Tutorial

Straightforwarding Scoring Suite (3*S*)

Version 1.0.0

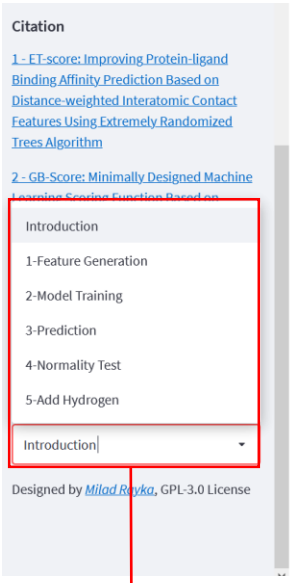
Author: Milad Rayka

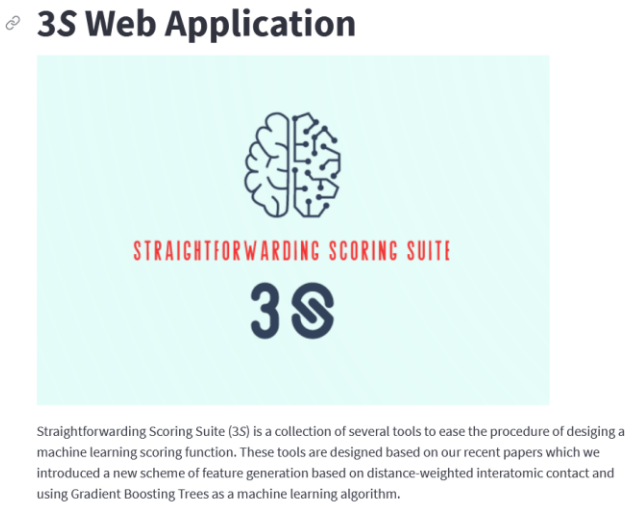
Table of contents

1 – Introduction

Straightforwarding Scoring Suite (3*S*) is a collection of several tools to ease the procedure of desiging a machine learning scoring function. These tools are designed based on our recent papers which we introduced a new scheme of feature generation based on distance-weighted interatomic contact and using Gradient Boosting Trees as a machine learning algorithm.

So far, this suite contains five tools: 1-Feature Generation 2-Model Training 3-Prediction 4-Normality Test 5-Add Hydrogen.

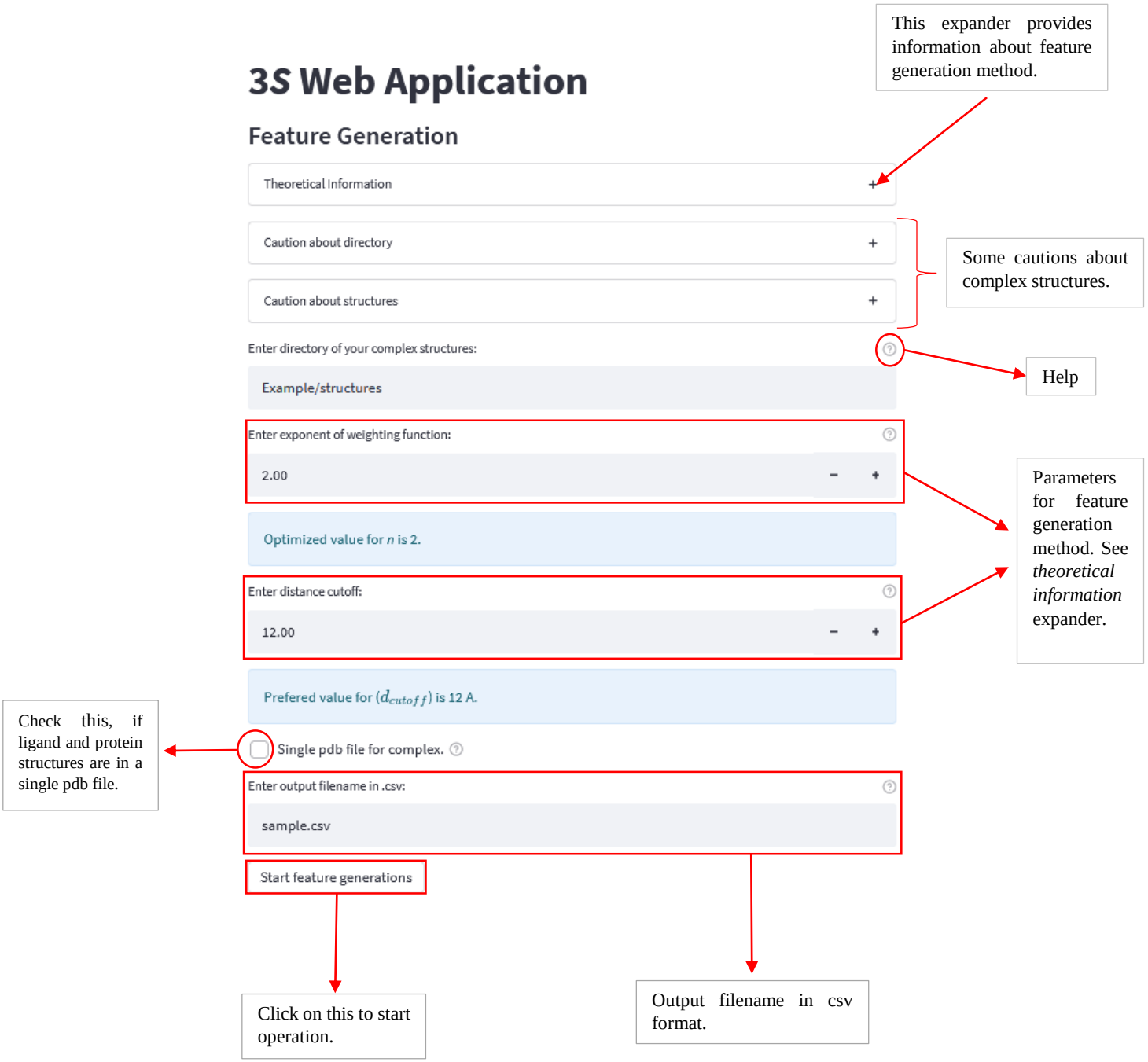




Different operations can be choosen from this side pandel.

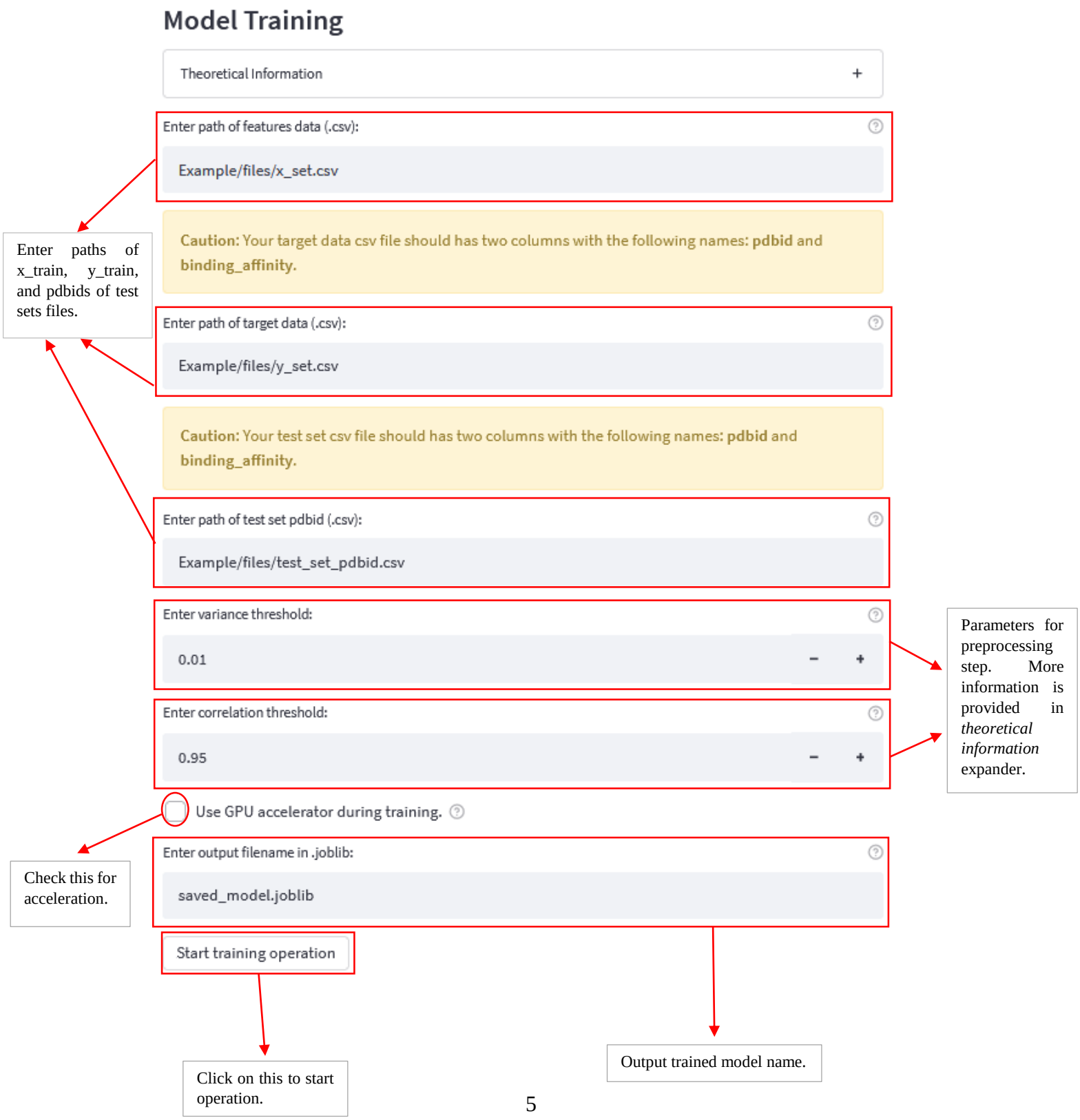
2 – Feature Generation

In this mode, features for different structure of complexes based on aforementioned method are genereted.



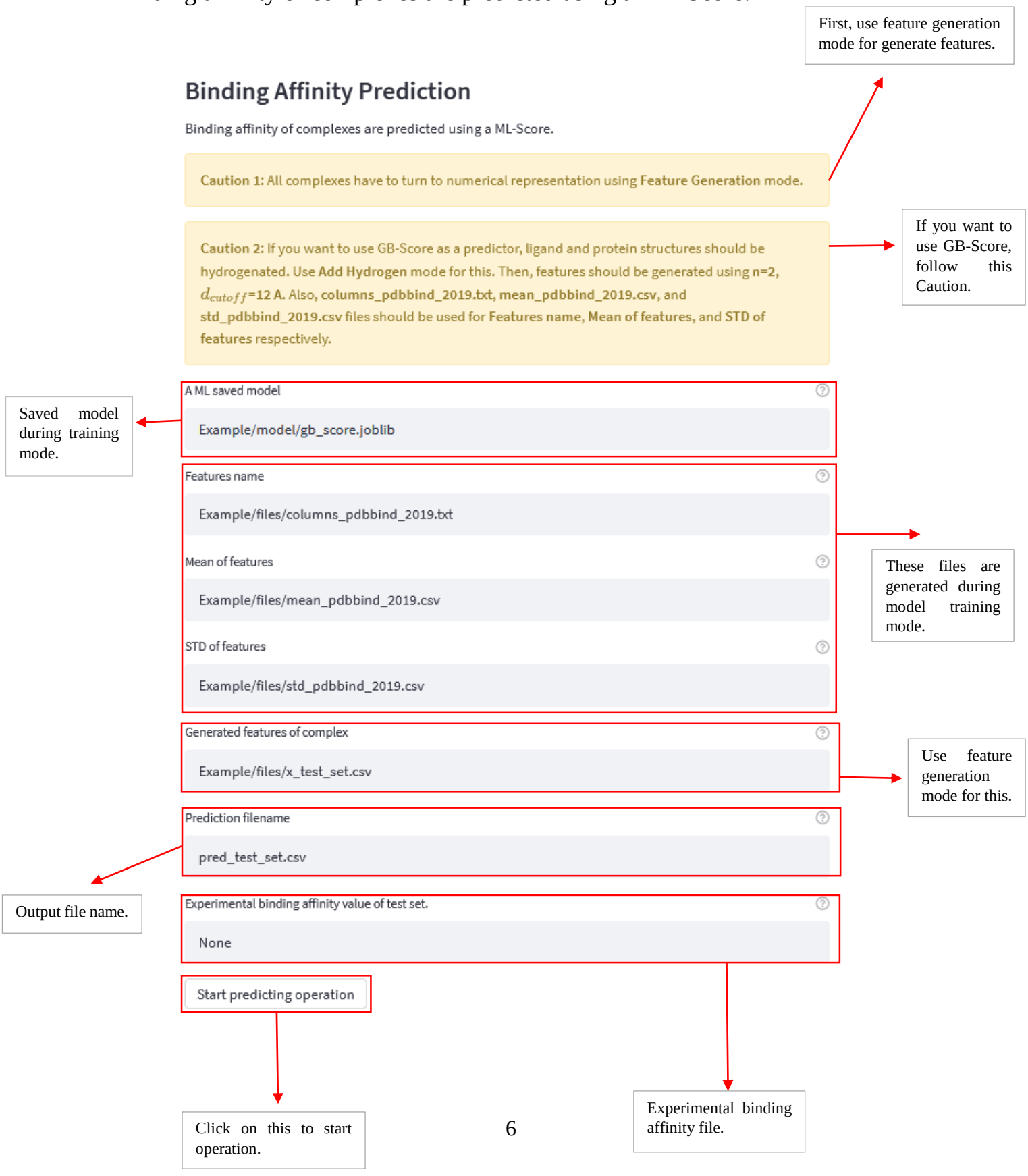
3 – Model Training

In this mode, a machine learning scoring function (Gradient Boosting Trees) is designed for a dataset of provided complex structures.



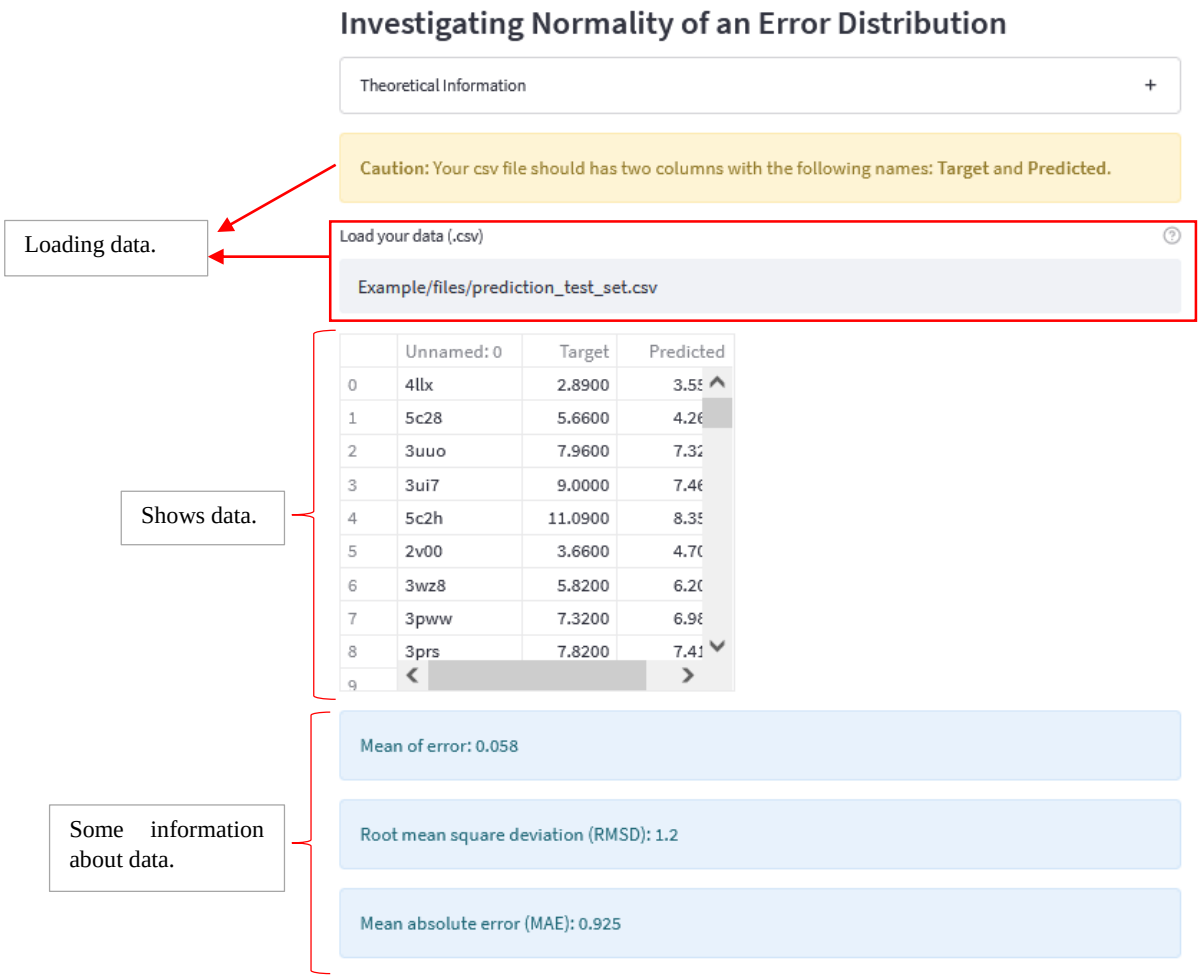
4-Prediction

Binding affinity of complexes are predicted using a ML-Score.



5-Normality Test

In this mode, if the test data has binding label, normality property of errors is analysed.



Shapiro-Wilk (W) Test

Theoretical Information +

W

0.9936

Pr

0.273

So, error distribution is normal ($Pr > 0.05$).

Skewness and Kurtosis Tests

Theoretical Information +

Skew

-0.0149

Kurtosis

3.3

Error distribution is asymmetric and leptokurtic.

95^{th} quantile of the absolute errors distribution

Theoretical Information +

Q95

2.5812

Tests about normality of data.

Save the plot and change the color of plot.

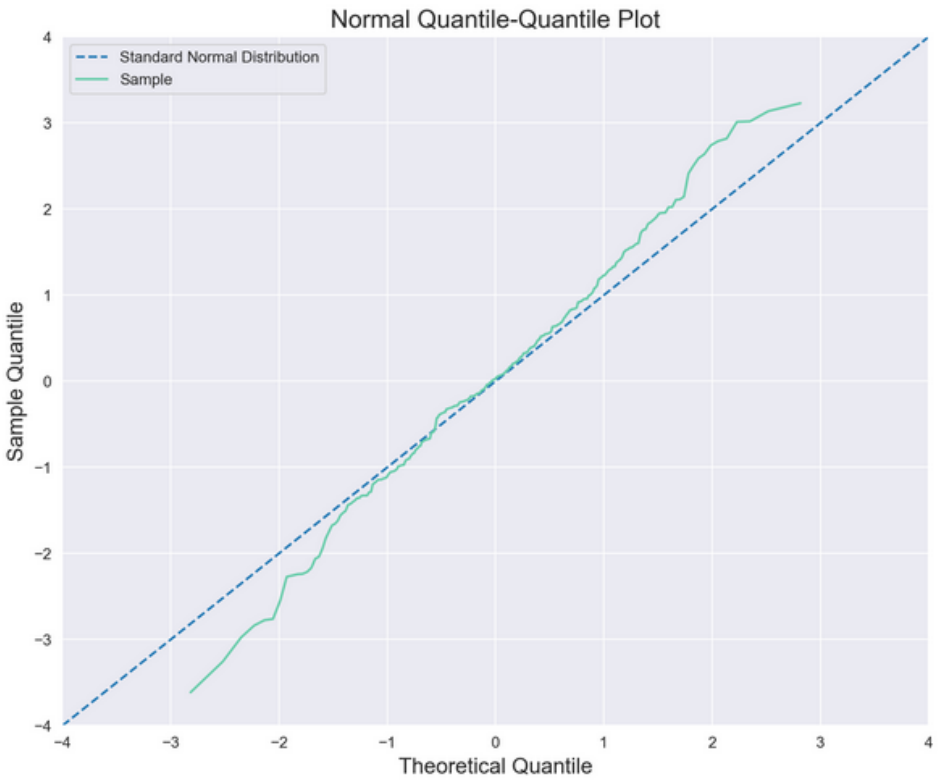
Normal Quantile-Quantile Plot

Theoretical Information +

Save QQ-plot in .png

Choose color for QQ-plot:

Show QQ-plot.



Caution: To download plot, first check *Save QQ-plot in .png* then the download button appears.

Histogram of Error Distribution

Theoretical Information +

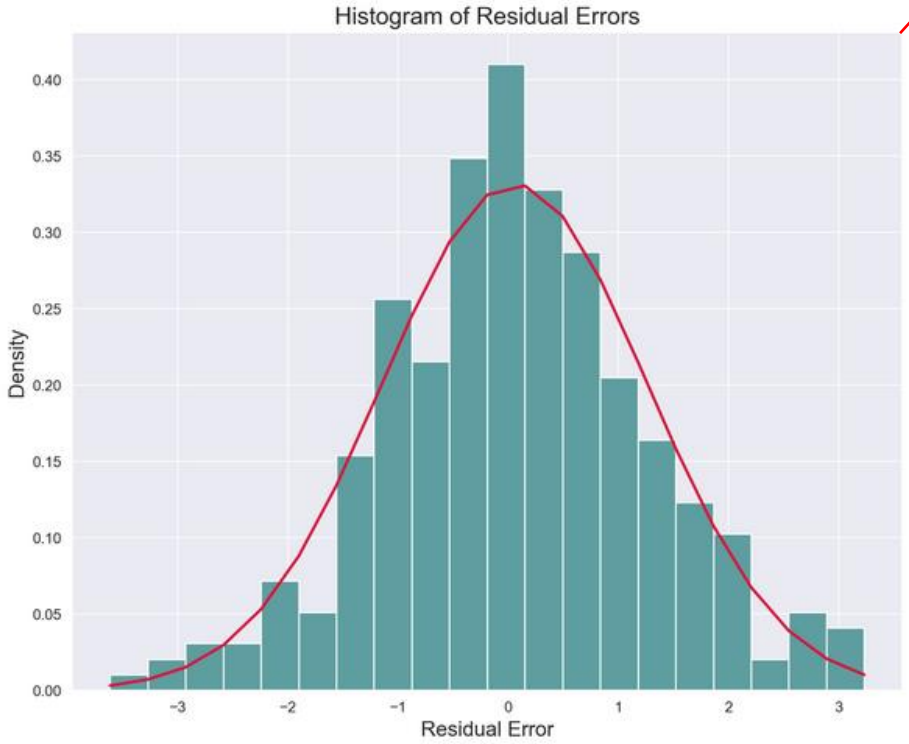
Save histogram plot in .png.

Choose color for histogram plot:

Choose color for guassian plot:

Save and change the colors of plot.

Shows histogram of residual errors.



Caution: To download plot, first check *Save histogram plot in .png* then the download button appears.


Outliers Plot
Theoretical Information
+
Choose a specific percentile
90
0
100
Silder for choosing percentiles.
Save outliers plot in .png
Choose color for non-outliers:
Choose color for outliers:
Outliers plot.
Save and change the colors of the plot.
Error Distribution
Residual Error
Experimental
3
2
1
0
−1
−2
−3
2
4
6
8
10
12
Caution: To download plot, first check save outliers plot in .png then the download button appears.

6-Add Hydrogen

Add hydrogens to ligand and protein at pH=7.4 using PDB2PQR and Openbabel.