The inductive bias (also known as learning bias) of a learning algorithm is the set of assumptions that the learner uses to predict outputs of given inputs that it has not encountered.

In machine learning, one aims to construct algorithms that are able to learn to predict a certain target output. To achieve this, the learning algorithm is presented some training examples that demonstrate the intended relation of input and output values. Then the learner is supposed to approximate the correct output, even for examples that have not been shown during training. Without any additional assumptions, this problem cannot be solved since unseen situations might have an arbitrary output value. The kind of necessary assumptions about the nature of the target function are subsumed in the phrase inductive bias.

A classical example of an inductive bias is Occam's razor, assuming that the simplest consistent hypothesis about the target function is actually the best. Here consistent means that the hypothesis of the learner yields correct outputs for all of the examples that have been given to the algorithm.

Approaches to a more formal definition of inductive bias are based on mathematical logic. Here, the inductive bias is a logical formula that, together with the training data, logically entails the hypothesis generated by the learner. However, this strict formalism fails in many practical cases, where the inductive bias can only be given as a rough description (e.g. in the case of artificial neural networks), or not at all.

The following is a list of common inductive biases in machine learning algorithms.

1. Maximum conditional independence: if the hypothesis can be cast in a Bayesian framework, try to maximize conditional independence. This is the bias used in the Naive Bayes classifier.

2. Minimum cross-validation error: when trying to choose among hypotheses, select the hypothesis with the lowest cross-validation error. Although cross-validation may seem to be free of bias, the "no free lunch" theorems show that cross-validation must be biased.

3. Maximum margin: when drawing a boundary between two classes, attempt to maximize the width of the boundary. This is the bias used in support vector machines. The assumption is that distinct classes tend to be separated by wide boundaries.

4. Minimum description length: when forming a hypothesis, attempt to minimize the length of the description of the hypothesis. The assumption is that simpler hypotheses are more likely to be true. This is NOT what Occam’s razor says. Simpler models are more testable, not “more likely to be true.” See Occam's razor.

5. Minimum features: unless there is good evidence that a feature is useful, it should be deleted. This is the assumption behind feature selection algorithms.

6. Nearest neighbors: assume that most of the cases in a small neighborhood in feature space belong to the same class. Given a case for which the class is unknown, guess that it belongs to the same class as the majority in its immediate neighborhood. This is the bias used in the k-nearest neighbors algorithm. The assumption is that cases that are near each other tend to belong to the same class.