Q1-1

Q1-2

(1)precision與recall是兩個不同的指標，有時會需要trade off。

所以當我們不只考慮true positive/retrieved set(precision)，同時也希望考量覆蓋率(recall)的時候，就會用到F1這個綜合指標。

(2)為了非線性轉換。

(3)在ML裡，

bias過大代表underfitting，Variance過大代表overfitting。

(4)RF是由多個DT組合成的model，根據其背後推倒過程，可知model增加DT數量時所增加的一般誤，上界受1.樹之前的相關性2.樹的強度所控制。

不剪枝是為了確保單一樹的強度。

(5)就是由{0,1}組成的稀疏矩陣。

(6)early stopping，人為設定迭代次數。

regularization，添加loss function的懲罰項。

validation set，在training set, test set之外另設validation set，以避免調整模型時的人為overfitting。