Animal Behaviour 124 (2017) 203–220

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Review

Applications of machine learning in animal behaviour studies

John Joseph Valletta [a, *], Colin Torney [a], Michael Kings [b], Alex Thornton [b], Joah Madden [c]

[a] *Centre for Mathematics and the Environment, University of Exeter, Penryn Campus, Penryn, U.K.*
[b] *Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn, U.K.*
[c] *Centre for Research in Animal Behaviour, University of Exeter, Exeter, U.K.*

ARTICLE INFO

Article history:
Received 15 August 2016
Initial acceptance 8 September 2016
Final acceptance 15 November 2016
Available online 23 January 2017
MS. number: 16-00718

Keywords:
animal behaviour data
classification
clustering
dimensionality reduction
machine learning
predictive modelling
random forests
social networks
supervised learning
unsupervised learning

In many areas of animal behaviour research, improvements in our ability to collect large and detailed data sets are outstripping our ability to analyse them. These diverse, complex and often high-dimensional data sets exhibit nonlinear dependencies and unknown interactions across multiple variables, and may fail to conform to the assumptions of many classical statistical methods. The field of machine learning provides methodologies that are ideally suited to the task of extracting knowledge from these data. In this review, we aim to introduce animal behaviourists unfamiliar with machine learning (ML) to the promise of these techniques for the analysis of complex behavioural data. We start by describing the rationale behind ML and review a number of animal behaviour studies where ML has been successfully deployed. The ML framework is then introduced by presenting several unsupervised and supervised learning methods. Following this overview, we illustrate key ML approaches by developing data analytical pipelines for three different case studies that exemplify the types of behavioural and ecological questions ML can address. The first uses a large number of spectral and morphological characteristics that describe the appearance of pheasant, *Phasianus colchicus*, eggs to assign them to putative clutches. The second takes a continuous data stream of feeder visits from PIT (passive integrated transponder)-tagged jackdaws, *Corvus monedula*, and extracts foraging events from it, which permits the construction of social networks. Our final example uses aerial images to train a classifier that detects the presence of wildebeest, *Connochaetes taurinus*, to count individuals in a population. With the advent of cheaper sensing and tracking technologies an unprecedented amount of data on animal behaviour is becoming available. We believe that ML will play a central role in translating these data into scientific knowledge and become a useful addition to the animal behaviourist's analytical toolkit.

Recent technological advances mean that large data sets can be collected on the movement (Hussey et al., 2015; Kays, Crofoot, Jetz, & Wikelski, 2015; Tomkiewicz, Fuller, Kie, & Bates, 2010), fine-scale motion (Brown, Kays, Wikelski, Wilson, & Klimley, 2013), social interactions (Krause et al., 2013), vocalizations (Blumstein et al., 2011) and physiological responses (Kramer & Kinter, 2003) of individual animals. Conversely, the logistical difficulties of collecting replicated data, especially from wild populations, mean that sample sizes are small, even though data on each individual may be rich, with many hundreds (or even thousands) of factors to consider. These complex data sets, generated from different sources, such as images and audio recordings, may fail to conform to assumptions of many classical statistical models (e.g. homoscedasticity and a Gaussian error structure). Moreover, unknown nonlinear dependencies and interactions across multiple variables make it unclear what type of functional relationship one should use to describe such data mathematically. Animal behaviour researchers are thus in a position where automatically collecting detailed data sets is becoming commonplace, but extracting knowledge from them is a daunting task, mainly due to the lack of accessible analytical tools.

Machine learning (ML) offers complementary data modelling techniques to those in classical statistics. In animal behaviour, ML approaches can address otherwise intractable tasks, such as classifying species, individuals, vocalizations or behaviours within complex data sets. This allows us to answer important questions across a range of topics, including movement ecology, social structure, collective behaviour, communication and welfare. ML encompasses a suite of methodologies that learn patterns in the data amenable for prediction. A machine (an algorithm/model) improves its performance (predictive accuracy) in achieving a task

* Correspondence: J. J. Valletta, Centre for Mathematics and the Environment, University of Exeter, Penryn Campus, Penryn TR10 9FE, U.K.
E-mail address: jj.valletta@exeter.ac.uk (J. J. Valletta).

http://dx.doi.org/10.1016/j.anbehav.2016.12.005

(e.g. classifying content of an image) from experience (data). The objective is for the predictive model to generalize well, that is, to make accurate predictions on previously unseen data. For instance, when Facebook users upload their photos, the 'auto-tagging' ML algorithm extracts facial features and suggests names of friends in that photo. Facebook's predictive model generalizes from manually tagged photos (known as the training data set). It is impossible to 'show' a machine all the images of an individual (e.g. different facial expressions); instead the model uses the extracted features to learn patterns that best discriminate one individual from another. The generalization error or predictive performance is a measure of how many previously unseen images (known as the testing data set) the algorithm tags correctly.

Both statistical modelling and ML seek to build a mathematical description, a model, of the data and the underlying mechanism it represents; thus inevitably there is substantial overlap between the two (Breiman, 2001b; Friedman, 2001; Zoubin Ghahramani, 2015). However, historically they differ in their rationale as follows. Statistical models start with an assumption about the underlying data distribution (e.g. Gaussian, Poisson). The focus is on inference; estimating the parameters of the statistical model that most likely gave rise to the observed data, and providing uncertainty bounds for these estimates. For ML, the focus is typically on prediction; without necessarily assuming a functional distribution for the data, a model that achieves optimal predictive performance is identified. It is this hypothesis-free approach that makes ML an attractive choice for dealing with complex data sets. While in traditional statistical modelling a hypothesis (model) is put forward and is then accepted/rejected depending on how consistent it is with the measured observations, ML methods learn this hypothesis directly from the training data set.

ML can tackle a wide range of tasks, including classifying observations into predefined sets (Kabra, Robie, Rivera-Alba, Branson, & Branson, 2013), clustering data into groups that share an underlying process (Zhang, O'Reilly, Perry, Taylor, & Dennis, 2015) and regressing an outcome of interest against multiple factors and elucidating their contributory effect (Chesler, Wilson, Lariviere, Rodriguez-Zas, & Mogil, 2002; Piles et al., 2013). Owing to its versatility, ML has been applied across a broad set of domains in animal behaviour to ask and subsequently answer biologically meaningful questions. Next, we highlight some facets of animal behaviour where ML has already been deployed.

GPS, accelerometer and/or video data are routinely used to monitor movement patterns of individuals. Three-dimensional accelerometer loggers can generate over a million data points per hour of recording (at a sample rate of 100 Hz). ML is used to automate the classification of behaviours/activities (Kabra et al., 2013) and tracking movement trajectories (Dell et al., 2014). This knowledge can then be used to infer individual decision rules in collective motion (Katz, Tunstrom, Ioannou, Huepe, & Couzin, 2011; Nagy, Kos, Biro, & Vicsek, 2010) and to compute activity budgets for individuals without the need for continuous human observation or time-consuming video analysis. This is especially suitable for organisms that are hard to observe directly, such as nocturnal (badgers, *Meles meles*: McClune et al., 2014), pelagic (little penguins, *Eudyptula minor*: Carroll, Slip, Jonsen, & Harcourt, 2014) and aquatic species (great sculpins, *Myoxocephalus polyacanthocephallus*: Broell et al., 2013; whale sharks, *Rhincodon typus*: Gleiss, Wright, Liebsch, Wilson, & Norman, 2013), or those that are hard to follow continuously owing to their speed or covertness (e.g. cheetahs, *Acinonyx jubatus*: Grünewälder et al., 2012; pumas, *Puma concolor*: Wang et al., 2015).

Another context in which ML has been successfully employed is in vocalization studies. Vocalizations can be recorded remotely permitting assessments of population size and species composition, individual behavioural and inter/intraspecific interactions (Blumstein et al., 2011). A typical recording, made using pulse code modulation (PCM) at 24-bit and 48 Hz sampling, produces over half a gigabyte of data per hour. Consequently, inspection of these data and analysis of sound recordings can be time consuming and highly subjective when conducted by visual inspection of sonograms. Instead, ML has been applied to classify and count particular elements or syllables (Acevedo, Corrada-Bravo, Corrada-Bravo, Villanueva-Rivera, & Aide, 2009). Early work used ML techniques to adjudicate similarity between calls based on sets of such elements (Tchernichovski, Nottebohm, Ho, Pesaran, & Mitra, 2000). These approaches can also discern differences in calls. Classification of calls from different species and subspecies is robust (Fagerlund, 2007; Kershenbaum et al., 2016) and permits assessment of community structure (e.g. frogs: Taylor, Watson, Grigg, & McCallum, 1996; birds: Brandes, 2008). Finer scale discriminations are possible at both the individual level (Cheng, Xie, Lin, & Ji, 2012) and the bird song elements level (Ranjard & Ross, 2008).

The assessment of animal welfare and the emotional states that may reveal it can be highly subjective, and poor welfare is often only indicated by multiple interacting factors (Broom & Johnson, 1993). ML can assist in monitoring such behaviours by matching the human assessment in terms of treatment effects on laboratory mice, *Mus musculus* (Roughan, Wright-Williams, & Flecknell, 2009). Such methods have been extended to provide a diagnostic tool for psychopharmacological drugs based on mouse open-field behaviour (Kafkafi, Yekutieli, & Elmer, 2009). ML was used in a comparative assessment of welfare across multiple laboratory populations of mice (Chesler et al., 2002) permitting a wide range of potential explanatory factors, each with diverse distribution, to be considered simultaneously as well as the interactions between them. A potential novel use of ML would be to detect emotional state in animals based on facial expression, body posture or vocalization. Such techniques have already been used in humans looking at facial (Michel & El Kaliouby, 2003), physiological (Shi et al., 2010), vocal (Shami & Verhelst, 2007) and gestural (Castellano, Villalba, & Camurri, 2007) cues of emotions. ML also permits integration of multiple sets of these cues to further enhance emotion detection (Caridakis et al., 2007).

Elucidating the underlying social network structure of individuals within social groups can help address important ecological and evolutionary questions (Krause, James, Franks, & Croft, 2015). Passive integrated transponder (PIT) tags and proximity loggers now permit automated collection of large volumes of social interaction data containing both spatial and temporal elements (Krause, Wilson, & Croft, 2011). Translating such data into biologically realistic patterns of association is not trivial, and may depend on subjective decisions by researchers, especially when the instances of association are ambiguous. Co-occurrences in time could be determined by ML clustering methods with individuals in the same foraging event (cluster) considered to have a social affiliation (Psorakis et al., 2015). Such methods appear to be robust and capture real-life pair bonds well (Psorakis, Roberts, Rezek, & Sheldon, 2012). A second facet of association patterns that benefits from application of ML techniques is determining to which social grouping an individual belongs within a network. In many cases, group membership is ambiguous with individuals having weak or sporadic membership to multiple clusters of other individuals. A subjective decision of membership could be arrived at, with such weak affiliations being discounted. Alternatively, ML techniques could be deployed to account for such 'fuzzy overlapping' (Gregory, 2011), and individuals can have their relative membership of each group determined.

It is clear that ML can address different objectives in numerous distinct fields of animal behaviour and is thus becoming a staple