(请在列出公式时使用优美易读的方式)机器学习方法
1.讲解k近邻方法进行非参数估计的原理和步骤

k近邻（k-nearest neighbors, k-NN）是一种基于实例的非参数学习方法，用于分类和回归问题。

其原理和步骤如下：

1. 给定一个训练数据集$D={(x_1,y_1),(x_2,y_2),...,(x_n,y_n)}$，其中$x_i\in X\subseteq R^n$表示样本的特征，$y_i\in Y$表示样本的类别（或回归值）。
2. 给定一个测试样本$x$，用某种距离度量方式（如欧氏距离、曼哈顿距离等）计算该样本与训练集中所有样本的距离，并找到距离$x$最近的$k$个样本，记作$N_k(x)$。
3. 对于分类问题，使用简单多数表决的方法，将$N_k(x)$中出现次数最多的类别作为$x$的类别预测结果；对于回归问题，使用平均值或加权平均值来预测$x$的回归值。
4. k值的选择可以通过交叉验证等方法来确定，通常选择较小的k值可以增强模型的复杂度和灵敏度，但也容易受到噪声的影响。

k近邻方法的优点是简单、直观，适用于各种数据类型和分布，不需要假设数据的分布形式。缺点是需要大量的存储空间、计算时间和样本数量，并且在高维空间中会出现"维数灾难"问题，容易受到噪声和异常值的影响。