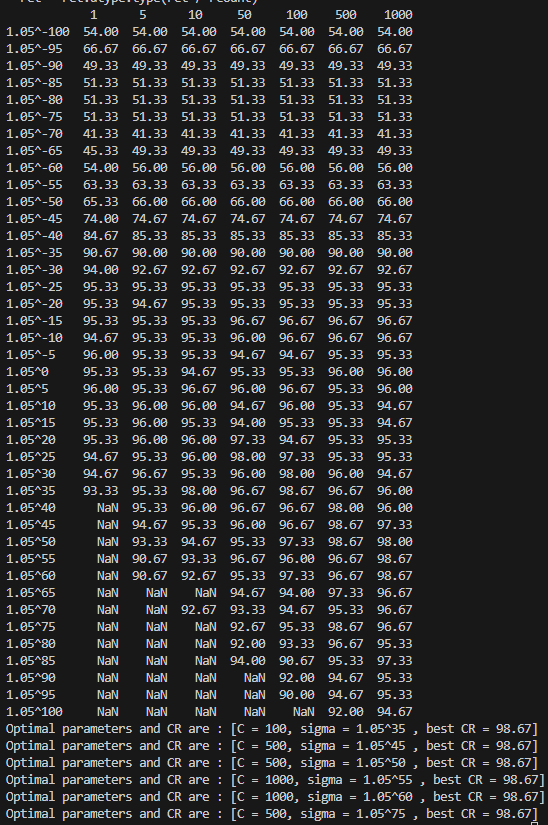
Machine Learning HW4

312512005 黃名諄

1. SVM grid search result:

在做grid search時，會有某些參數組合做one-against-one strategy時，3個模型中可能會有某幾個模型因沒有alpha在0和C之間，導致無法算出bias，模型無法預測，因此我的處理方式是若當組參數組合的3個model中只要有一個沒有辦法算出bias就放棄此組grid參數，令此組參數CR=nan來處理；且最後有多組參數組合都能達到最佳的CR，因此我都有印出來(row是sigma，column是C)

1. 結果討論:
   1. 請問在grid search的結果中，C的大小與分類率高低有何關係?

Ans : C的意義是懲罰分類錯誤，所以越大表示模型對於training data的分類錯誤越重視，也就是可能會過擬合training data；相反C越小泛化能力較大但也可能造成欠擬合發生。

從我做出來的結果來分析，可以發現其實C對分類率都是小幅度的影響，主要還是sigma影響較大，而在C由小到大排序下中後段的分類率較高，可驗證不能使C太大太小而過擬合或欠擬合，因此符合理想上C中後段會有個分類率最高的峰值。

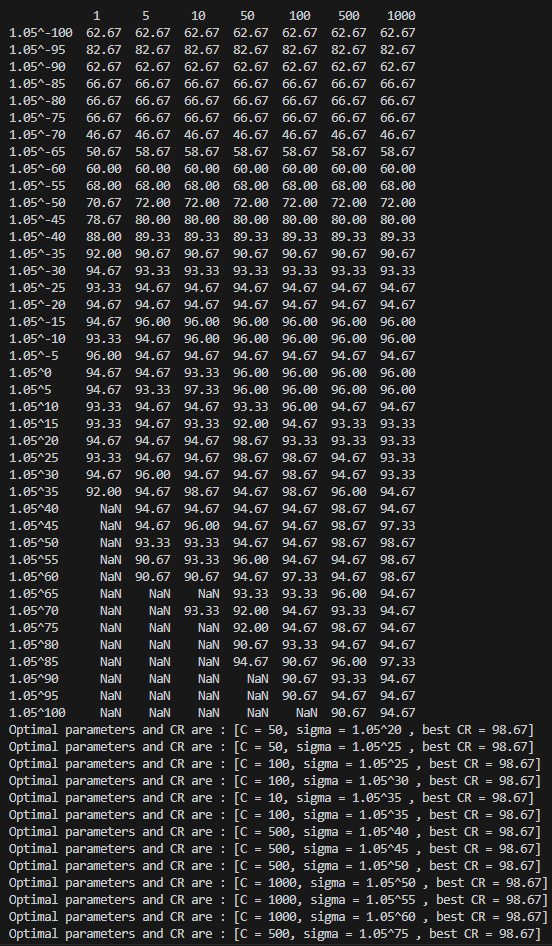
* 1. sigma的大小的改變與分類率是否有關係? 若有，請探討sigma的差異與特徵的數值有什麼關聯性?

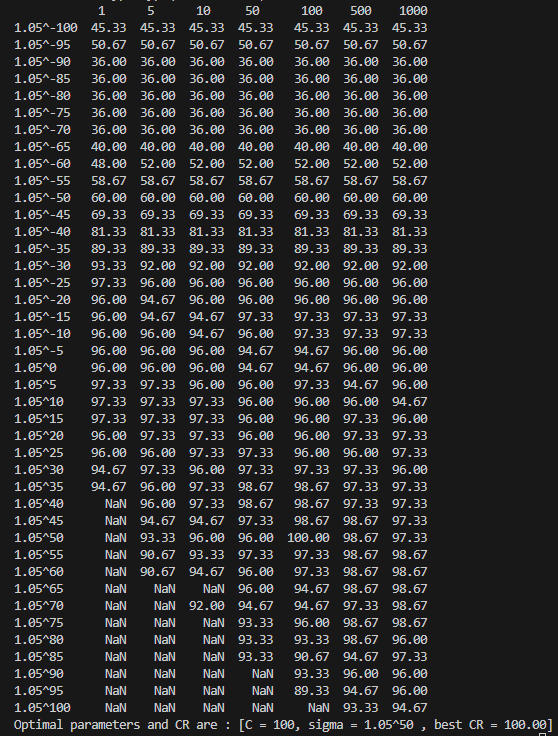
Ans : 從結果上來看，sigma對分類率的影響較大，sigma的影響我在hw3也有提到，sigma越大，標準差越大， RBF kernel更平緩而不集中，我的特徵數值帶入都分散，導致了泛化能力或許較高，但也可能欠擬合；另一方面若sigma越小，標準差越小， RBF kernel集中於局部區域，我的特徵數值帶入都集中於某處，導致過擬合發生，泛化能力低，在我的結果中可以發現最佳分類率也及終於中後段，也就是sigma一樣不能太大或太小導致過擬合及欠擬合發生。

* 1. 若分析過程不採用two-fold cross validation，則分類率是否會更高?

Ans : 以下是我分別用fold1和fold2當training data的各自分類率

1. Fold 1 當training ，CR1:



1. Fold 2 當training ，CR2:

由各自的分類率可看出，整體來說fold 1當training data的分類率CR1比fold 2當training data的分類率CR2來的高，但CR2最高分類率有到100%，所以不做two-fold cross validation，結果取決於擬採用哪個fold當training，像用fold 2當training data可能會有最高的分類率，但難以代表模型好壞，因為data使用上太單一沒有跟其他做比較，只是剛好此fold能得到好分類率，所以才需要做多折驗證來證明在各種組合下模型都有一定的分類水準，不然就只是剛好結果好看而已了。