Near-Optimally Teaching the Crowd to Classif 针对二元分类问题，**它将学生模拟为有限假设空间中的一个超平面。学习过程被模拟为马尔可夫链，假设学习者根据教师的反馈在假设空间中随机行走。预期错误率是选择教学集的标准。首先需要一个特征空间与假说空间**，假说空间中含有一个正确的假说（该假说对教学集中的所有图片打上正确的标签），用户自身会有个假说，每次教学时，会根据是否正确对图片进行分类，来改变是否要改变假说，最终的目的是使得用户尽可能接近正确的假说。

限制非常多：1.假设集合如何获得？ 2.论文中认为每个假设都是一个线性模型，那特征（W）该怎么获得？（实际的实验用的是合成图片或者手工制作的特征）

“Our teaching process requires a known feature space for image dataset X (i.e. the teaching set of images) and a hypothesis class H.”

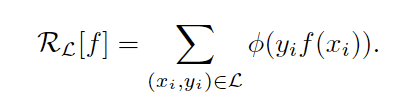
对于真实的图片集，如何获得特征空间与假说空间？

The Multidimensional Wisdom of Crowds

## MaxGrad在每次迭代中，教师都会从一个大型数据集中选择新的示例，以补充学生用于学习目标任务的小型示例集（称为教学集）。通过比较当前的学生模型和最优模型，教师可以选择最有助于学生学习的示例。教师的唯一作用就是优化学生学习的速度。

与之前算法的不同：

1. 对学生的建模：最优学生假设（最优是在机器学习意义上的定义）。student **learns the predictor that minimizes the risk**

(针对教学集的损失函数)

2.特征空间：采用CNN直接提取特征，无需手工制作