

# PREDICTING HORSE FEARFULNESS APPLYING SUPERVISED MACHINE LEARNING METHODS

Oleksandra Topal, Inna Novalija, Elena Gobbo,  
Manja Zupan Šemrov, Dunja Mladenić

This document is the results of the research project funded  
by the Slovenian Research Agency (ARRS).

# Return of mares and foals from the pastures



© Oliwia Chmielewska, 2015

[https://ich.unesco.org/en/7b-representative-list-01281?select\\_country=00203#table\\_cand](https://ich.unesco.org/en/7b-representative-list-01281?select_country=00203#table_cand)



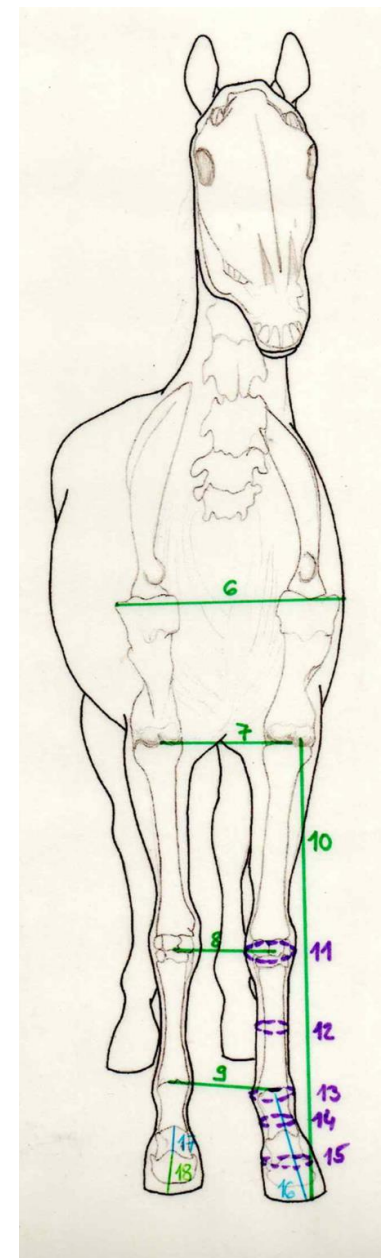
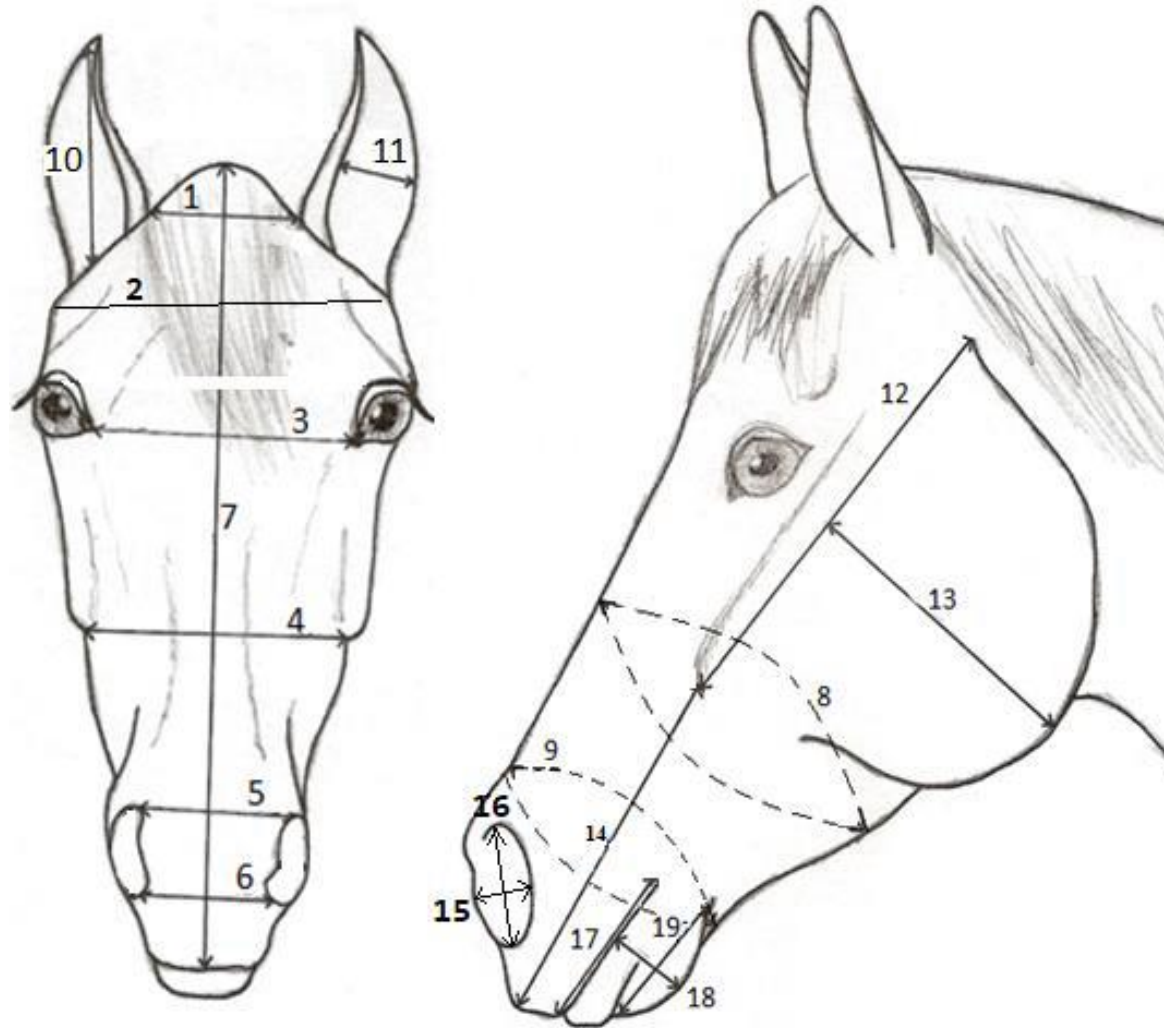
**The main goal** of the project is to identify the connection between the physiological and anatomical characteristics of Lipizzan horses and such a personality trait as fearfulness.

Our partners hypothesize that fear responses of individual horses are associated with specific body and head measures as well as with heart rate. An important result from their previous study shows that **a horse with a longer chest and a wider part between the nostrils is less fearful** while a horse with a larger inferior angle of the *right* nostril and outer edge of the *left* ear is less able to learn.

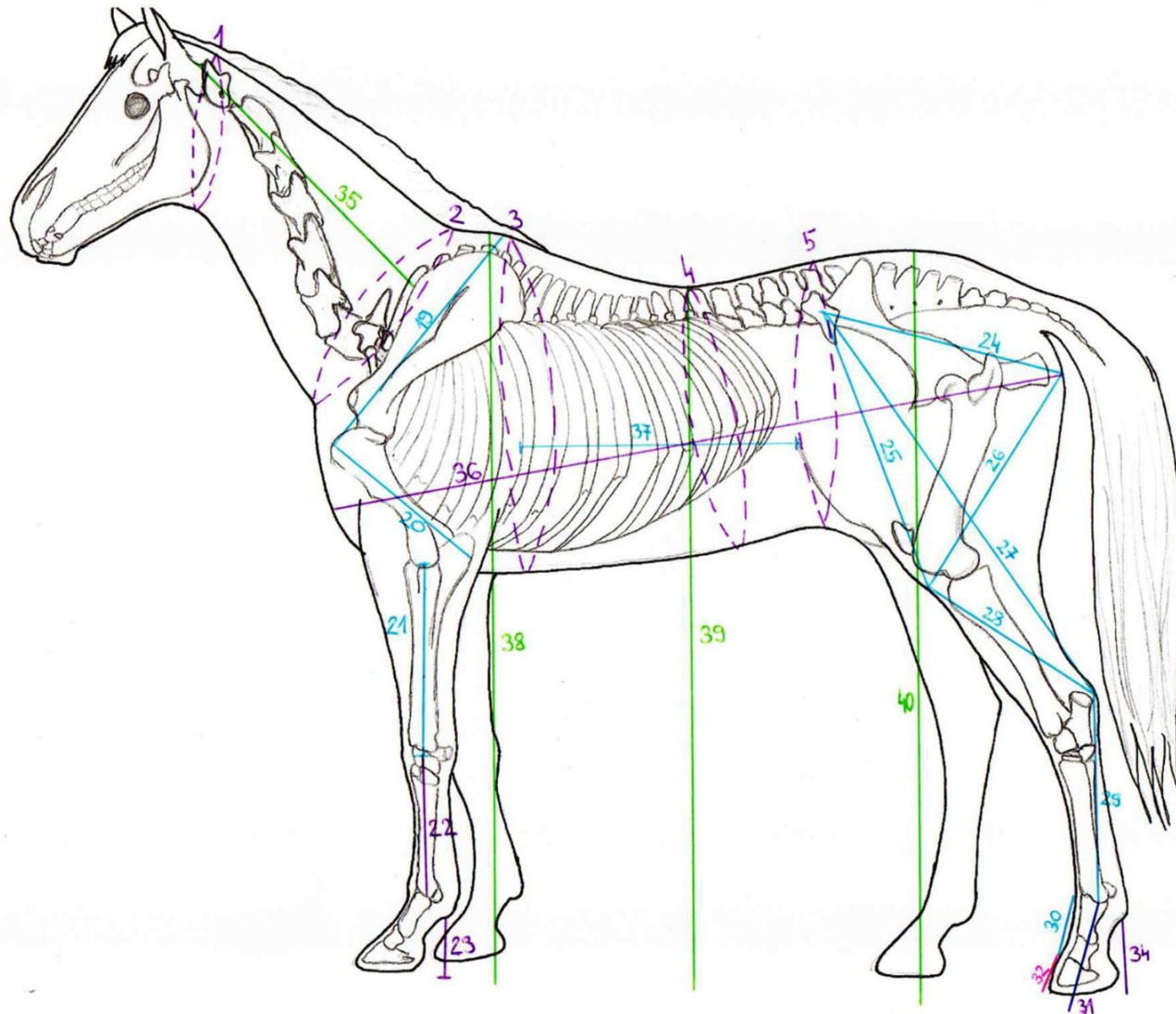
### Research tasks:

1. *Data analysis and preprocessing as well as feature engineering.*
2. *Statement of the problem and choice of solution methods for the current data set.*
3. *Solving the classification problem with ML models which help to predict the fear level of horses by their anatomical measurements, kinematic variables, and conditions of care and work.*
4. *Evaluation and visualization of the results of the models, selection the most important features for the fear score predicting of the Lipizzan horses and interpretation of research results.*

# Markers of head and body measurements of Lipizzan horses.



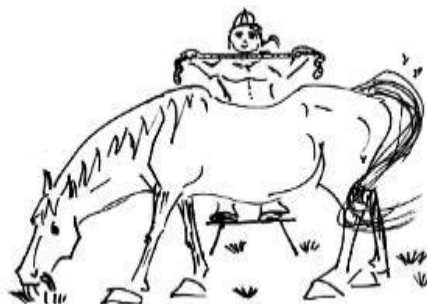
## Markers of body measurements of Lipizzan horses.



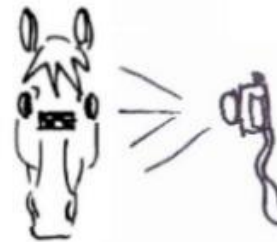
## Some features of anatomical measurements of Lipizzan horses (initial data).

ID	Age	Sex	Temperature	FH01	FH02	FH12L	FH12R	FH13L	FH13R	FB02	FB03	FB15L	FB15R	FB39	FB40
G001	26	M	37.3	15	20	24.75	24.75	22	22	113.5	182	38	39.5	144.5	153.25
K001	18	M	37.2	16	20.75	26.5	26.5	22	22	112.25	187.5	39	38.25	149.5	155
K002	17	G	36.8	16	21.25	22.5	22.25	21	20.75	112.25	183.5	41.75	42	146.5	153
K003	17	G	37.2	14.5	21.5	21.5	21.75	21.25	21.75	115	189.75	41.5	41.5	150.75	154.5
K004	13	G	37.1	14	25.25	24.25	23.75	21.25	21.75	115.5	186	39.75	40	152	159
K005	11	G	36.9	14.25	21	22.5	23	23.25	23	121	193	40.5	40.25	150	159
K006	9	M	37.2	13	23.5	21	21.75	19.5	20	113	178.5	38.25	38	142.75	151.5
K007	10	M	37.2	14	22.5	23.25	23.5	22	21.75	111.5	190	42	41.5	153.5	162.5
Z001	18	G	37.1	14	23	22.5	22.5	20.5	20	114	184.5	42.5	42.75	142.5	148
J001	13	G	37.1	13	23	22.5	22.5	20.25	20.25	110	175.25	37.75	38	143.5	150
J002	14	M	37.5	13.75	23	21.75	21	22	21.5	116.75	183	40	39.5	146.5	154
J003	21	M	37	12	21.75	20	20.25	18.5	18.5	113	174.75	36	35.5	141.5	150
J004	16	G	36.9	12.5	23.5	21	20.75	20	20	110.25	174.75	39	39	145.5	153.5
J005	18	G	37	14	23	21	20.75	19.5	19	109.5	175.5	39.75	40	147	152.5
J006	8	G	37.1	14	24.75	21.5	22	21	20.5	107.75	171.75	36.5	36.25	143.5	147.5
J007	15	M	37.1	14	24.75	21.25	21.75	20	20	119	177.25	35.75	35.5	141	151
J008	7	G	37.4	14.75	23	23.5	23.75	21.75	21	113.25	183	37	37.5	149.5	155
S001	19	M	37.1	13	23	21	21	20.25	21.5	102.75	171.25	37.75	32.5	144.5	151.5
S002	11	M	37.1	13.75	24	21	21.75	21	20	112.5	188	37.5	38	151	154.5
N001	13	M	37.4	13.25	24	22	22	22	22	121	186.75	37	36.75	144	154
N002	25	M	37.2	14.75	30	22	22	21	21.5	123.75	184	37.75	38	133.5	144
N003	15	M	37.1	14	23	21	21	20	20	114.5	181	35.5	35.5	136	146.5
H001	6	G	37.1	13.75	24.75	22	22	21	21.25	109	175	40.5	40	145.5	153.5
H002	5	G	37.1	13	23	20	20	20.5	20	118.25	185	37.75	38.5	144.5	148.5

Anatomical measurements, measured with:



a) sartorial meter



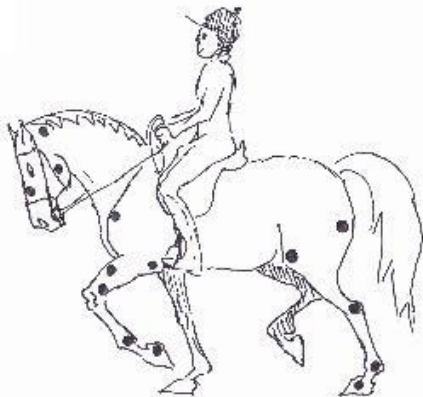
b) measuring stick

# Kinematic measurements and information about horses' housing and work.

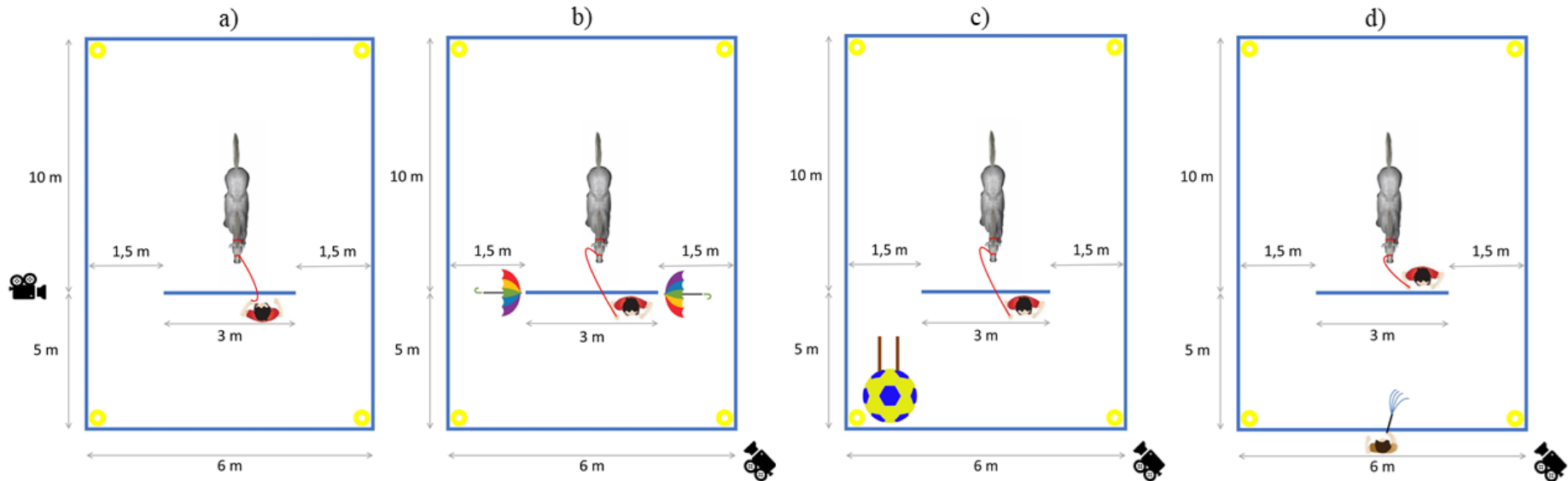
ID	Repetition	Test	Regularity	Simmetry	Cadence	DV power	Propulsion power	Stride length	Speed
G001	1	Walk	115	282	0.76	0.3	1.4	1.73	1.3
G001	1	Walk	178	215	0.76	0.4	1.4	1.76	1.3
G001	1	Walk	142	238	0.78	0.4	1.6	1.81	1.4
G001	1	Walk	167	192	0.73	0.3	1.2	1.92	1.4
G001	2	Walk	137	211	0.76	0.4	1.6	1.85	1.4
G001	2	Walk	162	200	0.76	0.3	1.3	1.77	1.4
G001	2	Walk	116	166	0.78	0.4	1.5	1.76	1.4
G001	2	Walk	135	281	0.73	0.3	1.2	1.8	1.3
G001	1	Trot	275	253	1.29	11.4	6.7	2.21	2.7
G001	1	Trot	240	200	1.22	8.5	4.6	2.09	2.6
G001	1	Trot	294	197	1.25	10.6	6.01	2.17	2.7
G001	1	Trot	288	250	1.25	10.9	4.8	2.19	2.7
G001	2	Trot	303	219	1.29	12.2	7.3	2.19	2.8
G001	2	Trot	280	400	1.25	9.3	4	2.01	2.5
G001	2	Trot	272	336	1.34	13.5	9.3	2.16	2.9
G001	2	Trot	297	196	1.29	11.6	6.1	2.3	3

K001	1	Walk
K001	1	Walk
K001	1	Walk
K001	1	Walk
K001	2	Walk
K001	2	Walk
K001	2	Walk
K001	2	Walk

ID	Work	School horse	Housing	Number of boxes	Equine activities	Pasture	Number of caretakers	Box openness	Injury history
G001	1	2	1	4	1	2	1	1	2
K001	2	1	1	9	2	1	8	2	2
K002	2	1	1	9	2	1	8	2	2
K003	2	1	1	9	2	1	8	2	1
K004	2	1	1	9	2	1	8	2	1
K005	2	1	1	9	2	1	8	2	1
K006	2	1	1	9	2	1	8	2	2
K007	2	1	1	9	2	1	8	2	2
Z001	1	1	2	0	1	2	2	1	2
J001	2	1	1	8	1	1	6	2	2
J002	2	1	1	8	1	1	6	2	2
J003	2	1	1	8	1	1	6	2	2
J004	2	1	1	8	1	1	6	2	2
J005	2	1	1	8	1	1	6	2	2
J006	2	1	1	8	1	1	6	2	2
J007	2	1	1	8	1	1	6	2	2
J008	2	1	1	4	1	1	6	1	2
S001	1	2	1	8	2	1	4	2	2
S002	1	1	2	0	2	1	5	1	1
N001	1	2	1	3	2	1	1	2	2
N002	1	2	1	3	2	1	1	2	1
N003	1	2	1	3	2	1	1	2	2
H001	1	2	2	0	2	2	2	1	2
H002	1	2	1	6	2	2	2	2	1



## The series of fear tests were conducted as part of the project about Lipizzan horses personality.



**a) Passive human test** the horse was held on a loose leash by a passive, unfamiliar handler who allowed the horse to initiate the interaction.

**b) Umbrella test** the horse was led through a passage formed by two identical colored umbrellas.

**c) Rolling ball test** the horse was led through the test area where a large blue and yellow ball was released from the ramp and rolled past the horse.

**d) Bag test** a blue plastic bag, attached to a whip was swung frontally from the horse.

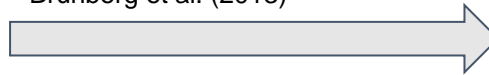


## Ethogram of motor reactivity for the definition of fear scores

Score	Reaction	Description
1	None	No fear reaction to stimulus
2	Head up	Head lifted high, no movement away, ears forward, possible approach toward the stimuli
3	Alert	Head high, quivering, stopping or moving one or two steps away from the stimulus, ears forward
4	Away	Moving backwards from the stimulus, not more than four strides
5	Flight	Jumping or turning away from the stimulus in a sudden movement, followed by trotting or galloping of more than four strides

Horse	ID	Repetition	Test	Fear score
1	G001	1	1	1
1	G001	1	2	1
1	G001	1	3	5
1	G001	1	4	4
1	G001	2	1	1
1	G001	2	2	2
1	G001	2	3	3
1	G001	2	4	4
2	K001	1	1	1
2	K001	1	2	2
2	K001	1	3	3
2	K001	1	4	4
2	K001	2	1	1
2	K001	2	2	1
2	K001	2	3	1
2	K001	2	4	3

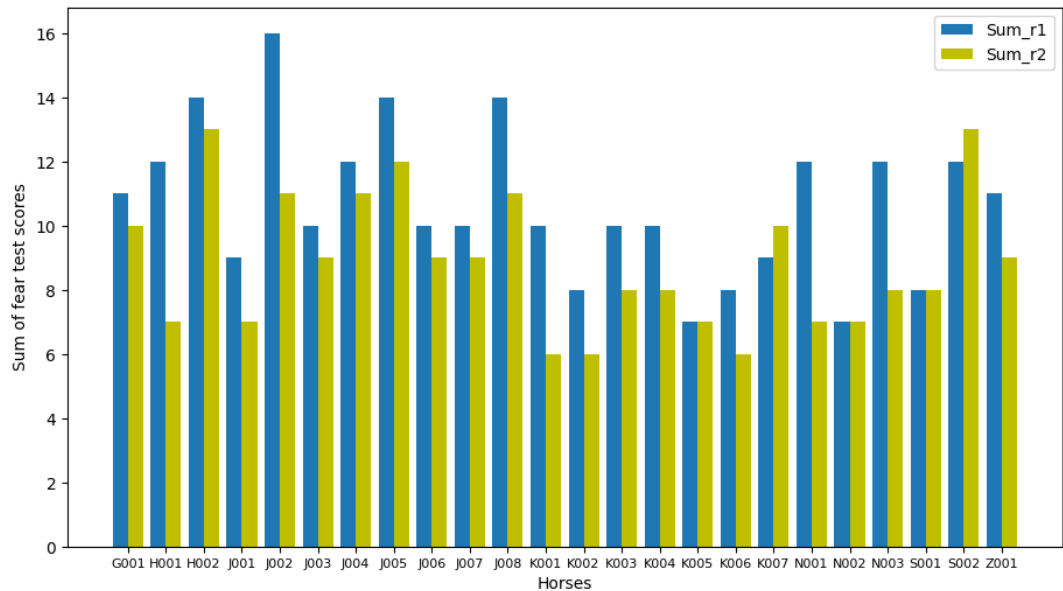
The sum of all the four scores  
will present the final fear score.  
Brunberg et al. (2013)



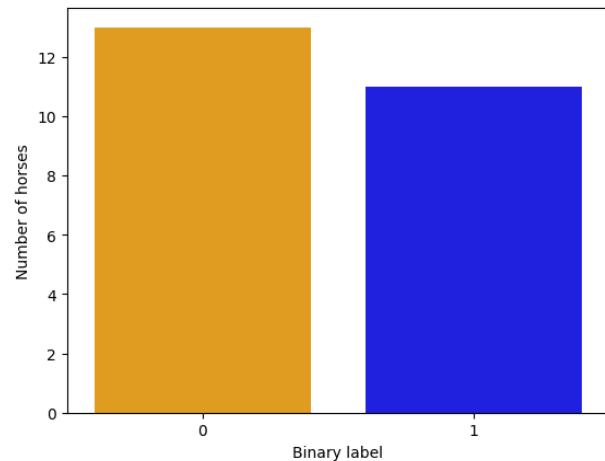
ID	TestR1_1	TestR1_2	TestR1_3	TestR1_4	Sum_r1
G001	1	1	5	4	11
H001	1	3	4	4	12
H002	2	3	5	4	14
J001	1	3	2	3	9
J002	3	3	5	5	16
J003	1	2	2	5	10
J004	1	2	5	4	12
J005	1	4	5	4	14
J006	1	2	3	4	10
J007	1	1	3	5	10
J008	2	3	4	5	14
K001	1	2	3	4	10
K002	1	3	1	3	8
K003	1	2	3	4	10
K004	1	3	2	4	10
K005	1	1	2	3	7
K006	1	2	1	4	8
K007	1	3	2	3	9
N001	1	2	5	4	12
N002	1	1	1	4	7
N003	1	1	5	5	12
S001	1	1	3	3	8
S002	1	2	5	4	12
Z001	1	2	4	4	11

## Sum of the fear score test by repetitions

ID	sum_r1	sum_r2	max	binary
G001	11	10	11	1
H001	12	7	12	1
H002	14	13	14	1
J001	9	7	9	0
J002	16	11	16	1
J003	10	9	10	0
J004	12	11	12	1
J005	14	12	14	1
J006	10	9	10	0
J007	10	9	10	0
J008	14	11	14	1
K001	10	6	10	0
K002	8	6	8	0
K003	10	8	10	0
K004	10	8	10	0
K005	7	7	7	0
K006	8	6	8	0
K007	9	10	10	0
N001	12	7	12	1
N002	7	7	7	0
N003	12	8	12	1
S001	8	8	8	0
S002	12	13	13	1
Z001	11	9	11	1



**Visualization of the division of horses into two classes according to the level of fear.**  
(Average to maximum 10.75)



## Fragment of correlation matrix of the anatomical features of Lipizzan horses

	FB07	FB08	FB09	FB10L	FB10R	FB11L	FB11R	FB12L	FB12R	FB13L	FB13R	FB14L	FB14R	FB15L	FB15R
FB07	1	-0.03432	-0.02094	0.490295	0.491822	0.286058	0.417504	0.365368	0.239598	0.29174	0.197423	0.222077	0.349661	0.223437	0.22986
FB08	-0.03432	1	0.85071	-0.1359	-0.17169	-0.24117	-0.24144	-0.19172	-0.24692	-0.30907	-0.28534	-0.27513	-0.27934	-0.08341	-0.06567
FB09	-0.02094	0.85071	1	0.061398	0.010233	-0.13723	-0.1658	-0.11801	-0.1526	-0.1695	-0.15326	-0.26542	-0.17966	0.047141	0.068838
FB10L	0.490295	-0.1359	0.061398	1	0.977362	0.50196	0.635156	0.470836	0.387981	0.545257	0.510463	0.399348	0.438238	0.653286	0.617876
FB10R	0.491822	-0.17169	0.010233	0.977362	1	0.501307	0.636463	0.445054	0.368965	0.533698	0.503706	0.38973	0.433218	0.651758	0.613943
FB11L	0.286058	-0.24117	-0.13723	0.50196	0.501307	1	0.897363	0.832354	0.829658	0.850536	0.909071	0.744374	0.816814	0.753767	0.829731
FB11R	0.417504	-0.24144	-0.1658	0.635156	0.636463	0.897363	1	0.777879	0.710853	0.829551	0.860415	0.698167	0.777901	0.749346	0.821167
FB12L	0.365368	-0.19172	-0.11801	0.470836	0.445054	0.832354	0.777879	1	0.929136	0.820273	0.828924	0.887231	0.860461	0.717697	0.732675
FB12R	0.239598	-0.24692	-0.1526	0.387981	0.368965	0.829658	0.710853	0.929136	1	0.813186	0.850518	0.878183	0.872015	0.657319	0.700307
FB13L	0.29174	-0.30907	-0.1695	0.545257	0.533698	0.850536	0.829551	0.820273	0.813186	1	0.915484	0.785372	0.848459	0.581985	0.669222
FB13R	0.197423	-0.28534	-0.15326	0.510463	0.503706	0.909071	0.860415	0.828924	0.850518	0.915484	1	0.811122	0.828891	0.667031	0.779044
FB14L	0.222077	-0.27513	-0.26542	0.399348	0.38973	0.744374	0.698167	0.887231	0.878183	0.785372	0.811122	1	0.879905	0.698511	0.700965
FB14R	0.349661	-0.27934	-0.17966	0.438238	0.433218	0.816814	0.777901	0.860461	0.872015	0.848459	0.828891	0.879905	1	0.680376	0.727273
FB15L	0.223437	-0.08341	0.047141	0.653286	0.651758	0.753767	0.749346	0.717697	0.657319	0.581985	0.667031	0.698511	0.680376	1	0.935612
FB15R	0.22986	-0.06567	0.068838	0.617876	0.613943	0.829731	0.821167	0.732675	0.700307	0.669222	0.779044	0.700965	0.727273	0.935612	1
FB16L	-0.10896	0.147352	0.207875	0.172239	0.218595	0.071212	0.100465	-0.10695	0.011692	0.189307	0.152227	-0.1595	-0.06586	-0.17265	-0.08644

## Fragment of the correlation of the features matrix of housing and working conditions of Lipizzan horses

	Work	School horse	Housing	Number of boxes	Equine activities	Pasture	Number of caretakers	Box openness	Injury history
Work	1	-0.82841687	-0.48795	0.790694367	-0.305505046	-0.57735	0.868845685	0.450341001	0.149071198
School horse	-0.82842	1	0.034648	-0.55460267	0.356385752	0.450956	-0.788315605	-0.122266666	-0.052925612
Housing	-0.48795	0.034647946	1	-0.59284436	0.063887656	0.507093	-0.310899075	-0.736788398	-0.072739297
Number of boxes	0.790694	-0.55460267	-0.59284	1	0.170439466	-0.49269	0.906912413	0.636043337	-0.057497442
Equine activities	-0.30551	0.356385752	0.063888	0.170439466	1	-0.07559	0.107438681	0.190764193	-0.487950036
Pasture	-0.57735	0.450956034	0.507093	-0.492694184	-0.075592895	1	-0.493267818	-0.596480908	0
Number of caretakers	0.868846	-0.788315605	-0.3109	0.906912413	0.107438681	-0.49327	1	0.391276836	-0.07915103
Box openness	0.450341	-0.122266666	-0.73679	0.636043337	0.190764193	-0.59648	0.391276836	1	-0.059234888
Injury history	0.149071	-0.052925612	-0.07274	-0.057497442	-0.487950036	0	-0.07915103	-0.059234888	1

## Lists of removed features

- **AllFeatures (120 variables):** removal of correlated features is not performed
- **Removed LeftCorr (89 variables):** 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 variables):** 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 variables):** 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 variables):** 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

```
LeftCorr=['FH10L','FH11L','FH12L','FH13L','FH14L','FH16L',  
'FH18L','FH19L','FB08','FB10L','FB11L','FB12L','FB13L',  
'FB14L','FB15L','FB18L','FB19L','FB20L','FB21L','FB22L',  
'FB23L','FB24L','FB26L','FB27L','FB28L','FB32L','FB34L',  
'FB35L','FB37L','FB38','FB39']
```

```
RightCorr=['FH10R','FH11R','FH12R','FH13R','FH14R','FH16R',  
'FH18R','FH19R','FB08','FB10R','FB11R','FB12R','FB13R',  
'FB14R','FB15R','FB18R','FB19R','FB20R','FB21R','FB22R',  
'FB23R','FB24R','FB26R','FB27R','FB28R','FB32R','FB34R',  
'FB35R','FB37R','FB38','FB39']
```

```
LeftCorr+=['FH10L','FH11L','FH12L','FH13L','FH14L','FH16L',  
'FH18L','FH19L','FB08','FB10L','FB11L','FB12L','FB13L',  
'FB14L','FB15L','FB18L','FB19L','FB20L','FB21L','FB22L',  
'FB23L','FB24L','FB26L','FB27L','FB28L','FB32L','FB34L',  
'FB35L','FB37L','FB38','FB39','FH18R','FB11R','FB12R',  
'FB13R']
```

```
RightCorr+=['FH10R','FH11R','FH12R','FH13R','FH14R','FH16R',  
, 'FH18R','FH19R','FB08','FB10R','FB11R','FB12R','FB13R',  
'FB14R','FB15R','FB18R','FB19R','FB20R','FB21R','FB22R',  
'FB23R','FB24R','FB26R','FB27R','FB28R','FB32R','FB34R',  
'FB35R','FB37R','FB38','FB39','FH18L','FB11L','FB12L',  
'FB13L']
```

# Machine learning methods and algorithms for solving classification task for predicting horse fearfulness

## Leave-One-Out Cross Validation

	X1	X2	...	X3	X4	Y				
1										
2		X1	X2	...	X3	X4	Y			
3	1		X1	X2	...	X3	X4	Y		
4	2	1								
5	3	2		X1	X2	...	X3	X4	Y	
6	4	3	1		X1	X2	...	X3	X4	Y
7	5	4	2	1						
8	6	5	3	2						
9	7	6	4	3						
10	8	7	5	4						
11	9	8	6	5						
12	10	9	7	6						
13	11	10	8	7						
14	12	11	9	8						
15	13	12	10	9						
16	14	13	11	10						
17	15	14	12	11						
18	16	15	13	12						
19	17	16	14	13						
20	18	17	15	14						
21	19	18	16	15						
22	20	19	17	16						
23	21	20	18	17						
24	22	21	19	18						
	23	22	20	19						
	24	23	21	20						
		24	22	21						
			23	22						
				24						
					24					

### Logistic Regression

`solver='liblinear', penalty='l1', C=0.2`

### Support Vector Machine

`C=1.5, kernel='sigmoid'`

### Decision Trees

Depends on the randomness of the estimator.

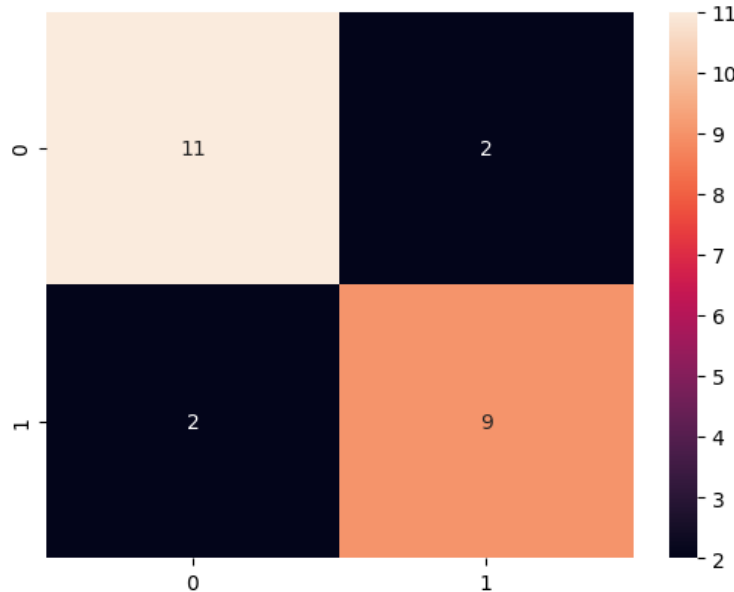
### Random Forests

`min_samples_split=3`

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

Confusion matrix and metrics by Decision Trees Classification results (LOOCV)

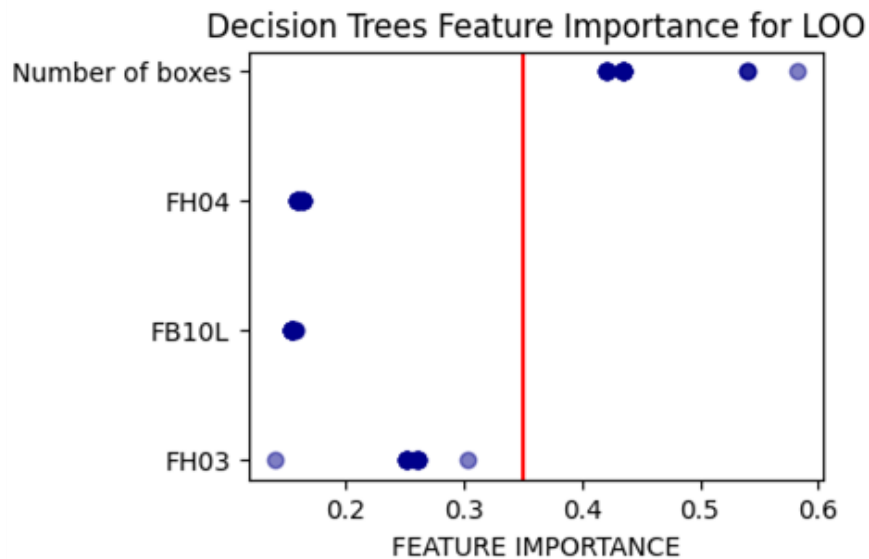


Accuracy= 0.8333333333333334

Precision= 0.8181818181818182

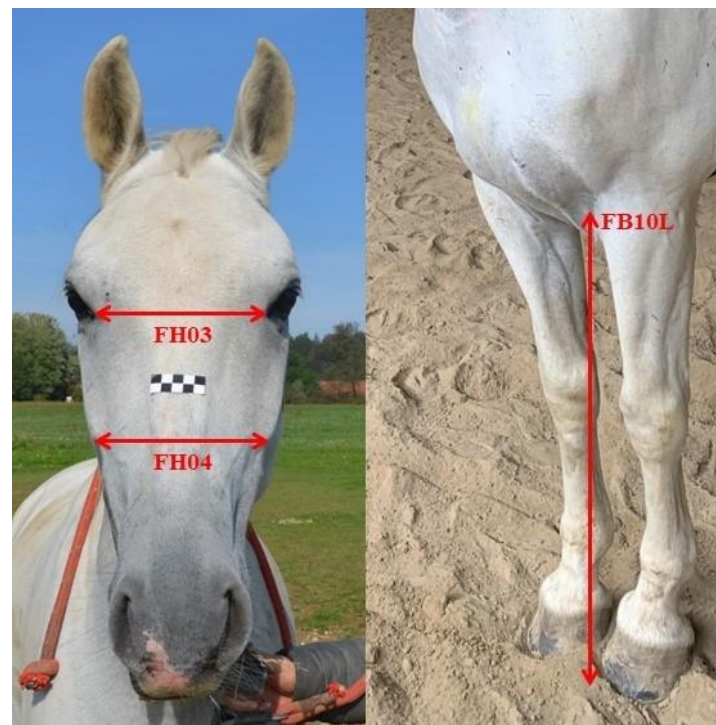
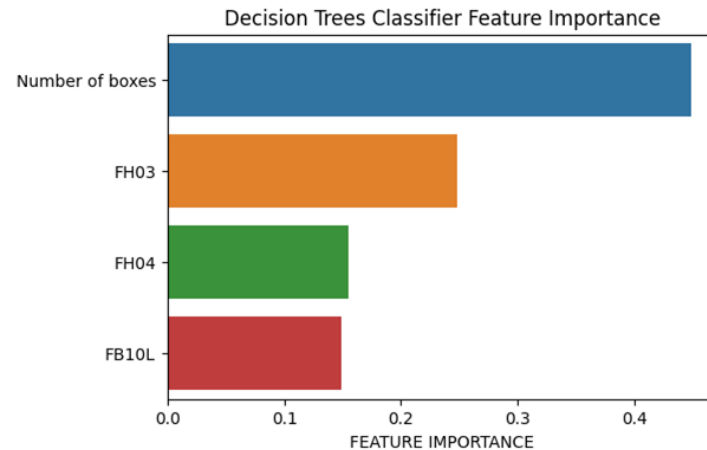
Recall= 0.8181818181818182

## Feature importance (LOOCV)

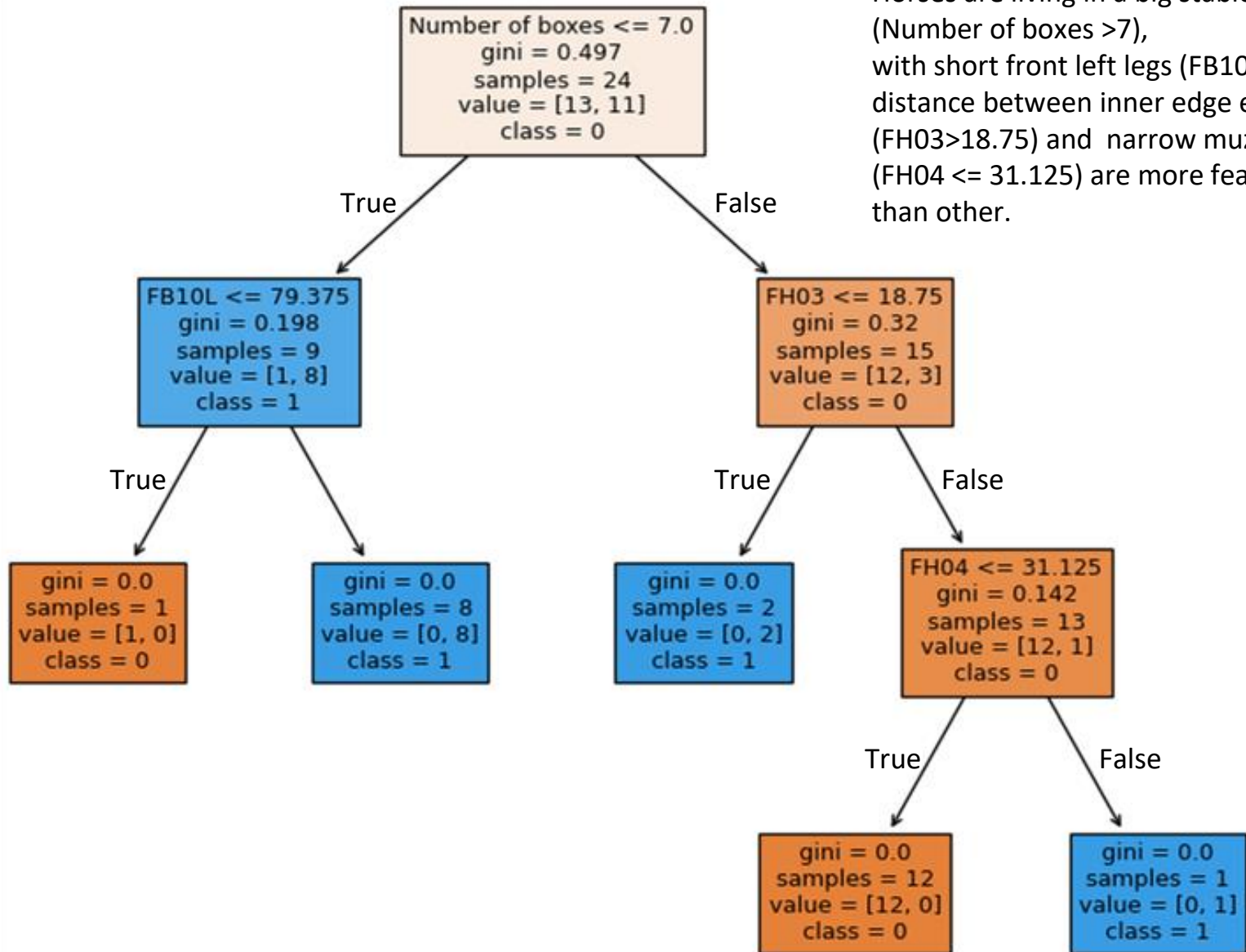


Features name	Numbers of LOOCV steps
Number of boxes	24
FB10L (Left leg length)	23
FH03 (Distance between the inner edges of the eyes)	21
FH04 (Horse muzzle width)	18

## Feature importance (all instances)



## Decision Tree trained on all the instances



Horses are living in a big stable (Number of boxes  $>7$ ), with short front left legs (FB10L), large distance between inner edge eyes (FH03 $>18.75$ ) and narrow muzzle (FH04  $\leq 31.125$ ) are more fearless than other.



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

**Thanks for your attention!**