The unmet potential of artificial intelligence in veterinary medicine

Parminder S. Basran, PhD, FCCPM1*, and Ryan B. Appleby, DVM, DACVR2

1Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY
2Department of Clinical Studies, Ontario Veterinary College, University of Guelph, Guelph, ON, Canada

*Corresponding author: Dr. Basran (psb92@cornell.edu)

https://doi.org/10.2460/ajvr.22.03.0038

ABSTRACT
Veterinary medicine is a broad and growing discipline that includes topics such as companion animal health, population medicine and zoonotic diseases, and agriculture. In this article, we provide insight on how artificial intelligence works and how it is currently applied in veterinary medicine. We also discuss its potential in veterinary medicine. Given the rapid pace of research and commercial product developments in this area, the next several years will pose challenges to understanding, interpreting, and adopting this powerful and evolving technology. Artificial intelligence has the potential to enable veterinarians to perform tasks more efficiently while providing new insights for the management and treatment of disorders. It is our hope that this will translate to better quality of life for animals and those who care for them.

V
terinary medicine is a broad and growing discipline that includes topics such as companion animal health, population medicine and zoonotic diseases, and agriculture. Similarly, artificial intelligence (AI) touches many fields of science including philosophy, mathematics, neuroscience, control theory and cybernetics, computer engineering, and the data sciences. The intersection of these 2 broad and growing fields has the potential for one to dramatically affect the other: the applications of AI in veterinary medicine are near limitless, and, in turn, advances in veterinary AI can affect and promote advances in AI in adjacent fields. In our companion article in the May issue of the JAVMA, we briefly describe the definitions and concepts of AI. In the present article, we provide more insight on the underpinnings of AI and how it is currently applied in veterinary medicine, and we discuss the opportunities in veterinary medicine. More specifically, we provide some insight on emerging trends in AI and veterinary medicine and other topics of interest to the readership of this publication.

Artificial Intelligence Fundamentals

Three broad categories encompass AI study: speech recognition, computer vision, and natural language processing (NLP). Most readers have direct experience with speech recognition, in which a computer system listens to and interprets audible commands or instructions. Speech recognition systems tackle the problem of ascribing to an action or function from voice commands by converting an audible signal to words and establishing a semantic understanding of those words for the purpose of extracting meaning or intent.1,2 Another major field of study is computer vision, in which a computer replaces functions of the human eye and brain.3 Computer vision applications can be adopted in 2-D, 3-D, and higher order imaging, such as video analysis. Common applications of computer vision include optical character recognition (eg, converting pixelated images of text into alphanumeric characters), real-time imaging tasks for the purposes of recognizing and detecting objects (eg, disease detection in medical images), and object movement tracking (eg, continuous monitoring of an object in a video). The third field of AI study is NLP, which concerns attempts to extract information from text, such as deciphering text from a discharge note or medical report.4 How AI might be leveraged for any of these purposes in veterinary medicine depends on the specific task and the available data types. Each field of study offers numerous potential applications for AI in veterinary medicine.

Machine learning basics

Machine learning (ML) is an important subfield of AI wherein algorithms are trained to execute a specific task by learning from patterns in the data.5 Any of a spectrum of ML algorithms, including supervised, semisupervised, and unsupervised algorithms, could be used for a given task and its associated data (Figure 1).6,7 Supervised ML algorithms...
are used when all data are labeled (or classified) and the task is to classify new input data to known output responses (also commonly referred to as structured prediction). The outputs could be categorical (e.g., classification of disease), ordinal (e.g., stage of disease), or continuous (e.g., prediction of a number). Unsupervised learning algorithms are used when the data sets are not labeled and the task is to determine whether there are patterns within the data for inference. As in supervised ML, data types can be either discrete or continuous (Figure 1).

In the structured prediction or supervised ML process, the entire data set is divided into training, validation, and testing data sets, typically in a percentage distribution like 70/20/10, respectively. As in the human domain, the more exposure the ML algorithm has to different data types and their corresponding range of possible outputs, the more adaptable and accurate the ML system becomes when exposed to new data. After a model is trained and hyperparameters (parameters associated with how the ML system learns from the data) are tuned, a model is finalized. The final performance of the ML model is evaluated by testing the model with the remaining (unseen) test data.

The ML algorithm itself can have many forms, such as support vector machines, random forest or decision trees, naive Bayes, neural networks, and (logistic) regression. Figure 1 provides a number of algorithms for supervised and unsupervised ML tasks. Each algorithm has tunable model parameters that are iteratively adjusted during the training process to achieve a desired result. At the heart of any ML model optimization is a cost or loss function that measures the extent of agreement or disagreement between a prediction on the basis of the current model parameters and actual outputs. If the cost function is high for a given set of model parameter values, the model parameters are modified, a new set of predictions are computed, and the cost function is recomputed. This process is repeated until a set of model parameters that minimize the cost function is obtained. A subset of supervised ML is deep learning, wherein training is performed on multiple layers in a neural network.

Unsupervised ML algorithms are used primarily for data inference or to draw reasonable conclusions or potential hypotheses from the data set. Unsupervised ML algorithms can group large data sets into meaningful and interpretable cohorts and subcohorts of the data (clustering). Many unsupervised ML methods are not strictly algorithms that learn from the data but ones that can reduce the dimensionality of the data to fewer—and hopefully more manageable—data elements. Common unsupervised ML algorithms that can reduce the dimensionality of numerical (interval and ratio) data include principal component analysis, independent component analysis, projection methods (e.g., t-SNE and Isomap), and feature selection methods. For categorical (nominal and ordinal) data sets, clustering approaches (fuzzy C-means) can be used to group unlabeled data sets. Standard statistical approaches (correlation, covariance, etc.) can then be used to explore relationships between variables for inductive purposes.

Other examples of ML include reinforcement learning and semisupervised ML. Semisupervised ML is the branch of AI that combines both supervised and unsupervised ML and is useful when there is a combination of labeled and unlabeled data sets. Reinforcement learning is different from supervised and unsupervised ML: it more closely resembles how the human brain works (Figure 1). In this approach, the final model is achieved through the maximization of rewards, as opposed to minimizing a cost or loss function. Unlike supervised ML methods, the data does not need to be labeled and the algorithm can be designed to generate labels for unlabeled data as an output. The algorithm attempts to solve a problem with different approaches and is rewarded (or punished) on the basis of the actions taken.

### Image Analysis and Deep Learning

Veterinary medical data can take many forms, including text in a medical report, 2-D image data (an ultrasound image, image of a blood smear, or pathology slide), 3-D volumetric data (CT, PET-CT, MRI), and even 4-dimensional images (4-dimensional CT). The choice of ML depends on the type of data to be analyzed and the desired task or output. As described in the companion piece, NLP can be used to analyze text data. Most readers are familiar with basic statistical parameter estimates (mean and SD) for describing and analyzing 1-dimensional data and
basic correlation analysis and regression modeling. Machine learning algorithms that generate an output may be thought of as sophisticated regression modeling. For 2-D and higher-dimensional data, more sophisticated data analysis and modeling are often required. Analysis of image data with ML has grown dramatically over the last decade, fueled by faster computers and image-based neural networks, which in its most common form is a convolutional neural network (CNN). While more detailed description of CNNs in veterinary medicine may be found elsewhere, we only briefly describe them here and discuss their importance in medical image analysis.

Simply put, CNNs are neural networks for images. Convolutional neural networks were initially designed for the purpose of classifying objects in an image acquired from a camera. A large database consisting of over 14 million images of 1,000 objects is used to train a CNN, which can later be used to classify objects in new, unseen images. These CNNs themselves are powerful tools for classifying everyday objects in images, including subtypes of animals. However, pretrained CNNs are not designed to classify medical images, such as those encountered in veterinary medicine. Instead, clever approaches can be used to retrain CNNs in a supervised ML manner by subjecting them to new and labeled images. This concept, known as transfer learning, has been applied in a wide range of veterinary applications. Whereas transfer learning can be used to produce predictive models with high accuracy, it is difficult for a researcher to understand which specific image features the CNN uses to classify an object. Another approach is for the researcher to manually extract features from relevant regions of interest within the image, much like a digital biopsy. This feature extraction and analysis, or radiomics, offers an opportunity to link image features with pathophysiologic expressions for disease classification (Figure 2).

**Applications in Veterinary Medicine**

Despite the recent surge in AI publications in veterinary medicine, many new and exciting clinical and research directions are yet to be explored with AI. This article does not purport to provide a comprehensive review article for the topic; rather, it provides several examples of AI that illustrate its potential in veterinary medicine. There are also a number of excellent review papers that provide more detail on AI, ML approaches in human medicine, and data science. Artificial intelligence in epidemiology and quantitative decision-making tools in veterinary medicine and animal health are also described in greater detail elsewhere.

**Companion animal care and image-based AI**

A large percentage of published works on veterinary medicine and AI focus on medical image analysis in the companion animal setting. An early publication that used the phrase “machine learning” in veterinary medicine is from McEvoy and Amigo. They developed an image-based neural network to classify image features in pelvic radiographs. With the advent of CNNs, a number of AI publications have explored the use of CNNs for detecting diseases and abnormalities from medical images, ranging from canine CT and thoracic radiographs, canine MRI images, canine ultrasound images, and dairy cow teat images to optical coherence tomography in dogs. Deep learning has also been recently applied in autosegmentation (the automated contouring of normal tissues in CT scans) in radiotherapy planning of dogs. Radiomics analysis has been explored in feline ultrasound images and micro-CT images of proximal sesamoid bones in Thoroughbred racehorses.

**Companion animals: disease and abnormality detection, sensors**

Awaysheh et al applied NLP and ML on structured and nonstructured pathology reports of cats with a variety of GI diseases. They reported that ML models trained on unstructured reports achieved higher diagnostic accuracy when compared with models trained on structured reports, suggesting that information was lost when pathologists used structured reporting. The same group demonstrated success in creating a supervised ML model designed to predict the presence and type of feline abdominal disorders from complete blood counts and serum chemistry. Machine learning models...
have also been developed to predict seizures from intracranial electroencephalogram signals in dogs and atrial fibrillation disease progression from endocardial electrograms of dogs.

Digital wearable technologies are prevalent in society, and while they are convenient and economical, their large streams of data can pose challenges for efficient animal tracking and abnormality detection. Vehkaoja et al. developed an ML model to detect 7 specific behaviors from collar-based sensor data in 17 dogs. Chambers et al. recently developed an ML model using collar-mounted accelerometer readings from approximately 2,500 dogs. Their ML algorithm was able to classify a variety of behaviors from accelerometer data, such as drinking water, licking, petting, rubbing, scratching, and sniffing, with reasonably high sensitivity and specificity. In another study, Kasness et al. developed an ML model to identify activities of search-and-rescue dogs using data from similar wearable technologies. While the system was validated in only 2 discrete search-and-rescue events, an interesting feature of their research is the integration of data and real-time data analysis software by use of a cloud-based system.

Pathology, infectious diseases, and animal health

Pathogen detection from blood smears can be challenging when collected from exotic animals, particularly when commercial blood smear image analysis solutions are not available. One approach is to leverage CNNs: Kittichai et al. used a CNN to classify an avian malaria parasite from stained blood films with an accuracy of 97%.

Indeed, AI-based automated feature detection systems that can automatically analyze blood smear or pathology-stained slides offers tremendous potential efficiencies and quality improvements in veterinary medicine. When employed to tackle the challenge of low reproducibility and poor inter-rater agreement of mitotic counts in histology analysis, a CNN-based ML model was shown to, on average, provide more accurate mitotic scores than those obtained by veterinary pathologists. A similar CNN-based approach was used to develop a model for detecting aggregate reticulocytes in microscopic images of cat blood smears, and the ML model was able to achieve an accuracy of 98.7%. Pijnaker et al. developed 4 different ML models for canis parasitemia screening using data collected on a commercial hematology analyzer and were able to achieve a sensitivity and specificity of 100% and 95.7%, respectively. Nagamori et al. reported their experience with the VETSCAN IMAGYST system, which uses a CNN model to classify feline and canine fecal parasites. Their ML model sensitivity ranged from 65.7% to 100% and specificity from 97.6% to 100%.

Through their nationalized veterinary database of over 180 clinics, Hur et al. sought to elucidate antimicrobial use in Australia by using NLP to analyze raw clinical note data. While they could develop ML models with high accuracy, they found that within their immense database, only 40% of records indicated a rationale for prescribing antimicrobials. This study demonstrates both the importance of complete records in the veterinary health record systems and the need for good data when researchers have access to large data sets. In their recent review, Ezanno et al. provided a perspective on AI-based animal health and offered several recommendations for addressing data challenges such as data standardization, harmonization, and the adoption of findable, accessible, interoperable, and reusable (also referred to as FAIR) data principles to improve reliability and reproducibility.

The Potential of AI

Even with the recent surge in AI publications, there remains new and exciting opportunities for AI in veterinary medicine and animal health. In our companion paper, we discussed several considerations and challenges the veterinary community face when adopting AI technologies. Here, we focus on the possibilities and potential of AI in veterinary medicine with an emphasis on clinical disease prediction and precision medicine and translational research and clinical trials.

Clinical disease prediction and precision medicine

Research on the use of NLP and AI with medical images to detect, predict, and classify disease will continue to grow alongside improvements in ML methods. Various -omics profiling, such as proteomics, metabolomics, genomics, transcriptomics, and dosiomics, are now feasible and attainable in veterinary medicine. When various -omics data are combined, data sets become larger and more challenging to process. Analysis of multomics data will require sophisticated data reduction and feature selection techniques, but it has the potential to offer improved diagnostics and more effective patient-specific treatment strategies (Figure 3).

![Multomics approaches in ML](image)

*Figure 3—Multomics approaches in ML. All of these -omics methods characterize the state of the patient at a given time. These features can be combined to generate potentially more accurate diagnostic predictions or treatment evaluation models.*

Unauthenticated | Downloaded 04/03/22 07:15 AM UTC
Speech recognition tools can be deployed in clinical triaging of animals in the telemedicine environment. Natural language processing is currently being explored in AI-based medicine by means of “chatbots” in human medicine, where patients can interact with a virtual clinician. Veterinary chatbots present an interesting potential, where owners would interact with an AI system to triage companion animal health concerns while providing the conveniences of telemedicine.

Practice workload, translational research, and clinical trials
Veterinary practice workloads have been steadily increasing, and this can contribute to increased workplace stressors and adverse health effects for veterinarians. Decision support tools that leverage AI may help relieve workplace stressors. Today, speech recognition software is commonly integrated with human health records systems to enable efficient text entry into the patient record. Similar approaches can be easily adopted in veterinary medicine in the companion animal hospital setting.

Translational research opportunities also exist in veterinary medicine, particularly within the One Health paradigm of human and animal care. Translational research is based on multidisciplinary collaborations among laboratory and clinical researchers and different communities in pursuit of more effective treatments and best practices. Unsupervised ML methods can be used to identify patterns in enhancer-gene connections in humans, mice, and flies, and similar approaches could be explored in common diseases of humans and companion animals. Such pattern detections in different diseases and phenotypes may be particularly useful in zoonotic and infectious disease modeling, and AI can be leveraged for these purposes.

Conducting clinical trials in veterinary medicine has a number of challenges, including patient recruitment, funding support, rapid adoption of technologies and treatment regimens that preclude their implementation even more crucial prior to clinical use. A potential pragmatic approach in the veterinary setting is the latter, where models are trained and tuned by use of pragmatic approach in the veterinary setting is the latter, where models are trained and tuned by use of

Data
Training Data (1)

Validation Data (1)

Test Data (1)

Data (2)

Trained AI Model

Figure 4—Schematic of the domain-shift problem. Training and validation data from institution 1 is used to develop an AI model, and the final performance is characterized with the test data from institution 1. However, for a variety of factors, the trained AI model may perform poorly when subjected to data from institution 2.

Good data = practical models
Opportunities to leverage AI hinge on access to good data. As demonstrated in 2 separate veterinary ML studies, the quality of a veterinary AI model can be heavily influenced by the quality of training and validation data. Unlike in a typical veterinary health record system, human medical health record systems comply with national and international data transfer and storage standards, and this compliance (at least in principle) provides researchers with indexed and unambiguous data fields within the health record system. With access to and an understanding of the data in the health record system, researchers can develop and test ML models with vetted data. Larger veterinary institutions may have the resources to tackle challenges associated in collecting, curating, and analyzing data for ML purposes, while smaller institutions may not. Addressing these and other data challenges in veterinary medicine requires a good working knowledge of database management and techniques to manage disparate data types.

However, even with access to good data, there remains the perplexing domain- or data-shift problems experienced in human medicine, where the distribution of data in the environment in which the model is developed is measurably different than the environment in which the model is deployed (Figure 4). Domain shift can be addressed by either updating and generalizing the model such that it incorporates multi-institutional data during training or, alternatively, customizing the model for a specific institution. While the former approach would offer opportunities for smaller institutions to extract the knowledge gained from models trained in larger data sets, testing and clinical implementation of these models in those environments become even more crucial prior to clinical use. A potential pragmatic approach in the veterinary setting is the latter, where models are trained and tuned by use of local patient populations.

Working smarter
With the advent and evolution of public resources and computing platforms, more flexible and simpler programming environments, and inexpensive computing power, access to AI technologies for the non-specialist is improving. This increase in access permits academics, veterinarians, and future veterinarians to
experiment and test AI technologies (one can simply visit the Colaboratory website\textsuperscript{15} to see the immense number of AI data sets and models). A module on the fundamentals of AI in veterinary medicine should become the norm in veterinary curricula, not only to empower future practitioners on the utility of AI in their everyday practice but also to provide them with the expertise to understand the limitations of AI and adopt best practices when using it. Professional life-long learning opportunities for AI and veterinary medicine will be instrumental in ensuring that members of modern-day veterinary professionals have the essential skills for enabling the safe implementation and interpretation of AI-based technologies. As seems to be the case in human medicine, AI will not be a panacea for veterinary medicine nor will it result in a collapsing workforce. Instead, AI-based decision support tools permit an opportunity to work smarter.

Working together

A key consideration in the successful adoption of AI technologies in veterinary medicine will be their safe introduction into clinical practice. Wilson et al.\textsuperscript{16} outlined 10 practical tips to guide success when deploying AI in human health care, and these are equally relevant in veterinary medicine. They propose a 4-phase approach to follow when introducing an AI-based solution in the hospital: (a) conceptualization of the AI project, (b) data acquisition and preparation, (c) AI application, and (d) translation. Central to this approach is the establishment of a collaborative team of experts to guide the implementation and steward the use of AI in the clinic. Such a team should include veterinary specialists and practitioners, data scientists, project managers, and software engineers. Also important is to identify those in leadership positions who can help remove or lower barriers to success. Finally, one must educate and respect the needs of end users—the animals and their owners—who benefit from this technology.

Conclusions

The recent surge in AI publications on veterinary medicine suggests new advances and applications of AI in veterinary medicine are forthcoming. This will pose challenges—and opportunities—for the veterinary community to understanding, interpreting, and adopting this powerful and evolving technology. As previously suggested, AI is not likely to become a panacea for veterinary medicine, nor will it replace the expertise of the veterinarian. More likely, AI promises opportunities for the veterinary community to work more efficiently and to provide new insights into managing and treating disorders, which will hopefully translate to improvements in the quality of life for animals and those who care for them.

Acknowledgments

No external funding was used in the preparation of this manuscript. The authors declare that there were no conflicts of interest.

The authors would like to acknowledge the American College of Veterinary Radiology/European College of Veterinary Diagnostic Imaging Artificial Intelligence Education and Development Committee for its role in discussions that have played a role in influencing the important features highlighted in this paper. Kim P. Nayyer is kindly acknowledged for her assistance with editing.

References


27. Boissady E, Comble A de L, Zhu X, Hespel AM. Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinary radiologists in conjunction with the artificial intelligence. Vet Radiol Ultrasound. 2020;61(6):619–627. doi:10.1111/vru.12912


