**Guideline for Running Code**

1. **Data**

**Xcont.csv**: GWAS data of Oryza Sativa

**cont\_LMM\_grain.weight\_p0.05\_sort.tsv**: grain weight data after filtering by biological characteristics,

**cont\_LMM\_time.flowering\_p0.05\_sort.tsv**: time to flowering data after filtering by biological characteristics,

**Ycont.txt**: output data including phenotypes of time to flowering and grain weight,

**indep\_1000\_10\_0.3.prune-43k.in**: input data for advanced regressor after filtering by biological characteristics.

We have compressed them into a file *Prepared\_data.rar.*

1. **Filtering Data**

Python code: **BI\_Rice\_SNPs\_Pvl\_Filtering.py**

Filtering the data with p value, then put it into the models for training. We have also saved the results in *Export\_data.rar.*

1. **Training Model**

Python code: **BI\_Rice\_SNPs.py**

The data from *Export\_data.rar* is used to train the model.

1. **Grid Search with cross-validation:** [sklearn.model\_selection.GridSearchCV](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)

Used with random forest regression and support vector regression.

1. **Radom forest regression:** [sklearn.ensemble.RandomForestRegressor](https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html)
2. **Support vector regression:** [sklearn.svm.SVR](https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html)
3. **Lasso with cross-validation:** [sklearn.linear\_model.LassoCV](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LassoCV.html)
4. **Multi-task Lasso with cross-validation:** [sklearn.linear\_model.MultiTaskLassoCV](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.MultiTaskLassoCV.html)
5. **Elastic Net with cross-validation:** [sklearn.linear\_model.ElasticNetCV](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.ElasticNetCV.html)
6. **Multi-task Elastic Net with cross-validation:** [sklearn.linear\_model.MultiTaskElasticNetCV](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.MultiTaskElasticNetCV.html)

The results have been compressed into the file *Results.rar.*

*Note: Python code with Python Version 3.7.1*