

bias and variance

bias就是偏差，可以理解为预测的值的平均点与预期的偏差

Variance就是预测的值的方差，如果越分散，方差就越大

在统计学和机器学习领域，"偏差-方差权衡"（Bias-Variance Tradeoff）是一个重要概念。用中文来解释：

偏差（Bias）：指的是模型在训练数据上的预测结果与真实结果之间的差异。**高偏差通常意味着模型过于简单，无法捕捉到数据的所有特征和复杂性，这种现象被称为“欠拟合”。**

方差（Variance）：指的是模型在不同的训练数据集上的性能波动。**高方差通常意味着模型过于复杂，过度拟合了训练数据的随机噪声，而不是底层数据的真实分布。**

偏差-方差权衡是指在减少一个（比如偏差）的同时，往往会增加另一个（比如方差），反之亦然。理想的模型应该在偏差和方差之间找到平衡，以达到最佳的泛化性能。

为了**降低偏差**，你可以：

- 选择更复杂的模型。例如，使用多项式回归而不是线性回归。
- 增加模型的参数。在神经网络中，增加网络的层数或每层的神经元数量。
- 训练时间更长，确保算法充分学习数据。
- 使用更少的正则化（例如，减小正则化参数）。

为了**降低方差**，你可以：

- 收集更多的训练数据。
- 采用数据降维技术，减少特征数量。
- 增加正则化（如L1或L2正则化）。
- 在训练过程中使用技术如交叉验证。
- 减少模型的复杂度。例如，在决策树中，减少树的深度或在神经网络中减少层数。

偏差-方差权衡

通常，降低偏差会增加方差，反之亦然。**这是因为一个极端简单的模型（高偏差）很少对数据的随机噪声作出响应（低方差），而一个极端复杂的模型（低偏差）会对数据中的每一个小波动都作出反应（高方差）。**