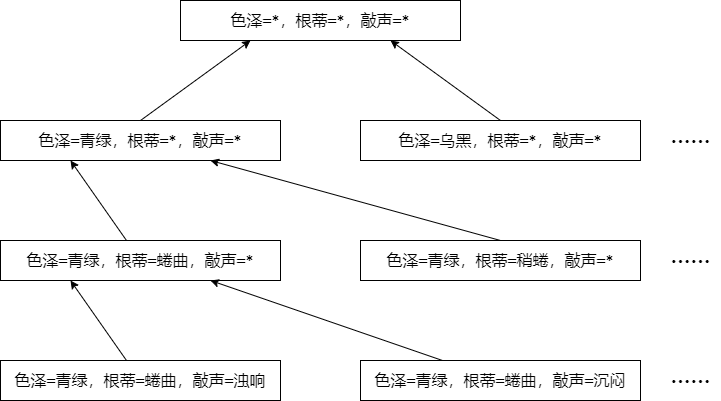
机器学习-2023春 作业1

***Steven***

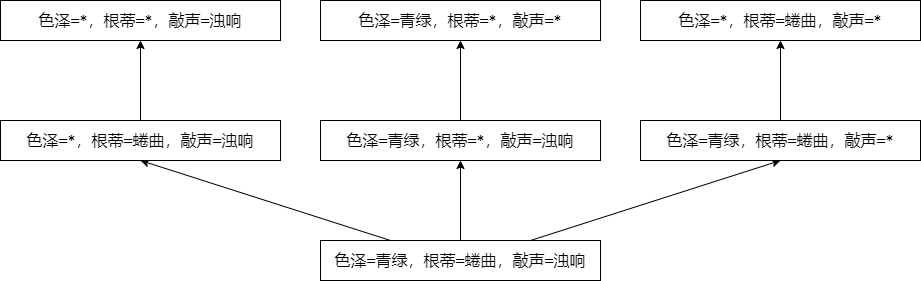
1. 若表1.1中只包含编号为1和4的两个样例，试给出相应的版本空间。



若仅包含1和4两个样例，则假设空间可表示为：



此时版本空间为：



1. 简述机器学习能在自动驾驶的哪些环节起什么作用。

自动驾驶原理通常分为四个部分：感知、定位、预测、决策，机器学习主要在预测和决策环节中发挥作用。具体来说，机器学习在训练期间对来自传感器的图像和点云数据进行建模，使得在正式道路环境中能够根据已训练模型，预测其他对象的行动，以便于提前做好准备。此外，自动驾驶还需要通过传感器时刻感知周围环境，当出现环境变化时，需要根据数据做出相应的决策，该步骤同样有赖于机器学习算法。

1. 数据集包含 1000 个样本，其中 500 个正例、 500 个反例，将其划分为包含 80% 样本的训练集和 20% 样本的测试集用于留出法评估，估算有多少种划分方式。

留出法要求测试集和训练集的类别比例相似，在本题中也就是要求训练集和测试集中都包括一半正例一半反例。由于训练集和测试集互斥，所以只需要确定了训练集，测试集便也随之确定。

训练集先从正例中抽取80%，即400个；再从反例中抽取80%，也是400个，此时测试集也随之确定，完成一次划分。故一共有种划分结果。

1. 试述真正例率（TPR）、假正例率（FPR）与查准率（P）、查全率（R）之间的联系。
2. 真正例率：所有正例中，能被预测为正例的比例。
3. 假正例率：被预测为正例中，反例的比例。
4. 查准率：被预测正例中，正例的比例。
5. 查全率：所有正例中，能被预测为正例的比例。

可见，真正例率=查全率，而假正例率与查准率互补。