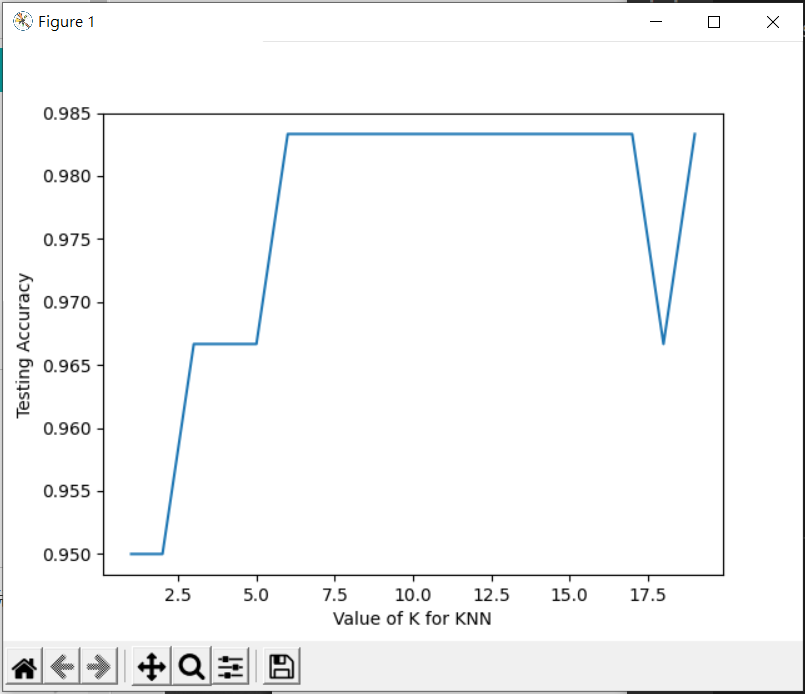
Project4\_108820006\_楊品賢 個人報告

本次作業要我們利用KNN與Adaboost分類器進行訓練，讓電腦透過只觀察花的訊息去預測花的品種。調整訓練集與測試集的比例，可以取得不同的結果，我認為這樣的過程是有趣的。

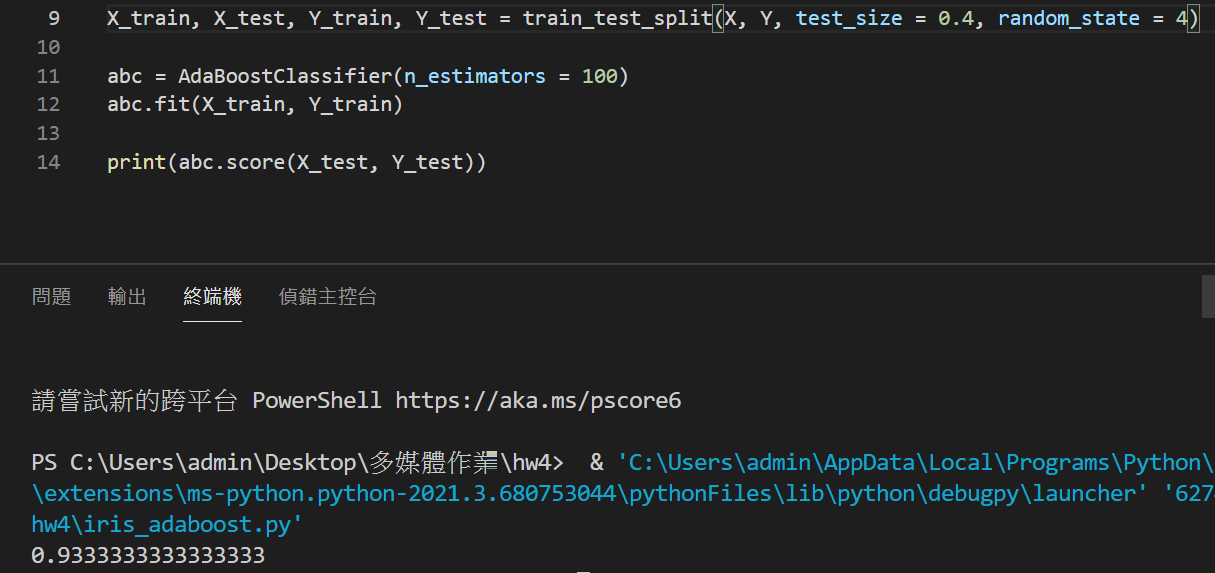
1. KNN：

(test\_size = 0.4, random\_state = 4, k = range(1, 20))



相同test\_size下，k值越高，準確率越高；針對test\_size比例作調整，test\_size比例越低，準確率越穩定(受k值的影響越小)。但當test\_size過低的情況下，平均準確率可能會因測試集樣本數過少而有較大的浮動(若test\_size = 0.05，測試集僅有8組數據。只要錯一個，準確率就不到90%，遠低於正常test\_size值的準確率)。因此，平均準確率與test\_size有著密切的關係。test\_size越低，給電腦練習的樣本數就越多，判斷會越精準；但相對的，進行test時答錯一次影響的準確率程度就越大。

1. Adaboost



test\_size僅會在比例較極端的情況下才會對準確率有明顯的影響(如KNN所述，test\_size會同時影響訓練集與測試集的樣本數。當test\_size比例在中間值，如0.3~0.8，準確率其實都差不多)。

n\_estimaters為分類器的迭代次數，迭代次數越多準確率越高。迭代次數較少時，增加迭代次數對準確率的影響較明顯(類似對數關係)。例如迭代2次及1次時的準確率差距很大，但迭代100次與99次的準確率就無明顯變化。