

小批量梯度下降法就是每次取 k 个样本来计算梯度。那么 k 也是一个超参数

在机器学习领用，有很多衍生处理的算法就是在多种极端的情况取折中的做法。





