

SHAP 中使用 AdaBoost

Guangyao Zhao

2023-01-07

Contents

SHAP 是一个非常好用的可解释性机器学习方法，但是原生并未对 AdaBoost 进行支持。此时需要自己在 SHAP 源代码中添加一些特定代码即可：

添加地址：`/Users/wenv/anaconda3/lib/python3.9/site-packages/shap/explainers/_tree.py`，大概添加在 1300 行左右，此处请改成自己的地址。

```
1 #TODO: 自己添加的对 AdaBoost 的支持
2 elif safe_isinstance(
3     model,
4     ("sklearn.ensemble.AdaBoostClassifier",
5      "sklearn.ensemble._weighted_boosting.AdaBoostClassifier",
6      "imblearn.ensemble.RUSBoostClassifier",
7      "imblearn.ensemble._weight_boosting.RUSBoostClassifier")):
8     assert hasattr(
9         model, "estimators_"
10    ), "Model has no `estimators_`! Have you called `model.fit`?"
11     self.internal_dtype = model.estimators_[0].tree_.value.dtype.type
12     self.input_dtype = np.float32
13     self.trees = [
14         SingleTree(e.tree_,
15                   normalize=True,
16                   scaling=weight,
17                   data=data,
18                   data_missing=data_missing) for e, weight in zip(
```

```

19         model.estimateds_, model.estimated_weights_ /
20         sum(model.estimated_weights_))
21     ]
22     self.objective = objective_name_map.get(
23         model.base_estimator_.criterion, None
24     ) #This line is done to get the decision criteria, for example gini.
25     self.tree_output = "probability"
26 elif safe_isinstance(
27     model,
28     ("sklearn.ensemble._weighted_boosting.AdaBoostRegressor",
29     "sklearn.ensemble.AdaBoostRegressor")):
30     assert hasattr(
31         model, "estimateds_"
32     ), "Model has no `estimateds_`! Have you called `model.fit`?"
33     self.internal_dtype = model.estimateds_[0].tree_.value.dtype.type
34     self.input_dtype = np.float32
35     self.trees = [
36         SingleTree(e.tree_,
37                    scaling=weight,
38                    data=data,
39                    data_missing=data_missing) for e, weight in zip(
40                    model.estimateds_, model.estimated_weights_ /
41                    sum(model.estimated_weights_))
42     ]
43     self.objective = objective_name_map.get(
44         model.base_estimator_.criterion, None
45     ) #This line is done to get the decision criteria, for example gini.
46     self.tree_output = "raw_value"

```