

wine_classification

October 31, 2020

```
[2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.neural_network import MLPClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import \
    classification_report, confusion_matrix, accuracy_score
```

Load Data

```
[4]: from sklearn.datasets import load_wine
data = load_wine(return_X_y=False)
print(data.DESCR)
```

```
.. _wine_dataset:
```

Wine recognition dataset

****Data Set Characteristics:****

```
:Number of Instances: 178 (50 in each of three classes)
:Number of Attributes: 13 numeric, predictive attributes and the class
:Attribute Information:
    - Alcohol
    - Malic acid
    - Ash
    - Alcalinity of ash
    - Magnesium
    - Total phenols
    - Flavanoids
    - Nonflavanoid phenols
    - Proanthocyanins
    - Color intensity
    - Hue
    - OD280/OD315 of diluted wines
    - Proline
```

```
- class:
  - class_0
  - class_1
  - class_2
```

:Summary Statistics:

```
=====
              Min    Max    Mean    SD
=====
Alcohol:      11.0   14.8    13.0    0.8
Malic Acid:   0.74   5.80    2.34    1.12
Ash:          1.36   3.23    2.36    0.27
Alcalinity of Ash: 10.6  30.0    19.5    3.3
Magnesium:    70.0  162.0   99.7    14.3
Total Phenols: 0.98   3.88    2.29    0.63
Flavanoids:   0.34   5.08    2.03    1.00
Nonflavanoid Phenols: 0.13  0.66    0.36    0.12
Proanthocyanins: 0.41  3.58    1.59    0.57
Colour Intensity: 1.3   13.0    5.1     2.3
Hue:          0.48   1.71    0.96    0.23
OD280/OD315 of diluted wines: 1.27  4.00    2.61    0.71
Proline:      278   1680    746     315
=====
```

:Missing Attribute Values: None

:Class Distribution: class_0 (59), class_1 (71), class_2 (48)

:Creator: R.A. Fisher

:Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)

:Date: July, 1988

This is a copy of UCI ML Wine recognition datasets.

<https://archive.ics.uci.edu/ml/machine-learning-databases/wine/wine.data>

The data is the results of a chemical analysis of wines grown in the same region in Italy by three different cultivators. There are thirteen different measurements taken for different constituents found in the three types of wine.

Original Owners:

Forina, M. et al, PARVUS -

An Extendible Package for Data Exploration, Classification and Correlation.

Institute of Pharmaceutical and Food Analysis and Technologies,

Via Brigata Salerno, 16147 Genoa, Italy.

Citation:

Lichman, M. (2013). UCI Machine Learning Repository
[<https://archive.ics.uci.edu/ml>]. Irvine, CA: University of California,
School of Information and Computer Science.

.. topic:: References

(1) S. Aeberhard, D. Coomans and O. de Vel,
Comparison of Classifiers in High Dimensional Settings,
Tech. Rep. no. 92-02, (1992), Dept. of Computer Science and Dept. of
Mathematics and Statistics, James Cook University of North Queensland.
(Also submitted to Technometrics).

The data was used with many others for comparing various
classifiers. The classes are separable, though only RDA
has achieved 100% correct classification.
(RDA : 100%, QDA 99.4%, LDA 98.9%, 1NN 96.1% (z-transformed data))
(All results using the leave-one-out technique)

(2) S. Aeberhard, D. Coomans and O. de Vel,
"THE CLASSIFICATION PERFORMANCE OF RDA"
Tech. Rep. no. 92-01, (1992), Dept. of Computer Science and Dept. of
Mathematics and Statistics, James Cook University of North Queensland.
(Also submitted to Journal of Chemometrics).

```
[5]: data = load_wine(return_X_y=True)
X = data[0]
t = data[1]
X=pd.DataFrame(X)
t=pd.DataFrame(t)

#print(X)
#print(t)
```

Setup division of data

```
[225]: #X_train, X_test, t_train, t_test = train_test_split(X, t, test_size = 0.3)
X_train, X_validate, X_test = np.split(X.sample(frac=1, random_state=42), [int(.
    ↳6*len(X)), int(.8*len(X))])
t_train, t_validate, t_test = np.split(t.sample(frac=1, random_state=42), [int(.
    ↳6*len(t)), int(.8*len(t))])
X_validate.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 36 entries, 139 to 155
Data columns (total 13 columns):
```

```

0      36 non-null float64
1      36 non-null float64
2      36 non-null float64
3      36 non-null float64
4      36 non-null float64
5      36 non-null float64
6      36 non-null float64
7      36 non-null float64
8      36 non-null float64
9      36 non-null float64
10     36 non-null float64
11     36 non-null float64
12     36 non-null float64
dtypes: float64(13)
memory usage: 3.9 KB

```

```

[226]: from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
scaler.fit(X_train)
X_train = scaler.transform(X_train)
X_validate= scaler.transform(X_validate)
X_test = scaler.transform(X_test)

```

```

[227]: mlp = MLPClassifier(activation='logistic', alpha=1e-5, solver = 'lbfgs',
    →hidden_layer_sizes=(30,30,30))

```

train the model

```

[228]: mlp.fit(X_train, t_train)
predictions = mlp.predict(X_validate)
print(classification_report(t_validate,predictions))
print(accuracy_score(t_validate, predictions))

```

	precision	recall	f1-score	support
0	1.00	1.00	1.00	12
1	1.00	0.93	0.96	14
2	0.91	1.00	0.95	10
accuracy			0.97	36
macro avg	0.97	0.98	0.97	36
weighted avg	0.97	0.97	0.97	36

```
0.9722222222222222
```

```

//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was

```

expected. Please change the shape of y to (n_samples,), for example using ravel().

```
y = column_or_1d(y, warn=True)
```

test the model

```
[229]: y=mlp.predict(X_test)
print(classification_report(t_test,y))
print(accuracy_score(t_test, y))
```

	precision	recall	f1-score	support
0	0.92	1.00	0.96	11
1	0.93	0.93	0.93	14
2	1.00	0.91	0.95	11
accuracy			0.94	36
macro avg	0.95	0.95	0.95	36
weighted avg	0.95	0.94	0.94	36

0.9444444444444444

Q1: tune the learning rate and batch size

```
[230]: # learning rate
learning_rate = np.arange(0.01, 0.3, 0.01)
accuracy = []
for i in learning_rate:
    clf = MLPClassifier(random_state=1, learning_rate_init=i).fit(X_train,
→t_train)
    accuracy.append(clf.score(X_test, t_test))
plt.title('Accuracy with Learning Rate')
plt.plot(learning_rate, accuracy)
print('The optimal learning rate is',learning_rate[accuracy.
→index(max(accuracy))], 'with accuracy = ',max(accuracy))
```

```
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
y = column_or_1d(y, warn=True)
```

```

//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)

```

```

//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)

```

```

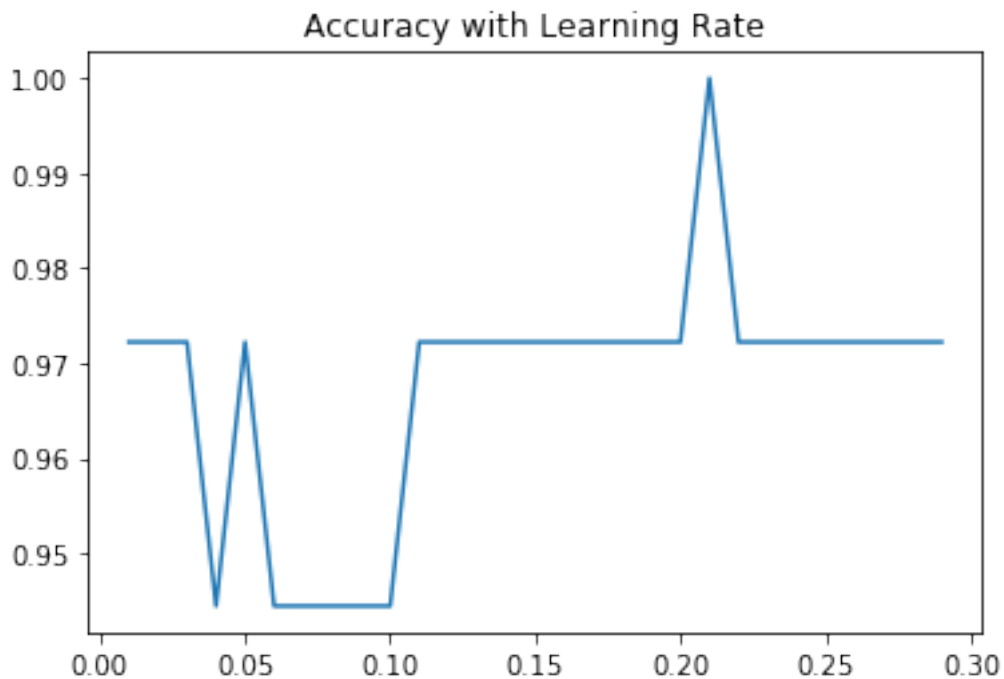
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)

```



```
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
    y = column_or_1d(y, warn=True)
```

The optimal learning rate is 0.21000000000000002 with accuracy = 1.0



```
[231]: #Batch Size#
Batch_Size = np.arange(10, 200, 10)
accuracy = []
for j in Batch_Size:
```



```

expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:350: UserWarning: Got
`batch_size` less than 1 or larger than sample size. It is going to be clipped
  warnings.warn("Got `batch_size` less than 1 or larger than "
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:350: UserWarning: Got
`batch_size` less than 1 or larger than sample size. It is going to be clipped
  warnings.warn("Got `batch_size` less than 1 or larger than "
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-

```

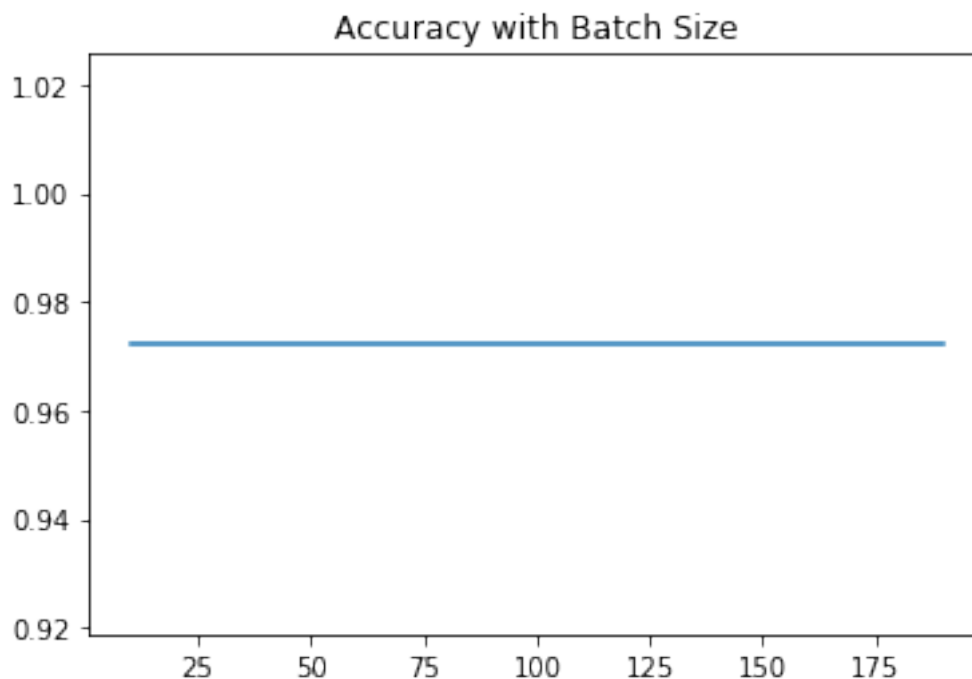


```

y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:350: UserWarning: Got
`batch_size` less than 1 or larger than sample size. It is going to be clipped
warnings.warn("Got `batch_size` less than 1 or larger than ")
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:350: UserWarning: Got
`batch_size` less than 1 or larger than sample size. It is going to be clipped
warnings.warn("Got `batch_size` less than 1 or larger than ")

```

The optimal Batch Size is 10 with accuracy = 0.9722222222222222



Q2. use 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of the total training dataset

```

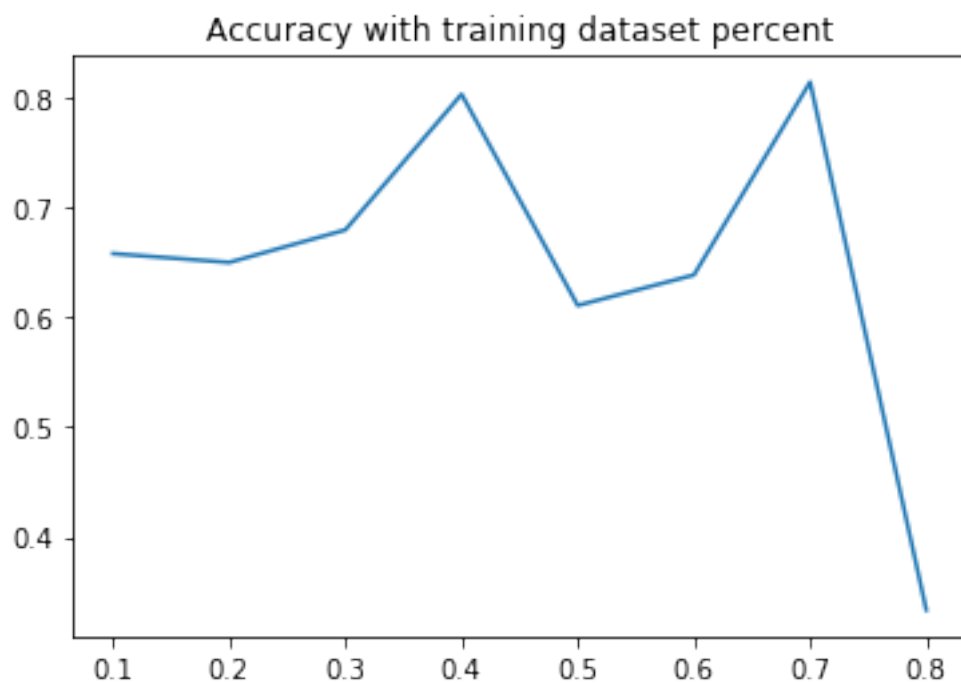
[289]: td_percent=np.arange(0.2,1,0.1)
accuracy = []
for a in td_percent:
    X_train, X_test, t_train, t_test = train_test_split(X, t, test_size = a)

```


DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().

```
y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
y = column_or_1d(y, warn=True)
```

The optimal training dataset percent is 0.7 with accuracy = 0.8148148148148148



Q3. use 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100% of the features

```
[299]: import random
frange = [0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0]
featuresnumber = [3,4,5,7,8,9,10,12,13]
accuracy=[]
for b in featuresnumber:
    data = load_wine(return_X_y=True)
    X = data[0]
    t = data[1]
    X=pd.DataFrame(X)
    t=pd.DataFrame(t)
```



```

ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)
//anaconda3/lib/python3.7/site-
packages/sklearn/neural_network/multilayer_perceptron.py:921:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples, ), for example using
ravel().
  y = column_or_1d(y, warn=True)

```

The optimal features percent is 0.7 with accuracy = 0.9444444444444444

