Predictive Models for Healthcare Analytics

A Case on Retrospective Clinical Study

Mengling 'Mornin' Feng
mfeng@mit.edu
mornin@gmail.com

Learning Objectives

* After the lecture, students should be able to:
  * Define what are predictive models
  * Apply appropriate metrics to assess the performance of predictive models
  * Explain some basic predictive models, such as logistic regression, decision tree, neural network etc
  * Conduct retrospective clinical studies with appropriate predictive models
  * Interpret results from predictive models
  * Understand the future trend of predictive models

What are Predictive Models?

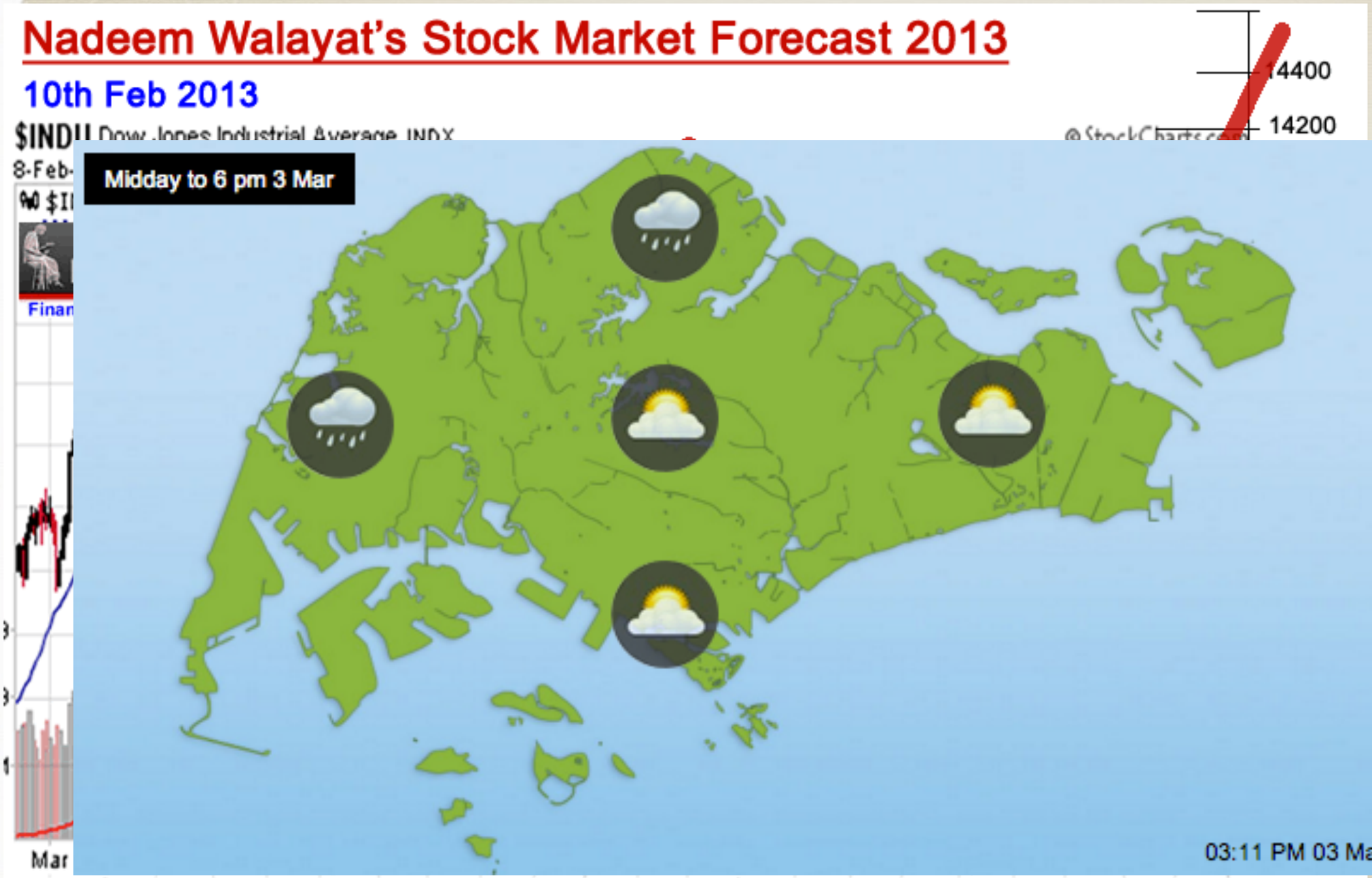
* To predict or to guess

What are Predictive Models?

* To predict or to guess

What are Predictive Models?

* To predict or to guess



What are Predictive Models?

* To predict or to guess

What are Predictive Models?

* To predict or to guess

What are Predictive Models?

* To predict or to guess

What are Predictive Models?

Classification (Supervised Learning)

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Classification (Supervised Learning)

VS

Assessment of Predictive Models

Assessment of Classification Models

Assessment of Classification Models

* Confusion Matrix

Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

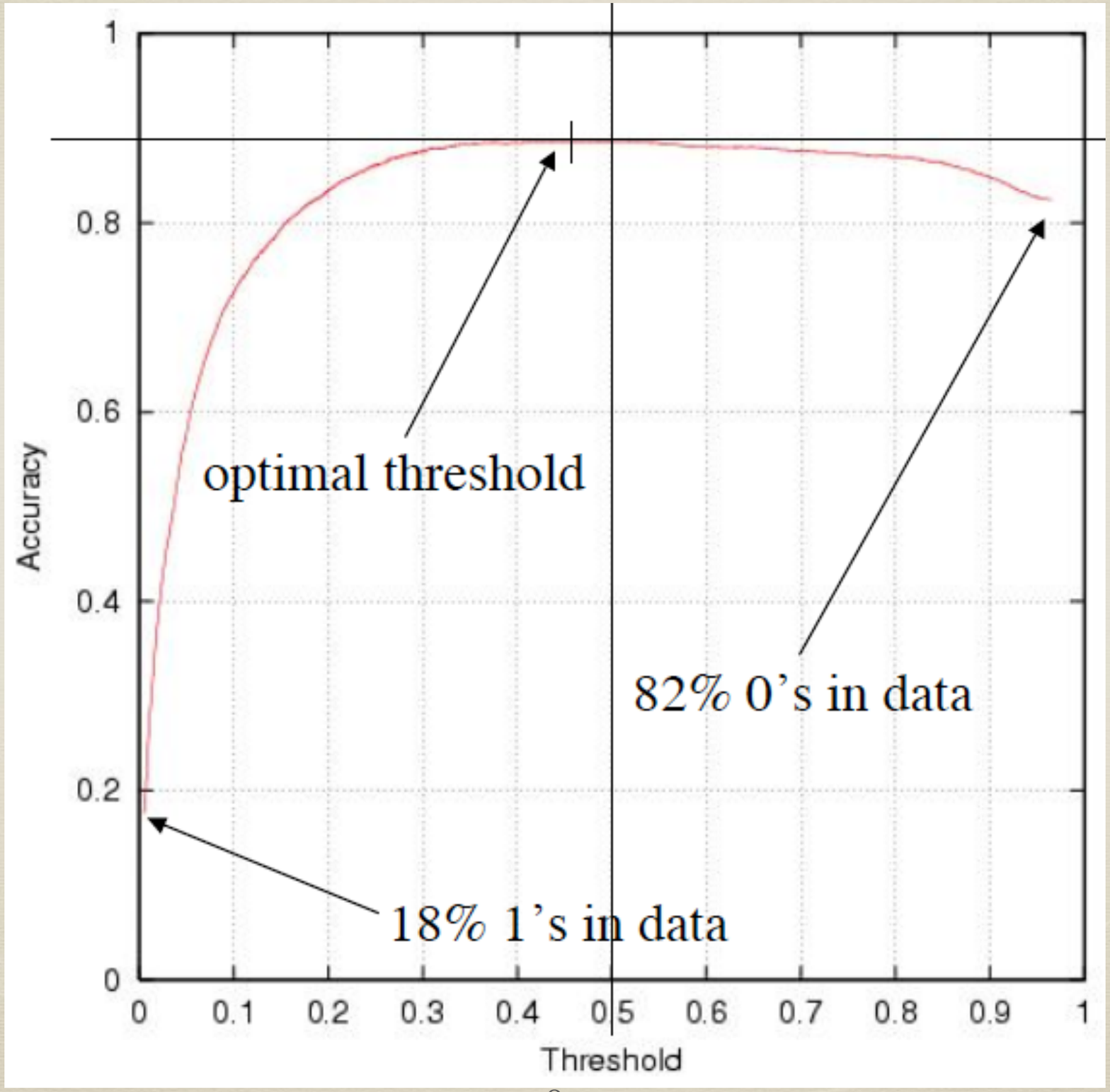
Assessment of Classification Models

* Confusion Matrix

	Predicted Yes	**Predicted No**
True Yes	**a** True Position (TP)	**b** False Negative (FN)
True No	**c** False Positive (FP)	**d** True Negative (TN)

* **Accuracy** = Correct Predictions / All Predictions
  = (a+d)/(a+b+c+d)

Prediction Threshold



Limitations of Accuracy

* Hard to interpret
  * Is 90% accuracy good?
  * Is 20% accuracy bad?
    * It depends! (on the base rate)
* Assume equal costs/weights on errors
  * Again it depends on the problem and applications
  * E.q.

Assessment of Classification Models

Assessment of Classification Models

* Error weights (costs)

Assessment of Classification Models

* Error weights (costs)

	Predicted Yes	**Predicted No**
True Yes	**a, w_a** True Position (TP)	**b, w_b** False Negative (FN)
True No	**c, w_c** False Positive (FP)	**d, w_d** True Negative (TN)

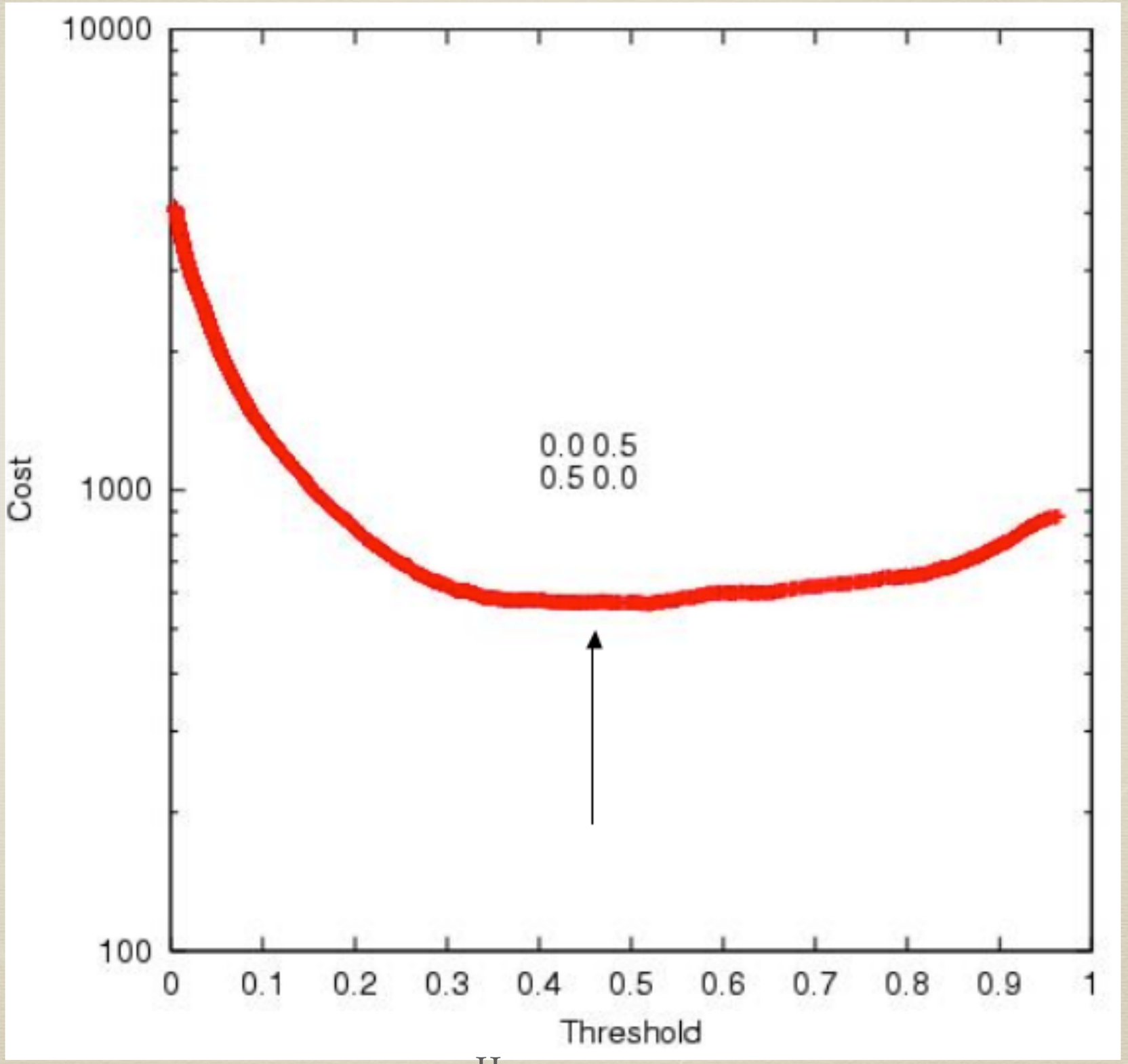
Assessment of Classification Models

* Error weights (costs)

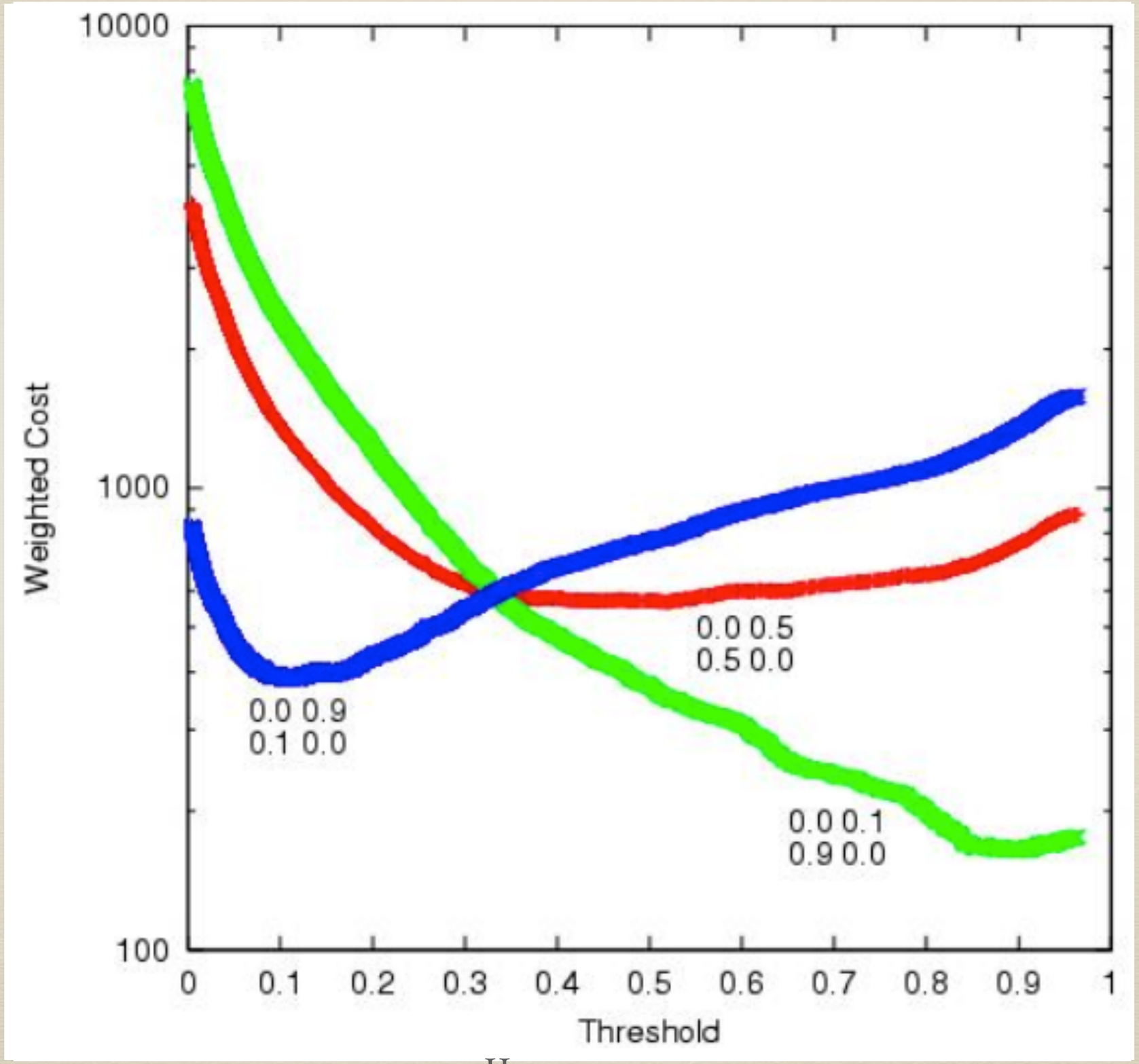
	Predicted Yes	**Predicted No**
True Yes	**a, w_a** True Position (TP)	**b, w_b** False Negative (FN)
True No	**c, w_c** False Positive (FP)	**d, w_d** True Negative (TN)

* **Error Costs** = $w_b * b + w_c * c$

Error cost


10000
1000
100
Cost
0.0 0.5
0.5 0.0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Threshold

Error cost


10000
1000
100
Weighted Cost
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Threshold
0.0 0.5
0.5 0.0
0.0 0.9
0.1 0.0
0.0 0.1
0.9 0.0

Assessment of Classification Models

* **Lift**
  * Commonly used in targeted marketing
  * Not interested in the entire population
  * How much more accurate the model is compared to random guessing when we predicted x% to be true

$$\frac{a/(a+b)}{(a+c)/(a+b+c+d)}$$

Assessment of Classification Models

* **Lift**
  * Commonly used in targeted marketing
  * Not interested in the entire population
  * How much more accurate the model is compared to random guessing when we predicted x% to be true

	Predicted Yes	Predicted No
True Yes	a	b
True No	c	d

$$\frac{a/(a+b)}{(a+c)/(a+b+c+d)}$$

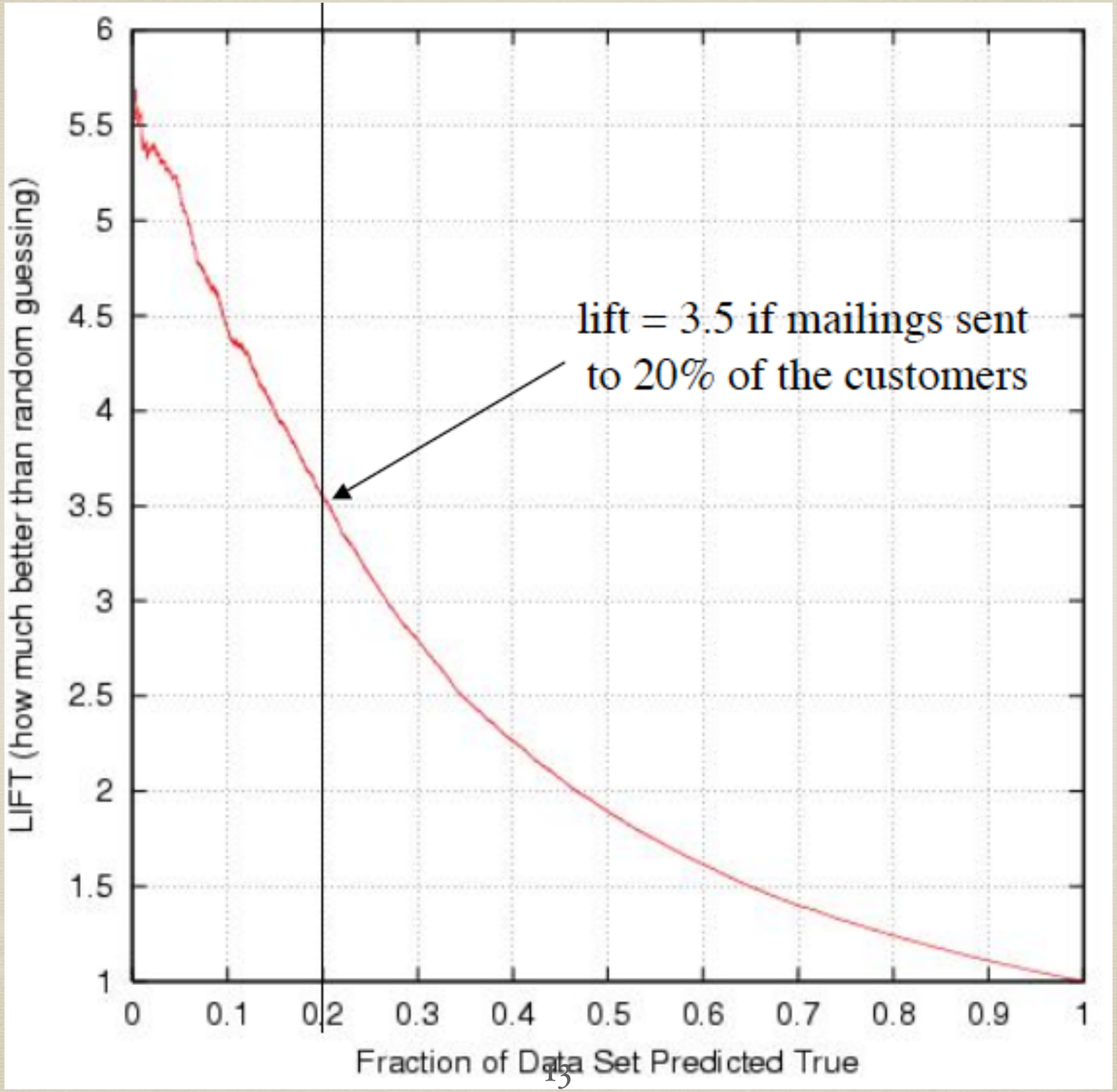
Assessment of Classification Models

* **Lift**
  * Commonly used in targeted marketing
  * Not interested in the entire population
  * How much more accurate the model is compared to random guessing when we predicted x% to be true

	Predicted Yes	**Predicted No**
True Yes	**a**	**b**
True No	**c**	**d**

* Given a threshold, **Lift** $= \dfrac{a/(a+b)}{(a+c)/(a+b+c+d)}$

Lift


lift = 3.5 if mailings sent
to 20% of the customers
LIFT (how much better than random guessing)
Fraction of Data Set Predicted True
6
5.5
5
4.5
4
3.5
3
2.5
2
1.5
1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Assessment of Classification Models

* **Precision, Recall, F-measure & Breakeven-Point**

$$\frac{2 * (precision * recall)}{precision + recall}$$

Assessment of Classification Models

* **Precision, Recall, F-measure & Breakeven-Point**

	Predicted Yes	Predicted No
True Yes	a	b
True No	c	d

$$\frac{2 * (precision * recall)}{precision + recall}$$

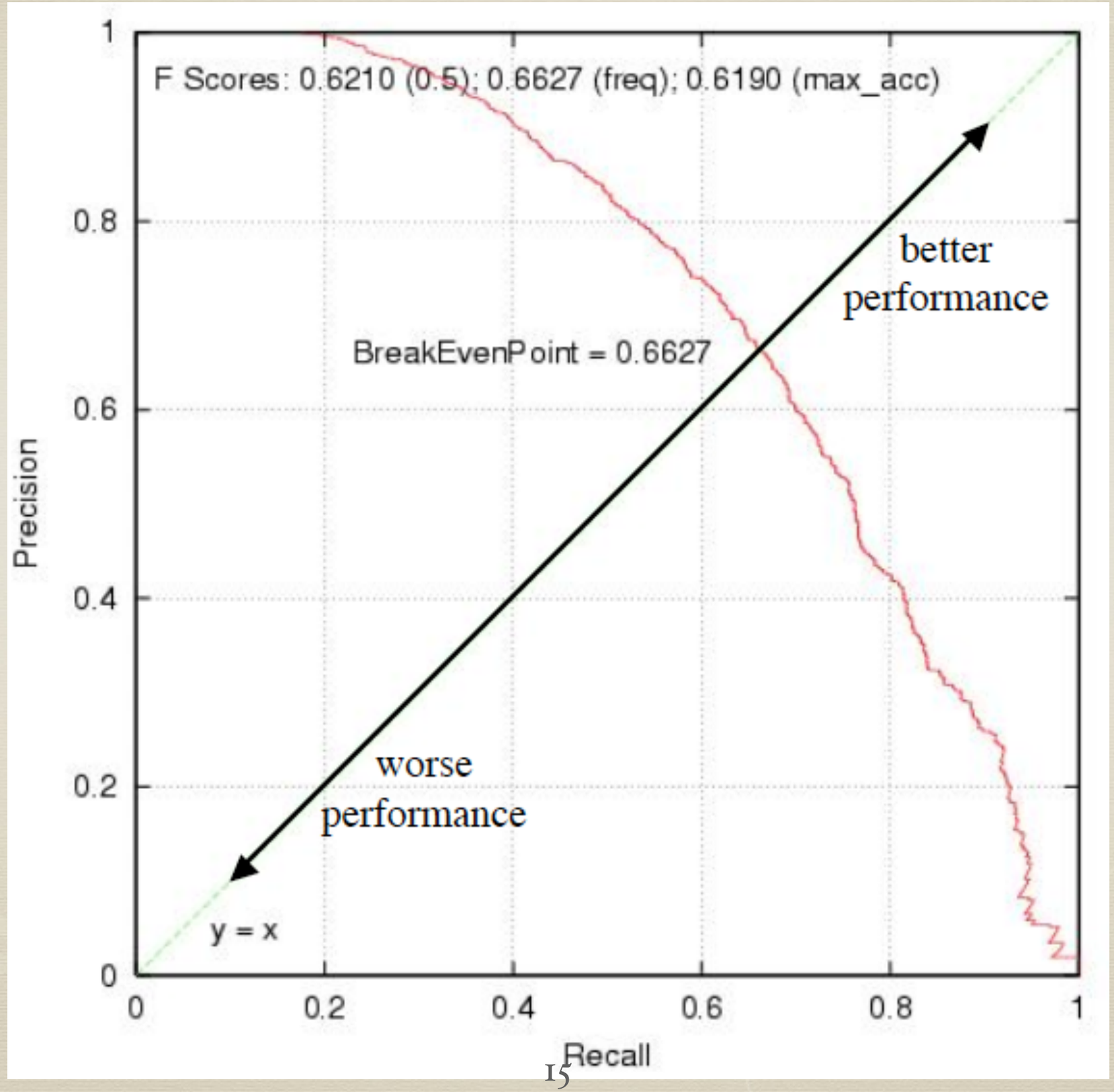
Assessment of Classification Models

* **Precision, Recall, F-measure & Breakeven-Point**

	Predicted Yes	**Predicted No**
True Yes	**a**	**b**
True No	**c**	**d**

* **Precision** = Pr(True|Predicted True) = a/(a+c)
* **Recall** = Pr(Predicted True|True) = a/(a+b)
* **F-measure** = $\frac{2*(precision*recall)}{precision+recall}$
* Breakeven Point: Precision = Recall

Precision & Recall



Assessment of Classification Models

* **Sensitivity, Specificity and ROC**

Assessment of Classification Models

* **Sensitivity, Specificity and ROC**

	Predicted Yes	Predicted No
True Yes	a	b
True No	c	d

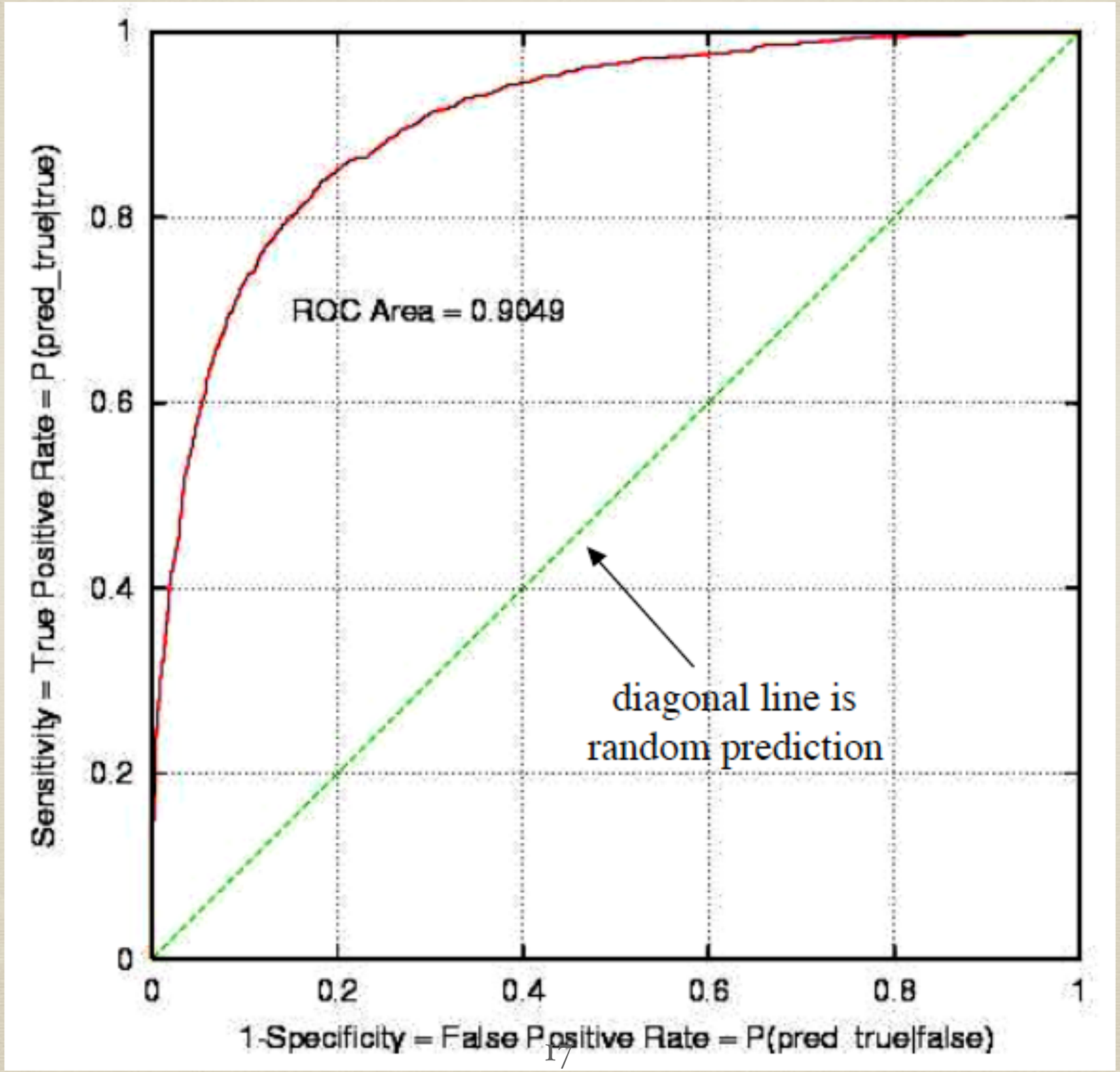
Assessment of Classification Models

* **Sensitivity, Specificity and ROC**

	Predicted Yes	**Predicted No**
True Yes	**a**	**b**
True No	**c**	**d**

* **Sensitivity** = Pr(Predicted True|True) = a/(a+b)
* **Specificity** = Pr(Predicted False|False) = d/(c+d)
* **Receiver Operator Characteristic (ROC)** Curve
  * Sensitivity vs (1-Specificity)

ROC curve


1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Sensitivity = True Positive Rate = P(pred_true|true)
1-Specificity = False Positive Rate = P(pred_true|false)
ROC Area = 0.9049
diagonal line is
random prediction

Assessment for Training

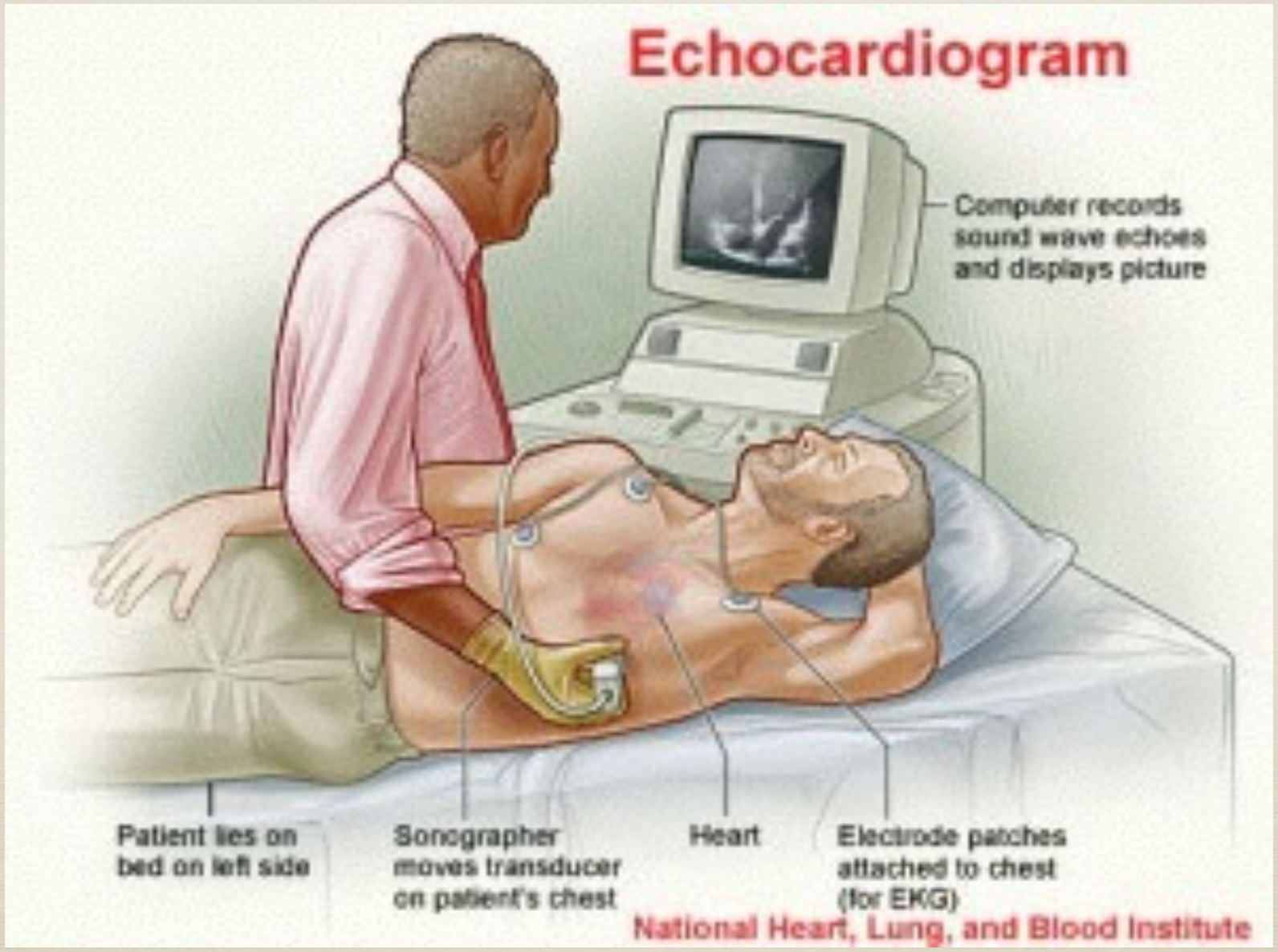
* Training and Testing Datasets
* n-Fold Cross-Validation
* Jackknife
* Bootstrapping

Applications of Predictive Models

Applications

* Outcome Predictions
  * Mortality, long-term outcome & quality of life
  * Quality assessment, decision support, cost-effectiveness
* Clinical event predictions
  * Early or timely intervention
  * Adverse effect/risk assessment
  * Treatment response
* Evidence generation: clinical studies

Case Study: Value of Echocardiogram for Critical Care Patients



The Question

* Whether Echocardiogram **independently** contribute to the improvement of critical care patients' outcomes?
    * Outcome: 28 days mortality
    * Patient cohort: MICU and SICU
* Data: the MIMIC data
* How?

Causal Inference with Predictive Model

Uni-variate Study

Co-founding Factors

* Demographics or admission info
    * Age, Gender, Weight, BMI, Service Unit, Severity at admission, Day of Admission, Hour of Admission
* Co-morbidity (chronic) conditions
    * CHF, Afib, Liver, Renal, COPD, Stroke, Cancer
* Vital Signs
    * Blood pressure, Heart Rate, Respiration Rate, Temptation, Oxygen Saturation
* Lab tests:
    * WBC, HGB, Creatinine, etc

Causal Inference with Predictive Model

Multi-variate Study

Causal Inference with Predictive Model

Propensity Score Study

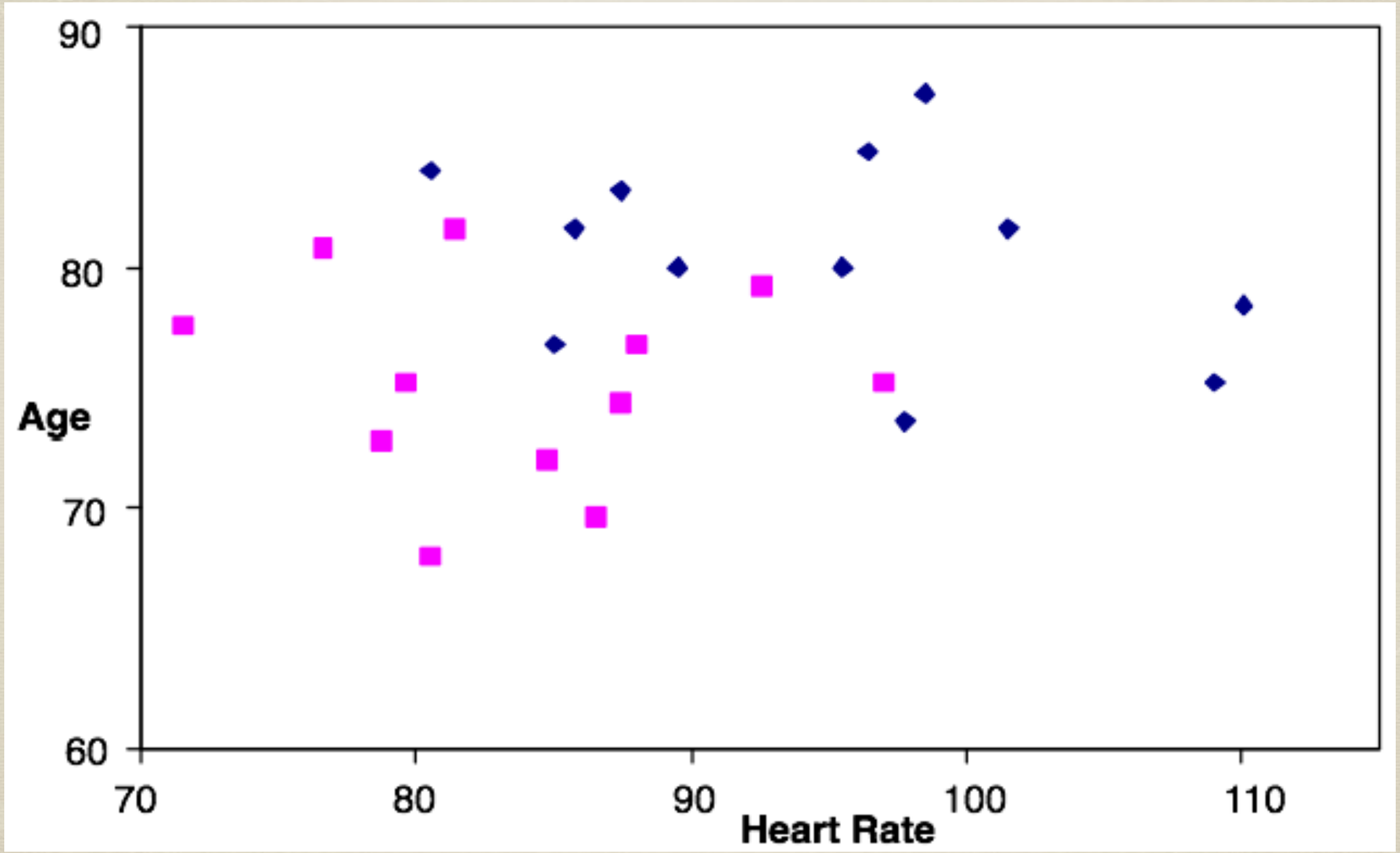
Propensity Score Based Study

* Step 1: Build the predictive model to estimate the likelihood of intervention
* Step 2: Assess the performance of the predictive model
* Step 3: Match up patients based on the predicted Propensity Score
* Step 4: Evaluate the balancing after matching
* Step 5: Compare the matched cohort

Estimation of Propensity Score with Gradient Boosting Model (GBM)

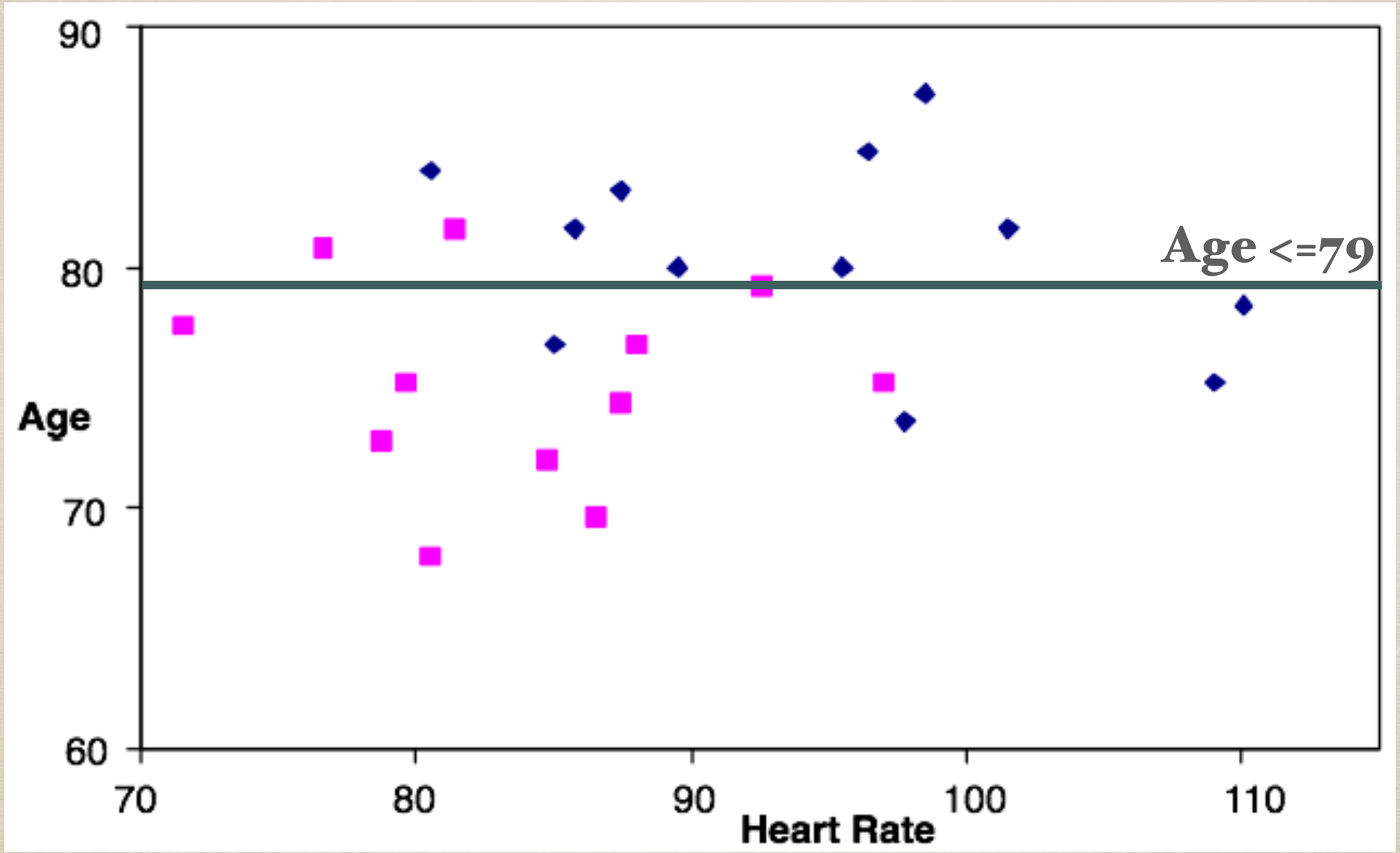
* Gradient Boosting Model (GBM)
    * An **ensemble learning** model based on **decision tress**

Decision Tress for Classification



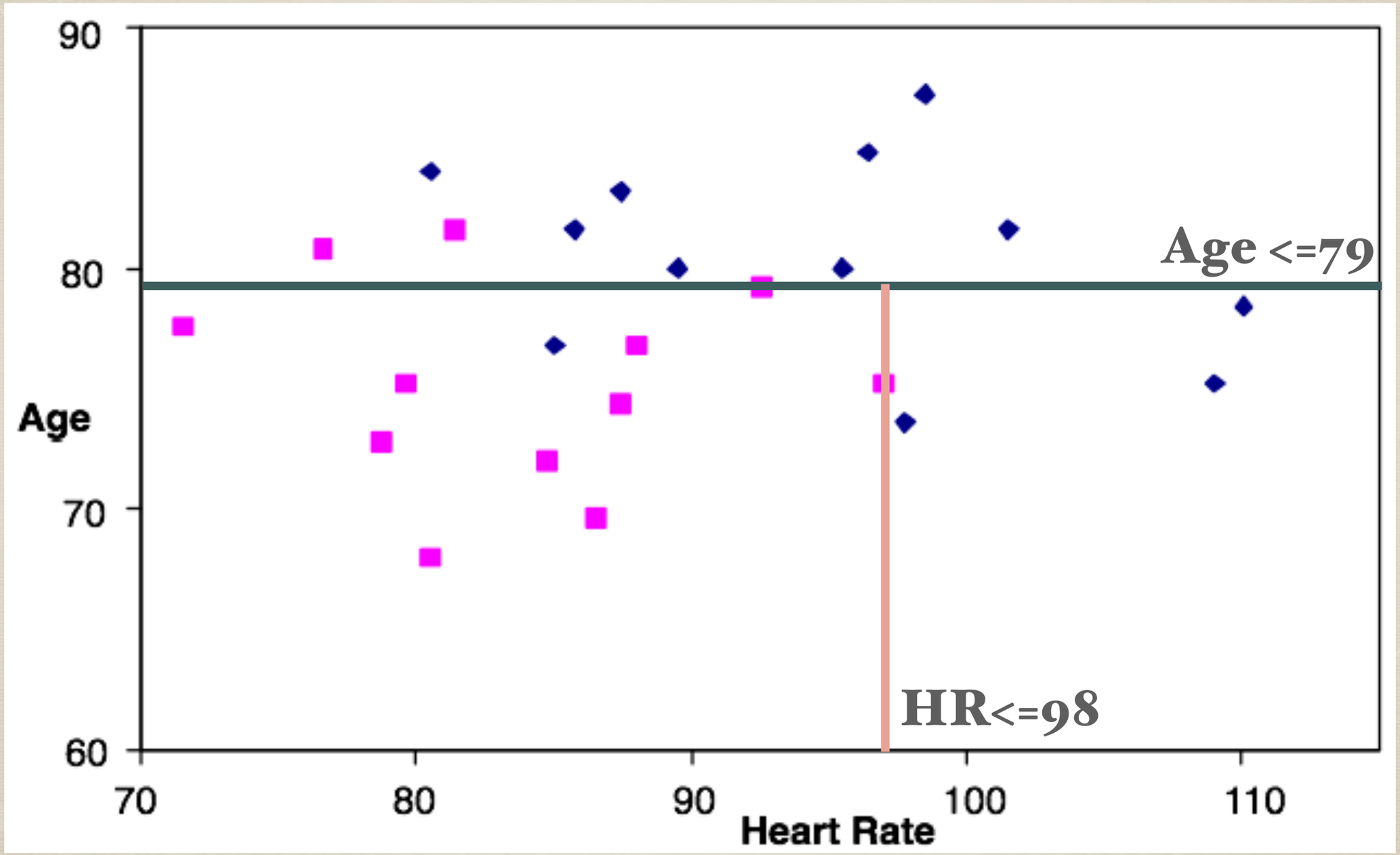
Echo Non-echo

Decision Tress for Classification



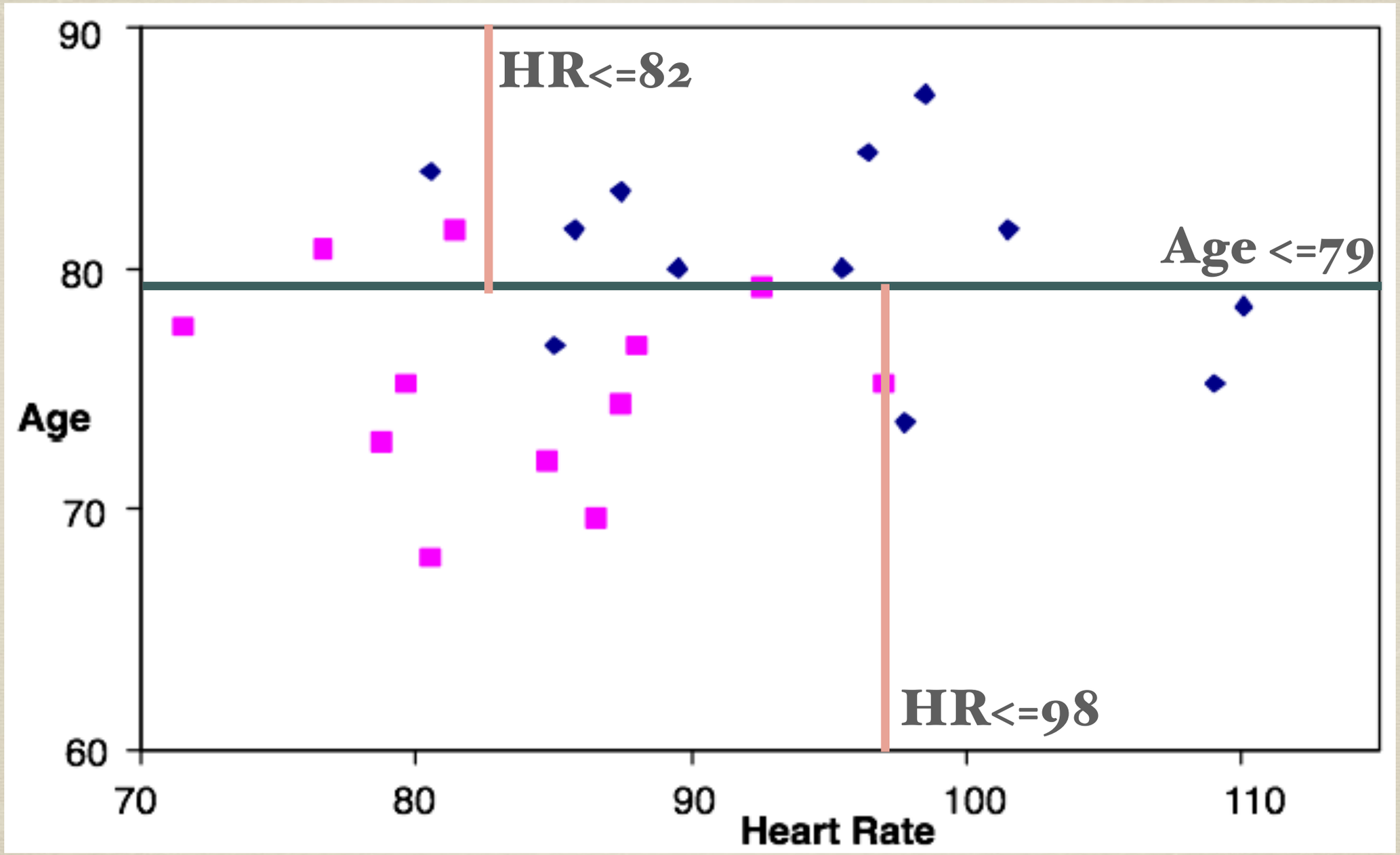
Echo Non-echo

Decision Tress for Classification



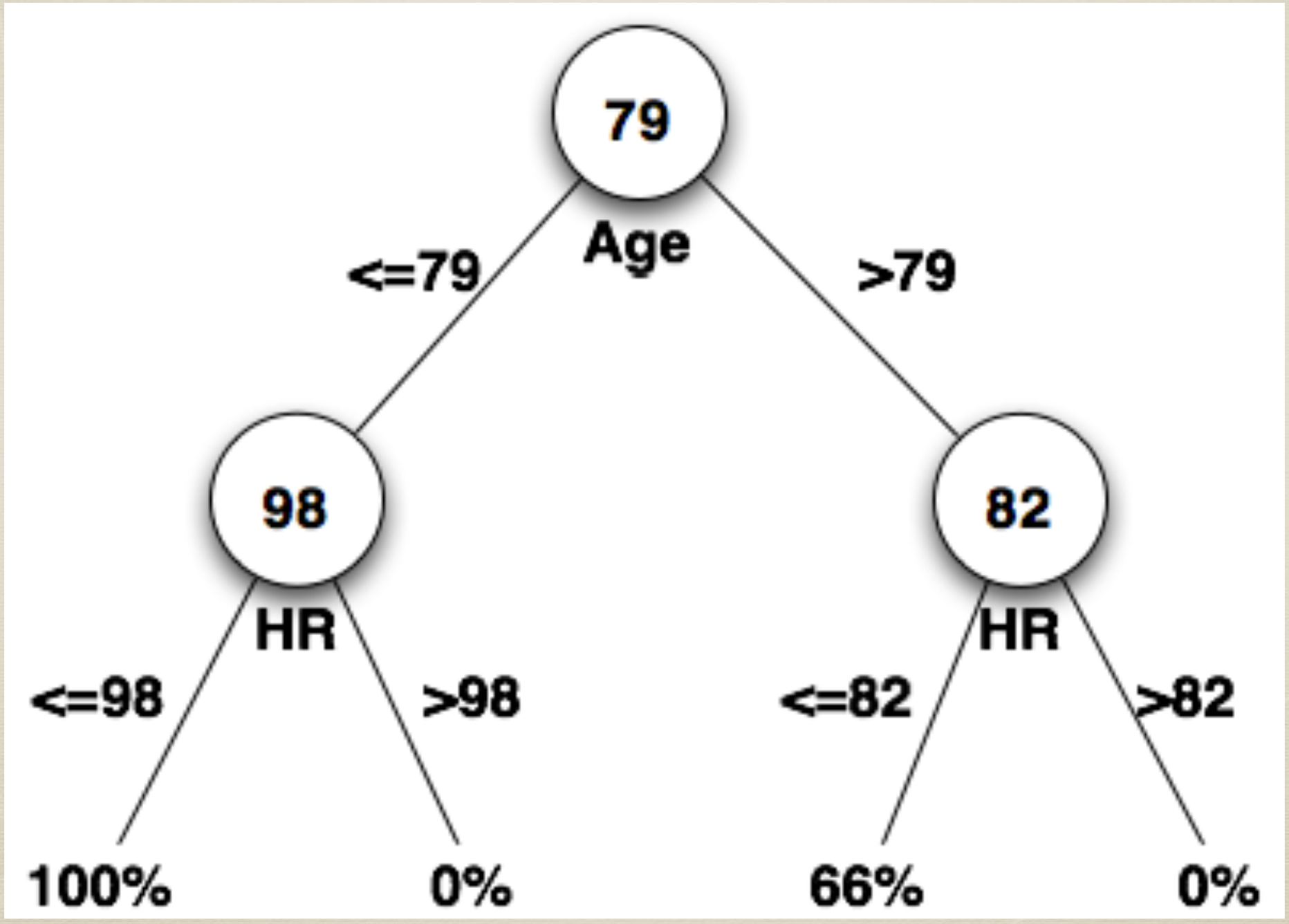
Echo Non-echo

Decision Tress for Classification



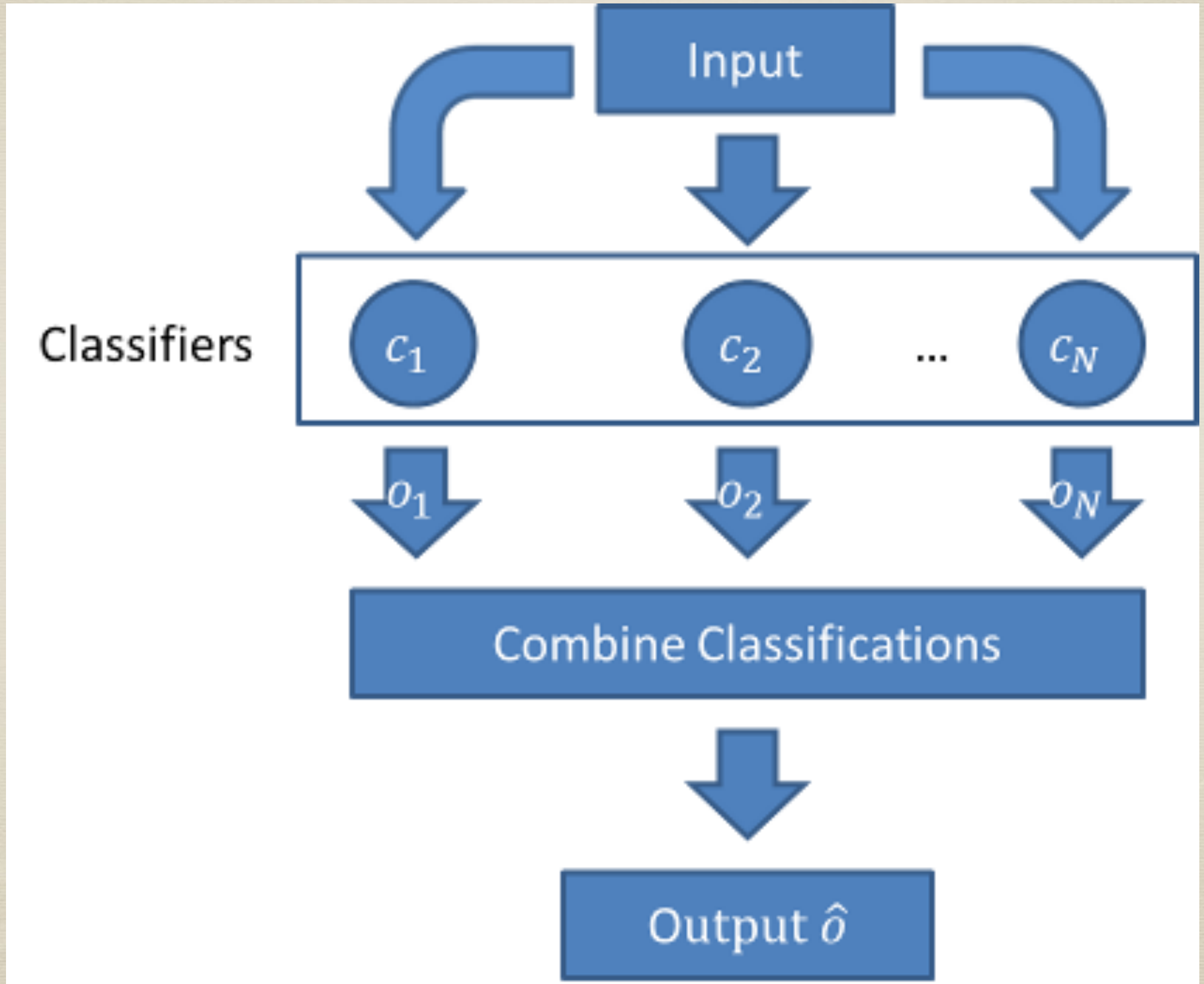
Echo Non-echo

Decision Tress for Classification



Ensemble Learning

Ensemble Learning



Propensity Score Study with Gradient Boosting Model (GBM)

What is Trending in Machine Learning

Deep Learning

What is Deep Learning?

https://www.youtube.com/watch?v=bHvf7Tagt18

What is Neural Network?

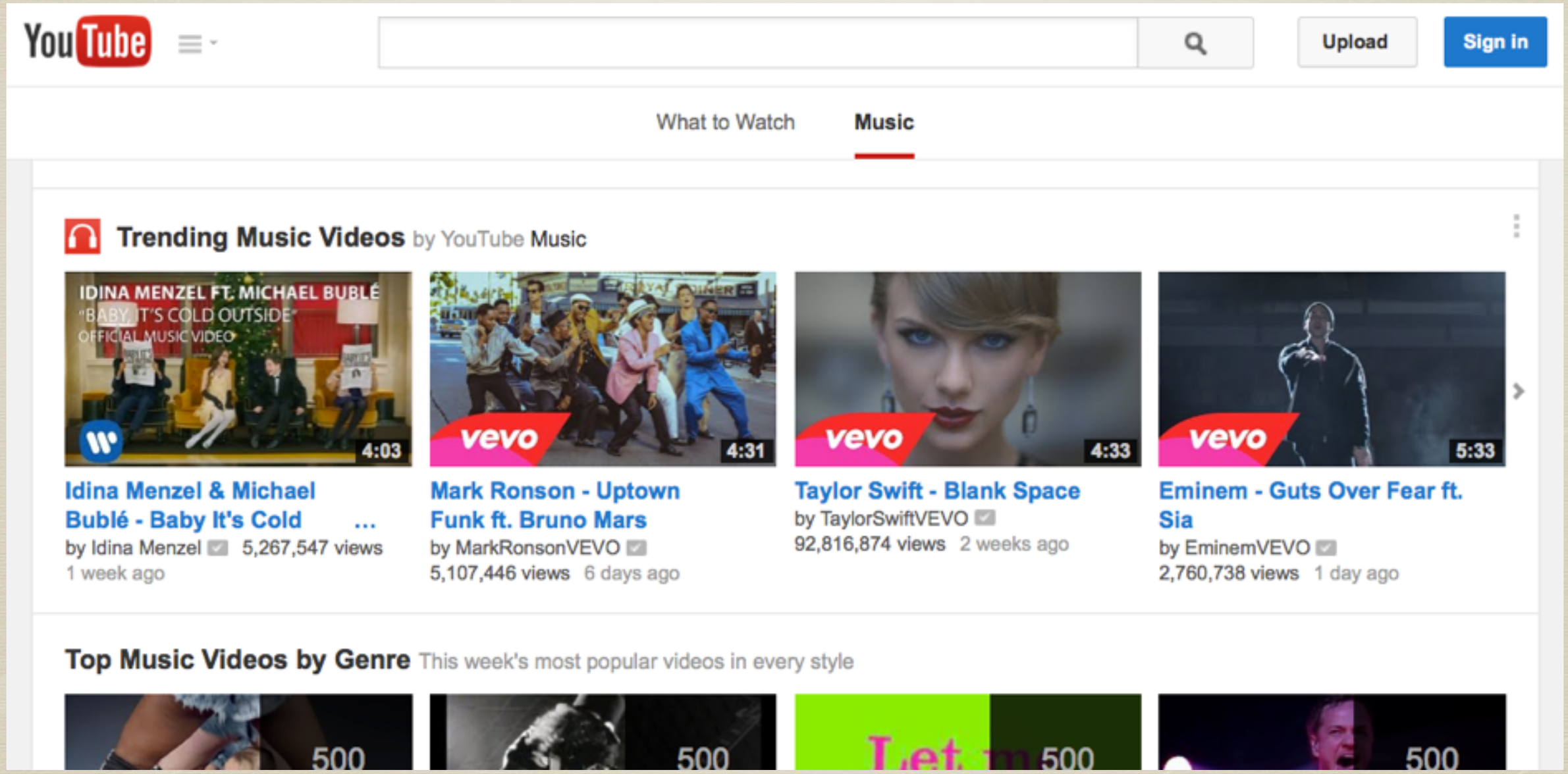
https://www.youtube.com/watch?v=DG5-UyRBQD4

Success Stories

https://www.youtube.com/watch?v=qv6UVOQ0F44

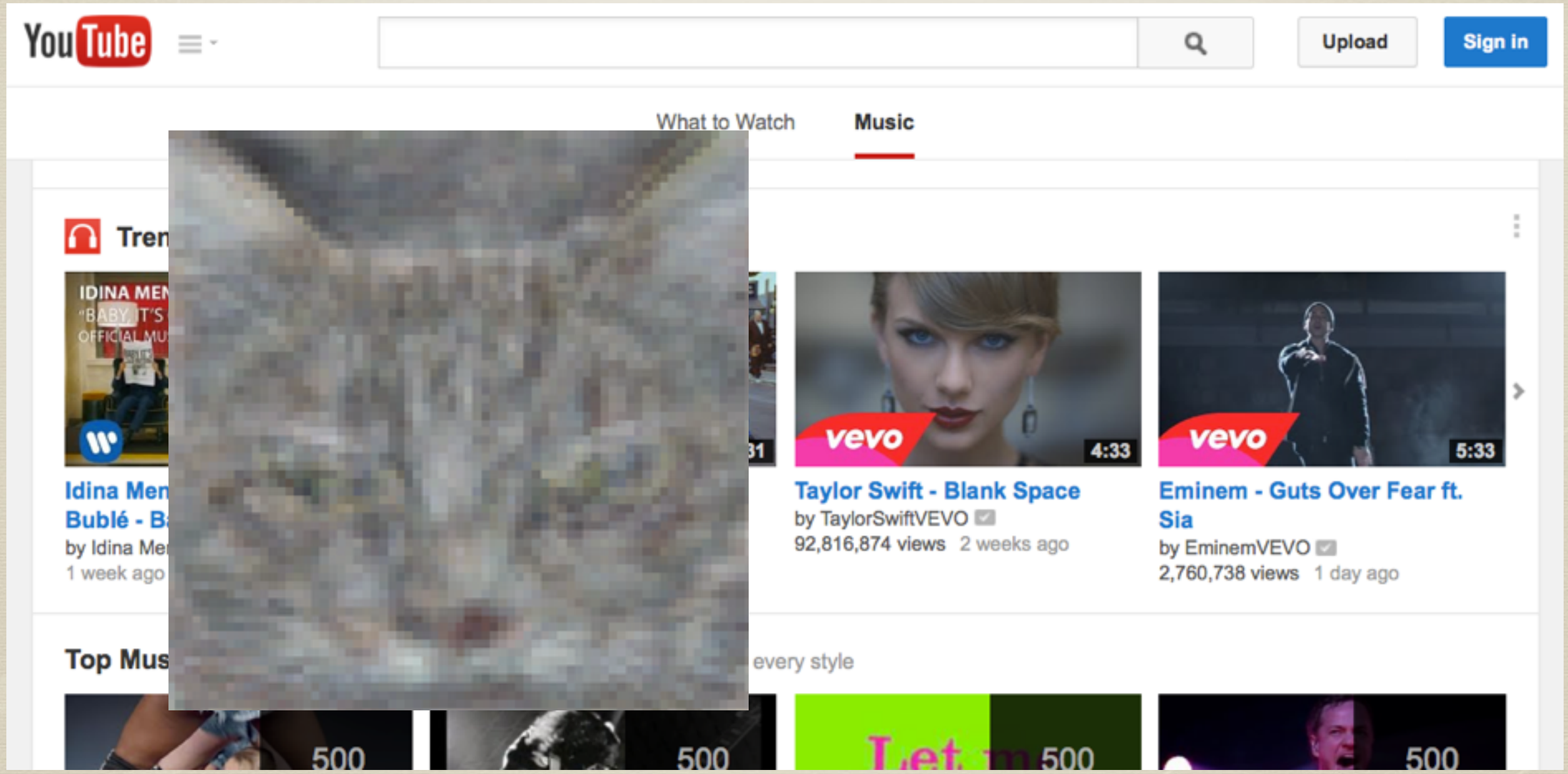
Success Stories

https://www.youtube.com/watch?v=qv6UVOQ0F44



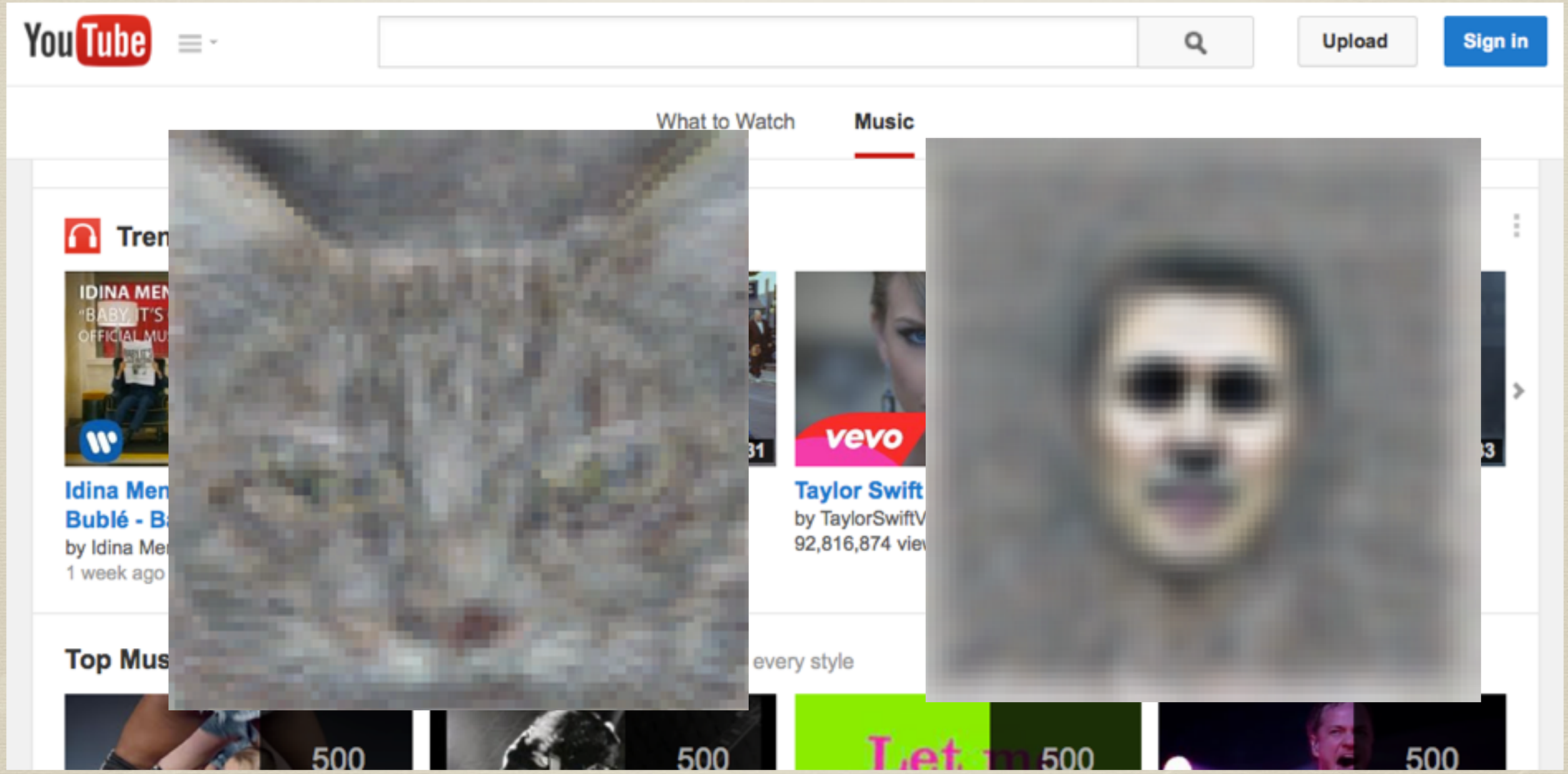
Success Stories

https://www.youtube.com/watch?v=qv6UVOQ0F44



Success Stories

https://www.youtube.com/watch?v=qv6UVOQ0F44



Success Stories

https://www.youtube.com/watch?v=qv6UVOQ0F44