

BÀI TẬP 01

TRƯỜNG MINH ÁNH - 1112010

Contents

Câu 1: 2

Câu 2: 2

Câu 3: 2

Câu 4: 3

Câu 5: 4

Câu 6: 5

Câu 7: 6

Câu 8: 6

Câu 9: 7

Câu 10: 7

Câu 1:

Trả lời:

[d] A quadratic programming problem with $d + 1$ variablesGiải thích: ta có w là ma trận $d \times 1$ **Câu 2:**

Trả lời: [a] 0 versus all

Hướng dẫn: Gọi hàm [err min max] = cau23(features_in)

Với: err – ma trận chứa Ein tương ứng từ 0->9

min – giá trị có Ein thấp nhất

max – giá trị có Ein cao nhất

features_in – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

```
>> [err min max] = cau23(features_in)
Accuracy = 83.6236% (6097/7291) (classification)
Accuracy = 98.4227% (7176/7291) (classification)
Accuracy = 89.9739% (6560/7291) (classification)
Accuracy = 90.9752% (6633/7291) (classification)
Accuracy = 91.0575% (6639/7291) (classification)
Accuracy = 92.3742% (6735/7291) (classification)
Accuracy = 90.8929% (6627/7291) (classification)
Accuracy = 91.1535% (6646/7291) (classification)
Accuracy = 92.5662% (6749/7291) (classification)
Accuracy = 91.1672% (6647/7291) (classification)

err =

    0.1638    0.0158    0.1003    0.0902    0.0894    0.0763    0.0911    0.0885    0.0743    0.0883

min =

    1

max =

    0
```

*Figure 1 - Câu 2***Câu 3:**

Trả lời: [a] 1 versus all

Hướng dẫn: Gọi hàm [err min max] = cau23(features_in)

Với: err – ma trận chứa Ein tương ứng từ 0->9

min – giá trị có Ein thấp nhất

max – giá trị có Ein cao nhất

features_in – ma trận chứa dữ liệu train(với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

```
>> [err min max] = cau23(features_in)
Accuracy = 83.6236% (6097/7291) (classification)
Accuracy = 98.4227% (7176/7291) (classification)
Accuracy = 89.9739% (6560/7291) (classification)
Accuracy = 90.9752% (6633/7291) (classification)
Accuracy = 91.0575% (6639/7291) (classification)
Accuracy = 92.3742% (6735/7291) (classification)
Accuracy = 90.8929% (6627/7291) (classification)
Accuracy = 91.1535% (6646/7291) (classification)
Accuracy = 92.5662% (6749/7291) (classification)
Accuracy = 91.1672% (6647/7291) (classification)

err =

    0.1638    0.0158    0.1003    0.0902    0.0894    0.0763    0.0911    0.0885    0.0743    0.0883

min =

    1

max =

    0
```

Figure 2 - Câu 3

Câu 4:

Trả lời: [c] 1800

Hướng dẫn: Gọi hàm [model1 model2] = cau4(features_in)

Với: model1 – thông số mô hình svm của giá trị 0(highest Ein)
 model2 – thông số mô hình svm của giá trị 1(lowest Ein)

features_in – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

```
>> [model1 model2] = cau4(features_in)

model1 =

Parameters: [5x1 double]
nr_class: 2
totalSV: 2390
rho: -1.0618
Label: [2x1 double]
sv_indices: [2390x1 double]
ProbA: []
ProbB: []
nSV: [2x1 double]
sv_coef: [2390x1 double]
SVs: [2390x2 double]

model2 =

Parameters: [5x1 double]
nr_class: 2
totalSV: 536
rho: 1.5323
Label: [2x1 double]
sv_indices: [536x1 double]
ProbA: []
ProbB: []
nSV: [2x1 double]
sv_coef: [536x1 double]
SVs: [536x2 double]
```

Figure 3 - Câu 4

Câu 5:

Trả lời: [a]: The number of support vectors goes down when C goes up

[d]: Maximum C achieves the lowest E_{in}

Hướng dẫn: Gọi hàm **cau56(features_in, features_out)**

Với: features_in – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

features_out – ma trận chứa dữ liệu test (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

Giải thích: [a]: Số support vector giảm từ 152 \rightarrow 54 \rightarrow 28 \rightarrow 25 khi C tăng từ 0.001 \rightarrow 0.01 \rightarrow 0.1 \rightarrow 1

```
Parameters: [5x1 double]
nr_class: 2
totalSV: 152
Parameters: [5x1 double]
nr_class: 2
totalSV: 54
Parameters: [5x1 double]
nr_class: 2
totalSV: 28
Parameters: [5x1 double]
nr_class: 2
totalSV: 25
```

Figure 4 - Câu 5(Thông tin về số support vector)

[d]: Ein khi C = 1 có giá trị thấp nhất là 0.006(bằng với Ein khi C = 0.01 và C = 0.1)

Thông tin do lỗi Ein theo thu tu tăng dần của C(từ 0.0001 đến 1):
0.0224 0.0070 0.0045 0.0045 0.0045

Figure 5 - Câu 5(Giá trị Ein)

Câu 6:

Trả lời: [b] When $C = 0.001$, the number of support vectors is lower at $Q = 5$

Hướng dẫn: Gọi hàm **cau56(features_in, features_out)**

Với: features_in – ma trận chứa dữ liệu train(với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

features_out – ma trận chứa dữ liệu test(với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

Giải thích: Số support vector khi Q = 2 là 152 còn khi Q = 5 là 28 với C = 0.001

Parameters: [5x1 double]	Parameters: [5x1 double]
nr_class: 2	nr_class: 2
totalSV: 152	totalSV: 28

Figure 6 - Câu 6(Q = 2 bên trái và Q = 5 bên phải)

Câu 7:

Trả lời: [c] $C = 0.01$ is selected most often

Hướng dẫn sử dụng: Gọi hàm `cau78(features_in)` như hình minh họa

Với: `features_in` – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x_1 và x_2)

```
>> cau78(features_in)
Ket qua thi nghiem voi Q = 2
So lan duoc chon cua cac C(tu 0.0001->1)
0
0
94
1
5

Gia tri do loi trung binh thap nhat:
0.0045
```

Figure 7 - Câu 7

Câu 8:

Trả lời: [c] 0.005

Hướng dẫn sử dụng: Gọi hàm `cau78(features_in)` như hình minh họa

Với: `features_in` – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x_1 và x_2)

```
>> cau78(features_in)
Ket qua thi nghiem voi Q = 2
So lan duoc chon cua cac C(tu 0.0001->1)
0
0
94
1
5

Gia tri do loi trung binh thap nhat:
0.0045
```

Figure 8 - Câu 8

Câu 9:

Trả lời: [e] 10^6

Hướng dẫn: Gọi hàm **cau910(features_in, features_out)**

Với: features_in – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

features_out – ma trận chứa dữ liệu test (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

```
>> cau910(features_in, features_out)
Ket qua thi nghiem voi Q = 2
Thong tin do loi Ein theo thu tu tang dan cua C (tu 0.0001 den 1):
    0.3562    0.0045    0.0032    0.0026    0.0006

Thong tin do loi Eout theo thu tu tang dan cua C (tu 0.0001 den 1):
    0.1025    0.0058    0.0051    0.0064    0.0064
```

Figure 9 - Câu 9

Câu 10:

Trả lời: [c] 100

Hướng dẫn: Gọi hàm **cau910(features_in, features_out)**

Với: features_in – ma trận chứa dữ liệu train (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

features_out – ma trận chứa dữ liệu test (với cột 1 là phân lớp dữ liệu, hai cột sau lần lượt là x1 và x2)

```
>> cau910(features_in, features_out)
Ket qua thi nghiem voi Q = 2
Thong tin do loi Ein theo thu tu tang dan cua C (tu 0.0001 den 1):
    0.3562    0.0045    0.0032    0.0026    0.0006

Thong tin do loi Eout theo thu tu tang dan cua C (tu 0.0001 den 1):
    0.1025    0.0058    0.0051    0.0064    0.0064
```

Figure 10 - Câu 10