letter_recognition_MultiLable_Classification

April 24, 2018

```
In [155]: # multilable classification demo
          import pandas as pd
          import numpy as np
          from sklearn.multiclass import OneVsRestClassifier
          from sklearn import svm ,naive_bayes
          from sklearn.model_selection import train_test_split
          # from sklearn.tree import DecisionTreeClassifier
          # from sklearn.neighbors import KNeighborsClassifier as KNN
```

```
In [94]: indexes = ['letter','x-box','y-box','width','height','onpix','x-bar','y-bar','x2bar','y
```

```
In [95]: df = pd.read_csv('https://archive.ics.uci.edu/ml/machine-learning-databases/letter-reco
```

```
In [96]: # df = dff
```

```
In [97]: dff = df
```

```
In [128]: df.head()
```

```
Out[128]:   letter  x-box  y-box  width  height  onpix  x-bar  y-bar  x2bar  y2bar  \
          0      T      2      8      3       5      1      8     13      0      6
          1      I      5     12      3       7      2     10      5      5      4
          2      D      4     11      6       8      6     10      6      2      6
          3      N      7     11      6       6      3      5      9      4      6
          4      G      2      1      3       1      1      8      6      6      6

             xybar  x2ybr  xy2br  x-ege  xegvy  y-ege  yegvx
          0      6     10      8      0      8      0      8
          1     13      3      9      2      8      4     10
          2     10      3      7      3      7      3      9
          3      4      4     10      6     10      2      8
          4      6      5      9      1      7      5     10
```

```
In [129]: y = df['letter']
          x = df
          x = x.drop('letter',axis=1)
          y.head()
```

```
Out[129]: 0    T
          1    I
          2    D
          3    N
          4    G
          Name: letter, dtype: object

In [100]: y = pd.get_dummies(y)

In [130]: y.head()

Out[130]: 0    T
          1    I
          2    D
          3    N
          4    G
          Name: letter, dtype: object

In [132]: train_x, test_x, train_y, test_y = train_test_split(x,y)

In [135]: test_y.head()

Out[135]: 2835     Z
          17868    F
          10780    M
          13502    L
          11657    O
          Name: letter, dtype: object

In [153]: # model = OneVsRestClassifier(KNN()) #0.9436
          model = OneVsRestClassifier(svm.SVC()) # 0.9666
          # model = OneVsRestClassifier(DecisionTreeClassifier()) # 0.8308
          # model = OneVsRestClassifier(naive_bayes.MultinomialNB()) # 0.6364 GNB, 0.5382 MNB, (
          model.fit(train_x,train_y)

Out[153]: OneVsRestClassifier(estimator=SVC(C=1.0, cache_size=200, class_weight=None, coef0=0.0,
              decision_function_shape='ovr', degree=3, gamma='auto', kernel='rbf',
             max_iter=-1, probability=False, random_state=None, shrinking=True,
             tol=0.001, verbose=False),
                    n_jobs=1)

In [154]: model.score(test_x,test_y)

Out[154]: 0.9666
```