REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

Applications of Machine Learning in Small Animal Medical Clinic: A Scoping Review

Vitória Souza de Oliveira Nascimento[a*]; Alexandre Dias Porto Chiavegatto Filho[b]
Faculty of Veterinary Medicine and Animal Science - University of São Paulo, São Paulo, Brazil.
* vitoriason@usp.br
[b]School of Public Health, University of São Paulo, Brazil.

Abstract
Artificial intelligence (AI) aims to develop computational strategies capable of performing human-like tasks. It has the potential of creating important tools to transform healthcare, and its applications are advancing faster every day. The objective of this study was to identify recent applications of machine learning in veterinary medicine, more specifically in the medical clinic of small animals. The search of studies published was conducted in Pubmed, CAB and Scopus databases. Studies were selected based on inclusion and exclusion criteria previously defined according to PRISMA guidelines. A total of 20 research articles were identified. Several different types of algorithms were used, with structured and unstructured data as input data and the most used performance metric was the area under the receiver operating characteristic (AUROC). Applications of machine learning can bring important benefits to the clinical practice of small animals, especially regarding diagnosis and prognosis.

Keywords: algorithms; cats; dogs; evidence-based veterinary medicine.

Introduction

The name Artificial Intelligence (AI) was first used by John McCarthy in 1956 with the software The Logic Theorist, which was created to simulate the problem-solving ability of human beings (Thakur et al, 2019). Before this event, Alan Turing had also previously raised the possibility that machines could simulate human behavior, and developed the Turing test to differentiate them from humans (Mintz & Brodie, 2019).

Methods based on AI have the potential to transform healthcare (Liang et al, 2019), and in some areas they are already transforming clinical practice and management (He et al, 2019). With the increase of digitized data, machine learning and computing infrastructure, AI technologies are expanding into areas that were previously considered to be the domain of only human specialists (Yu; Beam; Kohane, 2018).

Machine Learning (ML) is an area of AI research that aims to develop algorithms with the ability to learn how to perform a given task from experience (Faceli et al, 2011). Algorithms (a sequence of reasoning, instructions or operations to achieve a pre-established goal) are able to find hidden patterns in data, to classify objects based on the measurement of characteristics and to associate patients with specifict clinical and/or laboratory data to a certain health outcome (Altman, 2017).

Machine learning models have been frequently applied for risk assessment (Parikh; Kakad; Bates, 2016) and for predicting individual results, helping to advance personalized medicine. Its use offers an opportunity to develop algorithms for early diagnosis, analysis of medical images, and predictions of disease incidence, such as cancer and neurodegenerative diseases (Gultepe et al, 2016; Mintz & Brodie, 2019; Thakur et al, 2019).

In veterinary medicine, the use of these technologies will impact veterinarians as much as they are starting to impact physicians, helping to provide more efficient, accurate and assertive care to patients (Topol, 2019).

The use of machine learning methods in veterinary medicine has been very limited when compared to medicine (Anholt et al, 2014). With the increased use of these tools in veterinary medicine, it is expected that better evidence-based outcome assessments will become available (Awaysheh et al, 2019).

This article aims to provide a systematic review of recent machine learning applications in the small animal medical clinic, covering their specialties and summarizing the main findings and algorithmic strategies.

Metodology

Research strategies and articles eligibility

The systematic review was carried out in PUBMED, CAB Abstracts and Scopus. The search strategy was based on the use of boolean descriptors and operators (veterinary OR veterinary medicine) AND (dog OR canine OR cat OR feline OR small animals) AND (artificial intelligence OR deep learning OR machine learning). Subsequently, all studies found in the literature were screened according to the title/abstract and full text.

All studies published in peer-reviewed journals were considered, without language restriction. The systematic research was carried out between January 29 and 30, 2021.

Inclusion and exclusion criteria

To select eligible studies, the following criteria were used: i) the data reported should belong to primary research that directly applies machine learning; ; ii) did not use proteomics, genomics or metabolomics data; iii) studies with small animals (dogs and cats); iv) applications exclusively in medical clinic. In the case of poorly explanatory abstracts or concerns about the available data, studies were inspected for further consideration. Studies that met the criteria were fully evaluated.

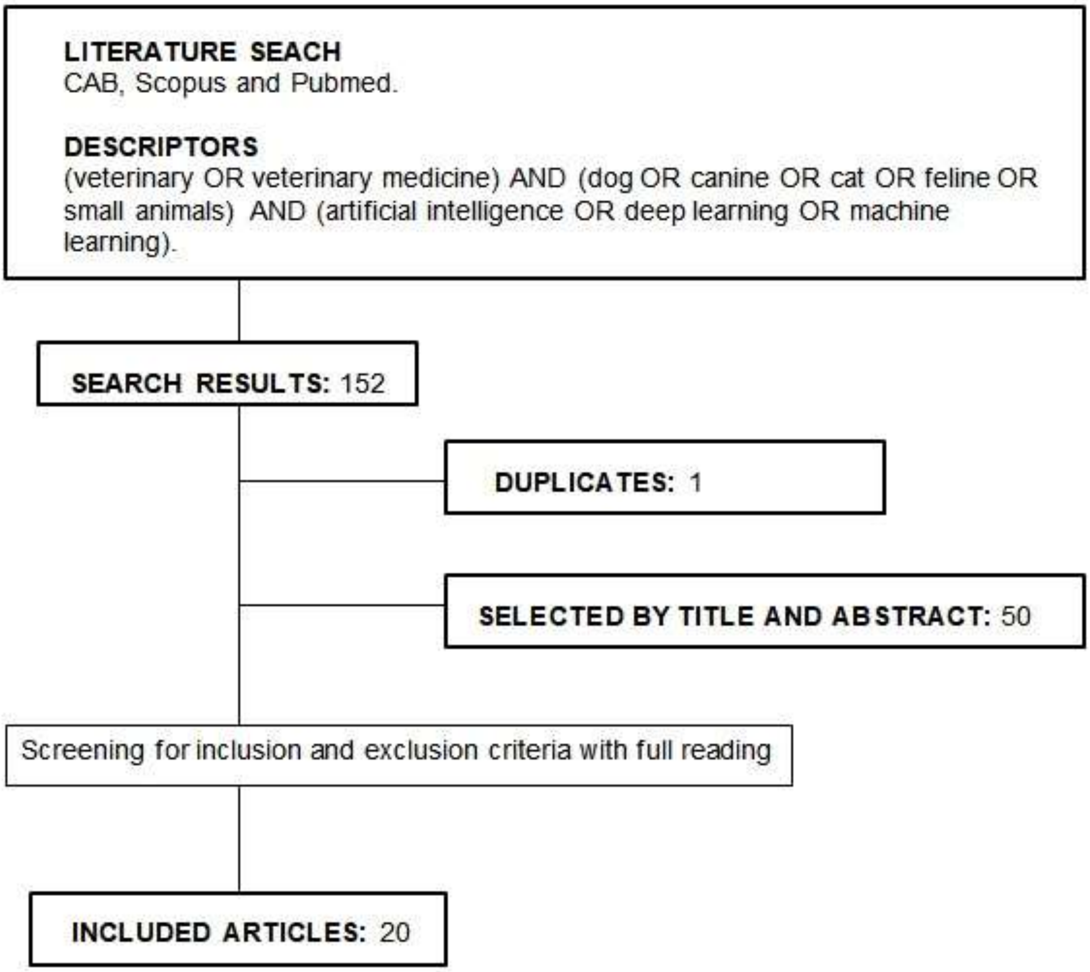
Data extraction and analysis

The studies were classified regarding authorship, title, year of publication, objective of the study, type of algorithm used and result, in order to guarantee the reliability of the collected data.

Results

A total of 152 articles were found in the three electronic databases: Pubmed (n = 151), CAB (n = 1) and Scopus (n = 0). The occurrence of duplicity was analyzed, causing the exclusion of 1 article, leaving 151 for individual evaluation. After screening for title and abstract, 50 articles were selected for full reading (Figure 1). Then a further 30 articles were excluded as they were not applications in medical clinic or because there was no application of machine learning, leading to the inclusion of 20 articles for the detailed analysis (Table 1).

Figure 1 - Flowchart of search, selection and inclusion of studies.



REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

Table 1 - Summary of selected studies.

Author	**Objective of the study**	**Algorithm**[a]	**Performance Metrics**[b]	**Training and test**
McEvoy & Amigo (2012)	Classify dog radiographs	*ANN* and PLS-DA	ANN: Sen.=86%; Spe.=100%; PLS-DA: Sen.=100% Spe.=89%	Training = 78% Test= 22%
Mirkes et al (2014)	Predict lymphoma diagnosis, screening and prognosis	RF, KNN and PDE	AUROC: 0.77 – 0.91	Training = 80% Test= 20%
Awaysheh et al (2016)	Differentiate inflammatory bowel disease from lymphoma in cats	NB, DT and ANN	AUROC: NB= 0.95; DT= 0.86 ANN= 0.94	Training = 90% Test= 10%
Brinkmann et al (2016)	Identification of pre-ictal states	SVM	AUROC: 0.72	Training = 80% Test= 20%
Torrecilha et al (2017)	Predict parasitic load on lymph nodes	ANN	Ac. = 0.869	Ranged from 10 to 55 (changing input data) with 18 different combinations
Banzato et al. (2018a)	Differentiate meningiomas and gliomas	CNN	Sen.=91% Spe.= 91%	Training= 70%, Validation= 15% Test= 15%

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022;
Publication: 30 March 2022

Banzato et al (2018b)	Detect degenerative liver disease	CNN	AUROC: 0.91	Training = 70%, Validation = 15% Test = 15%
Banzato et al (2018c)	Canine meningiomas graduation prediction	CNN associated or not with AlexNet	Ac.= 82,2% 68%	Training = 60%, Validation = 10% Test= 30%
Yoon; Hwang; Lee (2018)	Distinguish between normal and abnormal radiographic findings	CNN and BOF	Ac.=: Ranged from 79.6% to 96.9% in both	Training = 75% Test= 25%
Bradley et al. (2019)	Predict chronic kidney failure in cats	RNN	Sen.= 90.7% Spe.= 98.9%	Training = 63% Test= 37%
Kim et al. (2019)	Determine the severity of corneal ulcer in dogs	CNN	Ac.=: >90%	Training= 74% Test = 26%
Spiteri et al. (2019)	Identify neuromorphological change and biomarkers	*SFS*	AUROC: CMP = 0,72 SMS= 0,82	Training = 85% Test= 15%
Biourge et al.(2020)	Predict chronic kidney failure in elderly cats	ANN - MLP	Sen.= 87% Spe.= 70%	Training = 6% Test= 94%

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

Boissadyet al (2020)	Track thoracic injuries	CNN	AUROC: 0,86	Training = 90% Test= 10%
Burrai et al (2020)	Predict the diagnosis of breast cancer in dogs	SGB	Ac.=: 63%	Training = 80% Test= 20%
Burti et al., (2020)	Detect cardiomegaly through chest radiography	CNN	AUROC: >0,9	Training = 78,5% Test= 21,5%
Franzo et al (2020)	Predict the probability of survival	RF and LR	AUROC: RF= 0.997 E RL= 0.831	Training = 80% Test= 20%
Li et al (2020)	Detect left atrial enlargement using chest radiography	CNN	Ac.= 82.71% Sen.= 68.42% Spe.= 87.09%	Training = 90% Test= 10%
Reagan; Reagan; Gilor (2020)	Predict diagnosis of hypoadrenocorticism in dogs	AdaBoost	AUROC: 0,994	Training = 80% Test= 20%

Bohannan et al (2021)	Prediction of the effectiveness of chemotherapy drugs for canine lymphoma	RF	AUROC: >0,95	Training = 70% Test= 30%

[a]: Random Forest (RF); k-Nearest Neighbors (KNN); Support Vector Machine (SVM); Artificial Neural Network (*ANN*); Logistic Regression (LR); Stochastic Gradient Boosting (SGB); Sequential forward selection (SFS); Decision Tree (DT); Probability Density Evaluation (PDE); Partial Least Squares-Discriminant Analysis (PLS-DA); Bag of features (BOF); CNN= Convolutional Neural Network; Artificial Neural Network Multi Layer *Perceptron (ANN-MLP); Recurrent Neural Network (RNN).*
[b]: Sen.= Sensitivity; Spe. = Specificity; Ac. = Accuracy; AUROC = Area Under the Receiver Operating Characteristic.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

Summary of the results

The studies tested the predictive performance of a wide range of algorithms, depending on the type of data (10 studies used structured data, and 10 used images). Input data for the algorithms included the result of anamnesis, physical, laboratory and image exams (radiography, computed tomography and ultrasound), in addition to brain signals from electroencephalograms. The results show possible applications in several clinical specialties, such as medical clinic (4 studies), diagnostic imaging (6 studies), oncology (6 studies), nephrology (2 studies), neurology (1 study) and ophthalmology (1 study).

Types of Algorithms used according to data type

Structured data

The random forest algorithm was tested by Mirkes et al (2014), Franzo et al (2020) and Bohannan et al (2021), for predictions using clinical and laboratory data, obtaining good results from the AUROC, while Reagen et al (2021) used AdaBoost.

Image identification

Convolutional artificial neural networks were the algorithm most frequently used to detect patterns and changes in images, in addition to bag-of-features and SFS.

Countries where the studies were carried out

The studies were carried out in the United States of America (6/20 = 30%), United Kingdom (4/20 = 20%), Italy (5/20 = 25%), France (1/20 = 5%), South Korea (2/20 = 10%), Denmark (1/20= 5%) and Brazil (1/20= 5%). In four of them, researchers from different countries collaborated to carry out the study.

Programming languages

The programming languages used in the study were R (6/20 = 30%), Python (5/20 = 25%) and Java (9/20 = 45%). The languages were used to run the algorithms in their respectives softwares, in addition to MATLAB.

Performance metrics

In 4 studies, the authors used sensitivity (ratio between true positives and total cases) and specificity (ratio between true negatives and total cases), in 6 accuracy (ratio of true positives to true negatives), and in 10 the main metric of performance was the area under the receiver operating characteristic (AUROC).

Discussion

Satisfactory performance results were found in all studies that applied ML, both in studies that worked with the results of laboratory tests and in those that analyzed images for screening and diagnosis. The most used programming languages were R and Python and and the most used performance metric was the AUROC. Neural networks were the frequent

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

algorithms, mainly *recurrent neural network (*CNN) and *recurrent neural network (*RNN). In addition, almost half of the studies were published in 2020, with a noticeable increase in the number of publications starting in 2018.

In the study by Mirkes et al (2014), random forest (RF) performed better in the three outputs tested (differential diagnosis and screening), obtaining an AUROC of 0.771, 0.879 and 0.917 respectively. Franzo et al (2020) were able to predict the survival of dogs after infection by the Parvovirus using RF (AUROC = 0.997), which obtained a better result when compared to logistic regression (AUROC = 0.831).

RF is an algorithm based on decision trees, which have an organization similar to a flowchart, where conditions are checked, and if the flow is met, it follows one branch until the end of the flow (Breiman, 2001). RF is created by an ensemble of many decision trees, selected through bootstrap aggregating.

AdaBoost is quick and simple to program and work, creating a highly accurate prediction rule by combining many relatively weak rules, selecting the best among the weak classifiers and combining them to obtain a strong classifier (Abualkibash et al, 2013). It was used to make predictions by Bohannan et al (2021), obtaining an AUROC of 0.95, with a simple and inexpensive entry data in clinical routine, managing to predict the occurrence of canine lymphoma.

The two previous algorithms are ensemble techniques that combine the result of multiple models in order to produce a better predictive model (Rokach, 2010). There are several algorithms for ensemble classifiers, such as bagging, boosting, bayesian averaging, among others. In ensemble methods, the predictive performance of the model increases with the combination of weaker models (Zhou; Tang, 2006).

CNNs have dominated some activities, especially those related to medical images, mainly radiology (Yamashita et al, 2018). The algorithm is a deep neural network capable of receiving an input image, assigning weights to various aspects of the image and apllying convolutions techniques. Its wide use in image studies can be explained by its high predictive performance and its lower need for pre-processing of the images when compared to other techniques (Suganuma; Shirakawa; Nagao, 2017).

Yoon; Hwang; Lee (2018) compared bag of features (BOF) with CNN and found that having both can be useful in improving the performance of double reading (interpretation of imaging exams by two specialists). Since BOF is a simpler model of ML, Christmann et al (2018) showed in their study that less complex techniques remain relevant in certain scenarios.

Bradley et al (2019) and Biourge et al (2020) both used neural networks to predict the occurrence of chronic renal failure in cats. The architecture of the algorithm was based on a perceptron multilayer artificial neural network (ANN) with backpropagation and a recurrent ANN, using similar data as input (creatinine, urea nitrogen blood and urinary density, with Bradley et al (2019) also adding age as a predictor).

Banzato et al (c) used neural networks associated or not with AlexNet (convolutional neural network with 8 layers of depth). According to the authors, the main difference between the two CNNs is that AlexNet was trained with a data set of millions of images, while the other CNN was trained with a small set of data.

The most frequently used programming languages in machine learning in health are Python and R (Chiavegatto Filho, 2015). Our review confirmed that this is also the case in studies of small animals, as half of the selected articles used one of the two languages.

The studies presented some limitations, such as the statistical validation of the methods on new samples, the very limited size of some of the databases, the presence of an imbalance between the number of control and cases of the samples, and the application of techniques that have high performance on small datasets, but may not show satisfactory results on larger samples. Contrary to what is found in medicine, most veterinary records are free text, without a clear tabular coding, which is possibly the main barrier to increase the amount of studies with machine learning applications and translational research (Zhang et al, 2019). Despite the positive results of the application of machine learning in the clinical practice of small animals, there are still important technical challenges to overcome, such as the inclusion of complex, heterogeneous and large data sets (Altman, 2017).

Conclusion

Overall, the studies found promising results regarding the use of laboratory tests and diagnostic imaging for predictive analyses in small animals, which can soon be an important tool for the daily clinical practice of veterinarians. Despite this potential, the vast majority of existing studies are still based on applications developed in research environments, using data collected retrospectively. There are still important challenges regarding data standardization and interoperability of machine learning applications to facilitate their routine use in small animals.

Declaration Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Abualkibash, M., ElSayed, A., Mahmood, A., 2013. Highly scalable, parallel and distributed AdaBoost Algorithm using light weight threads and web services on a network of multi-core machines. International Journal of Distributed and Parallel Systems,4, 29-40. Doi: 10.5121/ijdps.2013.4303

[2] Altman RB., 2017. Artificial intelligence (AI) systems for interpreting complex medical datasets. Clin Pharmacol Ther., 101, 585-586. doi: 10.1002/cpt.650.

[3] Anholt, RM, Berezowski, J, Maclean, K,, Russel, M.L., Jamal, I., Stephen, C., 2014.The application of medical informatics to the veterinary management programs at companion animal practices in Alberta, Canada: A case study . Prev Vet Med., 113, 165 - 174 https://doi.org/10.1016/j.prevetmed.2013.11.005

[4] Athey, S., Tibshirani,J.,; Aposta, S., 2019. Generalized random forests. Ann. Estatista., 2, 1148--1178. doi:10.1214/18-AOS1709

[5] Atkins, C., Bonagura, J., Ettinger, S., Fox, P., Gordon, S., Haggstrom, J., Hamlin, R., Keene, B., Luis-Fuentes, V., Stepien, R., 2009. Guidelines for the diagnosis and treatment of canine chronic valvular heart disease. J Vet Intern Med, 23, 1142-1150.https://doi.org/10.1111/j.1939-1676.2009.0392.x

[6] Awaysheh A, Wilcke J, Elvinger F, Rees L, Fan W, Zimmerman KL., 2019. Review of Medical Decision Support and Machine-Learning Methods. Veterinary Pathology, 56, 512-525. doi:10.1177/0300985819829524

[7] Awaysheh, A., Wilcke, J., Elvinger, F., Rees, L., Fan, W., Zimmerman, K.L., 2016. Evaluation of supervised machine-learning algorithms to distinguish between inflammatory bowel disease and alimentary lymphoma in cats. J Vet Diagn Invest., 28, 679-687. doi: 10.1177/1040638716657377.

[8] Banzato, T., Bernardini, M., Cherubini, G.B., Zotti, A., 2018. A methodological approach for deep learning to distinguish between meningiomas and gliomas on canine MR-images. BMC Vet Res., 1-6.. doi: 10.1186/s12917-018-1638-2.

[9] Banzato, T., Bonsembiante, F., Aresu, L., Gelain, M.E., Burti, S., Zotti, A. 2018. Use of transfer learning to detect diffuse degenerative hepatic diseases from ultrasound images in dogs: A methodological study. Vet J., 233, 35-40. doi: 10.1016/j.tvjl.2017.12.026.

[10] Banzato, T., Cherubini, G.B., Atzori, M., Zotti, A., 2018. Development of a deep convolutional neural network to predict grading of canine meningiomas from magnetic resonance images. Vet J., 235, 90-92. doi: 10.1016/j.tvjl.2018.04.001.

[11] Bartges, J.W., 2012. Chronic kidney disease in dogs and cats. Veterinary Clinics: Small Animal Practice 42, 669-692. https://doi.org/10.1016/j.cvsm.2012.04.2008

[12] Biourge, V., Delmotte, S., Feugier, A., Bradley, R., McAllister, M., Elliott, J., 2020. An artificial neural network-based model to predict chronic kidney disease in aged cats. J Vet Intern Med. 34, 1920-1931. https://doi.org/10.1111/jvim.15892.

[13] Bohannan, Z., Pudupakam, R.S., Koo, J., Horwitz, H., Tsang, J., Polley, A., Han, E.J., Fernandez, E., Park, S., Swartzfager, D., Qi, N.S.X., Tu, C., Rankin, W.V., Thamm, D.H., Lee, H.R., Lim, S., 2021 Predicting likelihood of in vivo chemotherapy response in canine lymphoma using ex vivo drug sensitivity and immunophenotyping data in a machine learning model. Vet Comp Oncol., 19, 160-171. doi: 10.1111/vco.12656.

[14] Boissady, E., de La Comble, A., Zhu, X., Hespel, A.M., 2020. Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinarians in conjunction with the artificial intelligence. Vet Radiol Ultrasound., 61, 619-627. doi: 10.1111/vru.12912.

[15] Bradley R, Tagkopoulos I, Kim M, Kokkinos Y, Panagiotakos T, Kennedy J, De Meyer G, Watson P, Elliott J., 2019. Predicting early risk of chronic kidney disease in cats using routine clinical laboratory tests and machine learning. J Vet Intern Med. 33, 2644-2656. https://doi.org/10.1111/jvim.15623.

[16] Breiman, L., 2001. Random Forests. Machine Learning, 45, 5-32. https://doi.org/10.1023/A:1010933404324

[17] Brinkmann, B.H., Patterson, E.E., Vite, C., Vasoli, V.M., Crepeau, D., Stead, M., Howbert, J.J., Cherkassky, V., Wagenaar, J.B., Litt, B., Worrell, G.A., 2015. Forecasting Seizures Using Intracranial EEG Measures and SVM in Naturally Occurring Canine Epilepsy. PLoS One. 4, 10- 18. doi: 10.1371/journal.pone.0133900.

[18] Buchanan, J.W., 1999. Prevalence of cardiovascular disorders. In: Fox, P.R., Sisson, D., Moise, N.S. (Eds.) Textbook of Canine and Feline Cardiology: Principles and Clinical Practice. Philadelphia, PA: WB Saunders, Philadelphia, PA, 457-470.

[19] Burti S, Longhin Osti V, Zotti A, Banzato T., 2020. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. Vet J. 262, 01-07.
https://doi.org/10.1016/j.tvjl.2020.105505.

[20] Chengsheng, Tu & Huacheng, Liu & Bing, Xu. (2017). AdaBoost typical Algorithm and its application research. MATEC Web of Conferences. 139. 00222. 10.1051/matecconf/201713900222.

[21] Chiavegatto Filho, A.D.P., 2015. The use of big data in healthcare in Brazil: perspectives for the near future. Epidemiol. Serv. Saúde, 24,. 325-332. http://dx.doi.org/10.5123/S1679-49742015000200015

[22] Christmann, G., Montenegro, F., Grando, R., & Guerra, R., 2018. Implementação de um Modelo Bag of Features para Classificação de Frutas. *Revista ComInG - Communications and Innovations Gazette,3,* 70-80. doi:https://doi.org/10.5902/2448190430228

[23] Fan, J.,Upadhye, S.,Worster, A., 2006. Understanding receiver operating characteristic (ROC) curves. CJEM, 8, 19-20. 10.1017/S1481803500013336.

[24] Fernandes, F.T., Chiavegatto Filho, A.D.P., 2019. Data mining and machine learning perspectives for occupational safety and health. Rev. bras. saúde ocup., 44, 1-12. https://doi.org/10.1590/2317-6369000019418

[25] Franzo, G., Corso, B., Tucciarone, C.M., Drigo, M., Caldin, M., Cecchinato, M., 2020. Comparison and validation of different models and variable selection methods for predicting survival after canine parvovirus infection. Vet Rec. 31, 187-189 doi: 10.1136/vr.105283.

[26] Gultepe, E., Green, J.P., Nguyen, H., Adams, J., Albertson, T., Tagkopoulos, I., 2014. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. J Am Med Inform Assoc, 21, 315-25 10.1136/amiajnl-2013-001815

[27] He, J., Baxter, S.L., Xu, J., Xu, J., Zhou, X., Zhang, K., 2019. The practical implementation of artificial intelligence technologies in medicine. Nat Med 25, 30–36. https://doi.org/10.1038/s41591-018-0307-0

[28] Keene, B. W., Atkins, C. E., Bonagura, J. D., Fox, P. R., Häggström, J.,Fuentes, V. L., Oyama, M. A., Rush, J. E., Stepien, R., Uechi, M., 2019. ACVIM consensus guidelines for the diagnosis and treatment of myxomatous mitral valve disease in dogs. Journal Veterinary Intern Medicine, 33, 1127-1140.https://doi.org/10.1111/jvim.15488

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 23, No. 2 (2022)
http://www.veterinaria.org
Article Received: 28 December 2021; Revised: 20 January 2022; Accepted: 21 January 2022; Publication: 30 March 2022

[29] Kim, J.Y., Lee, H.E., Choi, Y.H., Lee, S.J., Jeon, J.S., 2019. CNN-based diagnosis models for canine ulcerative keratitis. Sci Rep.,9, 1-7. doi: 10.1038/s41598-019-50437-0.

[30] Li S, Wang Z, Visser LC, Wisner ER, Cheng H., 2020. Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. Vet Radiol Ultrasound. 61, 611-618. doi: 10.1111/vru.12901.

[31] Liang, H., Tsui, B.Y., Ni, H. *et al*, 2019. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med, 25, 433–438. https://doi.org/10.1038/s41591-018-0335-9

[32] Loncaric, F., Camara, O., Piella, G., Bijnens. B., 2021. La integración de la inteligencia artificial en el abordaje clínico del paciente: enfoque en la imagen cardíaca. Revista Española de Cardiología.,74, 72-80. https://doi.org/10.1016/j.recesp.2020.07.012

[33] McEvoy, F.J., Amigo, J.M., 2013. Using machine learning to classify image features from canine pelvic radiographs: evaluation of partial least squares discriminant analysis and artificial neural network models. Vet Radiol Ultrasound. 54, 122-1266. doi: 10.1111/vru.12003.

[34] Mintz, Y., Brodie, R., 2019. Introduction to artificial intelligence in medicine. Minim Invasive Ther Allied Technol., 28, 73-81. doi: 10.1080/13645706.2019.1575882.

[35] Mirkes, E.M., Alexandrakis, I., Slater, K., Tuli, R., Gorban, A.N. 2014. Computational diagnosis and risk evaluation for canine lymphoma. Comput Biol Med., 53, 279-90. doi: 10.1016/j.compbiomed.2014.08.006.

[36] Parikh R.B., Kakad M., Bates D.W., 2016. Integrating predictive analytics into high-value care: the dawn of precision delivery. JAMA 315, 651-652. https://doi.org/10.1001/jama.2015.19417

[37] Reagan, K.L., Reagan, B.A., Gilor, C.,2020 Machine learning algorithm as a diagnostic tool for hypoadrenocorticism in dogs. Domest Anim Endocrinol.,72, 1-9. doi: 10.1016/j.domaniend.2019.106396.

[38] Rokach, l., 2010. Ensemble-based classifiers. Artif Intell Rev 33, 1-39. https://doi.org/10.1007/s10462-009-9124-7

[39] Sagi, O., Rokach, L., 2018. Ensemble learning: A survey.. WIREs Data Mining Knowl Discov., 8, 1-18. https://doi.org/10.1002/widm.1249

[40] Shiraishi, J., Qiang, L., Appelbaum, D., Doi, K.,2011. Computer-aided diagnosis and artificial intelligence in clinical imaging. Semin Nucl Med, 41, 449- 462. doi: 10.1053/j.semnuclmed.2011.06.004.

[41] Spiteri, M., Knowler, S.P., Rusbridge, C., Wells, K., 2019. Using machine learning to understand neuromorphological change and image-based biomarker identification in Cavalier King Charles Spaniels with Chiari-like malformation-associated pain and syringomyelia. J Vet Intern Med., 33, 2665-2674. doi: 10.1111/jvim.15621.

[42] Streiner, D. L., Cairney, J., 2007. What’s Under the ROC? An Introduction to Receiver Operating Characteristics Curves. Can J Psychiatry, 52, 121–128.

[43] Suganuma, M.; Shirakawa, S.; Nagao, T. 2017. GECCO '17: Proceedings of the Genetic and Evolutionary Computation Conference, 497–504 https://doi.org/10.1145/3071178.3071229.

[44] Thakur, A., Mishra, A.P., Panda, B., Rodríguez, D. C. S., Gauray, I., Majhi, B., 2020. Application of Artificial Intelligence in Pharmaceutical and Biomedical Studies. Current Pharmaceutical Design, 26, 3569-3578. DOI: 10.2174/1381612826666200515131245.

[45] Topol, E.J., 2019. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25, 44–56. https://doi.org/10.1038/s41591-018-0300-7

[46] Torrecilha, R.B., Utsunomiya, Y.T., Batista, L.F., Bosco, A.M., Nunes, C.M., Ciarlini, P.C., Laurenti, M.D., 2017. Prediction of lymph node parasite load from clinical data in dogs with leishmaniasis: An application of radial basis artificial neural networks. Vet Parasitol., 30, 13-18. doi: 10.1016/j.vetpar.2016.12.016.

[47] Wyner, A.J., Olson, M., Bleich, J., 2017. Explaining the Success of AdaBoost and Random Forests as Interpolating Classifiers. Journal of Machine Learning Research, 18, 1-33

[48] Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K., 2018. Convolutional neural networks: na overview and application in radiology. Insights Imaging 9, 611-629. https//doi.org/10.1007/s13244-018-0639-9

[49] Yoon, Y., Hwang, T., Lee, H., 2018. Prediction of radiographic abnormalities by the use of bag-of-features and convolutional neural networks. The Veterinary Journal, 237, 43-48. https://doi.org/10.1016/j.tvjl.2018.05.009

[50] Yu, KH., Beam, A.L. & Kohane, I.S., 2018. Artificial intelligence in healthcare. Nat Biomed Eng 2, 719–731. https://doi.org/10.1038/s41551-018-0305-z

[51] Zhang, Y., Nie, A., Zehnder, A., Page, R. L., Zou, J., 2019. VetTag: improving automated veterinary diagnosis coding via large-scale language modeling. npj Digit. Med.,35, 1-8. https://doi.org/10.1038/s41746-019-0113-1

[52] Zhou, Z., Tang, W., 2006. Clusterer ensemble. Knowledge-Based Systems,19, 77-83. https://doi.org/10.1016/j.knosys.2005.11.003.