

JYOTHISHMATHI INSTITUTE OF TECHNOLOGY AND SCIENCE



MACHINE LEARNING

B.PRANITHA

ASSISTANT PROFESSOR

Bayesian Theory

Prior distribution over H

Given a sample S compute a posterior distribution:

$$\Pr[h|S] = \frac{\Pr[S|h] \Pr[h]}{\Pr[S]}$$

Maximum Likelihood (ML)

$$\Pr[S|h]$$

Maximum A Posteriori (MAP)

$$\Pr[h|S]$$

Bayesian Predictor

$$\sum h(x) \Pr[h|S].$$

Some Issues in Machine Learning

- What algorithms can approximate functions well, and when?
- How does number of training examples influence accuracy?
- How does complexity of hypothesis representation impact it?
- How does noisy data influence accuracy?

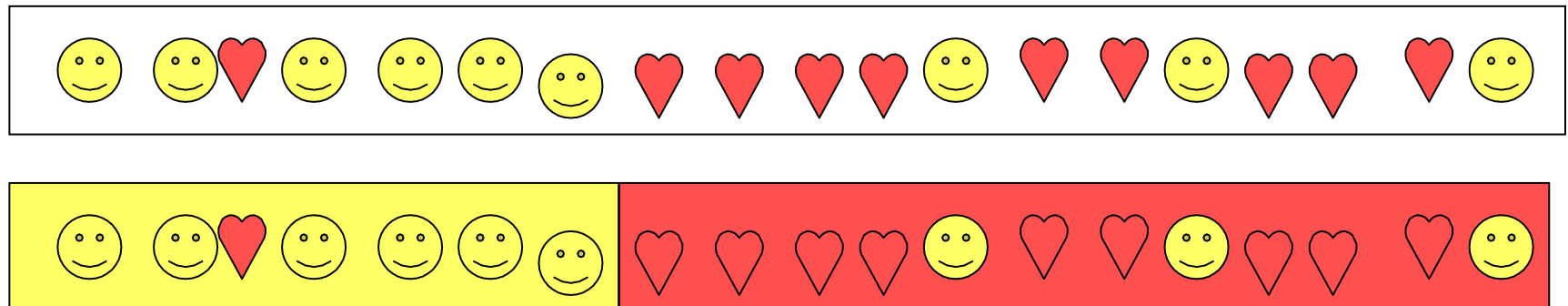
More Issues in Machine Learning

What are the theoretical limits of learnability?

- How can prior knowledge of learner help?
- What clues can we get from biological learning systems?
- How can systems alter their own representations?

Complexity vs. Generalization

- Hypothesis complexity versus observed error.
- More complex hypothesis have lower observed error on the training set,
- Might have higher true error (on test set).



Criteria for Model Selection

Minimum Description Length (MDL)

$$\varepsilon'(h) + |\text{code length of } h|$$

Structural Risk Minimization:

$$\varepsilon'(h) + \{ \log |H| / m \}^{1/2} \quad m \text{ \# of training samples}$$

- Differ in assumptions about a priori Likelihood of h
- AIC and BIC are two other theory-based model selection methods

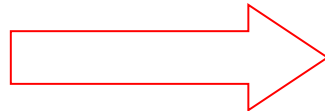
Weak Learning

Small class of predicates H

Weak Learning:

Assume that for *any* distribution D , there is some predicate $h \in H$ that predicts better than $1/2 + \epsilon$.

Multiple Weak Learning



Strong Learning

Boosting Algorithms

Functions: Weighted majority of the predicates.

Methodology:

Change the distribution to target “hard” examples.

Weight of an example is exponential in the number of incorrect classifications.

Good experimental results and efficient algorithms.

THANK YOU