Carnegie Mellon University
Heinz College

Unstructured Data Analysis

ROC & Precision-Recall Curves, and How to Use These For Model Selection

George Chen

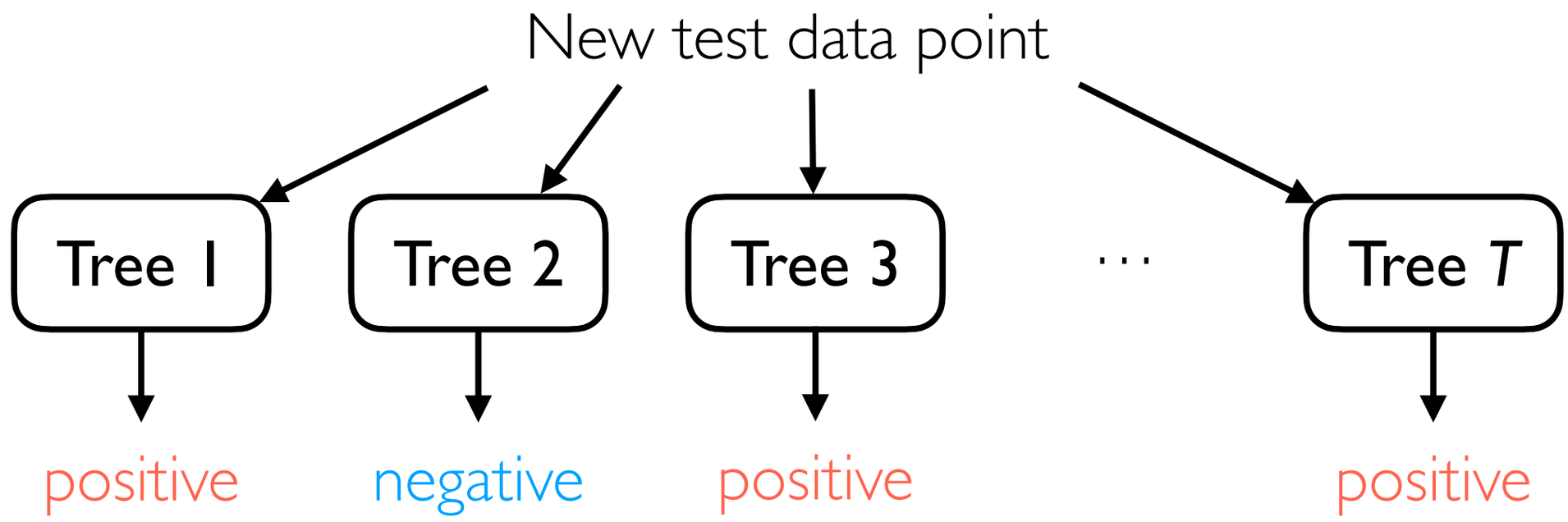
Another Way to Benchmark

- In the lecture demo on basic predictive data analysis: to assess model quality, we compare test set prediction accuracy across different models and also look at confusion matrices

- For binary classification, we can do a more detailed analysis

Binary Classification: ROC Curves

For simplicity, think of the random forest for now

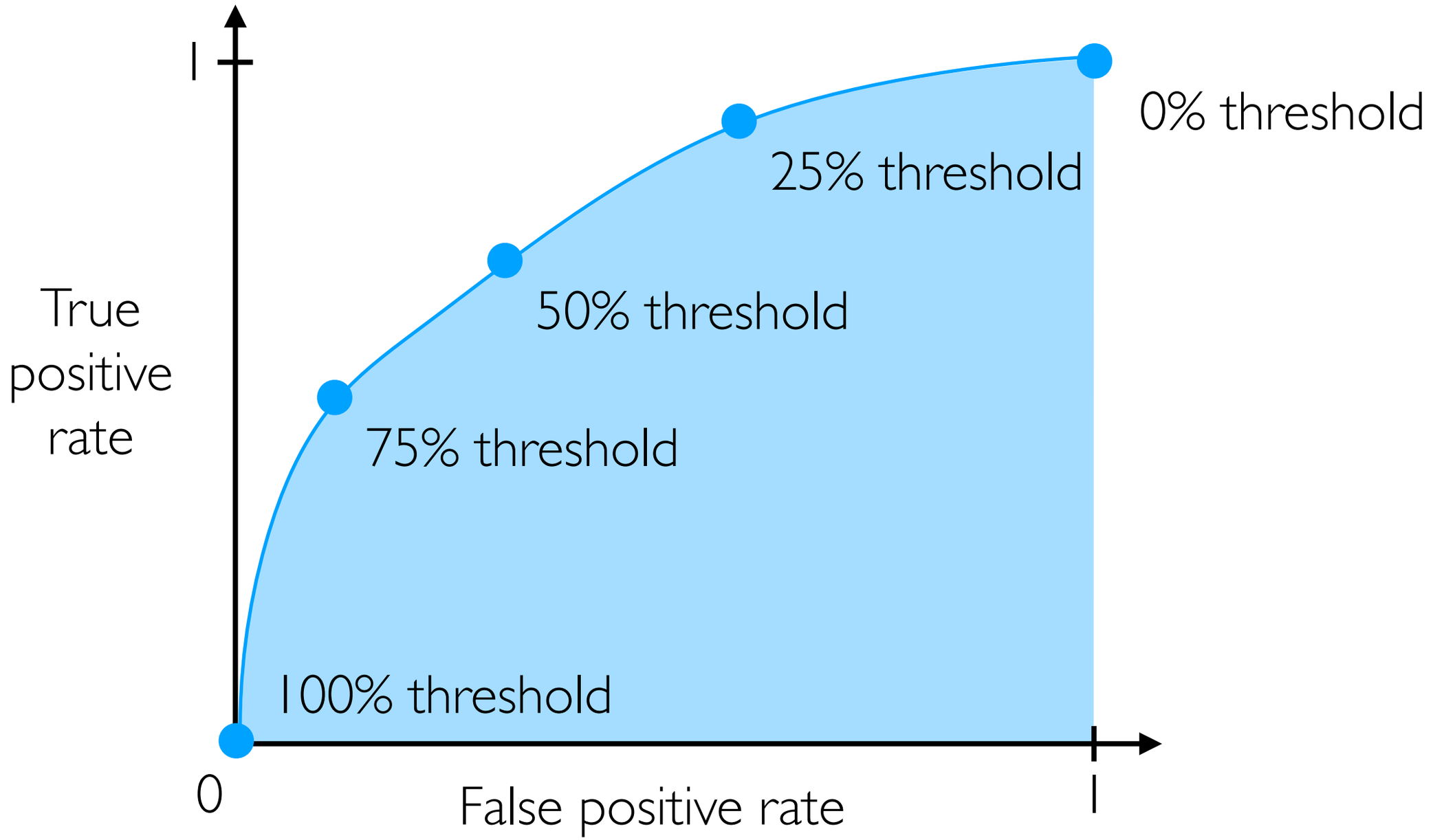


Final prediction: majority vote of the different trees' predictions

≥ 50% of trees need to say **positive** for final prediction to be **positive**

We can vary this 50% threshold!

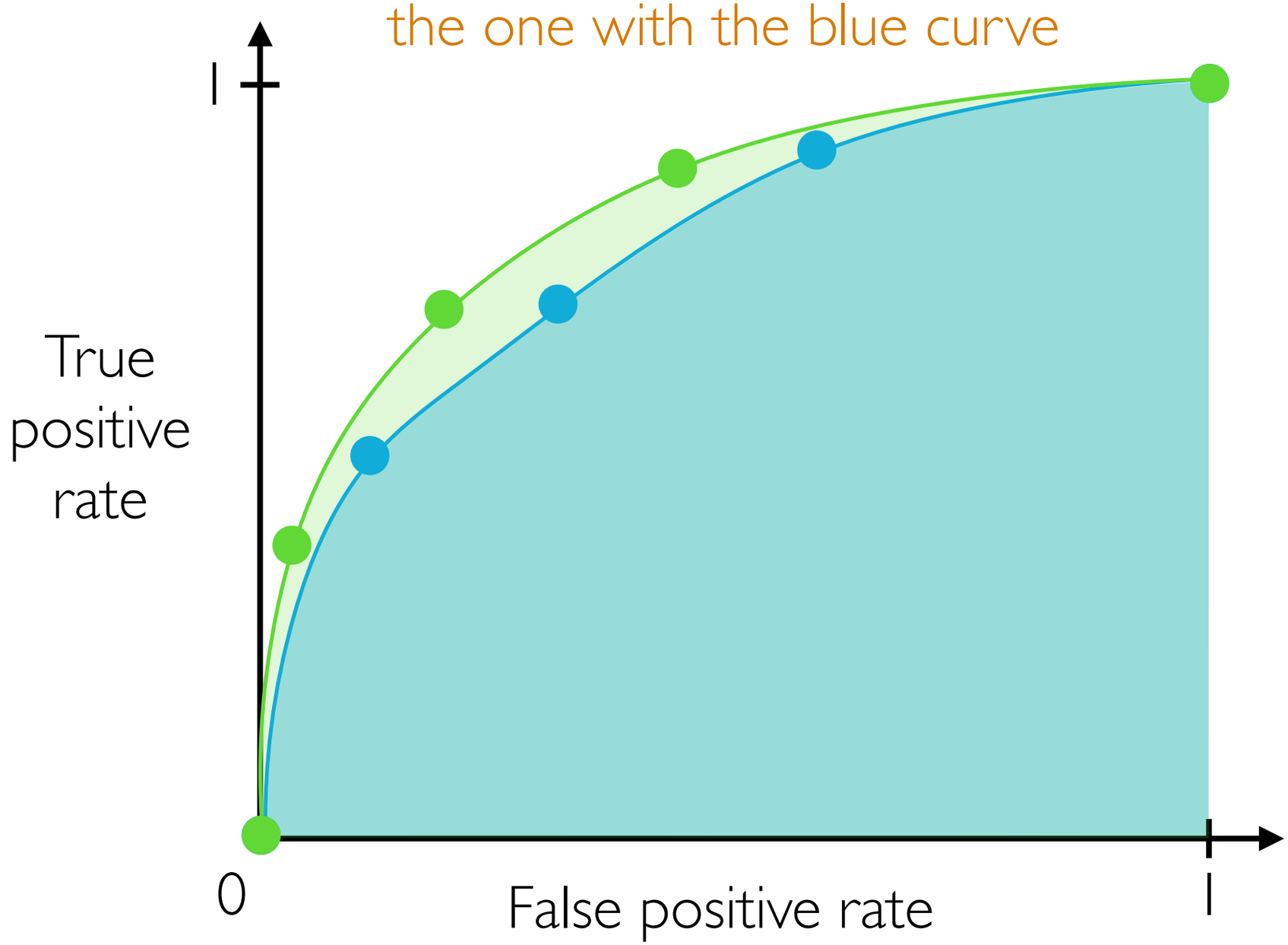
Binary Classification: ROC Curves



Error rates are computed on test data

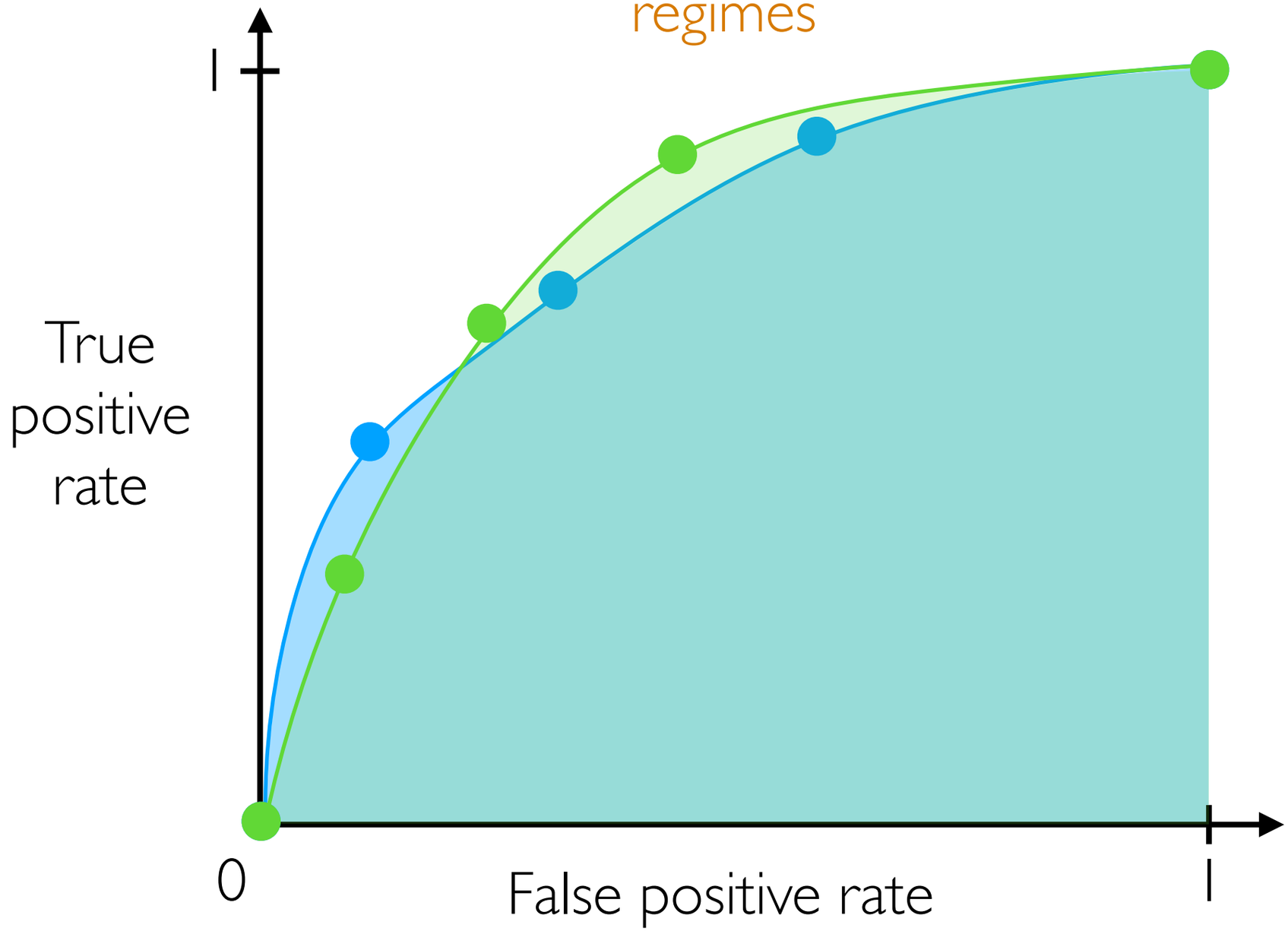
Binary Classification: ROC Curves

A classifier with the green curve is better than the one with the blue curve

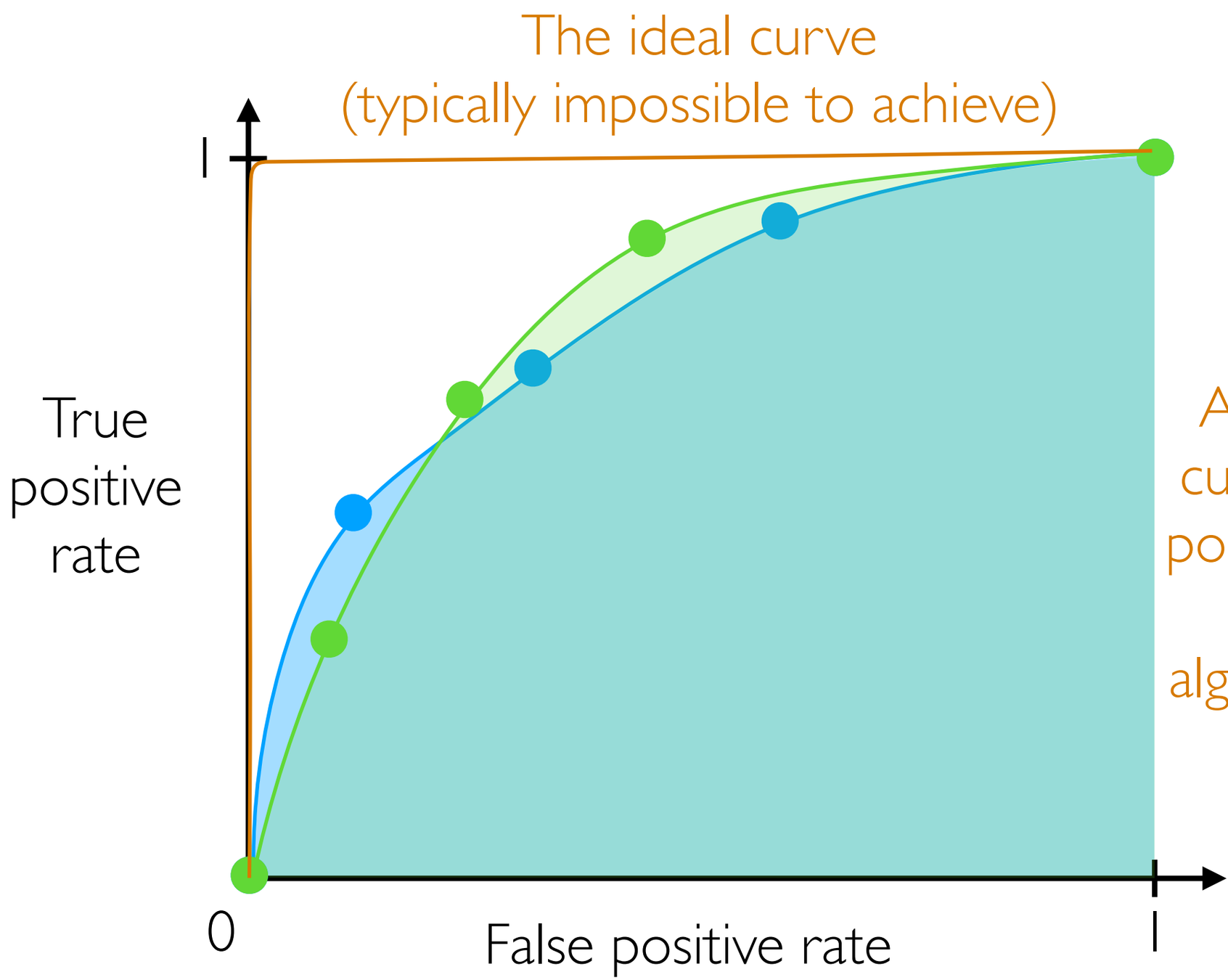


Binary Classification: ROC Curves

It's possible that algorithms are better in different regimes



Binary Classification: ROC Curves



Area under the curve (AUC) is a popular metric for comparing algorithms (higher is better)

Binary Classification: ROC Curves

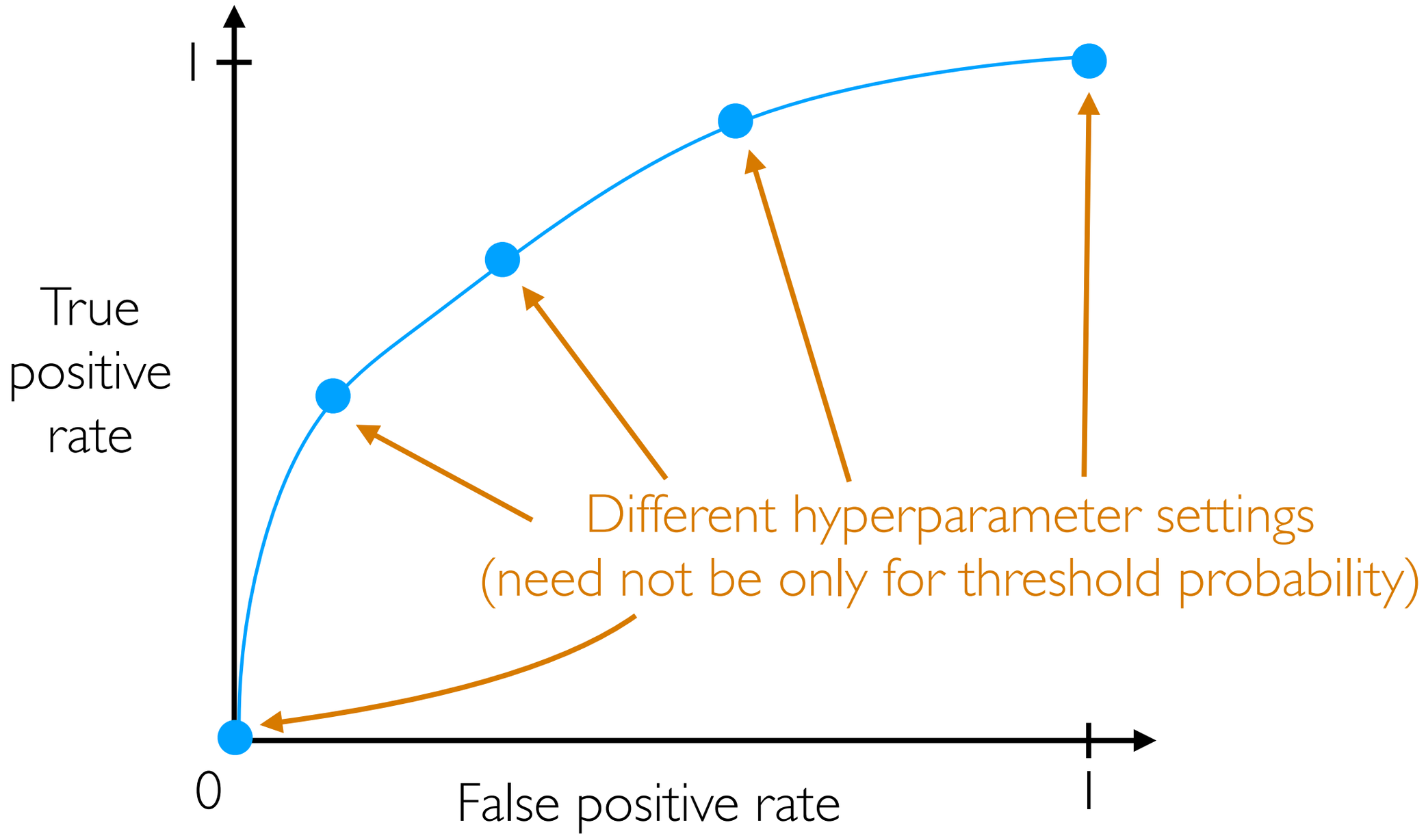
What we just saw:

- For a classifier that we can set the threshold probability to different values, we can plot an ROC curve
- True positive rate (TPR) and false positive rate (FPR) are evaluated on test data

Other variants are possible:

- Plot precision vs recall instead of TPR vs FPR
- Can actually plot ROC/precision-recall curves sweeping over hyperparameters aside from threshold probability!
- For ROC/precision-recall, rather than evaluating on test data, can evaluate on validation data during training *to help choose hyperparameters*

Binary Classification: ROC Curves



Can also be computed on validation data instead of test data!