Yinan Guo

guoy221@wfu.edu

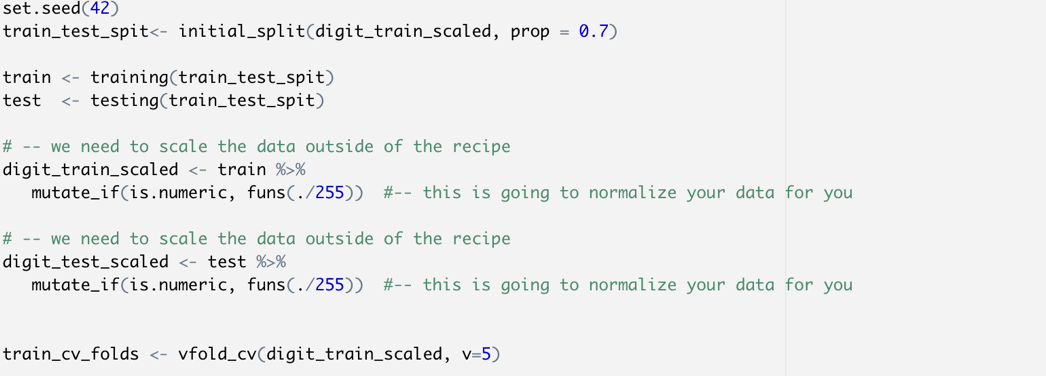
2021/11/21

# **Model Performance**

|  |  |  |
| --- | --- | --- |
|  | Accuracy | |
| Train | Test |
| MLP | 0.902 | 0.875 |
| Random Forest | 1.000 | 0.967 |
| XGBoost | 1.000 | 0.978 |

According to the accuracy of three models, XGBoost has the highest accuracy in test set, which is 0.978, meaning that 97.8% of the numbers in test pictures are predicted right.

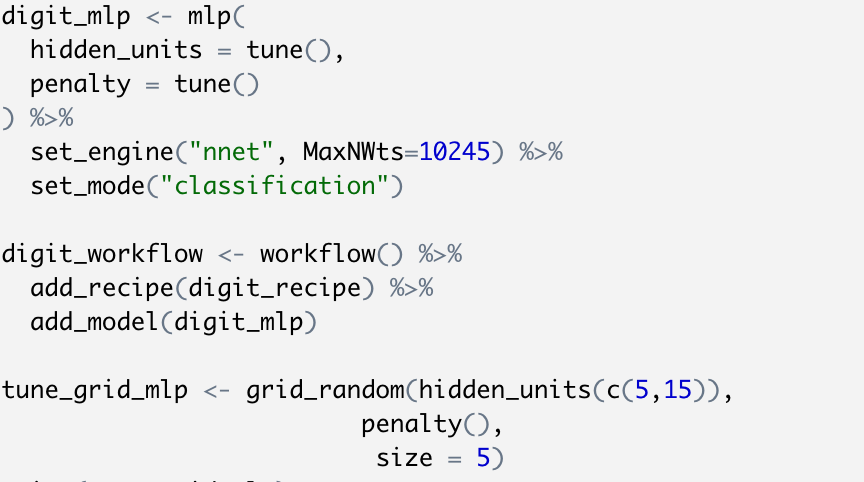
# **K-Fold Tuning**



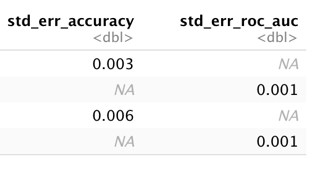
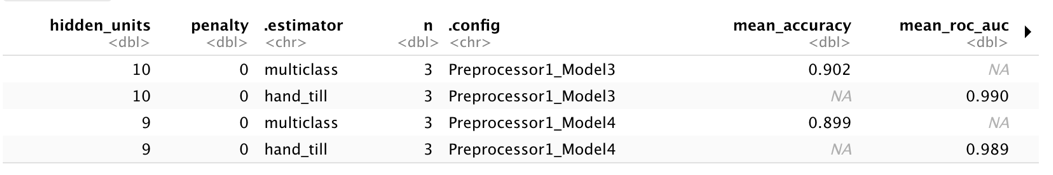
The train set is divided by 5 folds to make cross validation.

And in the tune grid part, the algorithm randomly selected 5 groups of different parameters for mlp, rf and xgb to build the model. The tuning results are shown as following:

* **Neural Network**

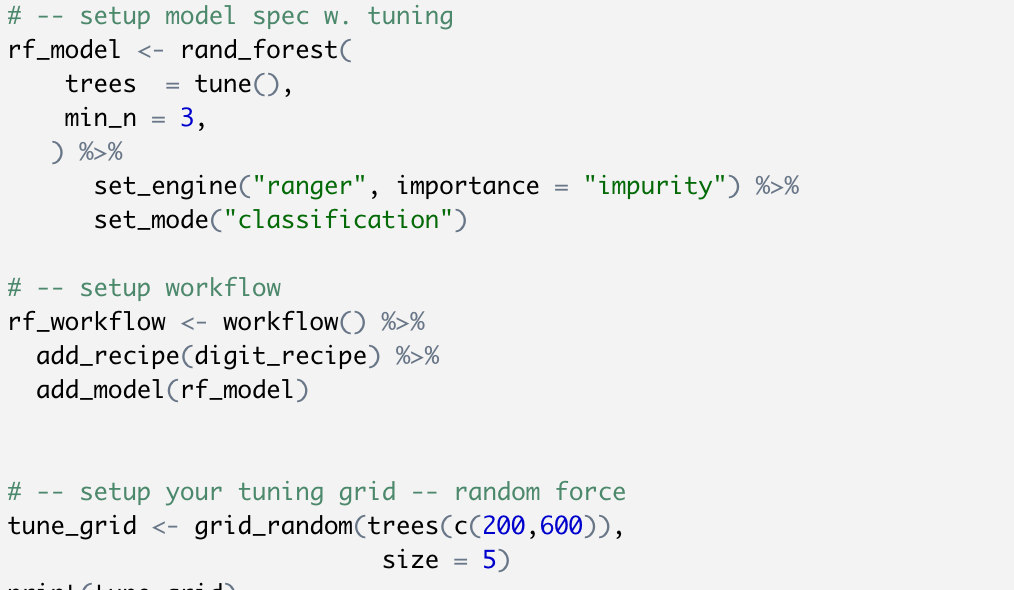


The tune grid part is focused on hidden\_units and penalty, of which hidden\_units is constrained between 5 and 15.

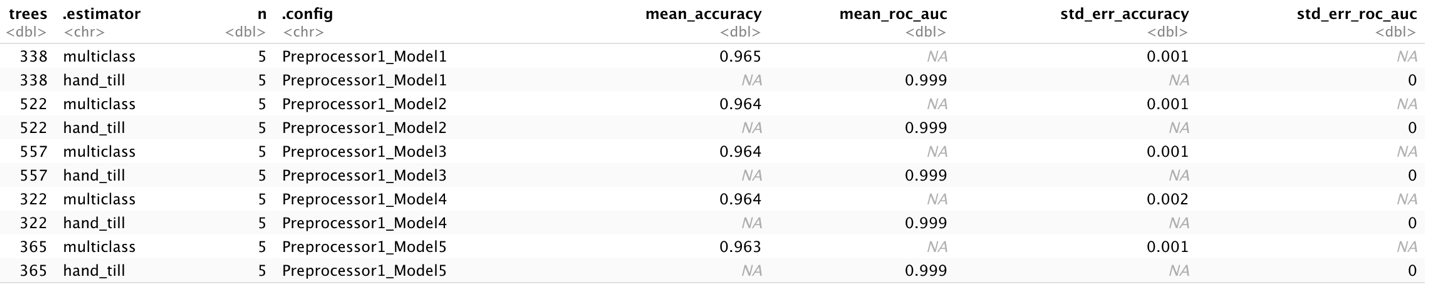


The best model in is the one with 10 hidden units and the penalty of 1.14298e-08, which gives out a mean accuracy of 0.902.

* **Random Forest**

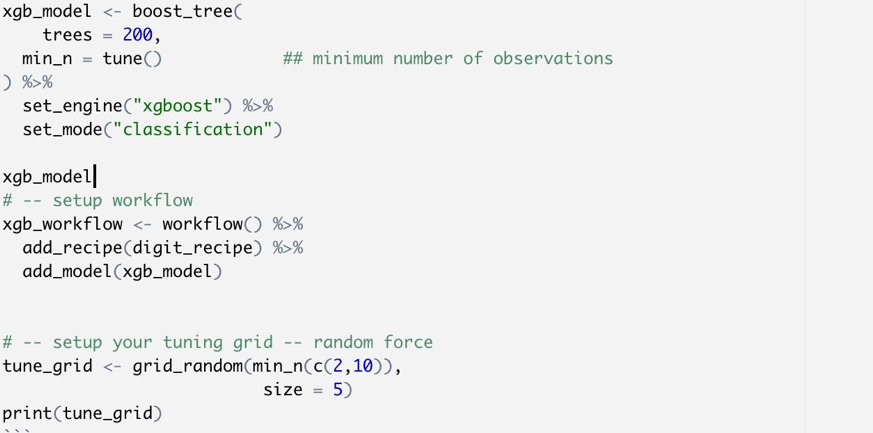


The tune grid part is focused on trees, which is constrained between 200 and 600.

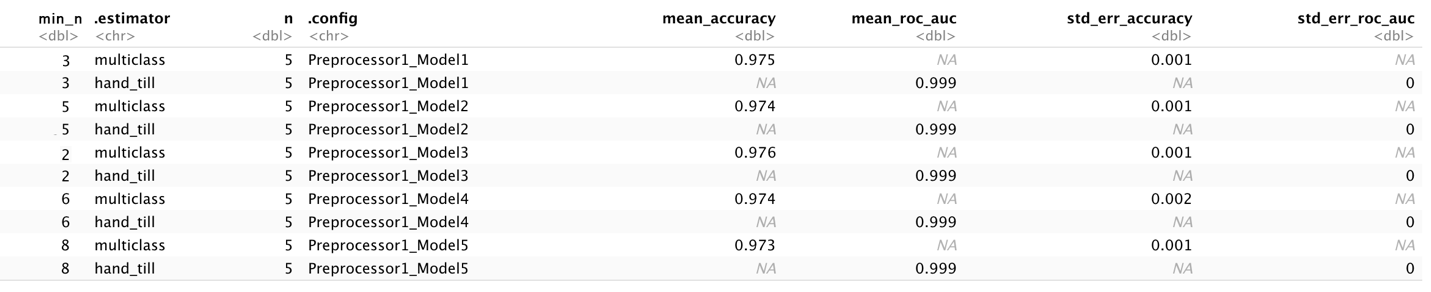


The best model in is the one with 338 trees, which gives out a mean accuracy of 0.965.

* **XGboost**



The tune grid part is focused on min\_n, which is constrained between 2 and 10.



The best model in is the one with min\_n of 2, which gives out a mean accuracy of 0.976.

# **Wrongly Predicted Numbers**

The examples in the 3 test sets of models are shown below. For example, in MLP model, the first number is actually 1 while it is predicted as 9 in the model.

|  |  |
| --- | --- |
| MLP | Random Forest |
|  |  |
| XGBoost |  |
|  |  |