

모듈 기능 정의

```
# 모듈이 사용할 패키지
from sklearn.model_selection import cross_val_score, cross_validate
from sklearn.model_selection import GridSearchCV
from pandas import DataFrame
```

```
def singleML(modelName, x, y=None, cv=5, **kargs):
    # 모델 생성
    model = modelName(**kargs)
    # 교차 검증
    score = cross_val_score(model, x, y, cv=cv).mean()
    # 결과 데이터 프레임
    df = DataFrame(cross_validate(model, x, y, cv=cv))
    return [model, score, df]
```

```
def gridML(modelName, x, y=None, params={}, cv=5, **kargs):
    model = modelName(**kargs)

    grid = GridSearchCV(model, param_grid=params, cv=cv)

    try:
        grid.fit(x, y)
    except:
        grid.fit(x)

    result_df = DataFrame(grid.cv_results_['params'])
    result_df['mean_test_score'] = grid.cv_results_['mean_test_score']
    result_df.sort_values(by='mean_test_score', ascending=False)

    return [grid.best_estimator_, grid.best_params_, result_df]
```

모듈 테스트

```
from pandas import read_excel
from sklearn.linear_model import LinearRegression
from sklearn.svm import SVC
```

```
origin = read_excel('https://data.hossam.kr/G02/breast_cancer.xlsx')
x = origin.drop('target', axis=1)
y = origin['target']
x.shape, y.shape
```

```
((569, 30), (569,))
```

```
model, score, df = singleML(LinearRegression, x, y)
model
```

▼ LinearRegression

LinearRegression()

score

0.7046861734644332

df

	fit_time	score_time	test_score
0	0.006001	0.002000	0.623595
1	0.003997	0.001000	0.698961
2	0.003999	0.002000	0.755933
3	0.004002	0.000998	0.773021
4	0.003000	0.002001	0.671920

```
model, score, df = singleML(SVC, x, y, kernel='linear', C=0.1, random_state=777)
model
```

▼ SVC

SVC(C=0.1, kernel='linear', random_state=777)

모듈 테스트 2

```
model, best, score = gridML(LinearRegression, x, y)
model
```

▼ LinearRegression

LinearRegression()

score

	mean_test_score
0	0.704686

best

```
{}
```

```
params = {  
    'C': [0.001, 0.01],  
    'kernel': ['linear', 'rbf'],  
}
```

```
model, best, score = gridML(SVC, x, y, params=params, random_state=777)  
model
```

▼ SVC

SVC(C=0.01, kernel='linear', random_state=777)

score

	C	kernel	mean_test_score
0	0.001	linear	0.940258
1	0.001	rbf	0.627418
2	0.010	linear	0.947291
3	0.010	rbf	0.789194

best

```
{'C': 0.01, 'kernel': 'linear'}
```