# Inductive bias, bias-variance tradeoff from Wikipedia

[*https://en.wikipedia.org/wiki/Inductive\_bias*](https://en.wikipedia.org/wiki/Inductive_bias):

The inductive bias (also known as learning bias) of a learning algorithm is the set of assumptions that the learner (model) uses to predict outputs of given inputs that it has not encountered. Inductive bias is anything which makes the algorithm learn one pattern instead of another pattern (e.g. step-functions in decision trees instead of continuous functions in linear regression models).

Learning involves searching a space of solutions for a solution that provides a good explanation of the data. However, in many cases, there may be multiple equally appropriate solutions. An inductive bias allows a learning algorithm to prioritize one solution over another, independently of the observed data.

In machine learning, the aim is 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.

[*https://en.m.wikipedia.org/wiki/Bias%E2%80%93variance\_tradeoff*](https://en.m.wikipedia.org/wiki/Bias%E2%80%93variance_tradeoff):

The bias-variance tradeoff describes the relationship between a model’s complexity, the accuracy of its predictions on training data, and how well it makes predictions on non-training data. In general, as the number of tunable parameters in a model increases, the model becomes more flexible and can better fit the training data – that is, the model has lower error or lower bias (bias of an estimator, not inductive bias). However, for more flexible models, there will tend to be greater variance to the model fit each time we take a set of samples to create a new training dataset. It is said that there is greater variance in the model’s estimated parameters.

The bias-variance problem is the conflict in trying to simultaneously minimize these two sources of error that prevent supervised learning algorithms from generalizing beyond their training set:

* The bias error is an error from erroneous assumptions (inductive bias). High bias can cause an algorithm to miss the relevant relations between features and target outputs (underfitting).
* The variance is an error from sensitivity to small fluctuations in the training set. High variance may result from an algorithm modeling the random noise in the training data (overfitting).

# My summary

* Inductive bias is a set of assumptions about the solution/target function/dataset. It constrains the solution space. Adding inductive bias means further constraining the solution space. Inductive bias is necessary for a learning algorithm to work.
* Simply choosing a model adds inductive bias. For example, logistic regression assumes that the decision boundary is linear w.r.t. the input features, while a decision tree does not assume a linear decision boundary.
* The learning algorithm has two sources of information: the model assumptions (inductive bias) and the training data.
  + When you have less data, you need more inductive bias (stronger assumptions) because the pattern will be weaker
  + When you have more data, you need less inductive bias (weaker assumptions) so your model can learn from the data
  + It’s important to choose the right inductive bias for your dataset
  + When datasets are small, your skill in making the model assumptions becomes much more important. Domain knowledge becomes more important.
* If your inductive bias is too high, your solution space is too constrained. Your model cannot learn from the training data. It will have very low variance because it is insensitive to the training data. This is underfitting. (You may say that your model has high bias, low variance.)
* If your inductive bias is too low, your solution space is not constrained enough. Your model will learn too much from the training data and will not generalize well. It will have very high variance because it is too sensitive noise/errors in the training data. This is overfitting. (You may say that your model has low bias, high variance.)
* Inductive bias is different from bias of an estimator, although both are used in discussions of bias-variance tradeoff.

Simple regression example. The training data is linear with Gaussian noise added (3 different randomly generated datasets with the same underlying slope and offset). When you choose a linear model, variance is low:

A graph with colored lines and dots

AI-generated content may be incorrect.

When you choose an 8th-order polynomial, variance is high:

A graph with lines and dots

AI-generated content may be incorrect.

# Inductive bias in different ML models (TBD)

Models with strong inductive bias make strong assumptions about the data and how it’s structured, require less data to train, and are less prone to overfitting.

* Linear regression: assumes a linear relationship between features and target variable
* Decision trees: assumes that data can be split into hierarchical, tree-like structures
* CNNs: assumes that nearby pixels are related and that features can be extracted hierarchically

Pros: high data efficiency, good generalization on limited data, less prone to overfitting

Cons: may underfit if the data doesn’t conform to the assumed structure

Models with weak inductive bias make fewer assumptions about the data and allow for more complex relationships to be learned.

* Transformers: lack strong build-in assumptions about data structure, making them highly flexible but requiring large amounts of data to train effectively
* Fully-connected neural networks: do not inherently prefer any specific structure, making them prone to overfitting if not regularized

Pros: high flexibility, can potentially model complex relationships

Cons: require significantly more data, prone to overfitting, may struggle on limited data

# References

* <https://medium.com/data-science/the-inductive-bias-of-ml-models-and-why-you-should-care-about-it-979fe02a1a56>
* CS229 notes
* <https://en.wikipedia.org/wiki/Inductive_bias>
* <https://en.wikipedia.org/wiki/Inductive_reasoning>
* <https://en.m.wikipedia.org/wiki/Bias%E2%80%93variance_tradeoff>
* <https://en.m.wikipedia.org/wiki/Bias_of_an_estimator>