



**Data Science
Academy**

www.datascienceacademy.com.br

Machine Learning

O Que São Métodos Ensemble?



Existe uma grande variedade de métodos para induzir modelos a partir de dados. Mas qual algoritmo irá ter o melhor desempenho para um determinado problema?

Com o objetivo de responder esta pergunta, Tom Michel, autor de um dos mais famosos livros sobre Machine Learning (que você encontra na seção de bibliografia), abordou esta questão executando um estudo com 23 algoritmos diferentes e construindo uma árvore de decisão para prever o melhor algoritmo a ser usado, dadas as propriedades de um dataset.

Embora o estudo estivesse inclinado em direção a árvores de decisão – pois 9 dos 23 algoritmos eram de árvores de decisão, o estudo revelou informações úteis, chegando a seguinte conclusão: a melhor forma de aumentar a acurácia de um modelo é juntar modelos diferentes a fim de criar um modelo final mais preciso. A isso deu-se o nome de Método Ensemble. Ensemble é uma palavra francesa que significa agrupamento.

Métodos Ensemble são uma categoria de Algoritmos de Machine Learning, que podem ser usados tanto em aprendizagem supervisionada quanto não supervisionada!

Construir um ensemble consiste em dois passos:

- 1- Construir diversos modelos.
- 2- Combinar suas estimativas.

Pode-se gerar modelos diferentes, variando por exemplo os pesos das observações, valores de dados, parâmetros, subconjuntos de variáveis ou aplicando diferentes técnicas de pré-processamento. A combinação pode ser feita por meio de votação, mas normalmente a combinação é feita por meio de pesos das estimativas dos modelos. Métodos Ensemble permitem aumentar consideravelmente o nível de precisão nas suas previsões.