

```
In [1]: #fixing OPEN_BLAS max pthreads reached issue
import os
os.environ['OPENBLAS_NUM_THREADS'] = '1'
os.environ['GOTO_NUM_THREADS'] = '1'
os.environ['OMP_NUM_THREADS'] = '1'

#imports
import pandas as pd
import numpy as np
import autosklearn.regression as autocl
from sklearn.model_selection import train_test_split
import sklearn.metrics

import sklearn.datasets
import matplotlib.pyplot as plt
from scipy import stats
import xgboost as xgb
from autosklearn.metrics import balanced_accuracy, precision, recall, f1
from sklearn import (manifold, datasets, preprocessing, model_selection, decomposition)
from sklearn.preprocessing import StandardScaler
import sklearn.preprocessing

CONSTRAINT = 0.3
```

```
In [2]: # select user input
# user_input = int(input("Which user's result do you want to look at?: "))
user_input = 0

# read the dataframe
data = pd.read_csv('data.csv')
df = data[data['user_number'] == user_input]

# drop all Null data (filtering null values)
df.dropna()

# gather the data
X = df.iloc[:, :-2].values
y = df["symptom_value"].values

n_samples = X.shape[0]
n_train = int(n_samples * 0.8)
indices = np.arange(n_samples)
np.random.shuffle(indices)
```

```
In [3]: def preprocessing(X):
    from sklearn.preprocessing import StandardScaler
    import sklearn.preprocessing
    std_scaler = StandardScaler()
    X = std_scaler.fit_transform(X)
    X = sklearn.preprocessing.normalize(X, norm='l2')
    return X
X = preprocessing(X)
```

```
In [4]: # split the dataset into training and test data
train_indices = indices[:n_train]
test_indices = indices[n_train:]

X_train = X[train_indices]
y_train = y[train_indices]

X_test = X[test_indices]
y_test = y[test_indices]
```

```
In [5]: # train the module
automl = automl.AutoSklearnRegressor(
    time_left_for_this_task = 120,
    per_run_time_limit = 30,
    disable_evaluator_output = False,
    resampling_strategy = "cv",
    resampling_strategy_arguments = {"folds": 10},
)
```

```
In [6]: automl.fit(X_train, y_train)
print("Before re-fit")
predictions = automl.predict(X_test)
print("Test R2 score:", sklearn.metrics.r2_score(y_test, predictions))
print("#" * 80)
print("After re-fit")
automl.refit(X_train.copy(), y_train.copy())
# pickle.dump(automl, open('model.sav', 'wb'))
predictions = automl.predict(X_test)
print("Test R2 score:", sklearn.metrics.r2_score(y_test, predictions))
print("#" * 80)

# automl.fit(X_train, y_train)
```

```

[WARNING] [2023-04-16 18:39:42,839:Client-EnsembleBuilder] No runs were available to
build an ensemble from
[WARNING] [2023-04-16 18:39:44,650:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:47,509:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:47,769:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:50,060:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:51,699:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:56,475:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:39:59,311:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:40:01,953:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:40:02,218:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:40:02,532:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
[WARNING] [2023-04-16 18:40:02,828:Client-EnsembleBuilder] No models better than ran
dom - using Dummy losses!
    Models besides current dummy model: 0
    Dummy models: 1
Before re-fit
Test R2 score: 0.733966672783318
#####
After re-fit
Test R2 score: 0.7398337410800013
#####

```

```

In [7]: # print the results
        print(automl.leaderboard())

```

```
# print the accuracy
predictions = automl.predict(X_test)

print("Test R2 score:", sklearn.metrics.r2_score(y_test, predictions))
```

	rank	ensemble_weight	type	cost	duration
model_id					
14	1	0.46	liblinear_svr	6.738509	1.562520
49	2	0.04	k_nearest_neighbors	9.433448	1.999790
22	3	0.50	gaussian_process	17.008425	1.969288
Test R2 score: 0.7398337410800013					