Received: 16 October 2022, Accepted: 05 November 2022

Review Article

Journal of Advanced Veterinary Research
(2022) Volume 12, Issue 6, 798-802

Machine Learning Based Prediction for Solving Veterinary Data Problems: A Review

Fardos A.M. Hassan[1] , Sherif A. Moawed[2], Iman E. El-Araby[1], Hagar F. Gouda[1*]

[1]Department of Animal Wealth Development, Faculty of Veterinary Medicine, Zagazig University, 44511, Egypt.
[2]Department of Animal Wealth Development, Faculty of Veterinary Medicine, Suez Canal University, 41522, Egypt.

***Correspondence**
Hagar F. Gouda
Department of Animal Wealth Development, Faculty of veterinary medicine, Zagazig University, 44511, Egypt.
E-mail address: hagarfathy@zu.edu.eg

Abstract

What if we could detect disease before it manifested itself? While it may appear to be fantasy, it is currently being used in human and animal health by combining advanced computing power with artificial intelligence (AI). Machine learning (ML), a subfield of AI, is a recent approach for developing predictions in many areas of medical science through classification and regression. Using ML to solve veterinary data problems is still unusual. It can aid in farm decision-making processes. ML outperformed statistical models in disease prediction, which produce a high bias in most cases and reduce model reliability when assumptions are violated. Some challenges of ML, such as data size, algorithm tunability, and feature selection, must be considered in order to develop a good predictive model. This review aimed to discuss the role of ML in solving veterinary problems and spotting the light on overcoming the possible challenges, and to encourage the researchers to increase the application of ML over conventional methods.

KEYWORDS
Machine learning, Veterinary, Prediction, Classification, Imbalance.

INTRODUCTION

Machine learning (ML), a branch of artificial intelligence (AI), focuses on the learning aspect of AI by using subsets of data to create algorithms that best represent the data. While traditional methods' algorithms can be explicitly coded using known features (Hastie *et al*., 2009; James *et al*., 2013).

ML models provide a great ability to analyze highly dimensional, non-linear data, and complex clinical scenarios (Jordan and Mitchell, 2015; LeCun *et al*., 2015). On the other hand, some studies Banerjee *et al*. (2021) and Navarro *et al*. (2021) noted that ML approaches are frequently used in tiny and low-dimensional situations.

ML algorithms are broadly classified into four types: supervised, unsupervised, semi-supervised, and reinforcement learning algorithms (Mohammed *et al*., 2016). Supervised learning entails mapping input data to output data using input-out pairs of data. Classification and regression are the two most common types of supervised ML. While unsupervised algorithms are based on unlabeled datasets and do not require human intervention. Semi-supervised learning is a hybrid of supervised and unsupervised ML (Han *et al*., 2022). Reinforcement learning (environment-driven approach) automatically evaluate behavior in a specific situation or environment in order to improve efficiency relying on software and machines (Kaelbling *et al*., 1996).

New data analysis tools may now offer new options to aid in synthesizing the strong framework of dairy farms, hence improving future management decisions. This shows the opportunities that ML techniques could provide for dairy production. Analyzing huge integrated datasets may enable farmers to provide improved decision support systems, assisting them to improve the wellbeing and efficiency of their animals (Cockburn, 2020). When compared to human medicine, the application of ML models in veterinary medicine is extremely limited (Wan and Bao, 2009; Parkhi *et al*., 2011). With few studies, applying ML to solve veterinary data problems has received little attention (Anholt *et al*., 2014; Santamaria and Zimmerman, 2011).

This review highlighted the application of ML models in the veterinary field in order to encourage researchers to use ML methods in solving veterinary problems rather than relying solely on traditional methods; to provide researchers with basic information for a successful application; and to highlight some challenges in the application.

The most common traditional methods in veterinary data analysis

Some of frequently used traditional statistical methods that applied to solve veterinary problems are, linear regression (Fernandes *et al*., 2020; Tirink and ÖNDER, 2022), cox regression (Holm *et al*., 2009), logistic regression (LR) (Holm *et al*., 2009; Mohamad Hifzan *et al*., 2015), non-linear regression (Mohamad

ISSN: 2090-6277/2090-6269/

Hifzan *et al*., 2015), and analysis of variance (El-Bayomi *et al*., 2019).

In epidemiology, LR is well known and commonly used (Freedman, 2009). Logistic regression is frequently employed in the analysis of binary output data. If the response has many categorical levels, LR grows more complex and incorporates multiple outcome functions (Yan *et al*., 2016). Overestimation of the effect size is a significant issue when using LR in research with small to moderate sample sizes. As a result, large sample sizes are required for LR to offer sufficient numbers in both outcome variable categories. Furthermore, the more independent variables are included, the larger the sample size required. Hosmer and Lemeshow advocate sample sizes of at least 400, with a minimum of ten cases per independent variable (Hosmer *et al*., 2013).

The implementation of the multiple linear regression (MLR) technique is complicated by the fact that its assumptions are not always met, which can result in biased results and failure to provide reliable prediction (Ernst and Albers, 2017).

ML's contribution to a better understanding of animal epidemiological schemes

Traditional statistical methods focus on inferences and are based on setting a hypothesis that might be accepted/rejected depending on how it fits with the measured data. While ML methods are hypothesis-free and do not require a specific data distribution to make a prediction (Valletta *et al*., 2017).

In the area of veterinary science researchers can use machine learning and data science, datasets for different purposes. Machado *et al*. (2015) applied random forests (RF) for prediction of bovine viral diarrhea. Bradley and Rajendran (2021) used artificial neural network (ANN), RF and other ML algorithms to increase taking on rates at animal shelters.

Disease diagnosis, animal welfare, population medication, education, and industry aspects of the veterinary field could all benefit from artificial intelligence and ML (Appleby and Basran, 2022).

A range of fish diseases are being diagnosed using ANN, case-based, rule-based, and fuzzy logic systems (Zeldis and Prescott, 2000). Chen *et al*. (2022) reported that the ANN model outperformed the MLR model in predicting dairy cattle animal waste nutrient levels.

The ANN, random forest regression (RFR) and support vector regression (SVR) can be applied to approximate complex relationship to evaluate the pattern of rumen fermentation Craninx *et al*. (2008), animal diet formulation Saxena and Parasher (2019), and milk production (Fernández *et al*., 2006). Bobbo *et al*. (2021) used ML methods to predict bovine mastitis in dairy cows.

Cushing's syndrome prediction in dogs can be improved by applying ML in disease diagnosis, that using ML as a decision-support tools is recommended (Schofield *et al*., 2021).

Simple and interpretable ML models are produced when applying decision tree in veterinary, clinical, and environmental fields (Kuhnert *et al*., 2000; Saegerman *et al*., 2011). Romero *et al*. (2020) classified cows at risk of tuberculosis using decision tree.

Successful implementation of machine learning

The nature and features of the input data and the performance of the learning algorithms are crucial in determining how effective the ML solutions are. Association rule learning and reinforcement learning techniques are available to effectively build ML data-driven systems for classification, regression, clustering, feature engineering, and dimensionality reduction (Han *et al*., 2011).

When applying ML models to produce successful applications with high model accuracy, several factors must be considered, including data size, feature selection criteria, train/test split ratio, hyperprameter tuning, data class imbalance, training the algorithm, and class imbalance.

Data preprocessing

ML are input-data-dependent to arrive the classification prediction. The quality of data reflected on ML prediction performance (Crone *et al*., 2006). A great role is attributed to preparation of the data and training process (Awaysheh *et al*., 2019).

Train /Test split

The importance of splitting data into train and test sets is to validate the model by comparing algorithm performance in train and test sets (Vabalas *et al*., 2019).

Overfitting and under-fitting are the two most common problems that may arise when validating the model. When algorithms perform well in the training set but fail to generalize to the test set, this is referred to as over-fitting. Under-fitting the model results in inaccurate predictions with higher error in the test set (Goodfellow *et al*., 2016).

Simple-random split

Data is divided randomly into a large train set for model building and training and small separate test set for model validation. Various split ratio previously used. One of the most common split ratio used in veterinary studies are 80:20% used in study of Machado *et al*. (2015) and Romero *et al*. (2022), 70:30% by Punyapornwithaya *et al*. (2022), and 60:40% by Keshavarzi *et al*. (2020). Other studies used 50:50% (Kwon *et al*., 2015).

Cross-validation (CV) split

Cross-validation is another common approach to perform train/test split. Commonly, many rounds or folds of cross-validation are accomplished, and the average of folds' performance is obtained do reduce variability (Kohavi, 1995). Depending on the number of folds, there are several cross-validation approaches. The most common approach is K-fold cross-validation, where K is the commonly K = 10 is used. Leave-one-out cross-validation is another common approach that represents K-fold cross-validation taken to its extreme, with K equal to the number of data instances (Awaysheh *et al*., 2019).

When validation with a separate dataset is not possible due to the small sample size, K-Fold CV is effective because it allows one to use all of the data for both training and validation. If validation were performed on a separate dataset, twice as much data would be required to have the same amount of data for training and validation. When compared to Train/Test, using all of the available data for validation should result in out of sample error estimates that are less influenced by noise and more representative of the population (Vabalas *et al*., 2019).

Variable selection

Variable selection is the inclusion process of the selected variables in a specific model from a large list of variables, removing those that are irrelevant or redundant (Ratner, 2010). Chi-square

widely used for qualitative data variable selection, it is reported in many recent studies for perdition and classification (Punyapornwithaya *et al*., 2022).

It is advised to start the variable selection process with univariate analysis. Significant variables with $p < 0.25$ are kept and used in multivariate analysis. A higher significant level is recommended to include more variables in multivariate analysis because univariate analysis ignores the fact that weakly associated variables have a significant contribution when they are combined (Hosmer *et al*., 2013). Keshavarzi *et al*. (2020) used LR as a selection method of the bulls' genetic information at $p < 0.05$ for inclusion the variables.

Several studies discussed the various types of variable selection and their advantages and drawbacks (Yoo *et al*., 2012; Li *et al*., 2020).

Sample size considerations in machine learning

ML is heavily reliant on sample size. A larger dataset that covers the greater variation in disease features is recommended for use in ML classification models to develop more robust models to be applicable in the real-world situations (Rashid and Calhoun 2020). However, several articles of prior ML studies have discovered an inverse relation between sample size and accuracy (Schnack and Kahn 2016; Varoquaux 2018). Predictions with high accuracy are usually confined to studies with small sample sizes (Arbabshirani *et al*. 2017). Recent large-scale inquiry results have raised a fundamental problem in the field. ML research with larger samples, in particular, did not provide better results, but consistently produced worse outcomes than studies with small samples (Flint *et al*., 2021).

When the dataset is too small to be used sufficiently in validation set, cross-validation techniques can be used to pick model hyper-parameters. After reserving a portion of the data for testing, k-fold validation entails dividing the training set into k subsets, k1 of which are utilized for training and the final one to evaluate performance. This process is done k times, and the performance scores from each subset averaged for each set of hyper-parameters to test. Random decision forests are appealing for smaller datasets (Vabalas *et al*., 2019).

Previous studies applied ML in small datasets and obtained respectable results. For example, Johnston *et al*. (2015) achieved a high classification accuracy of 85% in their study using the support vector machine learning method (SVM) with only 42 participants, 20 adults with disease and 21 controls.

The availability of the data rows and columns (variables) affect the optimal size of data (Wiemken and Kelley, 2019). ML can be implemented by data size ranged from thousands of records (Carslake *et al*., 2020; Song and Ying, 2015), to less than 500 records (Shaikhina *et al*., 2019) or even very small just less than 100 records (Burti *et al*., 2022; Cheng *et al*., 2020).

Class imbalance

Class imbalance is a serious issue that sparked the interest among ML researchers, often related to classification conditions. It refers to a condition when the instances of one class outnumber the instances of other classes. The predominant class is referred to as the majority class, while the class with lower cases is the minority class. However, in many applications, the class's lower instances are the more interesting ones, which exaggerates the problem (Saturuli *et al*., 2012). The class imbalance may be due to a variety of factors, the most prominent of which are economic constraints that many researchers face, or it may be due to the nature of some diseases that occur in populations with very low prevalence. Nguyen *et al*. (2009) attributed the possible causes of imbalance to intrinsic property or due to data acquisition restrictions such as cost, privacy, and time.

The imbalance ratio between minor and major classes ranges from 1:4 to 1:100; if it exceeds this ratio, the data is considered extremely imbalanced (Picek *et al*., 2018).

Handling class imbalance

Resampling technique

Data resampling techniques are typically divided into two types under-sampling and oversampling. Under-sampling reduces the number of events for a majority class so that it is the identical or comparable to the minority class. Over-sampling increases the number of occasions in the minority class in order to make it equal to the majority class. In an imbalanced multiclass setting, under-sampling limits the number of instances in all classes except the one with the fewest, while oversampling increases the number of instances in all classes except the one with the most (Picek *et al*., 2018). The third method is Synthetic Minority Oversampling Technique (SMOTE) which act by taking each minority class instance and introducing synthetic instances along the line segments connection any/all of the k minority class' nearest neighbors using Euclidean distance (Chawla *et al*., 2002).

Picek *et al*. (2018) evaluated the role of resampling techniques in their imbalanced data. Their results showed how various balancing techniques, over-sampling, under-sampling, and specially, SMOTE can help ML classifiers to achieve better results (improved entropy).

Romero *et al*. (2022) used down sampling for training data to treat class imbalance in their study which use RF and classification tree ML models to classify herds with inconclusive reactors at bovine tuberculosis surveillance tests as high- or low-risk. The classification tree model, trained on down-sampled data had higher sensitivity (74 % vs 70 %), lower specificity (89 % vs 93 %), and the same AUC (83 %) compared to the non-downsampled equivalent.

Weller *et al*. (2021) applied oversampling and SMOTE to alleviate the class imbalance due to the low prevalence (30%) in the training subset. The models built without resampling performed worse than models built using SMOTE or oversampling techniques.

Keshavarzi *et al*. (2020) investigated the impact of imbalanced data on ML algorithms performance. When algorithms applied to the balanced dataset, the results showed that up and down sampling resulted in significant ($P < 0.05$) upgrades in the abortion prediction models, with the down-sampling technique outperforming (higher AUC) the up-sampling technique. While the imbalanced data performance ranged from 32.3% (poor) to 69.2% (medium upward).

Tuning the algorithms' parameters

The tuning of hyperparameters is a critical issue in modern ML application problems. The default values of the model's hyperparameters can be obtained from specific software packages or manually configured using trial-and-error for adjusting the maximum values (Probst *et al*., 2019). Different hyperparameter initialization values may influence ML metrics such as accuracy (Nematzadeh *et al*., 2022).

Chen *et al*. (2022) performed ANN ML in their study, and to

set the optimal hyperparameters. ANNs were trained by various features, numbers of hidden layer(s) and neuron(s), training algorithms, learning rates and the error threshold as stopping criteria. Depending on trial-and error to reach the lowest RRMSE, a change in number of the hidden layers and neurons in each hidden layer were evaluated. The results showed they greatly affected the performance of ANN models. Therefore, a feedforward network with 2 hidden layers, the first layer with 3, and the second one with 6 hidden neurons is the best tuned hyperparameters for predicting nitrogen excretion in lactating dairy cows. Other studies obtained different parameters as there is no rule for the optimizing the numbers of hidden layer(s) and their neurons. (Craninx *et al*., 2008) developed an ANN model with one hidden layer and 6 hidden neurons for predicting rumen acetate, propionate and butyrate proportion.

The alternative method for tuning hyperparameters is using data-dependent tuning strategies Guyon *et al*. (2010), that attempt to minimize the expected generalization error of the inducing algorithm by evaluating predictions on an separate test set, or by running a resampling scheme such as cross-validation (Bischl *et al*., 2012). Simple grid search or random search strategies are available (Bergstra and Bengio, 2012).

Though few veterinary epidemiological studies adopt machine learning-based algorithms, and the majority of these studies ignore the importance of tuning for model parameters (Larison *et al*., 2014).

In a study of Chen *et al*. (2022), the results showed that the performance of ANN models were greatly improved by the turning process of selection of features and learning algorithms.

Tuning ML algorithms and using proper validation method are necessary to enhance generalizability and avoid overfitting problems, which smaller datasets are more especially susceptible to Charilaou and Battat (2022).

CONCLUSION

There is a significant impact of machine learning on veterinary medicine. This review focused primarily on disease prediction and only slightly on other occasions of solving veterinary issues. It is time for researchers and farm-decision makers to focus more on applying ML algorithms in the veterinary field. Conventional statistical methods necessitate establishing assumptions and comprehending hypotheses, which non-statisticians may be unfamiliar with.

The researchers that would apply ML models in their work should be cautious with several considerations to ensure that the ML prediction is reliable. The review summarized some of the challenges that may arise when implementing ML. The most important issue is the data quality used to train the models; the higher the quality, the better the model's predictive performance; otherwise, it's a case of garbage in, garbage out. Class imbalance should be addressed and treated; SMOTE technique provides more precise predictions. Another important point to note is that, while ML requires relatively large and high-dimensional data sets, it performs well when applied to small-sized data sets because it provides several techniques for dealing with small data sets as cross-validation techniques.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Anholt, R.M., Berezowski, J., Jamal, I., Ribble, C., Stephen, C., 2014. Mining free-text medical records for companion animal enteric syndrome surveillance. Prev. Vet. Med. 113, 417-422.

Appleby, R.B., Basran, P.S., 2022. Artificial intelligence in veterinary medicine. J. Am. Vet. Med. Assoc. 260, 819-824.

Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage 145, 137-165.

Awaysheh, A., Wilcke, J., Elvinger, F., Rees, L., Fan, W., Zimmerman, K.L., 2019. Review of Medical Decision Support and Machine-Learning Methods. Vet. Pathol. 56, 512-525.

Banerjee, A., Chen, S., Fatemifar, G., Zeina, M., Lumbers, R.T., Mielke, J., Gill, S., Kotecha, D., Freitag, D.F., Denaxas, S., Hemingway, H., 2021. Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility. BMC Medicine 19, 85.

Bergstra, J., Bengio, Y.J., 2012. Random search for hyper-parameter optimization. JMLR 13.

Bischl, B., Mersmann, O., Trautmann, H., Weihs, C., 2012. Resampling methods for meta-model validation with recommendations for evolutionary computation. Evolutionary Computation 20, 249-275.

Bobbo, T., Biffani, S., Taccioli, C., Penasa, M., Cassandro, M., 2021. Comparison of machine learning methods to predict udder health status based on somatic cell counts in dairy cows. Sci. Rep. 11, 13642.

Bradley, J., Rajendran, S., 2021. Increasing adoption rates at animal shelters: A two-phase approach to predict length of stay and optimal shelter allocation. BMC Vet. Res. 17, 1-16.

Burti, S., Zotti, A., Bonsembiante, F., Contiero, B., Banzato, T., 2022. A Machine Learning-Based Approach for Classification of Focal Splenic Lesions Based on Their CT Features. Front. Vet. Sci. 9.

Carslake, C., Vázquez-Diosdado, J.A., Kaler, J., 2020. Machine learning algorithms to classify and quantify multiple behaviours in dairy calves using a sensor: Moving beyond classification in precision livestock. Sensors 21, 88.

Charilaou, P., Battat, R., 2022. Machine learning models and over-fitting considerations. WJG. 28, 605-607.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. JAIR 16, 321-357.

Chen, X., Zheng, H., Wang, H., Yan, T., 2022. Can machine learning algorithms perform better than multiple linear regression in predicting nitrogen excretion from lactating dairy cows. Sci. Rep. 12, 12478.

Cheng, N., Schulte, A.J., Santosa, F., Kim, J.H., 2020. Machine learning application identifies novel gene signatures from transcriptomic data of spontaneous canine hemangiosarcoma. Brief. Bioinformatics 22, bbaa252.

Cockburn, M., 2020. Review: Application and Prospective Discussion of Machine Learning for the Management of Dairy Farms. Animals 10, 1690.

Craninx, M., Fievez, V., Vlaeminck, B., De Baets, B., 2008. Artificial neural network models of the rumen fermentation pattern in dairy cattle. Comput. Electron Agric. 60, 226-238.

Crone, S.F., Lessmann, S., Stahlbock, R.J., 2006. The impact of preprocessing on data mining: An evaluation of classifier sensitivity in direct marketing. Eur. J. Oper. Res. 173, 781-800.

El-Bayomi, K., Mohamed, F., Eltarabany, M., Gouda, H., 2019. Application of Different Biostatistical Methods in Biological Data Analysis. Z.V.J. 47, 203-212.

Ernst, A.F., Albers, C.J., 2017. Regression assumptions in clinical psychology research practice-a systematic review of common misconceptions. PeerJ 5, e3323.

Fernandes, A.F.A., Dórea, J.R.R., Rosa, G.J.d.M., 2020. Image analysis and computer vision applications in animal sciences: an overview. Front. Vet. Sci. 7, 551269.

Fernández, C., Soria-Olivas, E., Martín-Guerrero, J.D., Serrano, A.J., 2006. Neural networks for animal science applications: Two case studies. Expert Syst. Appl. 31, 444-450.

Flint, C., Cearns, M., Opel, N., Redlich, R., Mehler, D.M.A., Emden, D., Winter, N.R., Leenings, R., Eickhoff, S.B., Kircher, T., Krug, A., Nenadic, I., Arolt, V., Clark, S., Baune, B.T., Jiang, X., Dannlowski, U., Hahn, T., 2021. Systematic misestimation of machine learning performance in neuroimaging studies of depression. Neuropsychopharmacol. 46, 1510–1517.

Freedman, D.A., 2009. Statistical models: theory and practice. cambridge university press.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. Deep learning, Vol 1. MIT press Cambridge.

Guyon, I., Saffari, A., Dror, G., Cawley, G., 2010. Model selection: beyond the bayesian/frequentist divide. The Journal of Machine Learning Research 11, 61–87

Han, J., Pei, J., Kamber, M., 2011. Data mining: concepts and techniques, 3rd Edition. Elsevier, pp. 336-341.

Han, J., Pei, J., Tong, H., 2022. Data mining: concepts and techniques. Morgan kaufmann, 3rd edition, ISBN-10⁾: ⁾9780123814791.

Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H., 2009. The elements of statistical learning: data mining, inference, and prediction.2nd edition, Springer.

Holm, D.E., Thompson, P.N., Irons, P.C., 2009. The value of reproductive tract scoring as a predictor of fertility and production outcomes in beef heifers. J. Anim. Sci. 87, 1934-1940.

Hosmer, Jr.D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied logistic regression, Vol 398. John Wiley & Sons.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning, Vol 112. Springer.

Johnston, B.A., Steele, J.D., Tolomeo, S., Christmas, D., Matthews, K., 2015. Structural MRI-Based Predictions in Patients with Treatment-Refractory Depression (TRD). PLoS One 10, e0132958.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement learning: A survey. JAIR 4, 237-285.

Keshavarzi, H., Sadeghi-Sefidmazgi, A., Mirzaei, A., Ravanifard, R., 2020. Machine learning algorithms, bull genetic information, and imbalanced datasets used in abortion incidence prediction models for Iranian Holstein dairy cattle. Prev. Vet. Med. 175, 104869.

Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai, vol. 14 (Stanford, CA).

Kuhnert, P.M., Do, K.-A., McClure, R., 2000. Combining non-parametric models with logistic regression: an application to motor vehicle injury data. Comput Stat Data Anal. 34, 371-386.

Kwon, O.H., Rhee, W., Yoon, Y., 2015. Application of classification algorithms for analysis of road safety risk factor dependencies. Accid. Anal. Prev. 75, 1-15.

Larison, B., Njabo, K.Y., Chasar, A., Fuller, T., Harrigan, R.J., Smith, T.B., 2014. Spillover of pH1N1 to swine in Cameroon: an investigation of risk factors. BMC Vet. Res. 10, 1-8.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature, 521(7553), 436-444.

Li, W., Fang, H., Qin, G., Tan, X., Huang, Z., Zeng, F., Du, H., Li, S., 2020. Concentration estimation of dissolved oxygen in Pearl River Basin using input variable selection and machine learning techniques. Sci. Total Environ. 731, 139099.

Machado, G., Mendoza, M.R., Corbellini, L.G., 2015. What variables are important in predicting bovine viral diarrhea virus? A random forest approaches. Vet Res 46, 85.

Mohamad Hifzan, R., Izuan Bahtiar, A., Ariff, O., 2015. Describing growth pattern of Brakmas cows using non-linear regression models. Malays. J. Anim. Sci.18, 37-45.

Mohammed, M., Khan, M.B., Bashier, E.B.M., 2016. Machine learning: algorithms and applications. Crc Press.

Navarro, C.L.A., Damen, J.A., Takada, T., Nijman, S.W., Dhiman, P., Ma, J., Collins, G.S., Bajpai, R., Riley, R.D., Moons, K.G.J., 2021. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ 375.

Nematzadeh, S., Kiani, F., Torkamanian-Afshar, M., Aydin, N., 2022. Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases. Comput Biol Chem. 97, 107619.

Nguyen, G.H., Bouzerdoum, A., Phung, S.L., 2009. Learning pattern classification tasks with imbalanced data sets. pp. 193-208. https://ro.uow.edu.au/infopapers/792.

Parkhi, O.M., Vedaldi, A., Jawahar, C., Zisserman, A., 2011. The truth about cats and dogs. In: 2011 International Conference on Computer Vision, pp. 1427-1434.

Picek, S., Heuser, A., Jovic, A., Bhasin, S., Regazzoni, F., 2018. The Curse of Class Imbalance and Conflicting Metrics with Machine Learning for Side-channel Evaluations. TCHES 2019.

Probst, P., Boulesteix, A.L., Bischl, B., 2019. Tunability: Importance of Hyperparameters of Machine Learning Algorithms. JMLR 20, 1934-1965.

Punyapornwithaya, V., Klaharn, K., Arjkumpa, O., Sansamur, C., 2022. Exploring the predictive capability of machine learning models in identifying foot and mouth disease outbreak occurrences in cattle farms in an endemic setting of Thailand. Prev. Vet. Med. 207, 105706.

Rashid, B., Calhoun, V., 2020. Towards a brain-based predictome of mental illness. Hum Brain Mapp. 41, 3468-3535.

Ratner, B., 2010. Variable selection methods in regression: Ignorable problem, outing notable solution. J Target Meas Anal Mark. 18, 65-75.

Romero, M.P., Chang, Y.-M., Brunton, L.A., Parry, J., Prosser, A., Upton, P., Drewe, J.A., 2022. Machine learning classification methods informing the management of inconclusive reactors at bovine tuberculosis surveillance tests in England. Prev. Vet. Med. 199, 105565.

Romero, M.P., Chang, Y.-M., Brunton, L.A., Parry, J., Prosser, A., Upton, P., Rees, E., Tearne, O., Arnold, M., Stevens, K., Drewe, J.A., 2020. Decision tree machine learning applied to bovine tuberculosis risk factors to aid disease control decision making. Prev. Vet. Med. 175, 104860.

Saegerman, C., Porter, S.R., Humblet, M.F., 2011. The use of modelling to evaluate and adapt strategies for animal disease control. Rev. Sci. Tech 30, 555-569.

Santamaria, S.L., Zimmerman, K.L., 2011. Uses of informatics to solve real world problems in veterinary medicine. J Vet Med Edu. 38, 103-109.

Satuluri, N., Kuppa, M.R., 2012. A novel class imbalance learning using intelligent under-sampling. Int. J. Database Theory Appl. 5, 25-36.

Saxena, P., Parasher, Y., 2019. Application of artificial neural network (ANN) for animal diet formulation modeling. Procedia Comput. Sci. 152, 261-266.

Schnack, H.G., Kahn, R.S., 2016. Detecting neuroimaging biomarkers for psychiatric disorders: sample size matters. Front. Psychiatry 7, 50.

Schofield, I., Brodbelt, D.C., Kennedy, N., Niessen, S.J.M., Church, D.B., Geddes, R.F., O'Neill, D.G., 2021. Machine-learning based prediction of Cushing's syndrome in dogs attending UK primary-care veterinary practice. Sci. Rep. 11, 9035.

Shaikhina, T., Lowe, D., Daga, S., Briggs, D., Higgins, R., Khovanova, N., 2019. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. Biomedical Signal Processing and Control 52, 456-462.

Song, Y., Ying, L., 2015. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 27, 130.

Tirink, C., ÖNDER, 2022. Comparison of M, MM and LTS estimators in linear regression in the presence of outlier. Turkish J. Vet. Anim. Sci. 46, 420-128.

Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PLOS ONE 14, e0224365.

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.

Varoquaux, G., 2018. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 180, 68-77.

Wan, L., Bao, W., 2009. Animal disease diagnoses expert system based on SVM. In: International Conference on Computer and Computing Technologies in Agriculture, pp. 539-545.

Weller, D.L., Love, T.M.T., Wiedmann, M., 2021. Comparison of Resampling Algorithms to Address Class Imbalance when Developing Machine Learning Models to Predict Foodborne Pathogen Presence in Agricultural Water. Front. Environ. Sci. 9, Article 701288.

Wiemken, T.L., Kelley, R.R., 2019. Machine Learning in Epidemiology and Health Outcomes Research. Annu Rev Public Health. 41, 21-36.

Yan, L., Sun, Y., Boivin, M.R., Kwon, P.O., Li, Y., 2016. Revealing Facts and Avoiding Biases: A Review of Several Common Problems in Statistical Analyses of Epidemiological Data. Front Public Health 4, 207.

Yoo, S., Im, J., Wagner, J., 2012. Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. Landsc. Urban Plan 107, 293-306.

Zeldis, D., Prescott, S., 2000. Fish disease diagnosis program—problems and some solutions. Aquac Eng 23, 3-11.