

TAIKO MASTER

Content

- Data Collection
- Flow Chart
- Video information
- Play data
- Drum note
- Feature Engineering
- Model
- Experiment

DATA COLLECTION



Sensing platform

Right arm

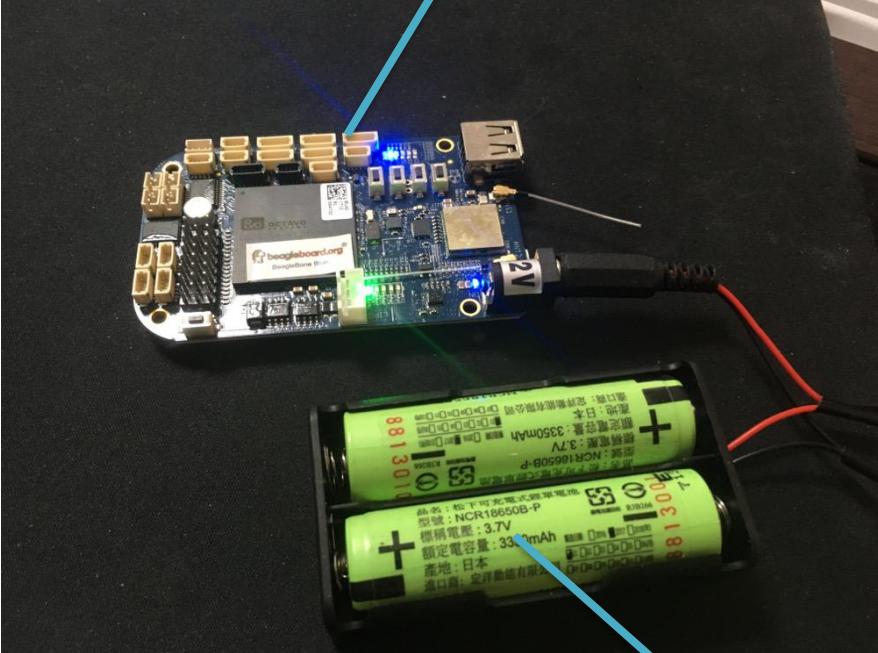


Left arm



Inside the bag

BeagleBone Blue



Put inside

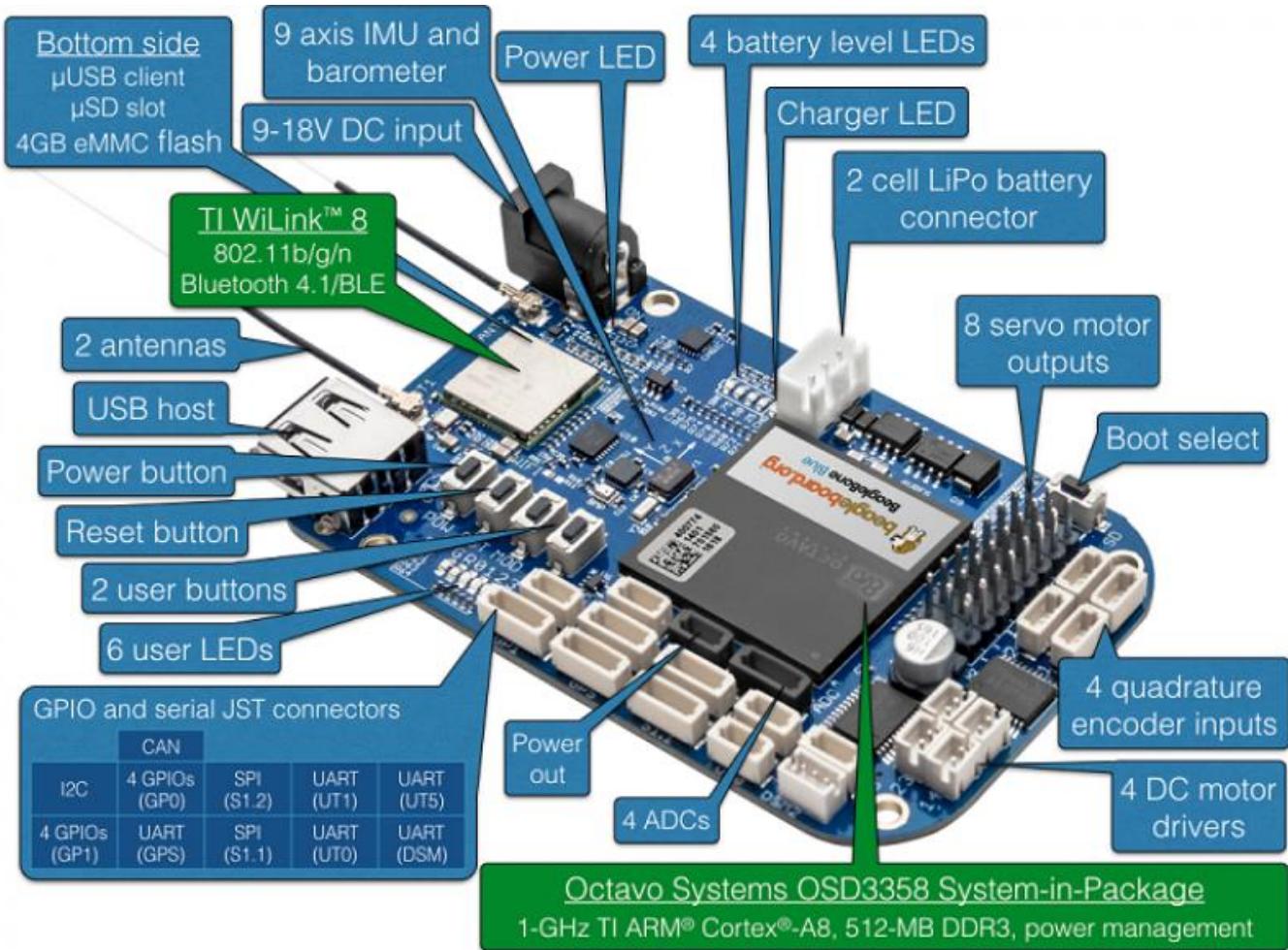
18650 Lithium-ion battery x2



BeagleBone Blue

Sensors:

- 9 axis IMU
 - Accelerators
 - Gyroscopes
 - Magnetometer
- Barometer
- Thermometer



Video Capture Device

Features:

- Support for capture resolutions up to 2048x2160
- Support for frame rates up to 120fps



Capture

*Magewell XI100DUSB One Channel
HDMI Gen 2 USB 3.0 Capture Dongle*



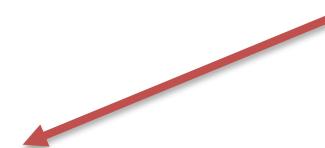
A frame of screenshot during playing

Subjects

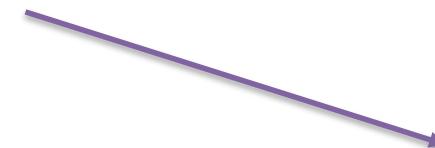
Drummer id	Handedness	Gender	Arm length
1	Right	Male	26.0
2	Right	Male	30.0
3	Right	Female	26.0
4	Right	Female	25.5
5	Right	Male	30.0
6	Right	Male	26.5
7	Left	Male	26.5
8	Right	Female	28.5



Handedness	Count
Right	7
Left	1



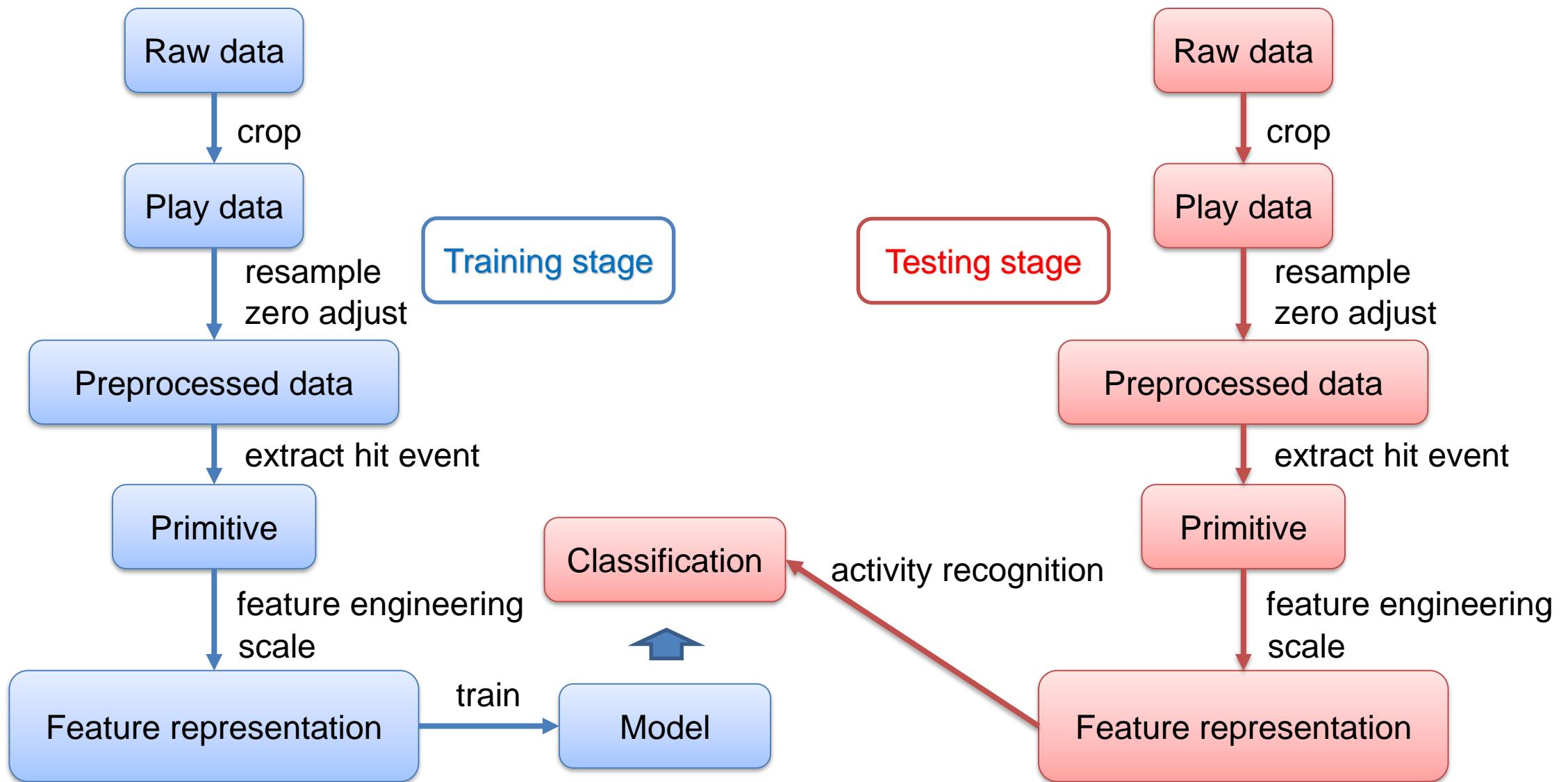
Gender	Count
Male	5
Female	3



FLOW CHART



Flow chart



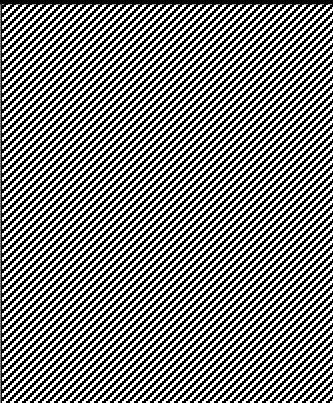
DRUM NOTE



Note

Hit Type	Description	Picture
0	pause	
1	don_small	
2	ka_small	
3	don_big	
4	ka_big	
5	roll_small	
6	roll_big	
7	roll_balloon	

Classification

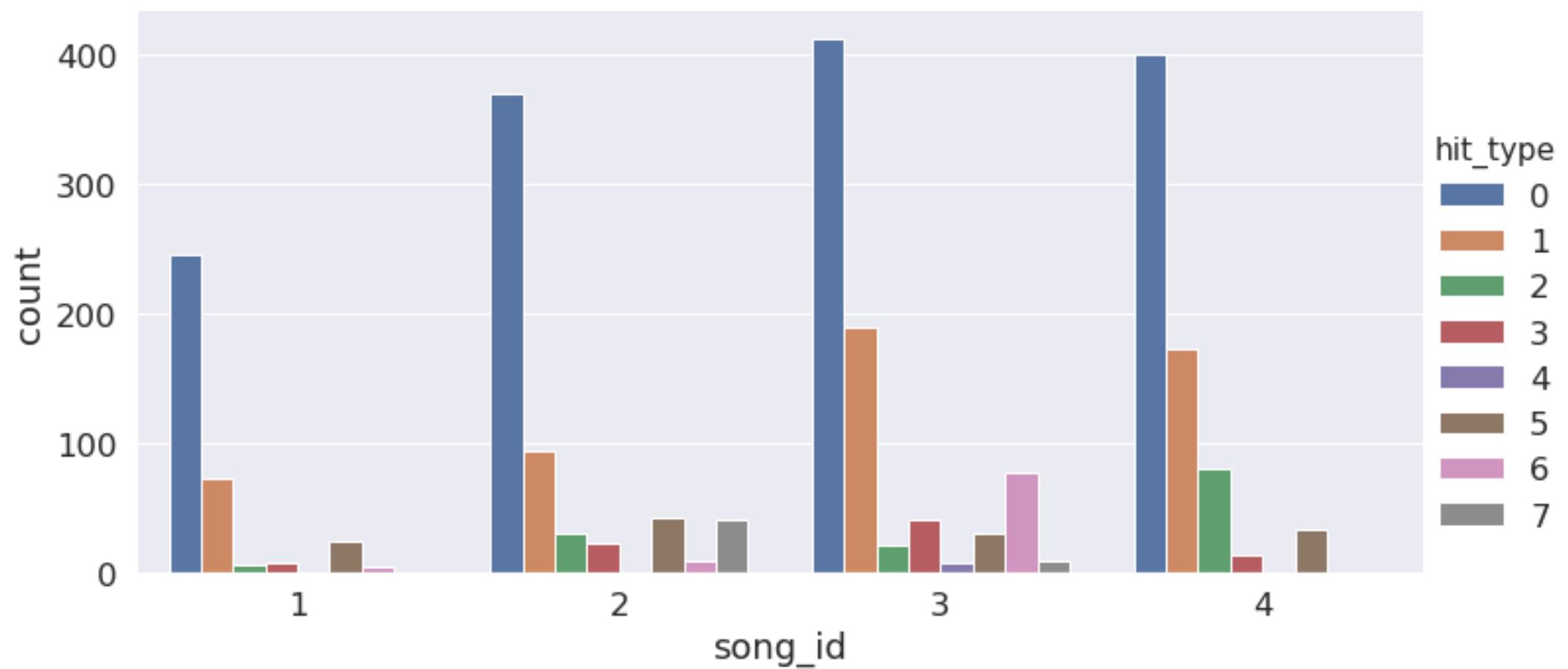
Group	Bag 0	Bag 1	Bag 2	Bag 3
single_stream		 	  	
dong_ka		 	 	  

Song picked

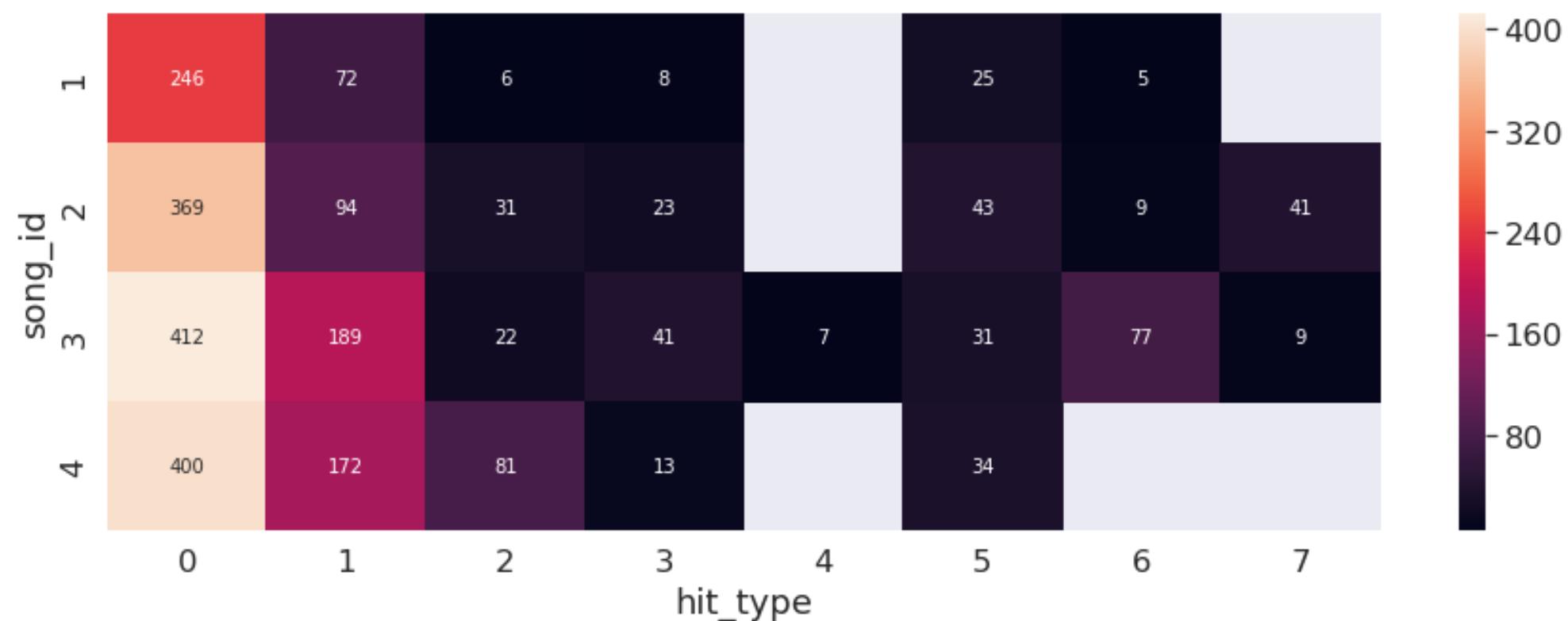
Song id	Name	Difficulty	Repeat play
1	<u>夢をかなえてドラえもん</u>	Easy	3
2	<u>ウイーアー！</u>	Normal	3
3	<u>六兆年と一夜物語</u>	Normal	3
4	<u>残酷な天使のテーゼ</u>	Hard	4

Notice **drummer 1** plays 夢をかなえてドラえもん in normal difficulty

Hit Type Statistics



Hit Type Statistics



Notice the order!

Hit Type Statistics

<code>song_id \ hit_type</code>	0	1	3	2	4	5	6	7
<code>single_stream</code>	Bag 0			Bag 1			Bag 2	
<code>dong_ka</code>	Bag 0		Bag 1		Bag 2		Bag 3	
1	246	72	6	8	0	25	5	0
SS	246			86			30	
DK	246		78		8		30	
2	369	94	31	23	0	43	9	41
SS	369			174			93	
DK	369		125		23		93	
3	412	189	22	41	7	31	77	9
SS	412			259			117	
DK	412		211		48		117	
4	400	172	81	13	0	34	0	0
SS	400			266			34	
DK	400		253		13		34	

VIDEO INFORMATION

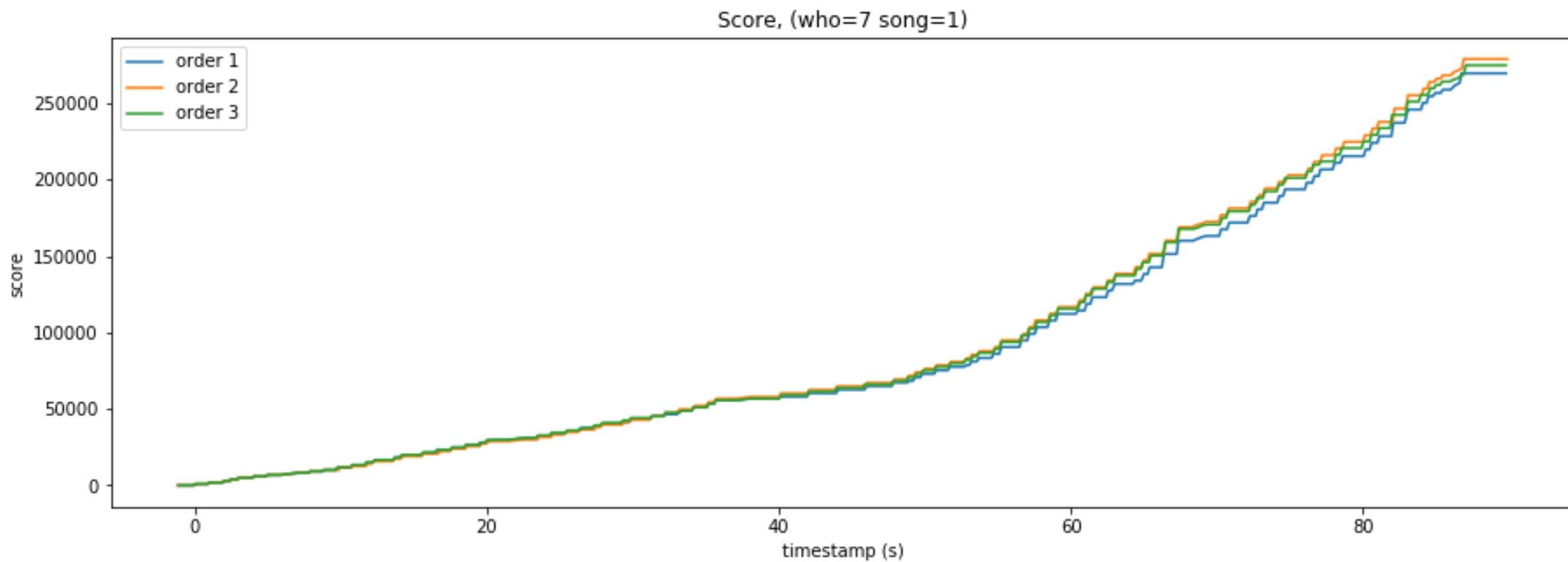


Result Screen

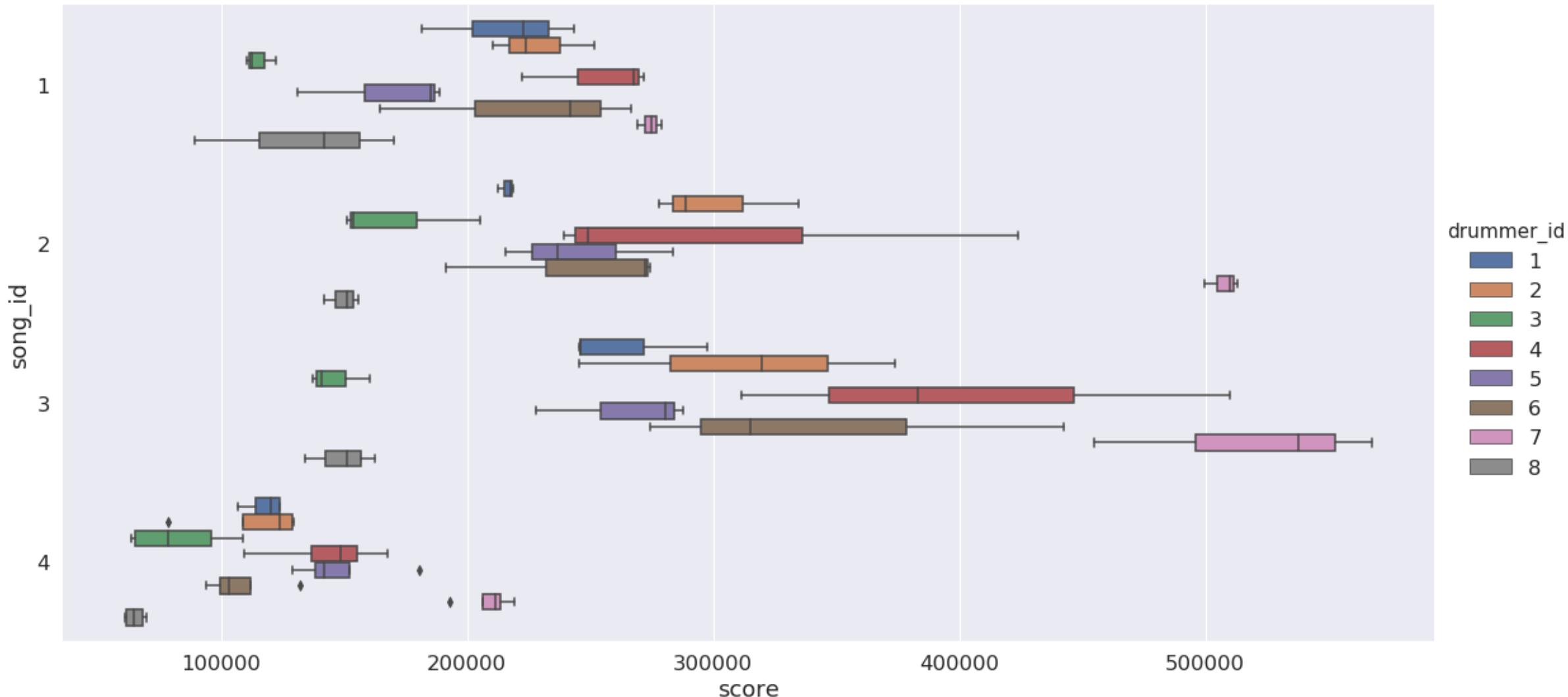


Score of Time Series

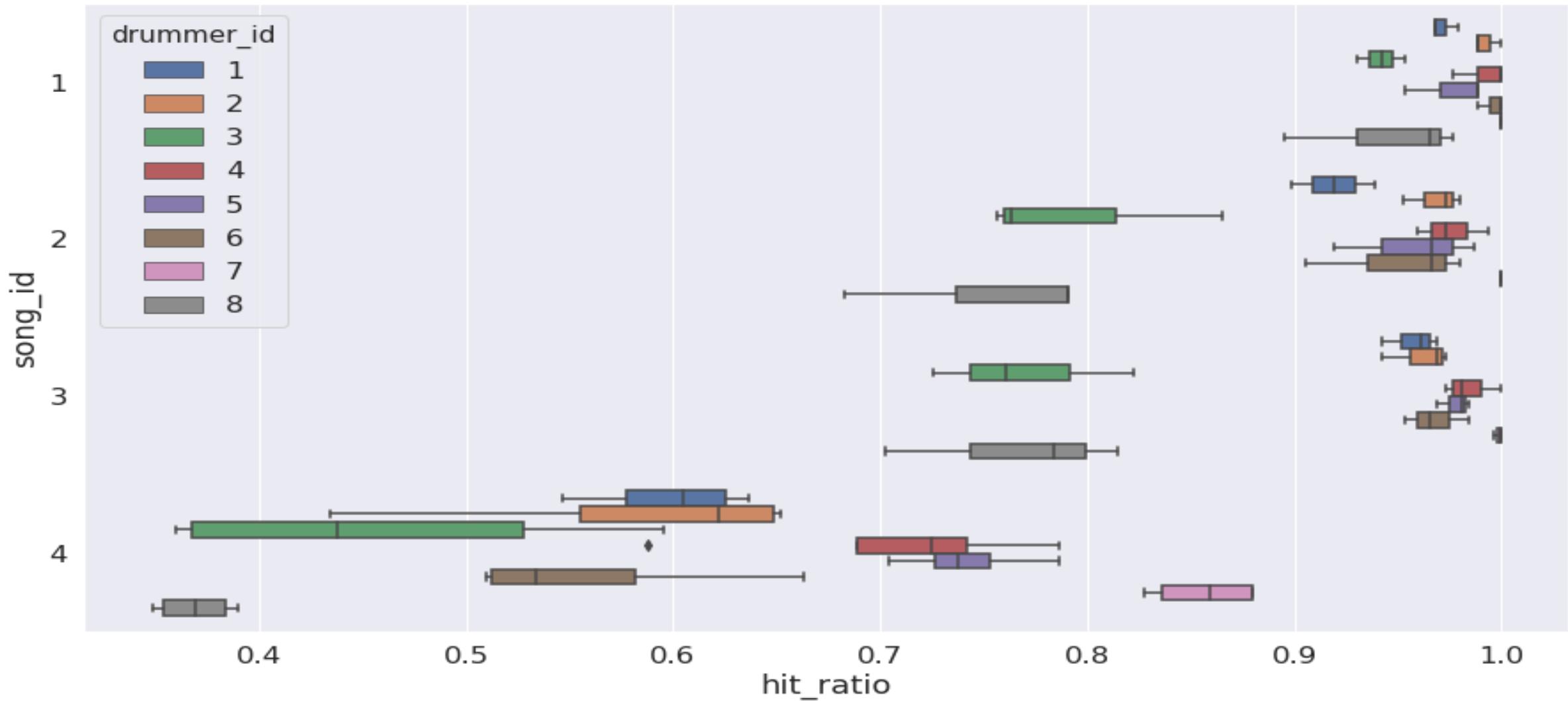
This information only gives a reference and not be put into model.



Score



Hit ratio

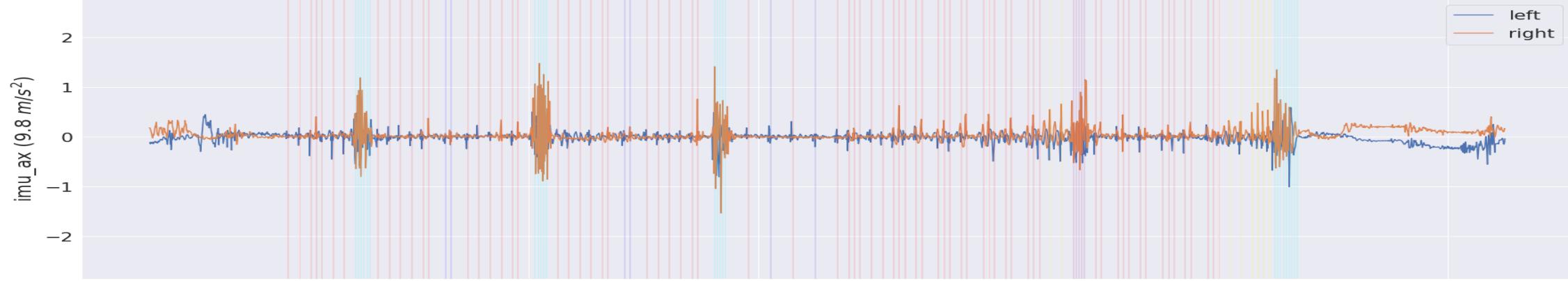


PLAY DATA

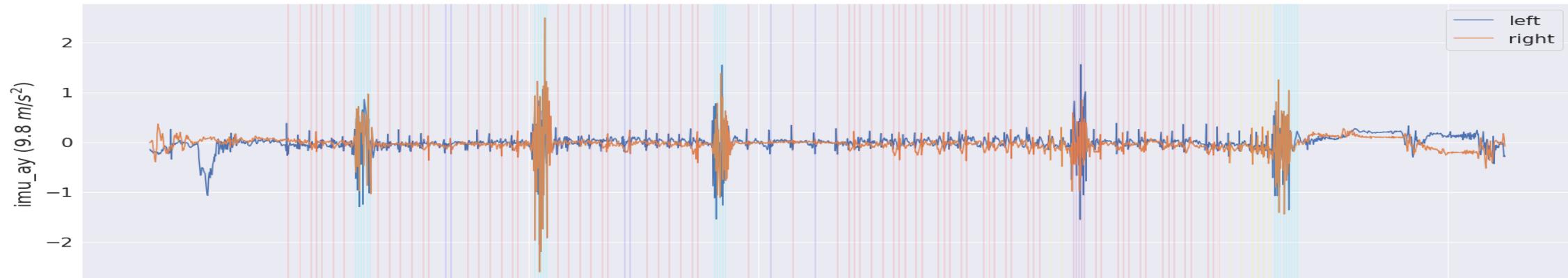


Raw data (Acc.)

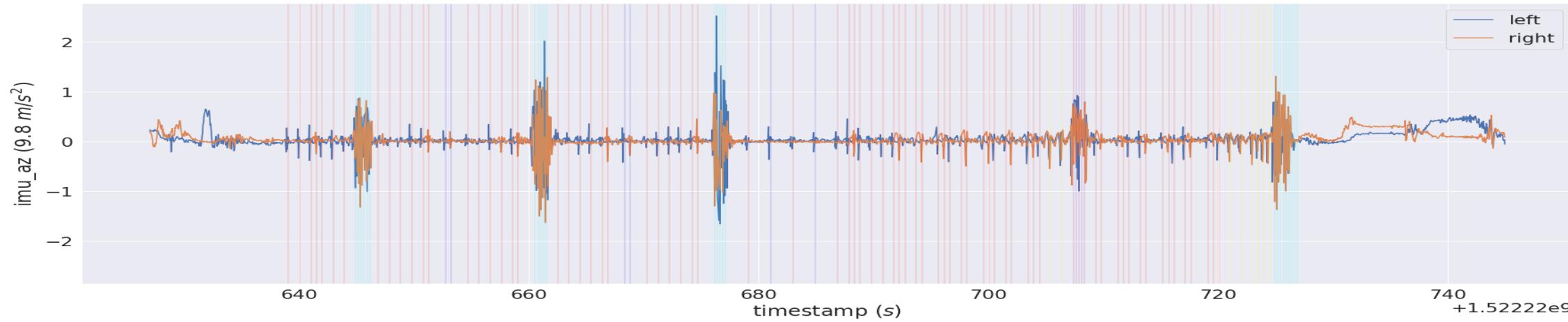
AX



AY

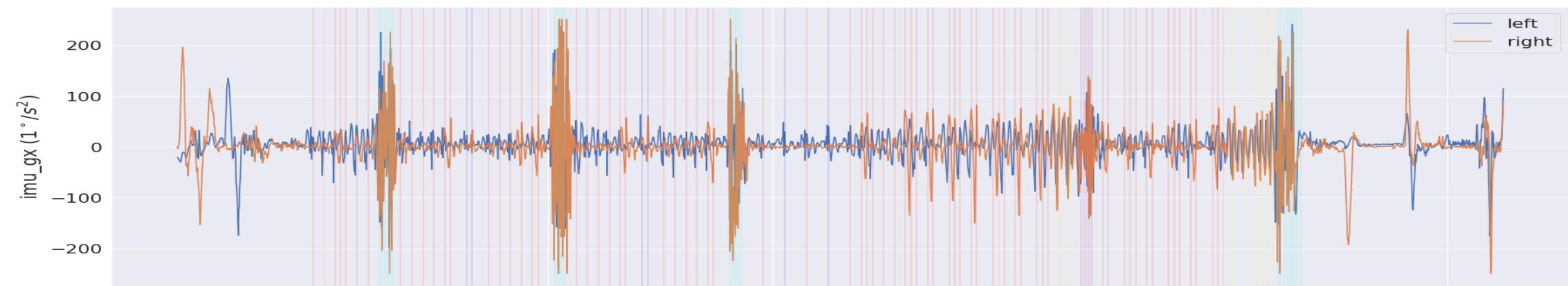


AZ

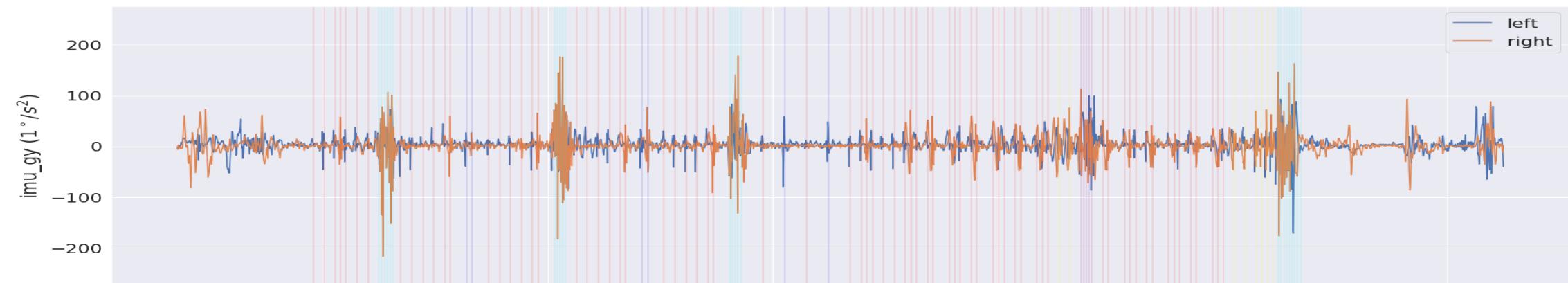


Raw data (Gyr.)

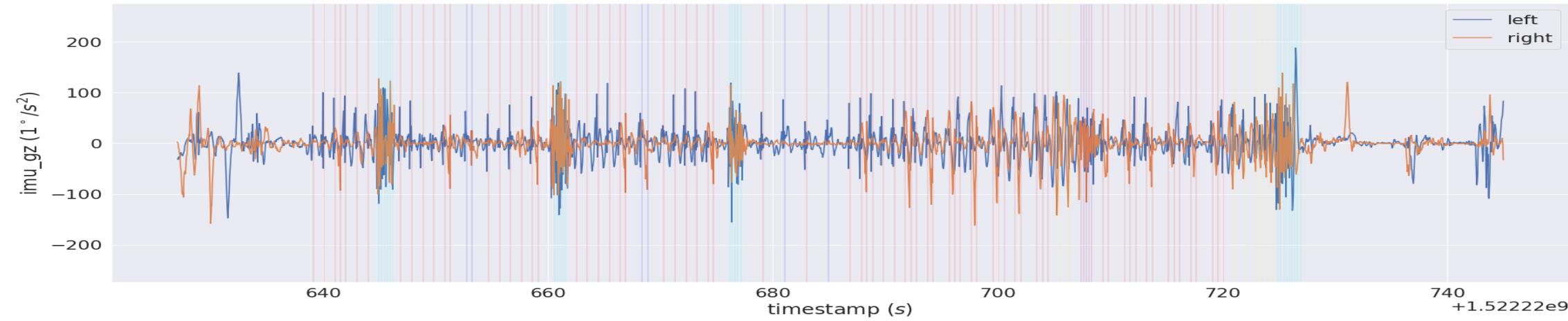
GX



GY



GZ



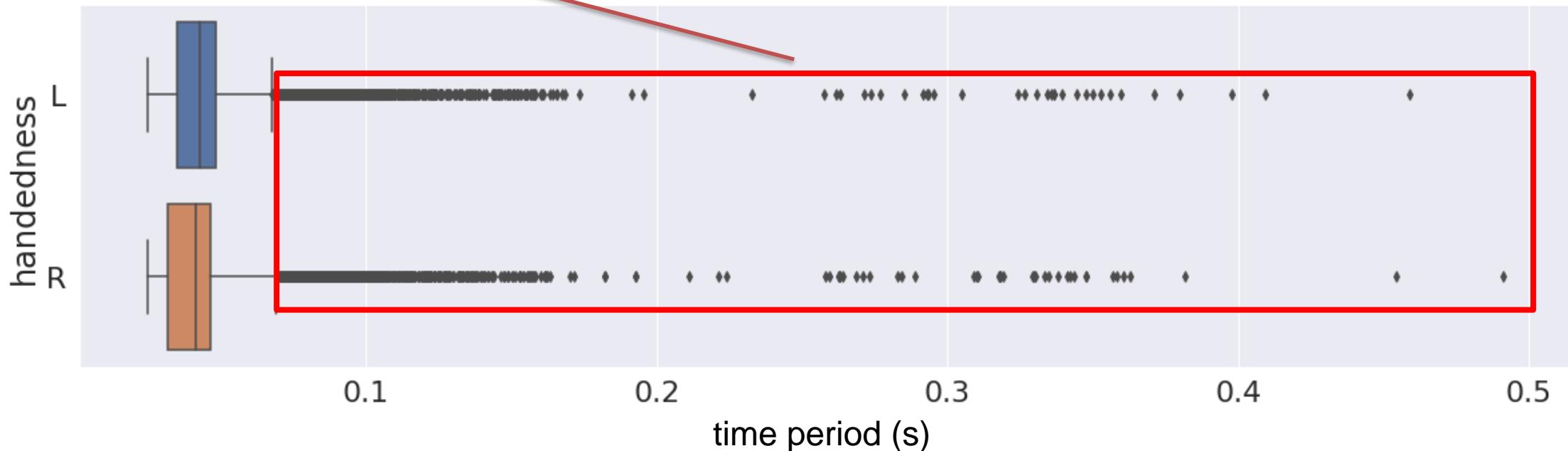
timestamp (s)

+1.52222e9

Sample rate

Since the minimal sample comes in **0.025 sec** time periods, then I **resample** raw data by **0.02 sec** time periods.

Notice there is unexpected sample rate during playing.



	L	R
count	21203.000000	22031.000000
mean	0.044498	0.042832
std	0.018627	0.019122
min	0.025040	0.024680
25%	0.034891	0.031506
50%	0.042842	0.041259
75%	0.048061	0.046604
max	0.459025	0.491320

PRIMITIVE



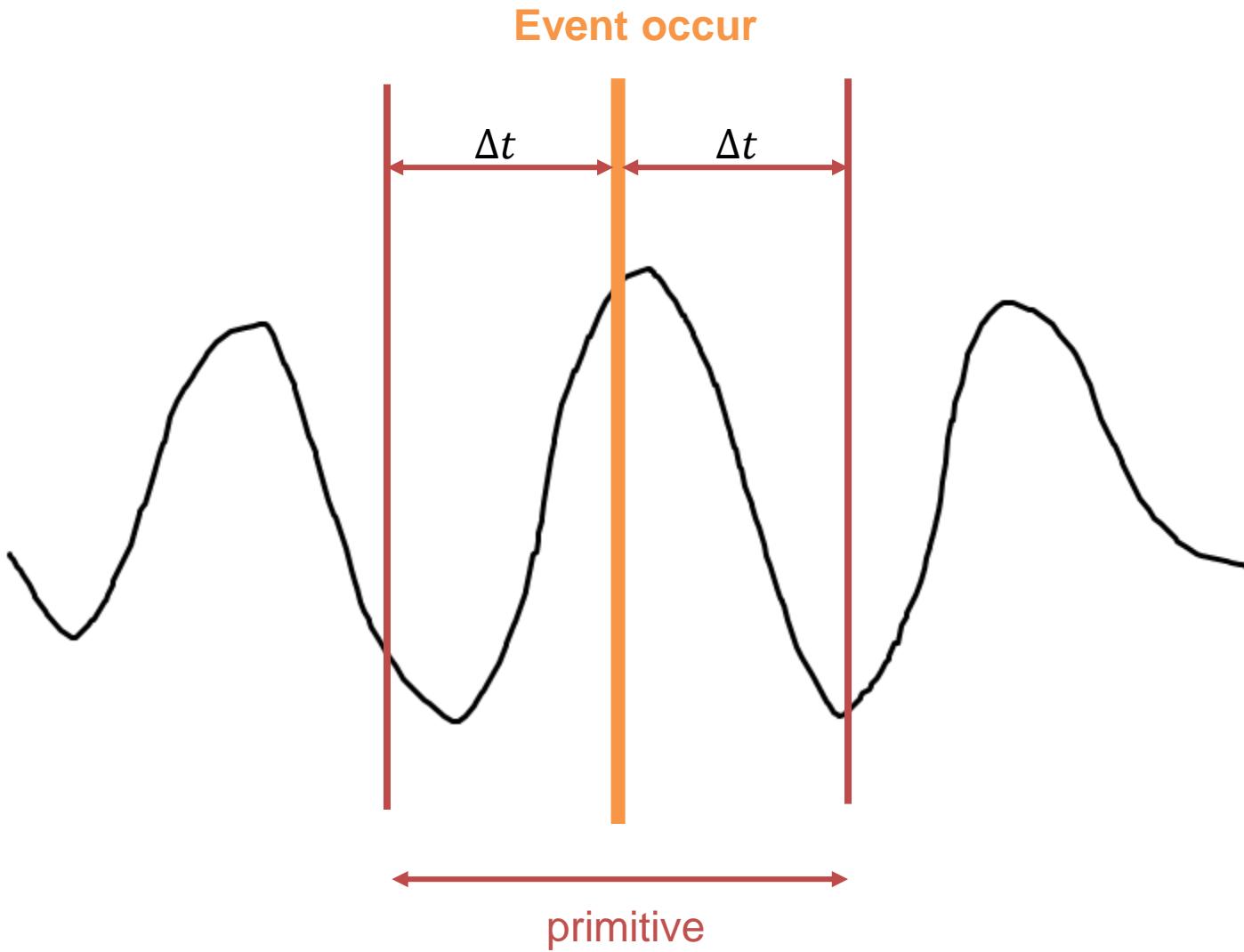
Hit event

This means we hit on time to get a “good” rate.
And the time called a **hit event**.



A frame of screenshot during playing

Primitive



FEATURE ENGINEERING



Feature extraction

Feature format: [Prefix]_[Middle]_[Suffix]

Prefix	Description
L	The signal collected from the left hand
R	The signal collected from the right hand

Sensor	Middle	Description
Accelerometer	A	The Euclidean distance of accelerometer signals in X, Y, Z axes
	AX	The X -axis signal from accelerometer
	AY	The Y -axis signal from accelerometer
	AZ	The Z -axis signal from accelerometer
Gyroscope	G	The Euclidean distance of gyroscope signals in X, Y, Z axes
	GX	The X -axis signal from gyroscope
	GY	The Y -axis signal from gyroscope
	GZ	The Z -axis signal from gyroscope

A primitive maps onto a row of features

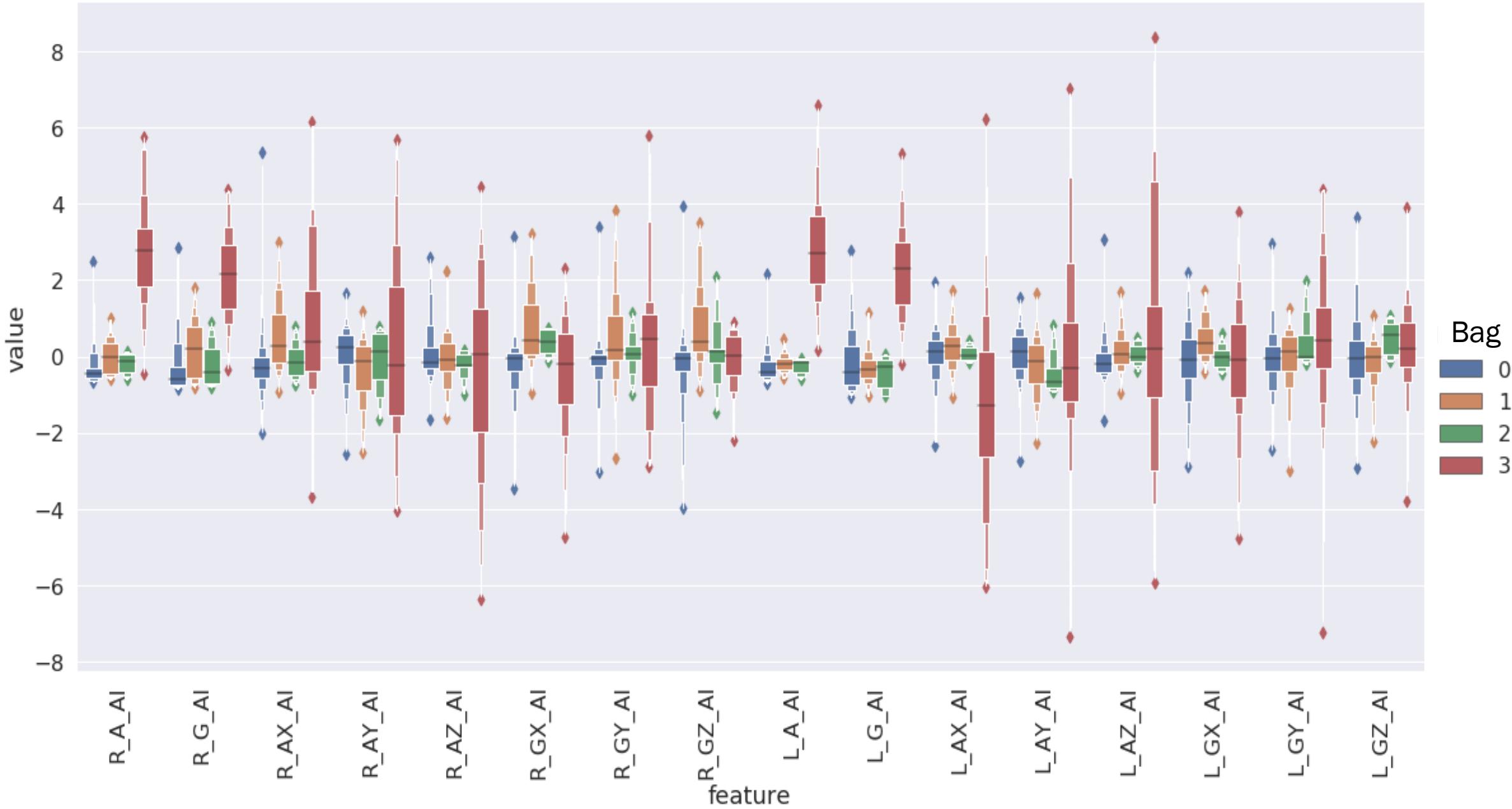
Feature extraction

Suffix	Description
AI	The mean of intensity
VI	The variance of intensity
MMI	The median of intensity
SMA	The signal magnitude area
AE	The average of energy
SDI	The standard deviation of intensity
IQR	The interquartile range of intensity
FR	The full range of intensity
FFT_COEF	The coefficient of fast Fourier transformation in frequency domain
MDCR	The total number of times the signal changes form below median to above median or vice versa
MCR	The total number of times the signal changes form below average to above average or vice versa
ZCR	The total number of times the signal changes form positive to negative or negative to positive or vice versa
DTW	The similarity between two signals
CORR	The Poisson correlation between two signals

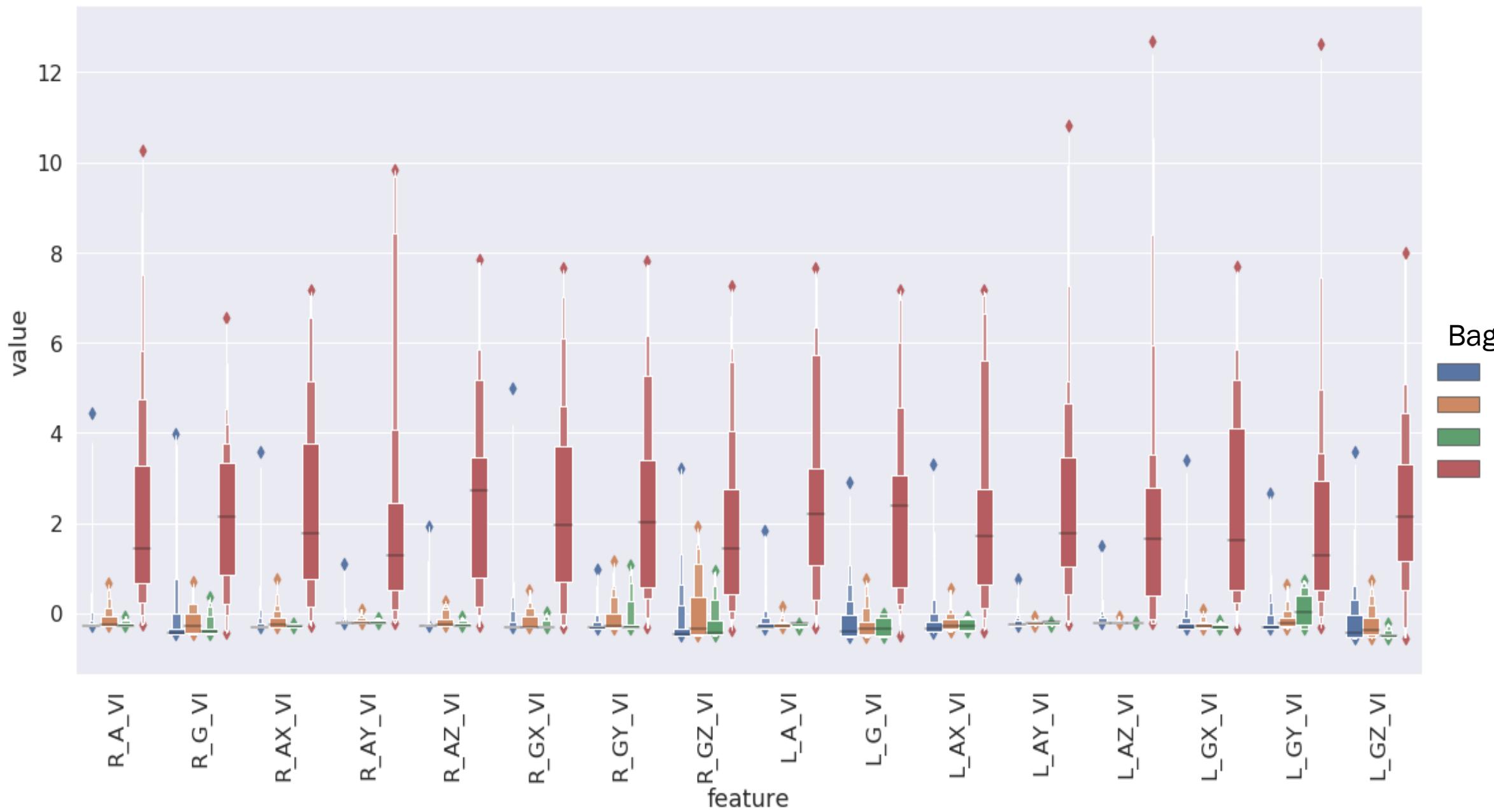
Feature examples

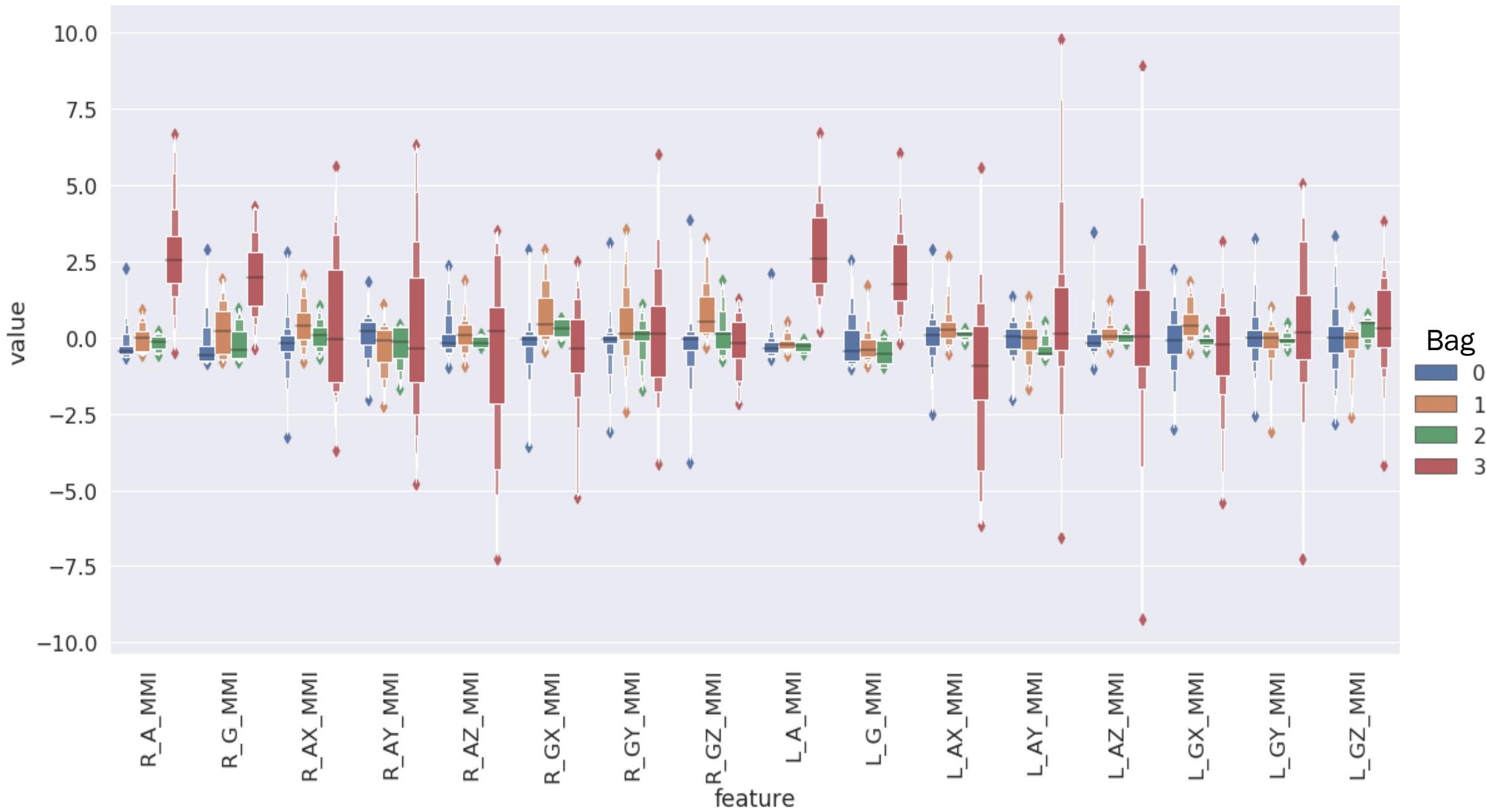
Subject	Who = 7, Song = 1, Order = 3
Group	dong_ka

AI

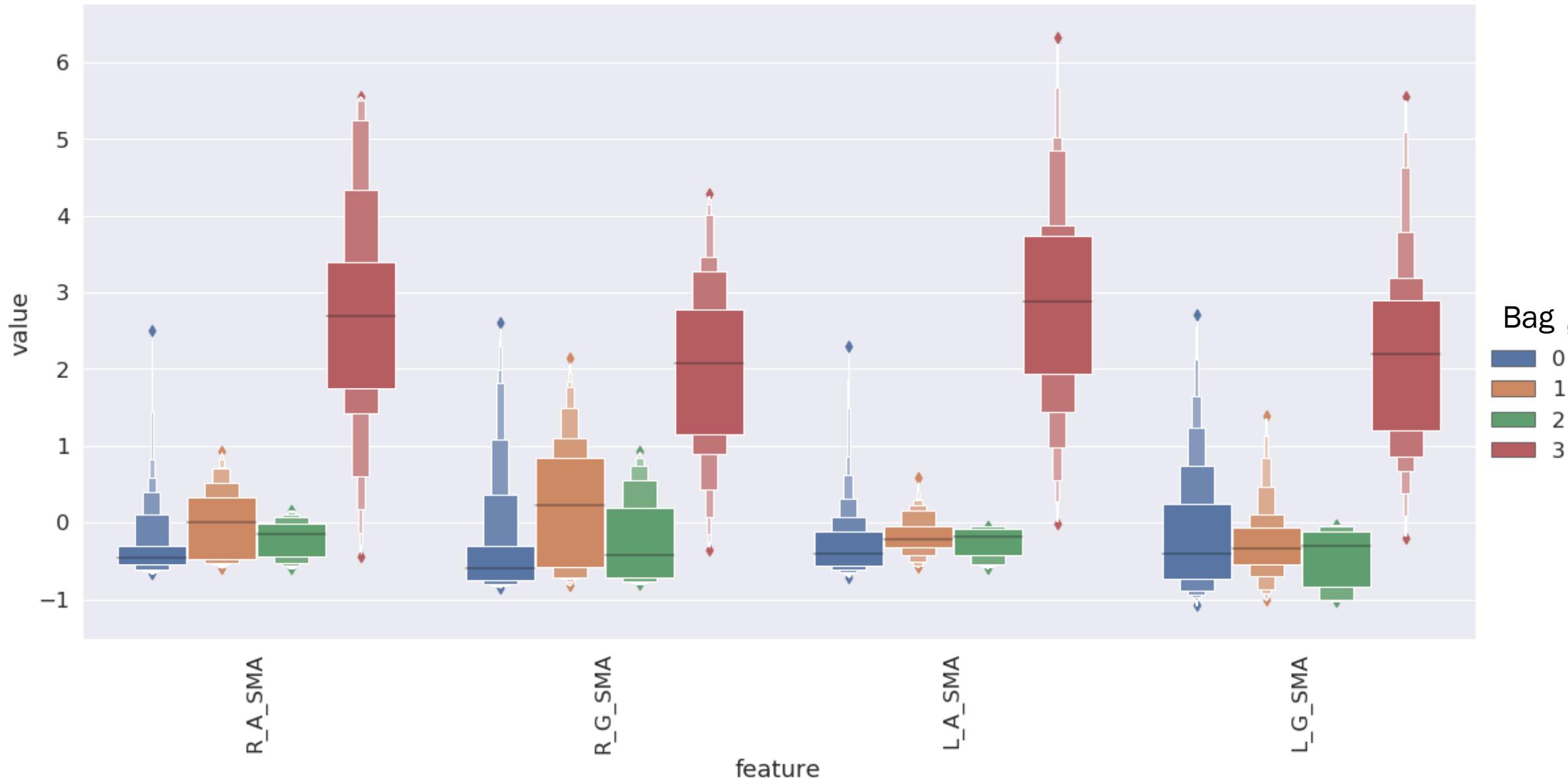


IA

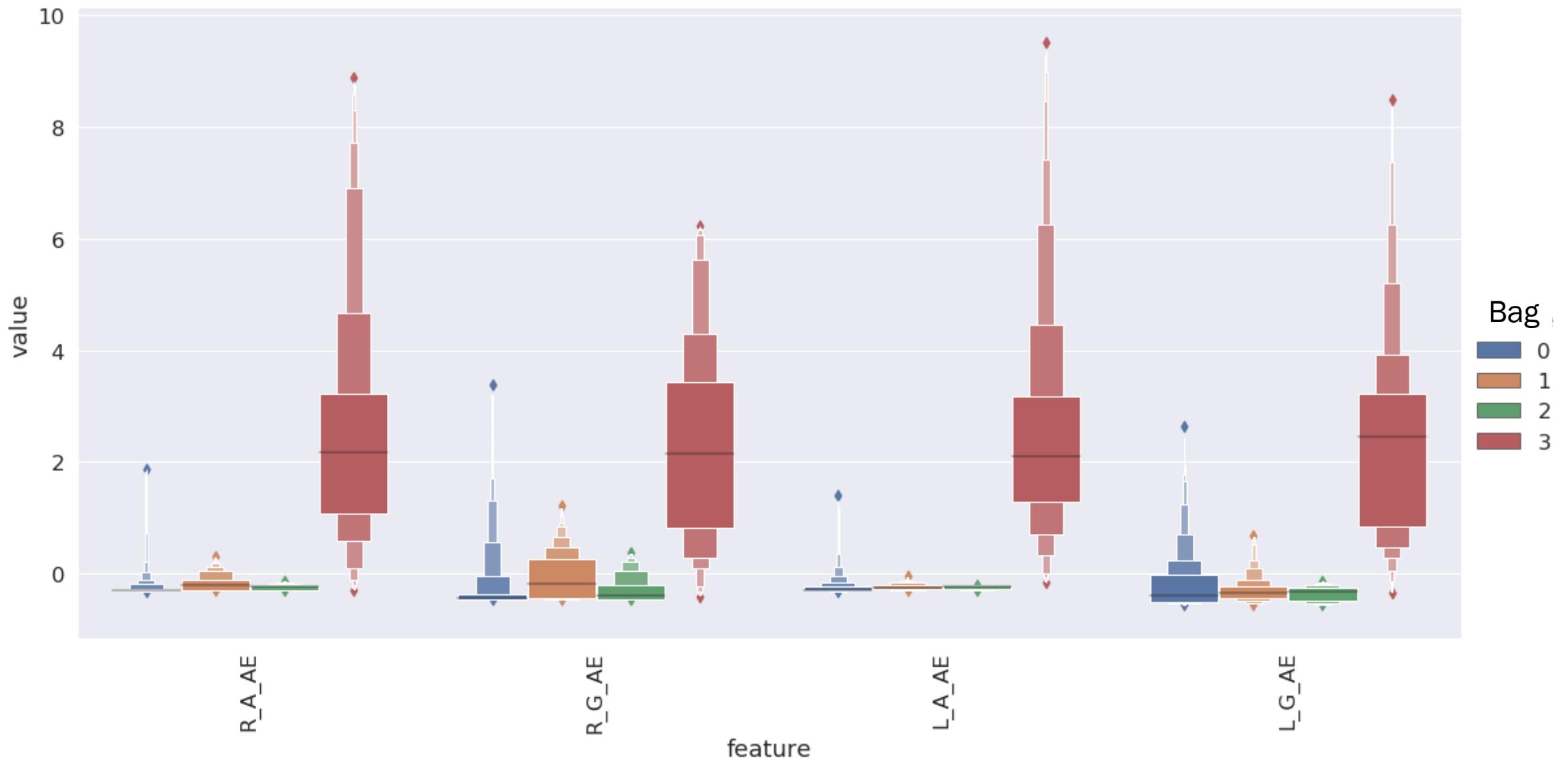


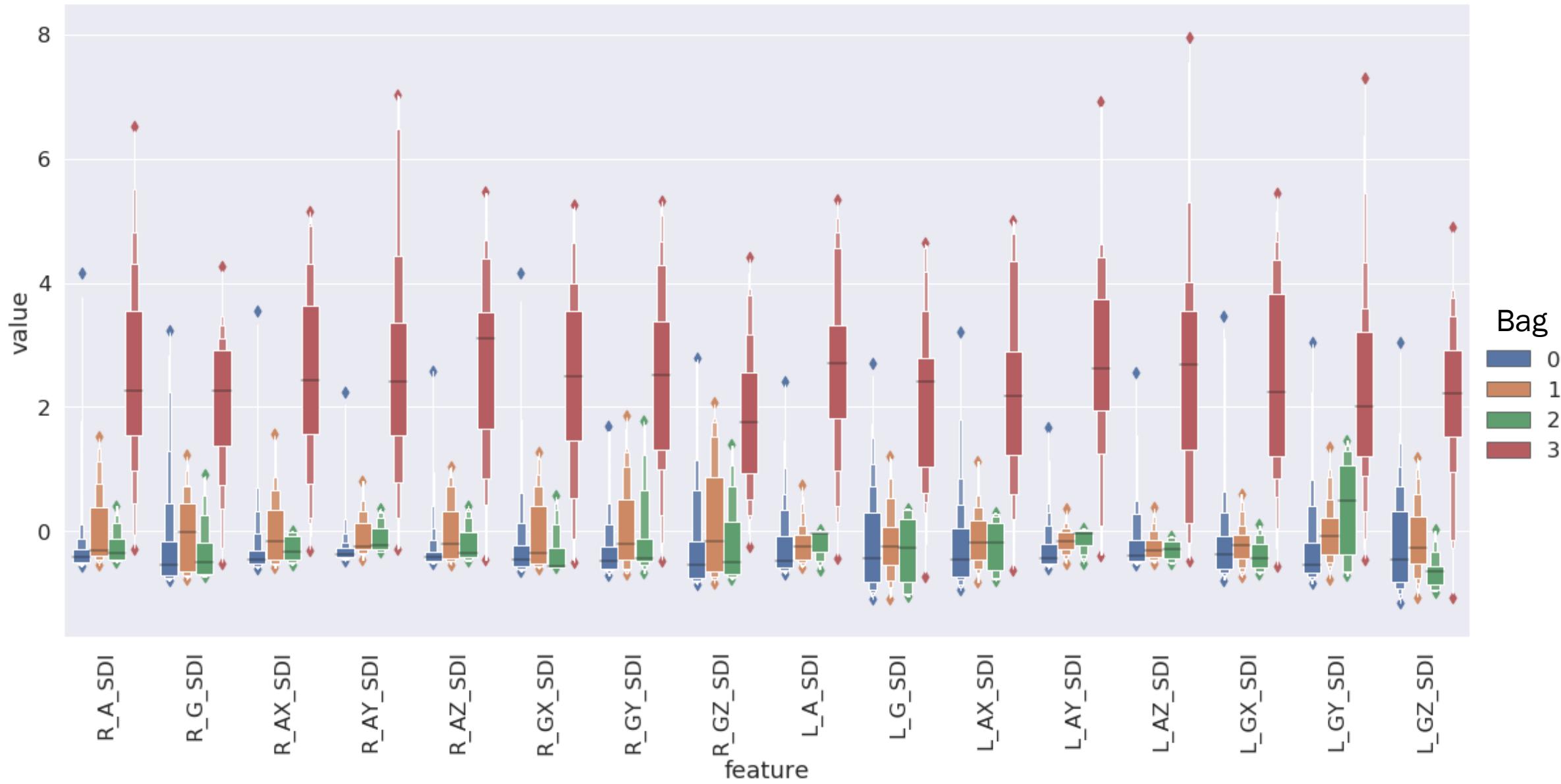


MAWS

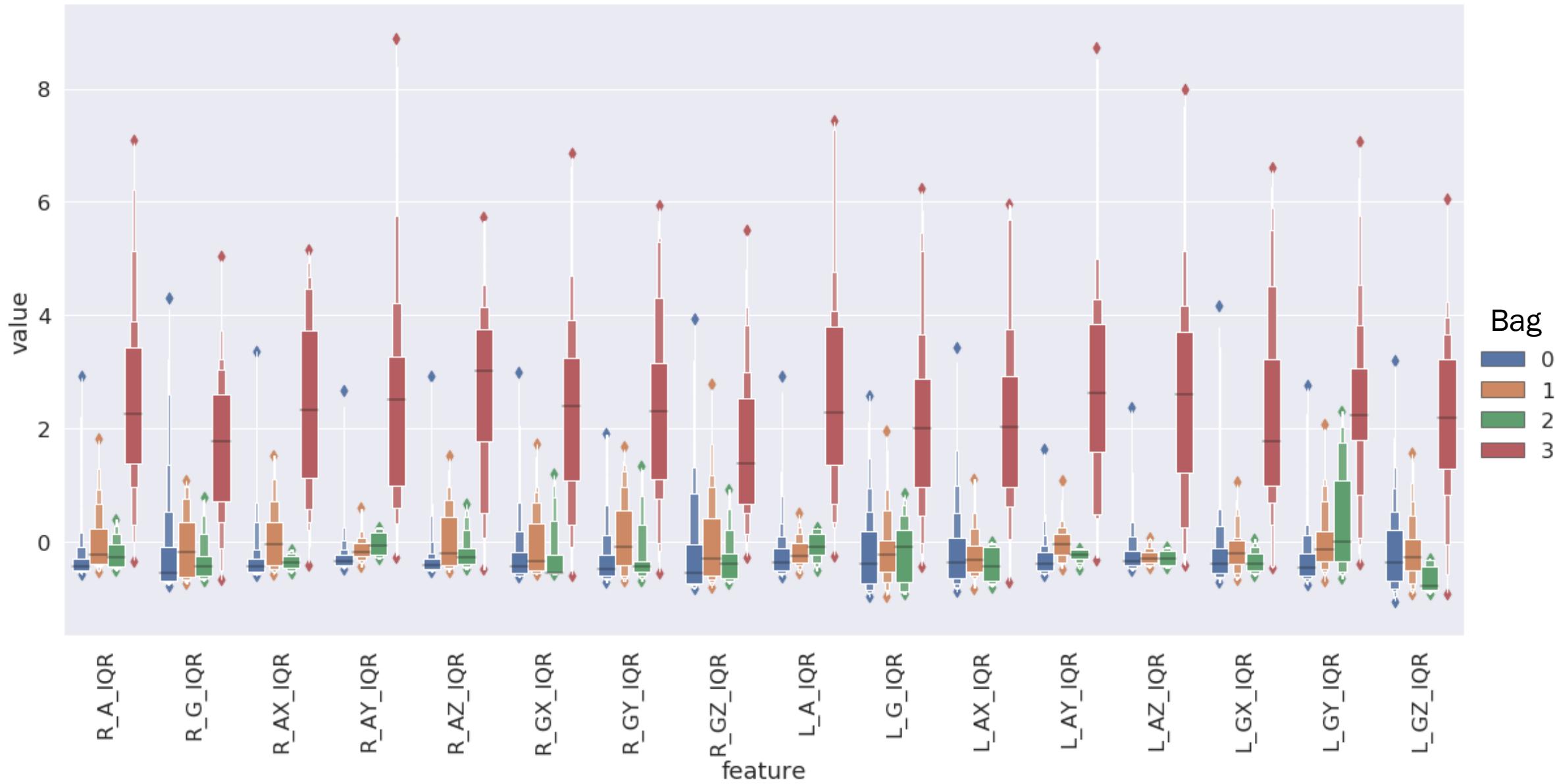


AE

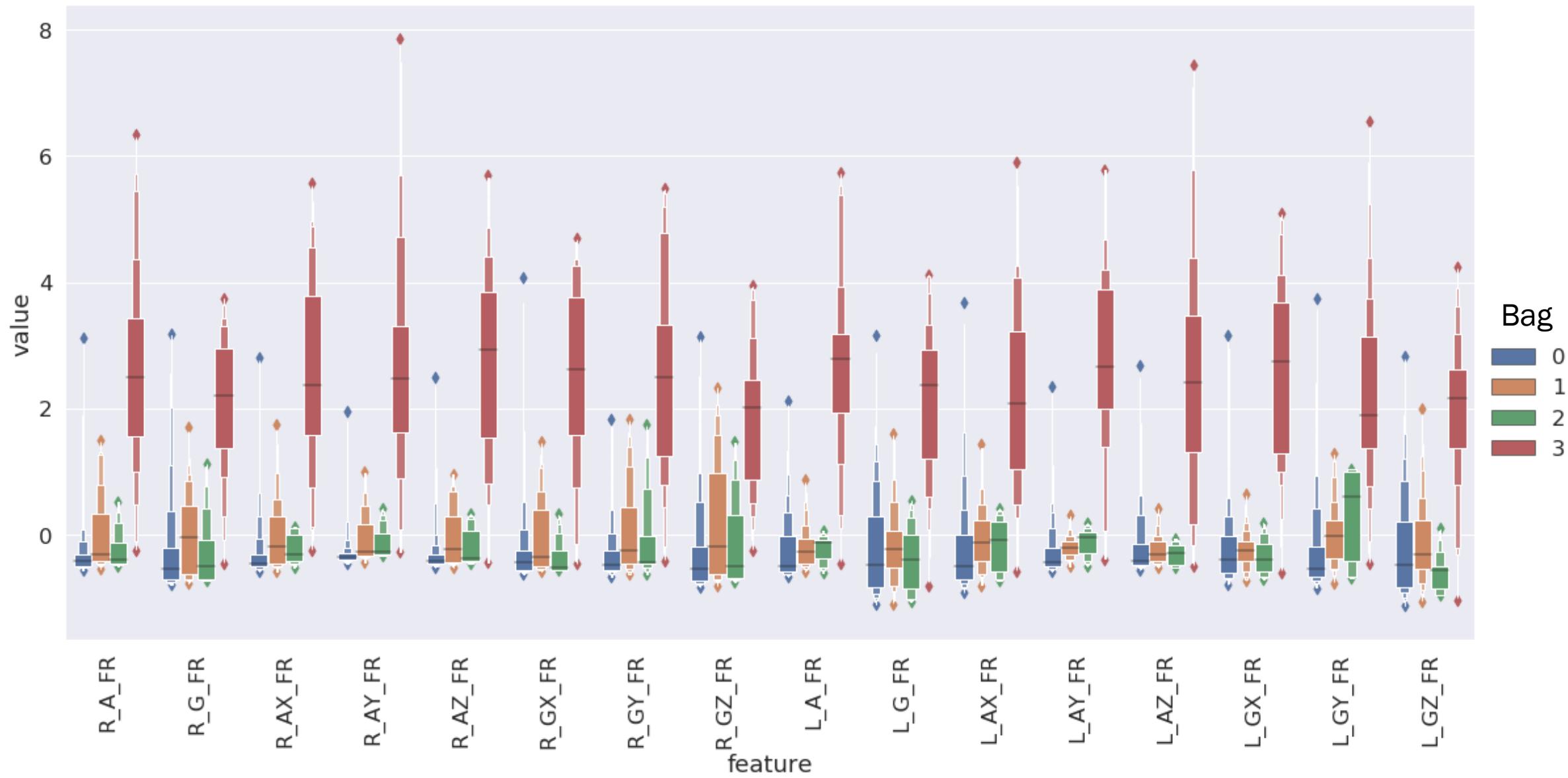




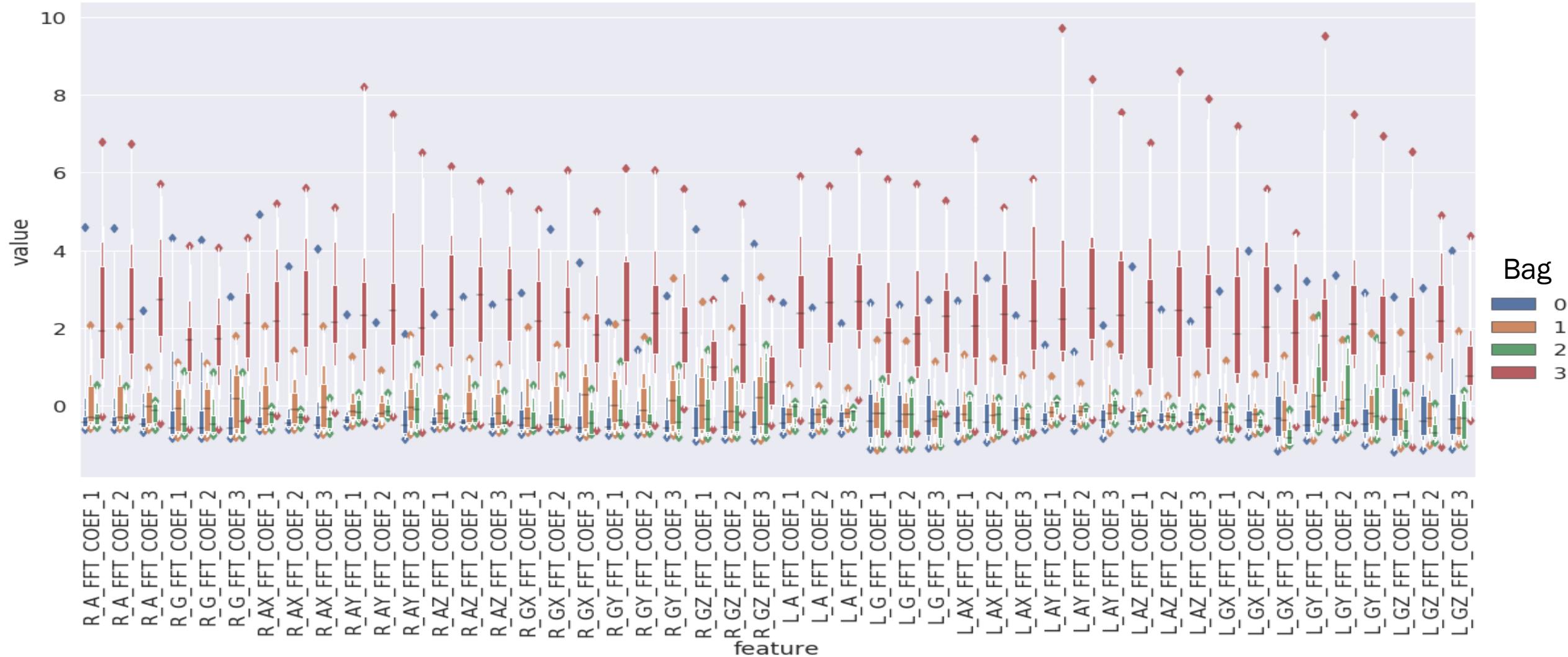
IQR



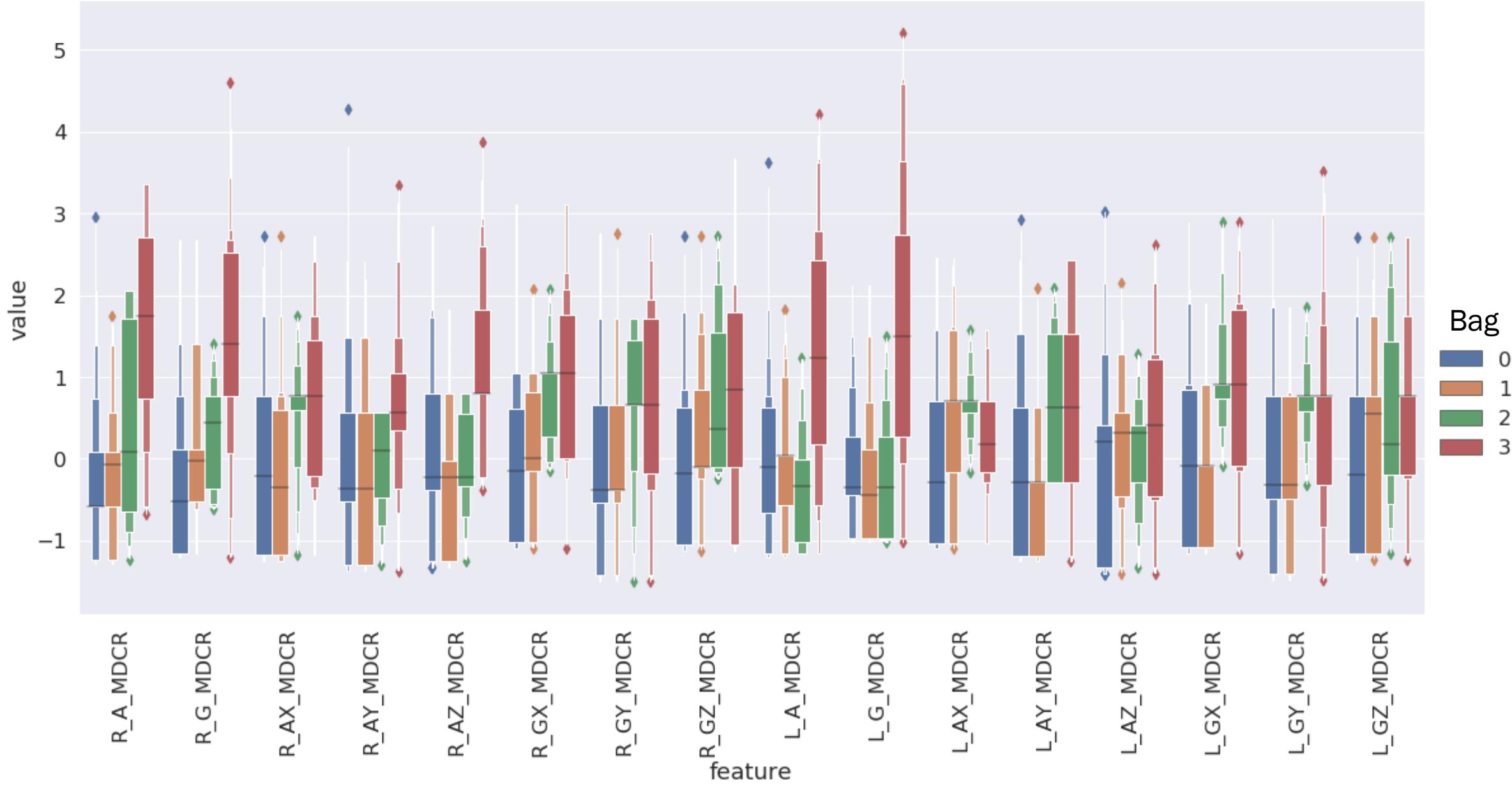
FR



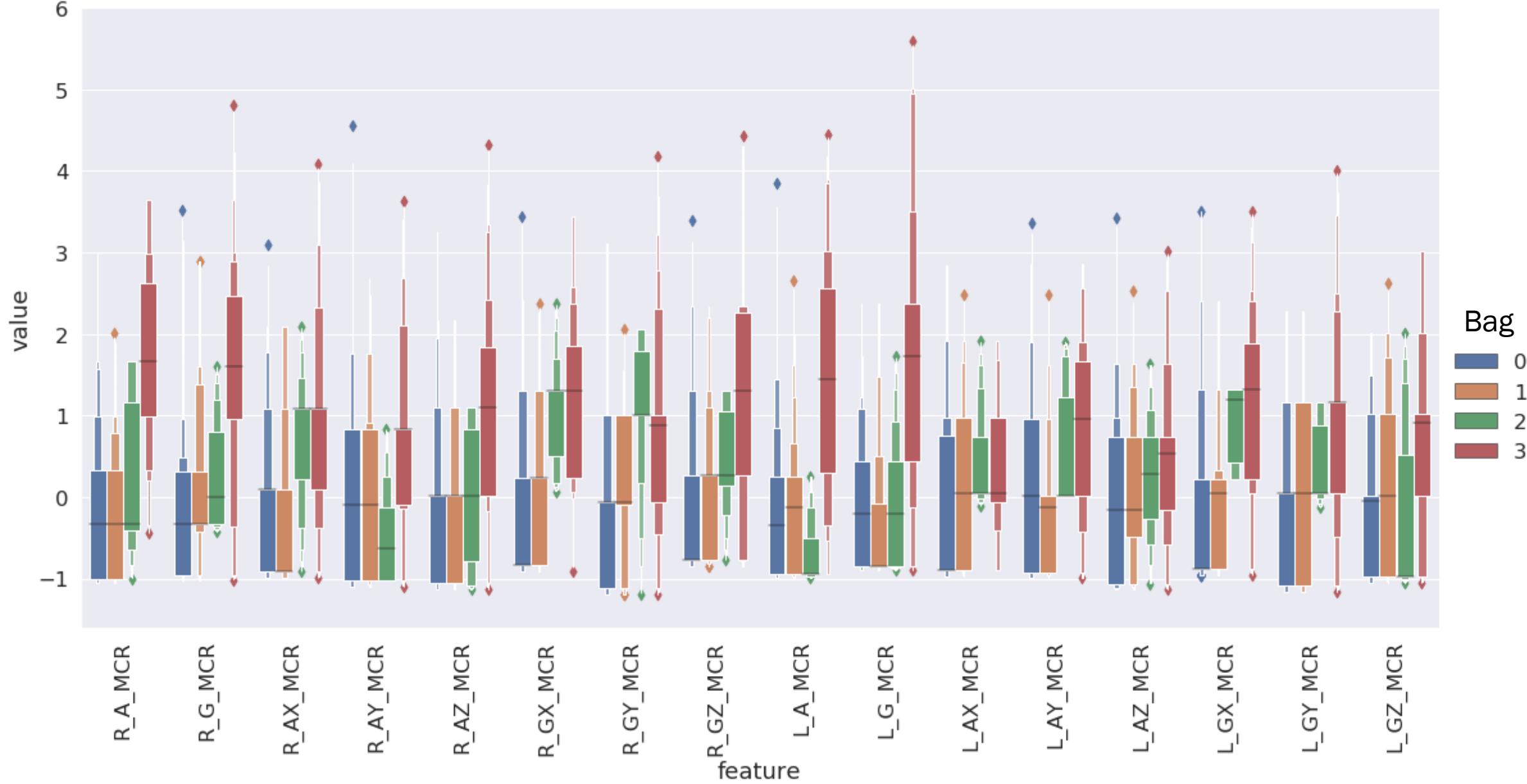
FFT_COEF



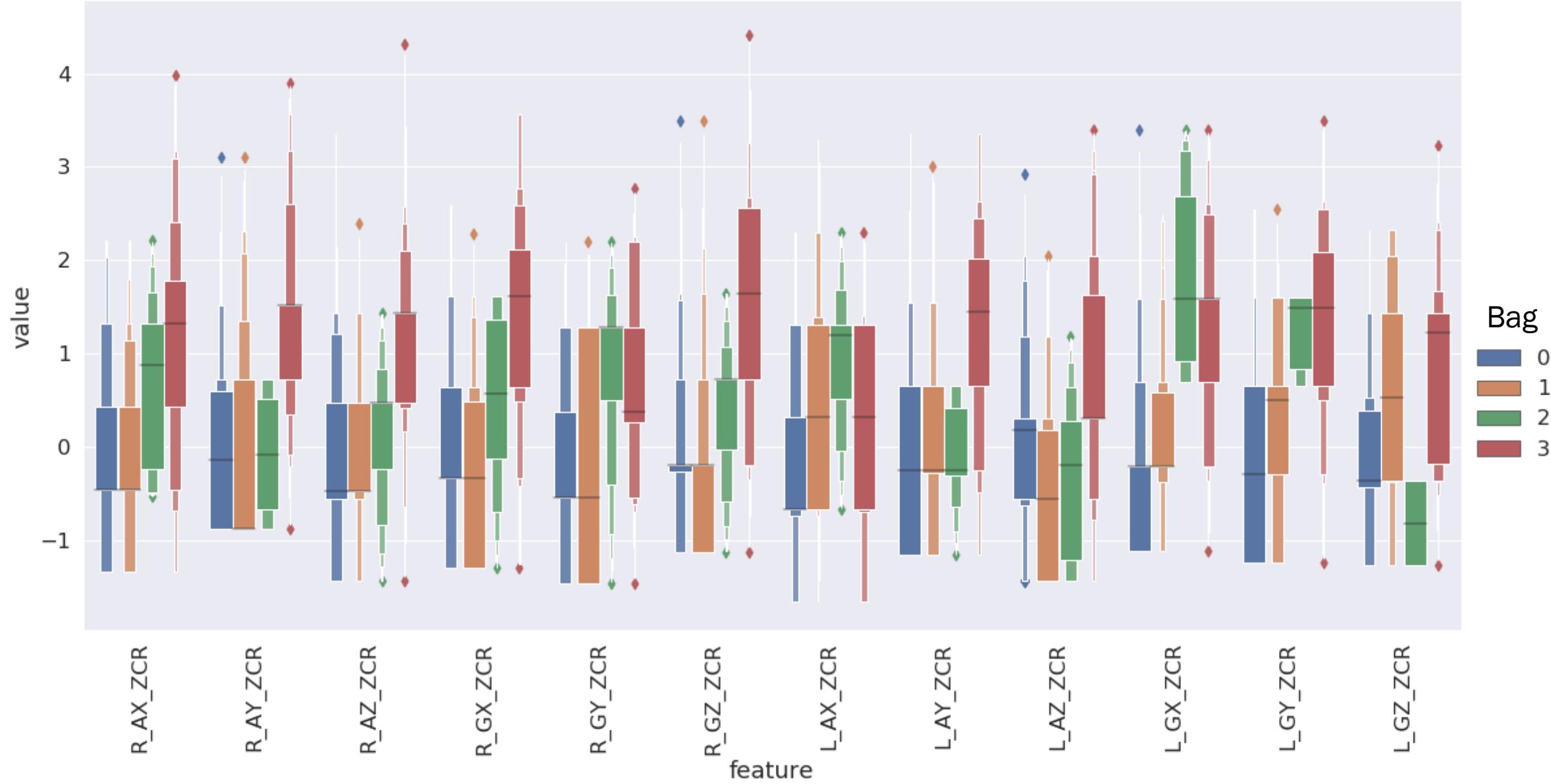
MDCR



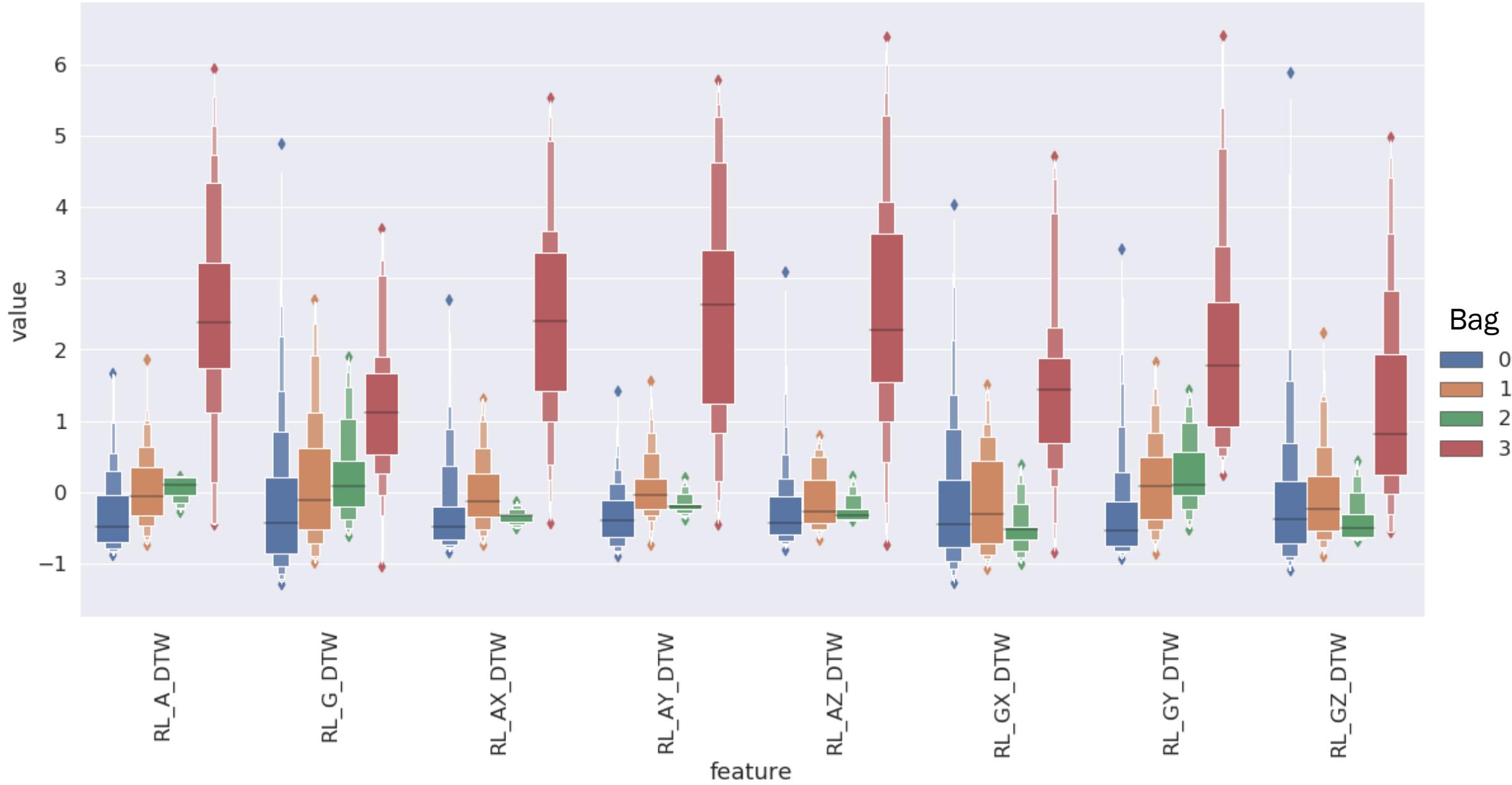
MCR



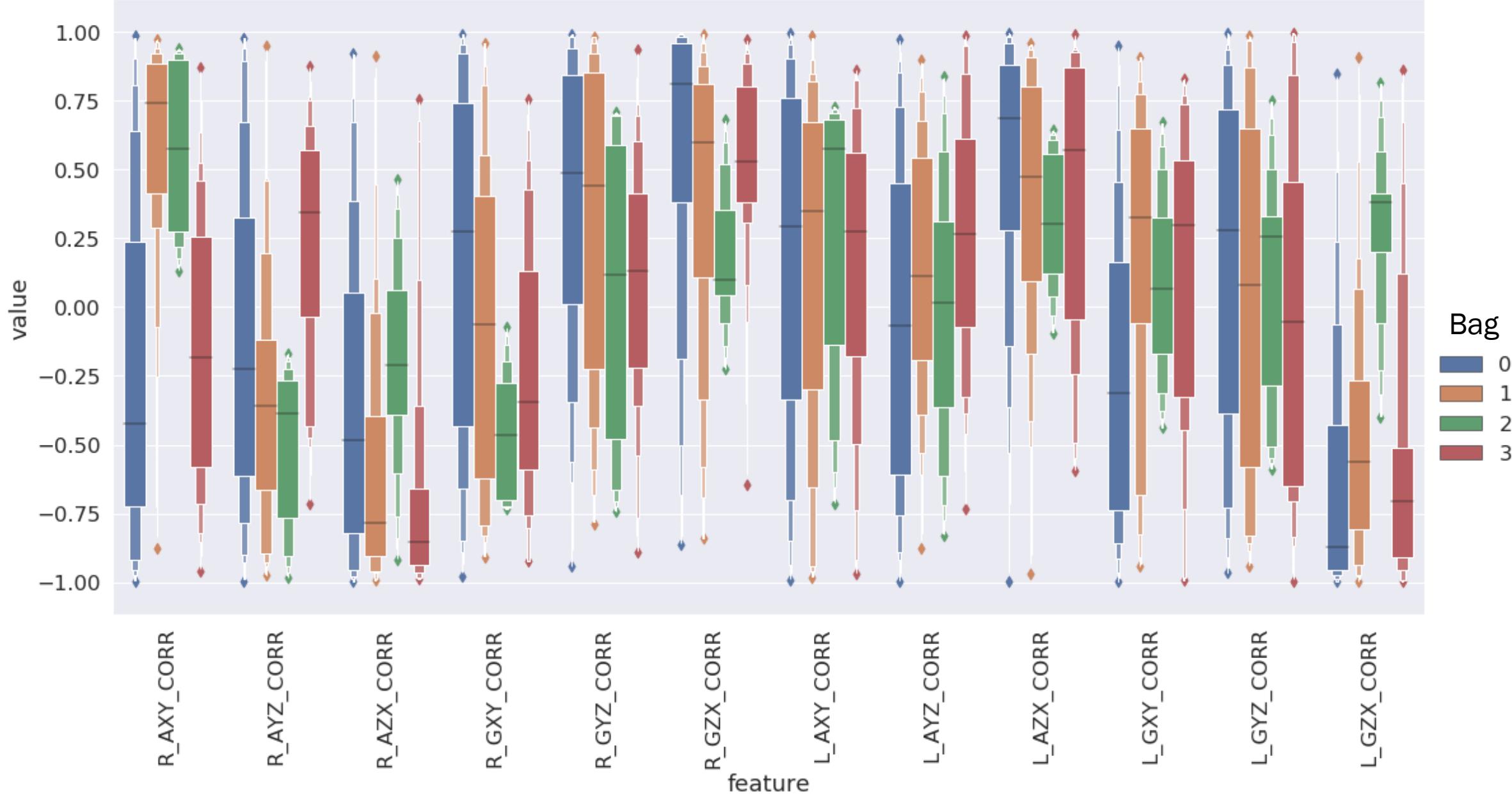
ZCR



WTD



CORR



SCALE



Scaling rule

We use z-score standard scale; i.e., $Z = \frac{x-\mu}{\sigma}$

Suffix	Scale or not
AI	Yes
VI	Yes
MMI	Yes
SMA	Yes
AE	Yes
SDI	Yes
IQR	Yes
FR	Yes
FFT_COEF	Yes
MDCR	Yes
MCR	Yes
ZCR	Yes
DTW	Yes
CORR	No

MODE SCENARIO



In the **same** song

one-to-one

Train Test

Example 1

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

Example 2

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

In the **same** song

rest-to-one

Train Test

Example 1

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

Example 2

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

In the **same** song

all-to-one

Train Test

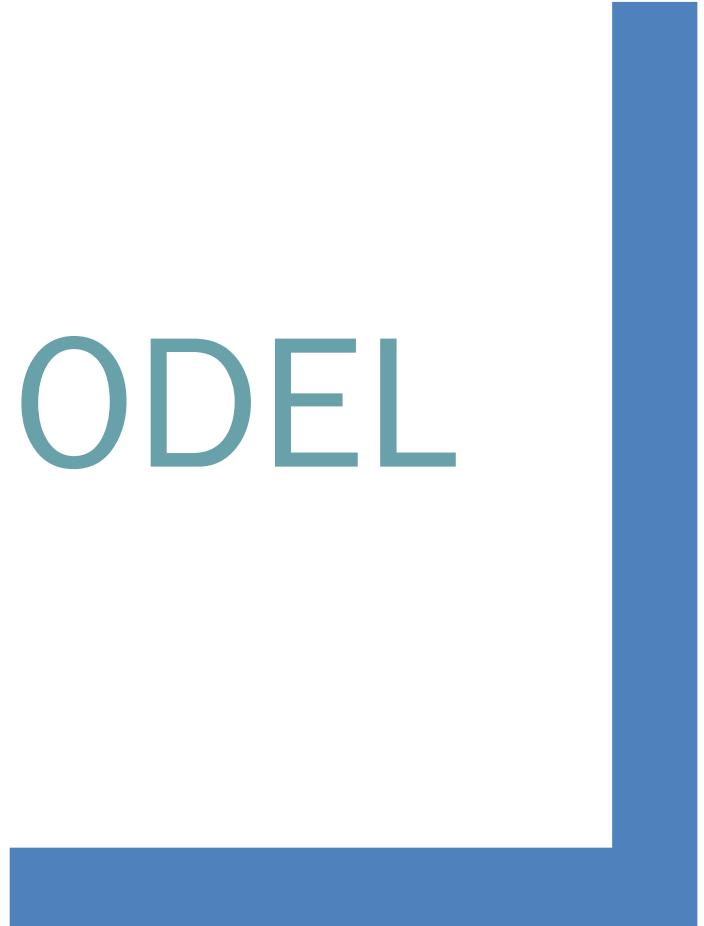
Example 1

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

Example 2

Order \ Who	1	2	3	4	5	6	7	8
1								
2								
3								
4								

MODEL



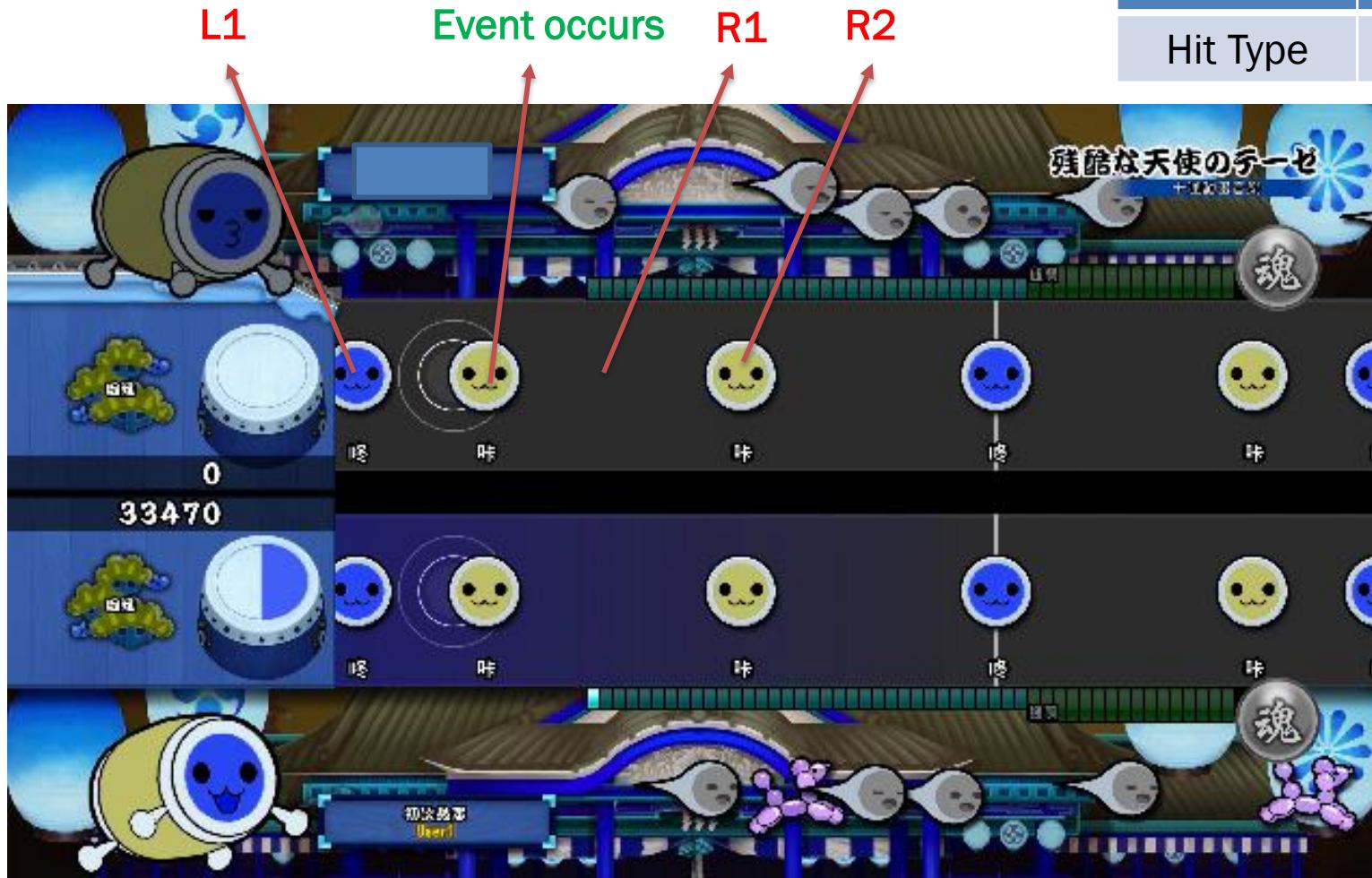
Model used

Model Abbr.	Description	Python Package
LGBM	light gradient boosting machine	lightgbm 2.0.10
KNN	k-nearest neighbors	scikit-learn 0.19.1
Linear SVM	support vector machine with linear kernel	scikit-learn 0.19.1
RBF SVM	support vector machine with RBF kernel	scikit-learn 0.19.1

They are all **supervised** learning.

LGBM special feature “near”

near	L1	R1	R2
Hit Type	1	0	2



Sensors parameter

Parameter	Description
acc	Use data collected from accelerometer
gyr	Use data collected from gyroscope
near	Use previous and next hit type from drum note

Evaluation

All objective functions and predictions are evaluated by **weighted F1 score**.

bag	precision	recall	f1-score	support
0	0.94	0.95	0.94	233
1	0.85	0.92	0.88	84
2	1.00	0.22	0.36	9
3	0.91	0.81	0.85	36
avg / total	0.91	0.91	0.91	362

$$\frac{0.94 \times 233 + 0.88 \times 84 + 0.36 \times 9 + 0.85 \times 36}{362} \cong 0.91$$

LGBM parameter

Parameter	Value
learning_rate	0.1
application	multiclass
max_depth	4
num_leaves	16
num_boost_round	200
early_stopping_round	100

KNN parameter

Parameter	Value
K	choose best from [5, 10, 20, 50]

Linear SVM parameter

Parameter	Value
kernel	linear
C	Choose best from [0.01, 0.1, 1, 10]

RBF SVM parameter

Parameter	Value
kernel	rbf
C	Choose best from [0.01, 0.1, 1, 10]

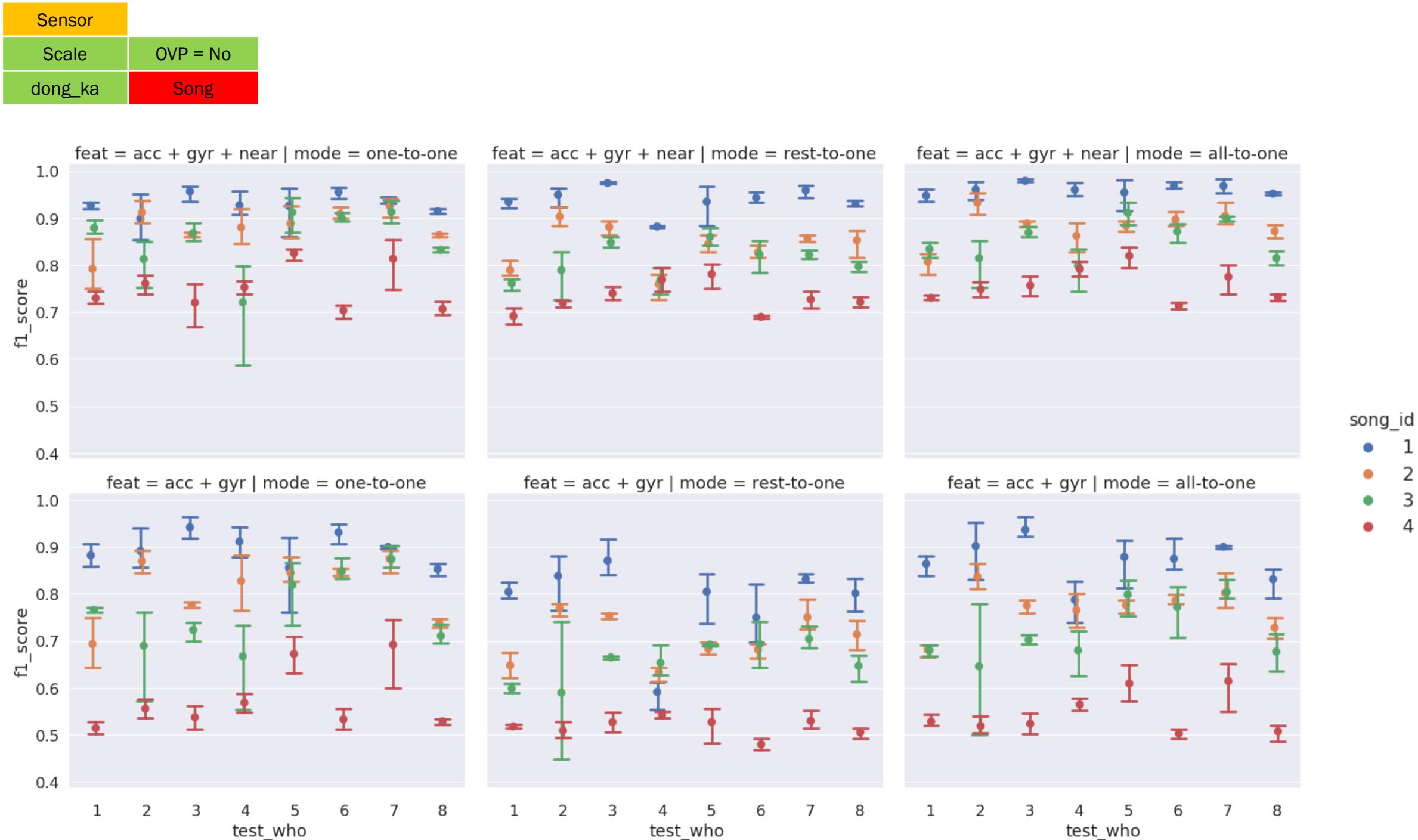
Sensor	Model
Scale	Oversampled
Group	Song

EXPERIMENT

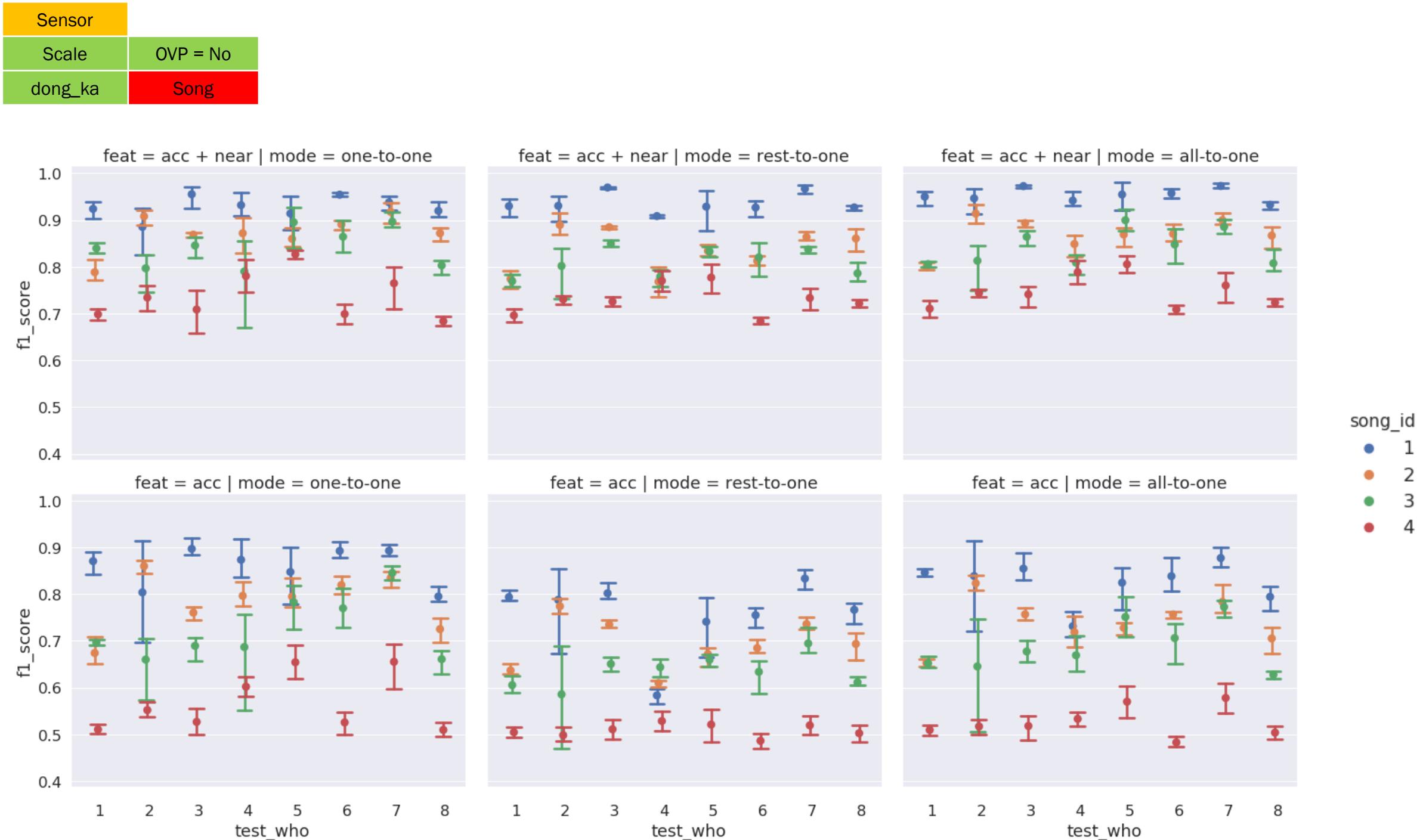
SONG PICKED



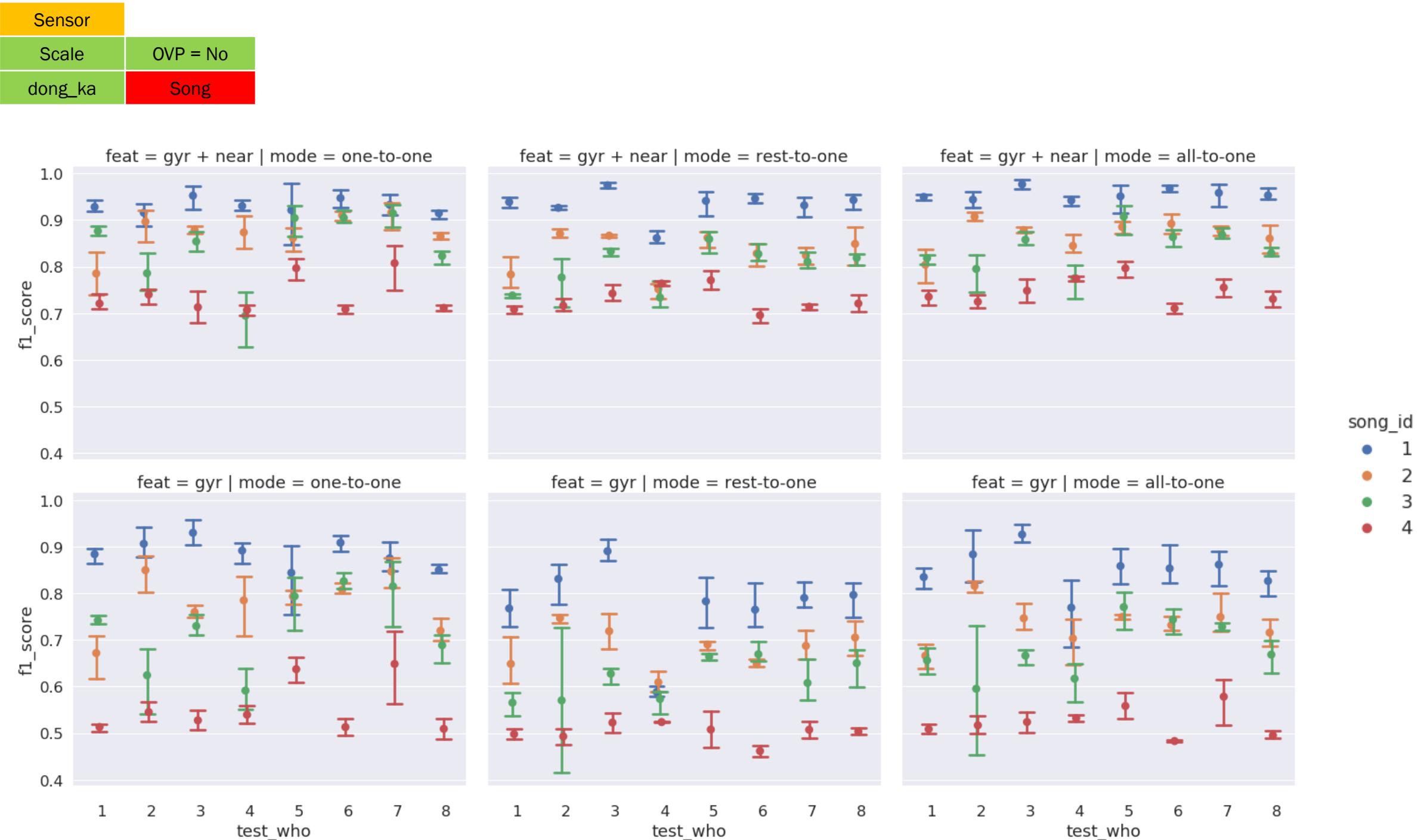
Song Picked (LGBM)



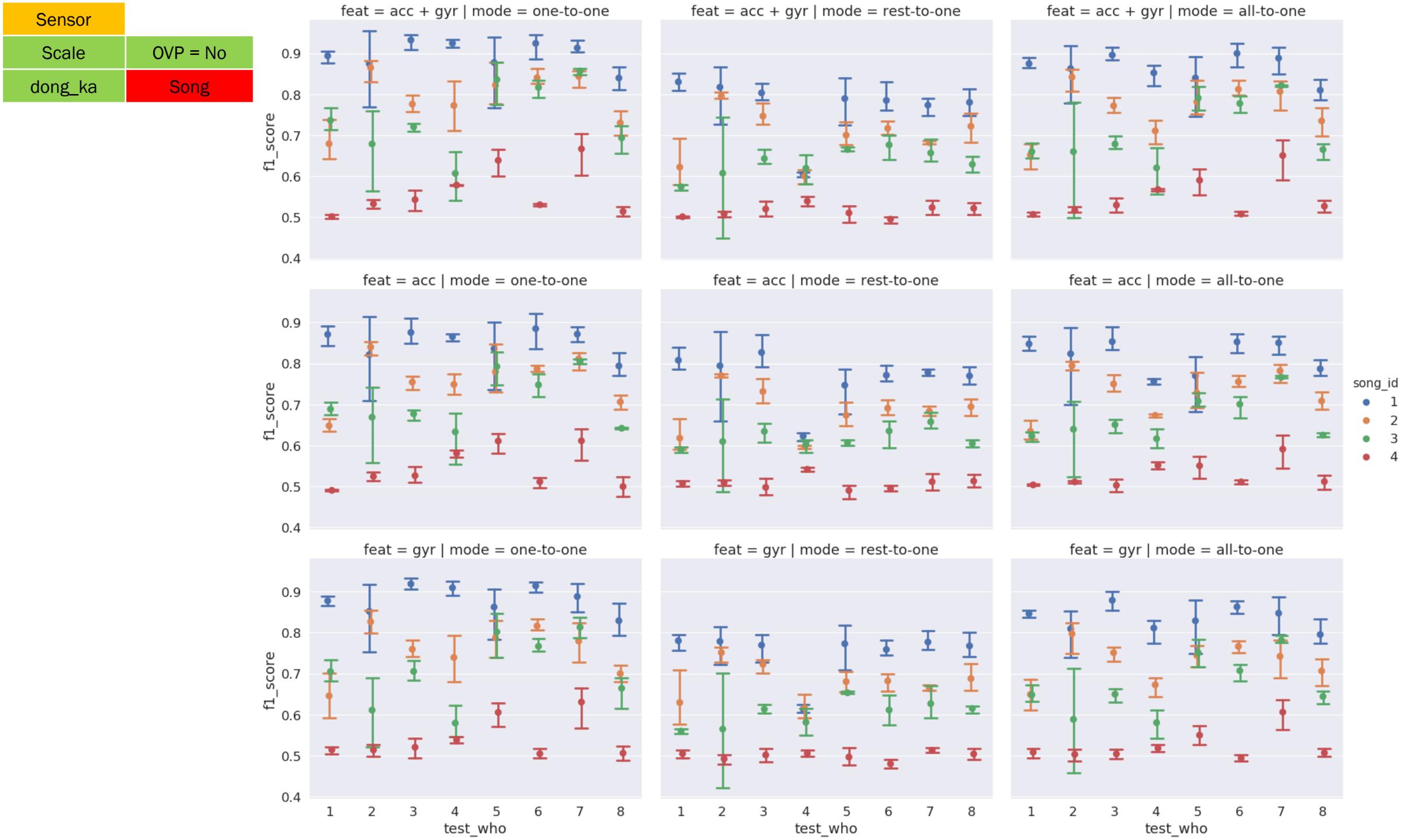
Song Picked (LGBM)



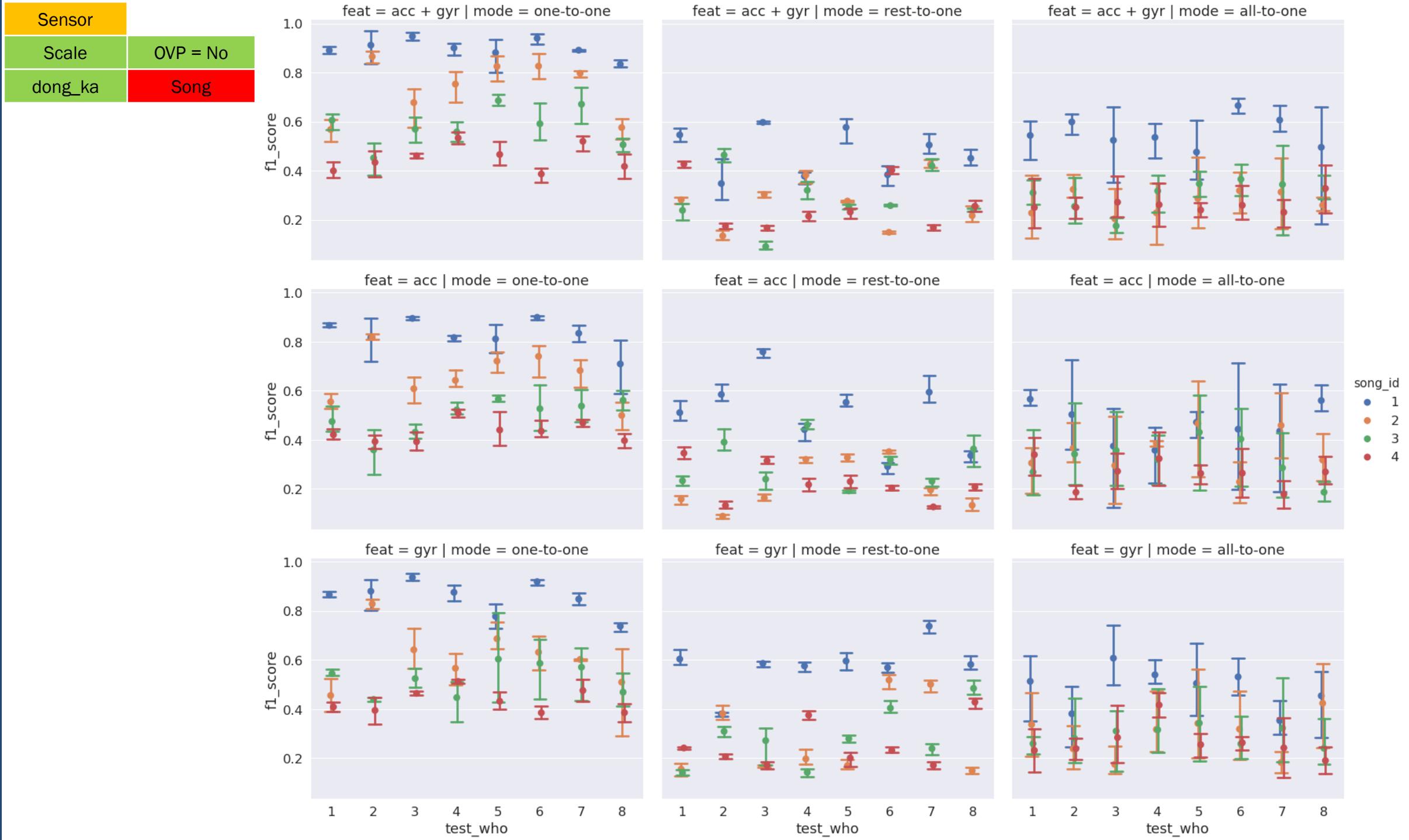
Song Picked (LGBM)



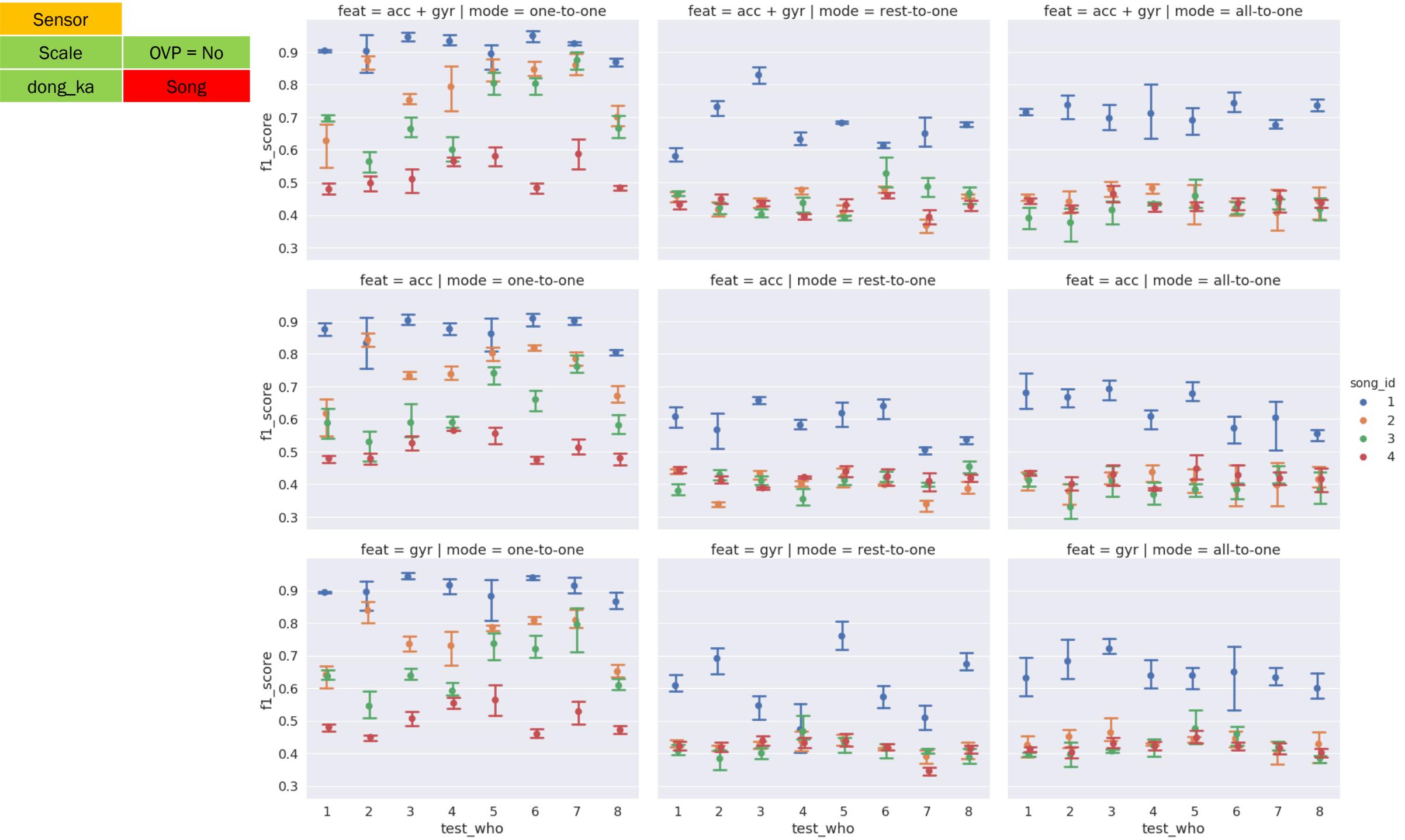
Song Picked (KNN)



Song Picked (Linear SVM)



Song Picked (RBF SVM)



Song Picked (remark 1)

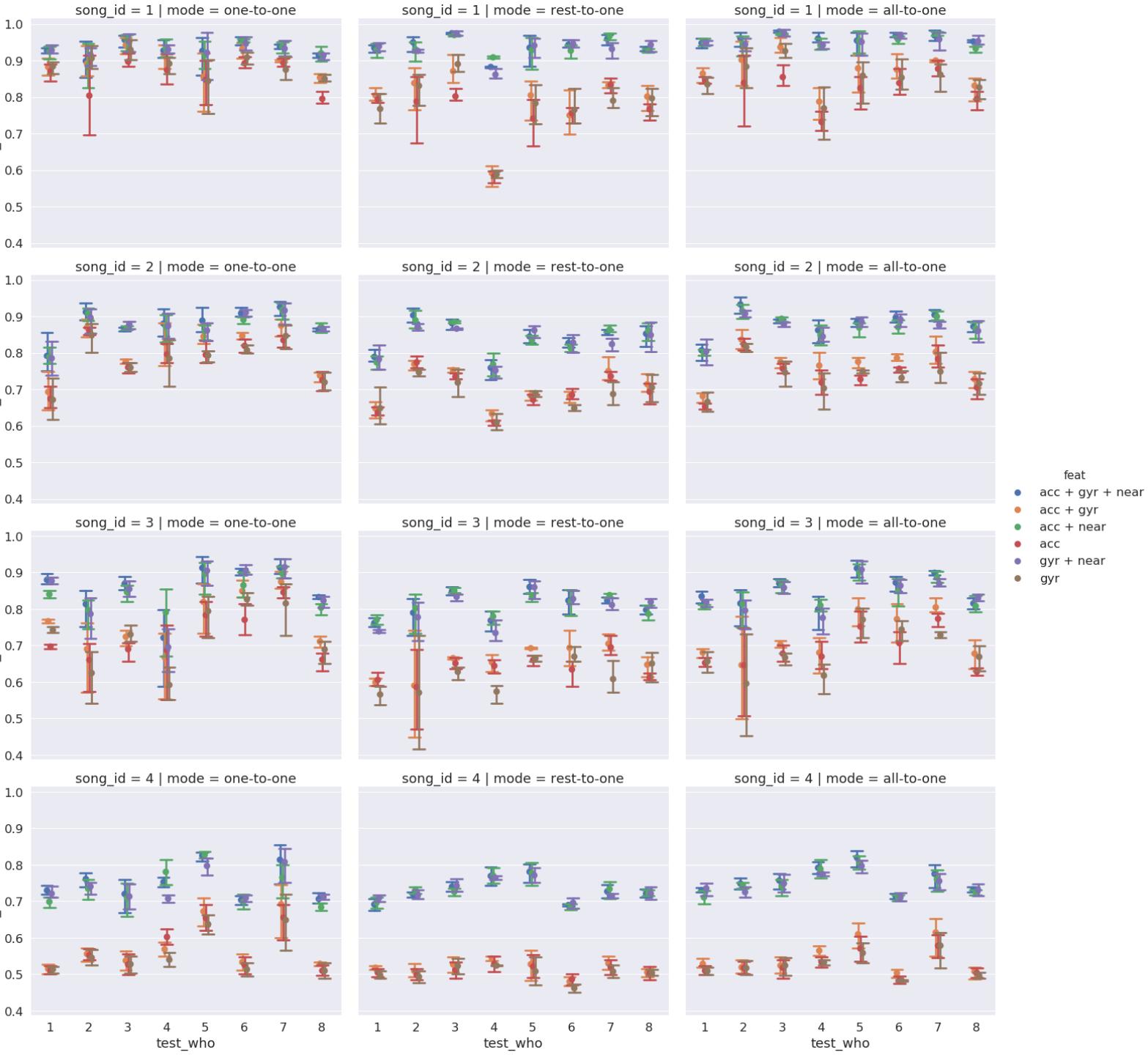
- Prediction performance has **positive correlation** with difficulty of the song;
i.e., “1”>”2”>”3”>”4”.

SENSOR PICKED



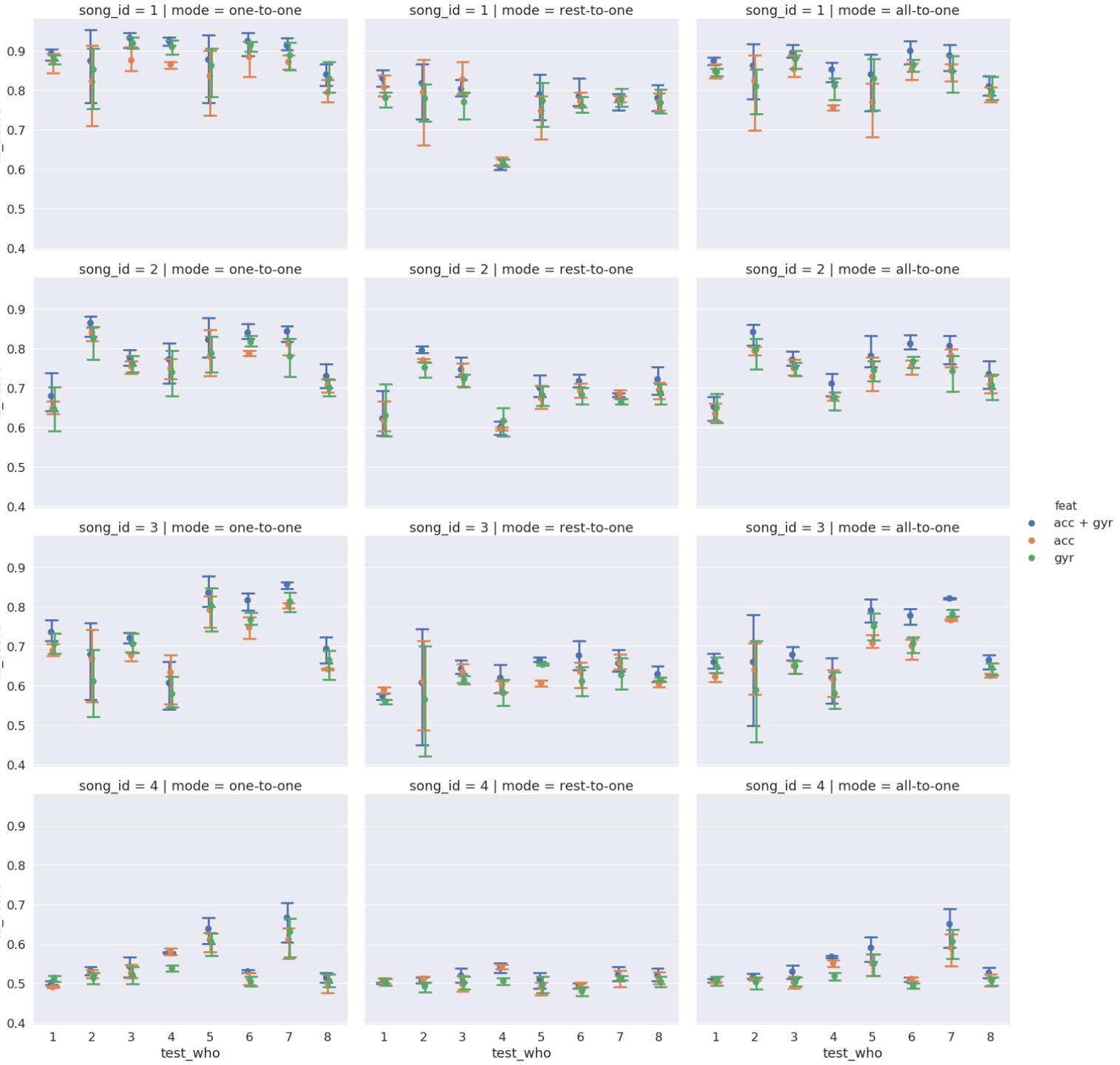
Sensor picked (LGBM)

Sensor	
Scale	OVP = No
dong_ka	Song



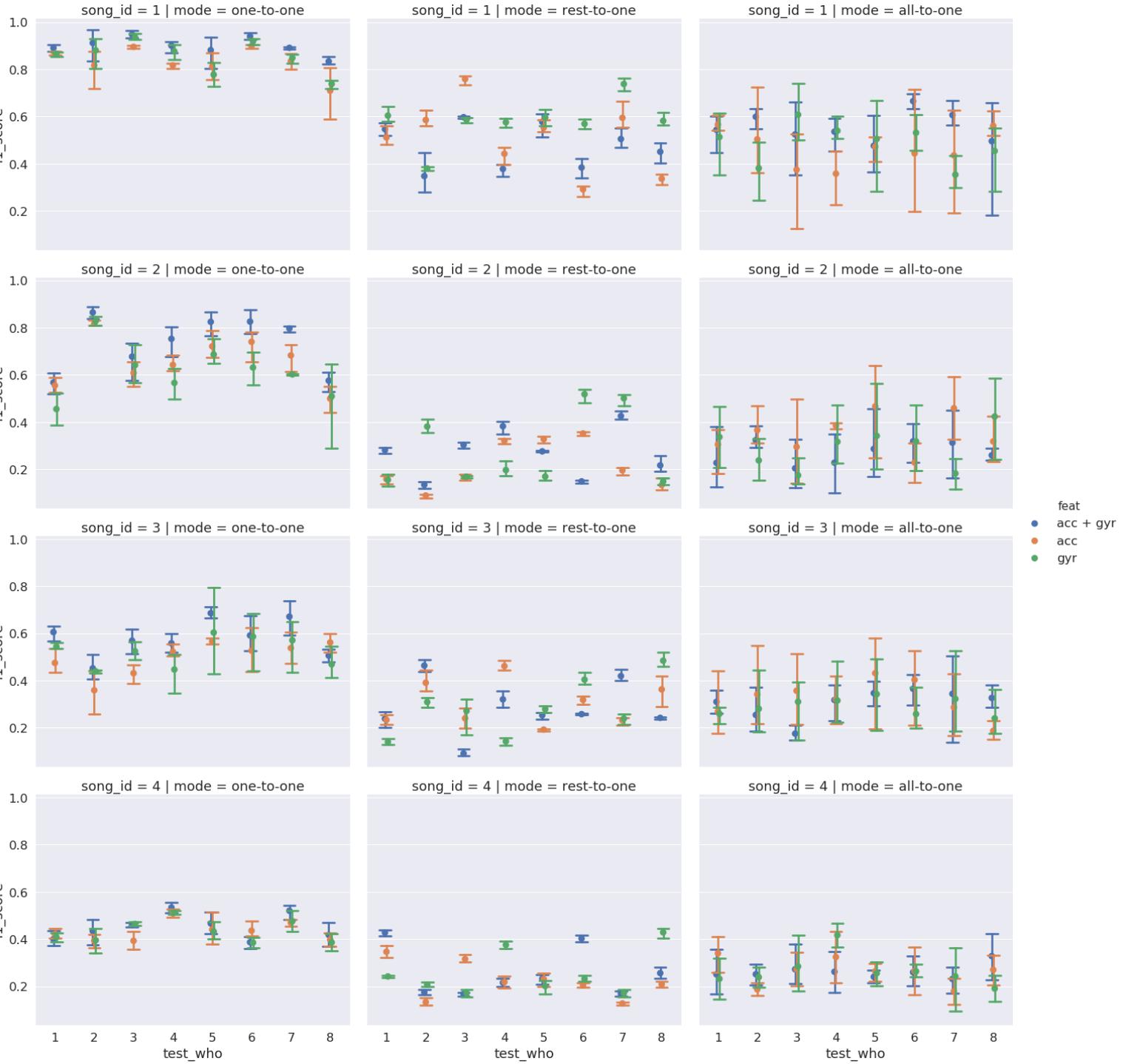
Sensor picked (KNN)

Sensor	
Scale	OVP = No
dong_ka	Song



Sensor picked (Linear SVM)

Sensor	
Scale	OVP = No
dong_ka	Song



Sensor picked (RBF SVM)

Sensor	
Scale	OVP = No
dong_ka	Song



Sensor picked (remark 2)

- “near” is powerful for LGBM
- Both acc and gyr are selected should be better for LGBM, KNN, RBF SVM

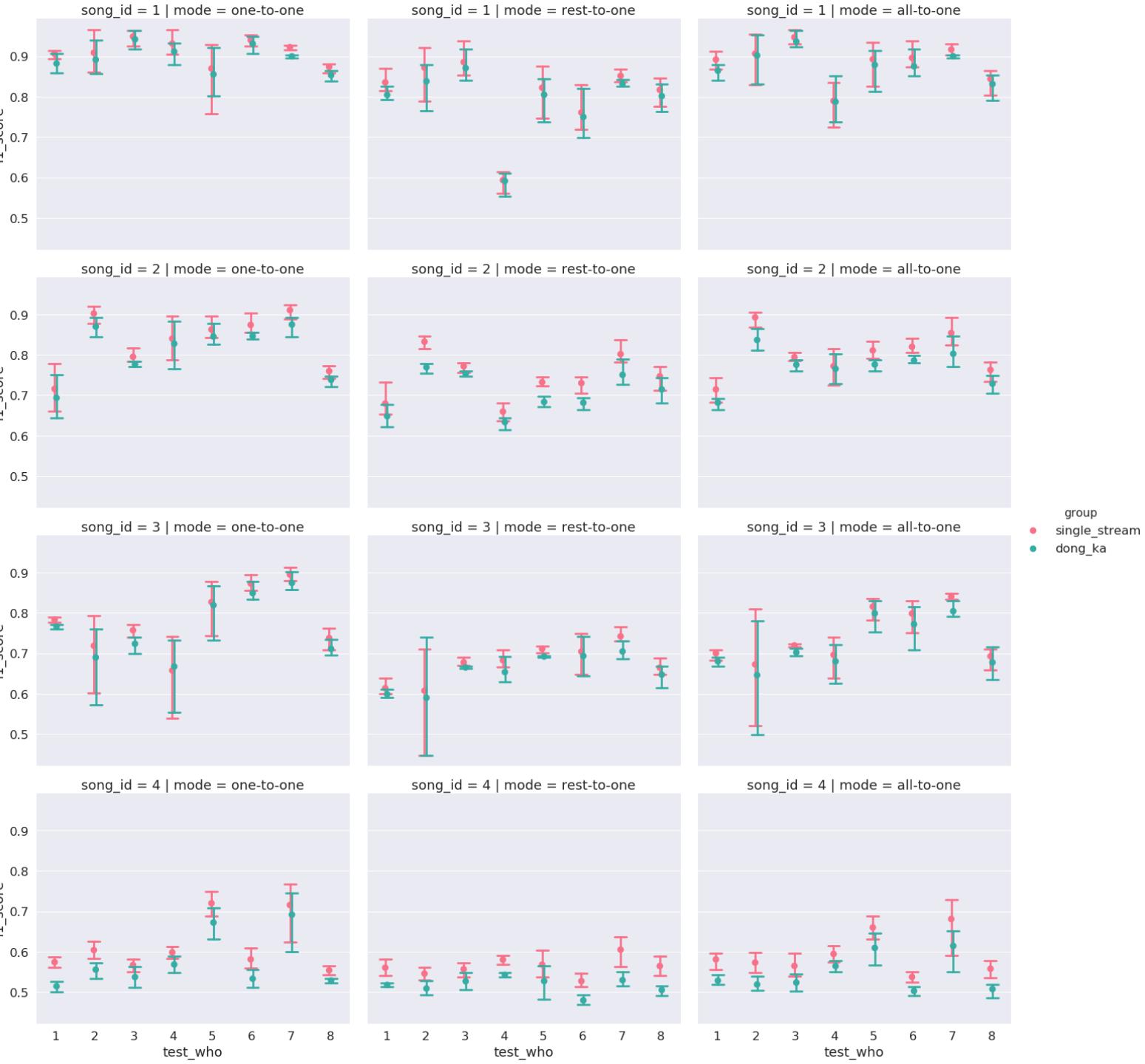
Mode \ Model	LGBM	KNN	Linear SVM	RBF SVM
One-to-one	Both	Both	Both	Both
Rest-to-one	Both	Both	Not definite	Both
All-to-one	Both	Both	Not definite	Both

GROUP PICKED



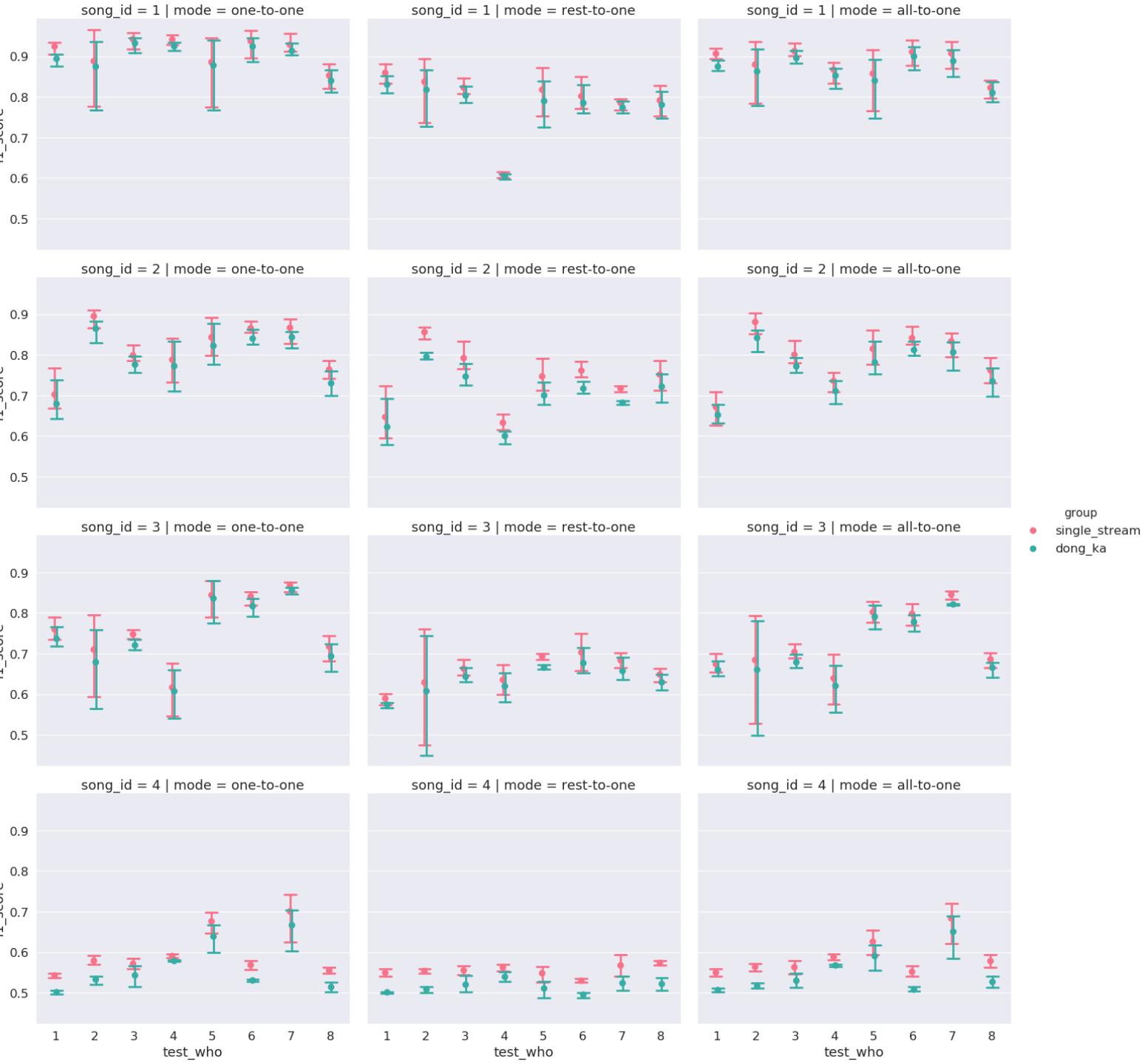
Group picked (LGBM)

acc + gyr	
Scale	OVP = No
Group	Song



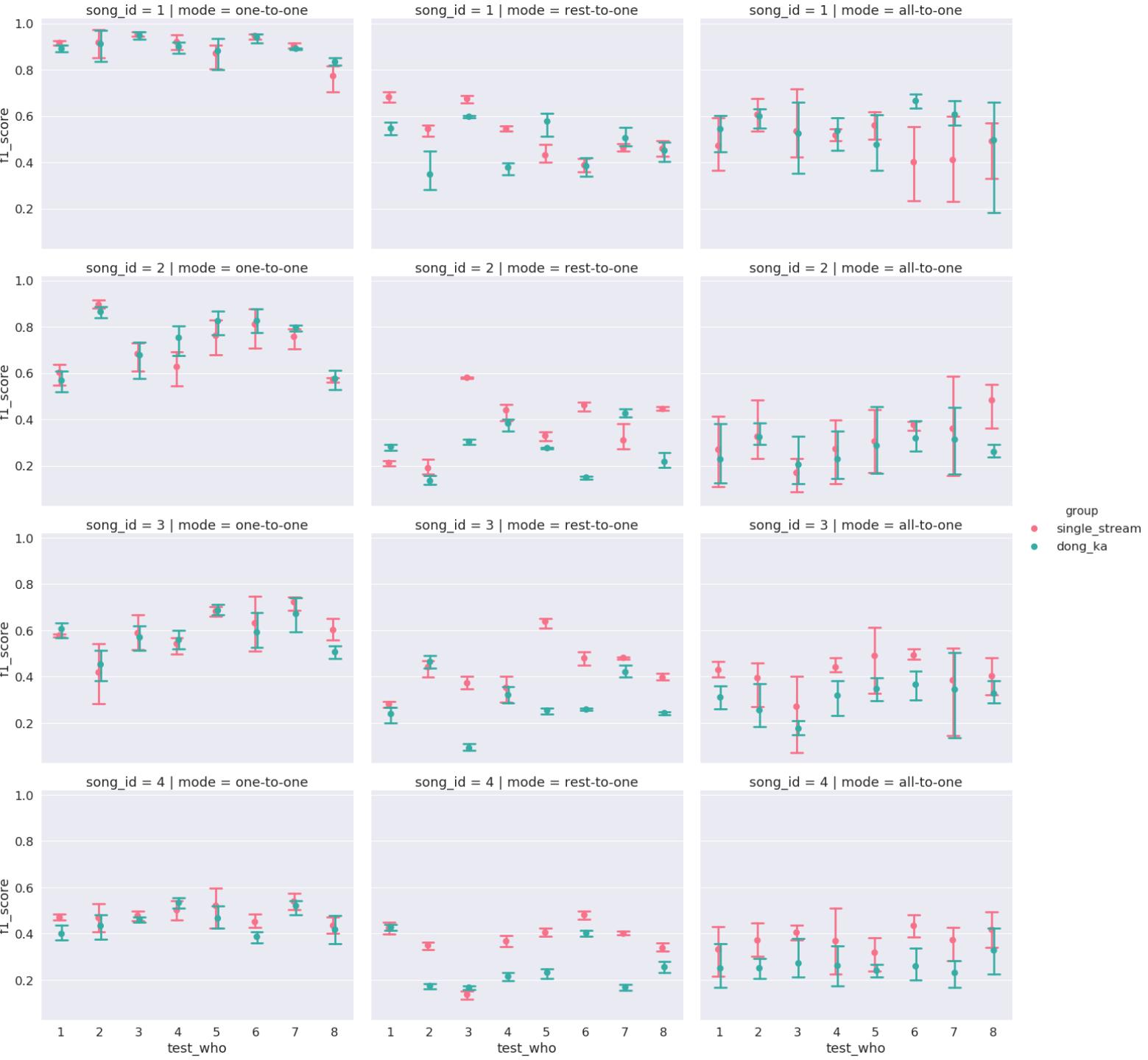
Group picked (KNN)

acc + gyr	
Scale	OVP = No
Group	Song



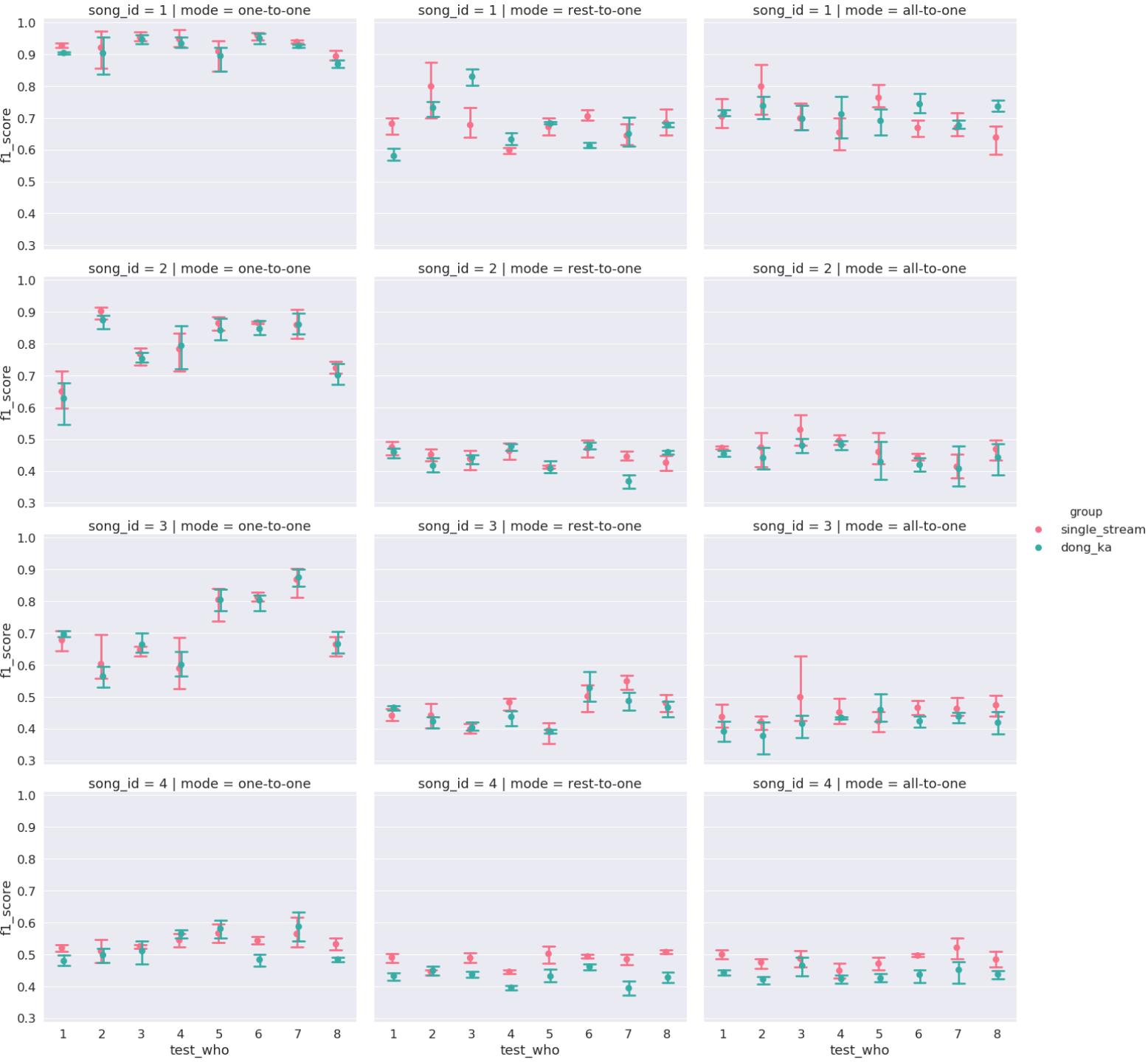
Group picked (Linear SVM)

acc + gyr	
Scale	OVP = No
Group	Song



Group picked (RBF SVM)

acc + gyr	
Scale	OVP = No
Group	Song



Group picked (remark 3)

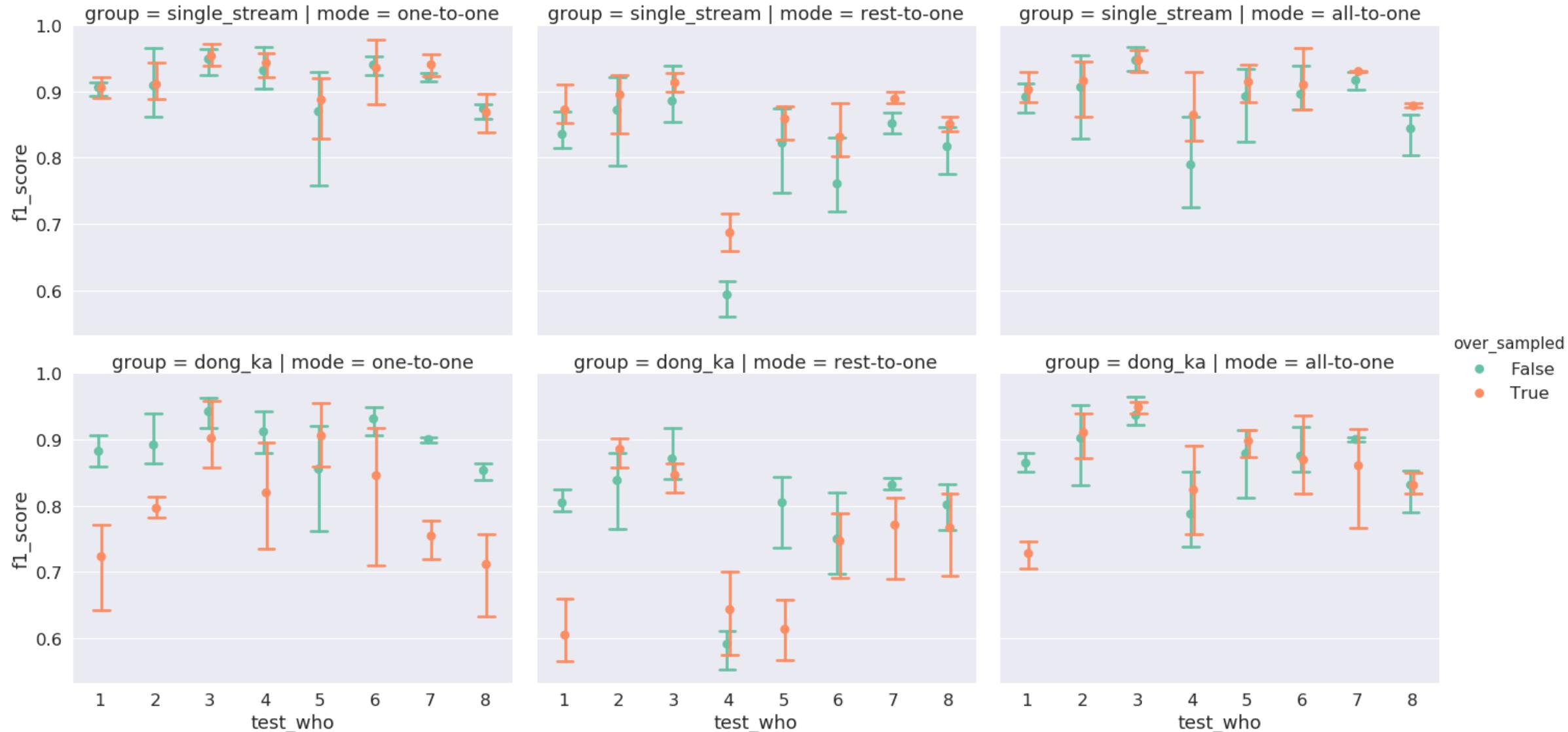
- In general, “single_stream” > “dong_ka”

OVERSAMPLED



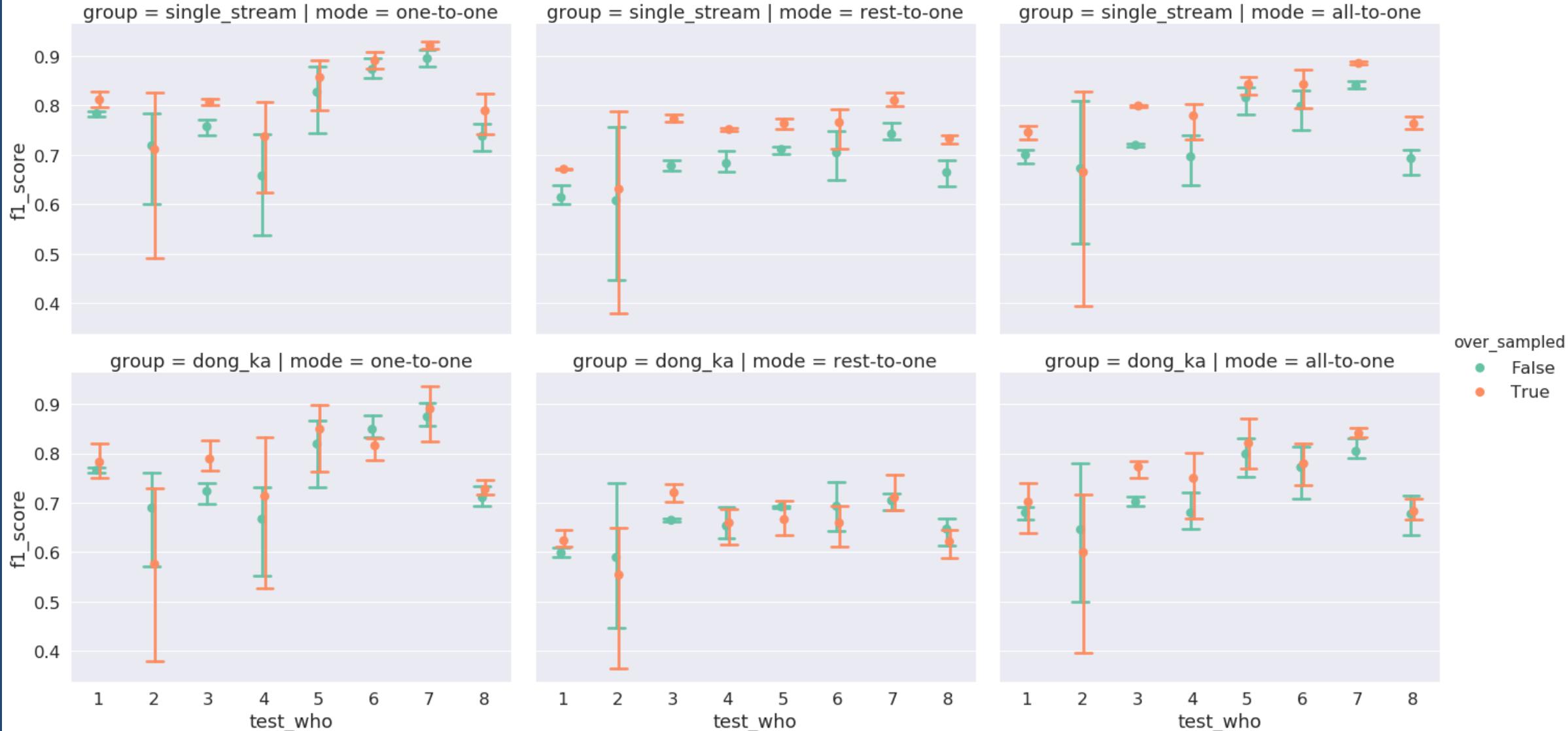
acc + gyr	
Scale	Oversample
Group	Song = 1

Oversampled (LGBM) Song 1



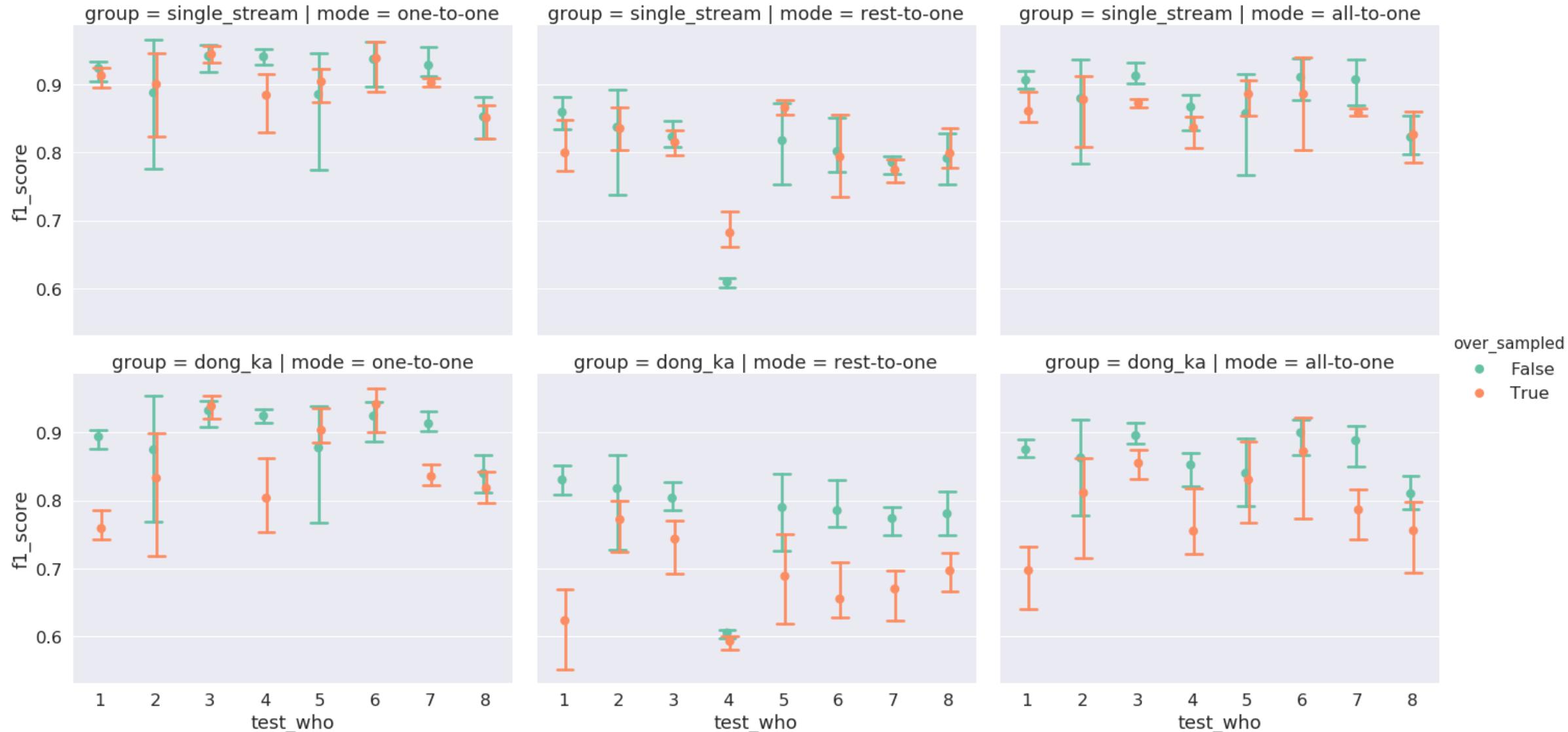
acc + gyr	
Scale	Oversample
Group	Song = 3

Oversampled (LGBM) Song 3



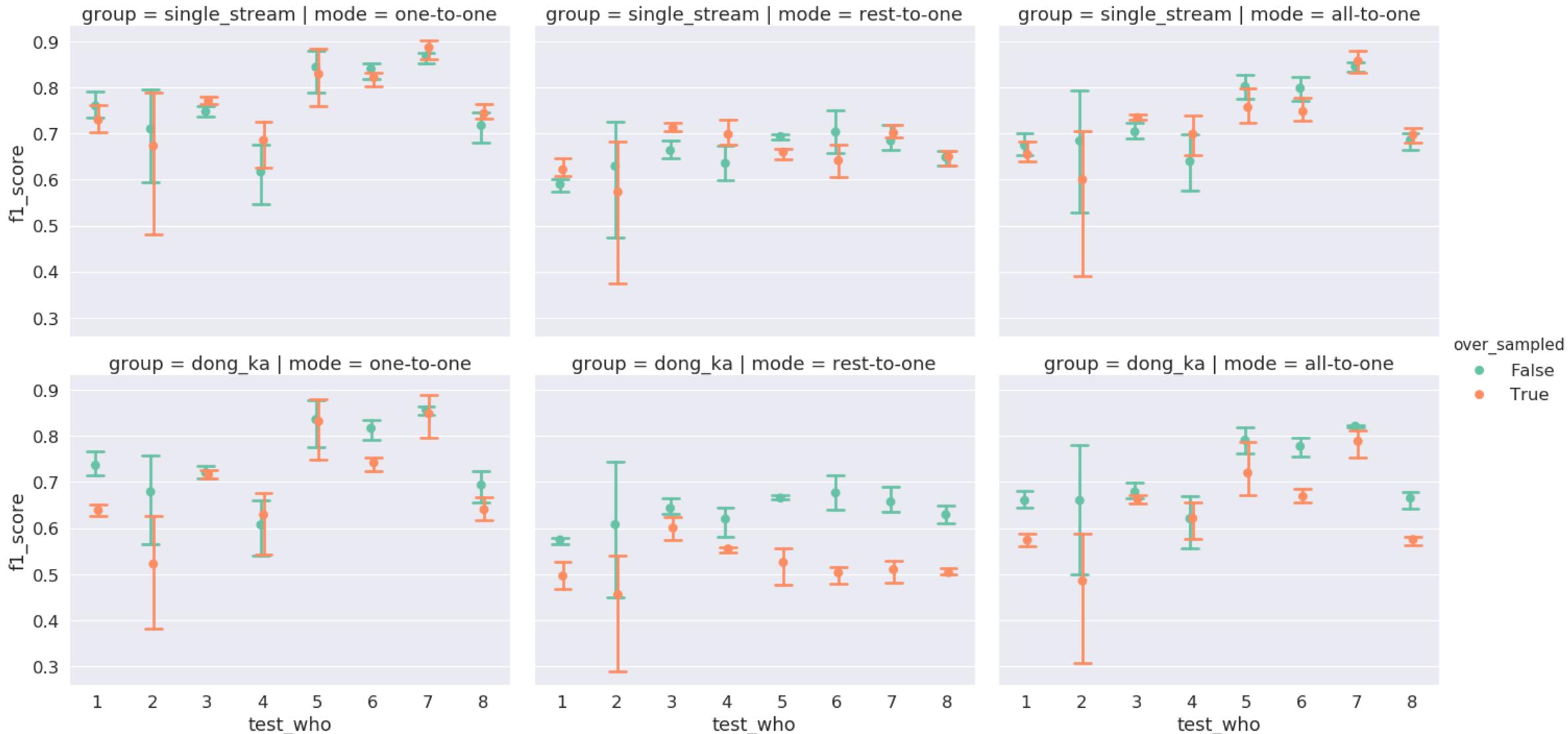
acc + gyr	
Scale	Oversample
Group	Song = 1

Oversampled (KNN) Song 1



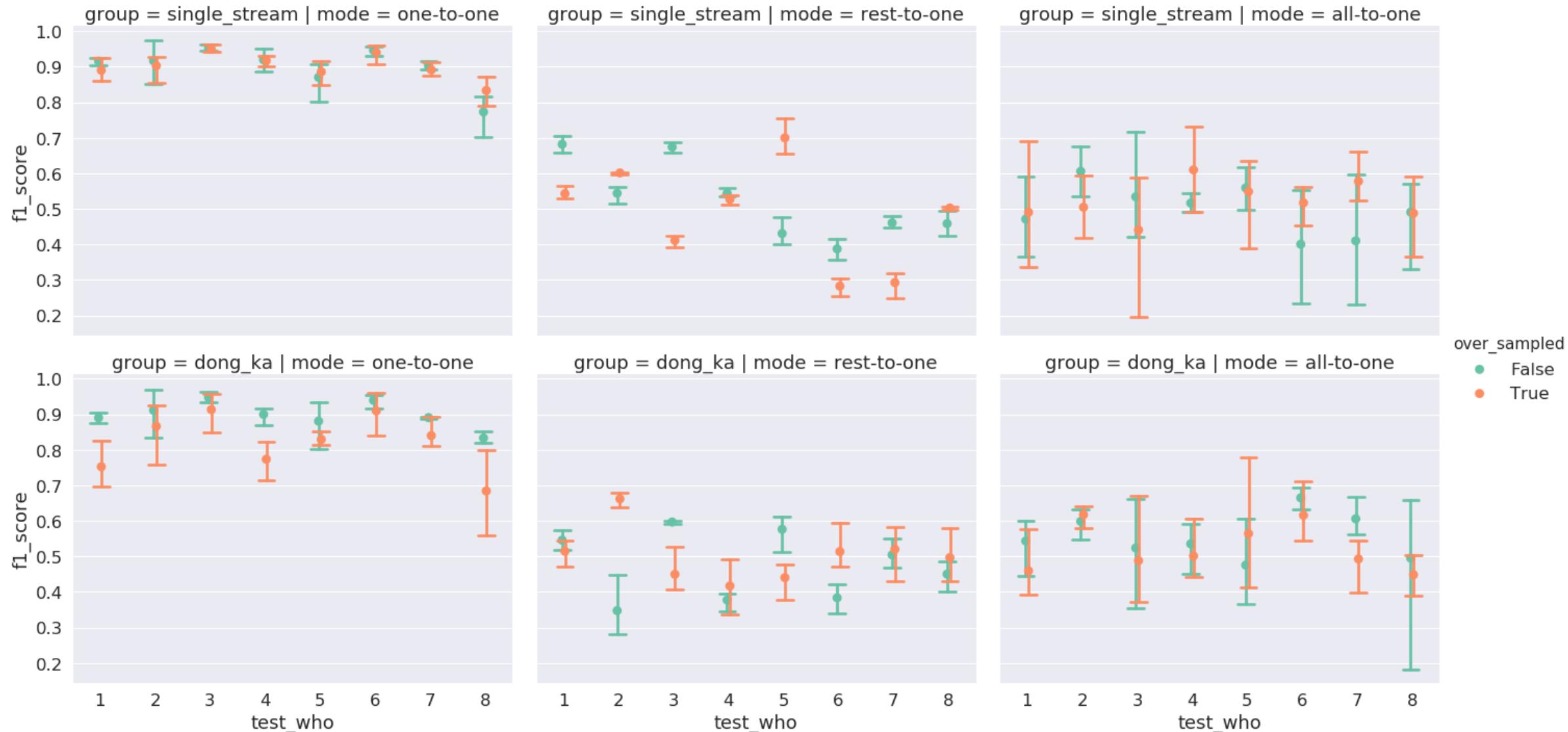
acc + gyr	
Scale	Oversample
Group	Song = 3

Oversampled (KNN) Song 3



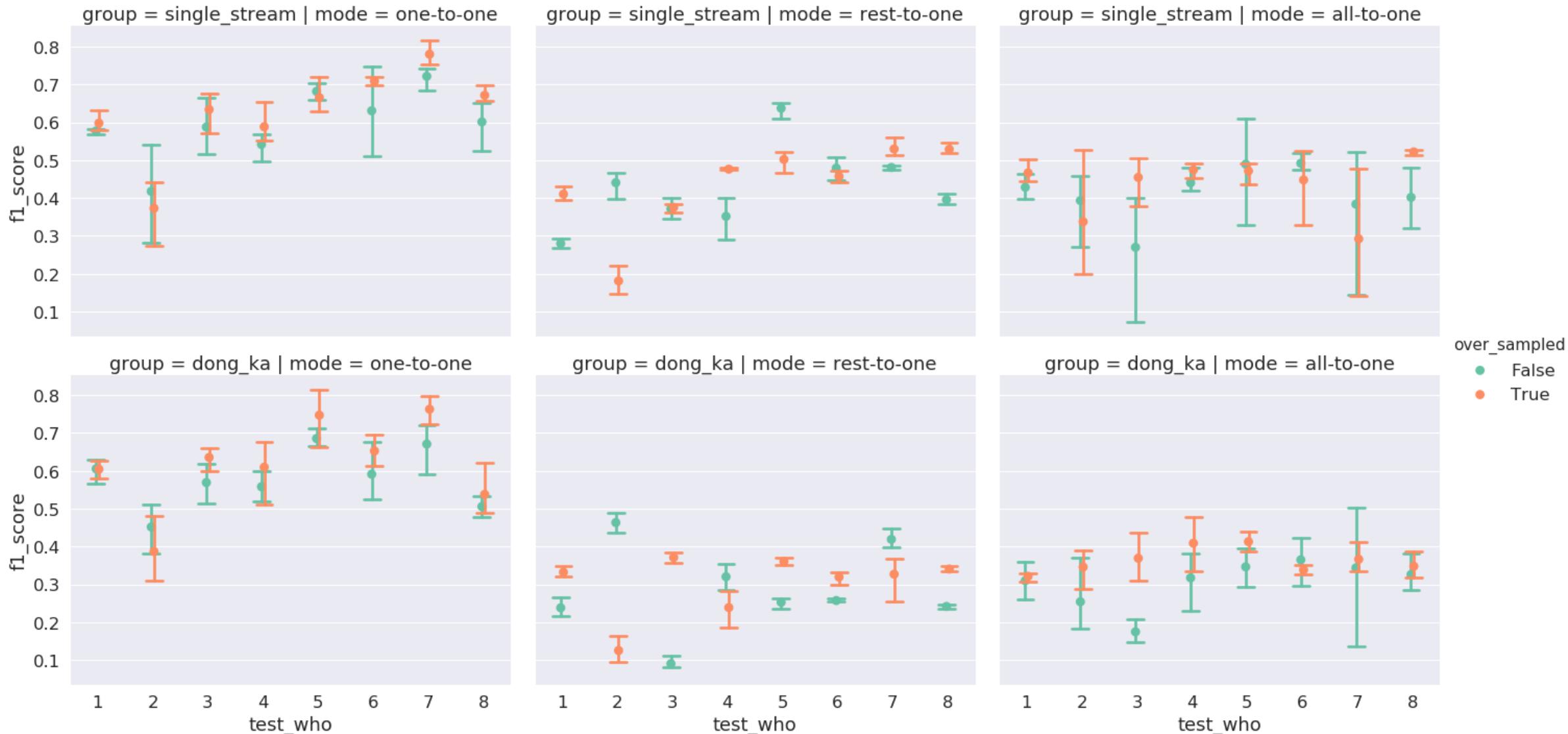
acc + gyr	
Scale	Oversample
Group	Song = 1

Oversampled (Linear SVM) Song 1



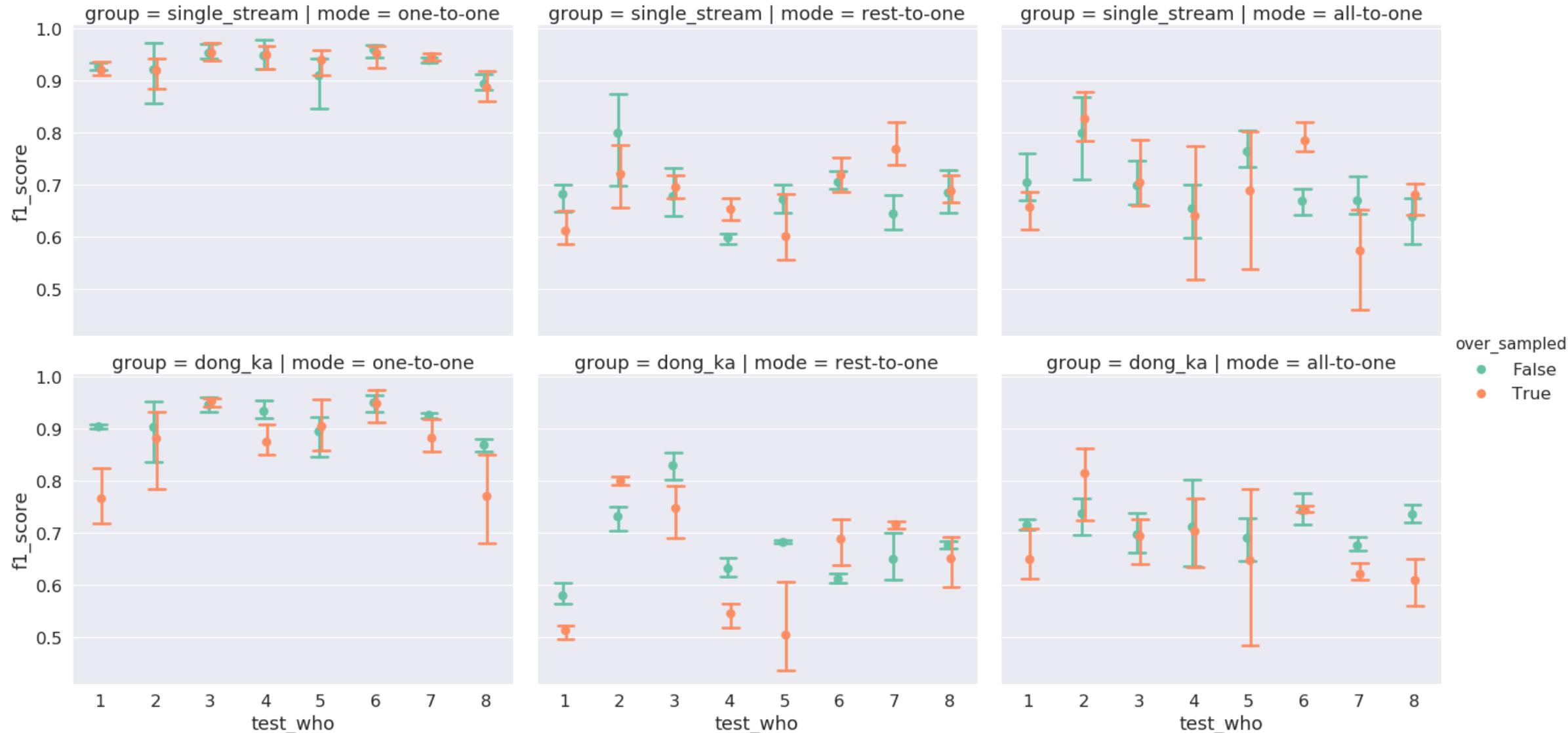
acc + gyr	
Scale	Oversample
Group	Song = 3

Oversampled (Linear SVM) Song 3



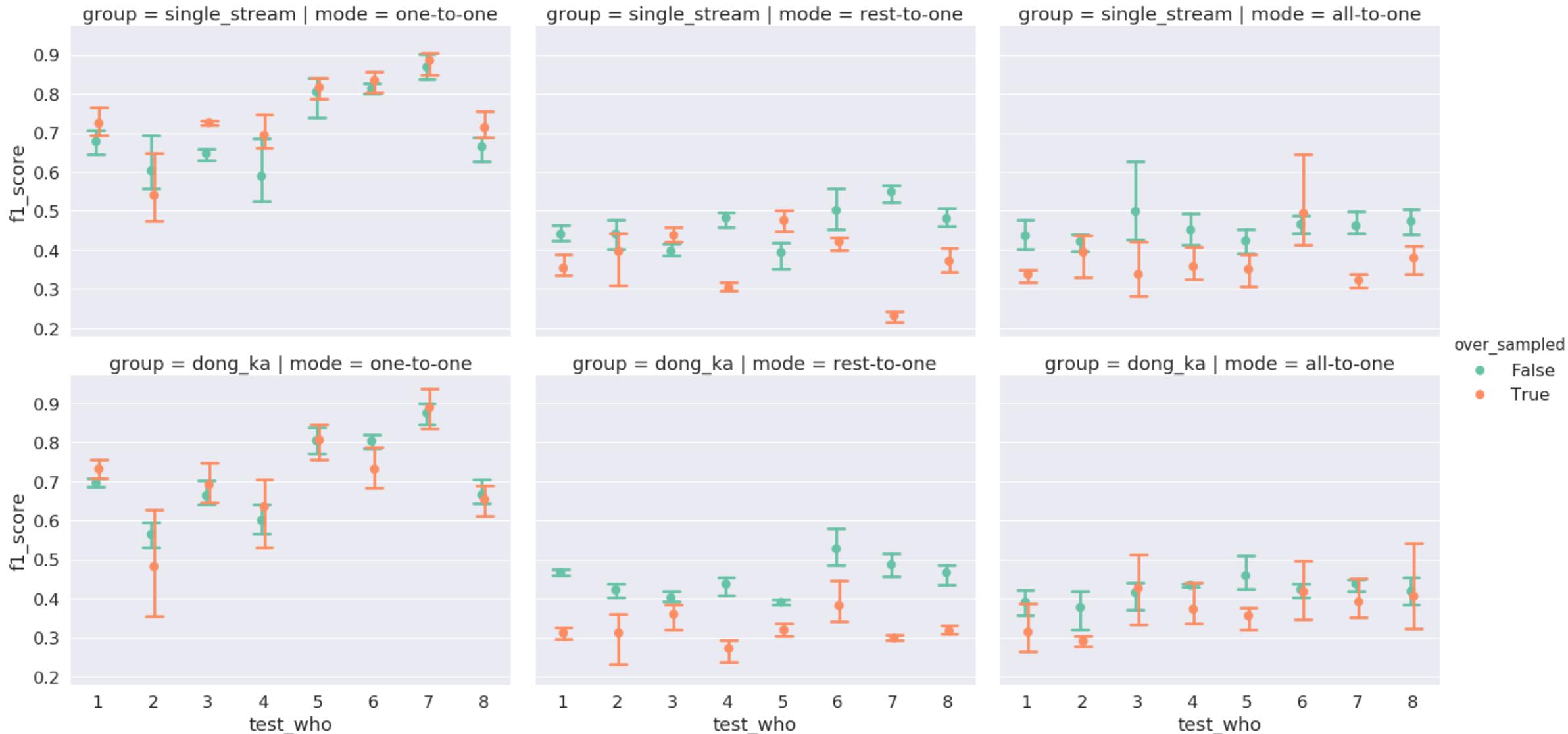
acc + gyr	
Scale	Oversample
Group	Song = 1

Oversampled (RBF SVM) Song 1



acc + gyr	
Scale	Oversample
Group	Song = 3

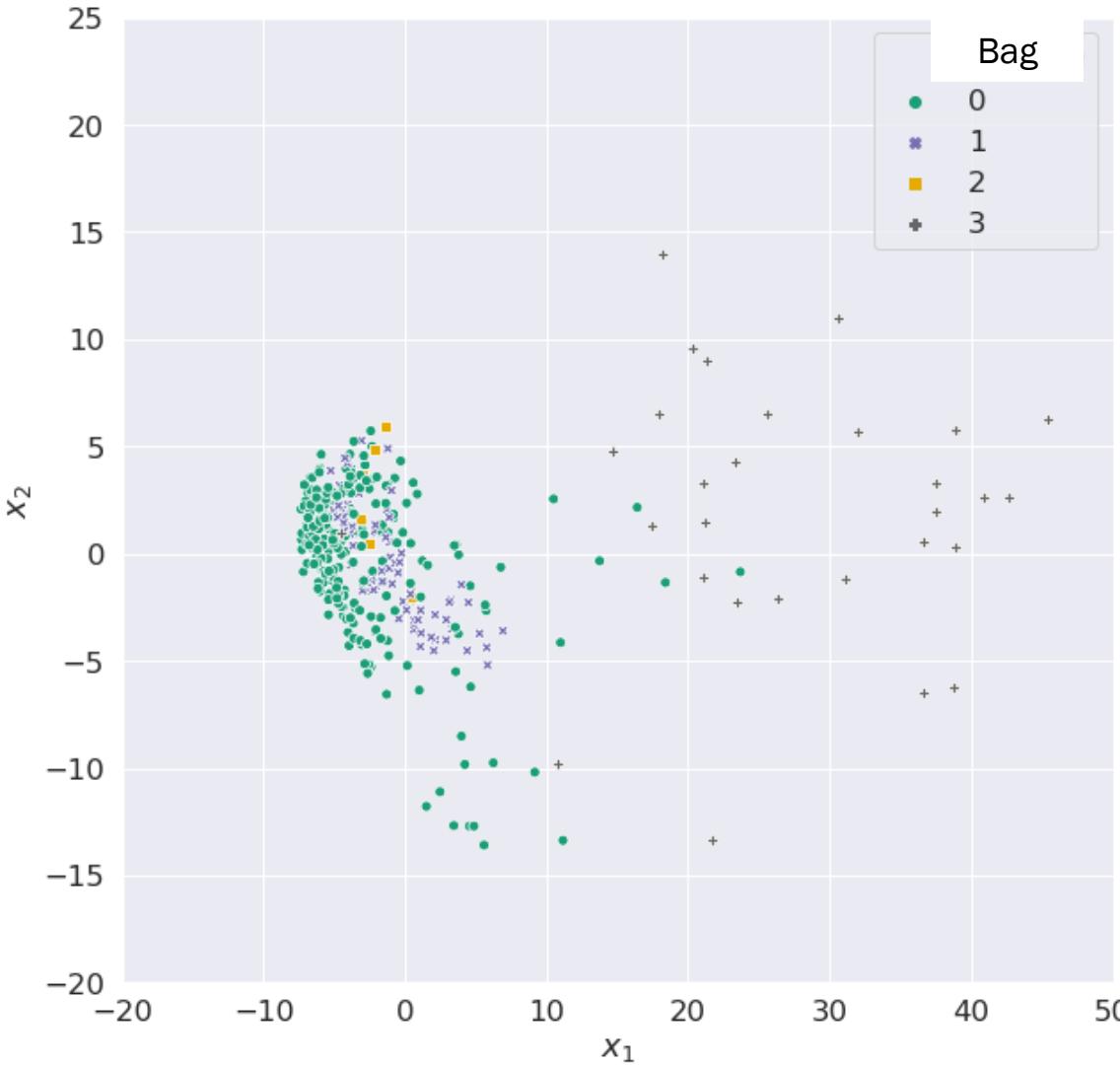
Oversampled (RBF SVM) Song 3



Oversampled PCA

acc + gyr	
Scale	Oversample
dong_ka	(7, 1, 3)

Before oversample



After oversample



Oversampled (remark 4)

- Maybe need to evaluated by “macro” F1-score.

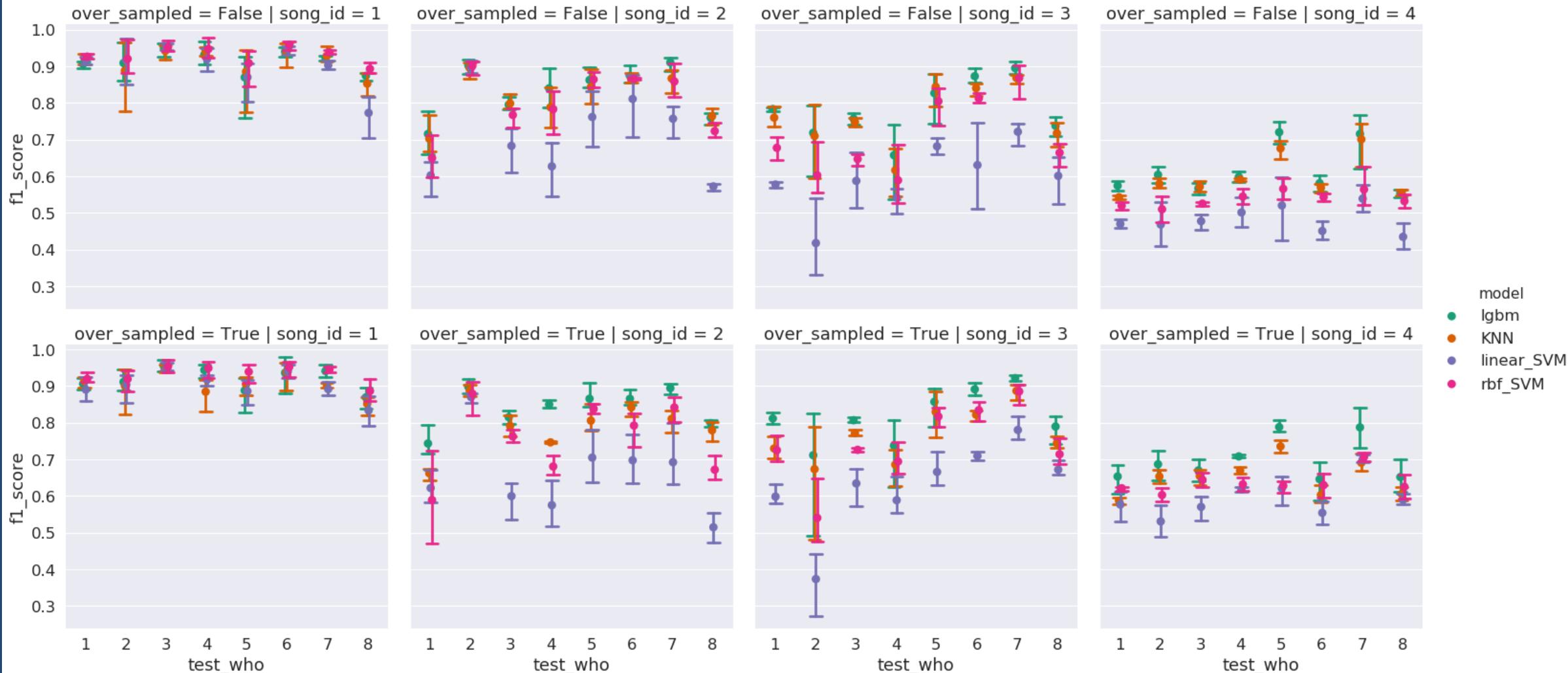
Model	LGBM	KNN	Linear SVM	RBF SVM
result	Better	Worse	Not definite	Worse

MODEL COMPARISON



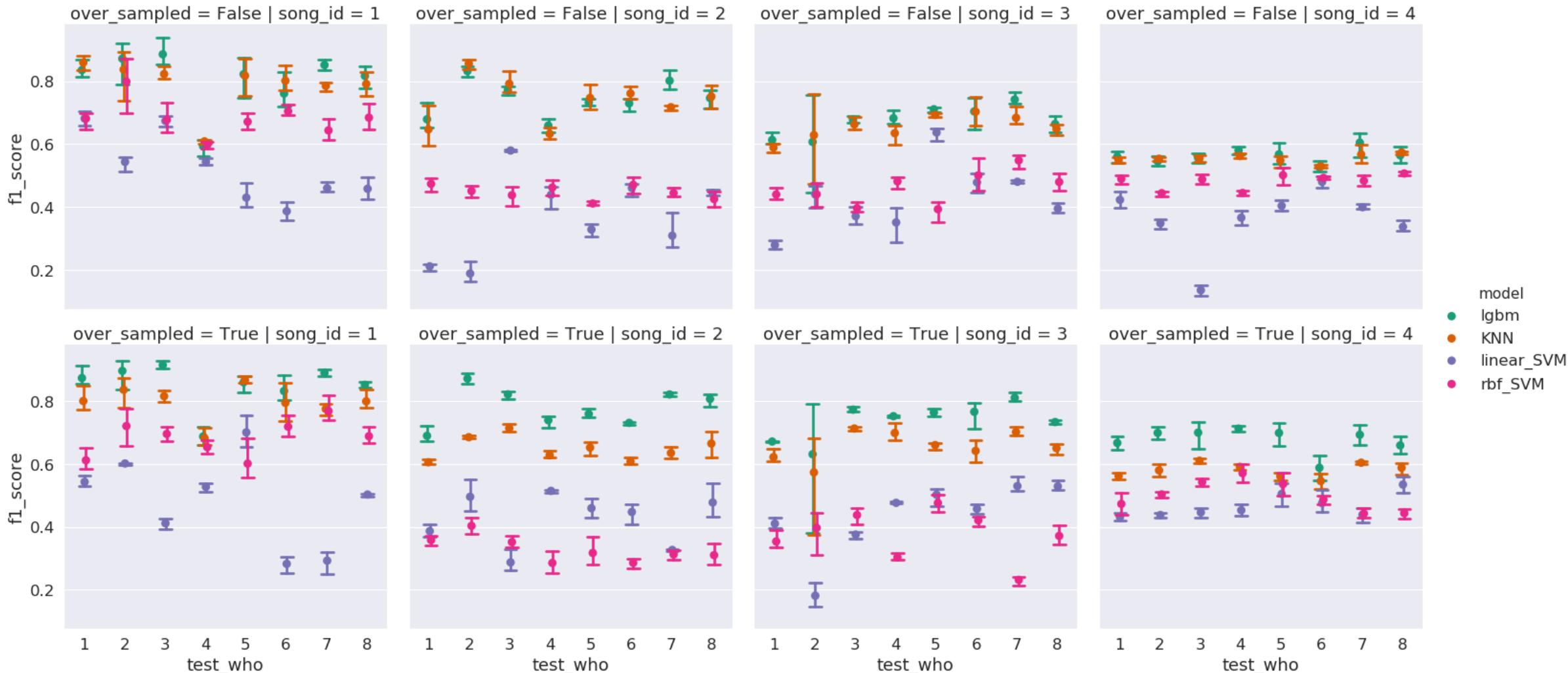
acc + gyr	Model
Scale	Oversample
SS	Song

Model comparison (one-to-one) SS



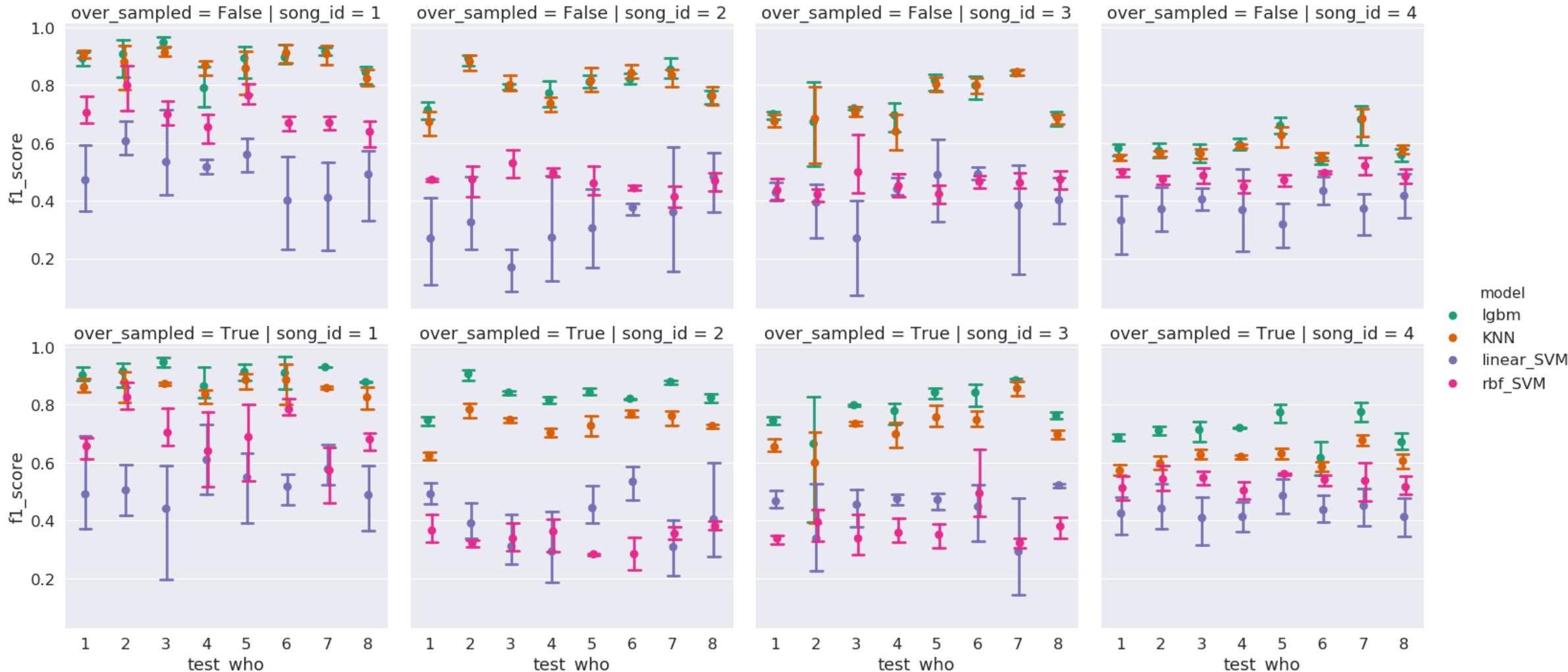
acc + gyr	Model
Scale	Oversample
SS	Song

Model comparison (rest-to-one) SS



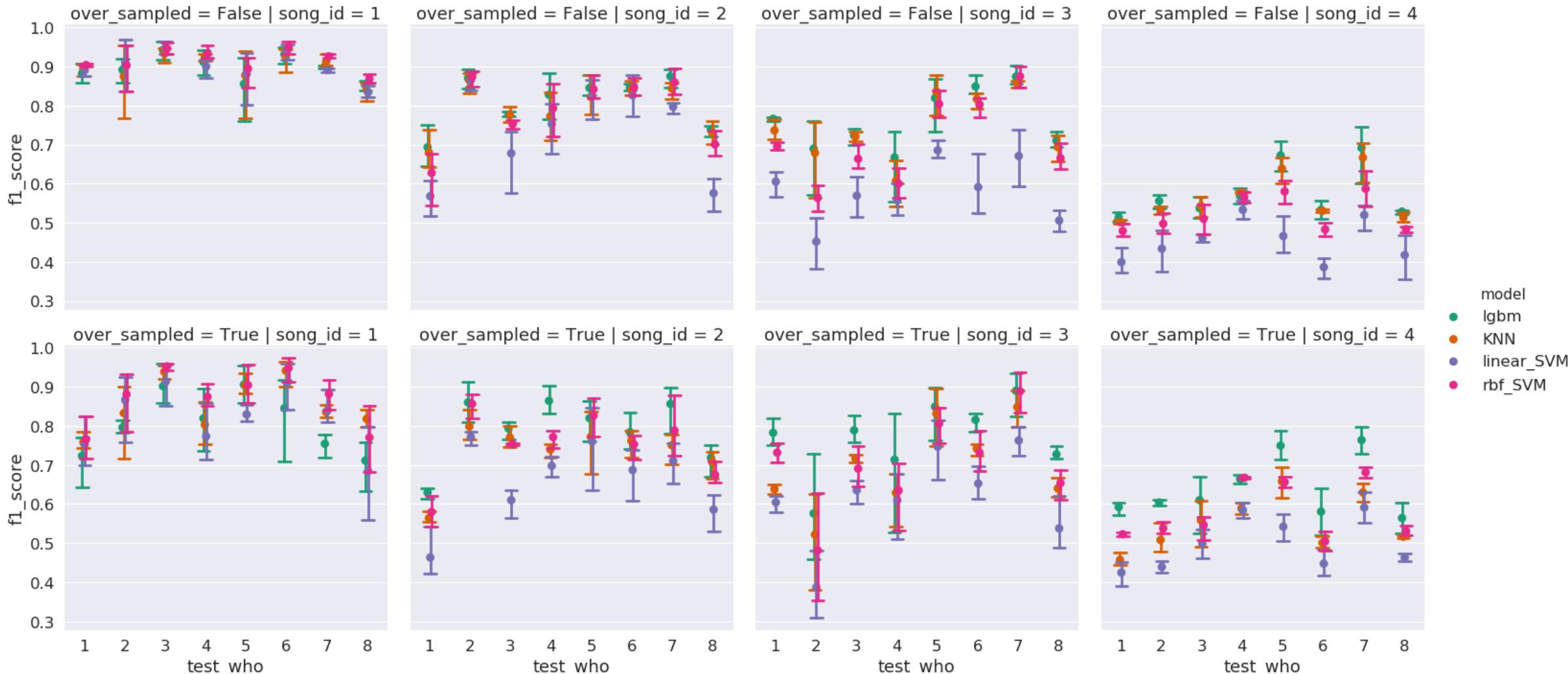
acc + gyr	Model
Scale	Oversample
SS	Song

Model comparison (all-to-one) SS



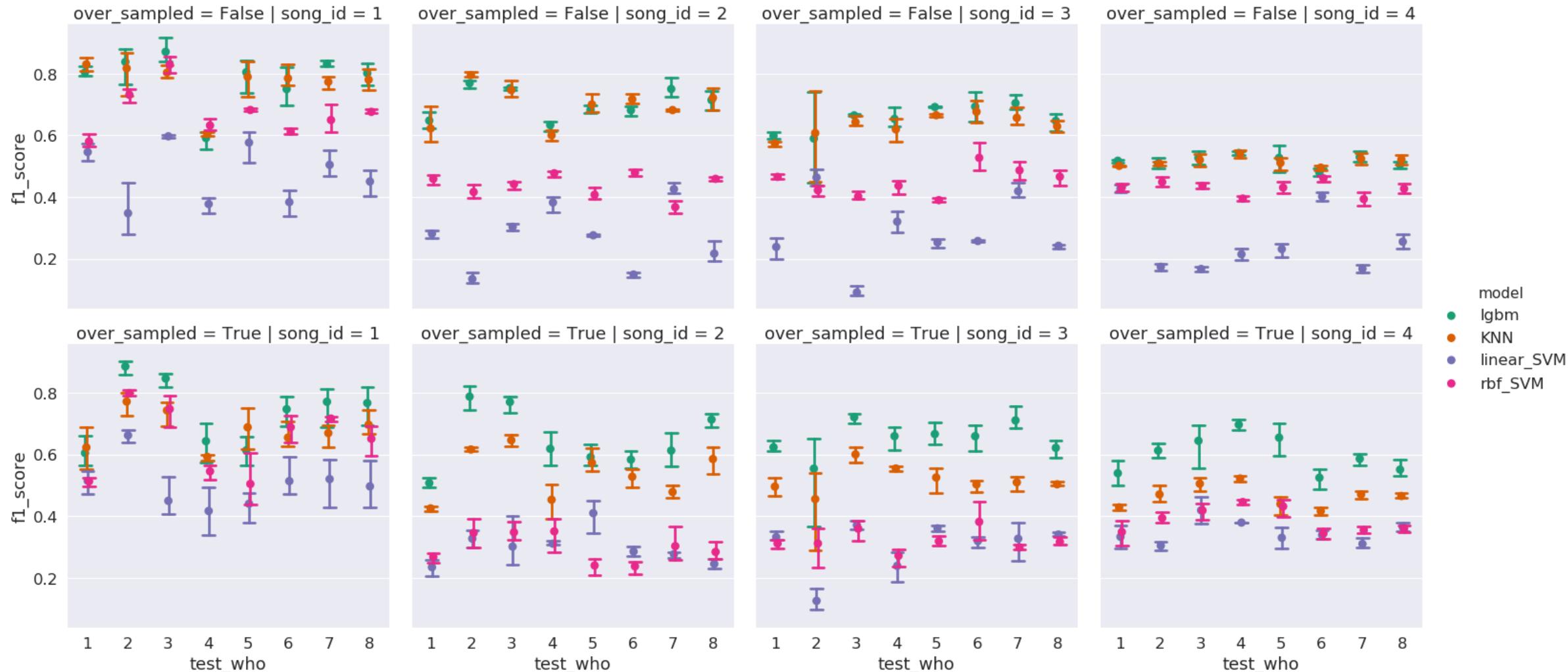
acc + gyr	Model
Scale	Oversample
DK	Song

Model comparison (one-to-one) DK



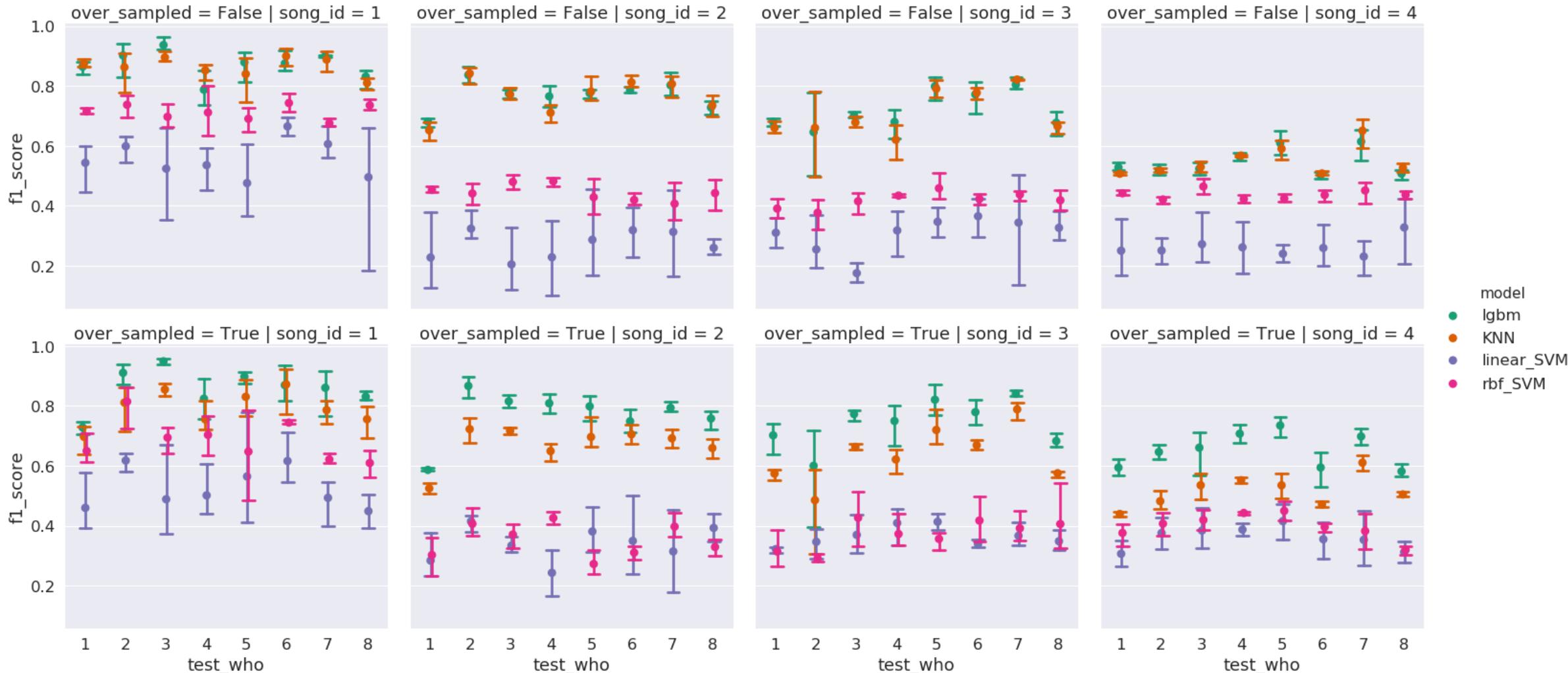
acc + gyr	Model
Scale	Oversample
DK	Song

Model comparison (rest-to-one) DK



acc + gyr	Model
Scale	Oversample
DK	Song

Model comparison (all-to-one) DK



Model comparison (remark 5)

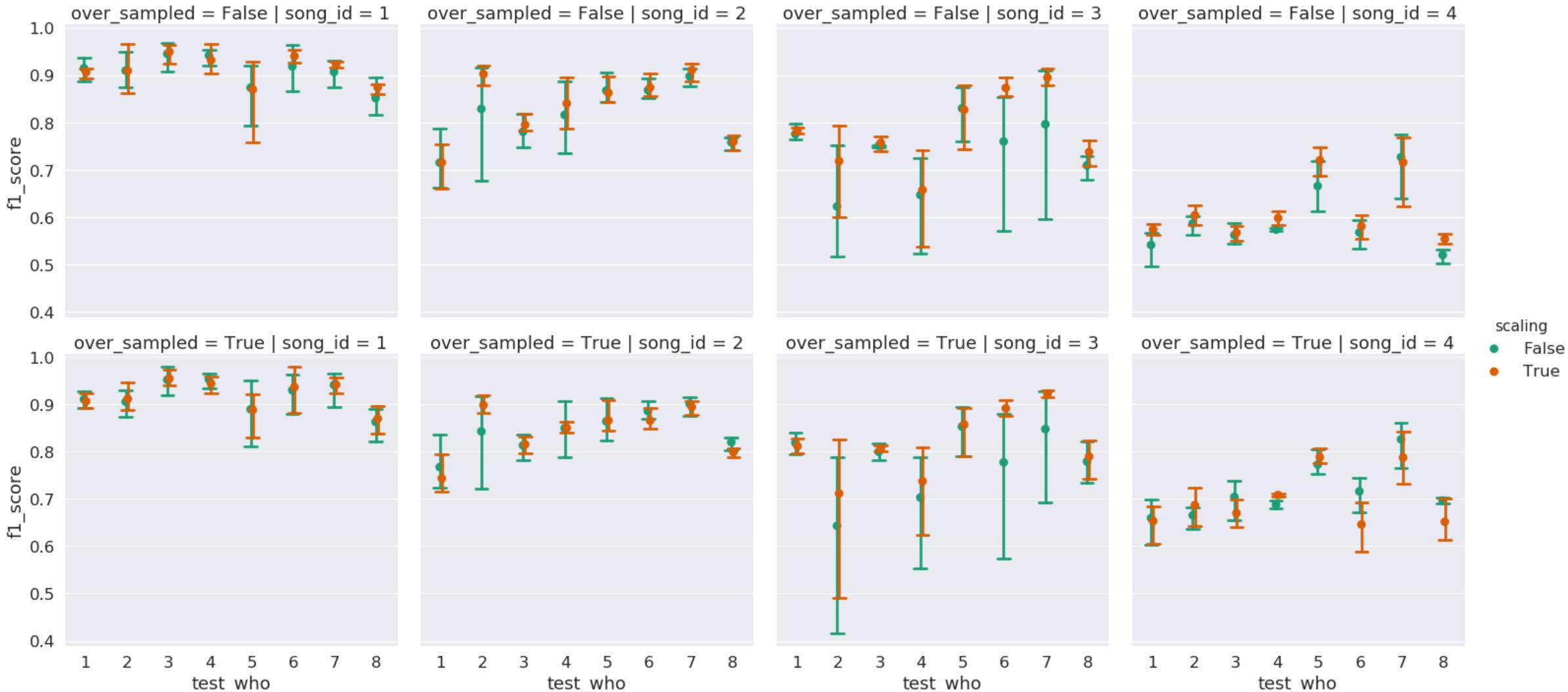
- In general, “LGBM” > “KNN” > “RBF SVM” > “Linear SVM”.

SCALING EFFECT



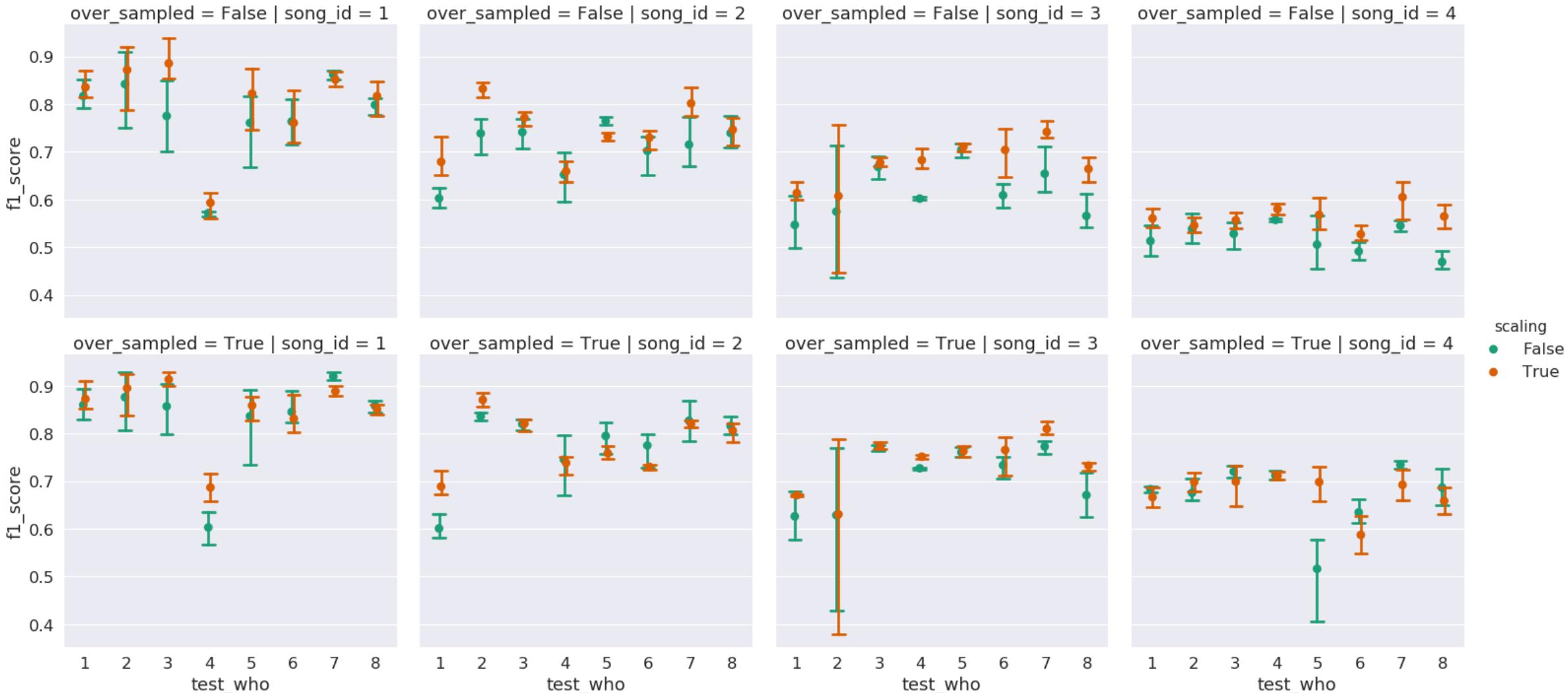
acc + gyr	LGBM
Scale	Oversample
SS	Song

Scaling effect (one-to-one) SS



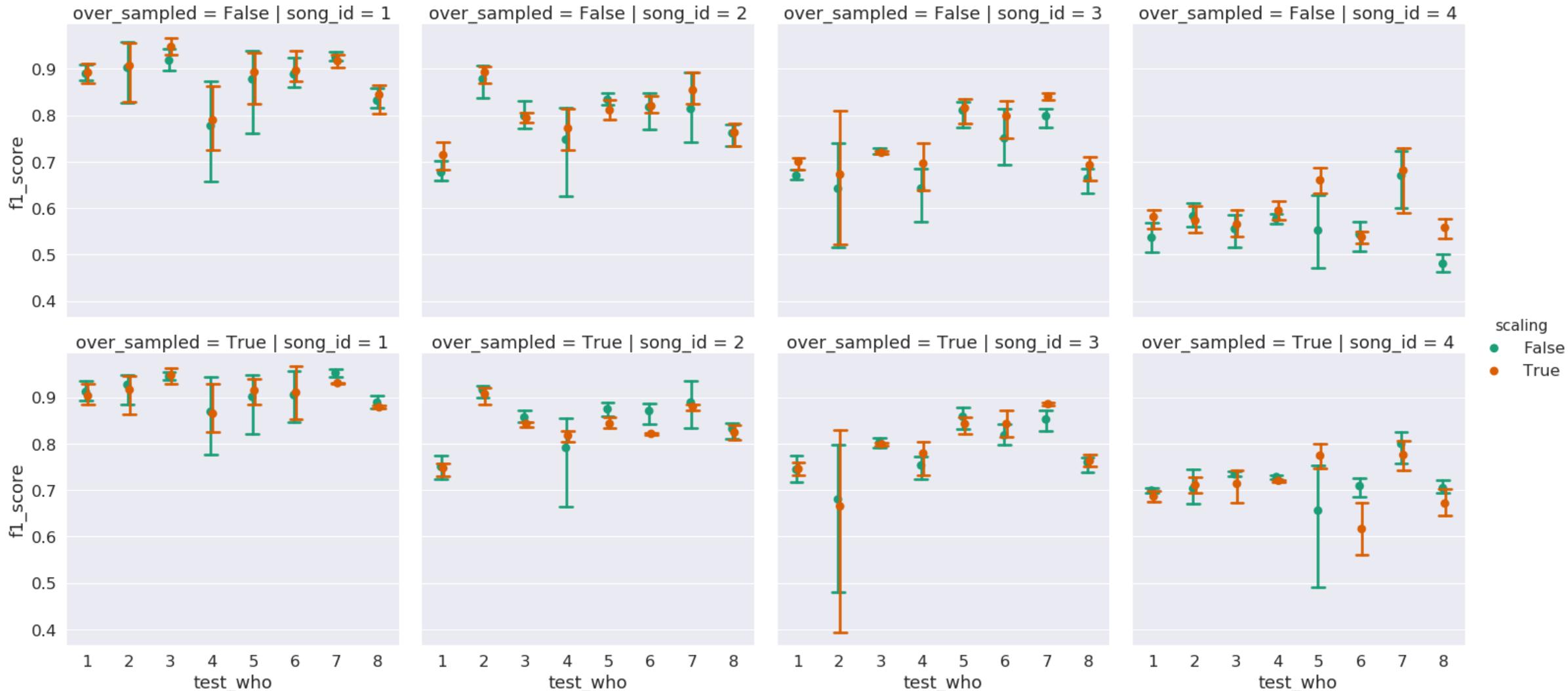
acc + gyr	LGBM
Scale	Oversample
SS	Song

Scaling effect (rest-to-one) SS



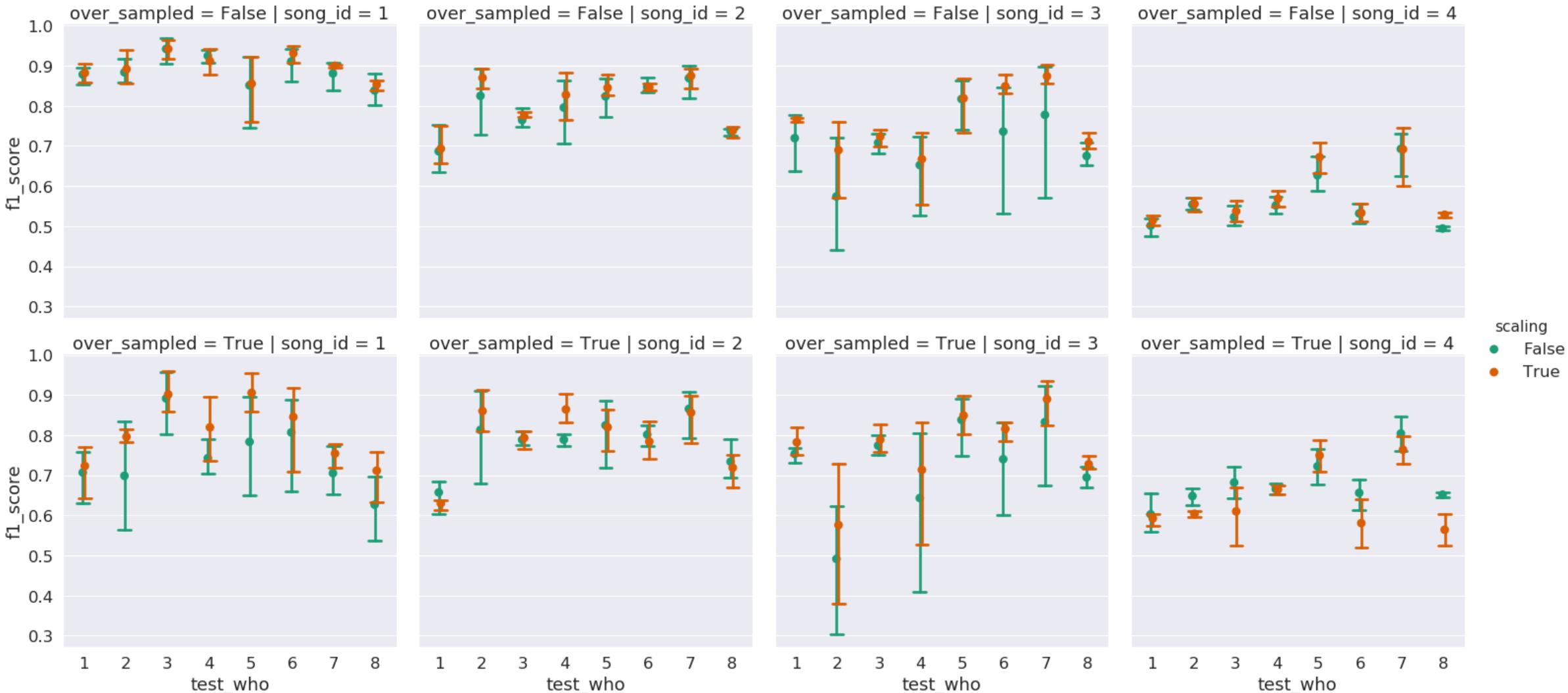
acc + gyr	LGBM
Scale	Oversample
SS	Song

Scaling effect (all-to-one) SS



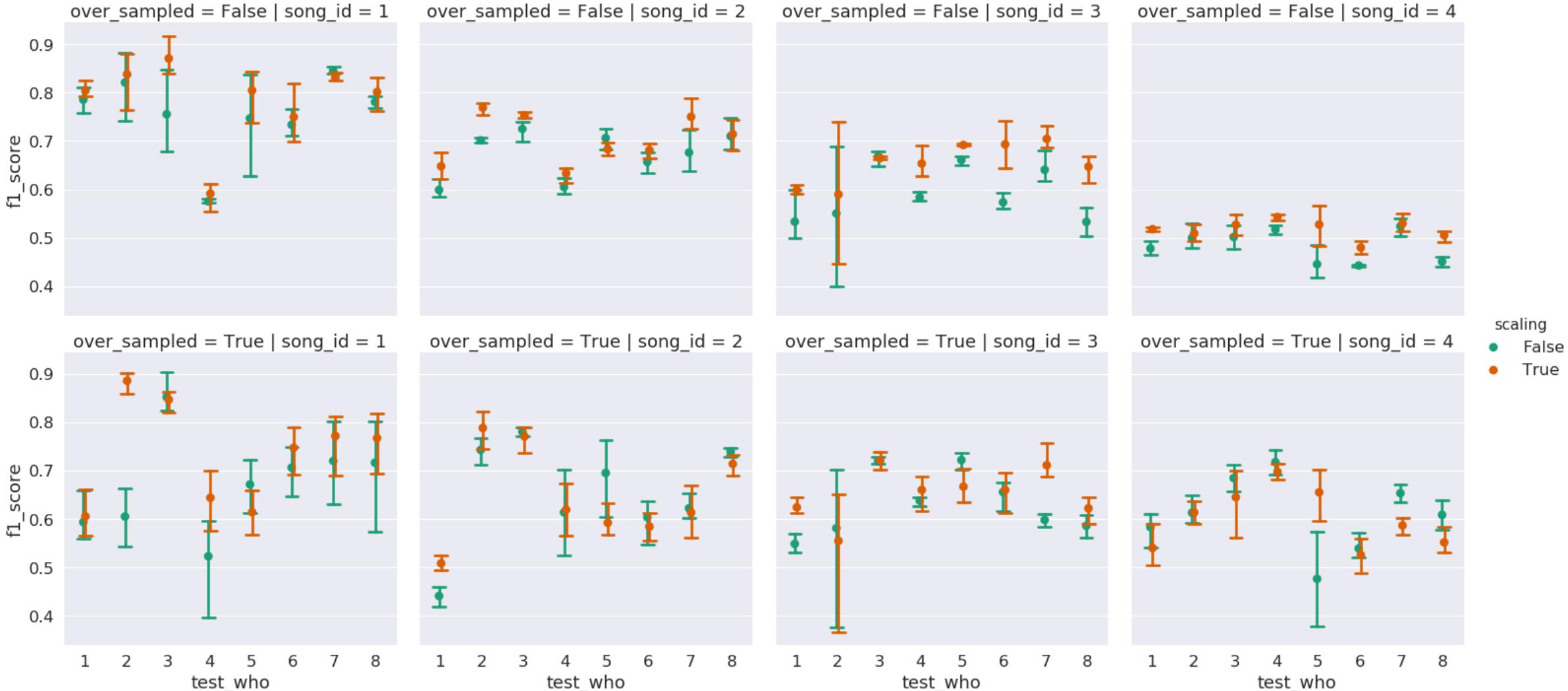
acc + gyr	LGBM
Scale	Oversample
DK	Song

Scaling effect (one-to-one) DK



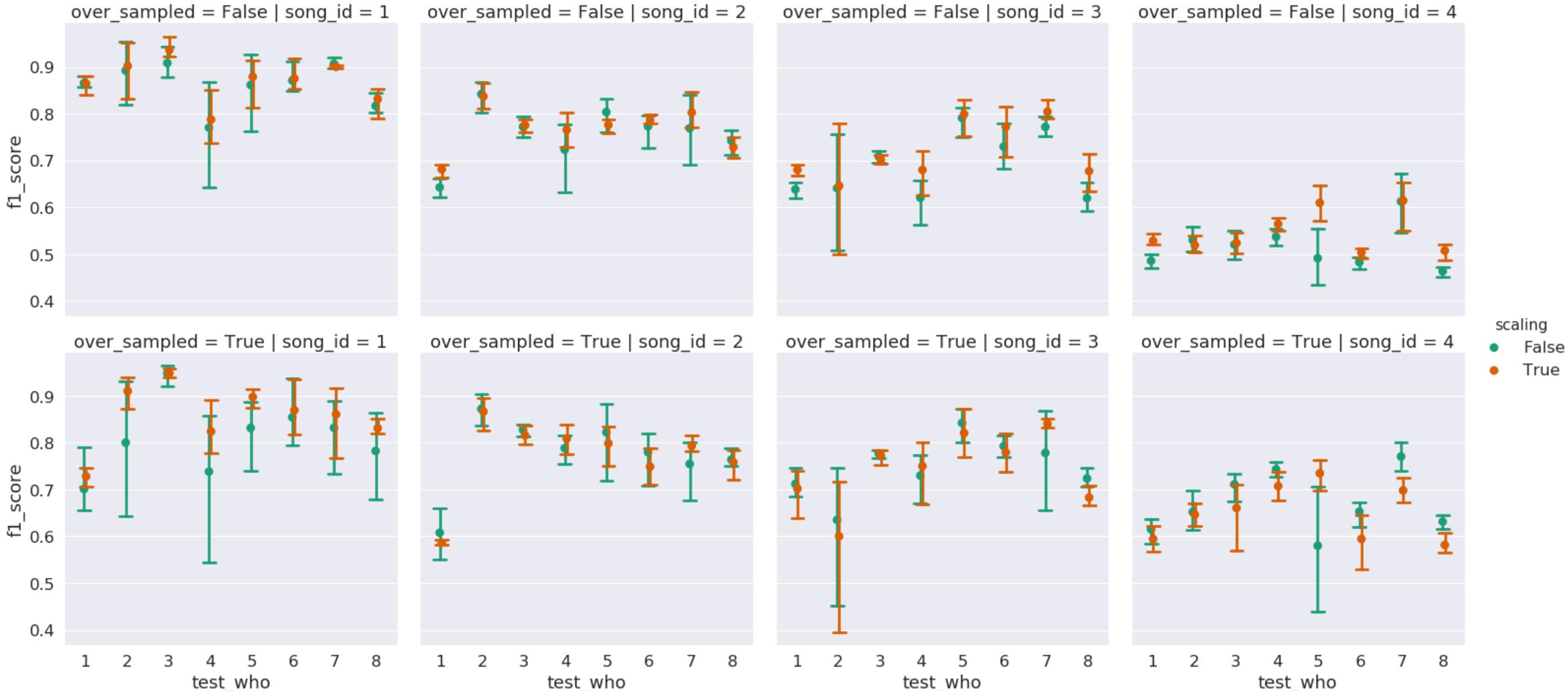
acc + gyr	LGBM
Scale	Oversample
DK	Song

Scaling effect (rest-to-one) DK



acc + gyr	LGBM
Scale	Oversample
DK	Song

Scaling effect (all-to-one) DK



Scaling effect (remark 6)

- It seems that oversampling leads to produce more noisy data.

	SS	DK
Before oversample	Better	Better
After oversample	Not definite	Not definite