**Web Content Mining: Classification**

kNN with inverted index

1) 在kNN算法中，对一个document分类可以搜索所有的documents，并找出其中距离最近的k个。试分析该过程的时间复杂度。

2) 如果我们使用document的tf.idf向量的余弦作为相似度，请思考是否可以使用倒排索引来查询一个document的k近邻。如果可以，请讨论该方法适用于何种情况，描述详细过程，并分析该方法和1中方法的区别。