Inferring accreted stellar mass fractions of central galaxies using random forest

Difference from Reference:

Data

Optimisation method:

Bootstrap

Criterion

n-estimators

No of features

Observable Features

Constraints

Arshia Anjum and Sibabrata Biswal

Supervisor- Dr. Subhankar Mishra

Machine Learning Project

Theory - Explained Before

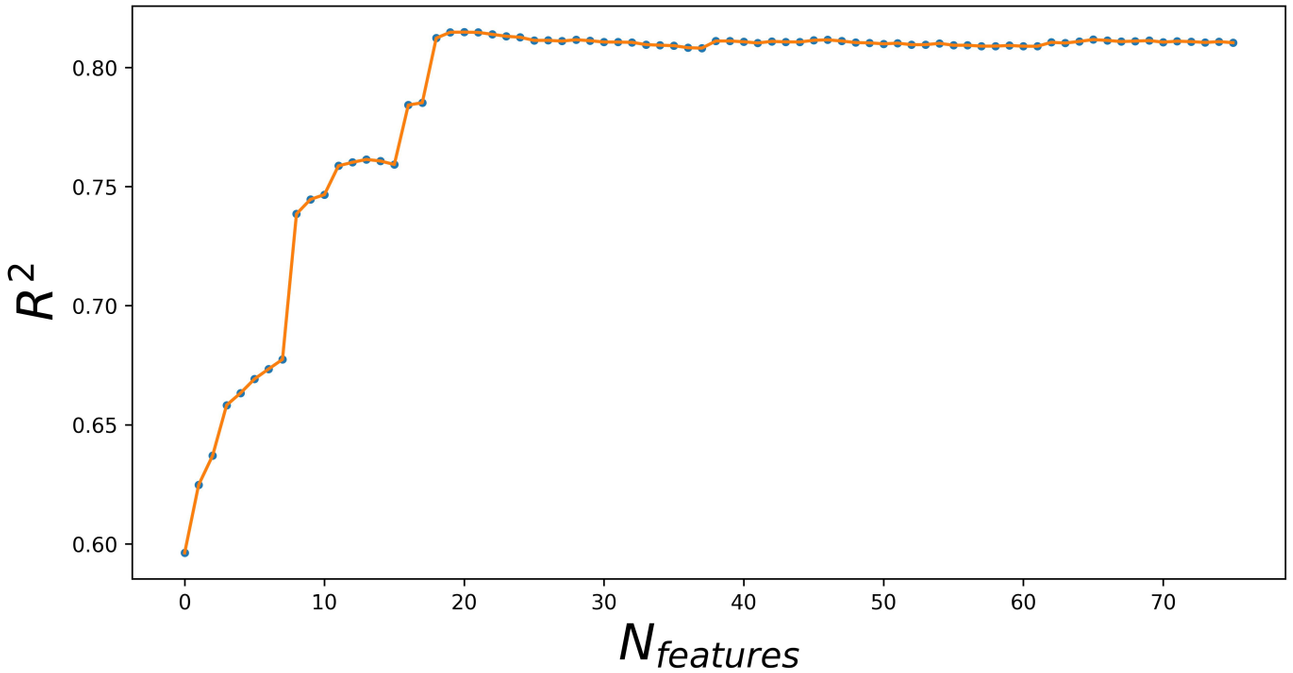
Data

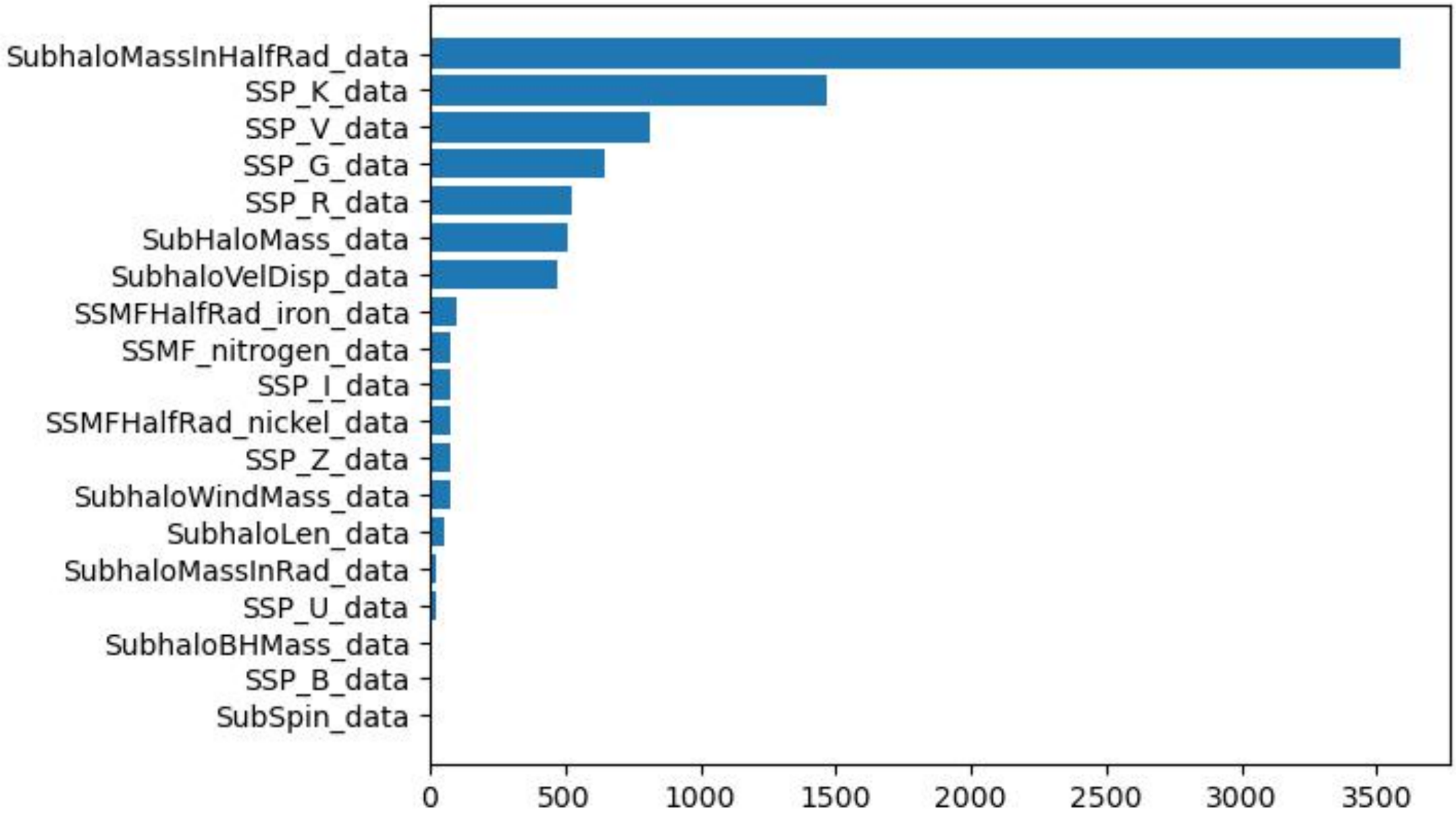
- The snapshots, the group catalogues, the offsets, SKIRT Imaging and the stellar assembly catalogues of the simulation.
- After reading the file and sorting for redshift = 0, we apply our first constraint, i.e. the mass of the central galaxies of the halos should be greater than $10^{10.16} M_{Sun}$.
- The second constraint had to be to check if the simulation gives faithful mock images for the chosen subhalos.
- The third constraint was to check if the central galaxy of the halo is at least 0.5 magnitudes brighter than the satellite galaxies. Skipped in our case.
- As we sort the stellar assembly data, we store the index of the subhalos (the subhalo ids).
- These ids are then checked in SKIRT Imaging data.
- Then they are used to get the galaxy features from the snapshot, group catalogue and offset files.
- The columns are then concatenated, and used for the model training.

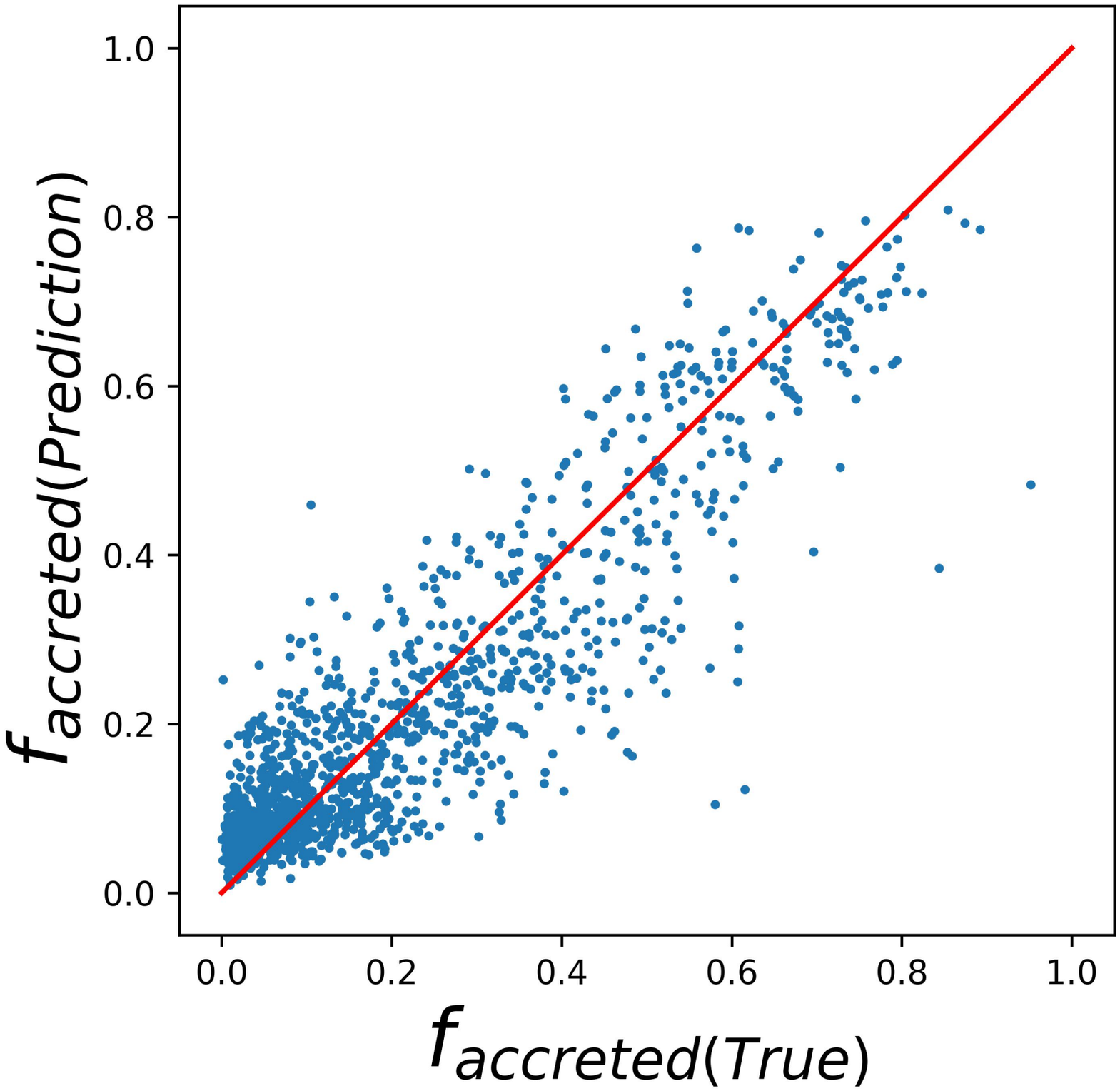
Features Used - Same as Before

Machine Learning Model

- TNG100-1
- R2 score saturates around 0.8136
- including 19 features in the model
- As a result of Grid Search, we have:
- n_estimators = 200,
- bootstrap = True,
- and criterion = mse.
- The dataset had 4586 galaxies included.



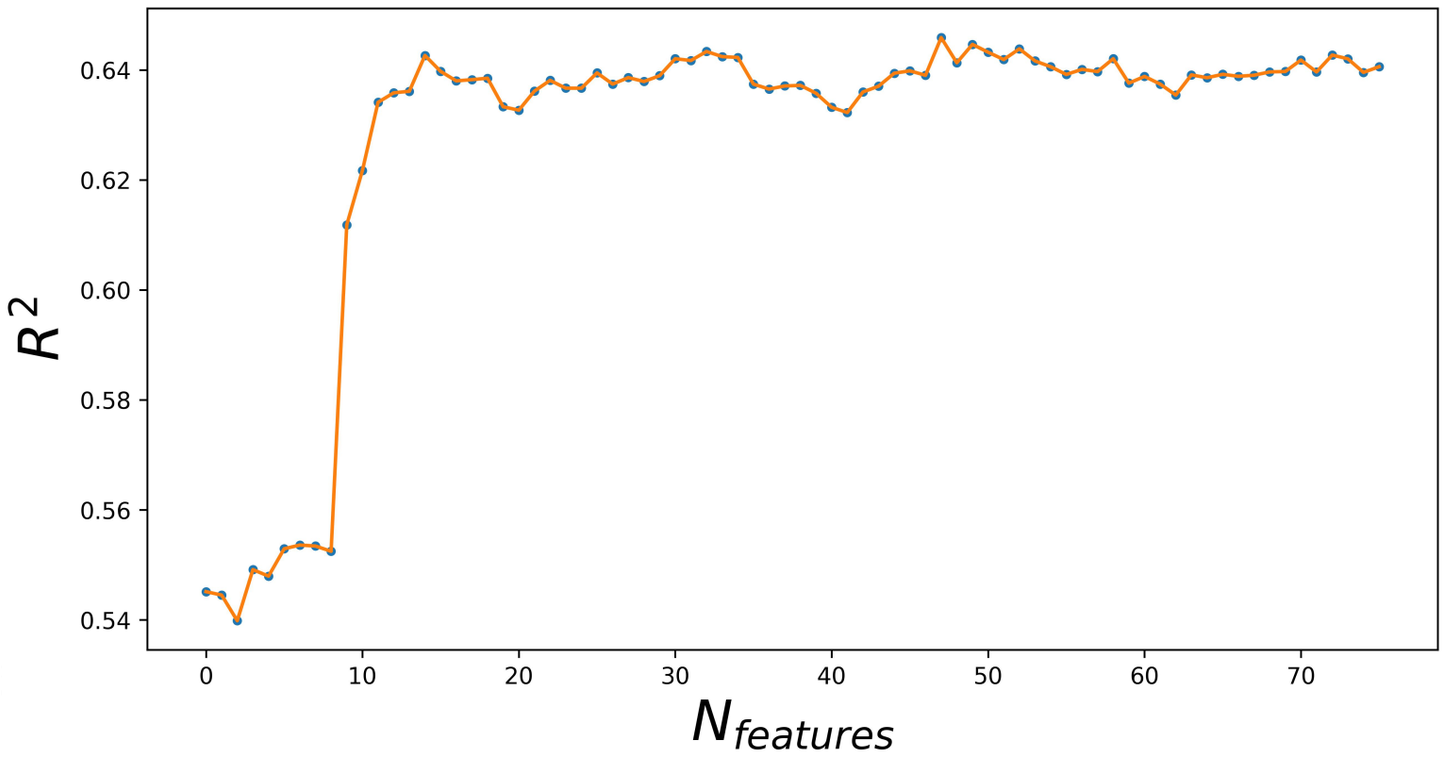

SubhaloMassInHalfRad_data
SSP_K_data
SSP_V_data
SSP_G_data
SSP_R_data
SubHaloMass_data
SubhaloVelDisp_data
SSMFHalfRad_iron_data
SSMF_nitrogen_data
SSP_I_data
SSMFHalfRad_nickel_data
SSP_Z_data
SubhaloWindMass_data
SubhaloLen_data
SubhaloMassInRad_data
SSP_U_data
SubhaloBHMass_data
SSP_B_data
SubSpin_data
0
500
1000
1500
2000
2500
3000
3500

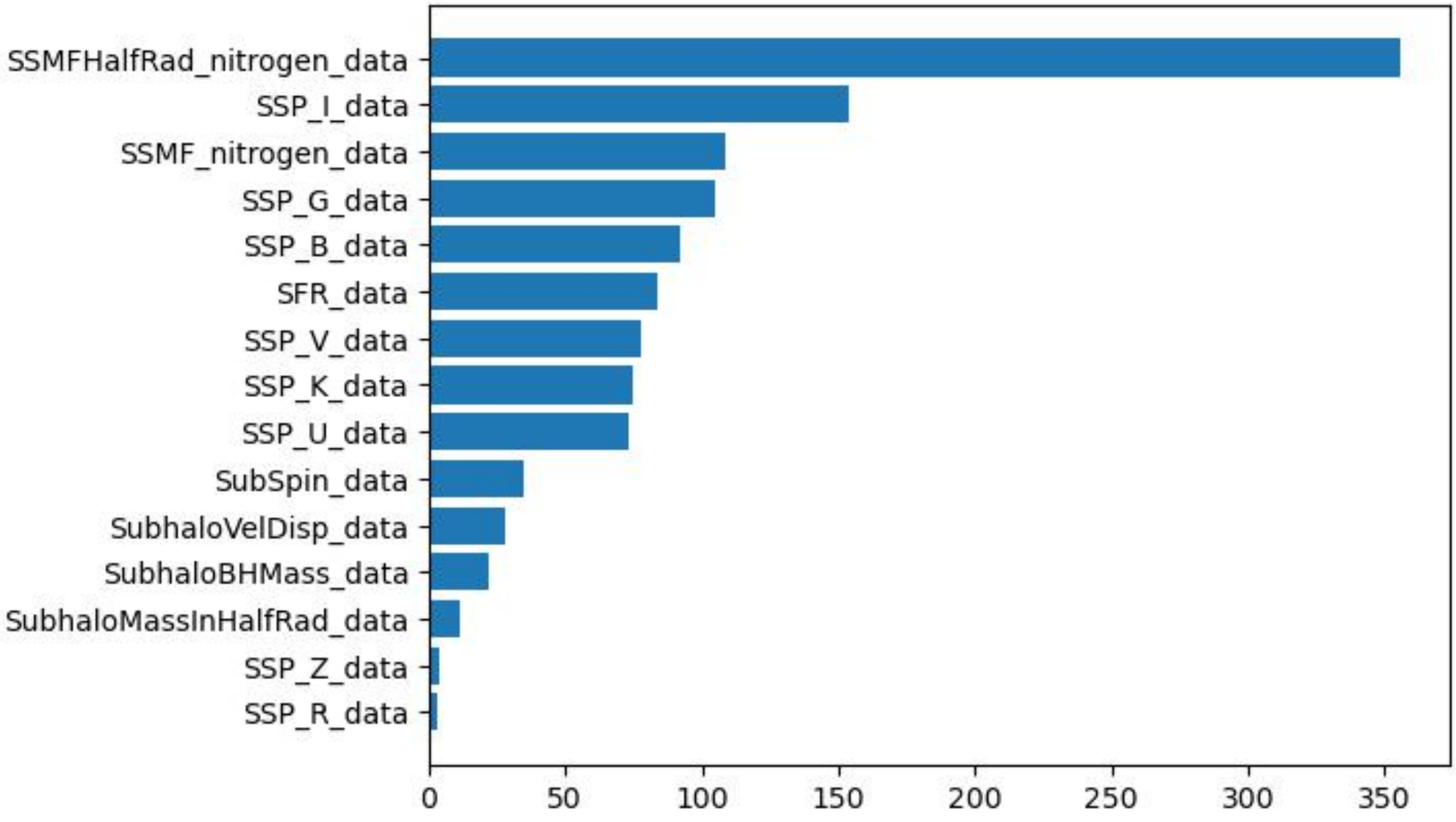

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
$f_{accreted(Prediction)}$
$f_{accreted(True)}$

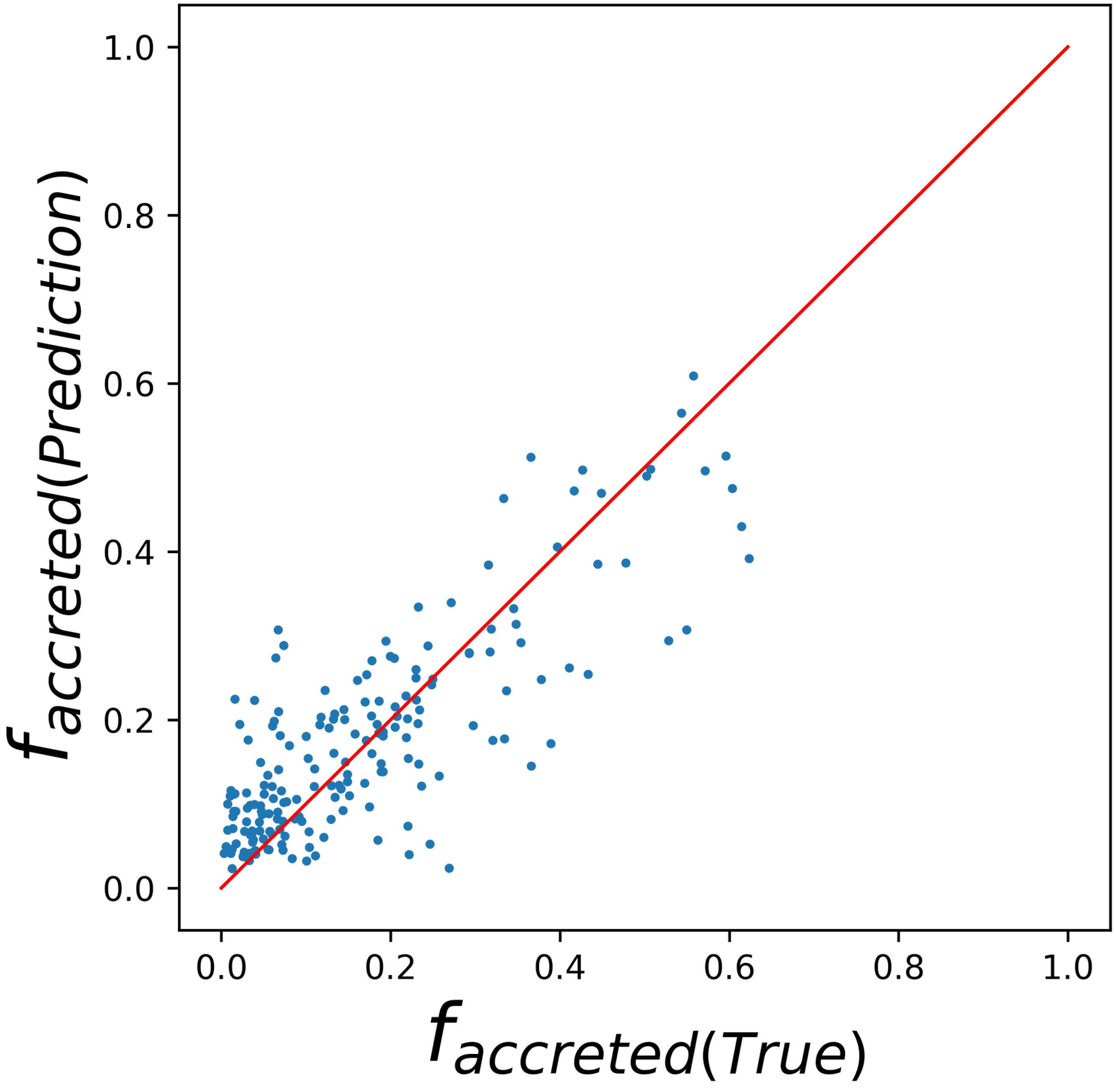
Features Used - Same as Before

Machine Learning Model

- TNG50-1
- R2 score saturates around 0.6636
- including 15 features in the model
- As a result of Grid Search, we have:
- n_estimators = 110,
- bootstrap = True,
- and criterion = mae.
- The dataset had 639 galaxies include




SSMFHalfRad_nitrogen_data
SSP_I_data
SSMF_nitrogen_data
SSP_G_data
SSP_B_data
SFR_data
SSP_V_data
SSP_K_data
SSP_U_data
SubSpin_data
SubhaloVelDisp_data
SubhaloBHMass_data
SubhaloMassInHalfRad_data
SSP_Z_data
SSP_R_data
0
50
100
150
200
250
300
350


1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
$f_{accreted}(Prediction)$
$f_{accreted}(True)$

Future Work

(For publishing)

- However, we shall notice that the dataset used can give a model with accuracy reaching ~85%.
- As for the publishing of this paper, more insight on the feature corelations shall be shown, alongwith analysed TNG300-1 data, both of which will be done within next few days.
- We will also be trying a model or two to better the accuracy as mentioned before, however no word for it.
- Using the mass constraints did not work for us, as it did not give a better model, however that can be added in the final submission for the sake of completeness. Same goes for the use of surface brightness profiles.

References

[1] R. Shi, W. Wang, Z. Li, J. Han, J. Shi, V. Rodriguez-Gomez, Y. Peng and Q. Li, Mon. Not. Roy. Astron. Soc. **515** (2022) no.3, 3938-3955 doi:10.1093/mnras/stac1541 [arXiv:2112.07203 [astro-ph.GA]].

[2] V. Rodriguez-Gomez, A. Pillepich, L. V. Sales, S. Genel, M. Vogelsberger, Q. Zhu, S. Wellons, D. Nelson, P. Torrey and V. Springel, *et al.* Mon. Not. Roy. Astron. Soc. **458** (2016) no.3, 2371-2390 doi:10.1093/mnras/stw456 [arXiv:1511.08804 [astro-ph.GA]].

[3] A. Pillepich, M. Vogelsberger, A. Deason, V. Rodriguez-Gomez, S. Genel, D. Nelson, P. Torrey, L. V. Sales, F. Marinacci and V. Springel, *et al.* Mon. Not. Roy. Astron. Soc. **444** (2014) no.1, 237-249 doi:10.1093/mnras/stu1408 [arXiv:1406.1174 [astro-ph.GA]].

[4] D. Montenegro-Taborda, V. Rodriguez-Gomez, A. Pillepich, V. Avila-Reese, L. V. Sales, A. Rodríguez-Puebla and L. Hernquist, doi:10.1093/mnras/stad586 [arXiv:2302.10943 [astro-ph.GA]].

[5] S. Tacchella, B. Diemer, L. Hernquist, S. Genel, F. Marinacci, D. Nelson, A. Pillepich, V. Rodriguez-Gomez, L. V. Sales and V. Springel, *et al.* Mon. Not. Roy. Astron. Soc. **487** (2019) no.4, 5416-5440 doi:10.1093/mnras/stz1657 [arXiv:1904.12860 [astro-ph.GA]].

[6] S. Wellons, P. Torrey, C. P. Ma, V. Rodriguez-Gomez, M. Vogelsberger, M. Kriek, P. van Dokkum, E. Nelson, S. Genel and A. Pillepich, *et al.* Mon. Not. Roy. Astron. Soc. **449** (2015) no.1, 361-372 doi:10.1093/mnras/stv303 [arXiv:1411.0667 [astro-ph.GA]].