

## Instruction for Decision Tree Model

- The model contains 3 sub-models for ploidy, protein weight, and RNA weight.
- The data of the simulation must be save as pickle.
- The result of the prediction is saved as csv file as the format:  
file\_path|type\_of\_mutant
- **tree\_dataprocessing.py** is used to initialised the training data for the sub-models. Modify the path of the simulations which used to train the model on line 193. Modify the path of the wild type simulation which used to create the scale for normalisation on line 197. If the length of the training simulations is 50001, remove the hastage for line 163.
- **svm\_models.py** contains the training processing of the sub-models and data processing of the new data for prediction. It the length of the training is 50,001, add hashtag from line 140 to line 144, and remove the hashtag for line 145.
- **tree\_classifier.py** is the tree model used to categorise the mutants of the simulations. This is the main program to start the classifier.
- **data\_protein.pickle**, **data\_rna.pickle** are the data from run1 for 1kos. Each pickle contains the first value, last value, gradient and standard deviation of simulations which have the length equal to the maximum runtime.
- **data\_ploidy.pickle**, **data\_pinched.pickle** are the data from run1 for 1kos. Each pickle contains all data point in the simulations which have the length equal to the maximum runtime.
- **scale.pickle** records the scale from the wild type simulation. It is used to normalize the data from protein and RNA weight.

### How to create the training data?

Step 1: Set the paths and the MAX\_LENGTH of the simulations

Step 2: In the terminal change the directory where the tree\_dataprocessing.py is located

Step 3: Type 'python3 tree\_dataprocessing.py' in the terminal

### How to use the model?

Step 1: Start the terminal at the folder containing all the files above

Step 2: Type 'python3 tree\_classifier.py' in the terminal

Step 3: Input the path of the simulations you want to predict