



**Data Science
Academy**

www.datascienceacademy.com.br

Machine Learning

Parâmetros x Hiperparâmetros

Alguns termos podem parecer confusos quando você começa a aprender Machine Learning e Inteligência Artificial. Há tantos termos a serem usados e muitos dos termos são intercambiáveis. Isto é especialmente verdadeiro se você veio de outro campo de estudo que pode usar alguns dos mesmos termos utilizados em aprendizagem de máquina, mas de forma diferente. E um bom exemplo disso são os termos "parâmetro" e "hiperparâmetro".

Não ter uma definição clara para estes termos é uma dúvida comum para iniciantes. Vamos então compreender as diferenças entre esses 2 termos.

O que é um parâmetro do modelo?

Um parâmetro do modelo é uma variável de que é interna ao modelo e cujo valor pode ser estimado a partir de dados. O parâmetro também costuma ser chamado de peso ou coeficiente. Suas principais características são:

- Eles são requeridos pelo modelo para fazer previsões.
- Os valores dos parâmetros definem a habilidade do modelo em resolver seu problema.
- Eles são estimados ou aprendidos a partir de dados.
- Muitas vezes, eles não precisam ser configurados manualmente.
- Eles geralmente são salvos como parte do modelo aprendido.

Os parâmetros são fundamentais para algoritmos de aprendizagem de máquina. Eles são parte do modelo que é construído com os dados históricos de treinamento.

Na literatura de aprendizagem de máquina, podemos pensar no modelo como a hipótese e os parâmetros como a adaptação da hipótese a um conjunto específico de dados. Muitas vezes, os parâmetros do modelo são estimados usando um algoritmo de otimização, que é um tipo de pesquisa eficiente através de possíveis valores de parâmetros. Em Estatística e Programação, o termo parâmetro pode ter definições ligeiramente diferentes:

Estatística: nas estatísticas, você pode assumir uma distribuição para uma variável, como uma distribuição Gaussiana. Dois parâmetros da distribuição Gaussiana são a média e o desvio padrão. Isso é válido no aprendizado de máquina, onde esses parâmetros podem ser estimados a partir de dados e usados como parte de um modelo preditivo.

Programação: na programação, você pode passar um parâmetro para uma função. Neste caso, um parâmetro é um argumento de função que poderia assumir um valor de um intervalo de valores. Na aprendizagem de máquina, o modelo específico que você está usando é a função e requer parâmetros para fazer uma previsão sobre novos dados.

Alguns exemplos de parâmetros do modelo incluem:

- Os pesos em uma rede neural artificial.
- Os vetores de suporte em uma máquina vetorial de suporte.
- Os coeficientes em uma regressão linear ou regressão logística.

O que é um hiperparâmetro de um modelo?

Um hiperparâmetro de um modelo é uma configuração externa ao modelo e cujo valor não pode ser estimado a partir de dados. Suas principais características são:

- Eles são frequentemente usados em processos para ajudar a estimar os parâmetros do modelo.
- Eles são frequentemente especificados pelo Cientista de Dados.
- Eles geralmente podem ser configurados usando heurísticas.
- Eles são frequentemente ajustados para um determinado problema de modelagem preditiva.

Não temos como conhecer o melhor valor para um hiperparâmetro em um determinado problema. Podemos usar regras básicas, copiar valores usados em outros problemas ou procurar o melhor valor por tentativa e erro.

Quando um algoritmo de aprendizagem de máquina é ajustado para um problema específico, usando por exemplo grid search ou pesquisa aleatória, você está ajustando os hiperparâmetros do modelo para descobrir os parâmetros do modelo que resultam nas previsões mais precisas.

Os hiperparâmetros do modelo são frequentemente referidos como parâmetros do modelo o que acaba gerando confusão. Uma boa regra geral para superar essa confusão é a seguinte:

Se você precisa especificar um parâmetro do modelo manualmente, então provavelmente, este é um hiperparâmetro.

Alguns exemplos de hiperparâmetros modelo incluem:

- A taxa de aprendizado para treinar uma rede neural.
- Os hiperparâmetros C e sigma para máquinas de vetor de suporte (SVMs).
- k em k-vizinhos mais próximos.