大家好!

因为上次一个同学的项目跟我很相似,所以我的展示会有比较快。

我的项目是关于拼音到汉字的转换方法。

我选择了这个项目是因为我一个人也有这个转换的问题。很多次我听别人说话的时候,我不知道对应的汉字是什么因为我的汉语还没有那么好。所以我很好奇我们的电脑是如何解决这个问题的。

大家因该已经明白了这个问题, 所以不用说太多。

最常用的方法是马尔可夫模型(mǎěrkèfūmóxíng)。它需要三个概率:

第一个是发射(fā shè)概率。它就是从汉字到拼音的概率。

第二个是转移(zhuǎn yí)概率。它是一个汉字到下一个汉字的概率。比如说从《我》的汉字到《门》和《是》的概率不同。

最后一个是开始概率。第一个汉字没有转移概率,所以需要通过开始概率选择第一个汉字的概率。

然后为了解码最常用方法是 Viterbi 算法。

给定一个拼音序列(xù liè),比如 zhong guo ren,Viterbi 将使用我们刚刚讨论过的概率矩阵(jǔ zhèn)来找到最有可能汉字。

最后我想展示如何使用这个转换方法。

代码就是这个文件里面。

然后我来跑他一下。应该都是正确。

最后一个。

好, 谢谢大家