Testing and Error Estimation

Machine Learning

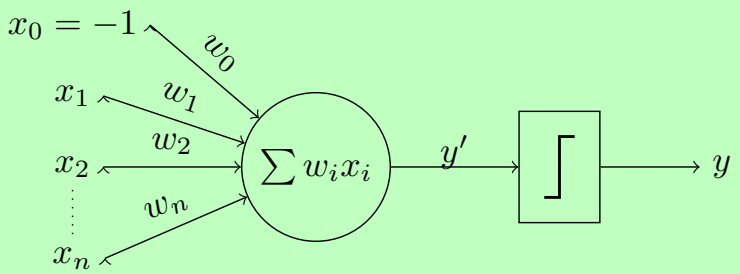
Prof Hans Georg Schaathun

Høgskolen i Ålesund

5th February 2016

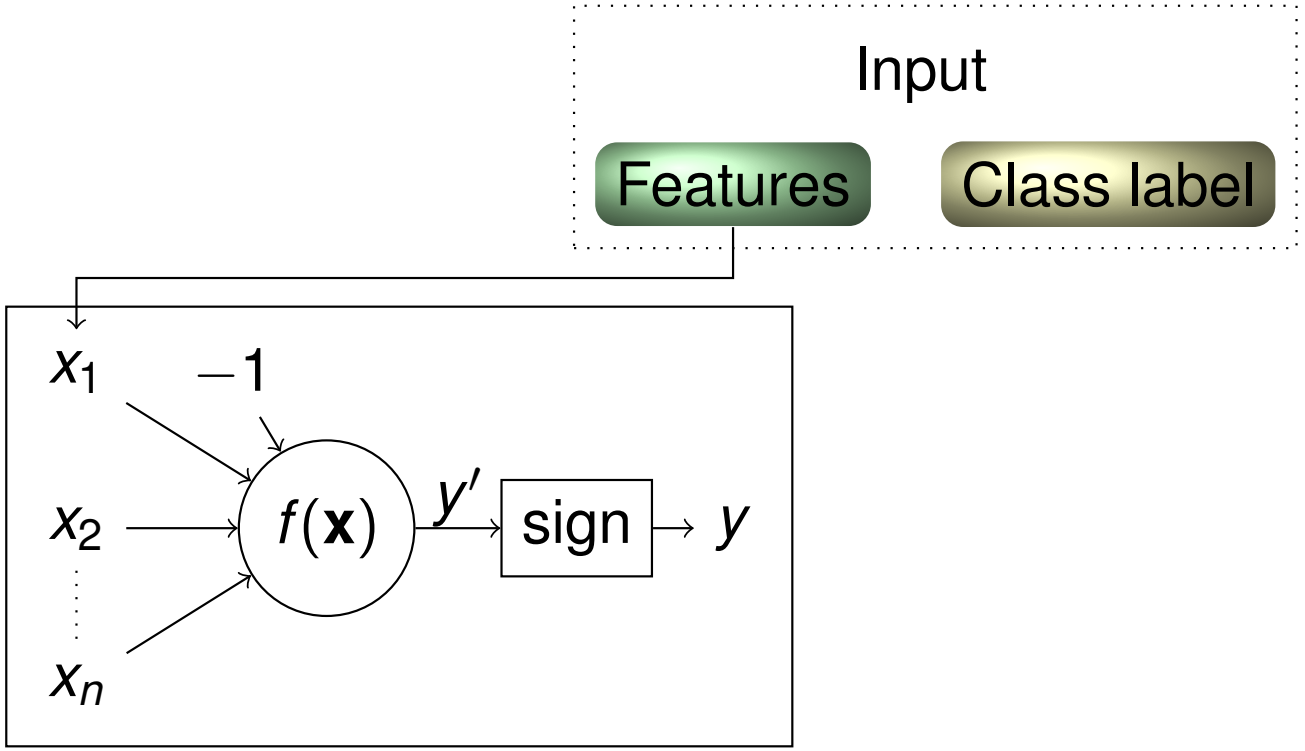
Data set

Training set
Used up in training

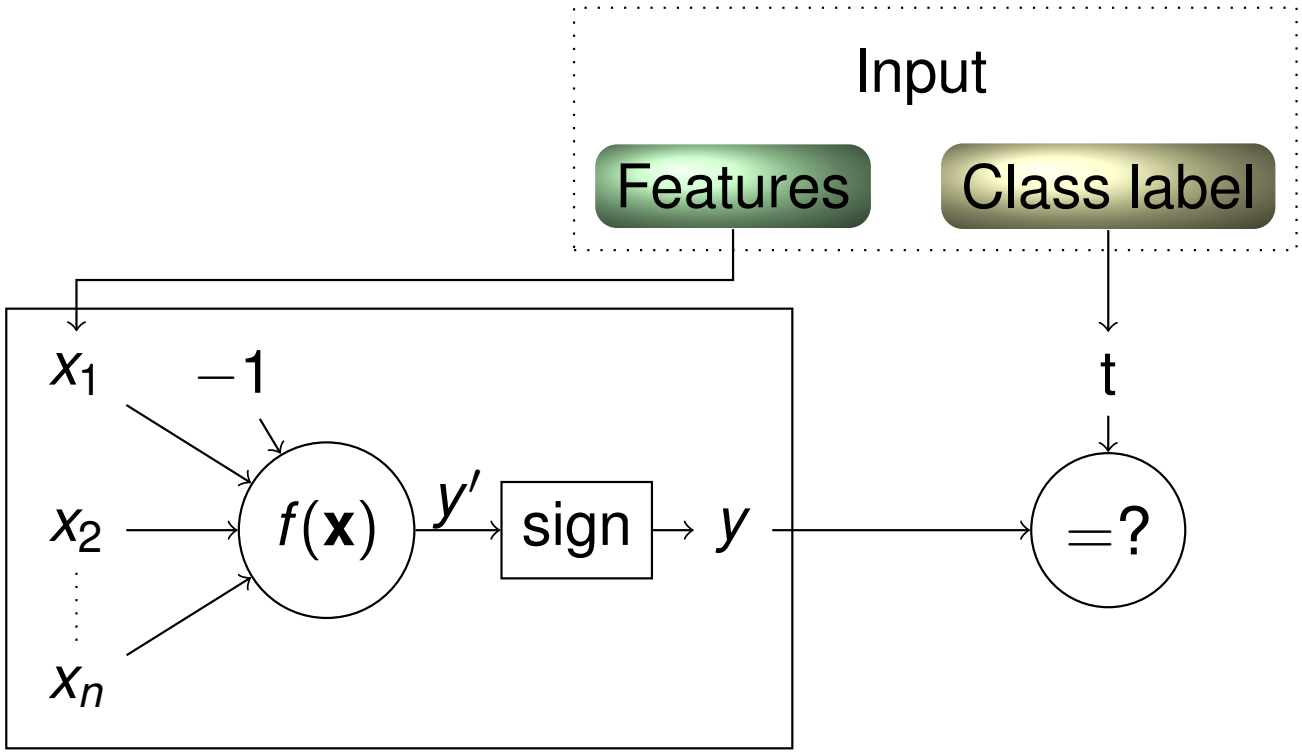


Test set
Find error probability

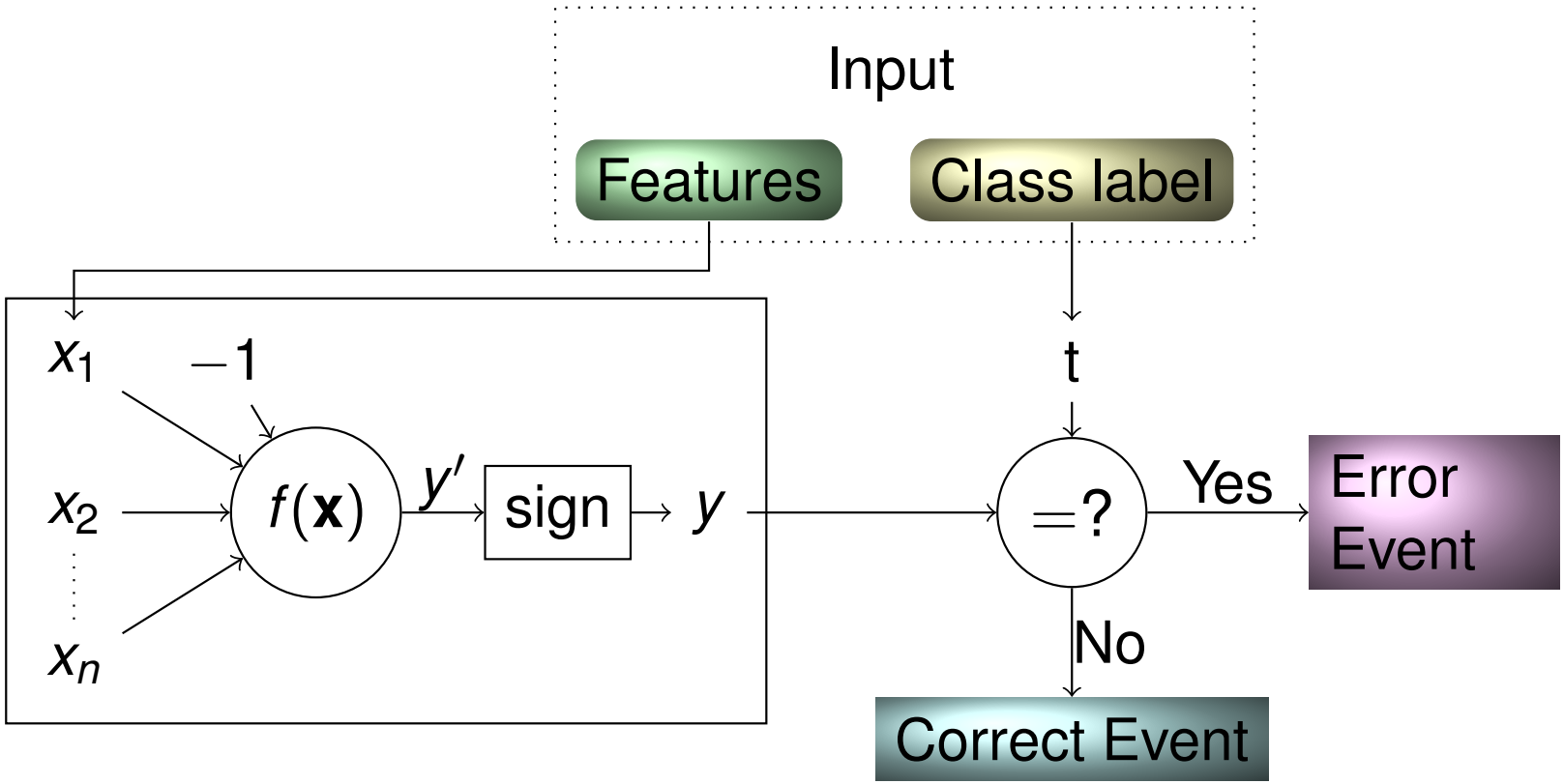
Testing the Neuron



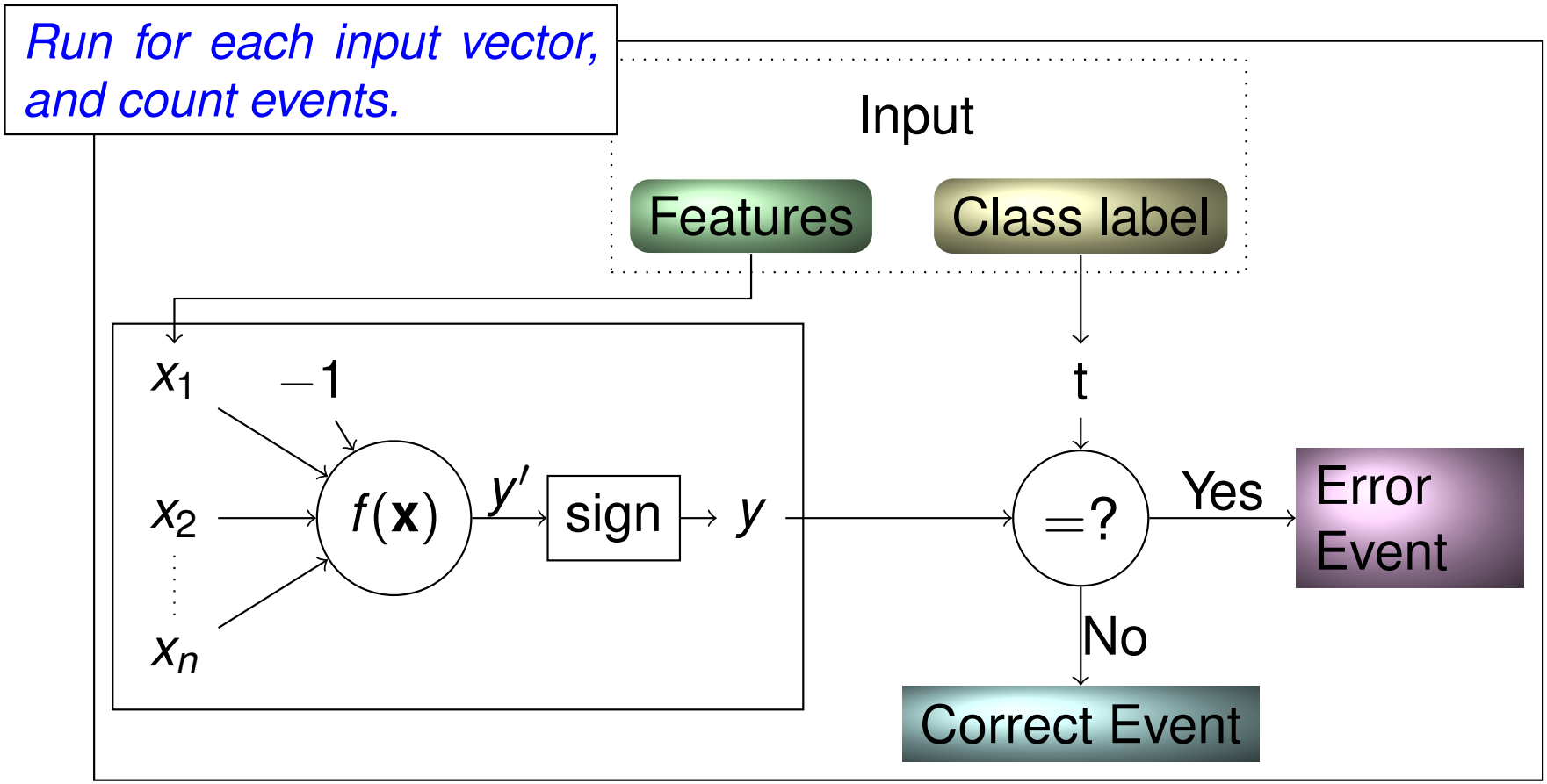
Testing the Neuron



Testing the Neuron



Testing the Neuron



Error Rate

1. error rate: $\frac{e}{e+c}$
2. e error events
3. c correct classifications

What is the error probability?

Error Rate

1. error rate: $\frac{e}{e+c}$
2. e error events
3. c correct classifications

What is the error probability?

— Error probability p_e

— Estimator: $\hat{p}_e$ (stochastic variable)

— Estimate: observation r_e (error rate)

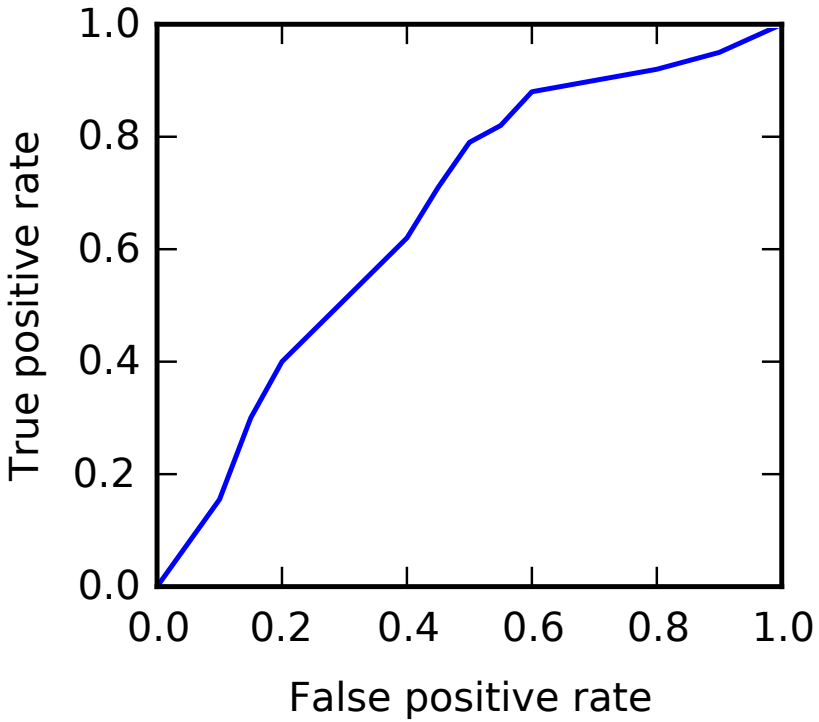
— Confidence interval: $(\hat{p}_l, \hat{p}_u)$

Confusion Matrix

		Prediction	
		Malign	Benign
Actual Class	Malign	True Positive	False Negative
	Benign	False Positive	True Negative

The ROC curve

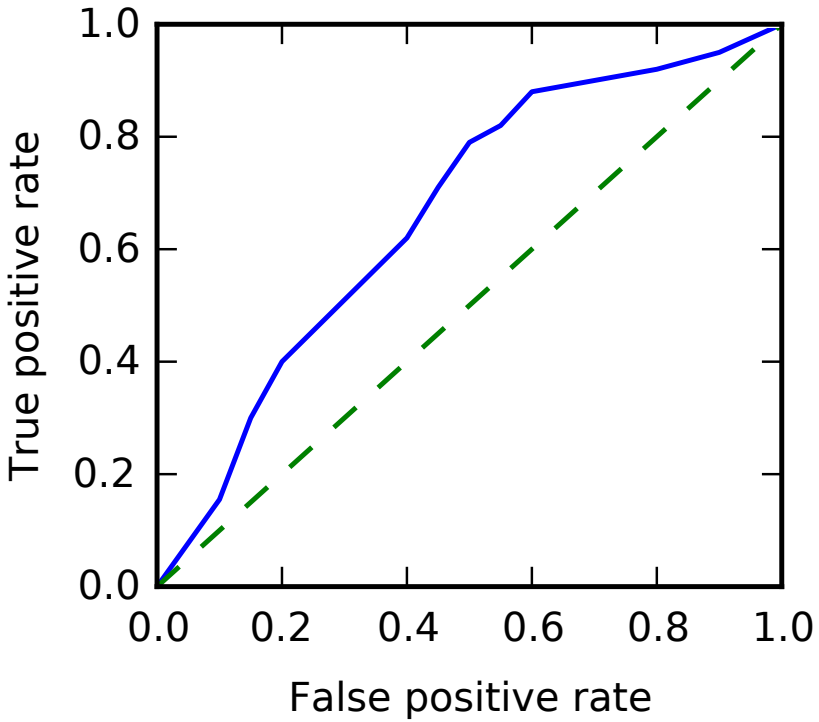
- We can vary w_0
- Receiver operating characteristic (ROC)



The ROC curve

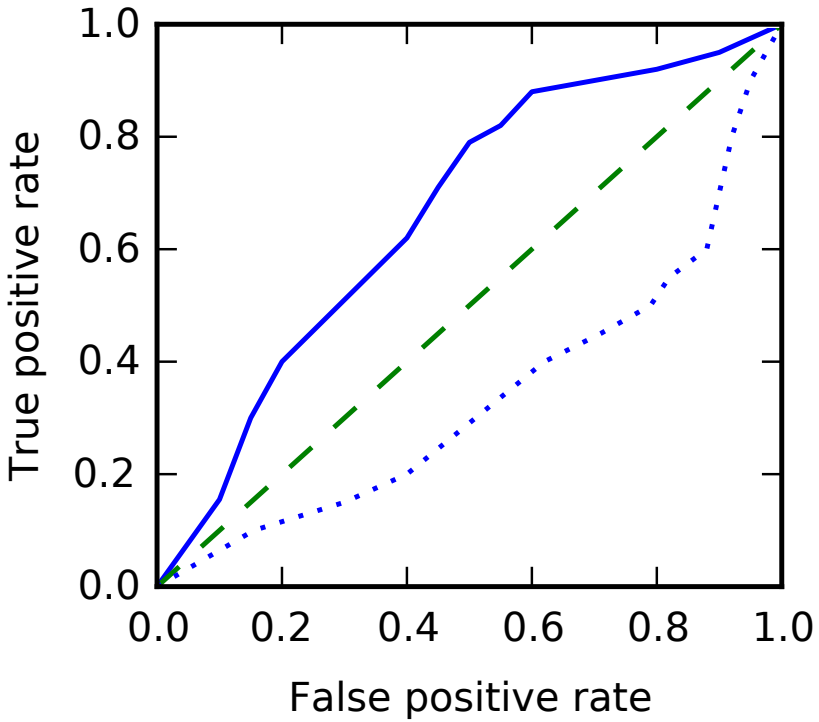
— We can vary w_0

— Receiver operating characteristic (ROC)



The ROC curve

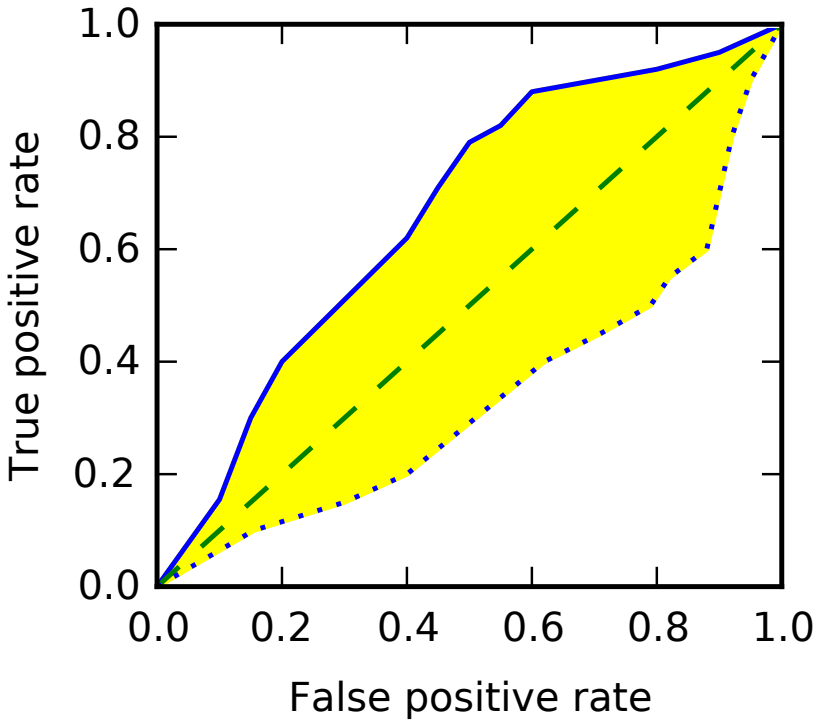
- We can vary w_0
- Receiver operating characteristic (ROC)



The ROC curve

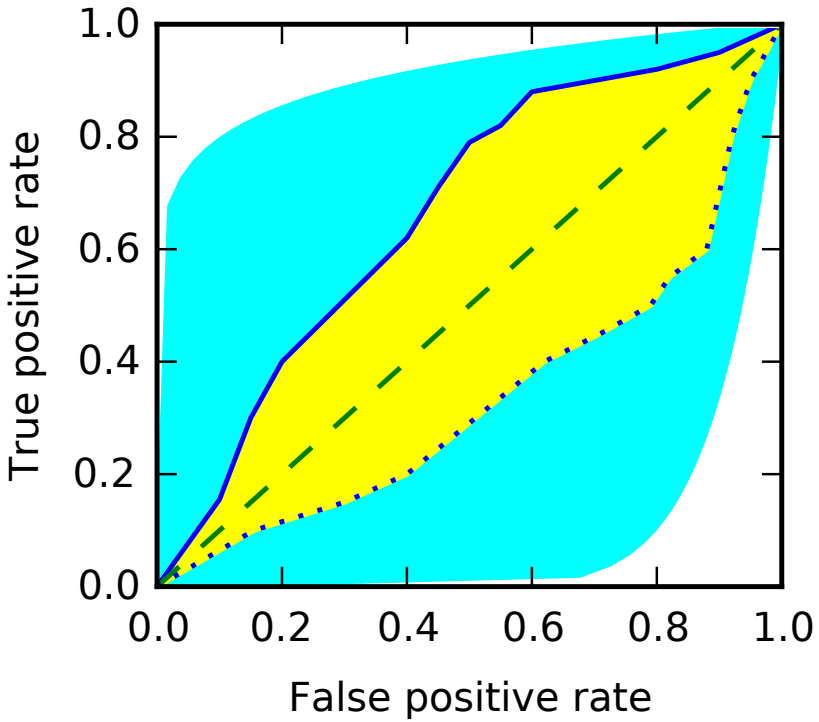
— We can vary w_0

— Receiver operating characteristic (ROC)



The ROC curve

- We can vary w_0
- Receiver operating characteristic (ROC)



Summary

Testing

1. Use a test set independent of the training set
2. Test set with known class labels
3. Do recall, and compare to known labels

Evaluation

1. Statistical analysis of test results
2. How large test set?