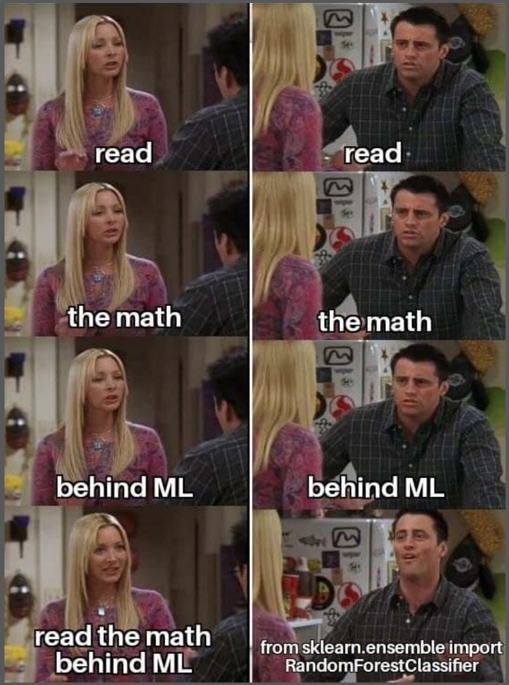
Discussion ROC

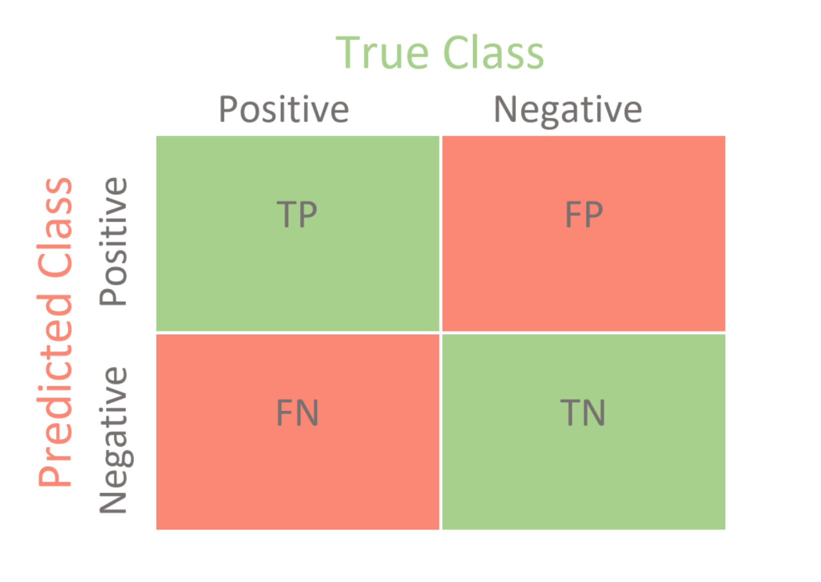


- We learned that in logistic regression, our goal is to create a model with maximum likelihood
- But we also have an additional way to further evaluate the model

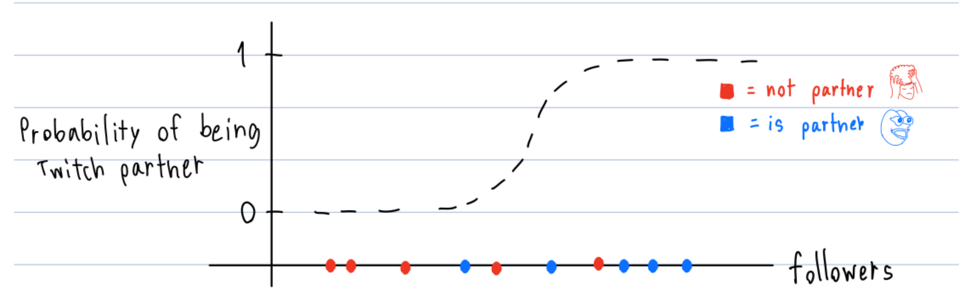
Introduce: Confusion Matrix

A confusion matrix is a table that is often used to **describe the performance of a classification model**

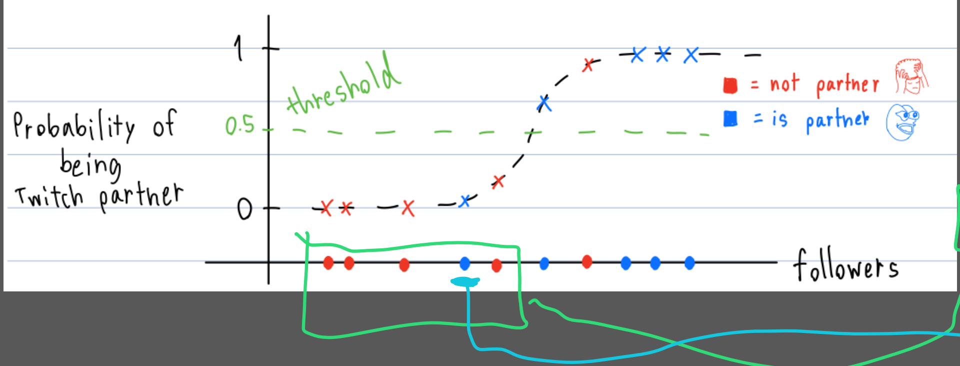
- True positive (TP): predict positive, and actually is positive
- True negative (TN): predict negative, and actually is negative
- False positive (FP): predict positive, and actually is negative
- False negative (FN): predict negative, and actually is positive



Example



Set threshold at 0.5

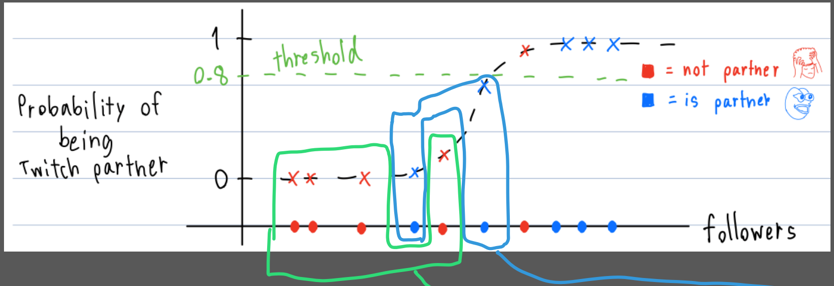


Predicted

Actual

	Is partner	Is not partner
Is partner	4	1
Is not partner	1	4

Now change threshold to 0.8

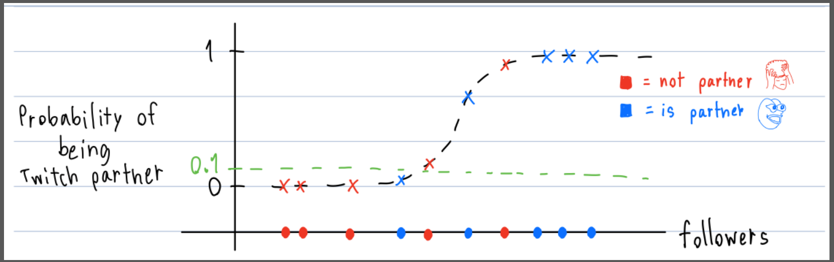


Predicted ↓

Actual

	Is partner	Is not partner
Is partner	3	1
Is not partner	2	4

Now change threshold to 0.1



Predicted ↓

Actual

	Is partner	Is not partner
Is partner	4	2
Is not partner	1	3

threshold at 0.5

Predicted		Actual	
		Is partner	Is not partner
	Is partner	4	1
	Is not partner	1	4

threshold at 0.8

Predicted		Actual	
		Is partner	Is not partner
	Is partner	3	1
	Is not partner	2	4

threshold at 0.1

Predicted		Actual	
		Is partner	Is not partner
	Is partner	4	2
	Is not partner	1	3

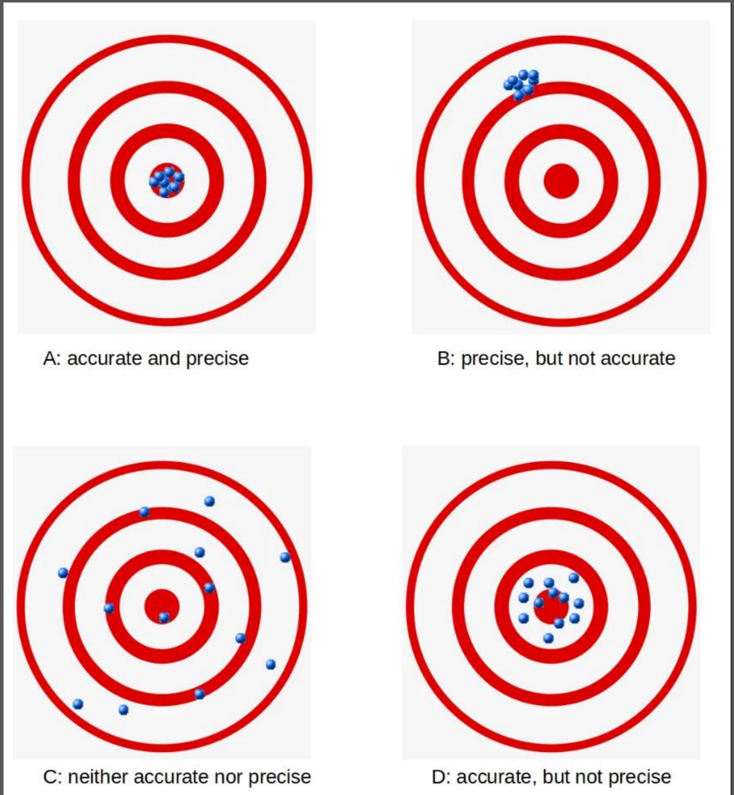
From these tables, what can we infer?

- **Accuracy:** Overall, how often is the classifier correct?

- **Precision:** When it predicts yes, how often is it correct? (It tells us how much you should trust the model when it says it found something)

- **F Score:** This is a weighted average of the true positive rate (recall) and precision

- **Recall**: (aka sensitivity, true positive rate) It tells us how much the model can find the thing you're looking for



By looking at the matrices, it is difficult to visualize, so let's make a graph!

threshold at 0.8

Actual

Predicted.

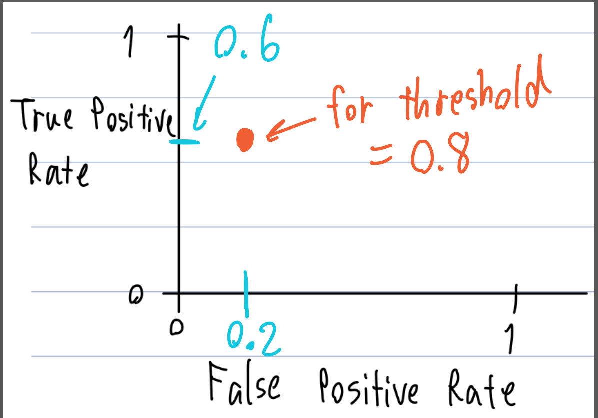
	Is partner	Is not partner
Is partner	3	1
Is not partner	2	4

True positive rate = TP / (TP + FN) = 3 / (3 + 2) = 0.6

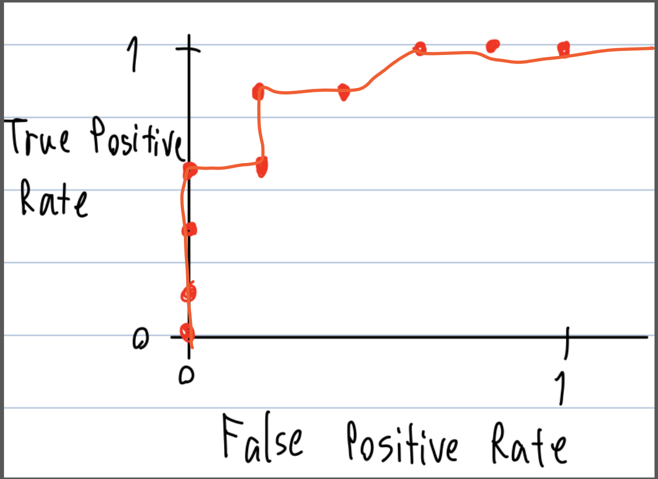
False positive rate = FP / (FP + TN) = 1 / (1 + 4) = 0.2

Positive ⟷ is partner

Negative ⟷ not partner



Keep plotting with different threshold, and we will have the following ROC curve graph for this model:

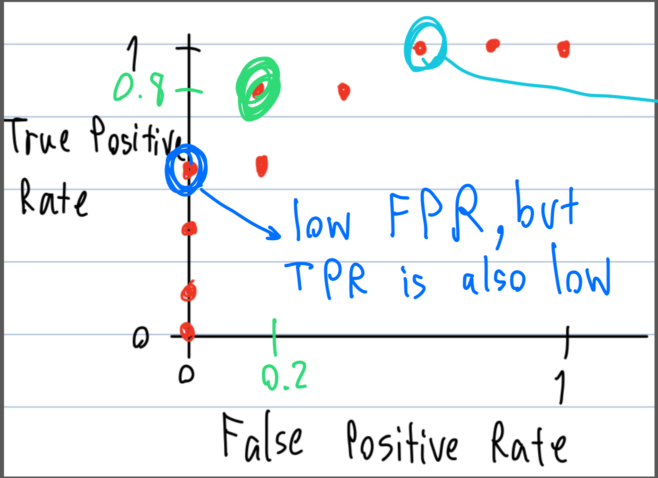


There are 2 things we can consider from ROC curve:

1. We can choose which threshold is good for our goal. For example, we want least false positive rate, we will choose the very left one

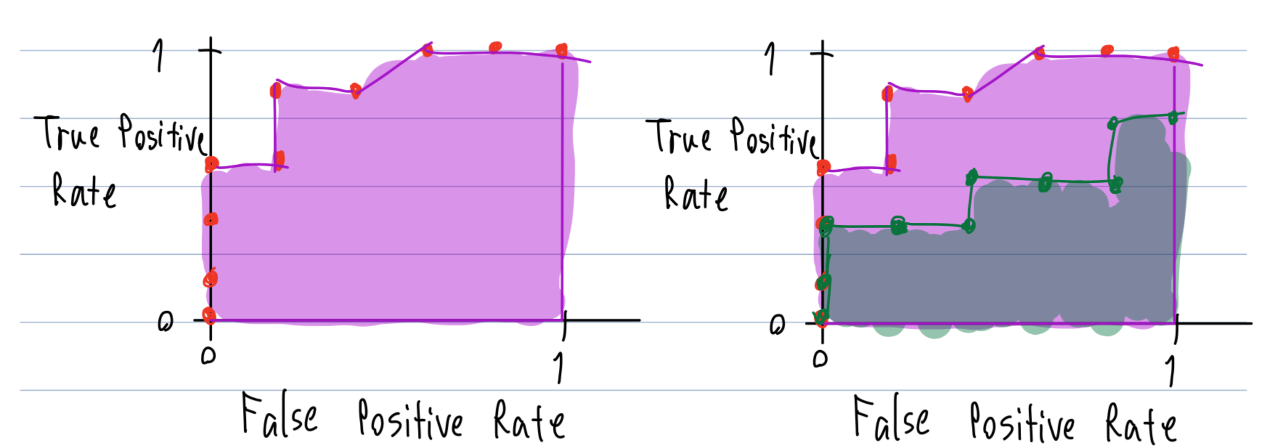
Note: always consider TP, FP.

For example, incorrectly classify that a person has a disease [FP]



Has high TPR, but FPR is also high

2. When we create multiple classification model, each model has its own ROC curve. In general, the best performance model has higher area under the curve (AUC)



Ex Pink → logistic regression , Green → Decision Tree

But you know, I learned something today

class 1 : 0.2
Class 2 : 0.6
Class 3 : 0.1
class 4 : 0.1

$\boxed{\text{Note: Multiclass}}$
↓
Generally, classify as class n, if it returns highest probability.

- We use confusion matrix to get TP, FP, TN, FN
- We can evaluate our model's performance from confusion matrix
- On our model, we have to weigh whether to include FP/FN at a cost of accuracy

$\boxed{\text{Break until 2:05}}$