

# 機器學習 HW4

系所:電控碩一

學號:313512022

姓名:呂靖樑

## 實驗結果:

C value Sigma (log scale)	1	5	10	50	100	500	1000
-100	54	54	54	54	54	54	54
-95	66.67	66.67	66.67	66.67	66.67	66.67	66.67
-90	72.67	72.67	72.67	72.67	72.67	72.67	72.67
-85	76.67	76.67	76.67	76.67	76.67	76.67	76.67
-80	78	78	78	78	78	78	78
-75	78.67	78.67	78.67	78.67	78.67	78.67	78.67
-70	51.33	51.33	51.33	51.33	51.33	51.33	51.33
-65	54	58.67	58.67	58.67	58.67	58.67	58.67
-60	62.67	64.67	64.67	64.67	64.67	64.67	64.67
-55	63.33	63.33	63.33	63.33	63.33	63.33	63.33
-50	66	66	66	66	66	66	66
-45	74	74.67	74.67	74.67	74.67	74.67	74.67
-40	84.67	85.33	85.33	85.33	85.33	85.33	85.33
-35	90.67	90	90	90	90	90	90
-30	94	92.67	92.67	92.67	92.67	92.67	92.67
-25	95.33	95.33	95.33	95.33	95.33	95.33	95.33
-20	95.33	94.67	95.33	95.33	95.33	95.33	95.33
-15	95.33	95.33	95.33	96.67	96.67	96.67	96.67
-10	94.67	95.33	95.33	96	96.67	96.67	96.67
-5	96	95.33	95.33	94.67	94.67	95.33	95.33
0	95.33	95.33	94.67	95.33	95.33	96	96
5	96	95.33	96.67	96	96.67	95.33	96
10	95.33	96	96	94.67	96	95.33	94
15	95.33	96	95.33	94	95.33	95.33	94.67
20	95.33	96	96	97.33	94.67	95.33	95.33
25	94.67	95.33	96	98	97.33	95.33	95.33
30	94.67	96.67	95.33	96	98	96	95.33
35	79.33	95.33	98	96.67	98.67	96.67	96
40	62	94.67	96	96.67	96.67	98	96
45	91.33	94.67	95.33	96	96.67	98	97.33
50	66.67	93.33	94.67	95.33	97.33	98.67	98
55	33.33	90.67	94	96.67	96	96.67	98.67

60	33.33	90.67	92.67	95.33	96.67	96.67	98.67
65	33.33	50	78.67	94.67	94.67	96.67	96.67
70	33.33	66.67	93.33	94.67	94.67	95.33	96
75	33.33	33.33	50	92.67	95.33	98	97.33
80	33.33	33.33	33.33	92	93.33	96	96.67
85	33.33	33.33	33.33	92	90.67	95.33	95.33
90	33.33	33.33	33.33	50	90	94.67	95.33
95	33.33	33.33	33.33	79.33	90	94.67	95.33
100	33.33	33.33	33.33	33.33	66.67	92.67	94.67

## 討論：

1. 小 C 更容忍錯誤，模型邊界平滑但泛化能力較弱；大 C 強調準確分類訓練數據，可能 overfitting。適中的 C 能在邊界平滑與分類準確之間取得平衡，通常效果最佳。在實驗結果中，我認為 C 大於 10 表現較佳。
2. 當 sigma 過小，模型過於局部化，容易 overfitting；當 sigma 過大，模型趨於線性，無法捕捉非線性特徵。適中的 sigma 能平衡全局與局部信息，通常分類效果最佳。特徵數值範圍也與 sigma 有關，建議進行特徵標準化，確保不同特徵對距離計算的影響一致。在實驗結果中，我認為 sigma (log scale) 介於 -20~70 表現較佳。
3. 我認為不採用 two-fold cross validation 分類率會提高，因為模型直接在訓練數據上測試，其邊界已經針對這些數據調整過，分類率自然會較高，但是無法評估其泛化能力。