可能是最简单的机器学习算法，因此讨论的内容相对不是很多。

K最近邻(K-Nearest Neighbors, KNN)是一种基于实例的学习方法(instance-based learning)，通过与已有实例的比较来对新的实例进行分类(classification)或回归(regression)。因为基于实例的学习方法在对新实例进行预测之前不需要进行训练，所以也是一种惰性学习方法(lazy learning)。

KNN基于这样一种假设，即相似的实例在某种距离度量上应该更为接近。因此在对新实例进行分类时，新实例的分类取决于与其最接近的K个已知实例中出现次数最多的分类(mode)；在对新实例进行回归时，则可使用最接近的K个已知实例的值的平均数。

优点

* 无需训练，只需将已知实例存储起来
* 实现相对简单，例如计算实例之间的距离就比其他动辄需要求解最优化问题的算法要简单得多

缺点

* 因为需要将已知实例存储起来，所以需要占用的空间直接取决于实例的数量。**[@bislearning](http://discussions.youdaxue.com/users/bislearning)** 同学提到在项目中作为成果物提交的分类器，如果是其它方法只需提交训练得到的模型和参数，而基于实例的方法就需要把所有的实例数据都打包，因为其数据即模型
* 因为预测时需要与所有已知实例比对，所以耗时也取决于实例的数量。可以用诸如对距离进行排序再进行二分查找的方式优化
* 维度灾难(curse of dimensionality)。随着维度的增加，需要维持准确预测的已知实例的数量会随之指数级增加。因为在增加了新维度之后，已知实例在新空间的投影就变得更稀疏，从而需要比原空间中多得多的实例进行填充(**[@mrbean](http://discussions.youdaxue.com/users/mrbean)** 补充)

@ x\_man165 补充：由于KNN训练的代价小(因为不作训练)，KNN或可被用于在线学习(online machine learning)中，即使用新数据不断训练和更新已有模型从而作出更好的预测