<u>Performance Metrics used to evaluate/compare Classifiers:</u>

Binary classification: Predicting Yes means Rejecting Null hypothesis, for Alternate
Confusion Matrix

	Pred Yes	Pred No
True Yes	True Positive	False Negative (Type 2)
True No	False Positive (Type 1)	True Negative

Overall **Accuracy**: Trace of matrix / Sum of elements (Matrix norm for p=1)

Accuracy on **positive** class: True Positive / (True Positive + False Negative) - Also called **Recall on positive class**
Accuracy on **negative** class: True Negative / (True Negative + False Positive) - **Recall** of **negative** class

Precision:
Precision on positive class: True Positive / (True Positive + False Positive)
Precision on negative class: True Negative / (True Negative + False Negative)

F1-score = Harmonic mean of Recall and precision. (For respective classes)
= Precision * Recall / (Precision + Recall)
By nature of it, F1 is high only if both Precision and Recall is high. This actually means, should minimize both FN and FP . More entries should be on the diagonals.

Sensitivity: Same as Recall on positive class. Also called Power. **TP / True Yes**
1 - type 2 error. (Type 2 error in statistics is failing to reject H0 when one should go with it)
True positive rate is another name

Specificity: Same as Recall on negative class. **TN / True No**
1- type 1 error (Type 1 error in statistics is rejecting H0 when one shouldn't)
True negative rate is another name (1-FP/True No)

Likelihood Ratio positive = sensitivity / (1 − specificity)

ROC plotting: TP and FP can be obtained for different thresholds on the probabilities calculated on positive class (same can be done for negative class as well)
A plot of TP in y axis, Vs FP in x axis is Called Receiver Operating Characteristics (ROC)
In terms of Sensitivity and Specificity, it is (Sensitivity * Positive class strength) Vs ((1-Specificity) * Negative class strength)

AUC: Area Under Curve (AUC) is one number to summarize ROC plot.

3 class classification:
Confusion Matrix

	Pred 0	Pred 1	Pred 2
True 0	True Positives on 0	False 0 to 1	False 0 to 2
True 1	False 1 to 0	True Positives on 1	False 1 to 2
True 2	False 2 to 0	False 2 to 1	True Positives on 2

Accuracy: Trace of the Matrix / Sum of elements (Matrix norm for p=1)

Sensitivity and Specificity for 3-class: Let m[i,j] denote matrix element at ith row and jth column, i and j ranging from 0 to 2

Sensitivity (recall) for 0th class: m[0,0] / (m[0,0]+m[0,1]+m[0,2])
Specificity for 0th class:
(m[1,1]+m[2,2]+m[1,2]+m[2,1])/(m[0,1]+m[0,2]+m[1,1]+m[2,2]+m[1,0]+m[2,0]) - note that this is different from recall, unlike in binary classification, where this was recall for negative class. Here negative class has 2 members.

Similarly define for the other two classes. Sensitivity numbers are same as recalls, and specificity is not to be interpreted that way.

Precision also can be extended similarly, and then f1 score is the same definition from corresponding precision,recall numbers

ROC plotting: Note that TP and FP are not so simple in this case. Each class has two other boundaries. There is a 3-D curved surface in this case, instead of a curved line on a 2-D plane. The simplest approximation is to take One Vs Rest classifier (thus have 3 classifiers) and vary thresholds, and get TP, FP numbers and plot them on 2-D. There are 3 such ROCs, and so 3 AUC numbers. But this is not the same as taking volume under the surface mentioned.

 Author: Aparna Shastry