

1.1

见 chapter1_exercise1.ipynb

1.2

见 chapter1_exercise2.ipynb

1.3

在训练过程中选择满足最多样本的假设。也可以对每个假设，求得其准确率。准确率=(符合假设的条件且为好瓜的样例数量)/(符合假设的条件的样例数量)。选择准确率最高的假设。

另一解答：通常认为两个数据的属性越相近，则更倾向于将他们分为同一类。若相同属性出现了两种不同的分类，则认为它属于与他最临近几个数据的属性。也可以考虑同时去掉所有具有相同属性而不同分类的数据，留下的数据就是没误差的数据，但是可能会丢失部分信息。

1.4

见 exercise4.pdf

1.5

- (1)在向搜索引擎提交信息的阶段，能够从提交文本中进行信息提取，进行语义分析。
- (2)在搜索引擎进行信息匹配的阶段，能够提高问题与各个信息的匹配程度。
- (3)在向用户展示搜索结果的阶段，能够根据用户对结果感兴趣的程度进行排序。