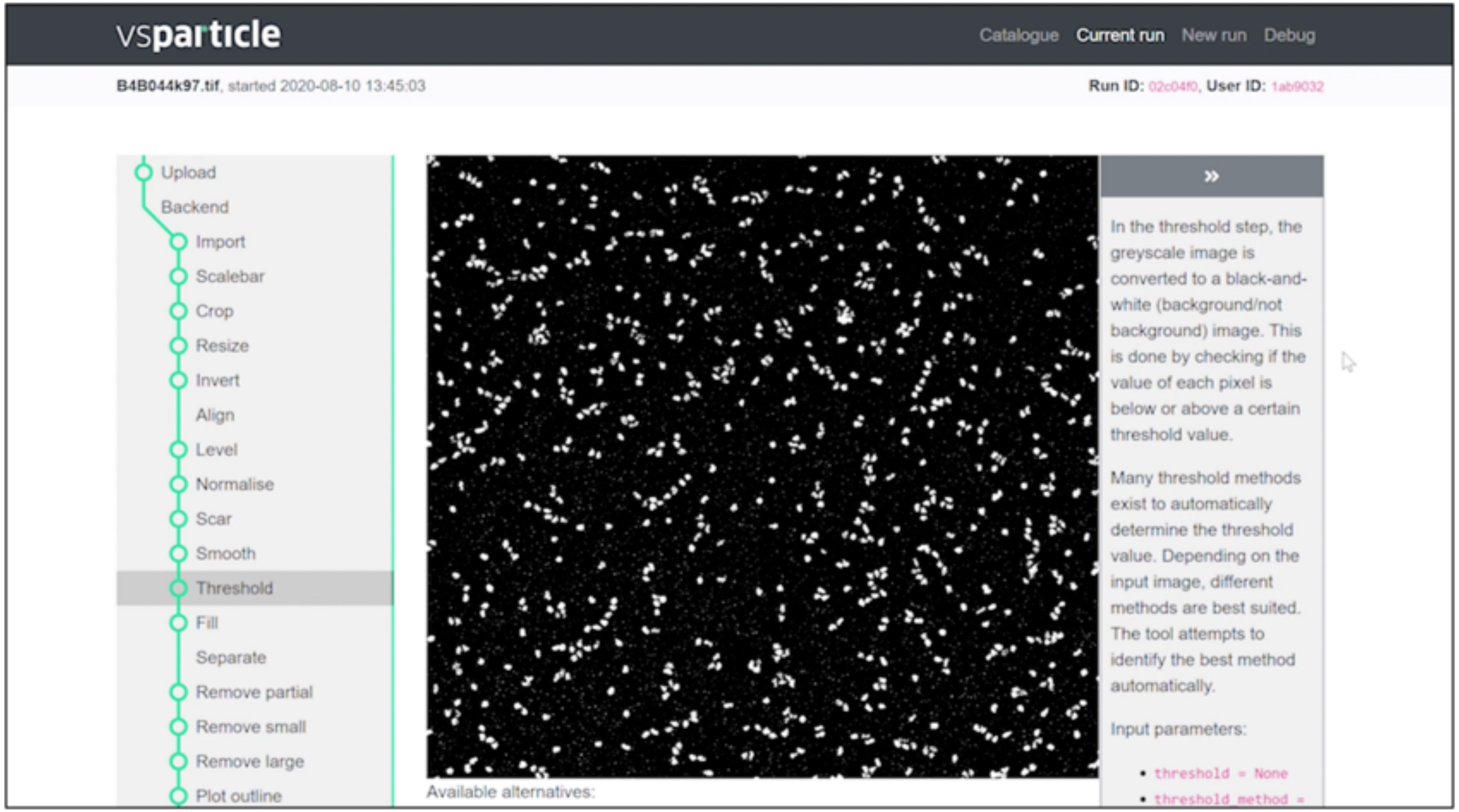
07/12 – 11/12

KLARA, YORAN, OSCAR

NANO — WEEK 14

OBJECTIVE

- VSParticle's image analysis software
- Create a model that predicts the thresholding method



B4B044k97.tif, started 2020-08-10 13:45:03

Run ID: 02c04f0, **User ID:** 1ab9032

In the threshold step, the greyscale image is converted to a black-and-white (background/not background) image. This is done by checking if the value of each pixel is below or above a certain threshold value.

Many threshold methods exist to automatically determine the threshold value. Depending on the input image, different methods are best suited. The tool attempts to identify the best method automatically.

Input parameters:

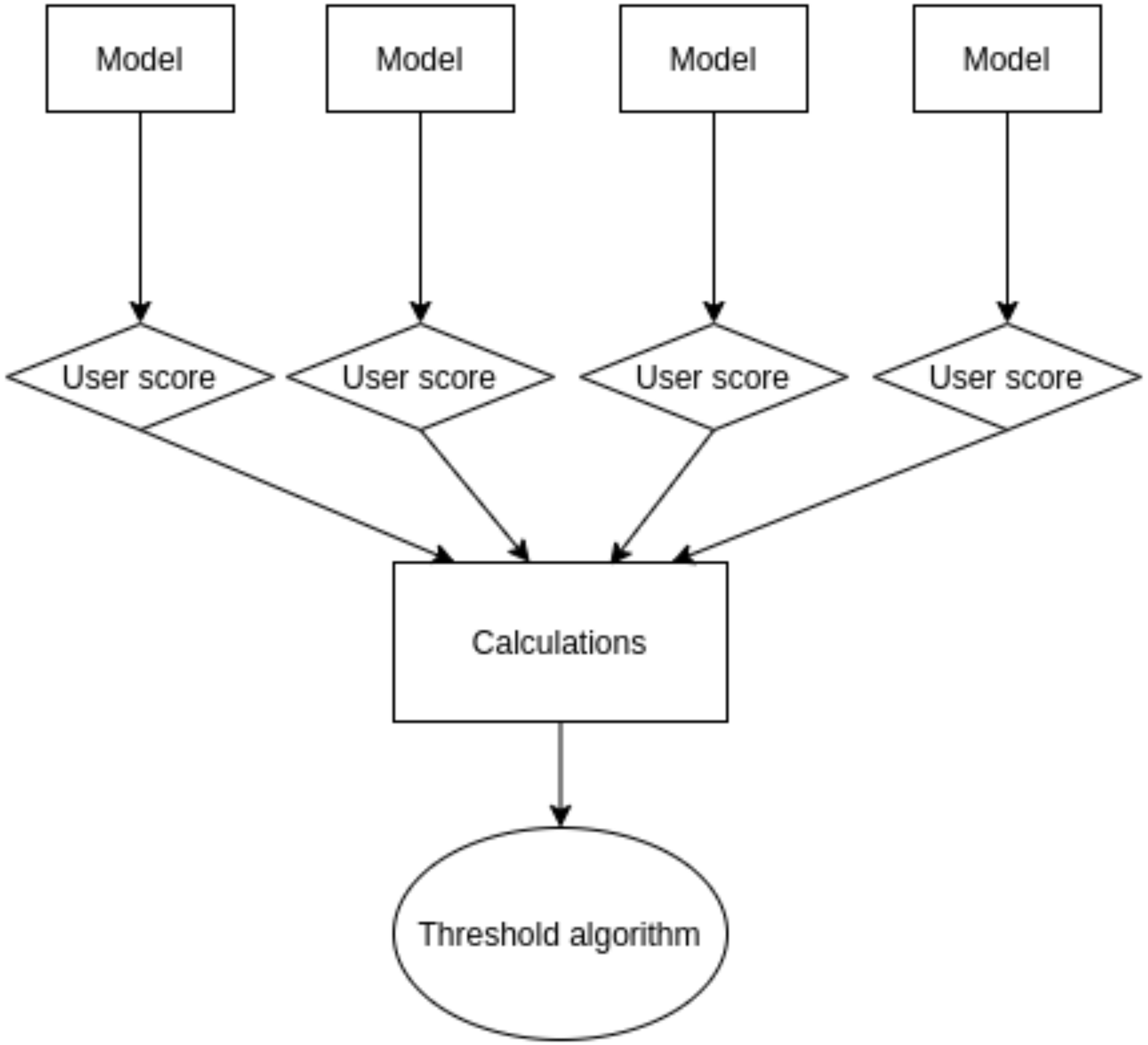
- `threshold = None`
- `threshold_method =`

Available alternatives:

DATASET

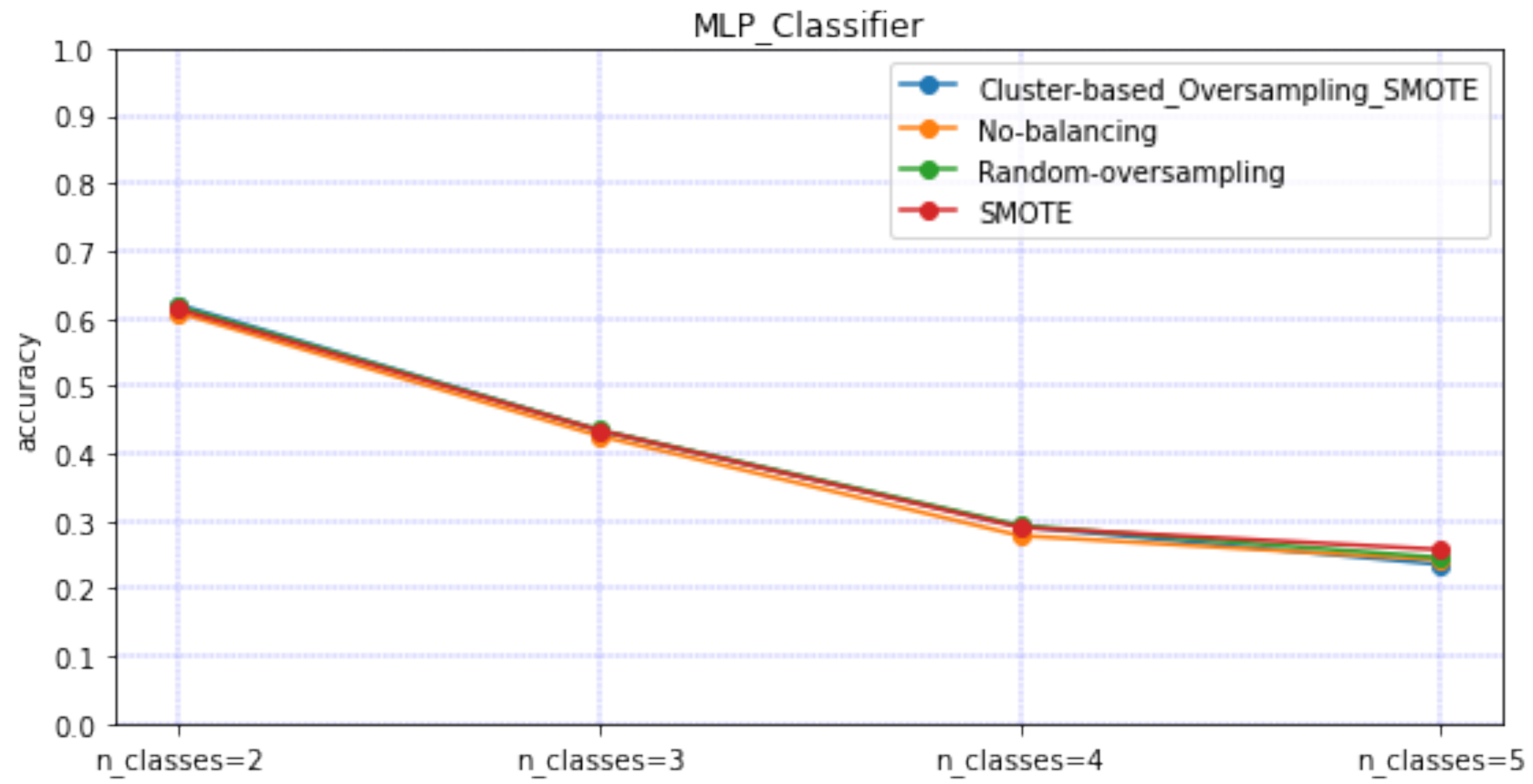
Contains:

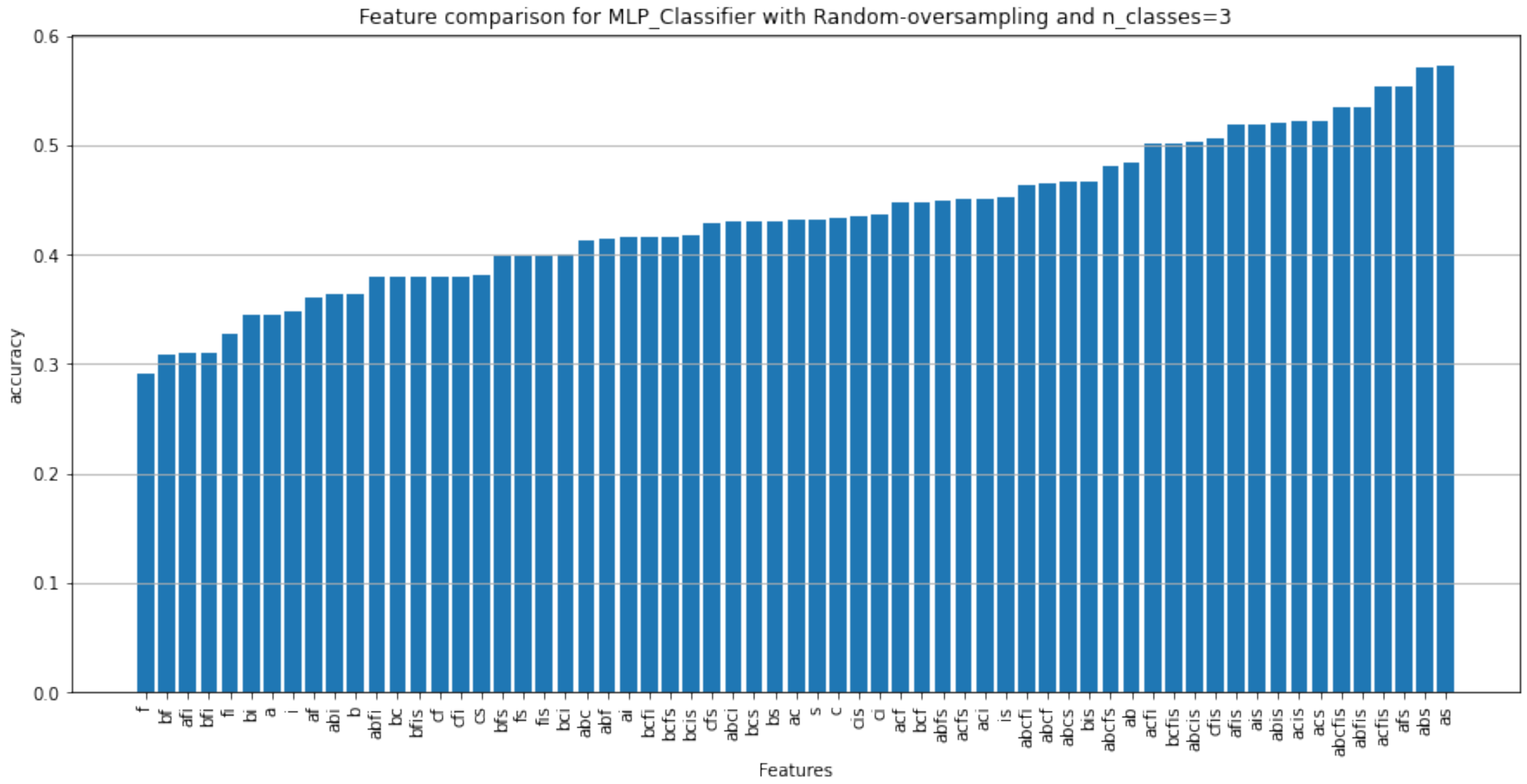
- Used threshold algorithm
- Used algorithm feature scores
- User score

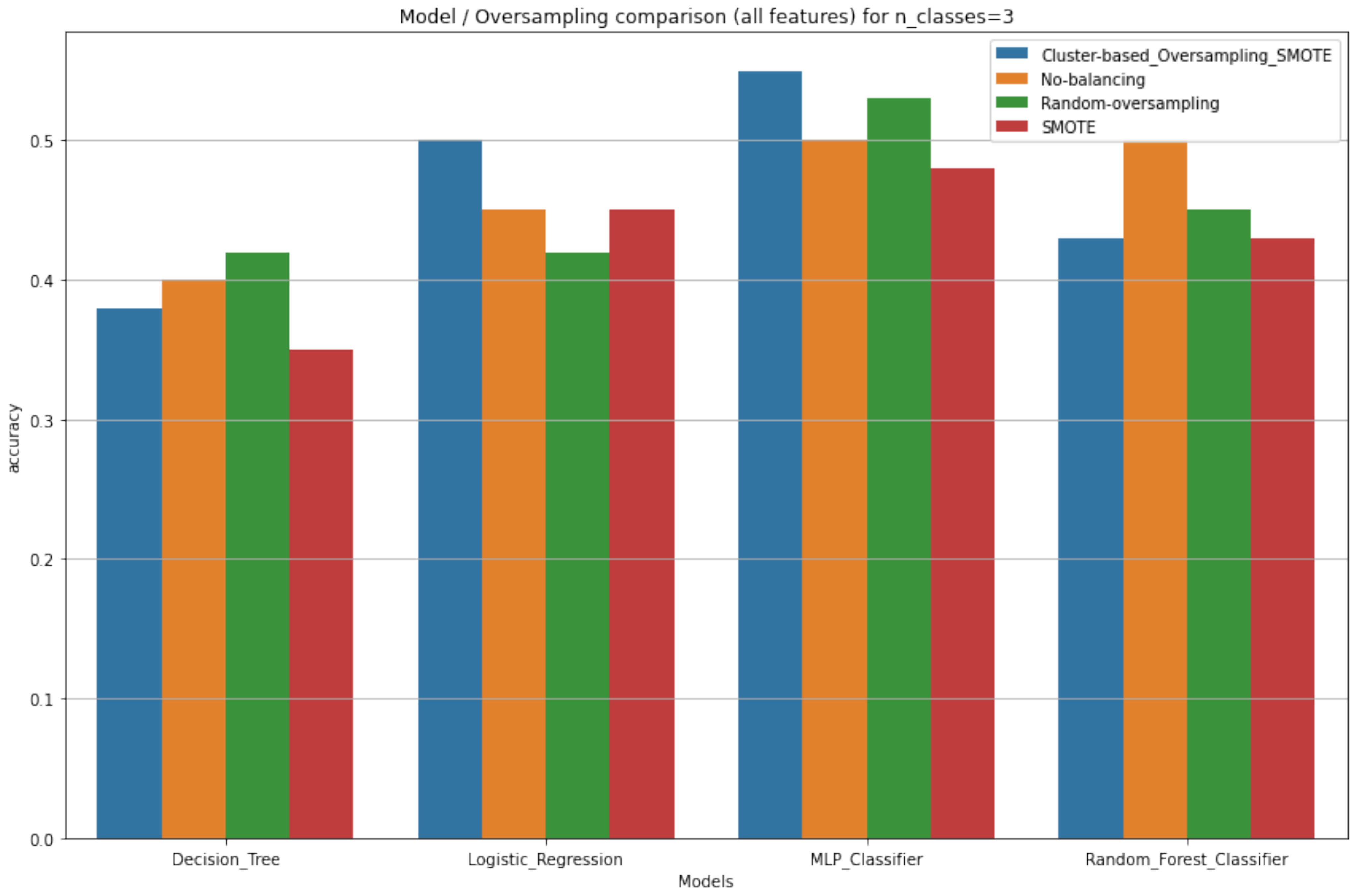


EXPERIMENT

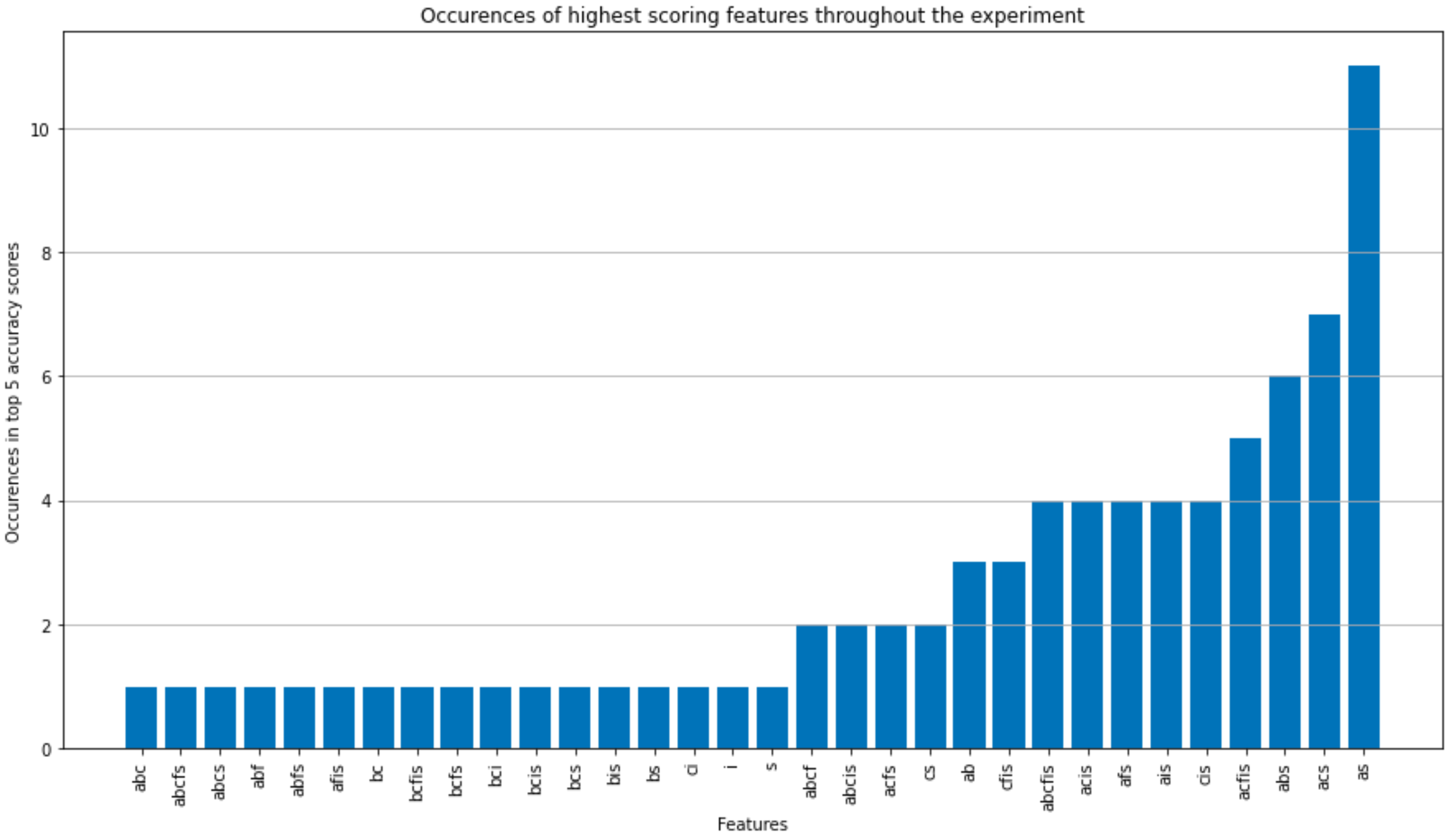
- Combinations of models, features, classes and balancing methods

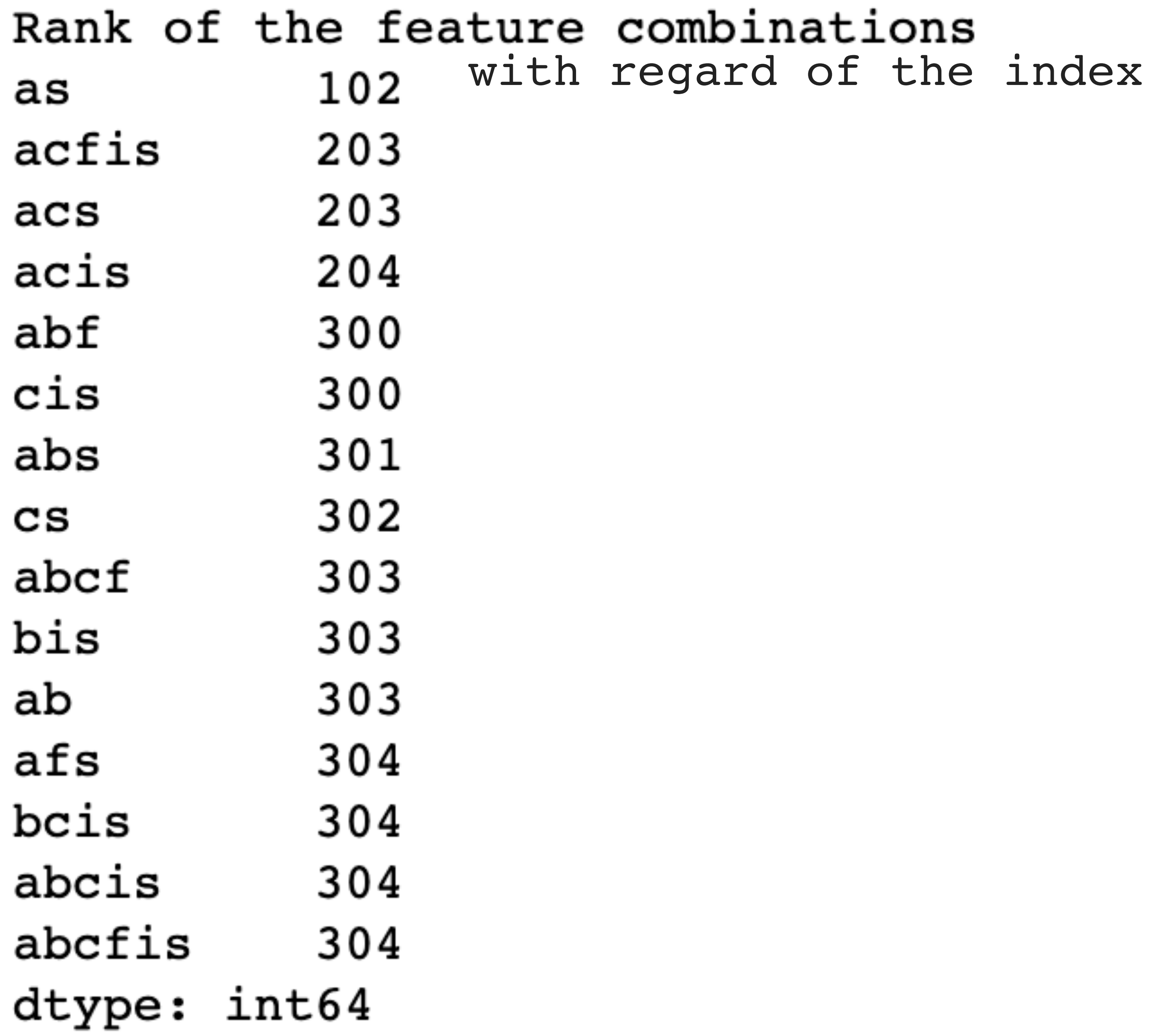






FURTHER ANALYSATION: FEATURE RANKING

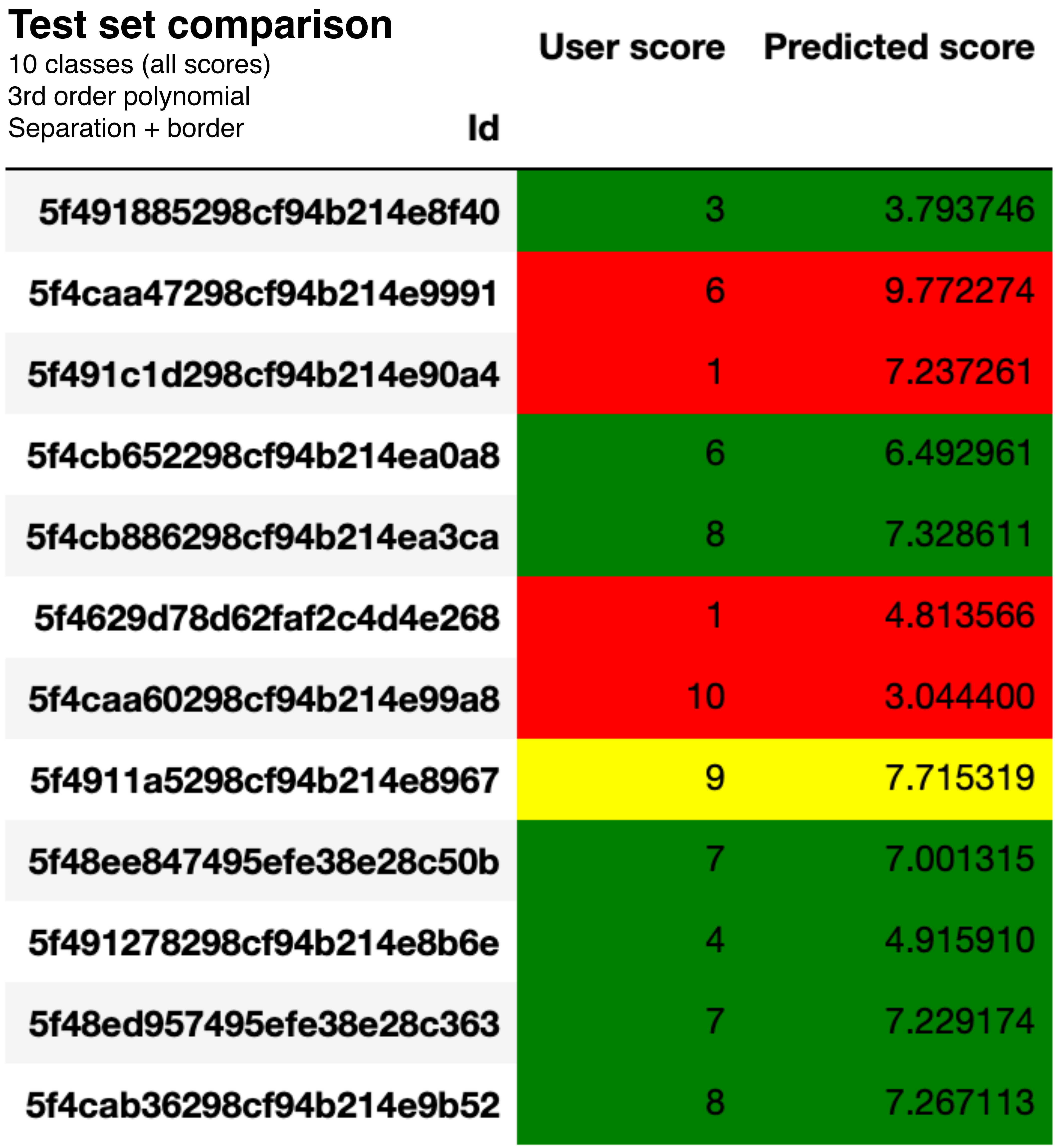




```
Rank of the feature combinations
as          102  with regard of the index
acfis       203
acs         203
acis        204
abf         300
cis         300
abs         301
cs          302
abcf        303
bis         303
ab          303
afs         304
bcis        304
abcis       304
abcfis      304
dtype: int64
```

POLYNOMIAL REGRESSION

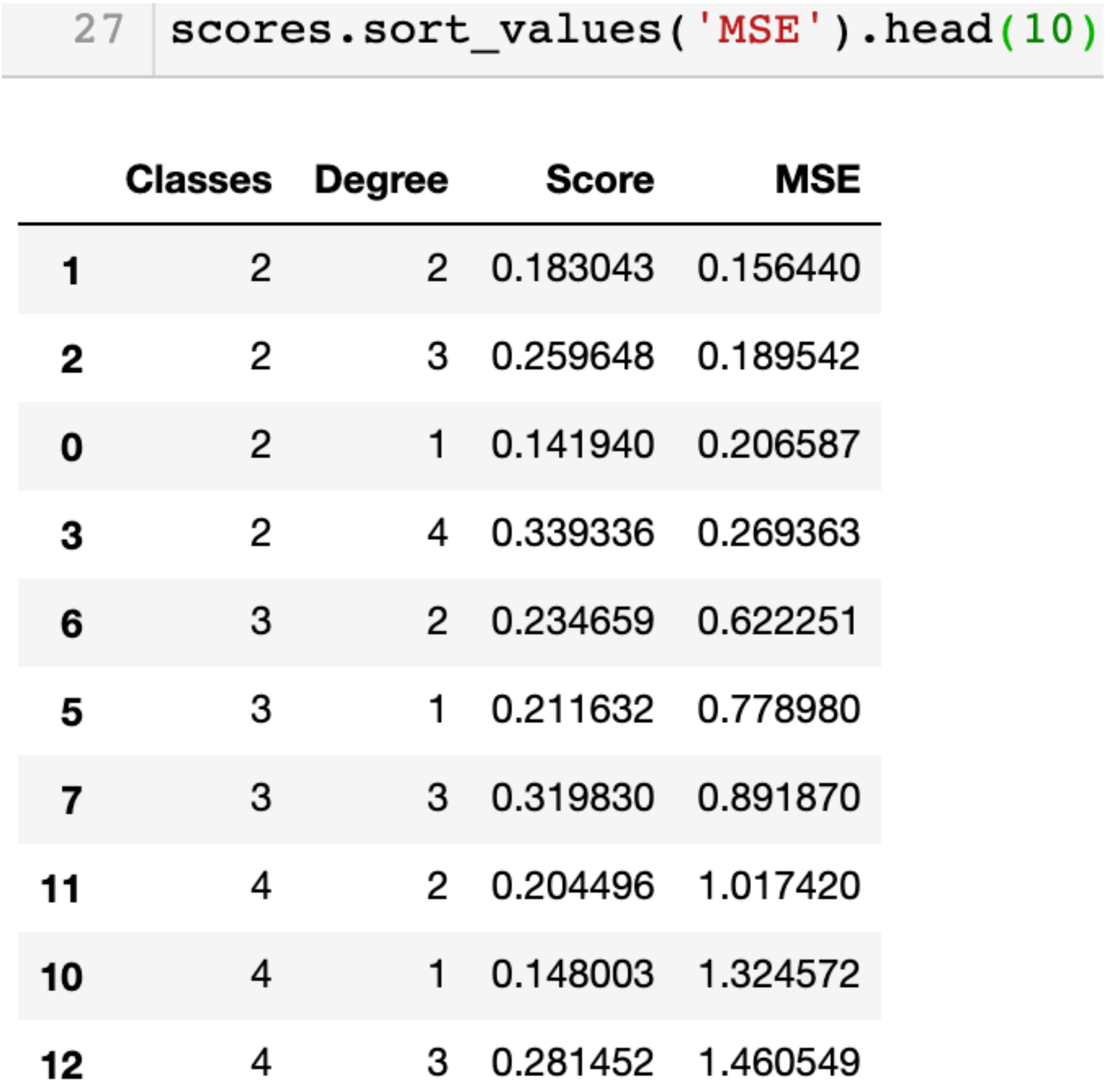
- Suggested by problem owner
- Regression ranks the user scores (10 is better than 4)
- Concern: not every user score has a record for every model



Test set comparison
10 classes (all scores)
3rd order polynomial
Separation + border

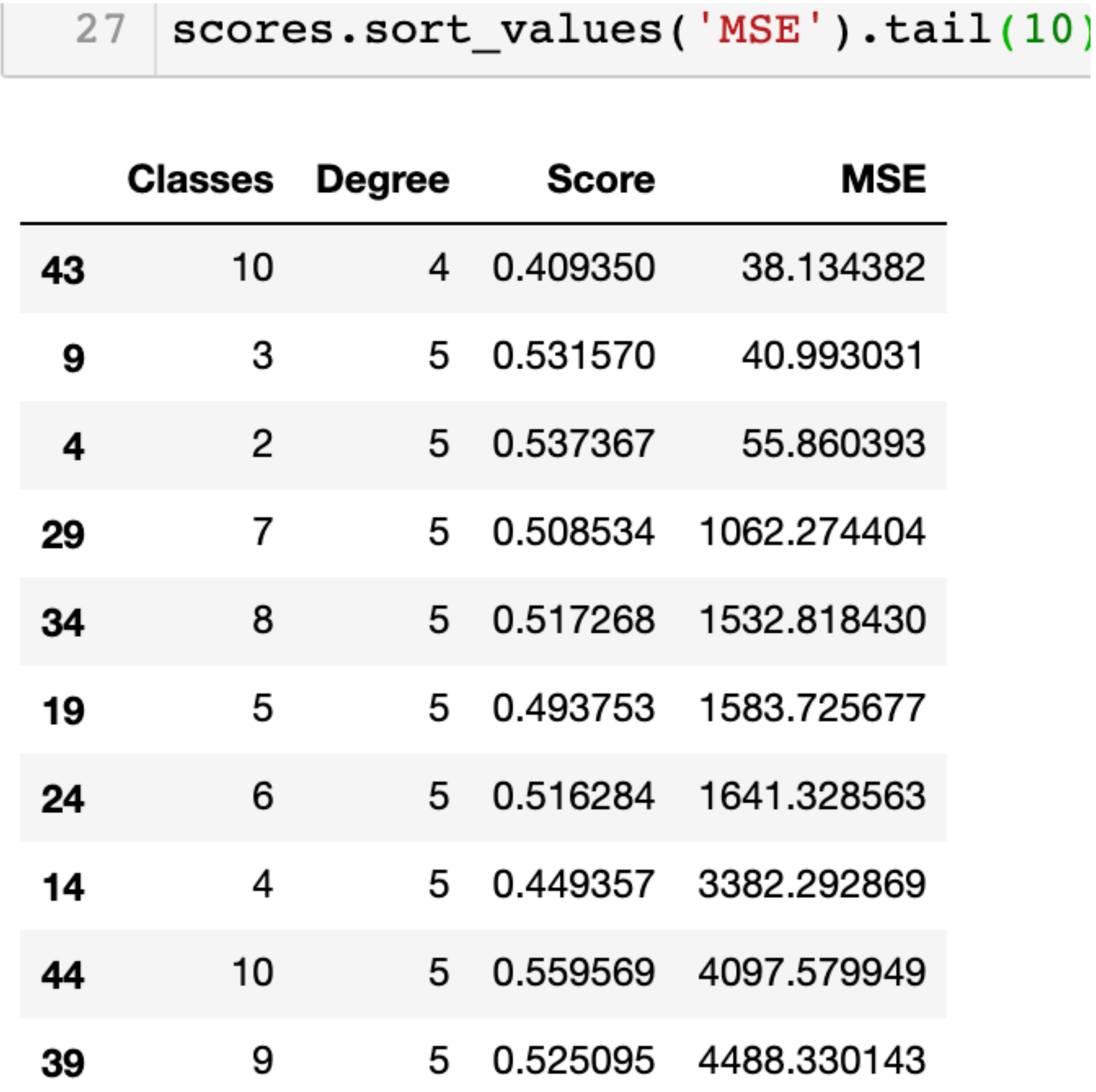
Id	User score	Predicted score
5f491885298cf94b214e8f40	3	3.793746
5f4caa47298cf94b214e9991	6	9.772274
5f491c1d298cf94b214e90a4	1	7.237261
5f4cb652298cf94b214ea0a8	6	6.492961
5f4cb886298cf94b214ea3ca	8	7.328611
5f4629d78d62faf2c4d4e268	1	4.813566
5f4caa60298cf94b214e99a8	10	3.044400
5f4911a5298cf94b214e8967	9	7.715319
5f48ee847495efe38e28c50b	7	7.001315
5f491278298cf94b214e8b6e	4	4.915910
5f48ed957495efe38e28c363	7	7.229174
5f4cab36298cf94b214e9b52	8	7.267113

POLYNOMIAL REGRESSION EXPERIMENT



```
scores.sort_values('MSE').head(10)
```

	Classes	Degree	Score	MSE
1	2	2	0.183043	0.156440
2	2	3	0.259648	0.189542
0	2	1	0.141940	0.206587
3	2	4	0.339336	0.269363
6	3	2	0.234659	0.622251
5	3	1	0.211632	0.778980
7	3	3	0.319830	0.891870
11	4	2	0.204496	1.017420
10	4	1	0.148003	1.324572
12	4	3	0.281452	1.460549



```
scores.sort_values('MSE').tail(10)
```

	Classes	Degree	Score	MSE
43	10	4	0.409350	38.134382
9	3	5	0.531570	40.993031
4	2	5	0.537367	55.860393
29	7	5	0.508534	1062.274404
34	8	5	0.517268	1532.818430
19	5	5	0.493753	1583.725677
24	6	5	0.516284	1641.328563
14	4	5	0.449357	3382.292869
44	10	5	0.559569	4097.579949
39	9	5	0.525095	4488.330143

COMING WEEK

- Have definitive best model
- Optimise model for each threshold method

THANK YOU!

ANY QUESTIONS OR FEEDBACK?