

Installation

sklearn: <https://scikit-learn.org/stable/install.html>

Dataset

Files

train_1.pkl, train_2.pkl, train_3.pkl
test_1.pkl, test_2.pkl, test_3.pkl

Description

number of train samples = 614
number of test samples = 154

The last column is the label/target, others are the attributes.

Usage

This is the sample python code to read .pkl files.

```
import pickle
pkl_file = open('train_1.pkl', 'rb')
data = pickle.load(pkl_file)
pkl_file.close()
print(data.shape) # 614 samples, 8 attributes, 1 label
print(data[:3]) # see the first 3 samples
```

Model

Logistic Regression

Manual: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

Example: https://scikit-learn.org/stable/auto_examples/linear_model/plot_iris_logistic.html#sphx-glr-auto-examples-linear-model-plot-iris-logistic-py