Improving Farm Management Information Systems with Data Mining

H. J. M. HOVING, A. KASSAHUN and C. CATAL*

Abstract— Over the past several years, farm enterprises have grown in size substantially while their number has steadily declined. As the size of their farms grow more and more farmers are deploying information systems, commonly called as Farm Management Information Systems (FMIS), to manage the day to day activities of their farms. The deployment of FMIS enable farmers to capture detailed data that can potentially be analysed by data mining tools to provide valuable information for optimizing the farm enterprises. However, data mining is generally not a common feature of many FMIS. In order to evaluate the suitability of data mining for use in FMIS, two case studies were performed using data captured in FMIS and applying various data mining algorithms. Microsoft Azure Machine Learning Studio is chosen because it provides a simple drag-and-drop visual interface that can be used by farm domain experts. In this study, two common problems were addressed in dairy farming: calving prediction of dairy cows and prediction of lactation value of milking cows. In both cases data mining models were built and experiments were run and results in both cases indicate that the required data is available from FMIS and data mining techniques provides acceptable performance. It was also shown that farm domain experts can easily use a user-friendly and drag-and-drop data mining tools with minimal initial training. Based on the insight from the two case studies and literature study, several decision problems that can be addressed with data mining such as heat prediction and lameness prediction were identified.

Index Terms— Farm management information systems, machine learning, calving prediction, lactation prediction.

I. INTRODUCTION

THE USE of data, and thus evidence, for decision making makes decisions systematic, unbiased, more accurate, and effective [1, 3, 6, 20].

While experience-based decision making by an expert may not require much data, but it definitely requires many years of observation and practice and is, therefore, very expensive. It is

HENK JAN HOVING, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands,(e-mail: henk-jan.hoving@wur.nl)
https://orcid.org/0000-0001-7990-4237
AYALEW KASSAHUN, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands, (e-mail: ayalew.kassahun@wur.nl) https://orcid.org/0000-0003-1066-7127
CAGATAY CATAL, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands,(e-mail: cagatay.catal@wur.nl)
https://orcid.org/0000-0003-0959-2930

also very time-consuming and labour-intensive as the expert has often to be physically present and make various measurements and observations.

When appropriate data is available, it is easier, cheaper, and often more accurate, to make decisions using computers and software systems. Moreover, the decision can be made automatically, and thus quickly. Therefore, data mining can be a very useful tool. Traditionally, many farmers make decisions mainly based on their personal experiences enhanced by a limited amount of data gathered in their paper-based diary. These decisions are called experience-based or simply expert decisions.

Recently more and more farmers are using management information systems that support simple and routine decisions. The management support systems made for farm enterprises are generally called *Farm Management Information Systems (FMIS)* [32]. FMIS generally gather a large amount of data on a daily basis from several sensors used in the farm such as milking robots and tag readers. Such a large amount of data, nowadays referred to as *big data*, can enable making non-trivial decisions, such as predicting expected calving date in dairy farming [25, 17]. In order to extract the right knowledge and patterns from big data and potentially make non-trivial decisions, several technologies including data mining, machine learning, and deep learning can be applied.

Data mining is defined as discovering new knowledge from the data [23]. It addresses a set of methods which are used to find correlations, patterns, and interesting relations between different data points in large databases. The use of data mining within FMIS can help automate decision making and even help discover new relationships among unanticipated factors. However, data mining is still not a common feature of FMIS, and it is not clear to what extent data mining can be integrated with FMIS and help address farm decision making problems and which problems it can solve effectively using the data that is commonly available within FMIS.

Data-driven decision making in comparison to expert decision making requires diverse kinds of data. An example of a decision problem that can be solved with data-driven decision making is the estimation of how long a lactation cycle of a cow will last based on the milk yield per cow [13]. Most farmers are currently not aware of how automated decision making will impact their daily life and the productivity of the cows. Naturally, there are several farmers who want to understand the underlying decision-making processes, gather the required data, deploy data mining models and take full advantage of the tools. When farmers understand how data-driven decision making works, they can make substantiated

decision on the investment of new sensors and devices and thereby, improve the state-of-the-art.

There are currently easy to use platforms for applying data mining algorithms by domain experts, of which Microsoft Azure Machine Learning Studio (MAMLS, https://studio.azureml.net/) is the most prominent of such tools. MAMLS provides many data mining algorithms out-of-the-box and requires no prior programming expertise to build data mining models and perform experiments. Expert users can customize MAMLS the data mining models using the R programming language but as one of the aims of this research is also to evaluate to what extent farm domain experts (specialists in farm and expert users of FMIS) can utilize data mining platforms, only the out-of-the-box algorithms of MAMLS has been used.

The uncertainties related to the above-mentioned problems make it difficult for farmers to make investments in data capturing devices and data mining tools. To address these problems, the following main research question and sub questions were formulated:

- *RQ:* How can FMIS be improved with data mining?
  - *RQ1:* Which problems at farms can be solved with data mining techniques?
  - *RQ2:* How can we these problems be mapped into data mining tasks?
  - *RQ3:* How can Azure ML Studio be applied by a domain expert?
  - *RQ4:* What kind of data is gathered at the farm and is more suitable for data mining?

The first sub question was addressed through literature search and supplemented with information from interviews and discussions with farmers. These interviews and discussions allowed us to get multiple points of view on the problems. The second sub question was addressed by using a data mining process and data mining tasks retrieved through literature search. Also, the case studies were used to help address this sub research question. The third sub-question is addressed by evaluating problems and data in detail and making several experiments by a domain expert, who is the first author of this article. The last sub question is addressed by both through literature research and discussion with different farmers, particularly, to investigate if data required for applying data mining is available in their FMIS.

To answer these research questions, the livestock sector and mainly on dairy farming were focused on. Particularly, two case studies were conducted in this research. The first case study was on the use of data mining algorithms for calving prediction problem, which is evaluated as a regression problem. The second case study focused on the estimation of milk production. Also, results of two interviews with farmers in the Netherlands are presented. Several experiments were conducted not only on public datasets, but also on our own dataset collected from a family dairy farm Veefokbedrijf Hoving that is associated with the first author.

The main contributions of this study are shown as follows:

- It was demonstrated that data mining provides acceptable results for calving prediction and the estimation of milk production.

- It was gathered a new dataset for predicting the milk production in a dairy farm and perform several experiments on Azure ML Studio platform.
- It was presented several tasks from dairy farming domain, which can be addressed by data mining techniques and tools.
- Based on our interviews in the Netherlands, it was concluded that farmers are already collecting sufficient data which can be used for data mining experiments.

The reminder of this paper is organized as follows: Section II explains the background and related work. Section III presents the methodology and the Section IV shows the experimental results. Section V presents the discussion and Section VI shows the conclusion and future work.

II. BACKGROUND AND RELATED WORK

A. Farm Management Information Systems (FMIS)

Essentially an FMIS is a Management Information System for farm enterprises. The key features of FMIS include financial management, reporting (also related to regulatory requirements), data acquisition, and the planning and management of farm operations, resources and people [39].

As in any other business, farmers too aim to minimize their production costs and maximize their yield [35]. The deployment of a management system helps to address this expectation by gathering, processing and synthesizing information following state-of-the-art management practices. Without an FMIS, managing a farm is a time-consuming, expensive, and labour-intensive task.

There are currently various types IoT (Internet of Things) devices such as sensors, actuators and machines used in farms. These devices gather vast amounts of data that are not captured or well-managed by the conventional FMIS systems that are widely used. Unfortunately, therefore, many FMIS do not make full use of the available data within farm enterprises [24].

In dairy farming a large amount of data is collected during the milking process of cows. The data includes the amount of milk, fat percentage, protein and lactose content, amount of urea, and somatic cell score. Based on this data, it is possible to predict how much milk the cows will be able to produce during the rest of the lactation cycle and make various decisions. For instance, a decision variable used in dairy farming is *a lactation value*. A lactation value is an indication for how well a cow is compared to the herd and can be computed based on the amount of fat, protein, and lactose content of the milk. An average cow has a lactation value of 100; if a cow performs worse than the average, its lactation value is less 100, otherwise its lactation value is equal to or above 100. Such information can be used to group the herd based on their milk production performance. The required data for determining lactation value is registered by software system that is associated with the milking machine.

Yet another example of decision variable is the Estimated Breeding Value (EBV). EBV is the combination of characteristics that are measured and marked by a livestock inspector. Unlike the lactation value, EBV is labour intensive activity, the results of which are registered in an FMIS. The

characteristics registered by the inspector include height, capacity, condition, legs, claws, and udder, and describe the phenotype of the cow. The goal for a farmer is to improve the EBV and the genetic superiority of individual cows and the overall herd.

The optimal use of data can be made when the various systems used at farm are integrated to the FMIS the farmer uses [20, 6]. The deployment of the systems and their integration with the FMIS requires that farmers understand what useful information can be extracted from the data and which problems the derived information can help address reliably.

B. Data Mining

Data mining field includes several techniques which make it possible to analyse large datasets in a short time. It helps to find different patterns in the data automatically but leaves the noise out [34]. Data mining has been used in many applications like market analysis and management, risk analyses, and fraud detection [23]. During the processing of the data, a part of the data is used to "train" the algorithm. This data is called the training data, which lets the algorithm to determine its parameters in order to identify different patterns, correlations, and anomalies within the data. After the training the algorithm, it is necessary to test it using the rest of the data set in order to make sure that training is reliable. After training and testing, the algorithm can be transformed into a software module and deployed as a web service in cloud platform.

Learning algorithms in data mining are divided into the following four categories:

1. *Supervised learning:* For supervised learning, training data should be labelled by considering different classes which exist in the data.
2. *Unsupervised learning:* In unsupervised learning, there is no output value, and only features / independent variables exist. The goal of unsupervised learning is to get a quick answer about the data [15]. It aims to infer a hidden structure in the data [26]. It also omits the noise which can influence the model [34].
3. *Semi-supervised learning:* Most of the time it is difficult to use supervised learning because not all the data is labelled. It is possible to give labels to all those data points, however, in practice this will be very expensive and time-consuming [14]. Semi-supervised learning algorithms can be used if there is very limited labelled data such as 10-20% and the remaining part of the dataset (80-90%) is unlabelled. The algorithm first tries to assign pre-labels to the unlabelled data and then, after several iterations, these labels are used in conjunction with the labelled data.
4. *Reinforcement learning:* During reinforcement learning, the agent observes the environment, performs some actions within the environment, and gets some rewards (either positive or negative) due to these actions [33].

One of the well-known and widely applied process is *Cross-Industry Standard Process for Data Mining (CRISP-DM).* The iterative process model contains several steps as shown in Figure 1 [26]. All the steps of this process are explained as follows and it starts with the business understanding step.
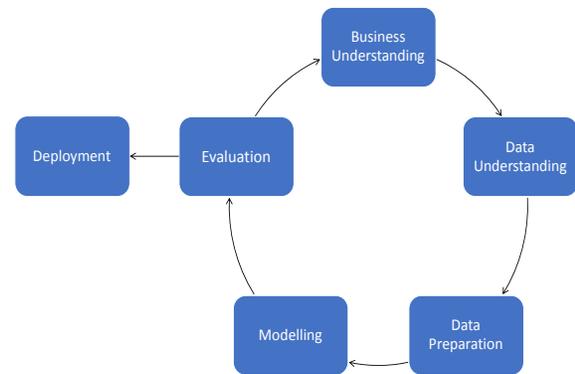

Fig. 1 CRISP-DM process (adapted from Provost & Fawcett [26])

1. *Business understanding:* This step helps to transform the business problem into one of data mining tasks.
2. *Data understanding:* After transforming the business problem into one of the data mining tasks, a dataset is required to build the models. In this step, the main task is to understand the format of the data and know how the relevant data can be collected.
3. *Data preparation:* In this step, all the raw data is prepared for modelling, data is cleaned up if necessary, and data is converted into the appropriate format.
4. *Modelling:* During this step, one of the learning algorithms is applied.
5. *Evaluation:* During the evaluation phase, the dataset is generally divided into two parts: a training set and testing set. This evaluation approach is called *hold-out*. In addition to the hold-out evaluation strategy, there is K-fold cross-validation approach which divides the dataset into K folds, which is mostly selected as 10. In each iteration, training is performed based on the (K-1) folds and the testing is done on the resting fold. This is repeated K times, and the average of values are calculated to determine the performance of the model.
6. *Deployment:* After the model is determined at the end of the evaluation phase, the model should be deployed into a web server or a cloud platform to let client applications easily access to this prediction model.

C. Data Mining Tasks

In data mining, there are hundreds of data mining algorithms. However, compared to the number of data mining algorithms, the number of data mining tasks are very less and very limited. There exist nine data mining tasks in data mining [26]. These tasks are explained as follows and are adaptations from [26]:

1. *Classification:* This task predicts for each individual data point in which class that point belongs to.
2. *Regression:* In regression, the aim is to predict a numerical a value which is actually a dependent variable.
3. *Similarity matching:* The goal of similarity matching is to find items (such as people, animals or other objects) that are similar to each other.

4. *Clustering:* Clustering techniques group individuals based on their common features, but are not driven by any specific purpose.
5. *Co-occurrence grouping:* Co-occurrence grouping finds associations between individuals, which is also known as *market-basket analysis*, *frequent item set mining*, and *association rule discovery*.
6. *Profiling:* Profiling, known as anomaly detection, characterizes the typical behaviour of an individual or a group and finds the abnormality.
7. *Link prediction:* Link prediction predicts the new links between groups or individuals.
8. *Data reduction:* In data reduction, most of the time the number of features of the dataset is reduced, but it is also possible to reduce the data points if they are detected as noisy instances.
9. *Causal modelling:* Causal modelling tries to understand what kind of events or actions influence certain events.

D. Related Work

Shahinfar et al. [29] predicted insemination outcomes in Holstein dairy cattle using data mining algorithms. Their objective was to create a user-friendly tool to help farmers make decisions. Rutten et al. [28] built a model to estimate the impact of the change of the first insemination. Fenlon et al. [12] worked on the estimation of insemination outcome by using data mining techniques for Irish dairy cows. Later, they used data mining algorithms to estimate the alignment for insemination in farms which use a seasonal calving system [12]. They stated that it is important that peak grass availability with peak lactating cow energy balances are aligned. Mahmoud et al. [22] used the RumiWatch device for estimating the calving time. By using the data about the rumination behaviour, it was possible to estimate the calving time. Zaborski et al. [38] used data mining techniques to detect dystocia in dairy cattle. Their approach gives a clear view which heifers get troubles during calving.

Borchers et al. [4] applied artificial neural networks for calving prediction. For gathering data, they used the HR Tag device of the company SCR Engineers and the data contained information about the rumination behaviour, neck activity, number of steps, and lying time.

For estimating the sickness in herds, data mining has been previously used. Different algorithms were used to predict the sick cows by using data containing grazing, standing, and rumination behaviour [27]. Yazdanbakhsh et al. [37] performed a research on sickness in an Alberta feedlot. They were capable of finding illness up to seven days in advance by using data mining techniques and an inferential sensor. Caraviello et al. [5] used several data mining methods to estimate the effect of factors on the reproduction of lactating Holstein cows. Kim and Heald [21] applied data mining techniques to get an idea of which bacterial was causing mastitis in a dairy herd. Kamphuis et al. [19] applied decision trees for the detection of clinical mastitis in farms using automatic milking. Alsaaod et al. [2] worked on the on lameness prediction problem, and predicted accurately the early state of lameness of cows.

Heat detection is also important in dairy businesses. Heat detection using accelerometers and unsupervised learning has been performed by Shahriar et al. [30]. Vanrell et al. [35] applied accelerometers to predict the heat events in dairy cows using decision trees algorithms. Chung et al. [7] used small microphones to collect data about the heat events and applied data mining to build models.

III.   METHODOLOGY

This research required data gathering for data mining, experimenting with various data mining methods and tasks, and finally, interviewing farmers to assess the suitability of data mining in supporting their decision making. Each of these methods are explained, starting with experimental design, followed by the data collected and then, the interviews. To design the experiments, MAMLS's graphical and experiment designer were used. Figure 2 shows the experimental design built for the calving prediction case study (Case Study-I).

The experimental design depicted in Figure 2 presents the following steps. The data set *days.csv* is uploaded to MAMLS. This dataset is split into two parts (one part is used for training and another for testing) with *Split Data* visual component. Training is performed in the *Train Model*. The inputs to Train Model are the outputs of the Split Data component and a data mining algorithm (in this case the *Neural Network Regression* algorithm). The result of the training is tested using the *Score Model* component. Besides the output of Train Model, the Score Model component requires the second output of Split Data. Finally, in *Evaluate Model* component uses the output of the Score Model to determine and present several evaluation parameters.

For regression tasks the following algorithms were applied:

- *Bayesian networks:* This algorithm applies probability theory for its calculations, which is based on random variables and their conditional independencies.
- *Decision tree:* This is one of the most simple but useful algorithms in data mining. It starts with a node at the base and extends to several leaf nodes which show the classes the tree can classify. Overfitting is one of the drawbacks of this algorithm. There are different implementations of this algorithm and one of them is boosted decision trees. Another example is the Random Forest algorithm which uses multiple trees for the decision.
- *Linear regression:* This algorithm aims to find a relationship between inputs and the output and it can deal with multiple input variables.
- *Neural Networks (NN):* NN is an algorithm inspired by the biological neural networks which are complex systems that contain many neurons. There are different topological models of NNs such as multi-layer perceptron and recurrent neural networks. Nowadays, deep learning algorithms, which are based on NNs, are widely used in different application domains.
- *Poisson Regression:* The Poisson regression, which is a type of generalized linear model, is based on the Poisson distribution.

The suitability of a data mining algorithm for a given problem and dataset is determined by evaluation parameters. The

following evaluation parameters have been used for classification and regression problems in this study:

- *Mean absolute error:* It measures the error between two continuous variables such as X and Y. It can take values between zero and infinity.
- *Root mean squared error:* This value provides the standard deviation of the residuals. It is between zero and infinity.
- *Relative squared error:* It gives a normalized squared error which means that the relative squared error takes the total squared error and normalizes it by dividing with the total squared error of the prediction. The result can be between zero and infinity.
- *Relative absolute error:* This is very similar to the relative squared error, but it is not squared. To make the outcomes positive, the formula contains absolute sign, therefore the results are between zero and infinity.
- *Coefficient of determination:* This parameter is also known as R^2. Adding more variables to the model will not necessarily reduce the coefficient of determination, but could indeed reduce prediction accuracy by introducing prediction variance [18]. It is between 0 and 1 and when it is near to 1, the model is said to be perfect. If it is negative, it means that the model cannot represent the underlying data.
- *Overall accuracy:* This parameter provides the non-weighed accuracy of all the classes combined.
- *Average accuracy:* The average accuracy is the weighed accuracy for all the classes, which indicates the goodness of a classification model as a proportion of true results to total cases.
- *Precision:* Precision provides an indication for the correct positive observations. This can be calculated from a confusion matrix, the following formula is used:
  Precision=true positives/(true positives + false positives)
- *Recall:* The recall is a synonym for sensitivity, which is the ratio of correctly predicted positive events. The formula is shown as follows:
  Recall=true positive / (true positives + false negatives)

A. Experimental Design for Calving Prediction

As shown in Figure 2, an experiment has been prepared for the calving prediction problem. In this experiment, 80% of the dataset has been used for the training and the rest has been used for testing. In Figure 3, it is shown how we can design an experiment to perform 10-fold cross validation analysis. This time *Partition and Sample* component was used instead of *Split Data* component.

B. Experimental Design for Predicting Lactation Value

For the second case study, data obtained from dairy farm Veefokbedrijf Hoving was used. At the beginning, it was aimed to estimate the lactation value of individual cows using regression algorithms. However, lactation values obtained were not sufficiently accurate; therefore, it was resorted to building a multi-class classification model.

The data points were divided into three classes: bad-0 (lactation values below 95), neutral-1 (lactation values between 95 and 105), and good-2 (lactation values over 105). In this experiment, data was obtained as a PDF file, which is

the only data format some FMIS use for data exchange. Therefore, first the data had to be transformed into a csv file.
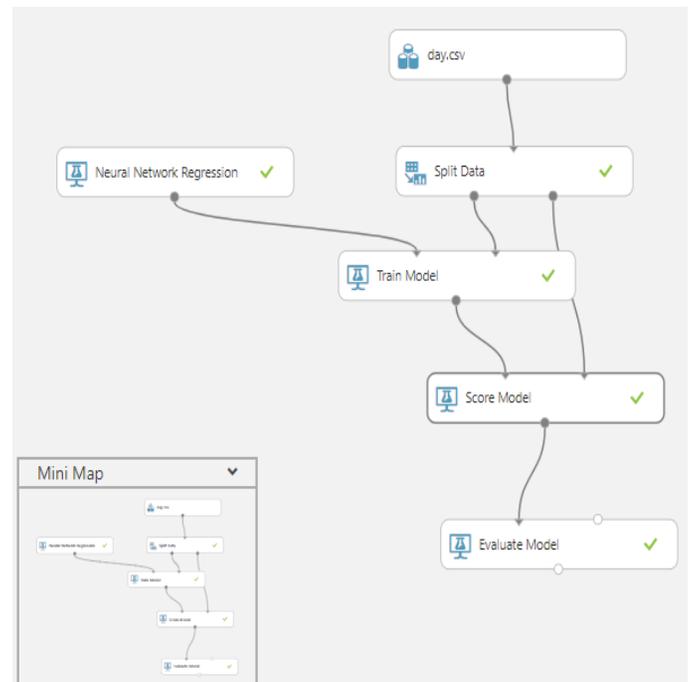

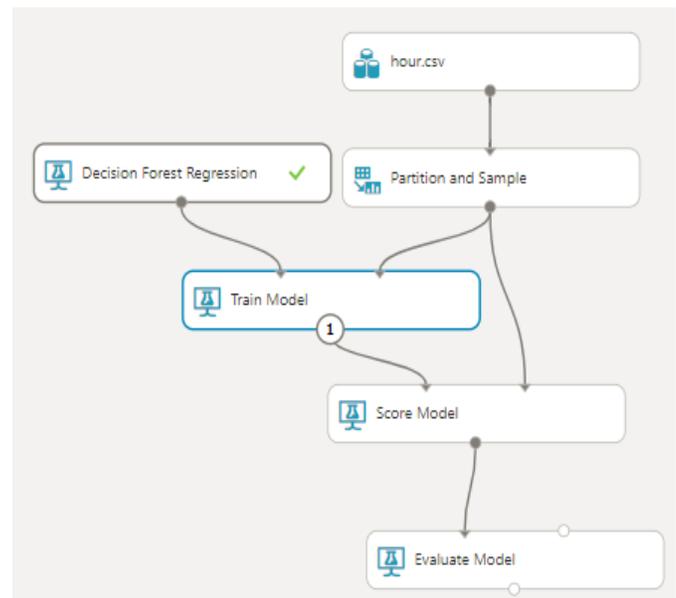Fig. 2. Experiment for calving prediction case study in Azure ML Studio

Fig. 3. Experiment for 10-fold cross-validation

Since many features were not useful, *Select Columns in Dataset* component has been used during the experimental design as shown in Figure 4. The following useful columns were selected: *Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, and class label*. The dataset was split into 80% training data and 20% testing data. First multiclass neural network algorithm has been used for the analysis, and then, multiclass forest, multiclass decision jungle, and multiclass logistics regression algorithms were applied.
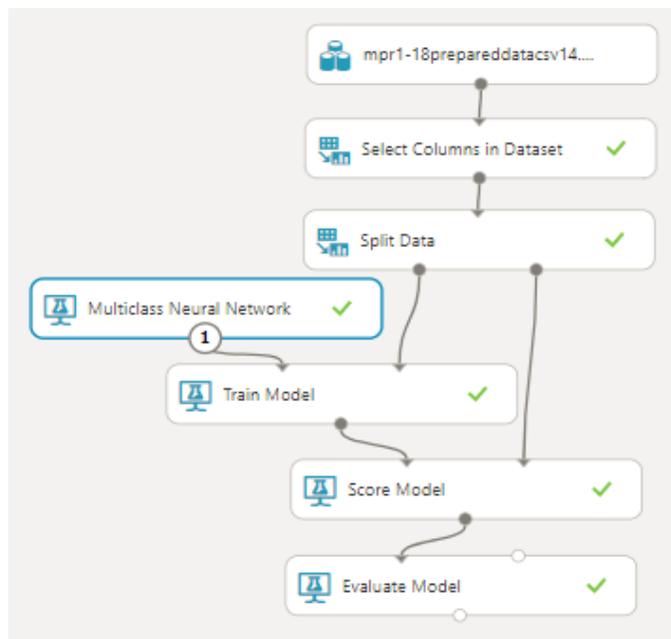
Fig. 4. Experiment for prediction lactation value (multi-class classification)

To get more information about the quality of lactation values, authors contacted CRV (https://www.crv4all.nl/), the company that managed the data. The company explained that the lactation values are partly measured and partly estimated. The company used basic lactation curve and added measurement values from individual cows. These measurements contain data about milk yield, percentage fat, percentage protein, percentage lactose, and the amount of urea. Furthermore, information about the cow, such as age, calving date, and lactation number are used with the model the company used to estimate lactation value. For estimating productions between two measurements, interpolation is applied [11]. By using the model output of the CRV, in this case study authors were measuring how well data mining results are matching with the classical methods of computing the productivity of a cow in the future and thereby showing the potential of data mining.

After building a classification model, it was aimed to build a regression model for this dataset. During the regression experiments, linear regression, Bayesian linear regression, boosted decision tree regression, decision forest regression, Poisson regression, and neural network algorithm were used.

C. Data Collection

For calving prediction experiments, two datasets (day.csv and hour.csv) shared by Borchers et al. [4] were used. One of them contains data for the prediction on which day the cow will calf and the other one contains data for estimating on which hour of the day the cow will calf. The data contains minute-by-minute data log of activity, lying, and ruminating behaviour of every dairy cattle. This data is gathered using sensors that are capable of measuring all these different activities. The data contains the monthly values milk production of all the cows, including the fat, protein, lactose, protein, and urea contents of the milk. Dataset can be accessed from the following public github link: https://github.com/Mrborchers/Machine-learning-based-calving-prediction-from-activity-lying-and-rumination-behaviors

D. Interviews

The dairy farm Veefokbedrijf Hoving has a herd of 60 milking cows and an old-school milking parlour. The interviewee measures the milk production with a monthly MPR (milk production registration). At the time of the interview, the farmer was content with the FMIS he has in place but would also be happy if he has a way of predicting certain events because predicting events would make his life easier. Therefore, in order to understand the value of prediction, it was asked to him how much money he would spend per cow if such a system can be built for the farm. He replied that he would spend €500 per cow. Several tasks were listed for him and he provided the following order of priority: prediction of sickness, heat detection, calving prediction, and lameness detection.

The second interviewee works part-time in a family farm which has a herd of 50 cows and a milking robot. Therefore, he has daily information about the milk production and concentrate consumption of every individual cow. The data about milk production is comparable with data from MPR, however it is daily instead of monthly. Furthermore, the cows have activity sensors on their neck and he plots crowing curves for the calves. At the time of the interview, he uses data from milking robot [9, 10] to make decisions about the heat activities, sickness, and lameness. Using the calve growing curves, he follows the growth of all the calves and if they do not grow optimally he changes the feeding patterns. He would like to see some reports that show which cows will have a heat. Though visual heat detection is manageable for 50 cows, it will be difficult when the number of cows reaches to 200. In the current circumstances, he would like to identify cows in heat that needs special attention—*i.e.* he wants to identify the outliers in heat behaviour. When it was asked how much he can spend for each cow, he responded that this is dependent on how much the sensors would return and how the system will present the necessary early events and warnings. His ranking for the tasks was as follows: prediction of sickness, heat detection, lameness, and calving prediction. Based on our interviews, it was observed that the prediction of sickness and heat detection have higher priority compared to the other tasks.

IV.   EXPERIMENTAL RESULTS

A. Case Study 1- Calving Prediction

In Table 1, the values regarding the five evaluation parameters for the 6 regression algorithms are presented. Of the 6 different regression algorithms linear regression and Bayesian linear regression algorithms resulted a perfect coefficient of determination (which is 1), which means that the two algorithms work perfectly on the given datasets. Also, the mean absolute errors for these two algorithms are very low. The other regression algorithms are not as good as the first two as can be seen from the values of the coefficient of determination (less than one) and the mean absolute errors (greater than 1). In addition to the hold-out approach, Cross-

Validation (CV) analysis was also used. The results of cross validation are presented in Table 2. A similar performance for linear regression and Bayesian linear regression algorithms are observed, both of which reach to the value 1 for coefficient of determination parameter.

The results for the prediction in which hour the cows will calf are shown in Table 3. Based on this table, it can be stated that compared to the other algorithms the Boosted Decision Tree Regression and Decision Forest Regression algorithms perform the best. In Table 4 the results of predicting the hour of calving using 10-fold CV are presented. The results are very similar to the results of Table 3. Boosted Decision Tree Regression and Decision Forest Regression algorithms perform better than the other algorithms.

TABLE I.

CALVING PREDICTION RESULTS BASED ON DAILY DATA

Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.000247	0.000026	1.43	2.2	2.51	2.91
RMSqE	0.00032	0.000034	1.89	2.86	3.32	3.42
RelAbsE	0.000065	0.000007	0.375	0.549	0.658	0.765
RelSqE	0.	0.	0.184	0.392	0.566	0.619
CoefDet	1.	1.	0.816	0.608	0.434	0.381

LEGEND:

MAbsE	Mean Absolute Error
RMSqE	Root Mean Squared Error
RelAbsE	Relative Absolute Error
RelSqE	Relative Squared Error
CoefDet	Coefficient of determination
LR	Linear regression
BayLR	Bayesian Linear Regression
BooDTR	Boosted Decision Tree Regression
DFR	Decision Forest Regression
NN	Neural Network
PR	Poisson Regression

TABLE II.

CALVING PREDICTION BASED ON DAILY DATA USING 10-FOLD CV

Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.000288	0.000196	1.56	2.23	2.7	3.79
RMSqE	0.000456	0.000346	2.1	2.83	3.53	4.16
RelAbsE	0.000076	0.000052	0.412	0.591	0.711	0.904
RelSqE	0.	0.	0.231	0.422	0.654	0.817
CoefDet	1.	1.	0.769	0.578	0.346	0.183

TABLE III.

CALVING PREDICTION RESULTS BASED ON HOURLY DATA

Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.263	0.202	0.	0.0499	0.23	1.47
RMSqE	0.254	0.249	0.	0.11	0.281	1.78
RelAbsE	0.295	0.288	0.	0.0701	0.329	0.24
RelSqE	0.0921	0.0891	0.	0.0172	0.113	0.0648
CoefDet	0.908	0.911	1.	0.983	0.887	0.935

TABLE IV.

CALVING PREDICTION BASED ON HOURLY DATA - 10-FOLD CV

Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.217	0.214	0.	0.0421	0.243	1.6
RMSqE	0.265	0.262	0.	0.094	0.298	1.93
RelAbsE	0.325	0.321	0.	0.063	0.363	0.284
RelSqE	0.106	0.104	0.	0.0134	0.134	0.0851
CoefDet	0.894	0.896	1.	0.987	0.866	0.915

B. Case Study 2- Predicting Lactation Value

For the second case study, experiments using multi-class classification algorithms and then based on regression algorithms were performed. For the classification models, the overall accuracy parameter can be used for selecting the right model. For the regression models, the coefficient of determination parameter is the indicator for selecting the right model. In Tables 5, the results of classification algorithms based on single measurement are presented and in Table 6, the results of regression algorithms based on single measurement are shown. In Table 7 and 8, the results of regression algorithms based on multiple measurements and classification results based on multiple measurements are shown, respectively. Based on these results, it can be stated that the performance of the multi-class classification models improved when multiple measurements are added to the dataset. As shown in Table 8, the average accuracy of the multiclass decision tree is around 0.88 which is quite high. This indicates that it can be determined which cows perform relatively poor, and which ones perform good in terms of milk production. The performance of regression algorithms, on the other hand, is not high enough for practical use, therefore, it is recommended the use of classification models for decision making at dairy farms. However, note that since the data was generated based on the lactation value computed by CRV (instead of values that are measured), sufficient insights to what extent the dependent variable is accurate are not known, and as a result it is not expected to reach a high coefficient of determination values for regression models.

Table 5 shows the results of the lactation value estimation with a classification model using only single measurement. The performance is not acceptable due to the low overall accuracy value. The best models are multiclass decision forest and multiclass neural network having the same overall accuracy value (0.625).

TABLE V.

PREDICTING LACTATION VALUE (CLASSIFICATION RESULTS)

Used features	Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, classification			
Algorithms →	McDF	McDJ	McLR	McNN
OA	0.625	0.583	0.542	0.625
AVA	0.75	0.722	0.694	0.75
MICAVP	0.625	0.583	0.542	0.625
MACAVP	0.622	0.575	0.432	NaN
MICAVRC	0.625	0.583	0.542	0.625
MACAVRC	0.625	0.569	0.515	0.524

LEGEND:

OA	OVERALL ACCURACY
AVA	AVERAGE ACCURACY
MICAVP	MICRO-AVERAGED PRECISION
MACAVP	MACRO-AVERAGED PRECISION
MICAVRC	MICRO-AVERAGED RECALL
MACAVRC	MACRO-AVERAGED RECALL
McDF	Multiclass Decision Forest
McDJ	Multiclass Decision Jungle
McLR	Multiclass Logistic Regression
McNN	Multiclass Neural Network

TABLE VI.

PREDICTING LACTATION VALUE (REGRESSION RESULTS)

Used features	Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, lactation value					
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	11.8	15.1	13.1	12.1	12.3	16.6
RMSqE	20.2	19.9	18.2	20.3	18.5	29.9
RelAbsE	0.643	0.826	0.714	0.66	0.674	0.908
RelSqE	0.475	0.461	0.386	0.48	0.398	1.04
CoefDet	0.525	0.539	0.614	0.52	0.602	0.0398

The accuracies of lactation value estimation with regression algorithms are shown in Table 6. The best model is boosted decision tree regression with a coefficient of determination of 0.61. The second best performing one is neural network algorithm. The low values of the coefficient of determination indicate that the results are generally unacceptable.

In Table 7, the results of estimation accuracies of the lactation value with regression algorithms are presented, but this time utilizing multiple measurements. As it can be seen, the algorithms performed poorly. This indicates that these models cannot predict the lactation value accurately. For boosted decision tree regression and decision forest regression the values of coefficient of determination are positive, but still very low. Compared to the Table 6, adding more data has not improved the model performance in this case.

The results for estimating the lactation value based on classification models are presented in Table 8. In this case, multiple measurements have been used. The overall accuracy increased when more measurements are added. The best performing model is multiclass decision forest in this case.

TABLE VII.

RESULTS WITH MULTIPLE MEASUREMENTS (REGRESSION)

Used features	Milk yield, % fat, % protein, % lactose, urea, age, average lactation value					
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	12.8	11.8	8.05	8.33	10.6	12.2
RMSqE	17.3	15.8	10.3	10.8	13.7	13.8
RelAbsE	1.16	1.06	0.729	0.755	0.96	1.1
RelSqE	1.75	1.46	0.622	0.676	1.09	1.12
CoefDet	-0.751	-0.455	0.378	0.324	-0.0894	-0.117

TABLE VIII.

RESULTS WITH MULTIPLE MEASUREMENTS (CLASSIFICATION)

Used features	Milk yield, % fat, % protein, % lactose, urea, age, average classification			
Algorithms →	McDF	McDJ	McLR	McNN
OA	0.818	0.682	0.727	0.636
AVA	0.879	0.788	0.818	0.758
MICAVP	0.818	0.682	0.727	0.636
MACAVP	0.73	0.583	0.678	0.571
MICAVRC	0.818	0.682	0.727	0.636
MACAVRC	0.857	0.762	0.81	0.952

V. DISCUSSION

In this study it was set out which problems at farm data mining can address and if farm domain experts can get the required data from an FMIS and translate them into data mining problems. It was also shown that data mining tools can be utilized by farm domain experts with little training. However, due to the empirical nature of this study, potential threats to validity must be addressed [31]. The four validity dimensions introduced by [36] are addressed, which in turn follow the work of Cook et al. [8]. The four validity dimensions we address are *conclusion*, *internal*, *external*, and *construct* validity.

To address conclusion validity threat, several evaluation parameters were used in this study. By presenting all the major evaluation parameters it is considered that the risk of other researchers applying a different evaluation parameter that will lead to a different conclusion about performance is reduced. In addition to the dataset from an FMIS for predicting the milk production, two public datasets for calving prediction were used. To strengthen the conclusion, additional experiments using more datasets from diverse FMIS are needed. Regarding the construct validity, the features in the public datasets were used as-is because those features have previously been used by the other researchers successfully. With regard to the dataset from an FMIS, the features were selected based on the domain knowledge. In order to minimize the possibility that different researchers reaching different performance results, various combination of features were investigated. External validity threat is about the limit of generalizability of the results of a case study [16]. Since the datasets are limited—one dataset belongs to only one farm in the Netherlands and the two public datasets belong only to one country, the performance of the models might change when

more datasets become available from different countries. To limit internal validity threats, a number of algorithms were applied for each problem and therefore, results are not limited by one or a few algorithms. The experiments are limited with the algorithms which exist in MAMLS and other researchers might achieve better performance with new algorithms that do not present in MAMLS.

VI. CONCLUSION

It was shown that a number of decisions at farm enterprises that until today required human expert can be supported or automated by data mining algorithms using data available in today's FMIS. For instance, high performance was achieved for the calving prediction problem. To make a hard conclusion on predicting the lactation value, more accurate lactation values are required. The other identified problems in this study are heat detection and lameness prediction. It was demonstrated that MAMLS can be used by a farm domain expert easily. It is planned to let a number of domain experts apply MAMLS on their own datasets in future research to reach more conclusive results.

(1)   References

[1] Albersmeier, F., Schulze, H., Jahn, G., & Spiller, A. (2009). The reliability of third-party certification in the food chain: From checklists to risk-oriented auditing. Food Control, 20(10), 927-935. doi:https://doi.org/10.1016/j.foodcont.2009.01.010

[2] Alsaaod, M., Römer, C., Kleinmanns, J., Hendriksen, K., Rose-Meierhöfer, S., Plümer, L., & Büscher, W. (2012). Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior. Applied Animal Behaviour Science, 142(3), 134-141. doi:10.1016/j.applanim.2012.10.001

[3] Azzaro, G., Caccamo, M., Ferguson, J. D., Battiato, S., Farinella, G. M., Guarnera, G. C., . . . Licitra, G. (2011). Objective estimation of body condition score by modeling cow body shape from digital images. J Dairy Sci, 94(4), 2126-2137. doi:10.3168/jds.2010-3467

[4] Borchers, M. R., Chang, Y. M., Proudfoot, K. L., Wadsworth, B. A., Stone, A. E., & Bewley, J. M. (2017). Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. J Dairy Sci, 100(7), 5664-5674. doi:10.3168/jds.2016-11526

[5] Caraviello, D. Z., Weigel, K. A., Craven, M., Gianola, D., Cook, N. B., Nordlund, K. V., . . . Wiltbank, M. C. (2006). Analysis of Reproductive Performance of Lactating Cows on Large Dairy Farms Using Machine Learning Algorithms. J Dairy Sci, 89(12), 4703-4722. doi:10.3168/jds.S0022-0302(06)72521-8

[6] Choubey, M. K. (2011). IT Infrastructure and Management (For the GBTU and MMTU): Pearson Education India.

[7] Chung, Y., Lee, J., Oh, S., Park, D., Chang, H., & Kim, S. (2013). Automatic detection of cow's oestrus in audio surveillance system. Asian-Australasian journal of animal sciences, 26(7), 1030.

[8] Cook, T. D., Campbell, D. T., & Day, A. (1979). Quasi-experimentation: Design & analysis issues for field settings (Vol. 351): Houghton Mifflin Boston.

[9] Delaval. (14-3-2018a). BCS

[10] Delaval. (14-3-2018b). View BCS.

[11] Delta, C. (February 1998). lactatieproductie en 305 dagenproductie. Handboek NRS.

[12] Fenlon, C., O'Grady, L., Doherty, M., Butler, S., Shalloo, L., & Dunnion, J. (2016, 12-15 Dec. 2016). Regression Techniques for Modelling Conception in Seasonally Calving Dairy Cows. Paper presented at the 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW).

[13] Grzesiak, W., Błaszczyk, P., & Lacroix, R. (2006). Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). Computers and Electronics in Agriculture, 54(2), 69-83.

[14] Hady, M. F. A., & Schwenker, F. (2013). Semi-supervised Learning. In M. Bianchini, M. Maggini, & L. C. Jain (Eds.), Handbook on Neural Information Processing (pp. 215-239). Berlin, Heidelberg: Springer Berlin Heidelberg.

[15] Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning The elements of statistical learning (pp. 485-585): Springer.

[16] Hosseini, S., Turhan, B., & Mäntylä, M. (2017). A benchmark study on the effectiveness of search-based data selection and feature selection for cross project defect prediction. Information and Software Technology.

[17] Husemann, C., & Novković, N. (2014). FARM MANAGEMENT INFORMATION SYSTEMS: A CASE STUDY ON A GERMAN MULTIFUNCTIONAL FARM.

[18] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). An Introduction to Statistical Learning with Applications in R. doi:DOI 10.1007/978-1-4614-7138-7

[19] Kamphuis, C., Mollenhorst, H., Feelders, A., Pietersma, D., & Hogeveen, H. (2010). Decision-tree induction to detect clinical mastitis with automatic milking. Computers and Electronics in Agriculture, 70(1), 60-68. doi:https://doi.org/10.1016/j.compag.2009.08.012

[20] Kim, C.-H., Weston, R. H., Hodgson, A., & Lee, K.-H. (2003). The complementary use of IDEF and UML modelling approaches. Computers in Industry, 50(1), 35-56.

[21] Kim, T., & Heald, C. W. (1999). Inducing inference rules for the classification of bovine mastitis. Computers and Electronics in Agriculture, 23(1), 27-42. doi:https://doi.org/10.1016/S0168-1699(99)00003-4

[22] Mahmoud, F., Christopher, B., Maher, A., Jürg, H., Alexander, S., Adrian, S., & Gaby, H. (2017). Prediction of calving time in dairy cattle. Animal Reproduction Science, 187, 37-46. doi:https://doi.org/10.1016/j.anireprosci.2017.10.003

[23] Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

[24] Nikkilä, R., Seilonen, I., & Koskinen, K. (2010). Software architecture for farm management information systems in precision agriculture. Computers and Electronics in Agriculture, 70(2), 328-336. doi:https://doi.org/10.1016/j.compag.2009.08.013

[25] Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2017). Multi-level automation of farm management information systems. Computers and Electronics in Agriculture, 142, 504-514. doi:https://doi.org/10.1016/j.compag.2017.11.022

[26] Provost, F., & Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking: " O'Reilly Media, Inc.".

[27] Rahman, A., Smith, D. V., Little, B., Ingham, A. B., Greenwood, P. L., & Bishop-Hurley, G. J. (2018). Cattle behaviour classification from collar, halter, and ear tag sensors. Information Processing in Agriculture, 5(1), 124-133. doi:https://doi.org/10.1016/j.inpa.2017.10.001

[28] Rutten, C. J., Steeneveld, W., Vernooij, J. C. M., Huijps, K., Nielen, M., & Hogeveen, H. (2016). A prognostic model to predict the success of artificial insemination in dairy cows based on readily available data. J Dairy Sci, 99(8), 6764-6779. doi:10.3168/jds.2016-10935

[29] Shahinfar, S., Page, D., Guenther, J., Cabrera, V., Fricke, P., & Weigel, K. (2014). Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. J Dairy Sci, 97(2), 731-742. doi:10.3168/jds.2013-6693

[30] Shahriar, M. S., Smith, D., Rahman, A., Freeman, M., Hills, J., Rawnsley, R., . . . Bishop-Hurley, G. (2016). Detecting heat events in dairy cows using accelerometers and unsupervised learning. Computers and Electronics in Agriculture, 128, 20-26. doi:https://doi.org/10.1016/j.compag.2016.08.009

[31] Šmite, D., Wohlin, C., Gorschek, T., & Feldt, R. (2010). Empirical evidence in global software engineering: a systematic review. Empirical software engineering, 15(1), 91-118.

[32] Sørensen, C. G., Pesonen, L., Bochtis, D. D., Vougioukas, S. G., & Suomi, P. (2011). Functional requirements for a future farm management information system. Computers and Electronics in Agriculture, 76(2), 266-276. doi:https://doi.org/10.1016/j.compag.2011.02.005

[33] Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1): MIT press Cambridge.

[34] Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. Animal Behaviour, 124, 203-220.

[35] Vanrell, S. R., Chelotti, J. O., Galli, J., Rufiner, H. L., & Milone, D. H. (2014). 3d acceleration for heat detection in dairy cows. Paper presented at the XLIII Jornadas Argentinas de Informática e Investigación Operativa (43JAIIO)-VI Congreso Argentino de AgroInformática (CAI)(Buenos Aires, 2014).

[36] Wohlin, C., Runeson, P., Host, M., Ohlsson, M., Regnell, B., & Wesslen, A. (2000). Experimentation in software engineering: an introduction. 2000: Kluwer Academic Publishers.

[37] Yazdanbakhsh, O., Zhou, Y., & Dick, S. (2017). An intelligent system for livestock disease surveillance. Information Sciences, 378, 26-47. doi:https://doi.org/10.1016/j.ins.2016.10.026

[38] Zaborski, D., Proskura, W. S., Grzesiak, W., Szatkowska, I., & Jędrzejczak-Silicka, M. (2017). Use of random forest for dystocia detection in dairy cattle. Applied Agricultural and Forestry Research, 147.

[39] Tummers, J., Kassahun, A., & Tekinerdogan, B. (2019). Obstacles and features of Farm Management Information Systems: A systematic literature review. Computers and Electronics in Agriculture, 157, 189-204.

Appendix- Interview Questions

- What do you measure in your farm?
- What kind of features do you combine to make decisions?
- Do you try to make more decisions by yourself based on the measured data in your farm?
- What kind of data is missing in your farm?
- How much money would you invest per cow for a new system with several features?

Rank the following tasks:

- List from the most important to less: heat detection, lameness, calving prediction, sickness prediction

BIOGRAPHIES

HENK JAN HOVING is a master student at the Wageningen University in the Netherlands. He finished his Bachelor degree in Wageningen University and completed his thesis on the application of machine learning in life sciences domain. His research interests are big data analytics, data mining, farm management information systems, data science, and life sciences.

AYALEW KASSAHUN has a PhD in knowledge systems from Wageningen University in the Netherlands where he currently works as a researcher and lecturer. Before joining the university, he worked for five years as a software engineer, first for Baan company (now Infor), a large ERP developer, and later for InfoRay, a developer of a business intelligence software for the banking and insurance sector. His research publications and interests are diverse, including software architecture, management information systems, business process modelling, transparency and traceability systems, and applying data science in the agriculture and food supply chains. He has been involved in a number of EU sponsored research projects and supervises a large number of students on diverse thesis research topics.

CAGATAY CATAL is a member of Information Technology Group in Wageningen University in the Netherlands. He received his B.S. and M.S. degrees in computer engineering from the Istanbul Technical University, Istanbul, and the Ph.D. degree in computer engineering from Yildiz Technical University. He worked 8 years at the Scientific and Technological Research Council of Turkey as Senior Researcher. Later, he worked 6 years in Istanbul Kultur University, Department of Computer Engineering as Associate Professor and Head of the Department. On January 2018, he joined the Information Technology Group of Wageningen University in the Netherlands. His research interests are data science, machine learning, software quality and testing, and experimental software engineering.