关于版本空间的定义

版本空间：存在的一个与训练集一致的“假设集合”，也就是可能有多个假设与训练集一致，。

版本空间(version space)是概念学习中与已知数据集一致的所有假设(hypothesis)的子集集合。

<https://blog.csdn.net/qq_18433441/article/details/55682732>

版本空间：书中定义为：与训练集**匹配**的假设空间子集称为版本空间。

搜索过程中可以不断删除与正例不一致的假设、和（或）与反例一致的假设

以上都一些书上、博客种找到的说法，其实我自己本身对这个也有些理解，但是理解的不透彻，导致了很容易忘记是一回事，另一回事是怎么都无法记住这个版本空间（version space）到底是个什么东西，直到刚刚的时候忽然间灵光一闪，忽然就意识到了，有的时候懂了之后感觉就很简单，但是不懂的时候怎么都不懂。

版本空间，在我看来的说法都是正确的，原文中“与已知数据集一致的所有假设(hypothesis)的子集集合”，但是这个说法不够完整，我在后来考虑到文中一直在讲的“泛化”问题，忽然就醒悟了，“泛化”，怎么泛化呢？我们在这个就通过对于已知的训练集的数据然后对假设空间中的进行筛选，选出那些与正例一致和反例相反的，至于为什么倾向于选择和正例相同的，我在这里的看法是：这个模型更倾向于正例的判断（即我们要辨别出好瓜，而不是辨别出坏瓜，应该瓜是无限的，不可能辨别完坏瓜之后就是好瓜了），然后我们在这里继续看待这个版本空间对于“泛化”的作用，其实泛化是什么呢？使得更有普适性，而这里其实是比较极端的，或者说比较局限的，就是把我在训练集之中已经是正例的推而广之，然后把相同的留下来，而这个留下来的必然就是训练集的子集没错。

而在这里我以自己不太确切的语言来总结这个版本空间的实质：版本空间实质就是一个伪的局限的泛化空间，挑选出和训练集相适应的正例也就是他的子集。

若有不正确之处，或者有新的理解之后会再次更新。