



# Dataset: Horse Movement Data and Analysis of its Potential for Activity Recognition

Jacob Kammenga, Nirvana Meratnia, Paul Havinga



# Animal Activity Recognition (AAR)

- AAR provides large amount of information regarding the animal and its environment
  - Livestock management
  - E-health for pets
  - Detection of environmental effects such as forest fires and poaching
  - Movement ecology
  - Wildlife conservation

## Data Acquisition

- 14 horses and 4 ponies
- 7 days of data collection

# Sensors

## GCDC Human Activity Monitor (HAM)

- 3D Accelerometer, 100Hz
- 3D Gyroscope, 100Hz
- 3D Magnetometer, 12Hz





## Sensor Placement











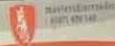
D

XX

GG

LESSCHER 4WD  
SAASVLD  
ISSAN & TOYOTA 4x4 SPECIALIST  
074-340 40 66

MASTERS



WISO  
www.wiso.nl

H  
autoSERVICE HERMAN  
www.autoserviceherman.nl  
Blaauwweg 171 | Sittard | 046-461040

# Labeling

data\_labeling

Files

- Label files folder location  
Measurements Horstlinde\Labels
- Video folder location  
F:\Measurements Horstlinde\2s
- Sensor data location  
F:\Measurements Horstlinde\2

Current video file: 20180515\_17-35-48

Select Video

accel: 20180515\_18-17

Select accel. file

Video Controls

**15-May-2018** Start: 17:35:48  
**18:44:49** Stop: 20:35:05

Camera name:Sony4K

|<< |>| >>|

Jump to time 18:43:50

Video offset (sec) 0

Sensor offset 0

Jump (sec) on "I" and "T" 8

Select an activity

- lying
- standing
- walking-rider
- walking-natural
- trotting-rider
- trotting-natural
- running-rider
- running-natural
- grazing
- eating
- fighting
- shaking
- scratch-biting
- breast-feeding
- rubbing
- unknown
- food-fight
- head-shake

OK Cancel

Labeling

Most recent label: 18:55:09 "unknown"  
Last label: 18:57:43 "unknown"

Remove most recent label

18:44:45 18:44:50 18:44:55

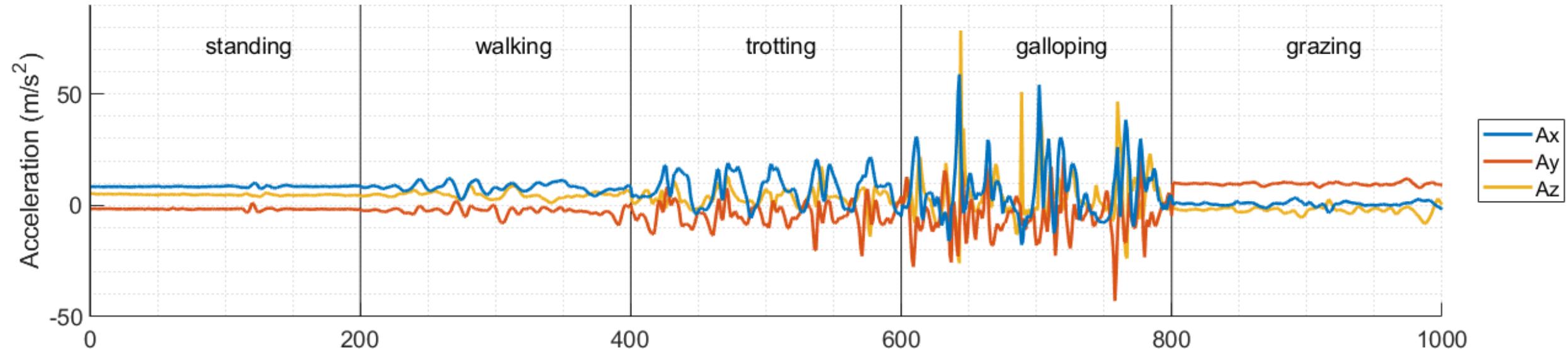
walking-rider standing



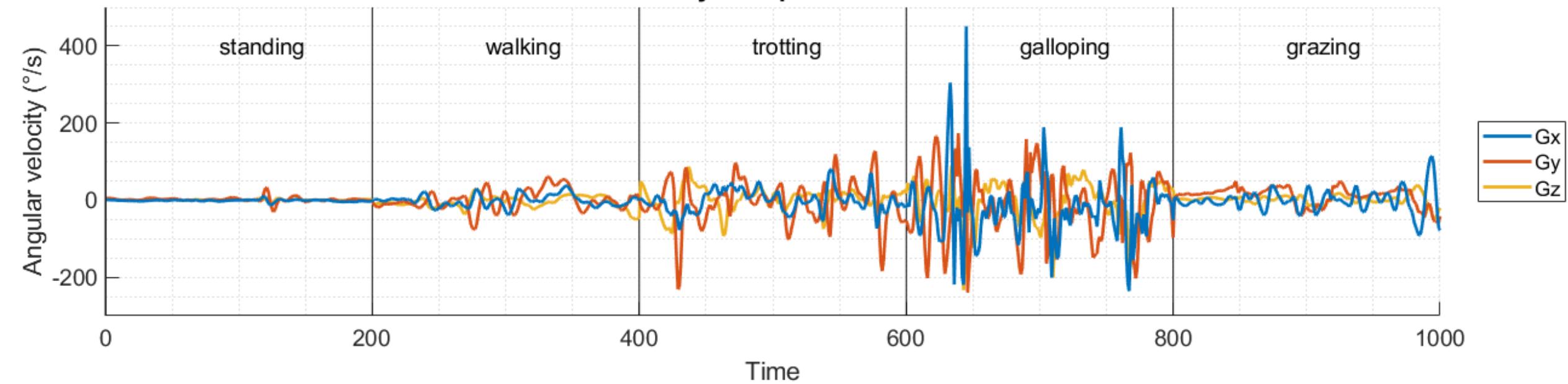
# Column description

Column name	Description
Ax	Raw data from accelerometer x-axis
Ay	Raw data from accelerometer y-axis
Az	Raw data from accelerometer z-axis
Gx	Raw data from gyroscope x-axis
Gy	Raw data from gyroscope y-axis
Gz	Raw data from gyroscope z-axis
Mx	Raw data from magnetometer x-axis
My	Raw data from magnetometer y-axis
Mz	Raw data from magnetometer z-axis
A3D	l2-norm (3D vector) of accelerometer axes
G3D	l2-norm (3D vector) of gyroscope axes
M3D	l2-norm (3D vector) of magnetometer axes
label	Label that belongs to datapoint
segment	Segment identifier
subject	Subject identifier

### Accelerometer data

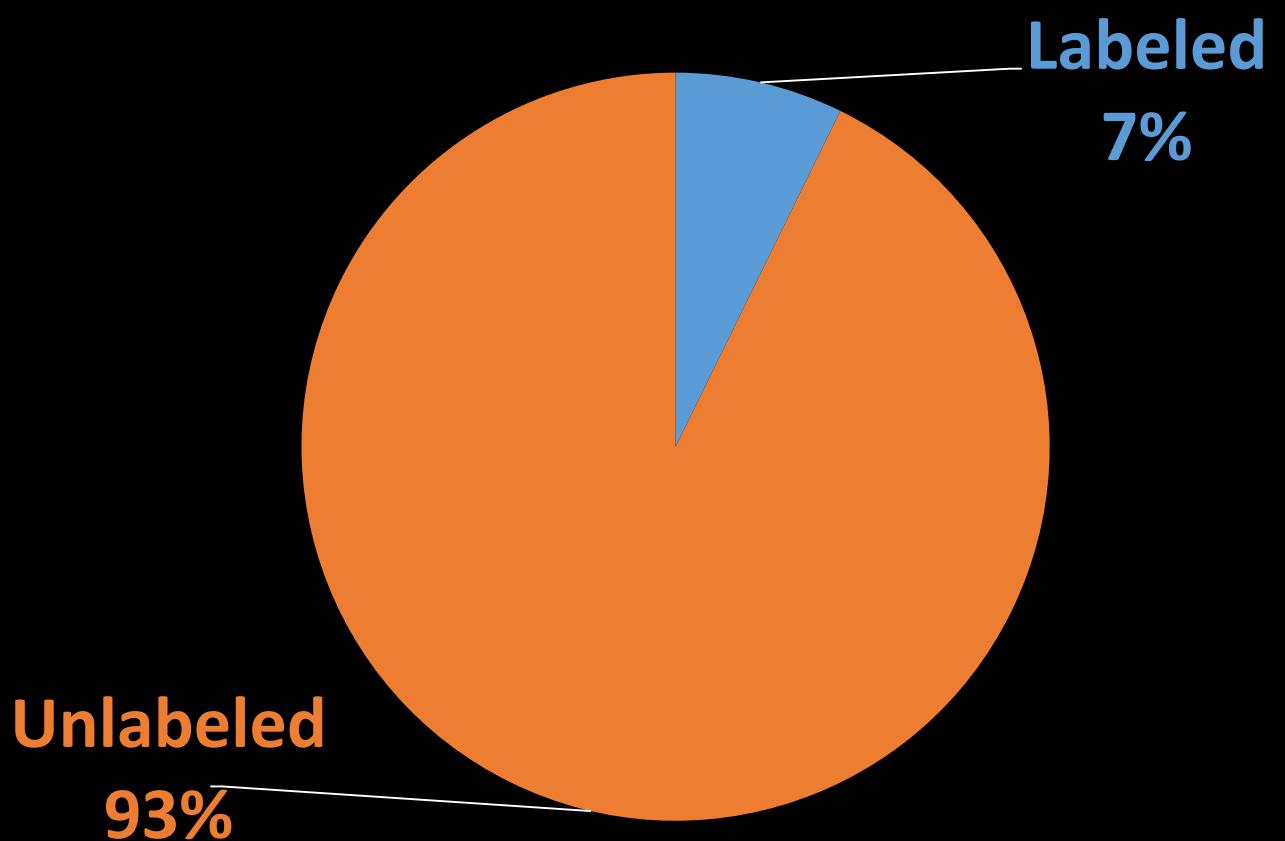


### Gyroscope data

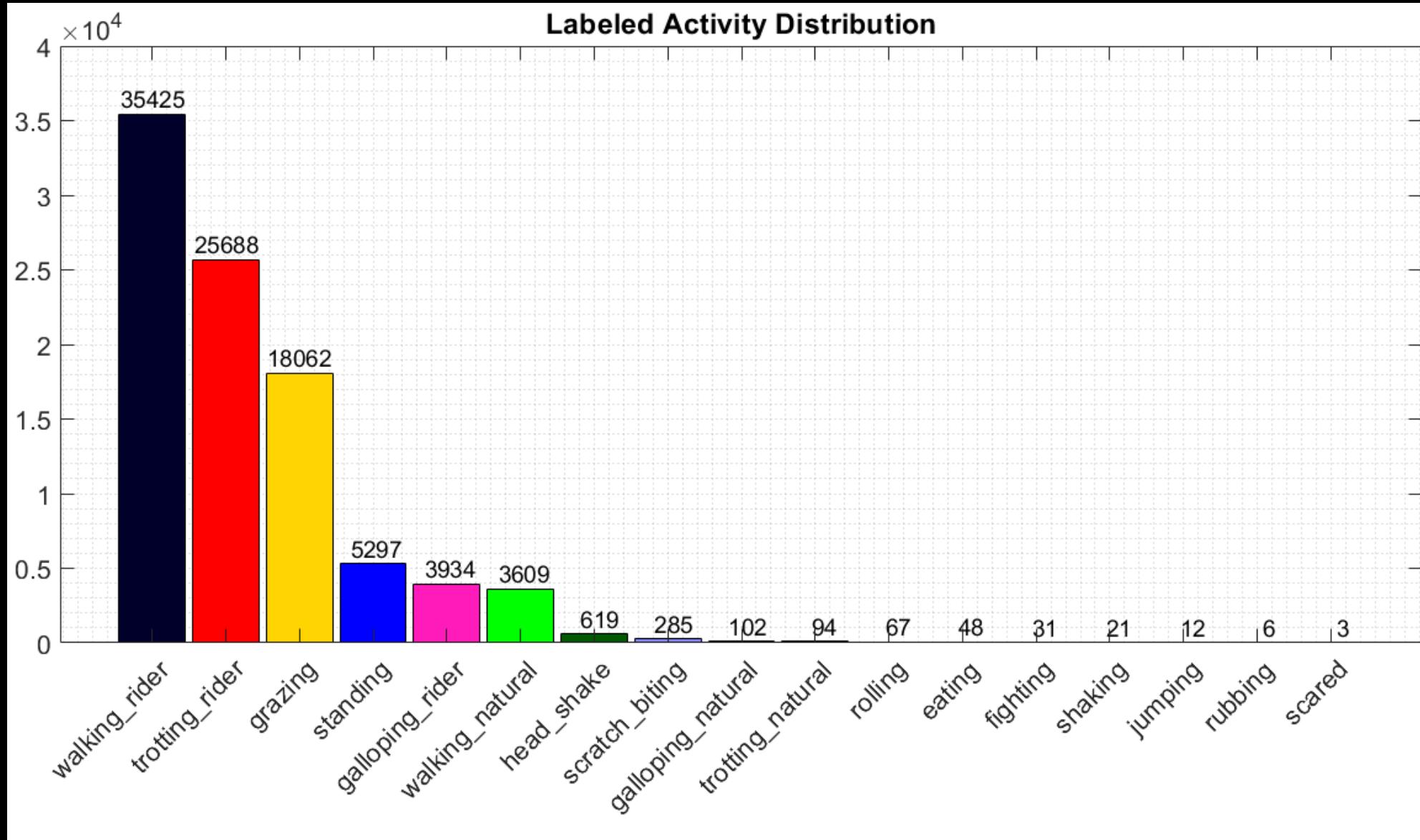


## Number of 2-second samples

Labeled	93303
Unlabeled	1191658



# Number of samples per activity



## Confusion Matrix with 6 activities

True Class

	eating		261	1	3552	248	74.7%	25.3%
eating	11986		261	1	3552	248	74.7%	25.3%
galloping		3487	34	415		3	88.5%	11.5%
standing	597		4499		16	1	88.0%	12.0%
trotting		2435	27	22374		240	89.2%	10.8%
walking-natural	461		5	27	1618	1216	48.6%	51.4%
walking-rider	403	4	36	526	8364	24785	72.6%	27.4%

eating  
galloping  
standing  
trotting  
walking-natural  
walking-rider

Predicted Class

## Dataset use cases

- Unsupervised representation learning
- Sensor-Orientation-Independent AAR
- Gait analysis and comparison
- Transfer learning
- AAR for similar animals such as donkeys or zebras



## Conclusions

- Large dataset, especially unlabeled data
  - Particularly suitable for unsupervised representation learning
- More labeled data for subset of subjects and activities
- Lesson learned: think about prospective dataset uses outside own research objective **prior** to collection

Thanks for your attention

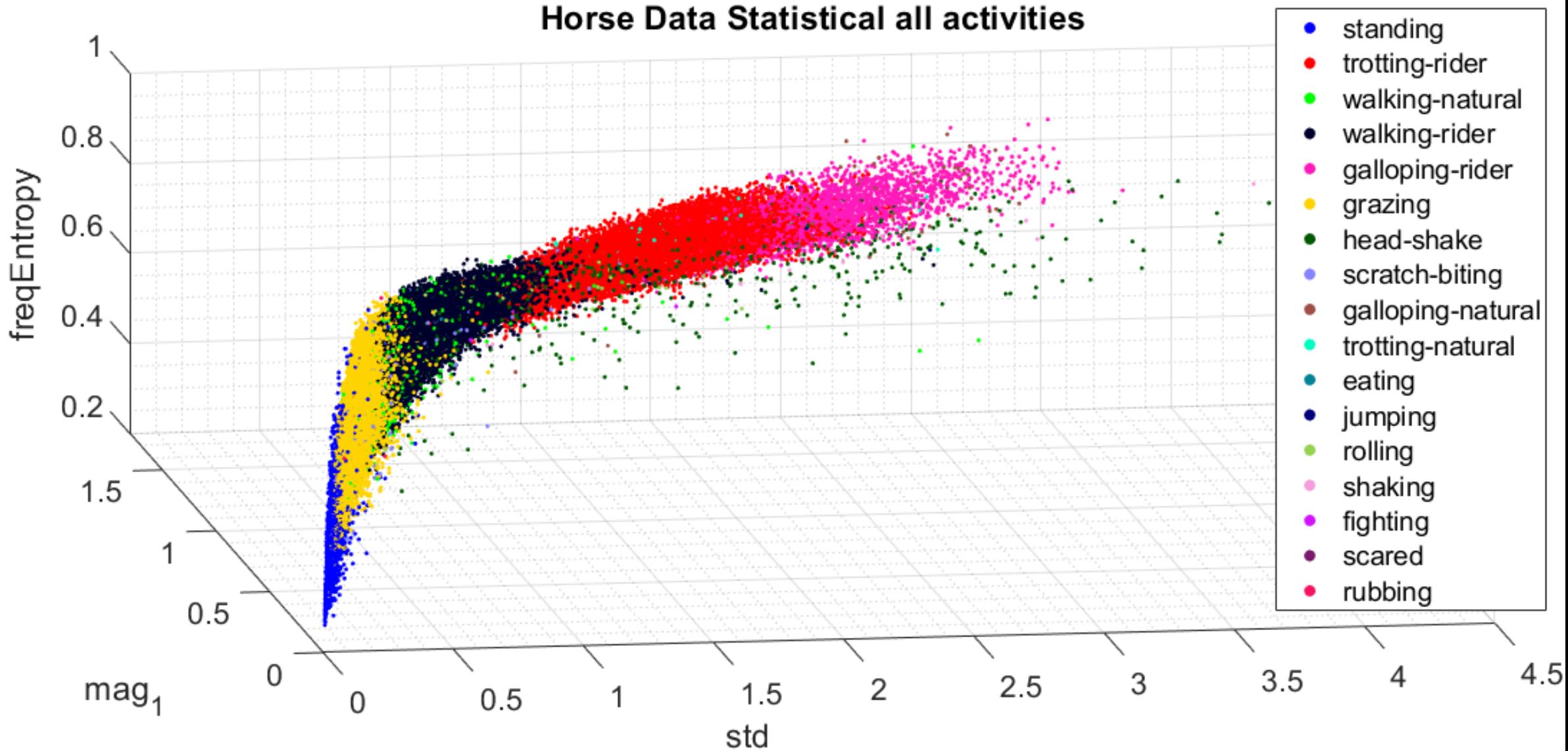


Jacob Kamminga  
University of Twente

## Number of 2-second samples per activity

Activity	null	standing	walking-rider	walking-natural	trotting	galloping	eating	other	total
nr samples	1191658	5297	35425	3609	25782	4036	18110	1044	1284961
fraction of labeled		6%	38%	4%	28%	4%	19%	1%	

## Horse Data Statistical all activities



## Evaluation

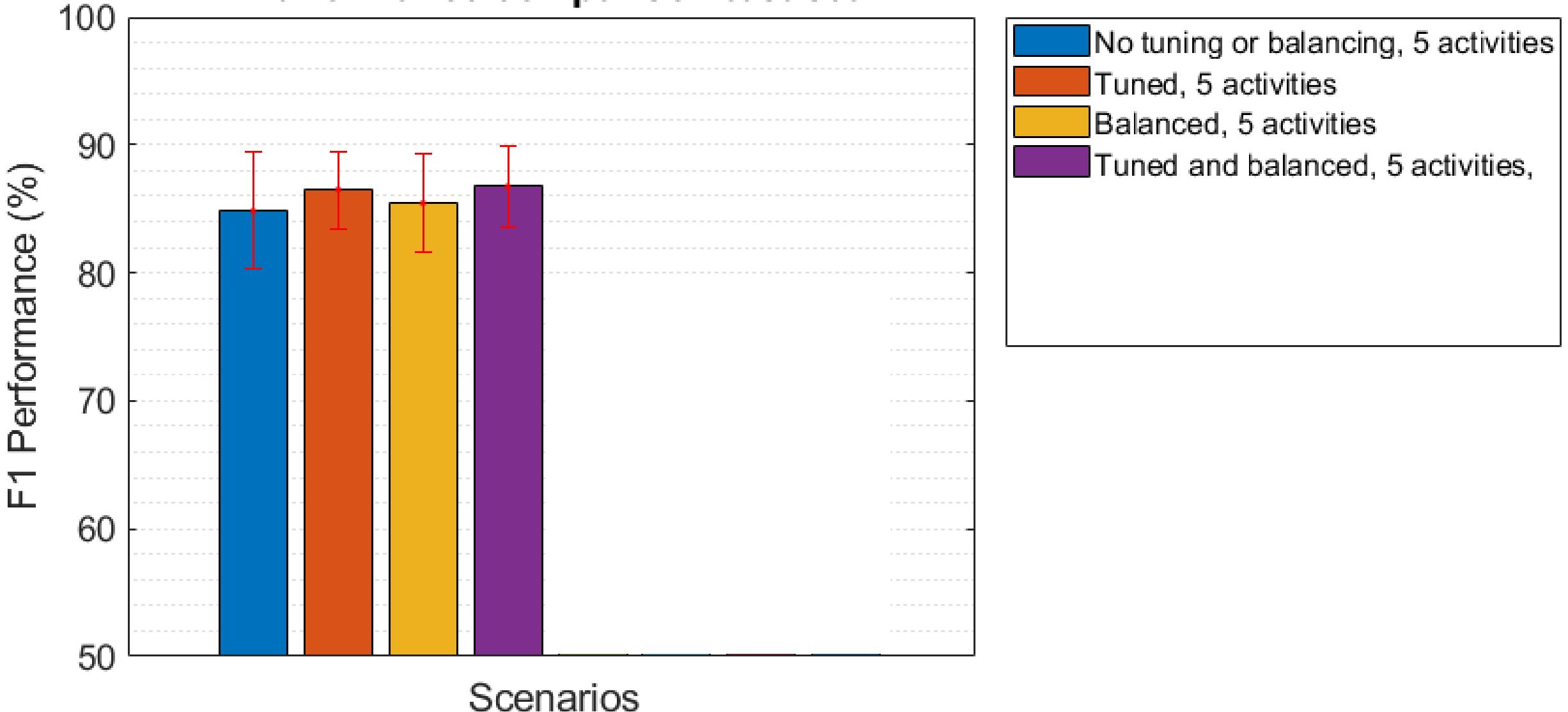
- Naïve Bayes classifier
- Accelerometer data from 6 subjects and 6 activities
- Leave-one-subject-out cross validation
- Transform 3 axes into 3D vector ( $\| \cdot \|_2$ -norm)
- Features: 21 summary statistics (mean, std, ... )

Activity	null	unknown	walking_rider	trotting_rider	grazing	standing	galloping_rider	walking_natural	head_s_hake	scratch_biting	galloping_natural	trotting_natural	eating	jumping	shaking	rolling	fighting	rubbing	scared	total
Galoway	62155	23264	9653	6374	4315	1750	1030	1402	59	170	13	49	16		4	13	25		110292	
Bacardi	92775	9850	1317	1981	1116	245	288	360		22	40			13			6		108013	
Driekus	85468	11271	4024	2670	2465	341	310	270	55	14	13	3	31		4	23			106962	
Patron	78536	15156	5150	3385	1951	1244	709	388	37		5	17				31			106609	
Happy	68468	13606	8896	7032	5062	1186	689	746	238	8	7	6	1						105945	
Zonnerante	90431																		90431	
Duke	81885																		81885	
Viva	69441	4413	1066	700		58	82	79	5	4								1	75849	
Flower	75741																		75741	
Pan	68628	1575	241			36		44											70524	
Porthos	67080																		67080	
Barino	66517																		66517	
Zafir	38424	10349	5078	3546	1091	347	826	161	105	23	9	13		12					59984	
Niro	43563	2740				85	20		2										46410	
Sense	38823	1569				1977	39		157	120	44	15	6				6	2	42758	
Blondy	31579																		31579	
Noortje	17777	2878					31												20686	
Clever	17696																		17696	
total	1094987	96671	35425	25688	18062	5297	3934	3609	619	285	102	94	48	12	21	67	31	6	3	1284961

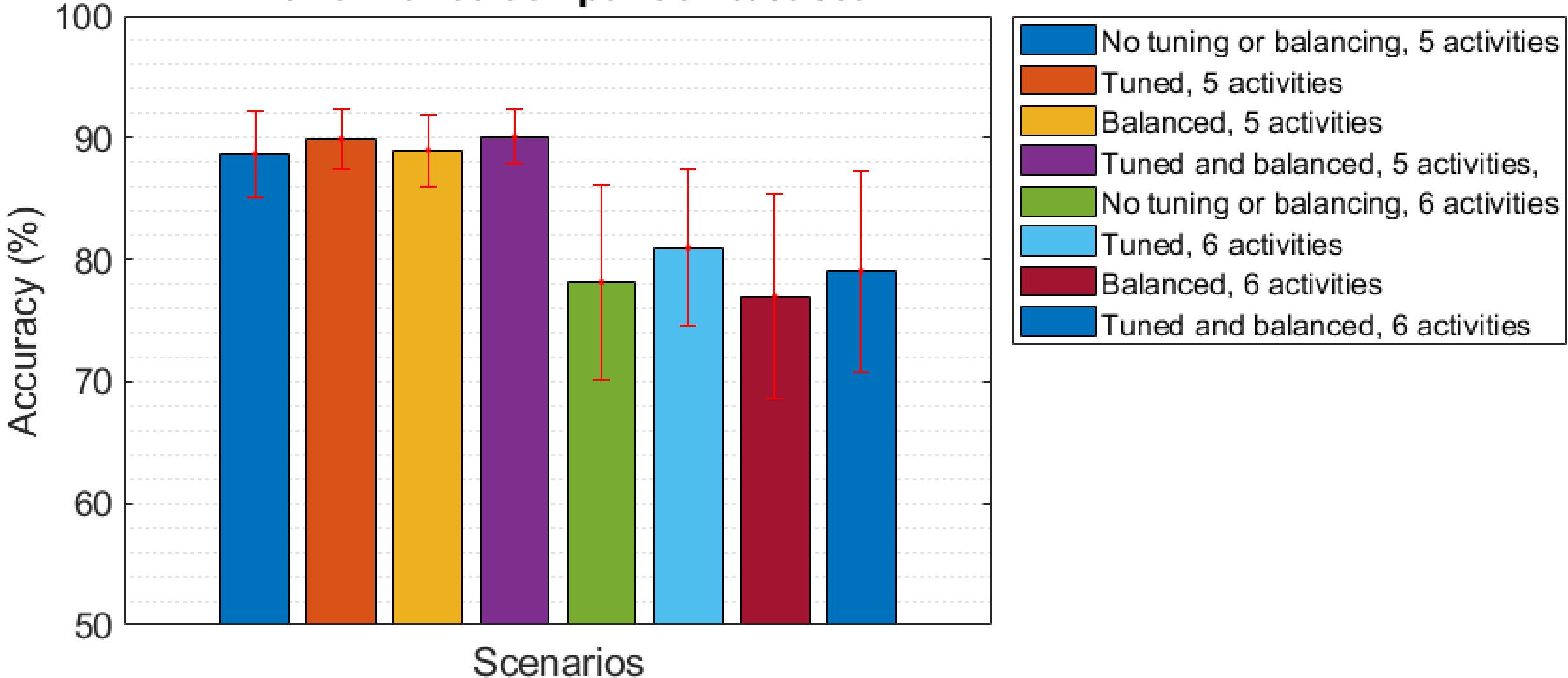
## Evaluation

- Naïve Bayes classifier
- Leave-one-subject-out cross validation
- Accelerometer data from 6 subjects and 6 activities
- Transform 3 axes into 3D vector ( $\| \cdot \|_2$ -norm)
- Features: 21 summary statistics (mean, std, ... )
- Balancing
  - Majority class: random under-sampling
  - Minority class: oversampling using SMOTE

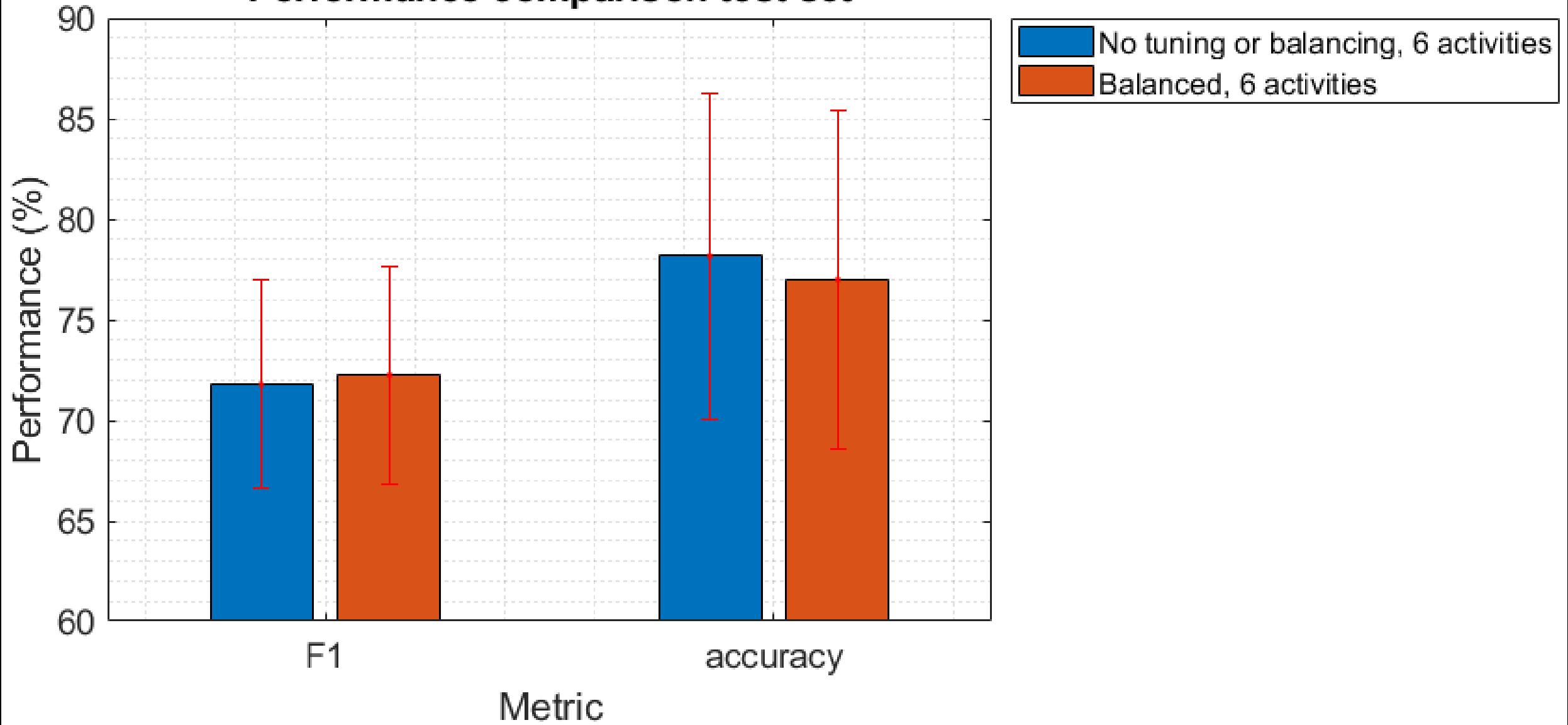
## Performance comparison test set



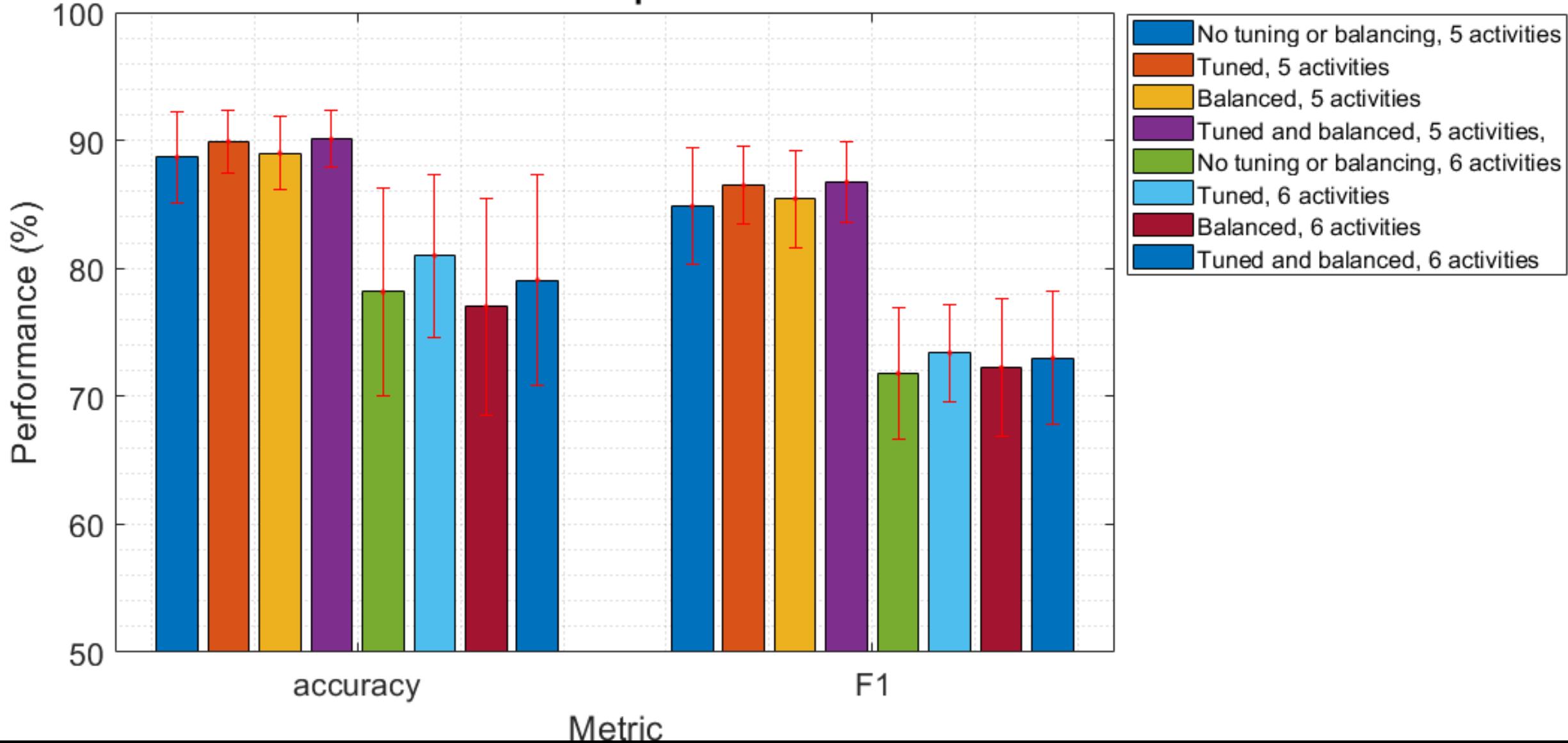
## Performance comparison test set



## Performance comparison test set



## Performance comparison test set



# Summary statistics

Domain	Feature	Description
Time	Maximum	Maximum value
	Minimum	Minimum value
	Mean	Average value
	Standard deviation	Measure of dispersion
	Median	Median value
	\nth{25} percentile	The value below which \SI{25}{\%} of the observations are found
	\nth{75} percentile	The value below which \SI{75}{\%} of the observations are found
	Mean low pass filtered signal	Mean value of DC components
	Mean rectified high pass filtered signal	Mean value of rectified AC components
	Skewness of the signal	The degree of asymmetry of the signal distribution
	Kurtosis	The degree of 'peakedness' of the signal distribution
	Zero crossing rate	Number of zero crossings per second
Frequency	Principal frequency	Frequency component that has the greatest magnitude
	Spectral energy	The sum of the squared discrete FFT component magnitudes
	Frequency entropy	Measure of the distribution of frequency components
	Frequency magnitudes	Magnitude of first six components of FFT analysis

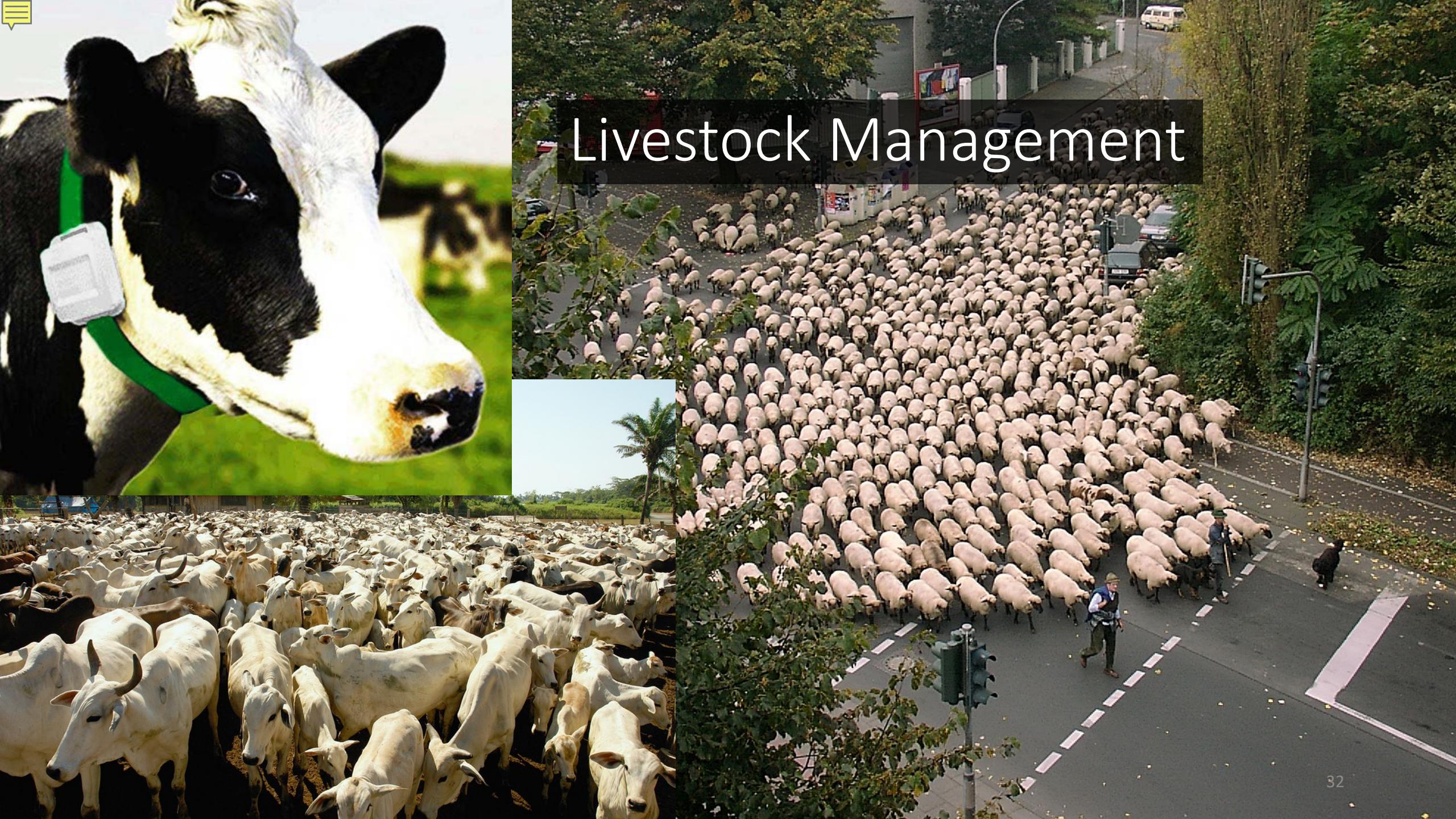
# Why Animal Activity Recognition?



# Detection of Environmental Events



Poaching detection



# Livestock Management



# Wildlife Conservation and Movement Ecology

