PREDICTING HORSE FEARFULNESS APPLYING SUPERVISED MACHINE LEARNING METHODS

Oleksandra Topal
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia
oleksandra.topal@ijs.si

Inna Novalija
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia
inna.koval@ijs.si

Elena Gobbo
Biotechnical faculty
University of Ljubljana
Jamnikarjeva 101 Ljubljana, Slovenia
elena.gobbo@bf.uni-lj.si

Manja Zupan Šemrov
Biotechnical faculty
University of Ljubljana
Jamnikarjeva 101, Ljubljana, Slovenia
manja.zupansemrov@bf.uni-lj.si

Dunja Mladenić
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia
dunja.mladenic@ijs.si

ABSTRACT

In this article, we present the first results of a study on the personality traits of Lipizzan horses focusing on their fearfulness. Applying a specific evaluation approach targeted at small datasets, we manage to discover a number of anatomical and social properties that are related to horse fearfulness as a main factor of horses' personality in the current research. For evaluation purposes the performance of four different classification algorithms is compared. Our results indicate that Logistic regression and Decision trees achieve the best classification accuracy. Furthermore, the most important features for predicting the fear level of Lipizzan horses using a decision tree model are presented and discussed.

KEYWORDS

Machine learning, classification problem, personality traits, Lipizzan horses.

1. INTRODUCTION

In the modern world, artificial intelligence provides powerful tools for solving many issues in various fields of research. The problems involving clustering, regression, and classification are the most commonly addressed problems in different types of biological studies. One of the actual topics of biological research where we can use artificial intelligence algorithms is the study of the animal personality.

In our work we are studying the personality traits of horses of the Lipizzan breed. Personality assessment can be used to select suitable training and weaning methods, choose or breed horses for police or therapeutic work, investigate underlying reasons for development of behavioral problems or assess how an unknown horse might react to a new or aversive situation or stimuli. According to a research study on animal behavior [1], it is possible to improve performance and horse welfare by identifying the right match between the horse's temperament, its rider's personality, housing conditions, management and by choosing the appropriate activity for an individual horse.

Number of experiments demonstrate that anatomical features may be associated with personality traits and behaviour in animals, mainly due to domestication and selection process that affected animals' morphology and personality. We can find a confirmation of this in Belyaev's domestication and selection experiment on foxes [2], also there is research on a number of species such as pigs and cattle [3], dogs [4], and horses [5]. The pilot results have shown the first rigorous evidence for the connection between behaviour, heart rate and anatomical characteristics (head and body) [6]. We therefore assume that various properties, such as anatomical and biomechanical as well as social environmental measurements, give us valuable objective insights to predict personality traits of Lippizan horses with an emphasis on fearfulness. We believe that this improved knowledge will help us understand the horse-human relationship, the complexity of animal personality in general and in relation to humans, as humans and horses share many emotional processes [7].

The main contribution of this research is assessment of the importance of different properties for predicting fearfulness of a horse as indicated by different traditional machine learning algorithms.

2. RELATED WORK

A number of animal studies researchers have tackled the topic of animal personality. Animal personality could be defined as temporally stable inter-individual patterns of affect, cognition, and behavior [8]. Gobbo and Zupan [9] in their study on dogs state that analysis of animal personality traits is closely linked to the safe human-animal interaction and animal's everyday behavior. Moreover, Buckley et al. [10] reported that personality of a horse should be considered as an important attribute and a key issue in horse health and performance. The most important personality trait in relation to human-horse relationship is suggested to be fearfulness [11].

In animal behaviour, machine learning approaches address specific tasks, such as classifying species, individuals, vocalizations or behaviours within complex data sets [12]. Machine learning has been used for clustering observations into groups [13] and for classification of animal related data [14].

In our work, we apply data mining and machine learning on the Lipizzan horse's dataset with broad anatomic, social, and biomechanical characteristics. In addition, the dataset used in the current research contains a small number of data points and requires using evaluation techniques for small datasets.

Similarly, to other related work approaches, we apply traditional machine learning classification methods for assessing a horse's personality and understanding which horse properties are the most important when predicting the fearfulness of a horse. Specifically, in our research, we investigate how feature selection method can influence the classification results for fear level prediction in horses.

3. PROBLEM DEFINITION

3.1 Data sources

For our study, we use a unique dataset that we have created and which contains anatomical measurements, biomechanics characteristics, housing conditions and fear score of Lipizzan horses. Based on our experience as experts in animal studies, we have collected and organized the data in four parts.

The first part contains age, gender, front, left and right (both sides need to be measured, because they are not identical [15, 16]) anatomical measurements of the horse head (FH) and body (FB). The second part contains the results of a study on the biomechanics of the Lipizzan horses. Biomechanical data were collected twice for two types of horse gaits, walking and trotting, so the table contains some redundant data. We have converted the table, so that the trot and walk data are separated by traits for each horse and can be used for modeling. The third part lists the conditions of keeping horses, such as the availability of pastures, the openness of stalls, the number of stalls, as well as equestrian activities, training and work of horses. The fourth part contains the results of fear test battery performed on each horse.

In our study, the explorative hypothesis is that anatomical-biomechanical-social properties of a horse may act as good indicators of fearfulness. We have many features describing different parameters of horses on the one side, and we have a horse fearfulness score on the other side, so we can use supervised machine learning methods to predict the horse's fearfulness levels.

3.2 Labeling data for the classification task

To label our dataset, we have had to transform a very complex fear rating table. During the experiment, two repetitions of each of the four fear tests of the individual horse have been carried out. We have compared the sum of the four scores of the first repetition (each score per individual fear test and a horse) with the sum of the four fear scores of the second repetition, and it turned out that the horses habituated to stimuli between the two repetitions (see Figure 1).

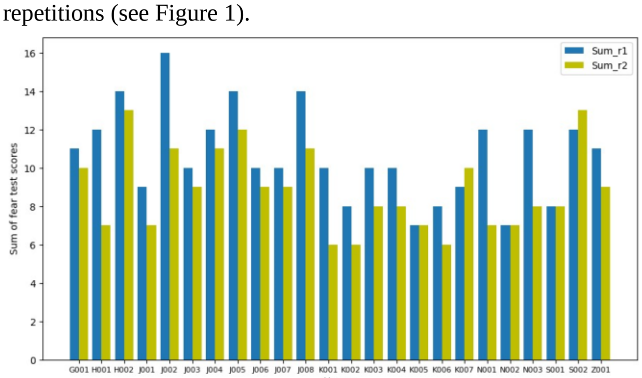

Figure 1 Comparison graph between two repetitions of fear tests.

We have made the decision to take the maximum value of the two sums in order to eliminate the habituation element. The task of classification assumes that the data is divided into classes, that's why we have found the average value of fear score, which was 10.75, and labeled the fearfulness variable with binary values as follows. If a horse has an above-average fear rating, then it corresponds to a value of 1 (class 1) - a fearful horse, if lower, then 0 (class 0) - a fearless horse. In this way we obtained a fairly balanced dataset, in which there are 13 fearful horses and 11 fearless horses (see Figure 2).

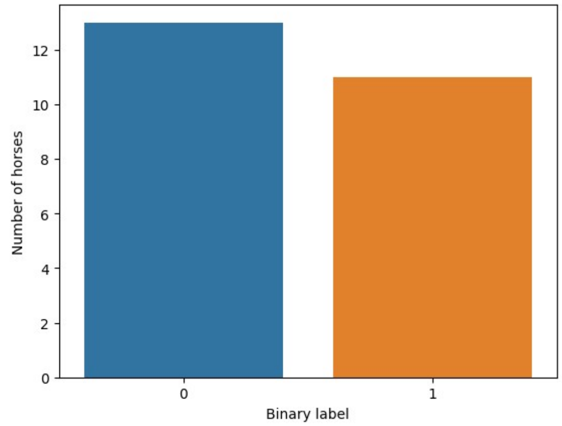

Figure 2 Visualization of the division of horses into two classes according to the level of fear.

4. METHODOLOGY

4.1 Data preprocessing.

Like almost all biological data, this dataset is very small, with only 24 instances, but more than 120 different features. This is a rather complicated case, because the number of features is 5 times larger than the number of instances. We conducted a correlation analysis using the Spearman coefficient which will allow us to reduce the dimensionality of the data. Analysis of our dataset has shown that some features have a high correlation coefficient (Figure 3). If correlation coefficient is more than 0.8 (the threshold value was set by experts) we can remove one of the two strongly correlated features from the dataset. Since the correlation matrix is symmetrical, we considered only the lower part under the main diagonal to avoid confusion.

	FB37L	FB37R	FB38	FB39	FB40
FB36L	0.478659	0.39878	0.47755	0.456626	0.539769
FB36R	0.537993	0.616558	0.501635	0.442362	0.455774
FB37L	1	0.932883	0.266347	0.114211	0.177381
FB37R	0.932883	1	0.306652	0.197601	0.189708
FB38	0.266347	0.306652	1	0.885471	0.827045
FB39	0.114211	0.197601	0.885471	1	0.891603
FB40	0.177381	0.189708	0.827045	0.891603	1

Figure 3 An illustrative fragment of the correlation matrix.

4.2 Evaluation method

For very small datasets, as in our study, we should find a suitable approach to evaluate machine learning models. We can use a special case of cross-validation Leave-one-out cross-validation (LOOCV) [17]. LOOCV is a type of cross-validation approach in which each observation is considered as the test set and the rest (N-1) observations are considered as the training set. In LOOCV, fitting of the model is done and predicting using one observation test set. Furthermore, repeating this N times, so each observation is taken once in the test set. This is a special case of K-fold cross-validation in which the number of folds is the same as the number of observations (K = N).

4.3 Classification methods

There are many machine learning algorithms suitable for solving the classification problem. We decided to take several different algorithms starting with Logistic Regression and Support Vector Machine as a simple model [18], Decision Trees and Random Forests.

For the completeness of the experiment, we have trained all the algorithms with the different sets of features (see follow bulleted list). The main results are presented in Table 1. The rows of Table 1 present different algorithms used, while the columns reflect feature selection methods:

- AllFeatures (120 features): removal of correlated features is not performed
- Removed LeftCorr (89 features): anatomical measurements from the left side of the horse head or body that correlate to the correspondent right side measurements are removed
- Remove RightCorr (89 features): anatomical measurements from the right side of the horse head or body that correlate to the correspondent left side measurements are removed
- Removed LeftCorr+ (85 features): anatomical measurements from the left side of the horse that correlate to the correspondent right side measurements are removed + anatomical measurements from the right side of the horse that correlate to other left side measurements are removed
- Remove RightCorr+ (85 features): anatomical measurements from the right side of the horse that correlate to the correspondent left side measurements are removed + anatomical measurements from the left side of the horse that correlate to other right side measurements are removed

Table 1 The accuracy of prediction of the horses' fear level of the different algorithms with different sets of features.

	AllFeatures	Removed LeftCorr	Removed RightCorr	Removed LeftCorr +	Removed RightCorr +
Logistic Regression	0.83	0.83	0.83	0.83	0.83
SVM	0.63	0.63	0.71	0.63	0.71
Decision Trees	0.75	0.75	0.79	0.71	0.83
Random Forests	0.67	0.67	0.71	0.63	0.67

As shown in Table 1, the best result has been obtained by Logistic Regression and Decision Trees.

If we look at the Logistic Regression coefficients, we find out that only one feature from 120 was chosen as significant and it is "Number of boxes" that means how many boxes were in the stable where the horse was housed. The number of horses housed in the same stable represents the horse's social environment, which may really affect its fearfulness.

In comparison to the other tested methods, Support Vector Machine and Random Forests show the lowest classification accuracy.

Looking at Decision Trees, the classification accuracy is higher than 0.7 for all sets of features. We can notice the difference in performance based on anatomical features. Removing the right correlated features gave better result than removing the left correlated features. Left measurements appear to be more significant for prediction in this model. We obtained the highest accuracy with Decision Trees (0.83) when we removed right correlated features + (**Removed RightCorr+).**

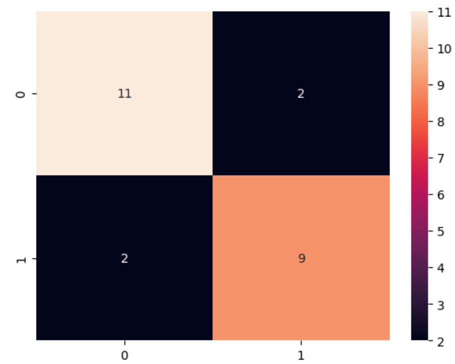

Figure 4 Confusion matrix by Decision Trees.

Figure 4 presents for Fearful (class 0) and Fearless (class 1) classes confusion matrix by Decision Trees.

In order to assess the learning outcomes of all models, we used LOOCV algorithm. We have noticed that the models during training chose different features as important in each validation step. In the following Table 2 we can see the most important features (see Figure 6 for more details) for the Decision Trees model and how many times they were chosen during the entire experiment (24 steps).

Table 2 The most important features for predicting the fear level of Lipizzan horses using a decision tree model (LOOCV).

Feature name	Numbers of times
Number of boxes	24
FB10L	23
FH03	21
FH04	18

Once we evaluated the decision tree model using the LOOCV algorithm and understood its performance, we were able to train the model on the **full set** without splitting it into a training and test set to obtain the most important features affecting the target variable (Figure 5).

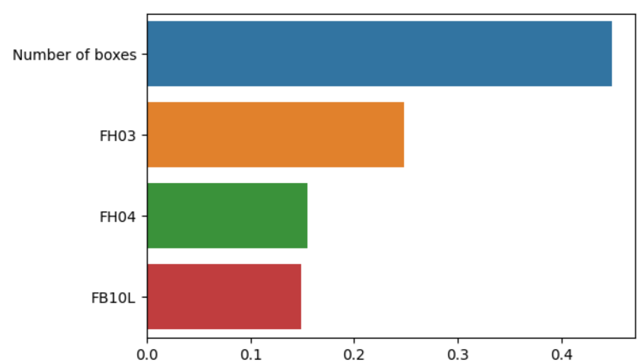

Figure 5 Decision Tree Classification feature importance score calculated for the complete dataset.

In our research, based on a small data sample of Lipizzan horses, we have been able to find out that social (Number of boxes) and anatomical (FH03, FH04, FB10L) features influence the fear score. We marked with the red lines the most important features on the Figure 6.

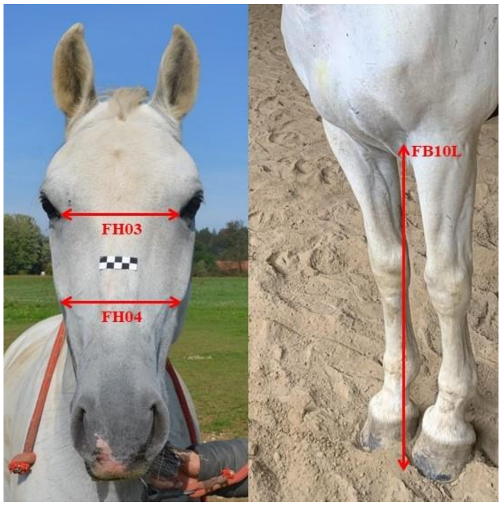

Figure 6 The most important measurements which can impact fear level of Lipizzan horses.

Figure 7 presents the Decision Tree obtained by the training the model on all available examples. In our study we have used the criterion Gini Impurity to help to choose the optimal split of the decision tree into branches.

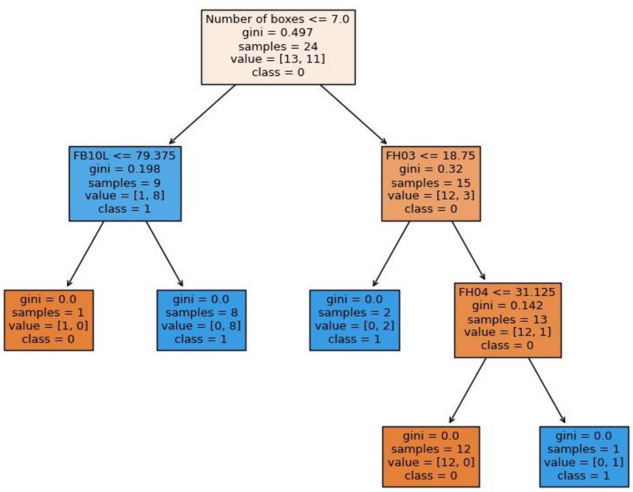

Figure 7 Decision Tree trained on all the examples

5. CONCLUSION AND FUTURE WORK

In this article, we have demonstrated some approaches to assessing and predicting the level of fear in Lipizzan horses. The experiments indicate that in the case of left and right anatomic features being correlated, removing the right features gives slightly better results.

We have found that social and anatomical features can explain the fearfulness level as a factor of horses' personality.

The future work will include the research with extended data set as well as exploring additional relevant features.

6. ACKNOWLEDGMENTS

This document is the result of the research project funded by the ARRS (J7-3154).

7. REFERENCES

[1] Hausberger M. et al. (2008) Applied Animal Behaviour Science. 109: 1–24.

[2] Trut, L. N. Early Canid Domestication: The Farm-Fox Experiment: Foxes bred for tamability in a 40-year experiment exhibit remarkable transformations that suggest an interplay between behavioral genetics and development. Am. Sci. 87, 160–169.

[3] Grandin, T. & Deesing, M. J. Genetics and the Behavior of Domestic Animals (2nd ed.) 488 p. (London: Academic Press, 2014).

[4] McGreevy, P. D. et al. Dog behavior co-varies with height, bodyweight and skull shape. PLoS ONE 8(12), e80529.

[5] Sereda, N. H., Kellogg, T., Hoagland, T. & Nadeau, J. Association between whorls and personality in horses. J. Equine Vet. Sci. 35, 428.

[6] Debeljak N, Košmerlj A, Altimiras J, Šemrov MZ. Relationship between anatomical characteristics and personality traits in Lipizzan horses. Scientific Reports. 2022 Jul 23;12(1):12618.

[7] Wathan J, Burrows AM, Waller BM, McComb K. EquiFACS: The equine facial action coding system. PLoS one. 2015 Aug 5;10(8):e0131738.

[8] Gosling, S.D. Personality in non-human animals. Soc. Personal. Psychol. Compass. 2008, 2, 985–1001.

[9] Gobbo, E. and Zupan, M., 2020. Dogs' sociability, owners' neuroticism and attachment style to pets as predictors of dog aggression. Animals, 10(2), p.315..

[10] Buckley, P., Dunn, T. and More, S.J., 2004. Owners' perceptions of the health and performance of Pony Club horses in Australia. Preventive veterinary medicine, 63(1-2), pp.121-133.

[11] McGreevy, P., & McLean, A. (2010). Equitation Science. Wiley-Blackwell, Chichester, West Sussex, UK.

[12] Valletta J.J, Torney C., Kings M., Thornton A., Madden J. (2017). Applications of machine learning in animal behaviour studies. Animal Behaviour. Volume 124: 203-220.

[13] Zhang J., O'Reilly K.M., Perry G.L.W., Taylor G.A., Dennis T.E. (2015). Extending the functionality of behavioural change-point analysis with k-means clustering: A case study with the little penguin (Eudyptula minor). PLoS One, 10 (4):e0122811.

[14] Kabra M., Robie A., Rivera-Alba M., Branson S., Branson K. (2013). JAABA: Interactive machine learning for automatic annotation of animal behavior. Nature Methods, 10 (1): 64-67.

[15] Wiggers N, Nauwelaerts SLP, Hobbs SJ, Bool S, Wolschrijn CF, et al. (2015) Functional Locomotor Consequences of Uneven Forefeet for Trot Symmetry in Individual Riding Horses. PLOS ONE 10(2): e0114836.

[16] Halsberghe, B.T., Gordon-Ross, P. and Peterson, R. (2017), Whole body vibration affects the cross-sectional area and symmetry of the m. multifidus of the thoracolumbar spine in the horse. Equine Vet Educ, 29: 493-499.

[17] Wong TT. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern recognition. 2015 Sep 1;48(9):2839-46.

[18] Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology. 2022 Jan;23(1):40-55.