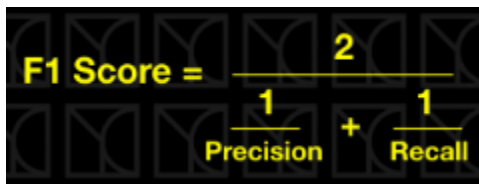


- When will we use F1-score instead of precision(accuracy)?

如果今天我覺得 Precision 和 Recall 都同等重要，我想要用一個指標來統合標誌它，這就是 F1 Score。Precision 看的是在預測正向的情形下，實際的「精準度」是多少，而 Recall 則是看在實際情形為正向的狀況下，預測「能召回多少」正向的答案。當 $\text{beta}=1$ 時就是 F1 Measure，代表 Precision 和 Recall 都同等重要，那如果我希望多看中一點 Precision，那 beta 就可以選擇小一點，當 $\text{beta}=0$ 時，F Measure 就是 Precision；如果我希望多看中一點 Recall，那 beta 就可以選擇大一點，當 beta 無限大時，F Measure 就是 Recall。


$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}$$

- Why don't we use binary classification function as the activation function in neural networks?

因為輸出 y 的值只會有兩種，即 1 或 0。

- What is the bias and variance of a machine learning algorithm?

bias 的最常見解釋是 bias–variance tradeoff。從本質上講，bias 是模型中的錯誤來源，導致它過度概括和不適合您的數據。相反，variance 是對數據中噪聲的敏感性，導致模型過度擬合。我們稱之為 tradeoff(權衡)，因為改進通常會使另一個指標變得更糟。

- When training a single tree in random forest, we don't prune the tree, why?

修剪(Pruning)是決策樹中用於減少過度擬合的合適方法。但是，通常隨機的森林會提供完整深度的良好表現。隨著隨機森林訓練使用自助聚合（或替換採樣）以及隨機選擇特徵進行分割，樹之間的相關性很低。如果您仍想控制隨機森林中的訓練，需要調整樹的深度而不是修剪。

- What is one-hot encoding?

one-hot encoding 是將分類變量轉換為可以提供給 Machine Learning 算法以在預測中做得更好的形式的過程。

- How to prevent overfitting in neural networks? (write down anything you know)

提早停止訓練、參數正規化、使用 Dropout