In [1]:

```
import numpy as np
from numpy import genfromtxt
import math as mt
import matplotlib.pyplot as plt
from scipy.spatial.distance import cdist
```

In [2]:

```
sytrain_data = np.genfromtxt('wine_train.csv',delimiter=',')
sytest_data = np.genfromtxt('wine_test.csv', delimiter=',')
```

In [3]:

```
sytrain_data
```

Out[3]:

```
array([[1.375e+01, 1.730e+00, 2.410e+00, ..., 2.900e+00, 1.320e+03,
        1.000e+00],
       [1.320e+01, 1.780e+00, 2.140e+00, ..., 3.400e+00, 1.050e+03,
        1.000e+00],
       [1.307e+01, 1.500e+00, 2.100e+00, ..., 2.690e+00, 1.020e+03,
        1.000e+00],
       ...,
       [1.253e+01, 5.510e+00, 2.640e+00, ..., 1.690e+00, 5.150e+02,
        3.000e+00],
       [1.288e+01, 2.990e+00, 2.400e+00, ..., 1.420e+00, 5.300e+02,
        3.000e+00],
       [1.373e+01, 4.360e+00, 2.260e+00, ..., 1.750e+00, 5.200e+02,
        3.000e+00]])
```

In [4]:

```
sytest_data
```

Out[4]:

```
array([[1.422e+01, 3.990e+00, 2.510e+00, ..., 3.530e+00, 7.600e+02,
        1.000e+00],
       [1.200e+01, 9.200e-01, 2.000e+00, ..., 3.120e+00, 2.780e+02,
        2.000e+00],
       [1.367e+01, 1.250e+00, 1.920e+00, ..., 2.460e+00, 6.300e+02,
        2.000e+00],
       ...,
       [1.247e+01, 1.520e+00, 2.200e+00, ..., 2.630e+00, 9.370e+02,
        2.000e+00],
       [1.362e+01, 4.950e+00, 2.350e+00, ..., 2.050e+00, 5.500e+02,
        3.000e+00],
       [1.245e+01, 3.030e+00, 2.640e+00, ..., 1.730e+00, 8.800e+02,
        3.000e+00]])
```

In [22]:

```
means = [];
variance = [];
sytrain_normalised = np.zeros((90,14));
for i in range(0,14):
    means.append(np.mean(sytrain_data[i]));
    variance.append(np.var(sytrain_data[i]));
for j in range(0,14):
    for i in range(0,89):
        sytrain_normalised[i][j] = ( sytrain_data[i][j] - means[j] )/(variance[j]);
```

In [25]:

```
means
```

Out[25]:

```
[104.35714285714286,
 85.36428571428571,
 83.245,
 85.53285714285714,
 89.84571428571428,
 100.96000000000001,
 74.265,
 66.22,
 110.33357142857143,
 101.87714285714286,
 85.20142857142856,
 102.67571428571429,
 89.20785714285715,
 115.44500000000001]
```

In [27]:

```
#built-in function
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
print(scaler.fit(sytrain_data))
StandardScaler(copy=True, with_mean=False, with_std=True)
print(scaler.mean_)
print(scaler.transform(sytrain_data))


```

```
StandardScaler(copy=True, with_mean=True, with_std=True)
[1.29653933e+01 2.27000000e+00 2.37629213e+00 1.96494382e+01
 9.89101124e+01 2.27235955e+00 2.02943820e+00 3.60674157e-01
 1.57617978e+00 5.09123596e+00 9.53213483e-01 2.56033708e+00
 7.29707865e+02 1.93258427e+00]
[[ 0.95735106 -0.48954541  0.12347021 ...  0.4694969   1.92139113
  -1.2017632 ]
 [ 0.28625935 -0.44421713 -0.86552621 ...  1.1606187   1.0425456
  -1.2017632 ]
 [ 0.12763767 -0.69805549 -1.0120442  ...  0.17922574  0.94489609
  -1.2017632 ]
 ...
 [-0.53125238  2.93727245  0.96594865 ... -1.20301785 -0.69887055
   1.3755121 ]
 [-0.10419401  0.65272721  0.08684072 ... -1.57622362 -0.65004579
   1.3755121 ]
 [ 0.93294772  1.89472205 -0.42597224 ... -1.12008324 -0.68259563
   1.3755121 ]]
```

In [ ]:

```

```