

CS478: Machine Learning

Inductive Bias

Limitations of the Conjunction Bias

	sky	air	humidity	wind	water	forecast	enjoy?
1.	sunny	warm	normal	strong	warm	same	yes
2.	cloudy	warm	normal	strong	warm	same	yes
3.	rainy	warm	normal	strong	warm	same	no

What's the version space produced by examples 1 and 2?

What happens when we see the third example?

What Justifies this Inductive Leap?

- + $\langle \text{Sunny Warm Normal Strong Cool Change} \rangle$
- + $\langle \text{Sunny Warm Normal Light Warm Same} \rangle$

$S: \langle \text{Sunny Warm Normal ? ? ?} \rangle$

Why believe we can classify the unseen?

$\langle \text{Sunny Warm Normal Strong Warm Same} \rangle$

An UNBiased Learner

Idea: Choose H that expresses every teachable concept (i.e., H is the power set of X)

Consider $H' =$ disjunctions, conjunctions, negations over previous H . E.g.,

$\langle \text{Sunny Warm Normal ? ? ?} \rangle \vee \neg \langle \text{? ? ? ? ? Change} \rangle$

What are S, G in this case?

S

G

A Biased Hypothesis Space

- Consider a hypothesis space that can represent every possible *hypothesis*.
- Such a hypothesis space must be able to represent every subset of instances. If there are x possible instances, then there are $2^{|x|}$ subsets!

For example, there are 96 possible instances for the EnjoySport concept. So there are 2^{96} , or approximately 10^{28} , distinct hypotheses (concepts)!

- The hypothesis space consisting only of conjunctive concepts contains only 973 distinct hypotheses.

CS478: Machine Learning

Inductive Bias

Sample Inductive Biases

- **Candidate Elimination:** New instances classified only if all members of the version space agree on the classification.

Inductive Bias?

- **Rote Learner:** Learning simply involves storing each example in memory. New instances classified only if they can be found in memory.

Inductive Bias?

- **Find-S:** New instances are classified based on the most specific hypothesis returned by the algorithm. If it matches, it's a positive instance. Else it's a negative instance.

Inductive Bias?

The stronger the inductive bias, the greater the inductive leap!

A fundamental property of inductive inference:

A learner that makes no a priori assumptions regarding the identity of the target concept has no rational basis for classifying any unseen instances.

Representational assumption of candidate elimination algorithm: concept can be represented as a conjunction of attribute values.

Defining Inductive Bias

Consider

- concept learning algorithm L
- instances X , target concept c
- training examples $D_c = \{(x, c(x))\}$
- let $L(x_i, D_c)$ denote the classification assigned to the instance x_i by L after training on data D_c .

The **inductive bias** of L is any minimal set of assertions B such that for any target concept c and corresponding training examples D_c

$$(\forall x_i \in X) [(B \wedge D_c \wedge x_i) \vdash L(x_i, D_c)]$$

where $A \vdash B$ means A logically entails B