A meta-learning approach in a cattle weight identification system for anomaly detection

Rodrigo García[1,2,3], Jose Aguilar[1,4,5]

[1]GIDITIC, Universidad EAFIT, Medellín, Colombia.

[2]Faculty of Engineering, Universidad de Córdoba, Montería, Colombia.

[3]Faculty of Engineering, Universidad del Sinú, Montería, Colombia.

[4]CEMISID, Universidad de Los Andes, Merida, Venezuela.

[5]IMDEA Networks Institute, Leganés, Madrid, Spain.

rjgarciah@eafit.edu.co, aguilar@ula.ve

Abstract

Weighing management in cattle farming is important for farmers, as it allows them to accurately monitor the growth and development of their animals. It is also a valuable tool that allows farmers to maximize the production and welfare of their animals. However, it is difficult for the farmer to detect if the herd of animals being weighed is gaining the ideal weight for a given breed and age. In addition, normally, when a new breed of cattle is introduced to a farm, there is very little data. This article proposes a meta-learning framework (MTL) for identification models used in the fattening process of animals to detect anomalies in cattle weight. The proposed MTL framework has a knowledge base of Meta-Models on Identification models based on machine learning techniques, which is used to select the identification model to use when a new breed of cattle arrives on the farm. This knowledge base is updated, either because a previous identification model has been successfully adapted to the new breed, or a new identification model has had to be generated, allowing the framework to continuously improve its performance over time. Particularly, this article presents in detail the process of adaptation of the previous identification models to new breeds carried out by our MTL framework. Besides, to test our approach, a case study is presented, using records of animals raised and fattened at the "El Rosario" farm, located in the municipality of Monteria (Córdoba-Colombia). The results are very encouraging in terms of the ability of our framework to adapt the identification models to different possible scenarios in the process of detecting anomalous weights. In general, the identification models generated with our proposal had an R^2 of 90.8%, which suggests that the models can explain the variability observed in the data.

keywords: Meta-Learning, Identification Models, Artificial Intelligence, Precision Livestock Farming, Rotational Grazing, Beef Production, Anomaly detection

1 Introduction

A bovine's weight can be an indicator of its health and well-being. Cattle that are abnormally underweight may have health problems, such as disease or nutritional deficiencies [1].

On the other hand, cattle that are abnormally overweight may have health problems related to obesity, which can cause problems in the joints or the cardiovascular system. Thus, detecting abnormalities in cattle weight is a major issue in the cattle industry [2]. Besides, good cattle weight management can improve efficiency in Beef Production (BP), which can translate into increased profitability for cattle producers.

Precision livestock farming (PLF) offers farmers a real-time monitoring and management system. PLF can provide a real-time warning when something goes wrong so that the farmer can take immediate action to solve the problem [3]. For this, it requires systems that allow it to carry out the identification process of abnormal situations. Particularly, identification techniques propose approximate models of a real system, based on linguistic or mathematical expressions or an algorithm [4]. System identification has had an important development [4], but many problems remain. One of these problems is the definition of models for control and yield adjustment in real-time [4].

On the other hand, Machine learning (ML) is a very useful tool for cattle farming [5]. It can be used to improve herd selection, and herd rotation management, among other things. Also, ML models can provide information about the health and performance of cattle, as well as the quality and quantity of food and water that must be supplied. ML can too help identify diseases early, allowing for faster treatment and better monitoring of a BP [6]. All these possible applications help livestock producers to make better decisions in the management of their BP [7]. Finally, it is important to emphasize that within the identification techniques there are those based on ML, which have produced very interesting results in different contexts [2].

Although ML can be a valuable tool in cattle ranching, it also presents some challenges [5]. One of the main challenges is the availability of high-quality data, as ML requires a large amount of data to operate efficiently. In this regard, MTL, also known as "learning about learning", is an ML paradigm that is used to improve the ability of an ML-based model to adapt to the context (new datasets, etc.) For example, MTL allows an ML model to "learn how to learn" from new datasets, which may even be very small [8]. In general, MTL allows an Automated-Machine-Learning (AutoML) process for the automatic selection, composition, and parameterization of ML models, to achieve optimal performance on a given task [9, 10].

Specifically, in the context of cattle weight anomaly detection, MTL can be an important tool for the definition of detecting anomalies models with a limited amount of data. In this paper, we present a novel MTL framework, with the aim of automating the selection and/or parameterization of ML-based models of cattle weight identification. Our approach proposed in this work is a novel approach to meta-learning in the context of livestock weight anomaly detection. Our proposal is innovative in that it is based on the construction and continuous adaptation of metamodels based on the results and prediction quality of individual models (in this case, detection of anomalies). These metamodels are not predefined, but are generated in real-time based on the ML models to be adapted and their quality metrics. Thus, our metamodels are built according to the context in where will be used the MTL framework. The underlying premise in the choice of MTL lies in the assumption that the ideal weight growth curve for a specific breed of cattle is largely similar across breeds and contexts. This assumption is based on the idea that, although variations may exist, certain fundamental patterns in cattle weight evolution are shared among populations of the same or different breeds.

In particular, the proposed solution seeks to improve anomaly detection when little data exists for a specific cattle breed. The main contributions of this work are:

- The proposition of the first MTL architecture for livestock weight anomaly detection in the livestock industry, which can be successfully adapted to different scenarios in livestock production.

- The definition of an MTL framework that uses a meta-model knowledge base of identification models based on machine learning techniques, which allows the selection of the appropriate identification model for each breed of cattle on the farm.
- The characterization of an MTL approach for precision livestock farming, which can be adapted to different contexts of production for anomaly detection, among other applications.

The remainder of this paper is structured as follows: Section 2 introduces the related work to this work. Section 3 describes our MTL framework and Section 4 its instantiation in a case study. Furthermore, Section 5 shows an analysis and discussion of the results, and Section 6, a comparison with other works. Finally, Section 7 presents the conclusions and future works.

2 Related Work

Regarding ML for anomaly detection, considerable progress has been made in the use of tools for routine monitoring and collection of animal information in BP [11]. However, although there are many applications of ML in livestock, the use of meta-learning in this industry is new, and there are no works about the utilization of MTL specifically to improve detection of cattle weight anomalies during rotational grazing. For this reason, in this section, we present some works on anomaly detection using MTL in other contexts, and on anomaly detection in PLF.

2.1 Anomaly detection using MTL

In this section, we will show studies on the use of MTL approaches in the field of anomaly detection, with the aim of improving the supervised processes.

Moon *et al.* [12] proposed a method for unsupervised anomaly detection in time-series sensor data of smart buildings. They used a model-agnostic MTL and the variational autoencoder technique to adapt the model to a new target task with few unlabeled anomaly data. Entezami *et al.* [13] defined an unsupervised MTL method for health monitoring of civil structures for challenges such as large data with missing values and severe environmental changes. The MTL method is based on locally robust Mahalanobis-squared distances for online anomaly detection

Peng *et al.* [14] used an approach to spectrum anomaly detection in cognitive radio using MTL. The proposed method addresses the problem of the inability of existing deep learning-based spectral anomaly detection algorithms when are used directly across different frequency bands. The method involves using pre-training to analyze the differences between frequency bands, constructing an MTL dataset to find optimal model parameters, and fine-tuning the model using a small amount of target band data to detect anomalies. Tan *et al.* [15] proposed a self-supervised anomaly detection method that uses MTL to increase adaptability. The proposed method aims to improve sensitivity to subtle irregularities while maintaining robustness. The method is relevant for screening applications, and its effectiveness is demonstrated through experimental results.

Tavares *et al.* [16] proposed a method to extract meta-features from the event log and to recommend the most suitable encoding technique to improve anomaly detection performance. Their results showed that event log characteristics have different impacts on the representational capabilities. Dogo *et al.* [17] proposed a method for detecting anomalies in water quality, which is formulated as a classification problem in the presence of class imbalance. Sixteen single and static ensemble classification methods embedded with resampling strategies are optimized and compared, and six dynamic selection techniques are proposed and evaluated using an MTL approach.

With respect to the previous works, the objective of our work is to detect anomalies in the weight gained during cattle fattening, using an identification model based on an MTL architecture. MTL techniques were used to evaluate and select different ML algorithms to find the most suitable model.

2.2 Detection of anomalies in PLF

This section analyzes relevant research on the detection of abnormalities in the BP process, which are presented according to the object of study. The studies include the use of environmental and body temperature data, motion sensors, sound data, and activity monitoring, to detect anomalies in the health of cattle.

Animal behavior: Cai *et al.* [18] proposed a monitoring system for analyzing daily hog activity and abnormal behaviors in hog farms. The system uses a passive infrared detector and a high-accuracy acquisition system to collect data on daily hog activity, and uses an improved K-means clustering method to detect abnormal behavior during the night. The developed system provides data for the analysis and evaluation of the health, diseases, and environmental conditions of hog farms, which can affect fertility and productivity rates.

Animal welfare: Perrin *et al.* [19] evaluated an anomaly detection algorithm used in an automated surveillance system of cattle mortality. The method combined temporal regression and spatial cluster detection to identify clusters of spatial units showing an excess of deaths compared to their own historical fluctuations. The study simulated 1,000 outbreaks of a disease causing extra deaths in the French cattle population and applied the algorithm to identify clusters of spatial units showing an excess of deaths. The results indicated that the algorithm was able to identify unusual mortality clusters caused by an outbreak in certain conditions.

Kramer *et al.* [20] developed a fuzzy-logic model for the classification and control of lameness and mastitis, in cows, using data from the Futterkamp dairy-research farm of the Schleswig-Holstein Chamber of Agriculture. The fuzzy-logic model generated disease alerts using milk yield as the output variable, and as input data, dry-matter intake, dry-matter intake behavior (number of visits at the feeding trough, time spent at the feeding troughs), water intake, activity, and information about preliminary diseases. Sai *et al.* [21] developed an artificial intelligence module to estimate in a non-contact manner the body temperature of cattle, allowing for efficient individual monitoring of the health status of cattle. The module collected data on environment temperature, humidity, illuminance, and infrared images of cattle in a real-life environment.

Haladjian *et al.* [22] presented an approach to automatically detect cow lameness by monitoring changes in their gait, using a wearable motion sensor attached to their hind left leg. For that, the approach builds a model of a cow's usual walking pattern and detects deviations from this model. Results from a controlled experiment show that the approach can detect deviations in cows' gait with an accuracy of 91.1%. Wagner *et al.* [23] conducted a study on dairy cows, in which their activities were captured as time series by an indoor tracking system. The state of cows (diseases, stress, no problem) was manually labeled by animal caretakers, or by a sensor for ruminal pH (acidosis). Then their approach used a Fourier-based method to detect anomalies in time series.

Chung *et al.* [24] proposed a data-mining solution for the detection of oestrus, using sound data from Korean native cows (Bos taurus coreanea). They extracted the mel frequency cepstrum coefficients from sound data, with a feature-dimension reduction technique, and used a support vector machine for anomaly detection. The results indicated that this method can be used to detect estrus both economically (even with a cheap microphone) with an accuracy greater than 94%. Finally, García *et al.* [2] proposed an approach to detect anomalies in the cattle fattening processes. This approach used the historical record of animal weight to identify whether animals have gained the appropriate weight over time. They compared several ML techniques (Decision Tree, Gradient Boosting, K-Nearest Neighbors-based regression and Random Forest) in the task of anomalous weight detection, using Mean Absolute Error as quality metrics.

Pastures: Calera *et al.* [25] presented two approaches to monitor pasture quality based on multi-source info, being DEIMOS-1 the major satellite contributor. The first approach is based on in-depth monitoring of the crop phenology, characterized by means of the weekly Normalized Difference Vegetation Index. The second approach is focused on drought and other anomaly detection in crops and pastures. For pastures, a specific module has been designed to detect drought occurrence, by comparing current Normalized Difference Vegetation Index values with historical ones. Besides, they detected maximum livestock stocking rate, need of supplementary feeding, and overstocking risk.

Farm equipment: Park *et al.* [26] defined a mechanism to detect anomalies in pig house equipment using a recurrent neural network learning model, with data from sensors and environmental controllers. They predict malfunctions of each equipment, and when something goes wrong with the sensor, they use the difference between the predicted value and the measured value.

Early detection of abnormalities in animal husbandry is important for several reasons. First, it can improve production efficiency by allowing farmers to intervene quickly in situations that affect animal growth and health. Second, it can improve animal welfare by detecting and treating diseases and injuries before they become serious problems. In addition, early detection of abnormalities can help prevent the spread of disease to other animals on the farm, and in the wider community. Therefore, early detection of abnormalities is significant both for the health and welfare of the animals and for the economic sustainability of the livestock industry.

3 MTL Framework

MTL paradigm can be used to predict the performance of an ML algorithm on a specific task [8]. Thus, one goal of the MTL paradigm is to find the correlation between the characteristics of a dataset and the performance of different learning algorithms. With this information, a predictive meta-model can be defined to estimate the performance of different ML approaches on a dataset.

3.1 Conceptual architecture

The general architecture of the proposed framework is depicted in Figure 1. The association engine selects the best algorithm and parameterization by using a knowledge base. The knowledge base is a collection of meta-features describing datasets, models, and hyperparameters. Each time a new dataset is presented to the system, the association engine provides a suggestion of the most appropriate model (predictors, with their hyperparameters configuration). Immediately, the quality of the prediction is evaluated, and in case it is acceptable, then it is added to the knowledge base which is located in the Meta-Model.

In this way, the identification system for anomaly detection is able to adapt to new data (few or many), using the MTL paradigm together with the metadata that is stored in the form of a meta-model base. In general, this identification system has been developed on the concept of MTL, so it consists of three main modules: learning, association, and adaptation.

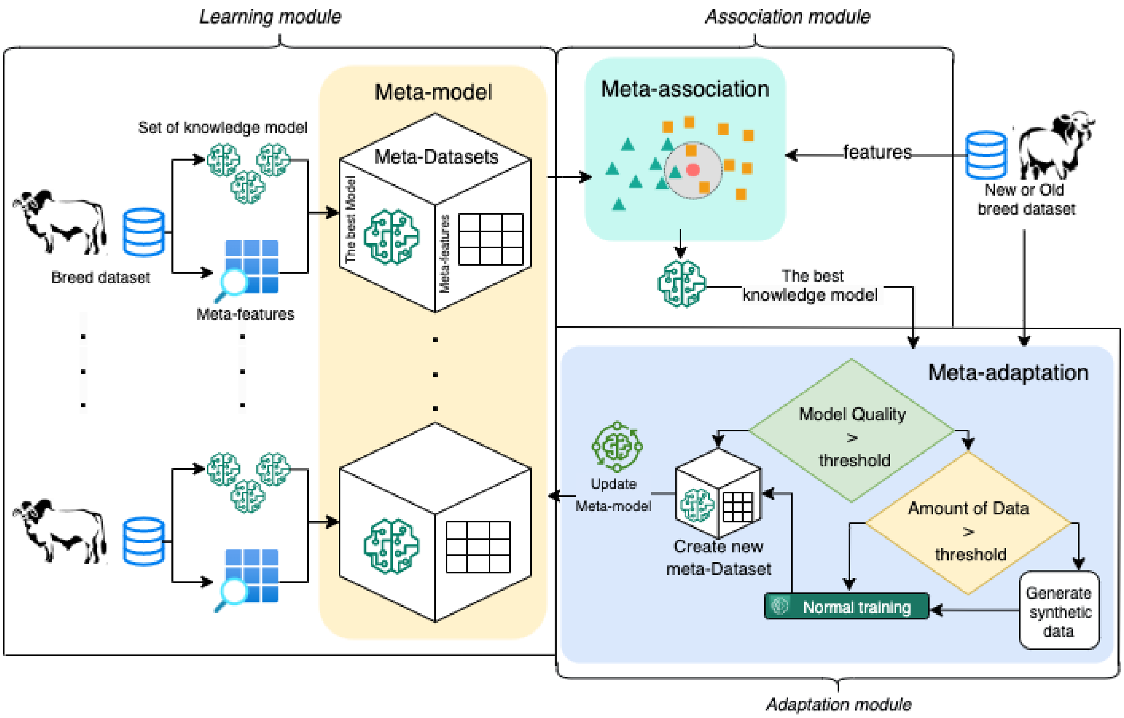

Figure 1: Meta Architecture for PLF

3.2 Learning module

It defines a knowledge base (metamodel) using metadata about previous learning tasks and learned models. It contains the characteristics of the datasets (e.g., mean, median, variance, and standard deviation of the variables), performance measures of each machine learning algorithm on those particular datasets, among other things. The metamodel represents the correspondence between the meta-characteristics describing the datasets, and the performance metrics obtained by the group of learning algorithms when applied to these datasets. Thus, this module keeps mainly the results of different learning algorithms on datasets, and meta-features of these datasets. In this context, the appropriate ML model for a new dataset can be selected using the metamodel.

Particularly, the metamodel is defined at the beginning based on the ML models to be adapted (predictive, diagnostic, or prescriptive models) in the context where will be used the MTL framework. For the definition are considered the parameters of the techniques used, and the metrics to assess the quality of ML models. With this information, our MTL framework can evolve and improve its performance over time as more information is gathered. Specifically, in the PLF context, the information in the metamodel is:

- Livestock Datasets: The primary inputs are different historical livestock datasets, which have been collected from various breeds and contexts.
- ML Models by type (predictive, diagnostic, or prescriptive models): For each type of

model, different ML algorithms (e.g. regression, random forest, support vector machine, etc.), and their parameters and relevant performance values by each dataset.

This information is used by the association module each time it receives a request (input) to find the most appropriate model for a type of learning (for example, to predict) for a new dataset. In turn, the metamodel is updated by the adaptive module each time it completes its task. This update can be including the new dataset and what techniques it was trained with (with its parameters and quality metrics); or if an existing model was used, the quality metrics and that new dataset are updated in said model.

In summary, the *input* of the module will be the updates that the adaptation module makes each time it finishes its task, either to incorporate new datasets with the ML models generated and the techniques used for it; or to update existing ML models built now with the new dataset and the techniques used for it. Also, the *output* of the module is an updated metamodel that incorporates information about which models are most effective for each dataset that describes a specific breed or context, which is used by the association module for the future selection of models in new situations.

3.3 Association module

The association module starts when a new dataset arrives for analysis (with little or a lot of information). At that moment, a set of meta-features describing the new dataset are extracted. Then, this module compares the features of the new dataset with the previous ones, using K-Nearest Neighbors (KNN) to place the new dataset in the closest cluster (for this, it is assumed that each metamodel in the knowledge base is a cluster). Now, it suggests the best model with its respective hyperparameters.

3.4 Adaptation module

In the adaptation module, the selected model is tested using the new data and its performance is evaluated. If the performance is not satisfactory (it does not pass a quality threshold), then the system evaluates whether the new dataset has enough data for a normal training. If it has too little data, then it proceeds to generate synthetic data and train. For both cases (enough data or synthetic data), then it ends up in normal training. Next, this module updates the meta-model (if any were selected), or builds and adds a new meta-model to the knowledge base (if it creates a new ML model). In general, this module acts when there is a new dataset (in this particular case, a new breed or crossbreed) or an old dataset that has time without being used. Then, it updates the meta-features and performance measures of current metamodels, or creates a new metamodel. In conclusion, this module creates/selects the best model and the best hyperparameters for a given dataset. In addition, the meta-model´s knowledge base is updated each time, allowing the system to continuously improve its performance over time.

4 Case Study

This section presents the experimental context of the case study and the instantiation of the MTL architecture for PLF in it.

4.1 Context

In cattle ranching, a system widely used by beef producers in the tropics is rotational grazing. Rotational grazing consists of dividing the entire surface area of a farm into more than two paddocks, while some remain occupied, the others are at rest [27]. This reduces the total grazing area, and forces cattle to consume forage uniformly, assigning batches of cattle to different paddocks (see Figure 2).

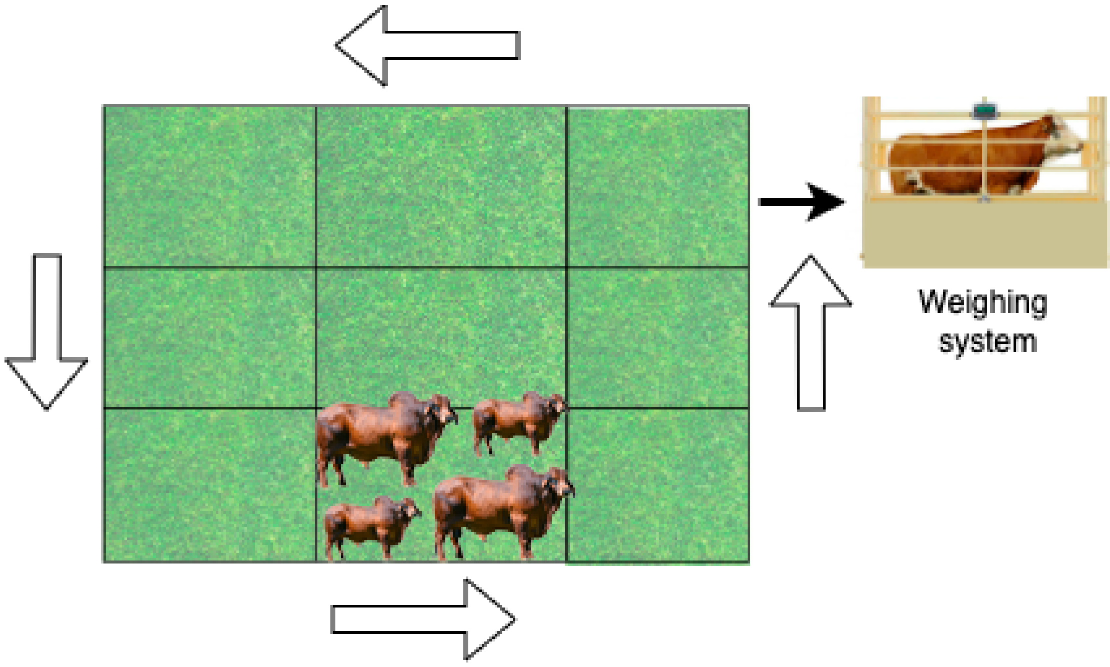

Figure 2: Rotational grazing system

Weighing cattle on rotational grazing is crucial for effective herd management because it is an important indicator of performance, and provides valuable information on grazing quality and other factors that can affect cattle growth and health. By monitoring cattle weight, producers can make informed decisions about grazing and feeding management, which can improve their farm's productivity and profitability.

In addition, cattle weights can also help detect health problems in cattle early, allowing producers to take preventative measures before they become serious problems. Regular monitoring of cattle weight can also help producers optimize grazing management, which can reduce the risk of overgrazing and improve long-term pasture and soil quality.

The age for slaughtering cattle will depend on the moment at which they reach the appropriate weight. The ideal is to be able to reach that appropriate weight between 450 kg to 500 kg at 24 months of age. Thus, regularly weighing livestock on rotational grazing is an important practice that improves livestock management and agricultural productivity. It is important to weigh cattle on rotational grazing for several reasons:

- Monitoring cattle performance: Cattle weight is an important indicator of the performance of the BP, which can help producers make decisions about herd management, including the timing of the sale.
- Evaluating grazing success: Cattle weight can also provide information on grazing quality, feed availability, and other factors that may be affecting cattle growth.
- Early detection of health problems: Fluctuations in weight can be an early sign of health problems in cattle, allowing producers to act before they become major problems.
- Improve grazing: Weight information can also help producers make decisions about grazing management, including pasture rotation and the number of cattle in each section, which can improve cattle growth and health, as well as farm productivity.

In particular, we use real datasets in our case study, which are the records of animals raised and fattened at the "El Rosario" ranch, located in the municipality of Monteria (Córdoba-Colombia).

4.2 Instantiation of the MTL architecture for PLF

This section describes the use of our MTL architecture in a monitoring system of the animal fattening process at the Rosario ranch for the detection of anomalies. In general, the system starts with the learning module where we have different animal breed datasets containing the variables gender, age, and weight. These breeds are crosses of the genetic groups Angus x Cebu (AC); Bon x Cebu (BC); Cebu x Angus x Cebu (CAC); Cebu x Cebu (CC); Holstein x Cebu (HC); Bon x Angus x Cebu (BAC), typical in Colombian farms. The data information for different animal breeds is described in Table 1, in which we see that their statistical metrics are quite close.

Table 1: Breed information

		Weight		
Breed	Count	**Mean**	**Min**	**Max**
AC	398	427.850369	165.219753	574.046996
BAC	398	410.001041	173.788039	541.459499
CAC	398	442.683480	181.033801	581.983351
BC	398	459.947979	156.834548	619.393828

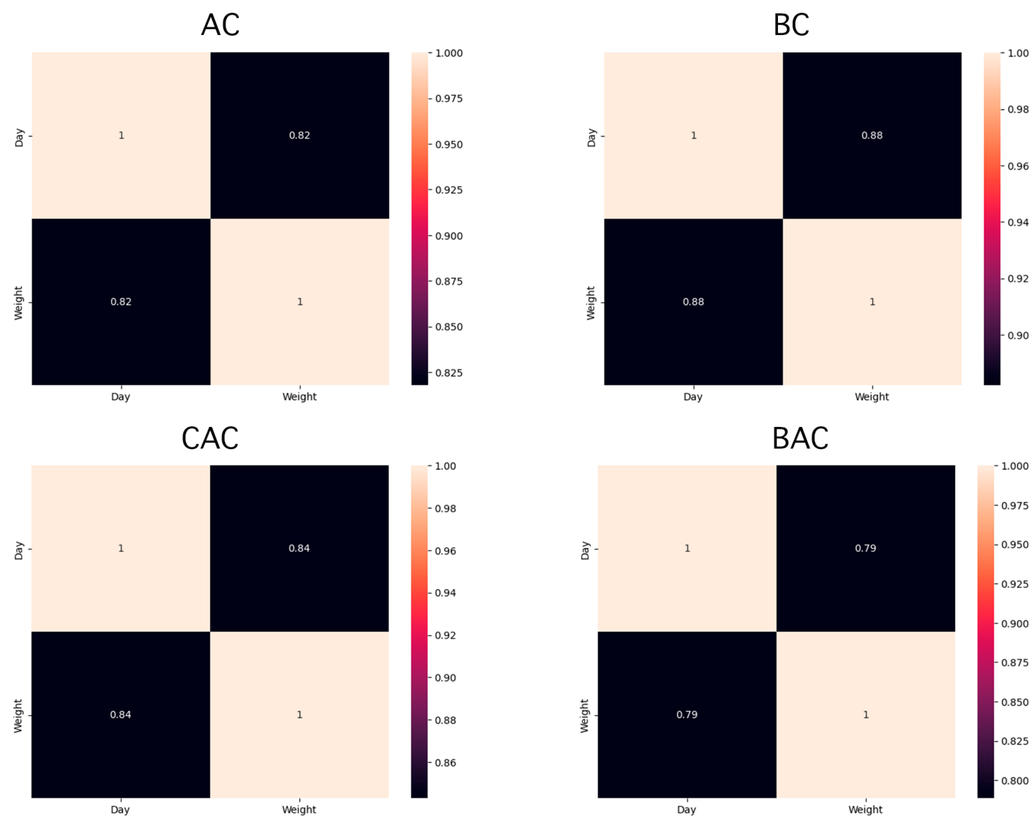

Figure 3: Correlation between age and weight

The relationship between cattle weight and age data is crucial in detecting abnormalities and building effective models for livestock management, as can be seen in Figure 3. These data are highly correlated, and analyzing them together provides valuable information on the health status and performance of cattle. Therefore, cattle age is a key factor influencing their weight because as animals age, it is natural to expect their weight to increase. Thus, understanding how this relationship evolves at different stages of growth is essential.

The learning module refers to the process of acquiring knowledge and skills that make it possible to learn to learn. The goal of this module is to understand the patterns and strategies used to learn, and apply that knowledge to future learning situations. That involves selecting and preparing the data to learn, which includes, among other things, choosing and tuning the appropriate ML model for the specific problem being addressed.

For our particular case, for each breed, a set of knowledge models is trained using different ML techniques with the dataset features. From the set of models, the one with the best quality metrics is selected. With this model, the metamodel of this dataset is created, which will be stored in the knowledge base (see Table 2).

Three ML techniques were used in this first module: K-Neighbors Regressor (KNN), Gradient Boosting Regressor (GB), and Random Forest Regressor (RF). GridSearchCV is used as a hyperparameter optimization tool. This tool exhaustively searches for the optimal combination of hyperparameters of a model, to improve its performance and accuracy. Moreover, the features of the variables of the dataset are calculated. Finally, a meta-model is assembled for each dataset, which is composed of the best ML model and their meta-features (see Table 2).

Table 2: Meta-Model

knowledge		*Meta-features*									
Breed	Model	Best params	R^2	Median	Mean	Std	Var	Kurtosis	Entropy	Variation
AC	GB	learningrate:0.1 maxdepth:5 minsamplessp:5 nestimators:40	92	451.6	427.8	64.6	4175.25	1.20	10.62	0.15
BAC	RF	maxdepth:20 maxfeatures:3 maxleafnodes:25 nestimators:40	89	430.24	410.0	57.1	3262.80	1.49	10.62	0.13
CAC	GB	learningrate:0.1 maxdepth:5 minsamplessp:5 nestimators:40	93	470.04	442.6	72.5	5265.85	0.86	10.62	0.16
BC	GB	learningrate:0.1 maxdepth:5 minsamplessp:5 nestimators:40	95	492.2	459.9	87.5	7658.87	0.33	10.62	0.19

The following sections will explain the operation of the other modules of our architecture.

5 Experiments

Our architecture is divided into 3 modules, the first one was already instantiated in the previous section, and the next two are explained below. In this section, we will describe the behavior of the previously trained ML models in the MTL architecture, using data from the farm “Ganadería El Rosario”. The duration of the experiments is a complete cycle of animal fattening on the farm before their sales.

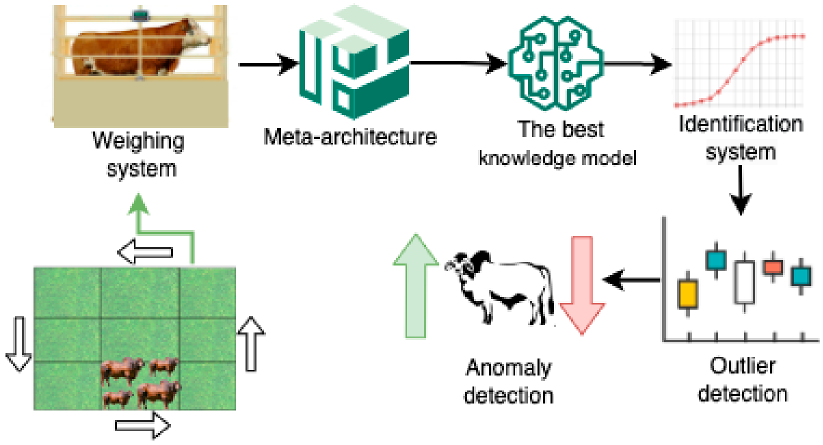

Figure 4: Schematic diagram of system behavior

The weighing process in rotational grazing involves the use of an animal weighing system (weigh scale) that is installed in a walkway or access corridor between the different grazing sections. The animals are guided to this weighing scale. During the process, the animals are weighed, and the data is recorded in a livestock management software. This data is input to our architecture that will select the best model to build an ideal weight identification system. Subsequently, it will detect anomalies using outlier detection techniques (see Fig 4).

MTL tells ML how to learn, in order to learn from previous learning experiences rather than starting from scratch for each task. To show the versatility of our MTL architecture, we pose different scenarios that show how our proposal adapts to new and old datasets.

Thus, there are two main scenarios, the first is when the breed is known (old dataset) in the knowledge base and the second scenario is when it is an unknown breed (new dataset). For both scenarios, 2 sub-scenarios can be presented, the first would be when the model selected in the knowledge base has a good quality metric when using that dataset (it is not necessary to retrain), and the second when the metric quality is not good.

5.1 Scenario 1: Breed known by the knowledge base

5.1.1 Good quality of the model

If the breed is known (e.g., AC breed), then our MTL architecture searches in its knowledge base (Meta-model) the corresponding information. Subsequently, it selects the best model for that breed and makes the prediction. For this case, the best model is GB (see Table 3). Finally, it compares the ideal weight growth curve described by the AC breed cattle identification model with the current weights to detect anomalies (point cloud in Fig 5). In this scenario, the decision to use the GB model was based on its performance in quality metrics (see Table 4).

Table 3: Meta-model in the case of a breed known by the knowledge model

Breed	Model	Best params	R^2	Median	Mean	Std	Var	Kurtosis	Entropy	Variation
AC	GB	learningrate:0.1 maxdepth:5 minsamplessp:5 nestimators:40	92	451.6	427.8	64.6	4175.25	1.20	10.62	0.15

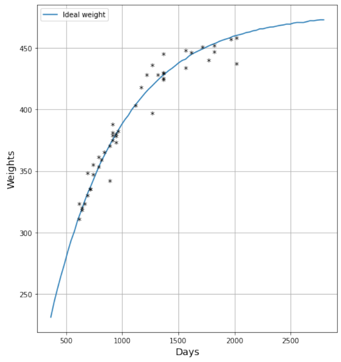

Figure 5: Ideal weight-curve versus current weight in AC breed

Since the curve represents the ideal temporal behavior of weight, points that deviate from it are potentially anomalous, but it is necessary to define the magnitude of the distance that separates an anomalous weight from one considered normal. To define this, multiple predictions of ideal weight are obtained for each combination of day, breed, and sex. Ideal weight intervals (not outliers) are then constructed to detect outliers using an isolation forest. This is an ML algorithm for anomaly detection that uses randomized decision trees to isolate anomalies in a dataset [28]. This allows for rapid detection of anomalous increases and decreases in animal weights at any desired time for decision-making.

Table 4: Decision metrics with good values

Breed	Model	R^2
AC	GB	92.5

5.1.2 Bad quality of the model

In the event that the quality of the metric is bad (see Table 5), a new training will be performed using new data (synthetic or this data) to adapt the selected model to these new situations described by the data. This means that the models can improve their predictive capability over time.

Table 5: Decision metrics with bad values

Breed	Model	R^2
AC	GB	57.2

Thus, the quality of the predictions is improved thanks to the meta-adaptation module. In this particular case, the new performance of the selected model after retraining is shown in Table 6.

Table 6: Improved decision metrics

Breed	Model	R^2
AC	GB	87.5

5.2 Scenario 2: Breed unknown by the knowledge base

In the case that a new breed arrives at the farm, or a new cross is created, the system will activate the second module (association) to select the best learned knowledge model for this specific dataset. This module extracts a set of meta-features from the dataset of the new breed to associate the model that best fits the specific features of the new data. Thus, this module automatically selects the best model based on the specific features of the dataset and its similarity with the features of the datasets with which previous models were trained.

In order to test the level of adaptation of our proposal, we removed the CAC, BAC and CC breeds from the knowledge base. Subsequently, the adaptation module is activated. Thus, first the quality of the model and the amount of data are evaluated. For this particular case, the model was not trained on the current BAC, CAC and CC breed, so the meta-model does not know these breeds. The association module is activated to search for the most similar group. Table 7 shows the input of the association module, and Table 8 shows to which group each breed is assigned.

Table 7: Feature inputs for the association module

Breed	Median	Mean	Std	Var	Kurtosis	Entropy	Variation
BAC	430.24	410.0	57.1	3262.80	1.49	10.62	0.13
CAC	470.04	442.6	72.5	5265.85	0.86	10.62	0.16
CC	471.631	455.190	45.666	2085.421	0.900	5.241	0.100

Table 8: Meta-Association

Breed	Median	Mean	Std	Var	Kurtosis	Entropy	Variation	Cluster
AC	451.6	427.8	64.6	4175.25	1.20	10.62	0.15	0
BAC	430.24	410.0	57.1	3262.80	1.49	10.62	0.13	0
CAC	470.04	442.6	72.5	5265.85	0.86	10.62	0.16	0
BC	492.2	459.9	87.5	7658.87	0.33	10.62	0.19	1
CC	471.6	455.1	45.6	2085.42	0.90	5.24	0.10	0

If the performance of the selected knowledge model is of bad quality for the new dataset, then this can have a negative impact on the decisions and results based on it. For this case, the MTL framework invoked by the fattening monitoring system will consider the following cases.

5.2.1 Good quality of the model

If the selected model has a quality higher than the quality threshold using the new dataset, then a new meta-model is created that contains this model as the best model and the meta-features of this new dataset. Subsequently, the knowledge base is updated to be aware of this new breed (see Table 9).

Table 9: New meta-model

Breed	Model	R^2	Median	Mean	Std	Var	Kurtosis	Entropy	Variation
AC	GB	92	451.6	427.8	64.6	4175.25	1.20	10.62	0.15
BC	GB	95	492.2	459.9	87.5	7658.87	0.33	10.62	0.19
BAC	GB	87	430.24	410.0	57.1	3262.80	1.49	10.62	0.13
CAC	GB	90	470.04	442.6	72.5	5265.85	0.86	10.62	0.16
CC	GB	90	471.6	455.1	45	2085.42	0.90	5.24	0.10

5.2.2 Bad quality of the model

For this case, if the performance of the selected model is of poor quality with the new data set, then two situations can occur:

- In the first situation, the amount of data is sufficient to train a model (see Table 10). Previously, we have removed the CC race from the knowledge dataset to test this situation. In this case, the system tests different ML techniques using a cross-validation process. Finally, a list of models with their quality metrics is obtained. The next step is to select the one with the highest performance (see Table 11) to update the meta-model with this new breed (see Table 12).

Table 10: Description of breed data CC

Breed	Median	Mean	Std	Var	Amount of data
CC	471.6	455.1	45.6	2085.42	976320

Table 11: List of models for the CC breed

Model	R^2	Select
RF	91.6	True
GB	89.2	False
DT	86.3	False
KNN	82.5	False

Table 12: New meta-model

Breed	Model	R^2	Median	Mean	Std	Var	Kurtosis	Entropy	Variation
AC	GB	92	451.6	427.8	64.6	4175.25	1.20	10.62	0.15
BC	GB	95	492.2	459.9	87.5	7658.87	0.33	10.62	0.19
BAC	GB	87	430.24	410.0	57.1	3262.80	1.49	10.62	0.13
CAC	GB	90	470.04	442.6	72.5	5265.85	0.86	10.62	0.16
CC	RF	91.6	471.6	455.1	45	2085.42	0.90	5.24	0.10

- In the second situation, the amount of data is not sufficient to train a model. Therefore, we proceed to generate synthetic data to be able to perform the training. Synthetic data has already been used in different areas. Once the amount of data necessary for the traditional training is completed, then the steps described for the first situation are executed.

5.3 General Analysis

Some features of the MTL framework are the following. It can start from a set of models built with available historical cattle weight data, allowing the system to learn patterns and relationships between the data. On the other hand, it has the ability to adapt to face unforeseen situations (for example, new breeds) or the emergence of new variables, which gives it great flexibility and the ability to respond to different scenarios. This is especially important in the cattle industry, where changes in the environment and conditions can affect the weight of cattle and therefore their health.

In general, our approach can consider different scenarios in the adaptation process of ML models. When models fit the new data well, simply update your metamodels to that context, and when it doesn't, retrain the model with that new data (even if it's not enough to retrain, eventually generate synthetic data). As can be seen, all possible adaptation scenarios of ML models are considered. It can also be seen that our approach does not depend on specific ML techniques (random forest, regression, etc.). New models based on other techniques can be added. The same is true for ML model types. In this study, identification models were considered, but future work could consider diagnostic and optimization models, among others.

Finally, our proposal had a general average of R^2 of 90.8% for the different scenarios proposed. MTL enhances learning capabilities according to new data presented to it. In other words, our MLT framework is capable of learning to learn. In this context, our proposal has the capacity to continuously improve its anomaly detection capacity, even if the characteristics of the cattle or the environment in which they are found change over time. Particularly, in the experiments with the data from the "Ganadería El Rosario" using our MLT-based framework, two types of anomalies were mainly detected: fattening problems of the animal because it was sick or because the quality of the pasture was not good.

6 Comparison with Other Works

In this section, we propose several criteria to compare our approach with similar works. Particularly, we define three criteria, which are:

- ***Criterion 1:*** The work includes adaptive modeling processes.
- ***Criterion 2:*** The work includes the transferability of the model to different domains or tasks.
- ***Criterion 3:*** The work includes incremental learning capabilities.

In Table 13, a qualitative comparison with related studies is made, based on previous criteria.

Table 13: Comparison with previous works.

	Criterion 1	***Criterion 2***	***Criterion 3***
Moon et al., 2023 [12]	✗	✓	✗
Entezami et al., 2023 [13]	✗	✗	✗
Peng et al., 2022 [14]	✗	✗	✗
Tan et al., 2022 [15]	✗	✗	✗
Tavares and Junior, 2021 [16]	✗	✓	✗
Dogo et al., 2021 [17]	✗	✗	✗

Table 13 continued from previous page

	Criterion 1	*Criterion 2*	*Criterion 3*
This work	✓	✓	✓

For criterion 1, our work is the only one that meets. This criterion is important since the inclusion of adaptive ML-based modeling processes is a valuable feature in any work that involves machine learning. Adaptive modeling processes allow the model to change and improve over time by taking into account new data or changes in the underlying system being modeled. Otherwise, it would be static, that is, it is designed to remain fixed and unchanged once it is trained.

For criterion 2, the work [12] can be used in real-time facility management systems in smart buildings for several applications, including early detection of equipment failure, automated monitoring, and energy conservation. In addition, [16] proposes detecting anomalous instances in business processes to avoid resource waste and mitigate security issues. Finally, in our work, the transferability of the model is observed in the model's ability to adapt and generalize to new tasks. In our approach, the models can be retrained in new datasets (can represent specific tasks), to include the knowledge and experience inside of them to the original model.

For criterion 3, our framework possesses incremental learning capability because our architecture can continuously learn when new datasets or information is received.

On the other hand, we have made a more specific comparison of works on the detection of anomalies in cattle weight. The paper of Segerkvist et al. [29] evaluated a method for monitoring the health of grazing cattle based on an unmanned, automatic precision weighing system that can be used on pasture. This system can generate alarms when animals show abnormal weight gain curves. The paper focused mainly on the detection of pasture-borne nematode parasite infections, which affect the weight gain of calves. Wagner et al. [30] used ML techniques to detect abnormal behaviour in cows with the Sub-Acute Ruminal Acidosis (SARA) disease, which is known to induce changes in behaviour. They used a positioning system to infer an animal's activity based on its position in relation to specific elements in the barn (alleys, feeder, and resting area), and defined ML models to predict activity on a given day. The work of Wagner et al. [23] detects the state of cows (diseases, estrus, no problem), which was trained using a dataset manually labeled by animal caretakers or by a sensor for ruminal pH (acidosis). They proposed a Fourier-based method to detect anomalies in this time series, which was compared with ML methods for time series classifications. As can be seen, the works are for specific diseases, or to determine behaviors, and do not propose a self-adaptive scheme of ML models, as is the basis of our proposal.

This work is relevant because it incorporates adaptive ML modeling processes, which allows the model to evolve and improve over time by including new data or changes in the underlying system. In addition, the model has the ability to transfer its knowledge to new tasks and can learn incrementally. These features make it useful for a variety of applications, including the detection of anomalies in cattle weights, which could have a significant impact on the cattle industry.

7 Conclusion and future work

While there are many applications of ML in livestock, the use of MTL in this industry is new and limited. In this paper, we have proposed an MTL architecture that can be used for PLF. Our architecture can use different ML models. The main advantage of our proposal is the ability to learn continuously as new data or information is received. In addition, it can update the existent models in its knowledge base, and its parameters, as new data is received. As well, our proposal can consider new datasets, maybe the reuse of trained models (it selects the most favorable one). Thus, one of the most outstanding qualities of our MTL architecture is its ability to adapt to new and old data over time.

Notably, our MTL architecture is the first to be used for cattle weight anomaly detection in the cattle industry. The metamodel knowledge base is continually updated to improve the identification/prediction models based on ML techniques over time. Our approach is effective in detecting livestock weight anomalies but can be adapted to different tasks.

However, our architecture also has some limitations that need to be addressed. First, it requires a large amount of data to train the initial models for the knowledge base, which may be difficult to obtain in some contexts. Second, it relies on the assumption that the ideal weight growth curve for a specific breed of cattle is similar in different contexts, which may not always be the case. Although the architecture we propose has some limitations, we believe that it can be used successfully in other scenarios of livestock management and animal productivity.

Future work is oriented to the development of specific autonomic cycles that allow self-monitoring and autonomous self-management processes in the context of PLF. In addition, another future work will be to test different synthetic data generation algorithms and evaluate their performance and behavior in our architecture.

Declarations

Data Availability Statement

Data will be made available in reasonable request

Funding and/or Conflicts of interests/Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Rodrigo García is supported by a PhD grant financed by Universidad EAFIT. All the authors would like to thank the "Vicerrectoría de Descubrimiento y Creación", of Universidad EAFIT, for their support on this research. Also, the authors want to thank the HAMADI 4.0 project, code 22-STIC-06, of the STIC-AmSud regional program, for the support of this work. Finally, the authors want to thank the student Manuela Zapata Mesa, from the Department of Mathematical Sciences, EAFIT University, for the scientific assistance provided during the realization of this work.

References

[1] N. C. Friggens, C. Disenhaus, and H. V. Petit, "Nutritional sub-fertility in the dairy cow: towards improved reproductive management through a better biological understanding," animal, vol. 4, no. 7, p. 1197–1213, 2010.

[2] R. García, J. Aguilar, M. Toro, and M. Jiménez, "Weight-identification model of cattle using machine-learning techniques for anomaly detection," in 2021 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 01–07, IEEE, 2021.

[3] D. Berckmans, "Precision livestock farming technologies for welfare management in intensive livestock systems," OIE Revue Scientifique et Technique, vol. 33, no. 1, p. 189 – 196, 2014. Cited by: 182.

[4] L. Ljung, "Perspectives on system identification," Annual Reviews in Control, vol. 34, no. 1, pp. 1–12, 2010.

[5] R. García, J. Aguilar, M. Toro, A. Pinto, and P. Rodríguez, "A systematic literature review on the use of machine learning in precision livestock farming," Computers and Electronics in Agriculture, vol. 179, p. 105826, 2020.

[6] H. F. Gouda, F. A. M. Hassan, E. E. El-Araby, and S. A. Moawed, "Comparison of machine learning models for bluetongue risk prediction: a seroprevalence study on small ruminants," BMC Veterinary Research, vol. 18, no. 1, 2022. Cited by: 0; All Open Access, Gold Open Access, Green Open Access.

[7] C. Benitez, R. García, J. Aguilar, M. Jiménez, and H. Robles, "Supervision system of the fattening process of cattle in rotational grazing using fuzzy classification," in 2022 XVLIII Latin American Computer Conference (CLEI), pp. 1–10, 2022.

[8] Y. Ma, S. Zhao, W. Wang, Y. Li, and I. King, "Multimodality in meta-learning: A comprehensive survey," Knowledge-Based Systems, vol. 250, p. 108976, 2022.

[9] J. Aguilar, M. Jerez, E. Exposito, and T. Villemur, "CARMiCLOC: Context Awareness Middleware in Cloud Computing," in 2015 Latin American Computing Conference (CLEI), pp. 1–10, 2015.

[10] J. Aguilar, A. Garcès-Jimènez, N. Gallego-Salvador, J. A. G. De Mesa, J. M. Gomez-Pulido, and J. Garcìa-Tejedor, "Autonomic Management Architecture for Multi-HVAC Systems in Smart Buildings," IEEE Access, vol. 7, pp. 123402–123415, 2019.

[11] G. Morota, R. V. Ventura, F. F. Silva, M. Koyama, and S. C. Fernando, "Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture1," Journal of Animal Science, vol. 96, pp. 1540–1550, 03 2018.

[12] J. Moon, Y. Noh, S. Jung, J. Lee, and E. Hwang, "Anomaly detection using a model-agnostic meta-learning-based variational auto-encoder for facility management," Journal of Building Engineering, vol. 68, 2023. cited By 0.

[13] A. Entezami, H. Sarmadi, and B. Behkamal, "Long-term health monitoring of concrete and steel bridges under large and missing data by unsupervised meta learning," Engineering Structures, vol. 279, 2023. Cited by: 0.

[14] C. Peng, M. Zhang, W. Hu, and L. Wang, "A fast cross-band spectrum anomaly detection algorithm based on meta-learning," p. 420 – 424, 2022. Cited by: 2.

[15] J. Tan, T. Kart, B. Hou, J. Batten, and B. Kainz, "Metadetector: Detecting outliers by learning to learn from self-supervision," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 13166 LNCS, p. 119 – 126, 2022. Cited by: 0; All Open Access, Green Open Access.

[16] G. M. Tavares and S. B. Junior, "Process mining encoding via meta-learning for an enhanced anomaly detection," Communications in Computer and Information Science, vol. 1450 CCIS, p. 157 – 168, 2021. Cited by: 2.

[17] E. M. Dogo, N. I. Nwulu, B. Twala, and C. Aigbavboa, "Accessing imbalance learning using dynamic selection approach in water quality anomaly detection," Symmetry, vol. 13, no. 5, 2021. Cited by: 5; All Open Access, Gold Open Access.

[18] Y. Cai, L. Ma, and G. Liu, "A night-time anomaly detection system of hog activities based on passive infrared detector," Applied Engineering in Agriculture, vol. 35, no. 4, pp. 481–493, 2019.

[19] J.-B. Perrin, B. Durand, E. Gay, C. Ducrot, P. Hendrikx, D. Calavas, and V. Hénaux, "Simulation-based evaluation of the performances of an algorithm for detecting abnormal disease-related features in cattle mortality records," PLoS ONE, vol. 10, no. 11, 2015. Cited by: 5; All Open Access, Gold Open Access, Green Open Access.

[20] E. Kramer, D. Cavero, E. Stamer, and J. Krieter, "Mastitis and lameness detection in dairy cows by application of fuzzy logic," Livestock Science, vol. 125, no. 1, pp. 92–96, 2009. cited By 49.

[21] S. Ma, Q. Yao, T. Masuda, S. Higaki, K. Yoshioka, S. Arai, S. Takamatsu, and T. Itoh, "Development of an anomaly detection system for cattle using infrared image and machine learning," Sensors and Materials, vol. 32, no. 12, p. 4139 – 4149, 2020. Cited by: 3; All Open Access, Gold Open Access.

[22] J. Haladjian, Z. Hodaie, S. Nüske, and B. Brügge, "Gait anomaly detection in dairy cattle," vol. Part F132525, 2017. Cited by: 7.

[23] N. Wagner, V. Antoine, J. Koko, M.-M. Mialon, R. Lardy, and I. Veissier, "Comparison of machine learning methods to detect anomalies in the activity of dairy cows," in Foundations of Intelligent Systems (D. Helic, G. Leitner, M. Stettinger, A. Felfernig, and Z. W. Raś, eds.), (Cham), pp. 342–351, Springer International Publishing, 2020.

[24] Y. Chung, J. Lee, S. Oh, D. Park, H. Chang, and S. Kim, "Automatic detection of cow's oestrus in audio surveillance system," Asian-Australasian journal of animal sciences, vol. 26, no. 7, p. 1030–1037, 2013.

[25] A. Calera, F. Lozano, I. Campos, and A. Osann, "Selected farming services with satellite assistance: Mapping crops, optimizing water resources and assessing crop and pasture drought stress," pp. 62–72, 2013. cited By 0.

[26] H. Park, D.-H. Park, and S.-H. Kim, "Deep learning-based method for detecting anomalies of operating equipment dynamically in livestock farms," in 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1182–1185, 2020.

[27] D. Briske, J. Derner, J. Brown, S. Fuhlendorf, W. Teague, K. Havstad, R. Gillen, A. Ash, and W. Willms, "Rotational grazing on rangelands: Reconciliation of perception and experimental evidence," Rangeland Ecology Management, vol. 61, no. 1, pp. 3–17, 2008.

[28] H. Bansal, B. Chinagundi, P. S. Rana, and N. Kumar, "An ensemble machine learning technique for detection of abnormalities in knee movement sustainability," Sustainability, vol. 14, no. 20, 2022.

[29] K. A. Segerkvist, J. Höglund, H. Österlund, C. Wik, N. Högberg, and A. Hessle, "Automatic weighing as an animal health monitoring tool on pasture," Livestock Science, vol. 240, p. 104157, 2020.

[30] N. Wagner, V. Antoine, M.-M. Mialon, R. Lardy, M. Silberberg, J. Koko, and I. Veissier, "Machine learning to detect behavioural anomalies in dairy cows under subacute ruminal acidosis," Computers and Electronics in Agriculture, vol. 170, p. 105233, 2020.