Train In Data

SFS

MLXtend - Sklearn

SFS – open source libraries

SFS – open source libraries


```python
from sklearn.feature_selection import SequentialFeatureSelector as SFS

sfs = SFS(
    estimator=RandomForestClassifier(
        n_estimators=10, n_jobs=4, random_state=0),
    n_features_to_select=10,  # the number of features to retain
    tol=None,  # the maximum increase or decrease in the performance metric
    direction='forward',  # the direction of the selection procedure
    scoring='roc_auc',  # the metric to evaluate
    cv=2,  # the cross-validation fold
    n_jobs=4,  # for parallelization
)

sfs = sfs.fit(X_train, y_train)
```


```python
from mlxtend.feature_selection import SequentialFeatureSelector as SFS

sfs1 = SFS(knn,
           k_features=3,
           forward=True,
           floating=False,
           verbose=2,
           scoring='accuracy',
           cv=0)

sfs1 = sfs1.fit(X, y)
```

Which library?

Sklearn

- Slighlty faster.
- Can stop based on model performance.
- Parameters we already know.

Train In Data

THANK YOU

www.trainindata.com