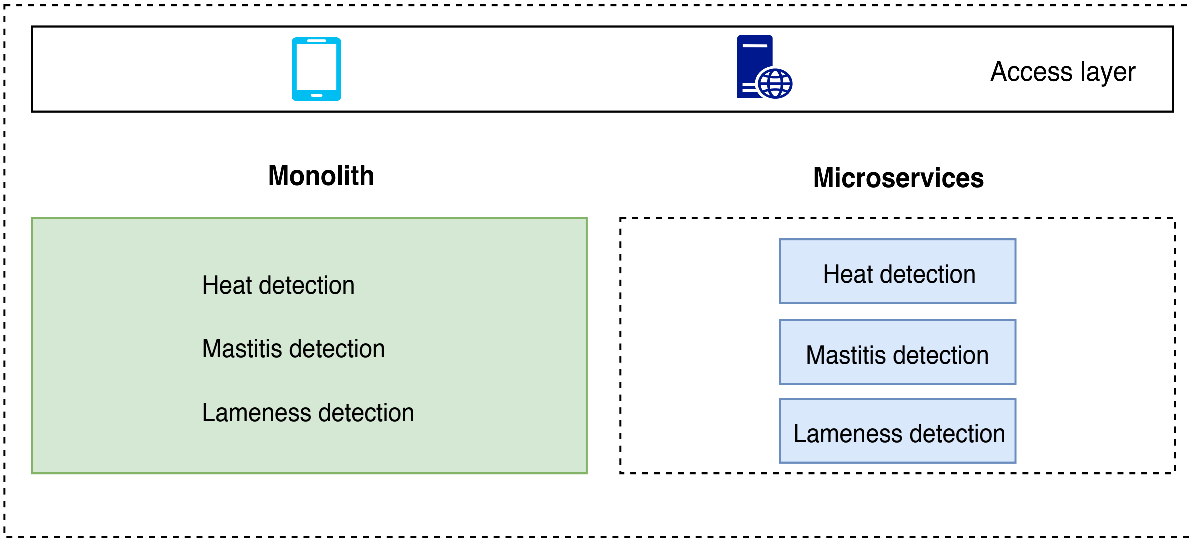


Figure 3.2: Difference between monolith and microservices

design approach ensures complex applications are broken down into small manageable services which are much faster to develop, and much easier to understand and maintain, an independent development can be assigned to a different service and let the other concentrate on other parts of the application. This software design approaches helps tackle the challenges highlighted above and several others. In the context of this work, microservices were considered because of three core reasons;

- Availability

  System availability is a measure of how often a system is capable of providing a service to its users. In a decoupled system, availability can be estimated in terms of the availability of the individual services that compose the system[56]. Authors in [57] conclude that, as size increases, so does a component's fault proneness. The goal of microservices in develop small components which in turn increases service availability.

- Maintainability

  Microservices implement a limited amount of functionalities, which makes their code base small and inherently limits the scope of a bug [56]. Moreover, since microservices are independent, a developer can directly test and investigate their functionalities in isolation with respect to the rest of the system.

- Scalability

  Microservices naturally lend themselves to containerisation [58], scaling a microservice architecture does not imply a duplication of all its components, and developers can conveniently deploy/dispose instances of services with respect to their load [59].

The only dependence imposed on a network of integrating microservices is the technology used to make them communicate, for example protocol . Apart from that, microservices impose no additional lock-in and developers can freely choose the optimal resources (languages, frameworks, etc.) for the implementation of each microservice [56]. This in turn leads to specialization. It also

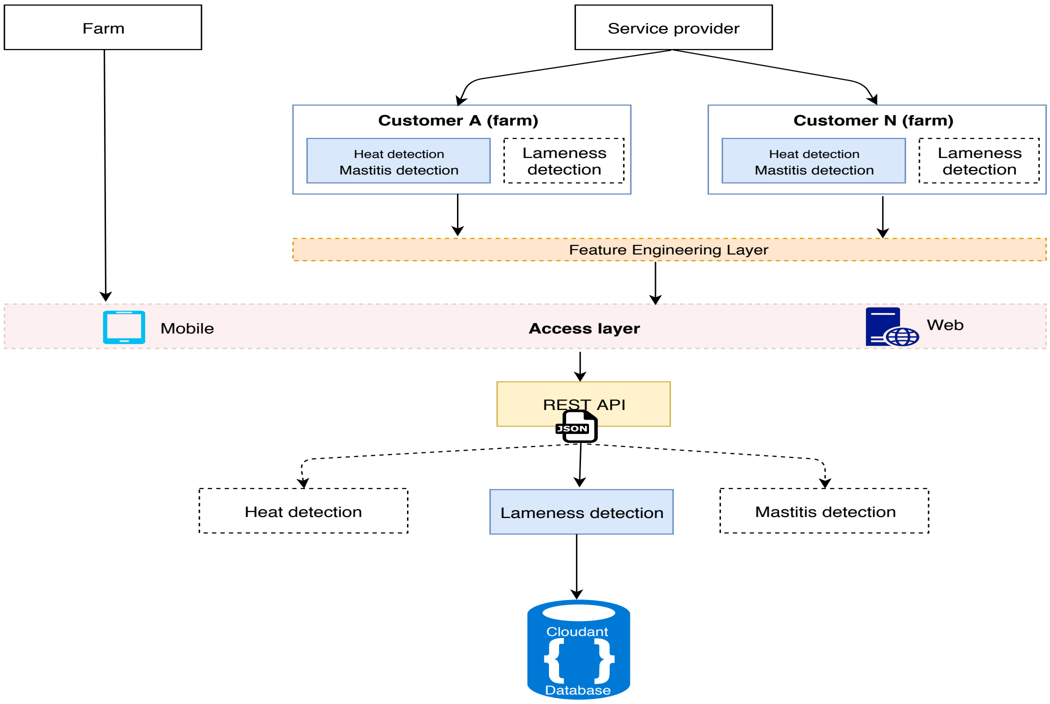


Figure 3.3: Proposed microservices design approach

tackles the problem of vendor lock-in because different service providers can consume services from other vendors. This leads to increased innovativeness since each provider will concentrate on what they can do best. Consider three common services provided by most agri-tech services providers that is; Heat detection, Mastitis detection and Lameness detection, the figure 3.2 shows the difference between monolith microservices design approaches.

3.1.3 Proposed System Design

Unlike all the available implementations that are based on a monolith design approach, the applications designed in this study follows a microservices approach[60]. 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 Figure 3.3 shows the design of the proposed system.

Figure 3.4: LRP attached as part of the experiment to the fore limb of the cows.

3.2 Data Collection

As part of the experiment, ethical approval was sought from Research Ethics Committee of Waterford Institute of Technology, Ireland prior to the deployment in July, 2017 (see Appendix 7). 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(50% on right leg and 50% on the left leg) of each cow, as shown in figure 3.4. 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. Also on-board the LRP is a memory with 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.

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[42]. 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 reported as lame (cows were checked for lameness by either an agricultural scientist or by the farmer). Because the number of none lame animals was small, splitting the data into training and testing folds was made in a such a way that atleast 75% of the lame animals was put in the training fold and the rest in the testing fold. This was a challenge as the dataset was imbalanced but, because this was a live experiment, we hoped to re-train the models after sometime. The initial performance on both the training and testing are reported in a later section.

CHAPTER 4

SYSTEM IMPLEMENTATION

This chapter presents the overall system implementation. The system was designed following a fog computing paradigm and each module is in turn designed as a micro service. In section 4.1, the overall architecture is presented, machine learning models are also discussed in this section and the design of the mobile app to aid farmer interaction and improve model learning is presented.

4.1 Design and Development

4.1.1 Architectural Design

The system is comprised of two functional units; the fog node, which is responsible for data preprocessing, data aggregation and sending critical alerts that need immediate attention. On the other hand, the cloud node is where most functions run; clustering of the herd, classification models, weather analytics and generating notification. Figure 4.1 show the overall system architecture.

4.1.1.1 Fog Node

As shown in Figure 4.1, 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; this is configured as (see appendix 2 for manufacture recommendations) - 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 12MB of unprocessed data. This is then preprocessed and aggregated to form behavarial activities. For this study, three of these are used for the analysis. Also on board the fog node is a dashboard which the farmer can interact with.

For communication between the fog node and cloud node, Message Queue Telemetry Transport (MQTT) [61] was used. This is an open source lightweight publish-subscriber model based protocol by IBM [62]. It is designed on top of the TCP-IP stack and it 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:

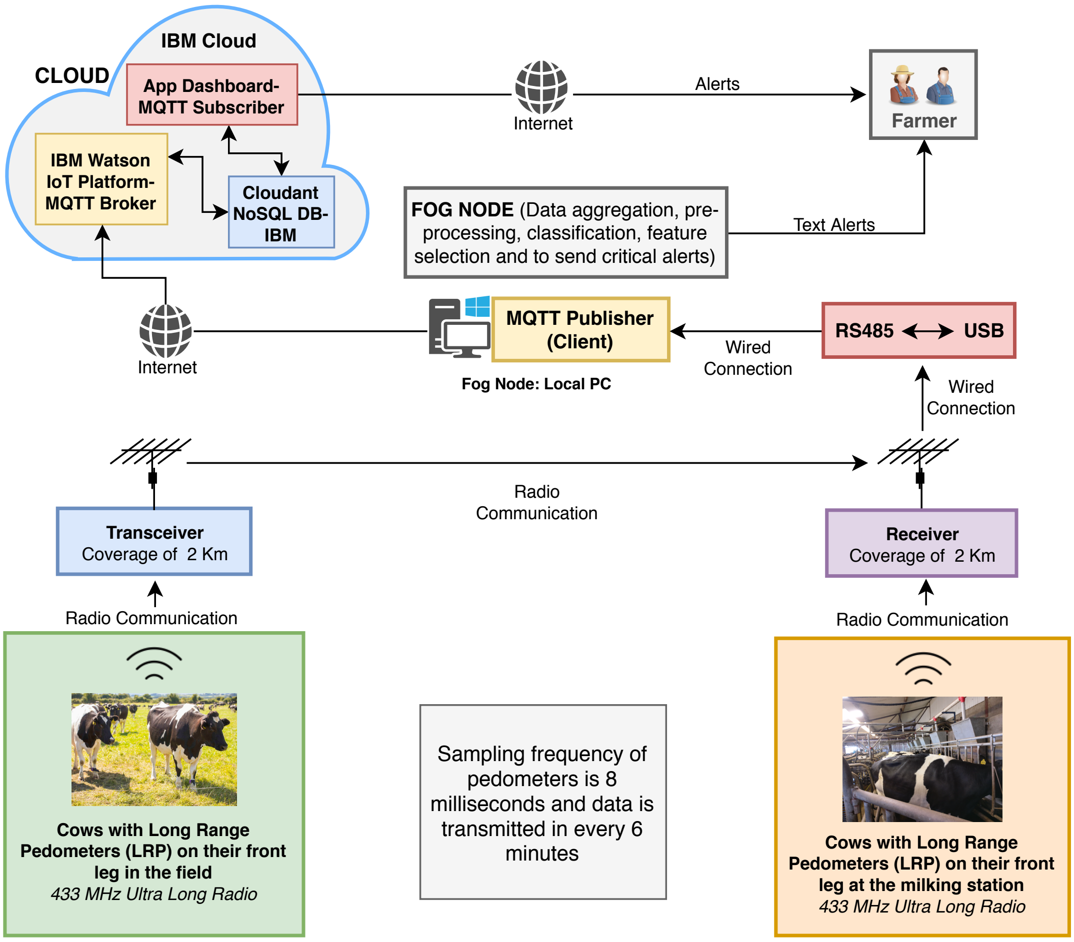


Figure 4.1: Overall systen architecture

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

4.1.1.2 Cloud Node

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. Watson IoT Platform facilitates powerful device management operations, storage and access to device data, connection of a wide variety of devices and gateway devices. It also provides secure communication to and from the devices by using MQTT and Transport Layer Security (TLS) [62]. In order for the fog node to communicate with the Watson IoT Platform, it is registered as a device onto the platform. A device can be anything that has a connection to the Internet and can push data to the cloud. The data is received in form of events by the IoT platform. Events are the mechanism by which devices publish data to the Watson IoT Platform. Devices control the content of their messages, and assign a name for each event that is sent. Figure 4.2 show a snapshot of the Watson IoT Platform as the fog node streams data.

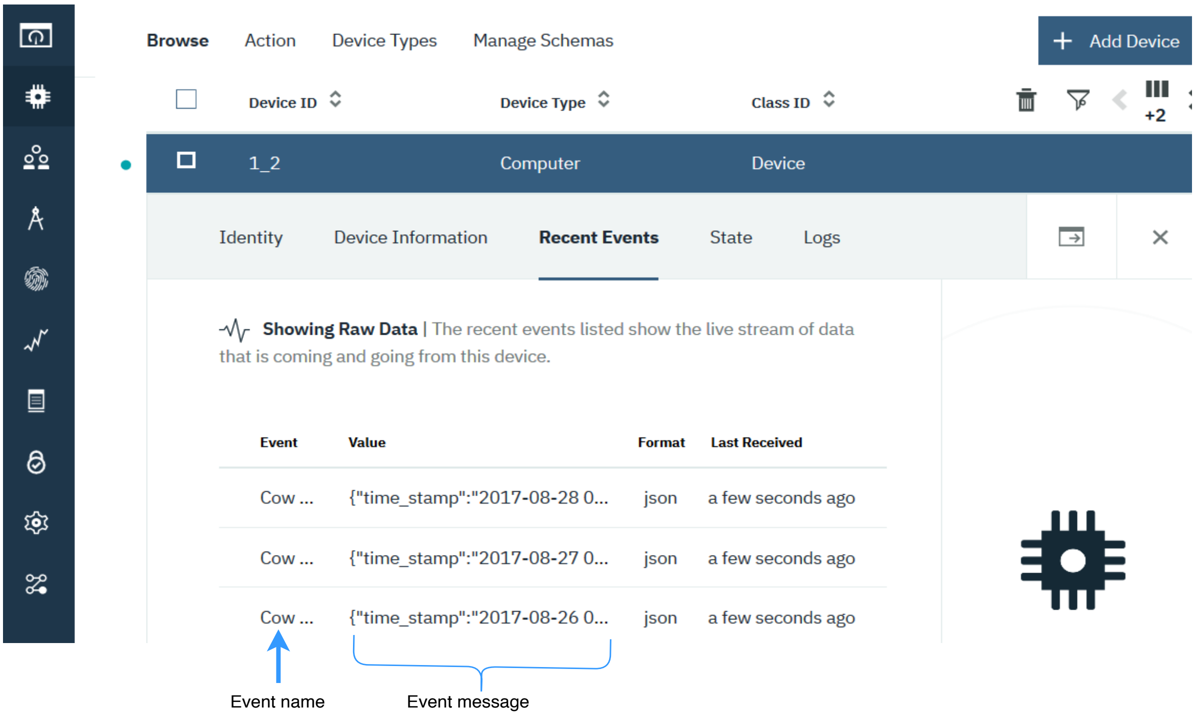


Figure 4.2: Snapshot of the Watson IoT Platform as data is being streamed

After the data is published at the Watson IoT platform, MQTT subscribers registered to these events pick up the event messages and saves them to the cloud database. Subscriber applications are anything that has a connection to the internet and interacts with data from devices and controls the behavior of those devices[62]. IBM Cloudant (cloud database) is a NoSQL JSON document store that is optimized for handling heavy workloads of concurrent reads and writes in the cloud; a workload that is typical of large, fast-growing web and mobile apps [63]. The cloud is also the site for fusion of the data from other sources, such as weather data.The Weather Company®, is an IBM Business that offers accurate weather forecasts globally with personalized and actionable insights. One of the services offered is the weather company data for advanced analytics which integrates weather-based insights from historical, current, and forecast conditions by accessing weather APIs. This data was also used to investigate the effects of weather on lameness. The overall work and data flow is show in figure 4.3.

4.1.2 Machine Learning

After the data is saved to the cloud database, then the next stage is to build and train the models. For the system to differentiate between normal and anomalous behaviour due to lameness, we must first form a cow profile to characterize normal behaviour. In this section we describe the underlying technique used in the formation cow profiles from the activity data. To that end, data for three activities; step count (The number of steps an animal makes per day), Lying time (Number of hours an animal spends lying down) and Swaps per hour (The number of times an animal moves from

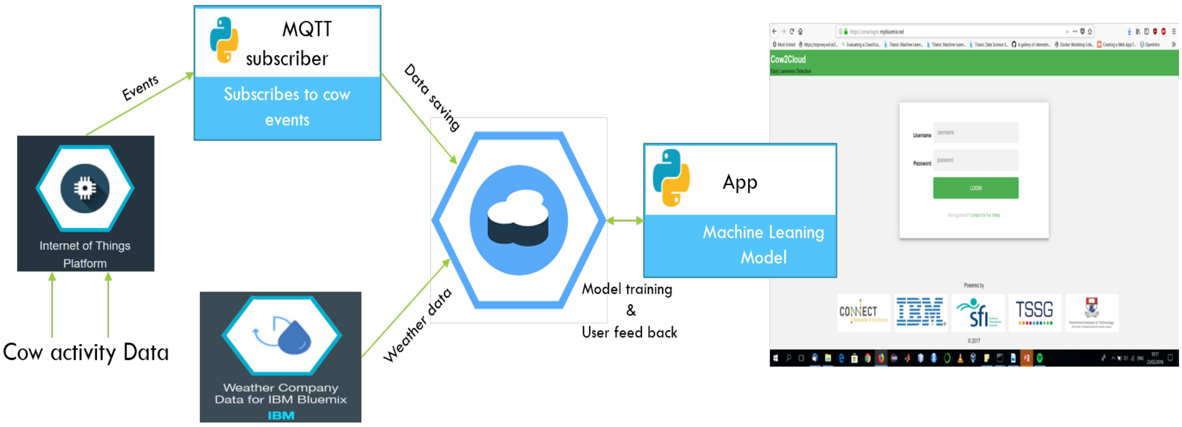


Figure 4.3: Work and data flow of the system

lying down to standing up) is used for form cow profiles. The statistical technique used to validate the differences between the cow profile and the transition from on profile to another is described.

4.1.2.1 Cow profiles

In order to build robust profiles that are distinguishable by the learning model, it is important 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, 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 is sometimes taken as a better measure. Figure 4.4 compares the mean and median of the herd. The results show that these almost trace out each other for all the three activities; lying time,step count and swaps per hour. Therefore it would not matter whether the mean or median is used.

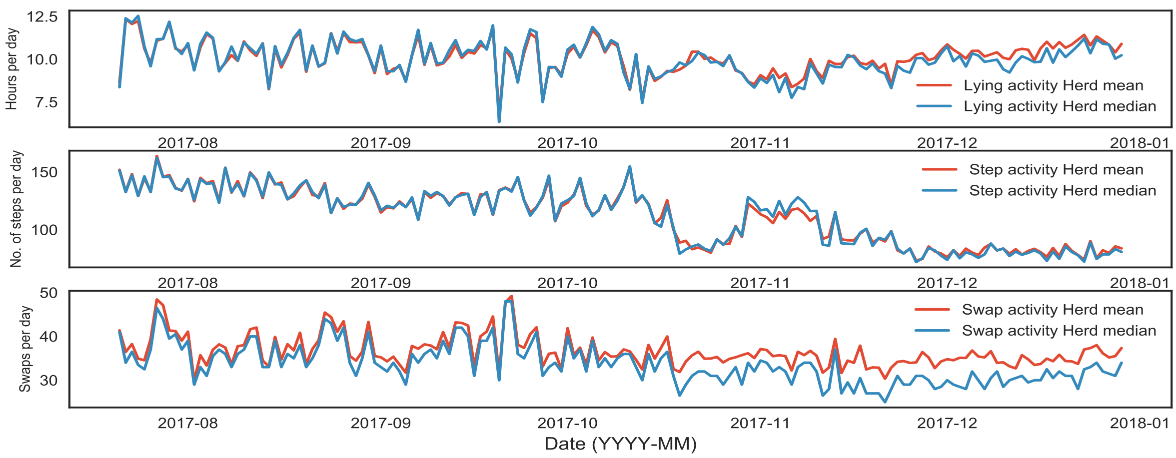


Figure 4.4: Comparing the Mean and Median of the various Animal Activities

Authors [64] in argued that animals grazing within the 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. It is also important to note that this will eliminate the effects of external factors as these will be affecting the whole herd and only leave the individual effects of lameness on the cow.

To define the Lame Activity Region (LAR) and the Normal Activity Region (NAR) as show in the figure 4.5 bellow, once a cow is identified as lame, we compare the herd mean for all the activities to the cow's activities and define a region $d_1 \leq D > d_2$, where d_1 is the day the activity starts to deviate from the herd activity mean, d_2 is the number of days after the cow is identified as lame and still lame, and D is the day the cow is identified as lame. The values of d_1 and d_2 will vary for each of the cows as others may have longer lameness cycles than others and depending when the cow is identified as lame. This is motivated by the fact that lameness is a transition from normal behaviour to lameness and back, it will probably start before it is seen and even continue after treatment until the cow becomes normal again. Once we define the LAR, the rest of the graph is treated as the NAR.

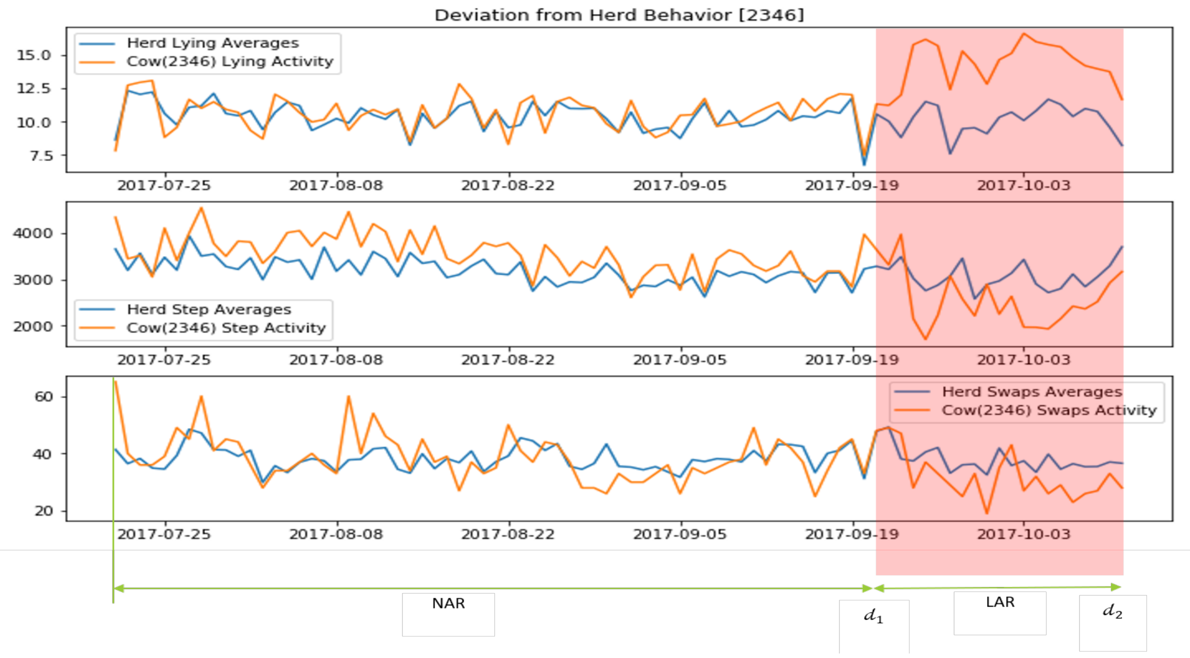


Figure 4.5: Relationship between herd mean and cow activity for 2346

Normal Profile,

To form the normal profile, we define a small window $\triangle wn$ within NAR for each of the normal cows and calculate mean absolute deviation N_{mad} for a given period of time.

$$N_{mad} = \frac{\sum_{n=\triangle wn}^{wn} |H_m - C_i|}{wn} \tag{4.1}$$

Where H_m, is the herd mean, C_i is the cow activity and wn is the window size of NAR.

Lame Profile,

To form the lame profile, we define a small window $\triangle wl$ within LAR for each of the lame cows and calculate mean absolute deviation L_{mad} for a given period of time.

$$L_{mad} = \frac{\sum_{n=\triangle wl}^{wl} |H_m - C_i|}{wl} \tag{4.2}$$

Where H_m, is the herd mean, C_i is the cow activity and wl is the window size of LAR.

The process is repeated for all the lame and non-lame cows. The results of this are plotted in a Density distribution plot.

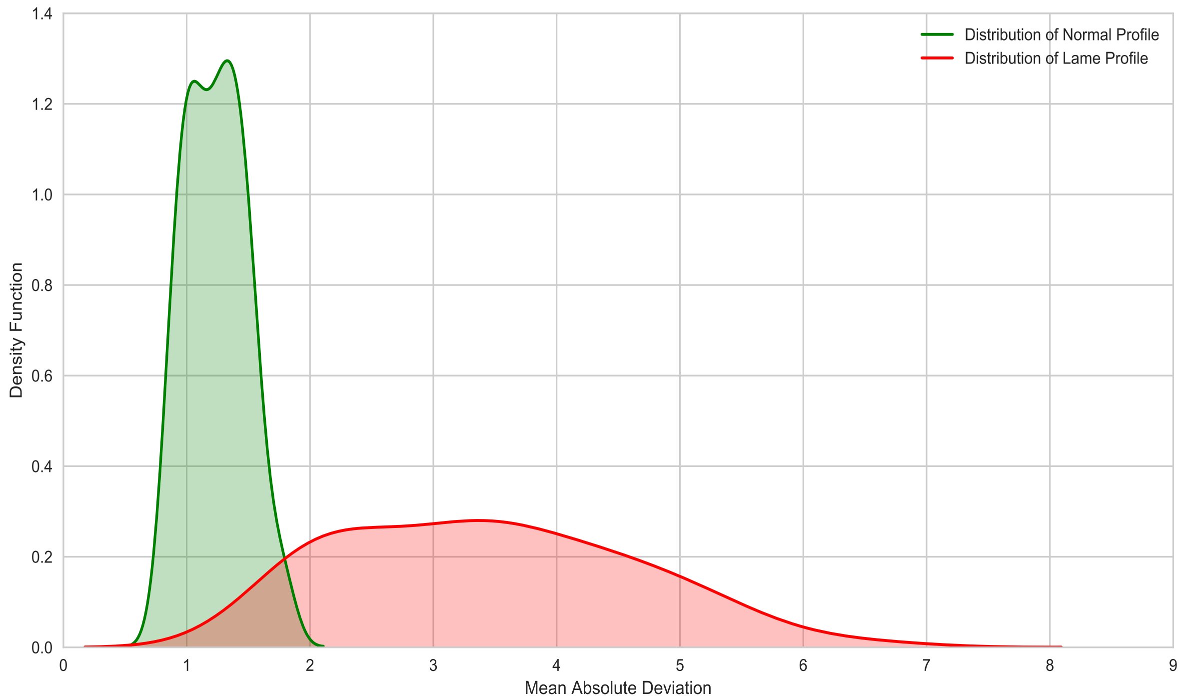


Figure 4.6: Density distribution plot comparing the Normal and Lame profiles

Relationship between individual cows and the Normal and Lame profiles

To test the viability of the profiles, randomly chosen cows that have at least been identified as lame at some point during the experiment where used (These were not used to form the profiles above). The goal was to go back in time and see how these relate to the constructed profiles before they get lame, when they are lame and when they transition back to normal behaviour. Using a equation 4.3, we define a window slice $\triangle ws$ (the optimal number of days was chosen after a repeative process) starting at the beginning of the experiment when the cow is not lame. We then slide through time as we calculate the Average Deviation (AD) within each $\triangle ws$ for each of the cows, and for each of the activity.

$$AD = \quad H_m - C_i \tag{4.3}$$

Where H_m, is the herd mean within $\triangle ws$ and C_i is the cow activity. This is repeated as we slide the window. The results of this are plotted as density distribution and compared with figure 4.6

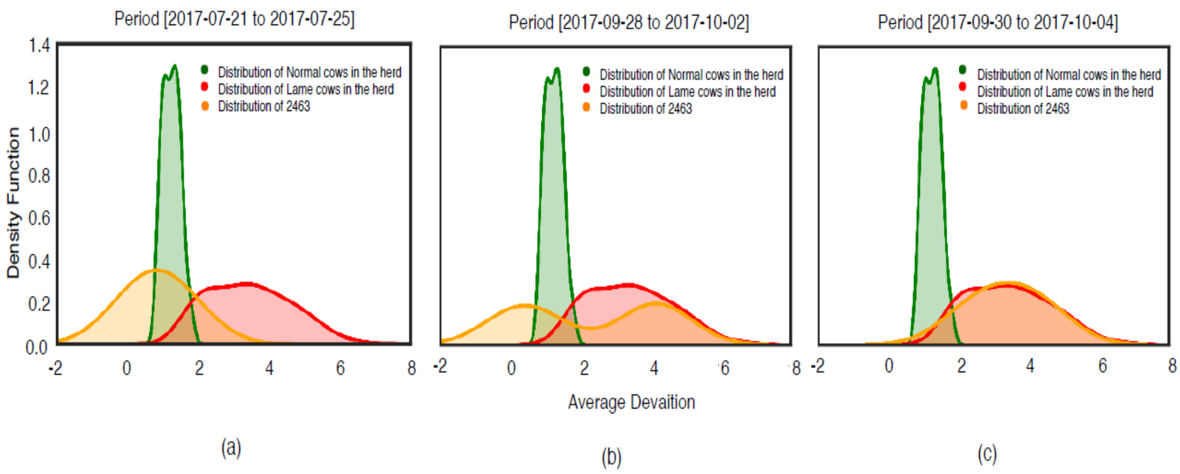


Figure 4.7: Comparing the distribution of cow 2463 against the normal and lame profile at three different stages

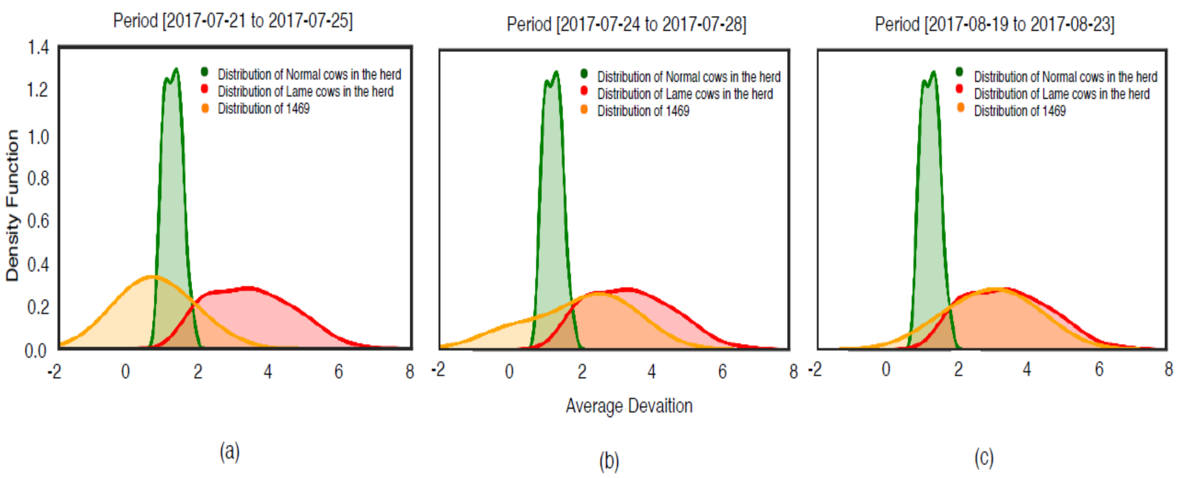


Figure 4.8: Comparing the distribution of cow 1469 against the normal and lame profile at three different stages

The three graphs (a,b and c) for both figures 4.7 and 4.8 show periods of transition from normal to lameness for the two cows 2463 and 1469. In figure 4.7 (a) and 4.8 (a), both animals are non-lame and the distributions relate to the Normal cows profile distribution. In 4.7 (b) and 4.8 (b) the distribution starts to shift to the right. Figure 4.7 (b) has two peaks. One relates more to the normal profile and the other to the lame profile. Figure 4.8 (b) on the other hand has one peak and this is mid way between both the normal and the lame profile. This kind of behaviour is justifiable because lameness is a transition. Perhaps this could be the best stage for the system to identify early lameness. Finally in figure 4.7 (c) and 4.8 (c), the distributions overlap with the lame profile distribution. It is important to note, even at this stage lameness is not yet visually detectable by the farmer for both cows.

Statistical Evaluation

The other way to investigate the relationship between the lame and normal profile was to conduct a statistical evaluation on the two distributions. Although from figure 4.6, it is clear that the two distributions are different, normally distributed, and that the normal cows have a small mean deviation as compared to the lame cows. There are many approaches that can be used to measure viability between two samples. Mainly T-Test and ANOVA are used. The student's T-Test (T-Test) compares the means while ANOVA compares the variance between samples. However for the T-Test, as the groups increases in number, one ends up with alot of pair plot comparisons. Well as ANOVA will give a single number (f-statistic) and one p-value. Also the T-Test among other things it requires the samples to have the same variance. For such reasons AONVA is usually preferred for large groups and varying variances.

Analysis of Variance ($ANOVA$) is a statistical method used to test differences between two or more samples. The null hypothesis is that there is no difference between two or more sample variances. The main assumption of the ANOVA is that the data is normally distributed. As highlighted above, this condition was met. Taking results from both equations 4.1 and 4.1 and setting $\alpha = 0.05$, we conduct a single factor ANOVA. The results of this are listed in the table 4.1 below;

Table 4.1: Results of the ANOVA on the Normal and Lame Profiles

Group	Counts	Sum	Average	Variance		
Lame Profile	78	267.7999	3.43332	1.348918		
Normal Profile	85	105.1237	1.236749	0.059152		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	Fcrit
Between Groups	196.2552	1	196.2552	290.3199	7.17E-38	3.899867
Within Groups	108.8354	161	0.675996			
Total	305.0906	162				

Since the F value is bigger than the Fcrit value, then we can reject the null hypothesis. Therefore, the Lame profile and the Normal profile are significantly different.

4.1.2.2 Clustering Model

From the above, it was 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:

initialization;
$x = MAD_{lying}$;
$y = MAD_{step}$;
$z = MAD_{swap}$;
$cluster$;
if *(x and y) < h || (x and z) < h || (y and z) < h* **then**
    $cluster \leftarrow NC$;
    return $cluster$
else
    reinitialization x,y,z;
    $x = MD_{lying}$;
    $y = MD_{step}$;
    $z = MD_{swap}$;
    if *(x and y)* < 0 || *(x and z)* < 0 || *(y and z)* < 0 **then**
        $cluster \leftarrow AC$;
        return $cluster$
    else
        $cluster \leftarrow DC$;
        return $cluster$
    end
end

Algorithm 1: Clustering algorithm

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.

The threshold was carefully chosen by a repetitive evaluation process. It's also important to note that these clusters are dynamic, that is the animals keep changing the clusters they belong to. This can be caused by many factors like age and weather. So it the role of the clustering model to keep regrouping the animals before selecting the appropriate classification model for that cluster (The best amount of time to re-cluster was found to be two weeks). Table 1 shows the distribution of the clusters as of writing. The total number used to build clusters was 146 as three of the animals where eliminated due other health related issues and one animal lost the tag during the experiment. This process is summarized in the algorithm 1.

Table 4.2: Distribution of the clusters

Active	Normal	Dormant
25	109	12

4.1.2.3 Classification Model

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. In general, to solve these kind of tasks, the learning model is usually tasked to produce a function $f : \mathbb{R}^n \longrightarrow \{1, ..., n\}$, where n is the number of labels. For example, let $\{X, Y\}$ denoted the data set (feature, label), θ the parameters, where;

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ . & . & . \\ . & . & . \\ . & . & . \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix} = \begin{bmatrix} Lyingtime & Stepcount & Swapsperhour \\ . & . & . \\ . & . & . \\ . & . & . \\ Lyingtime_{n1} & Stepcount_{n2} & Swapsperhour_{n3} \end{bmatrix}$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} Lame \\ Non-Lame \end{bmatrix}$$

When $y = f(x)$, decision function $f(x;\theta)$ together with a cost function $C(f(\cdot;\theta), X, Y)$ will assign an input depicted by vector x to a class label identified by y. It is important to note that there are other variants of functions f. For example f might out put a probability distribution as opposed to a class label. At the time of writing, the feature matrix X was made up 3 columns. Each of the columns is a feature. These were chosen because most literature suggests that they more representative of an animal transitioning to lameness or one that is already lame. The vector Y consists of labels 0 and 1, where zero indicates non-lame and one otherwise.

Since the activity data collected is time series based data, this makes such a problem different from traditional prediction problem. The addition of time as a constraint adds an order to the observations which must be preserved and can provide additional information for the learning algorithm. To retain this, and since we are interested in how the activity values relate to the herd mean as the cow transitions from normal behaviour to Lameness rather than values themselves, we take the Euclidean distance between the individual activities and the heard mean. This in turn also helps to scale the feature to avoid features with big values dominating those with very small values.

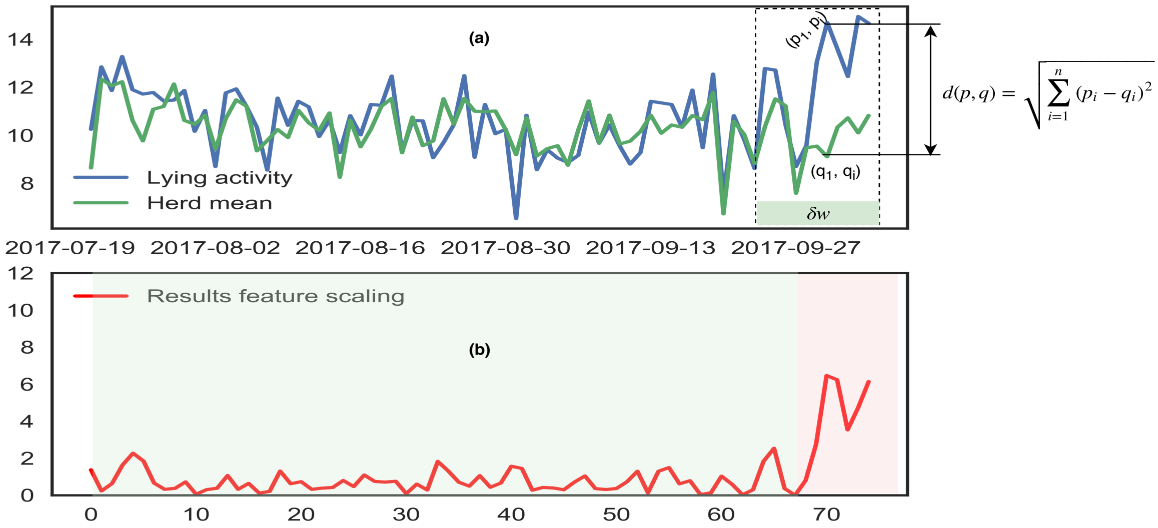


Figure 4.9: Feature transformation

We defined a small window δw (varied it's size until an optimal result was obtained) and calculate the euclidean distance between the points on the herd mean curve and the cow activity curve within the window. We then slide the window one day ahead. The process is repeated to the end of the curve. This is shown in figure 4.9 (a). The results of the above process is a more smooth curve that reflects the behaviour of both curves ass shown in figure 4.9 (b). We do this for all the three activities; Lying time, Step count and Swaps per hour.

To construct our training and testing dataset, we defined two regions on figure 4.9 (b). The region when the animal was not lame (green shed, NAR) and the region when the animal was lame (reg shed, LAR). The sets of values (S_n) within the green shed form the negative examples and the sets of values (S_l) within the red shed form the positive examples. And therefore combining both sets gives us the training and testing dataset for the learning model. It's however important to note that the training and testing sets S_n and S_l is different for each cluster (Active, Normal and Dormant).

The problem was formulated as a binary class problem with Lame as being the positive class and Non-lame as the negative class. The training process reported in this study is unique from the previous approaches because it has a feedback loop added. After model evaluation, an agricultural expert or farmer, re-annotates training data to improve the model accuracy. We experimented on a number of sklearn[19] classification algorithms ranging from Support Vector Machine (SVM), Random Forest (RF), K-Neighbors (K-NN) and Decision trees. Although a different model was trained and built for each of the three clusters, results reported in chapter 5 are only for the normal cluster as it not possible to efficiently evaluate the rest of the other clusters because testing data in these was very small. Figure 4.10 shows the overall work and data flow of training and testing of the models.

4.1.3 Offline first

Because these are usually in locations that are geographically remote, farm settings in most cases are network constrained areas. Therefore most of them are faced with limited cellular coverage. Existing solutions are mostly cloud based or completely offline. This limits the farmers' ability to interact with the application anytime anywhere. The applications designed in this work uses an offline first strategy via the mobile application and cloud dashboard. Once the model produces

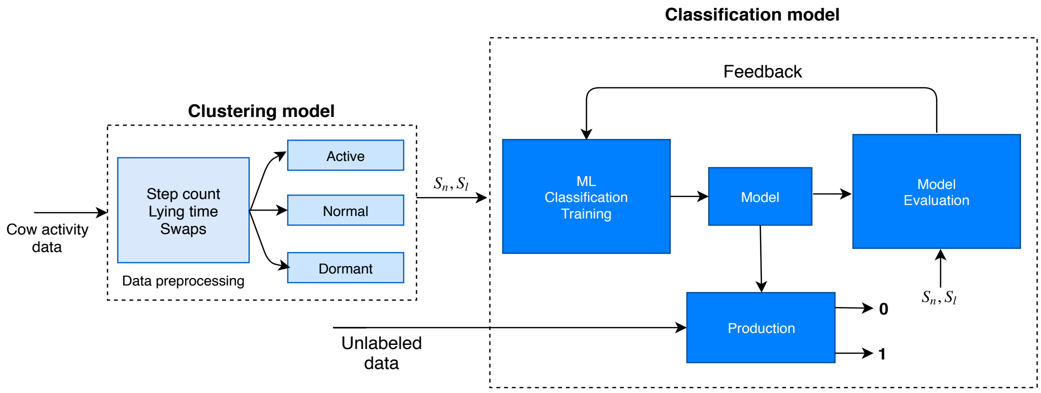


Figure 4.10: Model data and work flow

notifications, these are sent to the farmer's mobile device using push notification. On board the application is pouchdb which synchronizes with the cloud cloudant database using a REST API whenever a connection is established. The application in general helps to achieve the following tasks;

- Push notification, whether on WiFi or limited cellular network or the application is open or not, these will go through each time the status of the farm changes.
- Data annotation, during the training process, it was used by the human operator to annotate the data. This was done weekly by an agriculture student.
- Feedback to improve model learning, when a notification is generated, the farmer has the option of confirming if the said cow is actually lame, tag it as a false alarm or even report a missed alert. All this information is sent back to the model to improve it accuracy. Figure 4.11 show the data flow of the off line first design approach.

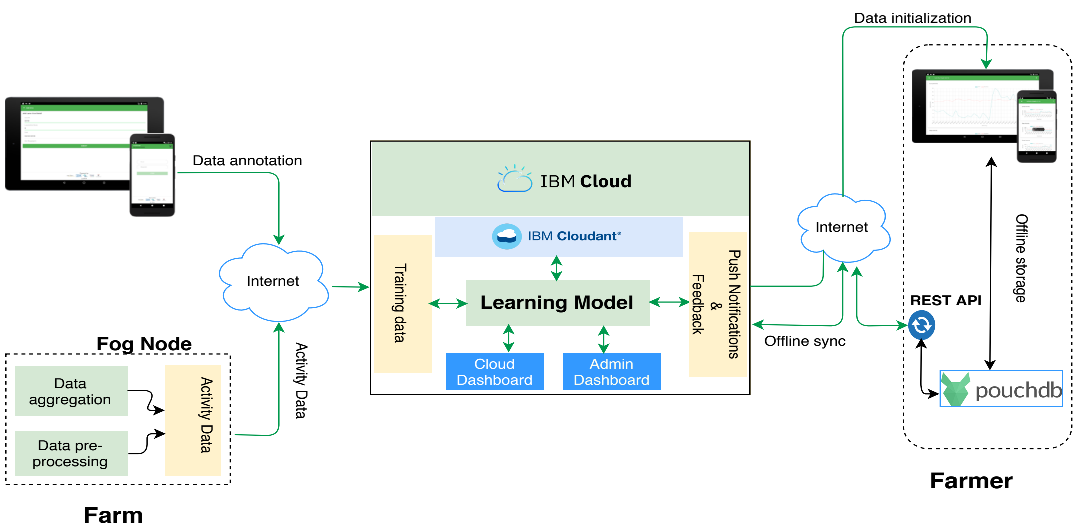


Figure 4.11: Offline first data flow diagram

Chapter 5

Discussion of Results

5.1 Effects of weather on lameness

Weather and climate influence both farm animal production and agronomic production [65]. Some scholars have studied how weather in general affects the prevalence of lameness among dairy cattle. Nigel et. al. [20], looked at the prevalence of lameness among dairy cattle in Wisconsin as a function of housing type and stall surface. In this study, he looked at the prevalence of lameness during the summer and winter. His results show that the Mean ± SD herd lameness prevalence was 21.1 ± 10.5% during the summer and 23.9 ± 10.7% during the winter. Authors in [66] studied the Correlation between lameness, rainfall and soil moisture and concluded that these had less effect on lameness as compared to other factors. Other scholars [67] looked how other external factors relate to lameness, like housing and other management practices. Although no one has studied how specific weather conditions affect lame dairy cows, some authors have suggested that incorporating weather data in a model to predict Lameness would improve the accuracy of such a model.

This study investigated how weather conditions (Temperature, Wind and Humidity) affected lame dairy cows. We calculate Pearson's Correlation coefficients between weather conditions and the animal activity (Lying time, Step count and swaps per hour).

Pearson Correlation Coefficient

Pearson's correlation coefficient measures the strength of a linear relationship between paired data. In a sample it is denoted by r and is by design constrained as; $-1 \le r \ge 1$

Interpretation:

1. Positive values denote positive linear correlation;
2. Negative values denote negative linear correlation;
3. A value of 0 denotes no linear correlation;
4. The closer the value is to 1 or –1, the stronger the linear correlation.

For the results of Pearson's correlation coefficient, a number of assumptions have to be tested. These inlcude; variables should be normally distributed. To test for this, the Shapiro-Wilk test for normality was used. Since the lying activity, step activity, swaps activity had already been tested above, here only results of temperature, wind and humidity are reported. Setting $\alpha = 0.05$, each of the variables was tested. If the any of the variables has a p values greater than the α , then they are normally distributed. From the tests, wind had the highest $p = 0.898$, followed by temperature with a $p = 0.861$ and humidity the lowest with $p = 0.377$. The other assumptions like no significant outliers, continuous, linear relationship and homoscedascity were also checked. The paired scatter graphs below show the relationship between the variables.

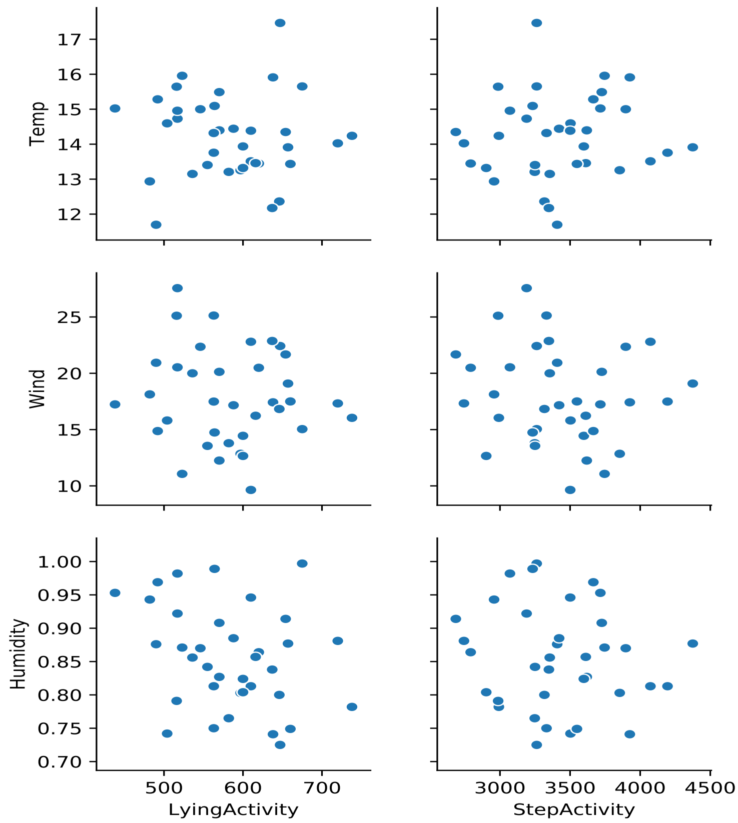


Figure 5.1: The paired relationship of the variables

All lame cows had an increase in correlation between weather (Temp, Wind and Humidity) and the activity (Lying time, Step count and swaps. There was an increase of 14.1%, 25.5 and 68.2 for temperature, wind and Humidity respectively against Lying time of Lame cows. Also, generally weather affects the activity of the cows hence slight increase in the correlation in Normal cow's graph. Figure shows the results of comparing some of the individual lame cows to the entire Normal group.

5.2 Rationale of the Clustering model

In a study about association patterns of visually observed cattle, Stephenson et al [64] 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

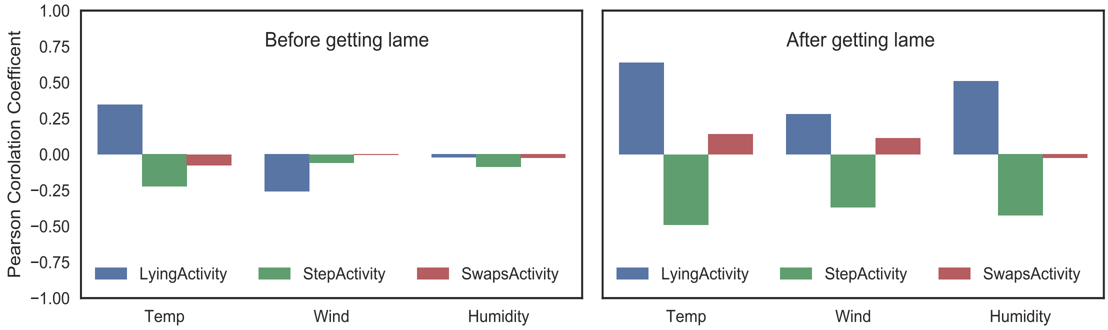


Figure 5.2: Effects of weather on 1469 before and after getting lame

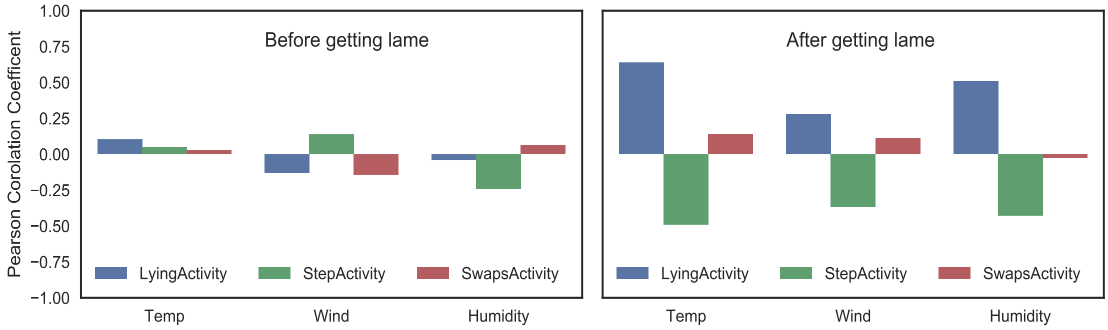


Figure 5.3: Effects of weather on 2213 before and after getting lame

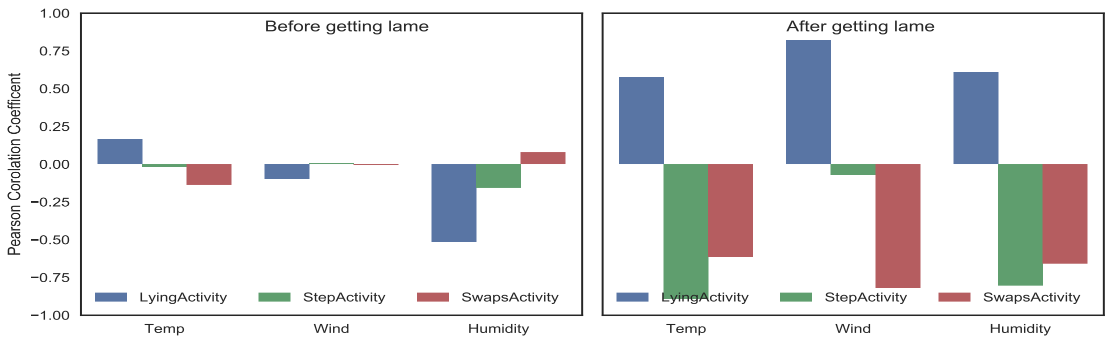


Figure 5.4: Effects of weather on 1348 before and after getting lame

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 5.5 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.

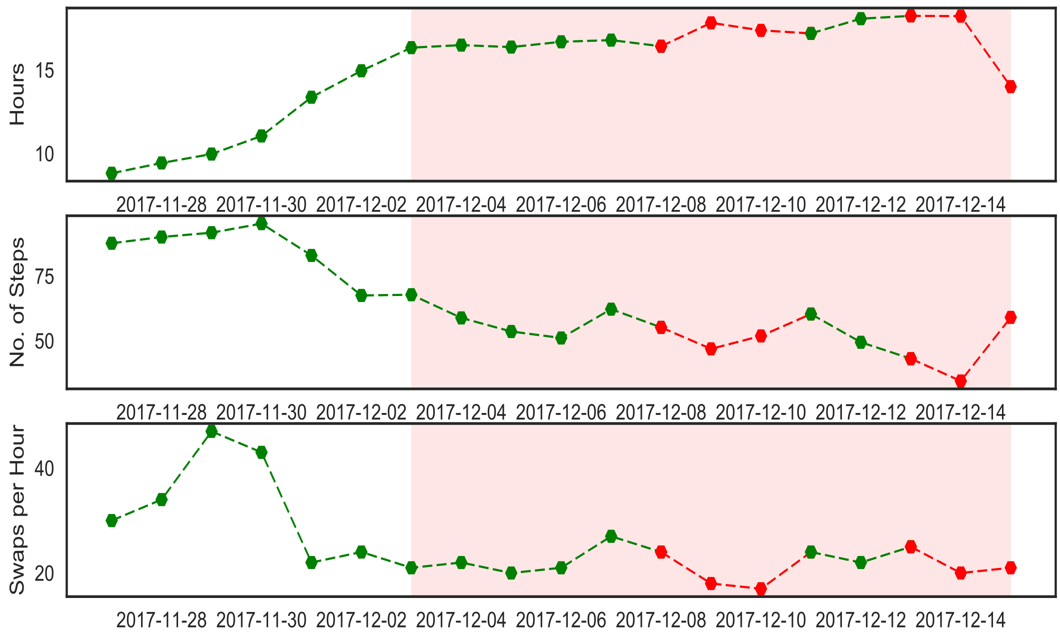


Figure 5.5: Animal confirmed as lame between 03/12/2017 to 15/12/2017 but could not be correctly identified by a one-size-fits-all model

5.3 Fog-Cloud data reduction

Some of the downsides of all the current approaches is that they are either fully cloud based, that is all the data is sent to the cloud cloud for processing of fully on premises, that is all the processing is done on the farm. The disadvantages of the first approach are many but the scope of this work focused on the reduction of data exchanged between the cloud and the devices (cows). The disadvantage of the second approach are the opportunities missed that come with fusion of data from different sources, for example weather data. This work applied fog architecture 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. This aspect of data reduction becomes even more crucial while scaling up the farm and the herd, as the amount of data collected and streamed would rapidly increase [68]. The system was also able to benefit from processing and notification at the fog node. Figure 5.6 show the daily data reduction between the fog and cloud node.

5.4 Resource utilization at the fog node

IoT solutions are usually deployed on resource constrained infrastructure from the sensors to fog nodes. On the side of the sensors, the mitigation is to take away any kind of processing from them. For fog nodes, to increase resilience of the system is to have coexistence of multiple fog nodes in the infrastructure, which are synchronized with each other, and data can be redirected between them in case of failure or overload [68]. In this system setup, one fog node was used. To ascertain the resilience of this and the scalability of the developed solution, it is necessary to determine how resources mainly in terms memory and CPU are utilized. The main processing component amongst the microservices

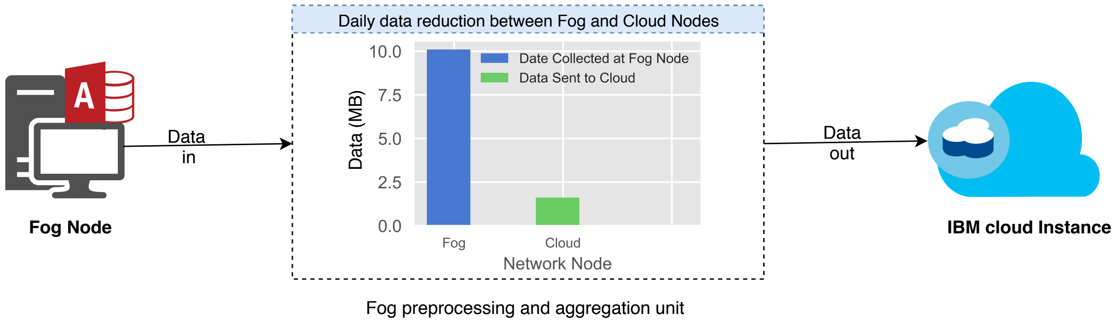
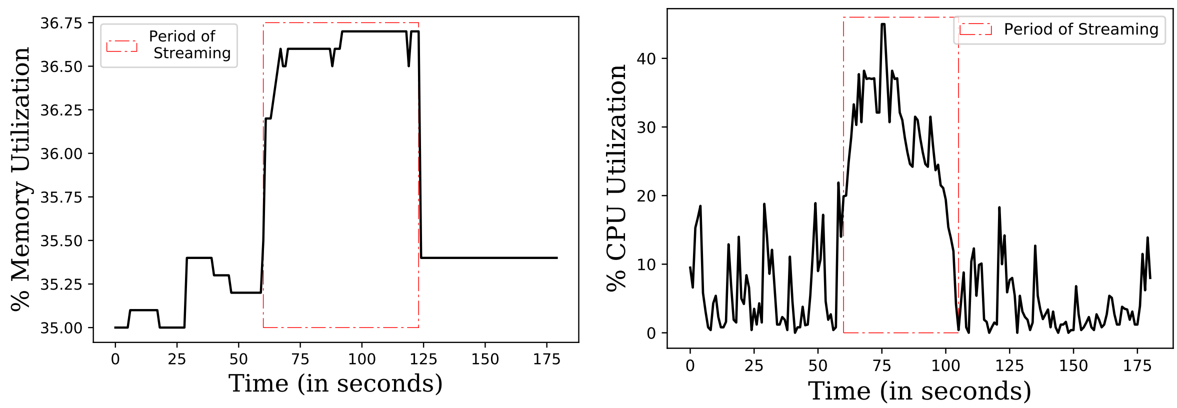


Figure 5.6: Daily data reduction between the fog node and the cloud node



(a) Memory utilization at fog node

(b) CPU utilization at fog node

Figure 5.7: Resource utilization at fog node

running on fog node is MQTT publisher which is responsible for data pre-processing, aggregation, streaming processed data to the cloud. We monitored resource (CPU and Memory) consumption on fog node in a time window of every 3 minutes when all the microservices were running on it. Using psutil (process and system utilities) [69] python library, we calculated the percentage increase in both CPU and Memory before, during and after streaming (i.e. before the MQTT publisher was run,during and after it was run). We used the percentage increase because this would form a baseline here i.e. irrespective of change in the fog node capabilities, the same increase in resource consumption would be expected.

Figure 5.7a, shows memory usage before, during and after the MQTT publisher is run. The region highlighted shows the period during which the MQTT publisher was streaming. As it can be seen, there is an increase from about 35.38% to 36.75%. The total usable memory was about 14.50GB. This would mean the impact of MQTT Component was only 0.12GB.

Figure5.7b, shows CPU usage before, during and after the MQTT publisher is run. The region highlighted shows the period during which the MQTT publisher was streaming. As it can be seen, there is an increase from about 20% to 45% in CPU utilization.

In all, the overall resource utilization and effects of MQTT component on the fog node are very minimal for it run on very resource constrained compute environments.

5.5 Lameness

We experimented on a number of sklearn[69] classification algorithms raging from Support vector machine, Random forest, K-Neighbors and Decision trees while noting the accuracy and number of days the classifier could catch anomalies before they could be seen by the farmer. 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. In all, the normal cluster model had a sensitivity of 89.7% and specificity of 72.5%.

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 (see table 5.1), which is the number of neighbors to consider while assigning the nearest class. We set the Distance metric to Minkowski with a power parameter of 2. The highest accuracy was obtained with $k = 2$ although this was over fitting the data.

Table 5.1: Settings of K-value and accuracy

K-value	Accuracy(%)	Number of Day
2	91	1
3	89	2
4	87	3
5	81	1

Optimal results were obtained with $k = 3$ and $k = 4$ which gave accuracies of 89% and 87% with 2 days before and 3 days before the visual signs could be seen by the farmer. So there was a trade off between accuracy and number of days before the visual signs.

RF

In order to setup a successful RF, two parameters need to be set up: the number of trees ($n\_estimators$) and the number of features in each split ($max\_features$). We used all these in their default state with $n\_estimators = 10$ and $max\_features = \sqrt{[}2]p$, where p is the number of features.

To the best of my knowledge, this is the first attempt to implement lameness detection based on custom models in turn based on activity levels of animals as opposed to a one-size fits all approach. Although several studies have shown that there exists independent groups within a herd. For example authors in [64] concluded that movement patterns of sub-sets of individual tracked cows may have levels of independence that are sufficient for analysis as individual experiment units. In figure 5.8 cow with ID could only be visually identified as lame on 07/11/2017 yet the model could pick it up as early as 04/11/2017 as shown by the red points. The model was also tested on historical data and it was able to show previous lameness cycles as shown by the red points in figure 5.9

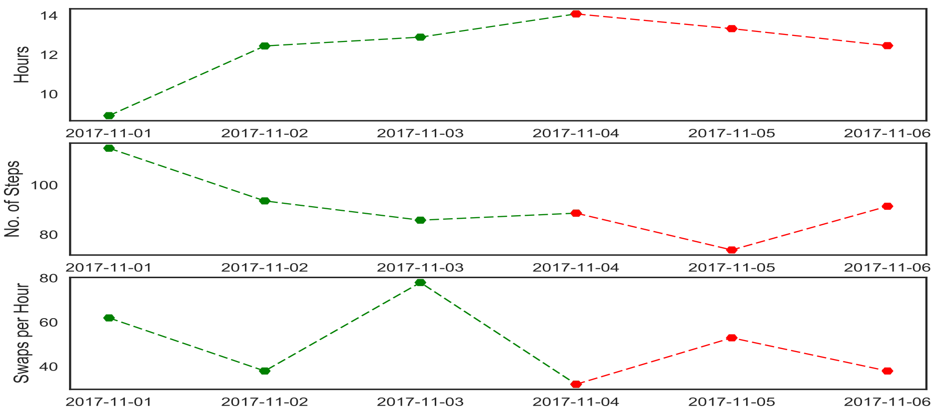


Figure 5.8: Red points indicating lameness anomalies for cow ID 1674

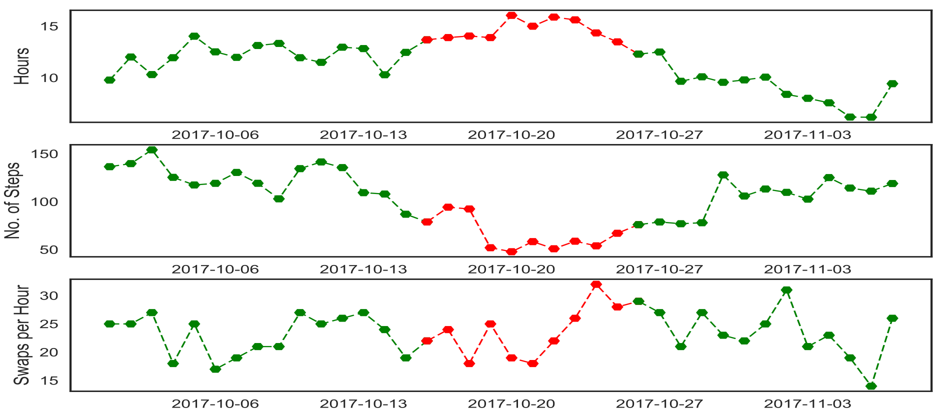


Figure 5.9: Red points indicating lameness anomalies identified in historical data for cow ID 1988

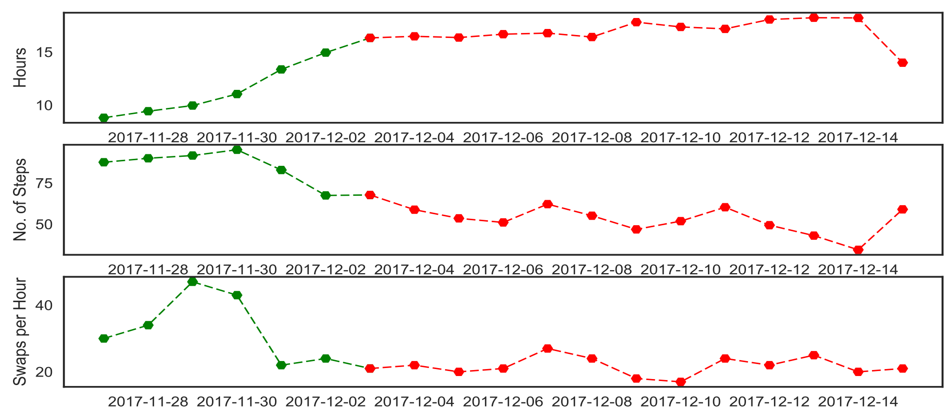


Figure 5.10: Red points indicating lameness anomalies for cow ID 2350

In figure 5.10 cow with ID 2350 was identified as lame as early as 03/12/2017 yet the farmer could only identify the same on 06/12/2017 and the locomotion scoring scientist could identify the same on 07/12/2017.

Mobile App

This is built with an offline first approach so that the farmer can keep up even in the worst network constrained environments. The app also provides the following features;

- Push notification, this to ensure that the farmer does not miss an alert whether the app is open or not.
- Interactive activity graphs. The farmer can monitor and visualize the activity of the cows on his mobile device, see figure 5.11

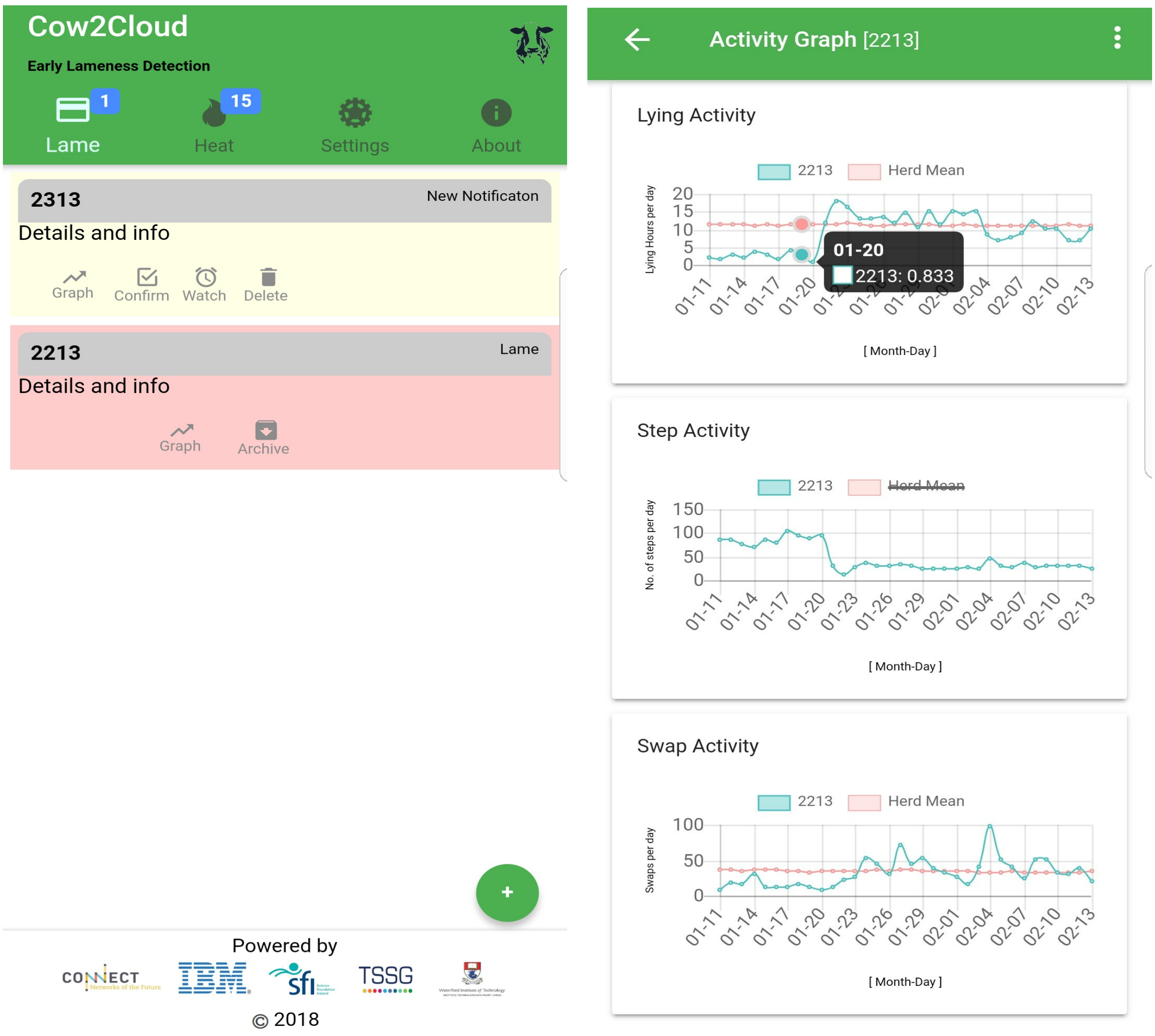


Figure 5.11: Mobile app screenshot with interactive graphs

Chapter 6

Summary

The work presented in this thesis describes how an end-to-end IoT application to predict lameness in dairy cattle was built. It is implemented as a fog enabled application and packaged as a microservice to avoid vendor lock-in, be sensor agnostic and have multiple end users. Data from the sensors is sent to the fog node using radio waves where it is preprocessed and aggregated. This is then sent to the IBM cloud using MQTT and stored in IBM cloudant NoSQL database. The study uses a novel technique to form animal profiles which eliminates effects external factors like weather and local farm conditions. This facilitates the formation of clusters in the herd to build cluster specific models for lameness prediction as opposed to a one-size-fits all approach used in all previous studies.The clustering technique forms three clusters and a machine learning model is learned for each of the clusters. An overall accuracy of 87% is achieved with K-NN with a classification error of 11.7%. This gives a prediction of three days before visual signs can seen by the farmer.

6.1 Conclusion

The aim of this thesis was to build and evaluate an early lameness prediction and detection system. Several approaches had been studied previously by other scholars. These include; Pressure Plate / Load cell techniques, Image processing techniques and activity based techniques. These have high initial setup costs, lack inter-vendor compatibility, require the animals to be in certain positions which biases the results and other technical challenges as highlighted above.

While evaluating the farmer's preferences for a lameness detection system, authors in [16] highlighted that farmers preferred activity based solutions that could easily be integrated with other existing or future systems. The findings in [70] indicate that considering small subsets of animals with similar behaviour for analysis and model building has the potential of improving the overall results of the learning model. It is from this that the basis of this thesis was formed.

In 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 (left leg for 50% of the cows and right leg for the other 50%) of each cow is aggregated at the fog node and analyzed in the cloud for lameness anomalies. The initial analysis looked at fussing external data sources like weather and evaluating the effects this has on lameness. The results of showed that weather did have an effect on lame cows as had been reported by other studies. Although this was the case, the technique used to form animal profiles eliminates the effects of external factors likes weather, location and farm conditions. This study is the first to offer a lameness prediction and detection service that is population and environmental agnostic.

From the animal profiles, it was discovered that animals behaved differently while normal in three distinct clusters and also got lame differently in their respective clusters. These are; Normal, Dormant and Active clusters. Although the existence of clusters in herds have been used before in cattle behaviours, this study is the first to use cluster specific model for lameness prediction as opposed to a one-size-fits-all.

Based on open source machine learning algorithms, several machine learning models were built and evaluated on the data collected. The results showed that the best two performing were 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. Although a different model was built for each of the clusters, it's only the normal cluster model that was evaluated because the two did not have enough lame cows for evaluation.

6.2 Future work

The work in this thesis shows that building custom models for small groups of animals that share similar features within the herd especially as the the herd size increases improves the accuracy of the lameness detection as opposed to a one-size fits all approach. 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. As the accuracy of the clustering model improved so did the lameness classification model. Although a good classification accuracy was obtained, the following key things need to be investigated more;

1. A more robust clustering technique as the current one is only based on threshold. Also expand the lameness detection model to do locomotion scoring automatically.

2. Since the clustering model would relate so much to the individual farm as the application scales, it could be possibly moved to the fog node as opposed to running in cloud.

3. In this work, only one aspect relates to fog computing, that is data reduction between the fog and cloud node was investigated. Also other aspects like Network resource allocation and bandwidth in smart agriculture could be investigated.

Bibliography

[1] K. A. O'Callaghan, P. J. Cripps, D. Y. Downham, and R. D. Murray, "Subjective and objective assessment of pain and discomfort due to lameness in dairy cattle," *Animal Welfare*, vol. 12, no. 4, pp. 605–610, 2003.

[2] N. Chapinal, A. de Passillé, D. Weary, M. von Keyserlingk, and J. Rushen, "Using gait score, walking speed, and lying behavior to detect hoof lesions in dairy cows," *Journal of Dairy Science*, vol. 92, no. 9, pp. 4365–4374, 2009.

[3] "Dairy Cow Mobility and Lameness - AHDB Dairy. Accessed on 1st Nov 2016."

[4] A. Poursaberi, C. Bahr, A. Pluk, D. Berckmans, I. Veermae, E. Kokin, and V. Pokalainen, "Online lameness detection in dairy cattle using Body Movement Pattern (BMP)," *International Conference on Intelligent Systems Design and Applications*, no. 1, pp. 732–736, 2011.

[5] M. Alsaaod, J. Niederhauser, G. Beer, N. Zehner, G. Schuepbach-Regula, and 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*, vol. 98, no. 9, pp. 6236–6242, 2015.

[6] A. García-Muñoz, G. Vidal, N. Singh, and N. Silva-del Río, "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*, vol. 128, pp. 33–40, 2016.

[7] P. G. Rajkondawar, U. Tasch, A. M. Lefcourt, B. Erez, R. M. Dyer, and M. A. Varner, "A system for identifying lameness in dairy cattle," *Journal of Agricultural Safety and Health*, vol. 18, no. 1, pp. 87–96, 2012.

[8] X. Song, T. Leroy, E. Vranken, W. Maertens, B. Sonck, and D. Berckmans, "Automatic detection of lameness in dairy cattle-Vision-based trackway analysis in cow's locomotion," *Computers and Electronics in Agriculture*, vol. 64, no. 1, pp. 39–44, 2008.

[9] N. Chapinal, A. de Passillé, M. Pastell, L. Hänninen, 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. 2895–2901, 2011.

[10] "IceRobotics | IceTag 3D™ (https://www.icerobotics.com/products/)."

[11] "StepMetrix (http://agromontnitra.sk/produkty/paznechty-dojnic-a-oviec/),"

[12] S. Neveux, D. Weary, J. Rushen, M. von Keyserlingk, and A. de Passillé, “Hoof Discomfort Changes How Dairy Cattle Distribute Their Body Weight,” *Journal of Dairy Science*, vol. 89, no. 7, pp. 2503–2509, 2006.

[13] S. Viazzi, C. Bahr, T. Van Hertem, A. Schlageter-Tello, C. E. B. Romanini, I. Halachmi, C. Lokhorst, and D. Berckmans, “Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows,” *Computers and Electronics in Agriculture*, vol. 100, pp. 139–147, 2014.

[14] A. Poursaberi, A. Pluk, C. Bahr, W. Martens, I. Veermäe, E. Kokin, J. Praks, V. Poikalainen, M. Pastell, J. Ahokas, A. Van Nuffel, J. Vangeyte, B. Sonck, and D. Berckmans, “Image based separation of dairy cows for automatic lameness detection with a real time vision system,” in *American Society of Agricultural and Biological Engineers Annual International Meeting 2009, ASABE 2009*, vol. 7, pp. 4763–4773, 2009.

[15] D. M. Weary, J. M. Huzzey, and M. A. G. Von Keyserlingk, “Board-invited Review: Using behavior to predict and identify ill health in animals,” *Journal of Animal Science*, 2009.

[16] T. Van De Gucht, W. Saeys, A. Van Nuffel, L. Pluym, K. Piccart, L. Lauwers, J. Vangeyte, and S. Van Weyenberg, “Farmers’ preferences for automatic lameness-detection systems in dairy cattle,” *Journal of Dairy Science*, vol. 100, no. 7, pp. 5746–5757, 2017.

[17] A. Van Nuffel, I. Zwertvaegher, S. Van Weyenberg, M. Pastell, V. M. Thorup, C. Bahr, B. Sonck, and W. Saeys, “Lameness detection in dairy cows: Part 2. Use of sensors to automatically register changes in locomotion or behavior,” 2015.

[18] A. Balalaie, A. Heydarnoori, and P. Jamshidi, “Microservices Architecture Enables DevOps: Migration to a Cloud-Native Architecture,” 2016.

[19] C. C. V. Ger and C. of XLVets), “Economic cost of lameness in Irish dairy herds,” *Forage & Nutrition*, 2012.

[20] N. B. Cook, “Prevalence of lameness among dairy cattle in Wisconsin as a function of housing type and stall surface,” *Journal of the American Veterinary Medical Association*, vol. 223, no. 9, pp. 1324–1328, 2003.

[21] L. Espejo, M. Endres, and J. Salfer, “Prevalence of Lameness in High-Producing Holstein Cows Housed in Freestall Barns in Minnesota,” *Journal of Dairy Science*, vol. 89, no. 8, pp. 3052–3058, 2006.

[22] M. von Keyserlingk, A. Barrientos, K. Ito, E. Galo, and D. Weary, “Benchmarking cow comfort on North American freestall dairies: Lameness, leg injuries, lying time, facility design, and management for high-producing Holstein dairy cows,” *Journal of Dairy Science*, vol. 95, no. 12, pp. 7399–7408, 2012.

[23] H. R. Whay, D. C. J. Main, L. E. Green, and A. J. F. Webster, “Assessment of the welfare of dairy caftle using animal-based measurements: direct observations and investigation of farm records,” *Veterinary Record*, vol. 153, no. 7, pp. 197–202, 2003.

[24] Z. E. Barker, K. A. Leach, H. R. Whay, N. J. Bell, and D. C. J. Main, “Assessment of lameness prevalence and associated risk factors in dairy herds in England and Wales,” *Journal of Dairy Science*, vol. 93, no. 3, pp. 932–941, 2010.

[25] J. Amory, P. Kloosterman, Z. Barker, J. Wright, R. Blowey, and L. Green, "Risk Factors for Reduced Locomotion in Dairy Cattle on Nineteen Farms in The Netherlands," *Journal of Dairy Science*, vol. 89, no. 5, pp. 1509–1515, 2006.

[26] C. Foditsch, G. Oikonomou, V. S. Machado, M. L. Bicalho, E. K. Ganda, S. F. Lima, R. Rossi, B. L. Ribeiro, A. Kussler, and R. C. Bicalho, "Lameness Prevalence and Risk Factors in Large Dairy Farms in Upstate New York. Model Development for the Prediction of Claw Horn Disruption Lesions," *PLOS ONE*, vol. 11, p. e0146718, 1 2016.

[27] N. Chapinal, A. de Passillé, J. Rushen, and S. Wagner, "Automated methods for detecting lameness and measuring analgesia in dairy cattle," *Journal of Dairy Science*, vol. 93, no. 5, pp. 2007–2013, 2010.

[28] S. R. van Amstel and J. Shearer, eds., *Manual for Treatment and Control of Lameness in Cattle*. Ames, Iowa, USA: Blackwell Publishing Professional, 11 2006.

[29] Nisha Angeline C. V. and R. Lavanya, "Fog Computing and Its Role in the Internet of Things," 2018.

[30] M. Pastell, H. Takko, H. Gröhn, M. Hautala, V. Poikalainen, J. Praks, I. Veermäe, M. Kujala, and J. Ahokas, "Assessing cows' welfare: Weighing the cow in a milking robot," *Biosystems Engineering*, vol. 93, no. 1, pp. 81–87, 2006.

[31] N. Chapinal and C. Tucker, "Validation of an automated method to count steps while cows stand on a weighing platform and its application as a measure to detect lameness," *Journal of Dairy Science*, vol. 95, no. 11, pp. 6523–6528, 2012.

[32] M. Pastell, L. Hänninen, A. de Passillé, and J. Rushen, "Measures of weight distribution of dairy cows to detect lameness and the presence of hoof lesions," *Journal of Dairy Science*, vol. 93, no. 3, pp. 954–960, 2010.

[33] W. Maertens, J. Vangeyte, J. Baert, A. Jantuan, K. C. Mertens, S. De Campeneere, A. Pluk, G. Opsomer, S. Van Weyenberg, and A. Van Nuffel, "Development of a real time cow gait tracking and analysing tool to assess lameness using a pressure sensitive walkway: The GAITWISE system," *Biosystems Engineering*, vol. 110, no. 1, pp. 29–39, 2011.

[34] A. Van Nuffel, M. Sprenger, F. A. Tuyttens, and W. Maertens, "Cow gait scores and kinematic gait data: Can people see gait irregularities?," in *Animal Welfare*, vol. 18, pp. 433–439, 2009.

[35] K. Zhao, J. Bewley, D. He, and X. Jin, "Automatic lameness detection in dairy cattle based on leg swing analysis with an image processing technique," *Computers and Electronics in Agriculture*, vol. 148, pp. 226–236, 2018.

[36] A. Poursaberi, C. Bahr, A. Pluk, A. Van Nuffel, and 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*, vol. 74, no. 1, pp. 110–119, 2010.

[37] Y. Guo, D. He, and L. Chai, "A machine vision-based method for monitoring scene-interactive behaviors of dairy calf," *Animals*, 2020.

[38] Y. Zhang, J. Cai, D. Xiao, Z. Li, and B. Xiong, "Real-time sow behavior detection based on deep learning," *Computers and Electronics in Agriculture*, 2019.

[39] A. Yang, H. Huang, B. Zheng, S. Li, H. Gan, C. Chen, X. Yang, and Y. Xue, "An automatic recognition framework for sow daily behaviours based on motion and image analyses," *Biosystems Engineering*, 2020.

[40] T. Van Hertem, E. Maltz, A. Antler, C. Romanini, S. Viazzi, C. Bahr, A. Schlageter-Tello, C. Lokhorst, D. Berckmans, and I. Halachmi, "Lameness detection based on multivariate continuous sensing of milk yield, rumination, and neck activity," *Journal of Dairy Science*, vol. 96, no. 7, pp. 4286–4298, 2013.

[41] L. Munksgaard, C. G. van Reenen, and R. Boyce, "Automatic monitoring of lying, standing and walking behavior in dairy cattle," *Animal Science*, vol. 84 (suppl., p. 304, 2006.

[42] V. M. Thorup, L. Munksgaard, P. E. Robert, H. W. Erhard, P. T. Thomsen, and N. C. Friggens, "Lameness detection via leg-mounted accelerometers on dairy cows on four commercial farms," *Animal*, vol. 9, no. 10, pp. 1704–1712, 2015.

[43] M. Alsaaod, C. Römer, J. Kleinmanns, K. Hendriksen, S. Rose-Meierhöfer, L. Plümer, and W. Büscher, "Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior," *Applied Animal Behaviour Science*, vol. 142, no. 3-4, pp. 134–141, 2012.

[44] C. Yunta, I. Guasch, and A. Bach, "Short communication: Lying behavior of lactating dairy cows is influenced by lameness especially around feeding time," *Journal of Dairy Science*, vol. 95, no. 11, pp. 6546–6549, 2012.

[45] K. O'Driscoll, L. Boyle, and A. Hanlon, "A brief note on the validation of a system for recording lying behaviour in dairy cows," *Applied Animal Behaviour Science*, vol. 111, no. 1-2, pp. 195–200, 2008.

[46] N. Blackie, E. Bleach, J. Amory, and J. Scaife, "Impact of lameness on gait characteristics and lying behaviour of zero grazed dairy cattle in early lactation," *Applied Animal Behaviour Science*, vol. 129, no. 2-4, pp. 67–73, 2011.

[47] "Dairy Masters (https://www.dairymaster.com/)."

[48] "ENGS-DAIRY (http://www.engs-dairy.com/)."

[49] S. V. W. S. D. C. A. V. N. Koen MertensKoen Mertens, Jürgen Vangeyte, "Automatic Lameness Detection in Dairy Cows using Artificial Neural Networks," in *Proceedings of the international conference of agricultural engineering CIGR-Ageng 2012*, 2012.

[50] D. Wu, Q. Wu, X. Yin, B. Jiang, H. Wang, D. He, and H. Song, "Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector," *Biosystems Engineering*, vol. 189, pp. 150–163, 1 2020.

[51] K. Bhargava, S. Ivanov, D. McSweeney, and W. Donnelly, "Leveraging fog analytics for context-aware sensing in cooperative wireless sensor networks," *ACM Transactions on Sensor Networks*, 2019.

[52] T. C. Hsu, H. Yang, Y. C. Chung, and C. H. Hsu, "A Creative IoT agriculture platform for cloud fog computing," *Sustainable Computing: Informatics and Systems*, 2018.

[53] D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors driven ai-based agriculture recommendation model for assessing land suitability," *Sensors (Switzerland)*, 2019.

[54] A. Tsipis, A. Papamichail, G. Koufoudakis, G. Tsoumanis, S. E. Polykalas, and K. Oikonomou, "Latency-Adjustable Cloud/Fog Computing Architecture for Time-Sensitive Environmental Monitoring in Olive Groves," *AgriEngineering*, vol. 2, pp. 175–205, 3 2020.

[55] M. Mičiaková, P. Strapák, I. Szencziová, E. Strapáková, and O. Hanušovský, "Several Methods of Estrus Detection in Cattle Dams: A Review," *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 66, no. 2, pp. 619–625, 2018.

[56] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, and L. Safina, "Microservices: Yesterday, Today, and Tomorrow," in *Present and Ulterior Software Engineering*, pp. 195–216, Cham: Springer International Publishing, 2017.

[57] K. El Emam, S. Benlarbi, N. Goel, W. Melo, H. Lounis, and S. N. Rai, "The optimal class size for object-oriented software," *IEEE Transactions on Software Engineering*, 2002.

[58] D. De Carvalho, M. Mazzara, B. Mingela, L. Safina, A. Tchitchigin, and N. Troshkov, "Jolie Static Type Checker: a Prototype," *Modeling and Analysis of Information Systems*, 2017.

[59] M. Gabbrielli, S. Giallorenzo, C. Guidi, J. Mauro, and F. Montesi, "Self-Reconfiguring Microservices," pp. 194–210, 2016.

[60] M. S.-S. Dmitry Namiot, "On Micro-services Architecture," *International Journal of Open Information Technologies*, 2014.

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

[62] "IBM Cloudant, https://www.ibm.com/cloud/cloudant (Accessed on 03/02/2018)."

[63] "IBM IoT Platform,https://console.bluemix.net/docs/services/IoT/iotplatform_overview.html#about_iotplatform, (Accessed on 03/02/2018."

[64] M. Stephenson and D. Bailey, "Do movement patterns of GPS-tracked cattle on extensive rangelands suggest independence among individuals?," *Agriculture (Switzerland)*, vol. 7, no. 7, 2017.

[65] G. R. da Silva, *Weather and climate and animal production*, vol. WMO Nr 134. 2007.

[66] L. A. Williams, G. J. Rowlands, and A. M. Russell, "Effect of wet weather on lameness in dairy cattle," *Veterinary Record*, vol. 118, no. 10, pp. 259–261, 1986.

[67] L. Solano, H. Barkema, E. Pajor, S. Mason, S. LeBlanc, J. Zaffino Heyerhoff, C. Nash, D. Haley, E. Vasseur, D. Pellerin, J. Rushen, A. de Passillé, and K. Orsel, "Prevalence of lameness and associated risk factors in Canadian Holstein-Friesian cows housed in freestall barns," *Journal of Dairy Science*, vol. 98, no. 10, pp. 6978–6991, 2015.

[68] M. Taneja, N. Jalodia, J. Byabazaire, A. Davy, and C. Olariu, "SmartHerd management: A microservices-based fog computing–assisted IoT platform towards data-driven smart dairy farming," *Software: Practice and Experience*, p. spe.2704, 5 2019.

[69] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "{API} design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122, 2013.

[70] "Cattle|Animal behaviour,http://www.animalbehaviour.net/cattle/ (Accessed on 21/03/2018)."

APPENDIX

Locomotion Scoring Sheet

Table 1: Sample locomotion scoring sheet used in data collection

Cow ID	Locomotion Score	Date Observed	Comments
1988	3	Nov 9, 2017	
2371	4	Nov 9, 2017	Severely Lame

Ethical Approval

Institiúid Teicneolaíochta Phort Láirge Waterford Institute *of* Technology

Port Láirge, Éire.
T: +353-51-302000
info@wit.ie

Waterford, Ireland.
T: +353-51-302000
www.wit.ie

Foirfe chun fónaimh

REF: 17/CM/TSSG/01

14th September, 2017.

Mr. John Byabazaire,
Apartment 403, Block C,
Railway Square,
Manor Street,
Waterford.

Dear John,

Thank you for submitting your amended documentation in relation to your project *'Smart Herd: Early Lameness Detection in Dairy Cattle'* to the WIT Research Ethics Committee.

Based on the revised WIT ethical approval application form and supporting documentation, I am pleased to inform you that we now fully approve the conduct of this project.

We will convey this decision to Academic Council.

We wish you well in the work ahead.

Yours sincerely,

Prof. John Wells,
Chairperson,
WIT Research Ethics Committee

cc: Dr. Alan Davy
Dr. Brendan Jennings

Presented Paper

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.

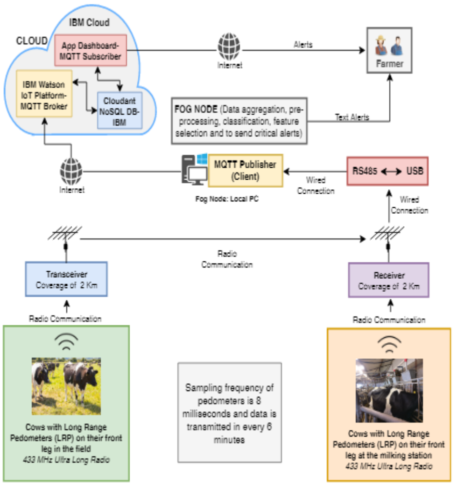


Fig. 1. Overall architecture of the test-bed

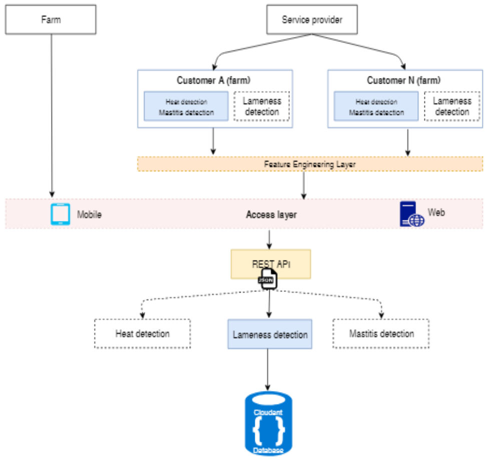


Fig. 2. Microservices based flow of the LDA

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

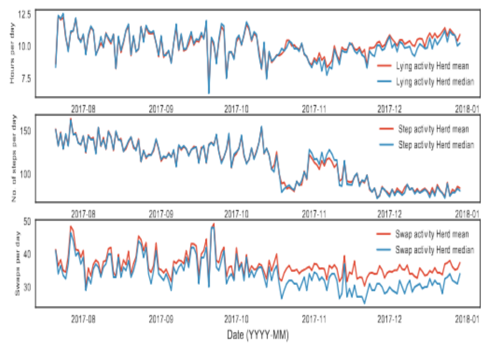


Fig. 4. Comparing the Mean and Median of the various Animal Activities

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.

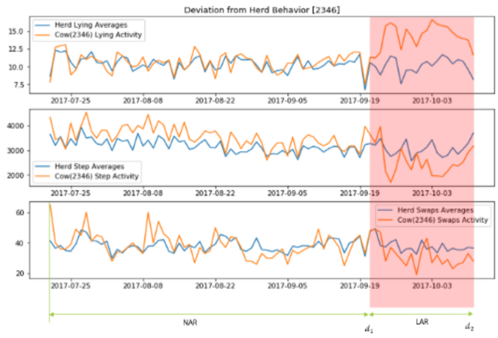


Fig. 5. Relationship between herd mean and cow activity for 2346

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

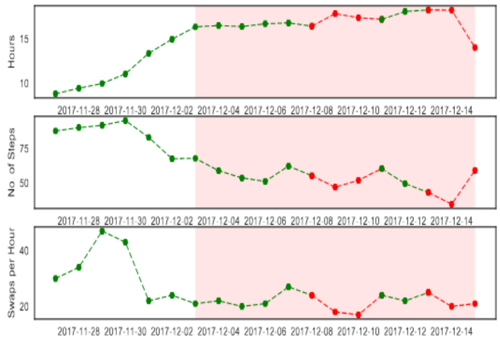


Fig. 6. Animal confirmed as lame between 03/12/2017 to 15/12/2017 but could not be correctly identified by a one-size-fits-all model

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.

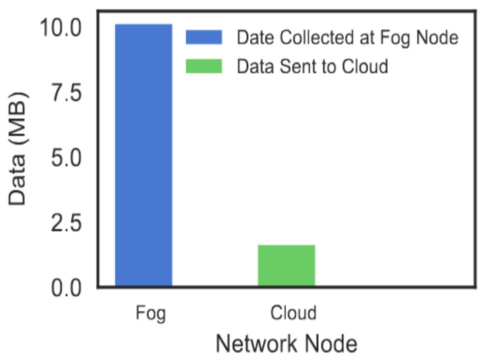


Fig. 7. Daily data reduction between the fog node and the cloud node

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.

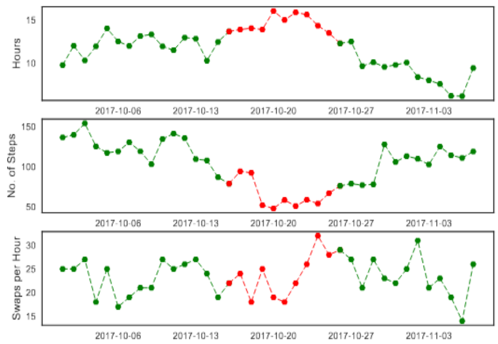


Fig. 8. Red points indicating lameness anomalies identified by the normal cluster model for cow ID 1988

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)