Machine learning algorithms can predict tail biting outbreaks in pigs using feeding behaviour records

Catherine Ollagnier[1*], Claudia Kasper[2], Anna Wallenbeck[3], Linda Keeling[3], Giuseppe Bee[1], Siavash A. Bigdeli[4]

[1] Swine Research Unit, Agroscope, Posieux, Fribourg, Switzerland

[2] Animal GenoPhenomics, Agroscope, Posieux, Fribourg, Switzerland

[3] Department of Animal Environment and Health, Swedish University of Agricultural Sciences, Uppsala, Sweden

[4] Edge Vision and AI, CSEM, Neuchâtel, Switzerland

*E-mail: catherine.ollagnier@agroscope.admin.ch

Abstract

Tail biting is a detrimental behaviour that impacts the welfare and health of pigs. Early detection of tail biting precursor signs allows for preventive measures to be taken, thus avoiding the occurrence of the tail biting event. This study aimed to build a machine-learning algorithm for real time detection of upcoming tail biting outbreaks, using feeding behaviour data recorded by an electronic feeder. Prediction capacities of seven machine learning algorithms (e.g., random forest, neural networks) were evaluated from daily feeding data collected from 65 pens originating from 2 herds of grower-finisher pigs (25-100kg), in which 27 tail biting events occurred. Data were divided into training and testing data, either by randomly splitting data into 75% (training set) and 25% (testing set), or by randomly selecting pens to constitute the testing set.

The random forest algorithm was able to predict 70% of the upcoming events with an accuracy of 94%, when predicting events in pens for which it had previous data. The detection of events for unknown pens was less sensitive, and the neural network model was able to detect 14% of the upcoming events with an accuracy of 63%. A machine-learning algorithm based on ongoing data

collection should be considered for implementation into automatic feeder systems for real time prediction of tail biting events.

Keywords

Tail biting, pig, behavioural change, prediction, monitoring behaviour, animal welfare

Background

Tail biting (TB) is an abnormal behaviour in pigs that is thought to have a multi-factorial origin. It can be triggered by factors like the lack of enrichment material, unfavourable environmental conditions, an unbalanced diet or poor health status. In addition to the welfare and ethical concerns associated with this cannibalistic behaviour, TB events cause pain, trigger infections, impair growth and devalue the carcasses.[1-5]

Routine tail docking is prohibited in Switzerland (455.1 OPAn, 2008) and in the EU (Council Directive 2008/120/EC, Annex 1), and farmers are requested to set up measures to prevent TB outbreaks. One strategy is to pinpoint the farm-specific risk factors for TB and to find solutions to reduce them.[4, 6] Another strategy is to monitor animals' behaviours to detect early signs of forthcoming outbreaks.[7-9] Early identification of TB indicators are important for efficient intervention. The behavioural monitoring can be done at the pen and at the individual animal level. Identification at the individual animal level can support preventive measures like for instance, by removing the biter or the bitten pigs. Observations at the pen level are more efficient to detect the TB event.[1]

To date, only a few behavioural indicators were studied at the pen level. These early indicators, like changes in activity level, tail posture and exploratory and feeding behaviour are present up to 63 days before the outbreak starts. Statham et al. reported that pigs spend less time lying and more time standing and sitting within 4 days before an outbreak[10], but this difference in postures was not observed by Wedin et al.[7] Lower tail position seems also to be indicative of an outbreak, and several authors have reported increased incidence of tucked or hanging tails in pens before and during TB outbreaks.[7-10,12] Lately, an automated system was developed that uses 3D cameras to

detect predictive indicators of TB.[8] With this new tool, the authors were able to automatically detect low tail posture, which was positively associated with more tail damage. However, the tail posture may not be very specific to a TB event, as a low tail posture has also been connected to negative emotional status in pigs [7], which can be caused by other factors like sickness. Exploratory behaviours could also be used as early indicators of TB events. For Zonderland et al., the amount of explorative behaviours decreased 6 days before an event[11], whereas for Statham et al., the environmental manipulations increase one day before the TB events.[10] However, Ursinus et al. couldn't find any change in the frequency of explorative behaviours prior to TB events.[14]

Electronic feeding systems with single-spaced feeders enable automatic monitoring of the feeding behaviours of each pig individually. Recording the identity of the pig, feeder entry and exit times and the amount of food consumed allow the calculation of the frequency of feeder visits per day, feeding rates, mean feeder occupation time, mean food intake per feeder visit, total food intake and total feeder duration per day for each pig. In 1994, Young and Lawrence found that pigs housed in groups and fed from automatic feeders showed a temporal patterning of feeding behaviour.[15] They also suggested that the feeding behavior might also be altered by social conditions. It has been later described that changes in feeding behaviours with automatic feeders were associated with negative events like aggressive behaviour or disrupted social dynamics.[8] If a TB event can be predicted from behaviour, as postulated by Statham et al. [9], then data from electronic feeders could be used to monitor in real time the feeding behaviour of the pigs. Indeed, the feeding behaviour of pigs assessed by electronic feeders appears to change before a TB event. Wallenbeck and Keeling, Munsterhjelm et al. and Tessier et al. describe changes in feeding behaviours before TB events.[17-19] These findings suggest that feeding behaviours recorded by electronic feeders could be a valid tool to detect early signs of TB event. Therefore, we aimed to use the feeding behaviour data combined with machine learning (ML) algorithms for real time prediction of TB outbreak. If successful, this will be the first demonstration that a model can predict TB in real time. Such a monitoring tool would open up new opportunities for the farmer to take targeted action in specific

pens to prevent the TB event. Being able to prevent TB would serve the welfare of the animals and provide economic benefits to the farmer. Since the tool requires only data that are already available from pig farms equipped with automatic feeders, it could be easily implemented in commercial practice as an additional management tool.

Material and methods

Data collection

This study analyses the feeding behaviours of two herds of grower-finisher pigs weighing between 25 and 100 kilograms. One data set originates from a testing boar station in Sweden and contains data collected from October 2004 to July 2007. The data set comes from a previous retrospective study that was published by Wallenbeck and Keeling.[10] The second data set contains data from the experimental pig farm of Agroscope and comprises recordings from November 2018 to April 2020. As tail docking is prohibited in Sweden and in Switzerland, the data are from pigs with intact tails.

The Swedish data set includes data from 42 pens (21 TB and 21 CTL) of boars (purebred Yorkshire, Landrace or Hampshire), recorded 70 days before and after the TB date. Boars were housed in groups of 7 to 14 animals per pen. Each pen measured 15.7 m^2 and had a slatted floor and plain resting area. All pigs had *ad libitum* access to the pelleted feed, which was optimised according to the Swedish nutrition norms for fattening pigs.[11] Water was provided *ad libitum* and straw was offered daily.

The Swiss data set consisted of 23 pens (6 tail biting and 17 control) of females and castrated male pigs (Swiss Large White), recorded 100 days before and after the TB date. Twenty pens (18 m^2) contained 11 to 15 pigs and were equipped with two automatic feeders; three pens (78 m^2) were equipped with eight automatic feeders for 31 to 55 pigs. All pens had straw in racks and sawdust on the floor. Water was available *ad libitum* through nipple drinkers. The pelleted finisher diet was formulated to have 20% lower dietary crude protein and essential amino acids compared to a standard diet formulated according to the Swiss feeding recommendations for pigs.[12]

For both study sites, data were collected by individual automatic feeders (ACEMO 48, Acemo, France; or MLP, Agrotronic Schauer, Austria) that recorded the number of visits to the feeder and the amount of feed consumed. The feeders were 0.6 m wide and 1.5-2.2 m long. Only one pig could enter the feeder at a time, and other pigs could not dislodge the pig feeding inside the feeder. Each pig had access to only one feeder.

A pen was assigned to the TB category if at least one pig had to be treated for tail damages. The TB date (day 0) was defined as the date at which the first treatment was recorded. For the Swedish data set, each pen in the TB category was paired to a pen in the control (CTL) category. For the Swiss data set, all the 23 farrowing batches reared under the same housing and feeding conditions were considered for analysis.

Definitions of the observations, analysis in time series and missing value imputation

The frequency of daily feeder visits (DFV), the daily feed consumption (DFC) and the standard deviation of the feed consumption at each visit (StdFC) were calculated per day and per pig (Table 1). These parameters were considered as 'observations' to predict TB events at the pen level.

Table 1: Observations used for tail biting predictions
The frequency of daily feed visit, daily feed consumption and standard deviation of the feed consumption were the observations used to predict tail biting (TB) or control (CTL) classes with the ML models. These parameters were derived from the data collected by the automatic feeder.

	definition	units	abbreviated
Frequency of daily feeder visits	Number of visits to the feeder (from 0:00 to 23:59:59 that date)	n	DFV
Daily feed consumption	Total feed consumption (from 0:00 to 23:59:59 that date)	g	DFC
Standard deviation of the feed consumption	Daily standard deviation of the feed consumption at each visit	g	StdFC

The time dependency of the observations was taken into account by analysing the data by groups of consecutive data points, called the 'analysis window'. The prediction model considered the analysis window to achieve a prediction at the pen level. The analysis window was first defined to contain

observations from 14 consecutive days (Fig. 1). The A_date was defined as the first day of the analysis window. The analysis window slides along the timeline and the A_date defines the class of the analysis window, i.e., “TB” or “CTL”. Analysis windows from control pens were always classified as “CTL”, whereas analysis windows from TB pens were considered as “TB” class only between day -35 and day 10 (TB window). The analysis window of a TB pen was classified as “TB” if the A_date was inside the TB window and “CTL”, if outside the TB window. Missing values were computed using median imputation (by data set) and a principal component analysis was performed before ML analysis.

In each pen, the analysis window contained observations from 10 pigs for 14 days, to standardize the size of the analysis window. Observations from 10 pigs was considered to give enough information on the pen, without creating too many missing data points, for the few occurrences that contained fewer than 10 pigs.

At the end of the data framing, each analysis window contained 420 observations {3 variables (DFV, DFC, StdFC) × 10 pigs × 14 days}, and one outcome (the class of the window: “TB” or “CTL”). The Swiss and the Swedish data set were merged into a third data set, called Swedish+Swiss, to incorporate more diverse observations and further increase the model’s generalizability for unseen data (pen/country). In total, the combined data set (Swedish and Swiss data) contained 6605 analysis windows, with 5479 and 1126 “CTL” and “TB” windows, respectively. The characteristics of the TB windows compared to CTL are presented in Table 2.

Table 2: Characteristics of the data sets. Mean and standard deviation (SD) of the observation are calculated for each data set depending on the class (TB, tail biting, or CTL, control) of the analysis window. The total data set (Swedish+Swiss data set) contained 6605 analysis windows, with 5479 and 1126 CTL and TB windows, respectively. Pigs characteristics (age, and bodyweight) are not balanced between analysis windows, thus the differences between CTL and TB windows should not be considered as traits depicting the change in feeding behavior between a control and a TB pen.

		Swedish		Swiss	
Observations		**CTL**	**TB**	**CTL**	**TB**
DFC[1]	**Mean**	2337.1*	2005.7	2282.3*	2438.8
	SD	757.4	708.6	600.2	678.2
DFV[2]	**Mean**	24.9*	24.0	12.3*	12.0
	SD	19.4	18.0	6.0	7.2
StdFC[3]	**Mean**	128.6*	113.8	166.8*	195.6
	SD	85.9	71.2	96.3	129.7

148 *****significant difference between CTL and TB analysis window classes ($p<0.0001$)
149 DFC[1]: frequency of daily feeder visits
150 DFV[2]: daily feed consumption
151 StdFC[3]: standard deviation of the feed consumption at each visit
152

153 ## *Models*

154 All data were analysed with R3.6.3, using the caret package to build ML models.[13]

155 Commonly used classification ML methods were first tested on all three data sets (i.e., Swedish, Swiss
156 and Swedish+Swiss). Table 3 presents a list of ML methods considered and implemented with R
157 packages in this study.

158 **Table 3: The seven machine learning methods used to predict tail biting events from feeding**
159 **behavior data**
160 Common ML approaches were tested in this classification problem with a binary outcome

Machine Learning	**R Package**	**Function**
Generalized Linear Model with Stepwise Feature Selection	glmnet	glmnet
Random forest	ranger	rf
Support Vector Machines with Radial Basis Function Kernel	kernlab	svmRadial
Bayesian Generalized Linear Model	arm	bayesglm
Neural network	nnet	nn
K-nearest neighbor	caret	kNN
Partial Least Squares Discriminant Analysis	pls	pls

161

162 For each data set, predictive models were first trained on a subset of data (training set), and the
163 models' performances on this training set were then compared. The predictive performances of the
164 models were further compared using the unseen data (test set). The test set contained either new

unseen analysis windows (cross validation, or CV, approach) or new pens (leave one out cross validation, or LOOP, approach).[14] The LOOP approach gives estimate metrics that are valuable when a new pen is presented to the model for prediction, as there is no pen overlapping between the training and testing data set. This represents the situation where the farmer tries to predict a TB event in a hitherto unknown pen. However, the CV resampling approach predicts TB events based on data previously recorded in the pen. This approach is correct when the model can be continuously updated with previous records of the pen, so that the prediction model already knows the feeding behaviour of the pen and tries to classify the analysis window of the testing set based on previous knowledge on this pen.

Model evaluation: performances metrics

This is a classification problem with binary outcomes (TB or CTL), and performances of the models should be assessed on parameters calculated with a confusion matrix.[15] Performance metrics definitions and confusion matrix are presented in Table 4. The sensitivity (rate of predicted TB class given the actual TB class) assesses the capacity of the model to detect an upcoming TB event. The positive predictive value (PPV) evaluates the capacity of the model to correctly predict a TB class. All models were optimized to maximize the sensitivity, as this study aimed to detect early warnings of TB events. The specificity (rate of predicted CTL given the actual CTL class) assesses the predictive performances of CTL in CTL class. The kappa statistic assesses how the model outperforms a random model that simply always predicts “CTL”. According to Landis and Koch, a kappa of 0-0.20 is slight, 0.21-0.40 is fair, 0.41-0.60 is moderate, 0.61-0.80 is substantial, and 0.81-1 is almost perfect.[25] The *p*-value assesses the statistical significance of the difference in accuracy between the random model and the tested model.

Table 4: The confusion matrix and performances metrics used to assess the performances of the models

A confusion matrix was applied to evaluate the prediction performances of the ML models for this classification problem with a binary outcome (tail biting, "TB" or control "CTL" class). The definition of the performances metrics are presented.

Confusion matrix

		Actual	
		TB	CTL
Predicted	TB	TP **True positive** predicted TB class that are actually TB class	FP **False positive** predicted TB class that are actually CTL class
	CTL	FN **False negative** TB class not detected by the model (predicted as CTL class)	TN **True negative** predicted CTL class that are actually CTL class

Performances metrics

$$Sensitivity = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$Positive\ Predicted\ Value\ (PPV) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

$$Specificity = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$$

$$Accuracy = \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TN} + \mathrm{FP} + \mathrm{TP} + \mathrm{FN}}$$

P value: statistical significance of the difference with a random model always predicting CTL class.

$\kappa = \frac{p_o - p_e}{1 - p_e}$ where ρ_o is the observed accuracy, and ρ_e is the expected accuracy of a random model always predicting CTL class.

Results

Model performances

Table 5 presents the model performances on the three training data sets. All models performed significantly better than the random prediction model (that simply always predict "CTL"), with kappas ranging from 0.30 to 1.00. Even if the criteria for optimization was the sensitivity, this performance criterion was always lower or equal to the specificity, which is most likely due to the imbalance between the numerous CTL and the rare TB.

Table 5: Models performances on <u>training data sets</u> for the Swiss, Swedish and Swedish+Swiss data sets. Seven machine learning models were trained on the training data set using a cross validation resampling approach. Performance metrics derived from the confusion matrix (table 4) are used to assess the models' performances. Models in bold are considered as the best predictive models.

				Swedish		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p-value
glmnet[1]	0.88	0.63	0.94	0.76	0.61	<0.0001
rf[2]	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**<0.0001**
svmRadial[3]	0.94	0.80	0.98	0.90	0.81	<0.0001
bayesglm[4]	0.88	0.63	0.94	0.76	0.61	<0.0001
nn[5]	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**<0.0001**
kNN[6]	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**	**<0.0001**
pls[7]	0.87	0.52	0.96	0.79	0.56	<0.0001
				Swiss		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p-value
glmnet	0.92	0.40	0.99	0.79	0.50	<0.0001
rf	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**<0.0001**
svmRadial	0.99	0.88	1.00	1.00	0.93	<0.0001
bayesglm	0.93	0.41	0.99	0.78	0.50	<0.0001
nn	**0.99**	**0.98**	**0.99**	**0.96**	**0.97**	**<0.0001**
kNN	0.95	0.54	1.00	0.98	0.67	<0.0001
pls	0.91	0.19	1.00	0.96	0.30	<0.0001
				Swedish + Swiss		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p-value
glmnet	0.86	0.30	0.97	0.67	0.35	<0.0001
rf	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**<0.0001**
svmRadial	1.00	0.99	1.00	1.00	1.00	<0.0001
bayesglm	0.86	0.30	0.97	0.67	0.35	<0.0001
nn	0.97	0.87	0.99	0.97	0.90	<0.0001
kNN	0.97	0.82	1.00	0.98	0.88	<0.0001
pls	0.85	0.13	0.99	0.75	0.18	<0.0001

[1]Generalized Linear Model with Stepwise Feature Selection; [2]Random forest; [3]Support Vector Machines with Radial Basis Function Kernel; [4]Bayesian Generalized Linear Model; [5]Neural network; [6]K-nearest neighbor; [7]Partial Least Squares Discriminant Analysis

207 *Model prediction performances*

208 Tables 6 and 7 summarize the performances of the models to predict unseen analysis windows (CV)
209 or unseen pens (LOOP), respectively. For the Swedish+Swiss data set, the performances of the model
210 on CV and LOOP were assessed on the combined testing set (Swedish+Swiss) and on subsets of the
211 Swedish or Swiss data set separately. The random forest (RF) model showed the best predictive
212 performances on both the unseen analysis windows and the unseen pens, with an average accuracy
213 of 84% and a sensitivity of 38% on all data sets. Predictive performances were higher on unseen
214 analysis windows than on unseen pens, and always lower than the model performances on the
215 training set. In the Swedish data set, the predictive performances of the RF model on unseen pens
216 were poor (kappa<0). The models trained on the Swedish data were unable to predict Swiss data
217 (kappa<0.06), and vice versa (data not shown). Models trained on the Swedish+Swiss data set
218 showed a poorer performance (low kappa) in predicting new analysis windows of a Swedish or Swiss
219 subset data set than the same model trained on the Swedish or Swiss data sets individually. Lower
220 kappas were also obtained for prediction on new pens, except for glmnet, bayesglm and pls models,
221 that had improved predictive performances (kappa) on the Swiss subset.

Table 6: Performances of models to predict <u>unseen analysis windows</u> (Cross Validation (CV) approach) of the Swedish, Swiss and Swedish+Swiss testing data sets. Seven machine learning models were trained on the training data set and their prediction performances to predict unseen analysis windows (testing set) were evaluated. Performance metrics derived from the confusion matrix (table 4) are used to assess the models' performances. Models in bold are considered as the best predictive models.

				Swedish		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
glmnet[1]	0.80	0.45	0.90	0.57	0.38	0.04
rf[2]	**0.95**	**0.80**	**0.99**	**0.98**	**0.85**	**<0.0001**
svmRadial[3]	0.86	0.54	0.95	0.75	0.54	<0.0001
bayesglm[4]	0.81	0.45	0.90	0.57	0.39	0.04
nn[5]	0.87	0.70	0.91	0.70	0.62	<0.0001
kNN[6]	**0.97**	**0.94**	**0.98**	**0.94**	**0.92**	**<0.0001**
pls[7]	0.81	0.36	0.93	0.60	0.35	-
				Swiss		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
glmnet	0.85	0.22	0.94	0.30	0.17	-
rf	**0.94**	**0.50**	**0.99**	**0.95**	**0.63**	**<0.0001**
svmRadial	0.91	0.28	1.00	0.92	0.39	0.006
bayesglm	0.85	0.22	0.93	0.30	0.17	-
nn	**0.89**	**0.47**	**0.95**	**0.54**	**0.44**	**-**
kNN	**0.93**	**0.41**	**1.00**	**1.00**	**0.55**	**<0.0001**
pls	0.89	0.11	0.99	0.56	0.15	-
				Swedish + Swiss		
	Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
Glmnet[a]	0.82	0.15	0.96	0.43	0.15	-
Swedish[b]	0.77	0.21	0.93	0.46	0.18	-
Swiss[c]	0.88	0.02	0.99	0.20	0.02	-
rf[a]	**0.94**	**0.70**	**0.99**	**0.96**	**0.78**	**<0.0001**
Swedish[b]	**0.95**	**0.80**	**0.99**	**0.95**	**0.83**	**<0.0001**
Swiss[c]	**0.94**	**0.48**	**1.00**	**1.00**	**0.62**	**<0.0001**
svmRadial[a]	0.88	0.38	0.98	0.80	0.46	<0.0001
Swedish[b]	0.86	0.48	0.97	0.80	0.52	<0.0001
Swiss[c]	0.90	0.16	1.00	0.76	0.23	-
bayesglm[a]	0.82	0.16	0.96	0.43	0.16	-
Swedish[b]	0.77	0.21	0.93	0.46	0.19	-
Swiss[c]	0.82	0.16	0.96	0.43	0.15	-
nn[a]	**0.88**	**0.55**	**0.94**	**0.67**	**0.53**	**<0.0001**
Swedish[b]	**0.87**	**0.63**	**0.94**	**0.75**	**0.61**	**<0.0001**
Swiss[c]	**0.88**	**0.37**	**0.95**	**0.47**	**0.35**	**-**
kNN[a]	**0.96**	**0.78**	**1.00**	**0.98**	**0.84**	**<0.0001**
Swedish[b]	**0.96**	**0.87**	**0.98**	**0.94**	**0.88**	**<0.0001**
Swiss[c]	**0.92**	**0.32**	**1.00**	**1.00**	**0.46**	**0.0006**
pls[a]	0.83	0.06	0.99	0.62	0.09	-
Swedish[b]	0.79	0.09	0.98	0.62	0.11	-
Swiss[c]	0.89	0.00	1.00	0.00	0.00	-

[1]Generalized Linear Model with Stepwise Feature Selection; [2]Random forest; [3]Support Vector Machines with Radial Basis Function Kernel; [4]Bayesian Generalized Linear Model; [5]Neural network; [6]K-nearest neighbor; [7]Partial Least Squares Discriminant Analysis

[a]Prediction performances of the models on testing data set containing Swedish and Swiss data.

[b]Prediction performances of the models on the Swedish subset of the testing data set.

[c]Prediction performances of the models on the Swiss subset of the testing data set.

Table 7: Performances of models to predict <u>unseen pens</u> (Leave one out cross validation (LOOP) approach) of the Swedish, Swiss and Swedish+Swiss data sets.

Seven machine learning models were trained on the training data set and their prediction performances to predict unseen pens (testing set) were evaluated. Performance metrics derived from the confusion matrix (table 4) are used to assess the models' performances. Bold models are considered as the most promising models, according to the authors. Models in bold are considered as the best predictive models.

				Swedish			
		Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
glmnet[1]		0.56	0.14	0.71	0.14	<0.00	-
rf[2]		**0.63**	**0.11**	**0.81**	**0.17**	**<0.00**	**-**
svmRadial[3]		0.54	0.20	0.66	0.17	<0.00	-
bayesglm[4]		0.56	0.14	0.71	0.14	<0.00	-
nn[5]		0.56	0.05	0.73	0.06	<0.00	-
kNN[6]		0.62	0.00	0.83	0.00	<0.00	-
pls[7]		0.58	0.10	0.75	0.12	<0.00	-
				Swiss			
		Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
glmnet		0.82	0.18	0.96	0.50	0.19	-
rf		**0.84**	**0.18**	**0.99**	**0.75**	**0.24**	**-**
svmRadial		0.81	0.02	0.99	0.25	0.01	-
bayesglm		0.82	0.18	0.95	0.47	0.18	-
nn		0.75	0.04	0.90	0.09	<0.00	-
kNN		**0.84**	**0.12**	**0.99**	**0.86**	**0.18**	**-**
pls		0.82	0.08	0.99	0.57	0.10	-
				Swedish + Swiss			
		Accuracy	Sensitivity	Specificity	PPV	Kappa	p value
Glmnet[a]		0.73	0.12	0.90	0.27	0.03	-
	Swedish[b]	0.62	0.00	0.84	0.00	<0.00	-
	Swiss[c]	0.87	0.37	0.99	0.86	0.45	0.008
rf[a]		0.74	0.00	0.96	0.00	<0.00	-
	Swedish[b]	0.69	0.00	0.93	0.00	<0.00	-
	Swiss[c]	0.82	0.00	1.00	0	0	-
svmRadial[a]		0.71	0.03	0.91	0.08	<0.00	-
	Swedish[b]	0.63	0.03	0.84	0.06	<0.00	-
	Swiss[c]	0.82	0.02	1.00	1.00	0.03	-
bayesglm[a]		0.73	0.13	0.91	0.28	0.04	-
	Swedish[b]	0.63	0.01	0.84	0.02	<0.00	-
	Swiss[c]	0.87	0.37	0.99	0.86	0.46	0.009
nn[a]		0.63	0.14	0.78	0.16	<0.00	-
	Swedish[b]	0.56	0.21	0.68	0.18	<0.00	-
	Swiss[c]	0.74	0.00	0.91	0.00	<0.00	-
kNN[a]		0.74	0.00	0.95	0.00	<0.00	-
	Swedish[b]	0.69	0.00	0.93	0.00	<0.00	-
	Swiss[c]	0.82	0.00	1.00	0.00	0.00	-
pls[a]		0.77	0.08	0.97	0.41	0.07	-
	Swedish[b]	0.70	0.00	0.94	0.00	<0.00	-
	Swiss[c]	0.86	0.24	0.99	0.92	0.34	0.04

[1]Generalized Linear Model with Stepwise Feature Selection; [2]Random forest; [3]Support Vector Machines with Radial Basis Function Kernel; [4]Bayesian Generalized Linear Model; [5]Neural network; [6]K-nearest neighbor; [7]Partial Least Squares Discriminant Analysis

[a]Prediction performances of the models on testing data set containing Swedish and Swiss data.

[b]Prediction performances of the models on the Swedish subset of the testing data set.

[c]Prediction performances of the models on the Swiss subset of the testing data set.

Impact of TB and analysis windows

We have tested the influence of the size of the analysis window and the size of the TB window on predictive performances for the Swedish data set. Accuracy of models with analysis windows of 7 or 21, or with different TB windows (-49 to +10 days, or -10 to +5 days) were compared with the default windows size (analysis window: 14 days; TB window: -35 to +10 days). The RF model was chosen for this analysis, as it was the model with the best predictive performances over the three data sets and for the two (CV and LOOP) approaches.

The predictive performances of the RF model increases when the analysis window contains 21 days (instead of 14 days) or when the TB window was larger (-49 to +10 days) (Table 8). Interestingly, the RF model never predicted TB class in CTL pens, thus ML models were able to discriminate CTL pen behaviour from TB pen behaviour. In addition, the prediction of a TB event for a TB pen was almost exclusively within the TB window (high specificity). The detection of an upcoming TB event was not better near the TB date (day 0) than at the beginning of the TB window (day -35, for default TB window size).

Table 8: Prediction performances of the random forest (rf) model depending on the TB and analysis window size on the Swedish data set.
The prediction performances of the rf model were assessed while varying the size of the TB window and of the analysis window. The prediction were made for unseen analysis windows or for unseen pens. The rf model was chosen as it seems to the authors that it is the most promising model for this ML prediction, based on previous results (Tables 6 and 7).

			Swedish					
	Analysis window size	**TB window size**	**Accuracy**	**Sensitivity**	**Specificity**	**PPV**	**Kappa**	**p value**
Unseen analysis windows (CV)	7 days	[-35;10]	0.94	0.76	0.99	0.96	0.81	<0.0001
	14 days	**[-35;10]**	**0.94**	**0.70**	**0.99**	**0.96**	**0.78**	**<0.0001**
		[-49;10]	**0.97**	**0.88**	**0.99**	**0.97**	**0.90**	**<0.0001**
		[-10;5]	0.95	0.45	1.00	0.94	0.58	<0.0001
	21 days	**[-35;10]**	**0.97**	**0.87**	**1.00**	**0.99**	**0.91**	**<0.0001**
Unseen pens (LOOP)	7 days	[-35;10]	0.64	0.00	0.87	0.00	<0.00	-
	14 days	**[-35;10]**	**0.63**	**0.11**	**0.81**	**0.17**	**<0.00**	**-**
		[-49;10]	**0.61**	**0.19**	**0.82**	**0.35**	**0.02**	**-**
		[-10;5]	0.91	0.00	1.00	0.00	0.00	-
	21 days	[-35;10]	0.66	0.00	0.88	0.00	<0.00	-

Discussion

In this study, TB events could be detected up to 35 days in advance using an ML model that analysed feeding behaviours recorded by electronic feeders. Early indicators of TB events were more easily identified when the model had access to previous records of the pen. Thus, one should consider performing continuous analysis of the data of each pen, even in the absence of TB events. To our knowledge this is the first time that an ML algorithm has been used to predict TB in pigs.

In the following discussion we compare our results using an ML model with statistical models from other articles attempting to characterize feeding behaviours before and after a TB event.

Munsterhjelm et al. previously investigated feeding behaviours in relation to TB, 70 days before to 28 days after a first wound was detected in the pen. Feeding behaviours in TB pens were compared to matched CTL pens. Pigs in TB pens tend to visit less frequently the feeder than pigs in CTL pens. The frequency of daily feeder visits (DFV) reduced with the approach of the TB date. The feeding behaviour in TB pens significantly changed 14 days before the detection of the TB event. Time spent at the feeder and daily feed consumption (DFC) were also reduced in TB pens compared to CTL pens. This tendency persisted after the TB date, and pigs continued to spend less time in the feeder and visit less frequently, even if the difference from the control group decreased with time. Pigs also tended to eat faster (more intake per second). They concluded that the rapid change in feeding behaviour suggests that TB behaviour escalates 14 days before the TB date.[18]

The change in feeding behaviour in TB pen was also observed by Tessier et al. during a TB outbreak in a pen. They studied the evolution of DFV, the DFC and the feeding time – 7 days before (pre-injury phase), 7 days during (acute phase) and 7 days after (recovery phase) the TB outbreak. The DFV decreased before the TB date, reached a minimum during the outbreak, and increased during the recovery phase. Consumption time (for a given amount of feed eaten) decreased while the TB outbreak was evolving (during the pre-injury and acute phases) and remained low during the recovery phase. This study corroborates that a change in feeding behaviour in a pen can indicate a

293 future TB event. This also confirms the gradual change of the feeding behaviour over time, reaching
294 its maximum at the TB date.[19]
295 The statistical analysis by Wallenbeck and Keeling on the Swedish data set also demonstrated a
296 significant decrease in DFV in TB pens compared to matched CTL pens, 42 to 63 days before the TB.
297 In addition, the DFC was always reduced in TB pens compared to CTL pens.[17] This difference in
298 feeding behaviour between CTL and TB pens must have been also detected in the present analysis by
299 the ML models that never predict TB class in CTL pens.
300 In contrast, Zonderland et al. did not find any change in feeding behaviour, especially in the
301 difference in time spent feeding between CTL and TB pens.[26] But they studied pigs fed *ad libitum*
302 from dry feeders with two feeding spaces, where pigs could feed at the same time from the same
303 feeder. The feeding behaviour in pens with feeder troughs is different from those with automatic
304 feeders.[16] The upcoming TB event may be responsible for this difference.
305 The comparison between TB and CTL windows in the present study (Table 2) doesn't reflect a general
306 tendency of the change in feeding behaviour. Indeed, the characteristics of the pigs in CTL and TB
307 windows are not balanced, so the differences in age and weight between the two categories could
308 explain the differences in DFV, DFC and StdFC and also the inconsistencies observed between the
309 Swiss and Swedish data sets.
310 Wallenbeck and Keeling described a change in DFV 42 to 63 days before a TB event, comparing TB
311 pens to CTL pens.[17] In the present study, the detection was limited to 35 days before the event (TB
312 window: -35 to +10 days). Nevertheless, increasing the size of the TB window to (-49 to +10 days)
313 improves the accuracy and the sensitivity of the RF model. In addition, the two analyses did not use
314 the same reference: the Wallenbeck analysis compared TB pens to matched CTL pens, trying to
315 characterize a feeding behaviour that would be typical of a future TB event.[17] For the present
316 analysis, each pen was also its own control, as a TB pen could be a TB class when inside the TB
317 window but a CTL when outside. The model could not only detect feeding behaviour that is typical of

318 a future TB event, but also detect changes in feeding behaviour over time that are indicative of TB
319 events.
320 Indeed, as demonstrated by Munsterhjelm et al. and Tessier et al., a drastic change in feeding
321 behaviour in each pen is also indicative of an upcoming TB event.[18-19] In addition, the present
322 analysis considers a combination of three feeding behaviour parameters (DFV, DFC, StdFC), thus
323 increasing the possibility to detect TB events. In a comprehensive book chapter on TB, Valros
324 describes the gradual change in feeding behaviour until the TB date. DFV can start decreasing
325 months to weeks before, whereas DFC seems to be affected only six days before the TB date.[5]
326 Thus, the combination of these two indicators should improve the specificity of detection.
327 With modified feeding behaviours being present one to several months before the event, the TB
328 events observed in the present study seemed to result from a long process and therefore to belong
329 to the two-stage type, as described by Taylor et al.[17] This type is described as starting as a harmless
330 behaviour that escalates to actual TB outbreak. A suggested cause was chronic but moderate stress.
331 In general, TB appears to be connected to an increased stress level in pigs.[18, 19]
332 In the different studies, the first changes in feeding behaviour linked to TB occurred 42 to 63
333 days[10], or 70 days[20] before the first TB event. Depending on the monitoring frequency, the TB
334 event may be detected sooner or later. Therefore the TB date may be determined when the TB
335 process has just begun or when it has become an outbreak. This would also explain the discrepancies
336 between how long before the TB event the feeding behaviour is described to have changed.
337 In addition, it seems that the feeding behaviour is not only specific to a pen, but each site has its own
338 characteristics (e.g., breeds, climates and feeding compositions). The model trained on one data set
339 (Swedish or Swiss) was unable to detect events in the other data set, and the combination of the
340 data sets (Swedish + Swiss) did not greatly improve model performance. Thus, the increase in size
341 resulting from the combination of the Swedish and Swiss data did not provide much advantage. One
342 explanation could be linked to site-specific risks factors. As acknowledged by the European Food
343 Safety Authority 2014, one of the difficulties in preventing TB resides in the fact that every farm is

different and has its own risk factors. Prevention strategies need therefore to be designed at a farm-specific level.[3] Depending on the chronic risk factors of the farm (e.g., breed, sex, feeding, access to manipulable material, space allowance and group size), the feeding behaviour associated with a TB event may vary. This observation was already acknowledge by Taylor et al. and Valros, when they defined the four types of TB (two-stage, sudden-forceful, obsessive and epidemic), associated with four putative causations. The two-stage type was brought about by chronic and moderate stress. The sudden-forceful type was suggested to be caused by competition for resources. The epidemic type stemmed from a sudden change in the pig's life (e.g., feeding disturbance or temperature).[5,28]

In conclusion, an ML model can be implemented in an automatic feeder system on a farm to detect early signs of TB behaviours at least 35 days before the actual TB event. Thanks to these early warnings, farmers could implement measures to prevent the occurrence of the TB event – for example, by adding more straw as occupational material. Farmers could also start a TB risk analysis to identify the reasons why pigs are disturbed. Continuous implementation of the model on farms would also lead to improved prevention of TB events, serving the welfare of the pigs and bringing an economic benefit to the farmers. Finally, this ML approach could also be a useful tool by allowing the systematic study of the effectiveness of different intervention strategies under controlled conditions.

References

1. Larsen MLV, Andersen HML, Pedersen LJ. Can tail damage outbreaks in the pig be predicted by behavioural change? Vet J. 2016;209:50-6. doi: 10.1016/j.tvjl.2015.12.001.

2. Brunberg EI, Rodenburg TB, Rydhmer L, Kjaer JB, Jensen P, Keeling LJ. Omnivores going astray: A review and new synthesis of abnormal behavior in pigs and laying hens. Front Vet Sci. 2016;3(57). doi: 10.3389/fvets.2016.00057.

3. European Food Safety Authority. Scientific opinion concerning a multifactorial approach on the use of animal and non-animal-based measures to assess the welfare of pigs. EFSA J. 2014:12(5):3702. doi: 10.2903/j.efsa.2014.3702.

4. European Food Safety Authority. Pig welfare risks associated with tail biting in pigs and possible means to reduce the need for tail docking considering the different housing and husbandry systems - scientific opinion of the panel on animal health and welfare. EFSA J. 2007;611:1-13. doi: 10.2903/j.efsa.2007.611.

5. Valros A. Chapter 5 - Tail biting. In: Špinka M, editor. Advances in Pig Welfare. Woodhead Publishing; 2018. p. 137-66.

6. D'Eath RB, Arnott G, Turner SP, Jensen T, Lahrmann HP, Busch ME, et al. Injurious tail biting in pigs: how can it be controlled in existing systems without tail docking? Animal. 2014;8(9):1479-97. Epub 2014/08/19. doi: 10.1017/s1751731114001359. PubMed PMID: 25130712.

7. Wedin MB, Baxter EM, Jack M, Futro A, D'Eath RB. Early indicators of tail biting outbreaks in pigs. Appl Anim Behav Sci. 2018;208:7-13. doi: 10.1016/j.applanim.2018.08.008.

8. D'Eath RB, Jack M, Futro A, Talbot D, Zhu Q, Barclay D, et al. Automatic early warning of tail biting in pigs: 3D cameras can detect lowered tail posture before an outbreak. PLoS One. 2018;13(4):e0194524. doi: 10.1371/journal.pone.0194524.

9. Lahrmann HP, Hansen CF, D'Eath R, Busch ME, Forkman B. Tail posture predicts tail biting outbreaks at pen level in weaner pigs. Appl Anim Behav Sci. 2018;200:29-35. doi: 10.1016/j.applanim.2017.12.006.

386 10. Statham P, Green L, Bichard M, Mendl M. Predicting tail-biting from behaviour of pigs prior
387 to outbreaks. Appl Anim Behav Sci. 2009;121:157-64. doi: 10.1016/j.applanim.2009.09.011.
388 11. Zonderland JJ, van Riel JW, Bracke MBM, Kemp B, den Hartog LA, Spoolder HAM. Tail posture
389 predicts tail damage among weaned piglets. Appl Anim Behav Sci. 2009;121(3):165-70. doi:
390 10.1016/j.applanim.2009.09.002.
391 12. Wallgren T, Larsen A, Gunnarsson S. Tail posture as an indicator of tail biting in undocked
392 finishing pigs. Animals (Basel). 2019;9(1):18. doi: 10.3390/ani9010018. PubMed PMID: 30626058.
393 13. Reimert I, Bolhuis JE, Kemp B, Rodenburg TB. Indicators of positive and negative emotions
394 and emotional contagion in pigs. Physiol Behav. 2013;109:42-50. Epub 2012/11/20. doi:
395 10.1016/j.physbeh.2012.11.002. PubMed PMID: 23159725.
396 14. Ursinus WW, Van Reenen CG, Kemp B, Bolhuis JE. Tail biting behaviour and tail damage in
397 pigs and the relationship with general behaviour: predicting the inevitable? Appl Anim Behav Sci.
398 2014;156:22-36. doi: 10.1016/j.applanim.2014.04.001.
399 15. Young R, Lawrence A. Feeding behaviour of pigs in groups monitored by a computerized
400 feeding system. Anim Prod. 1994;58:145-52. doi: 10.1017/S0003356100007182.
401 16. Brunberg E, Wallenbeck A, Keeling L. Tail biting in fattening pigs: associations between
402 frequency of tail biting and other abnormal behaviours. Appl Anim Behav Sci. 2011;133. doi:
403 10.1016/j.applanim.2011.04.019.
404 17. Wallenbeck A, Keeling LJ. Using data from electronic feeders on visit frequency and feed
405 consumption to indicate tail biting outbreaks in commercial pig production. J Anim Sci.
406 2013;91(6):2879-84. Epub 2013/03/13. doi: 10.2527/jas.2012-5848. PubMed PMID: 23478818.
407 18. Munsterhjelm C, Nordgreen J, Heinonen M, Janczak A, Valros A, editors. Feeding behaviour
408 and performance in relation to injurious tail biting in boars – a longitudinal study. 7th Eurpoean
409 Symposium of Porcine Health Management (ESPHM), Dublin, Ireland; Jun 2016.
410 19. Tessier F, Maikoff G, Bee G, Ollagnier C, editors. Tail biting in switzerland: a restrospective
411 study. 51eme Journées de la Recherche Procine (JRP), Paris; Feb 2019.

412 20. Andersson C, Lindberg JE. Forages in diets for growing pigs 1. Nutrient apparent digestibilities
413 and partition of nutrient digestion in barley-based diets including lucerne and white-clover meal.
414 Animal Science. 1997;65(3):483-91. Epub 2010/09/02. doi: 10.1017/S1357729800008687.
415 21. Agroscope. Apports alimentaires recommandés pour les porcs (livre jaune) 2005 [cited 2020].
416 22. Kuhn M. Building predictive models in R using the caret package. J Stat Softw. 2008;1(5). doi:
417 10.18637/jss.v028.i05.
418 23. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling
419 methods. Bioinformatics. 2005;21(15):3301-7. Epub 2005/05/21. doi: 10.1093/bioinformatics/bti499.
420 PubMed PMID: 15905277.
421 24. Tharwat A. Classification assessment methods. Applied Computing and Informatics.
422 2020;17(1):168-192. doi: 10.1016/j.aci.2018.08.003.
423 25. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of
424 majority agreement among multiple observers. Biometrics. 1977;33(2):363-74. doi:
425 10.2307/2529786.
426 26. Zonderland J, Schepers F, Bracke MBM, Hartog LA, Kemp B, Spoolder HAM. Characteristics of
427 biter and victim piglets apparent before a tail-biting outbreak. Animal. 2011;5:767-75. doi:
428 10.1017/S1751731110002326.
429 27. Nielsen BL, Lawrence AB, Whittemore CT. Feeding behaviour of growing pigs using single or
430 multi-space feeders. Appl Anim Behav Sci. 1996;47(3):235-46. doi: 10.1016/0168-1591(95)00649-4.
431 28. Taylor NR, Main DC, Mendl M, Edwards SA. Tail-biting: a new perspective. Vet J.
432 2010;186(2):137-47. Epub 2009/10/07. doi: 10.1016/j.tvjl.2009.08.028. PubMed PMID: 19804997.
433 29. Munsterhjelm C, Brunberg E, Heinonen M, Keeling L, Valros A. Stress measures in tail biters
434 and bitten pigs in a matched case-control study. Anim Welf. 2013;22(3):331-338. doi:
435 10.7120/09627286.22.3.331.
436 30. Zupan M, Janczak AM, Framstad T, Zanella AJ. The effect of biting tails and having tails bitten
437 in pigs. Physiol Behav. 2012;106(5):638-44. doi: 10.1016/j.physbeh.2012.04.025.

438

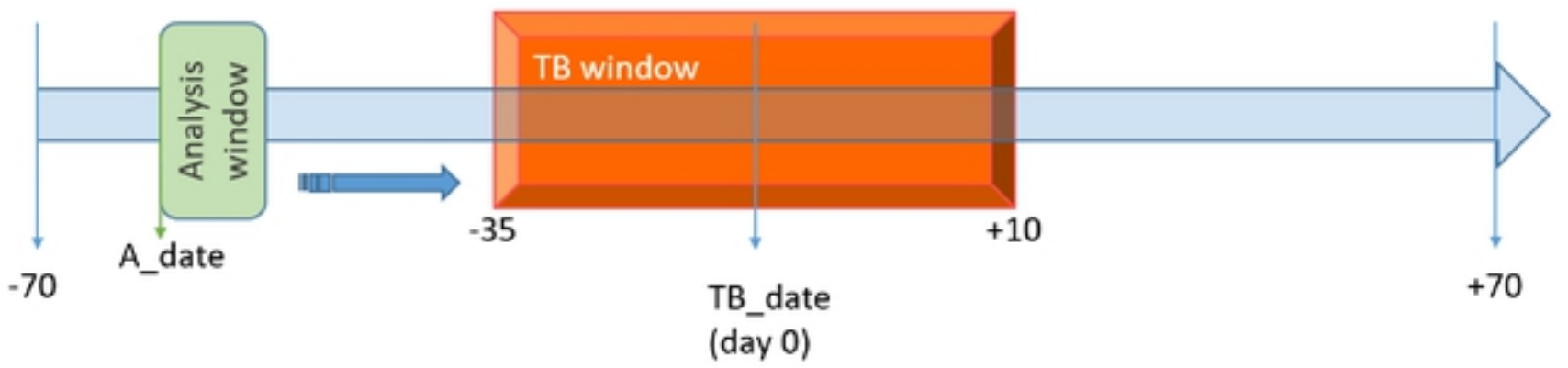

Figure 1: Analysis of the time dependency of the data thanks to the analysis window approach. Analysis window slides along the time line (Blue arrow). The tail biting (TB) pen will be classified as TB class, when the A_date (first day of the analysis window) enters the TB window (orange block) and as "CTL" class when it is outside the TB window. Control pen will always be classified as "CTL".

Figure 1