

第七讲 图像分类(5)

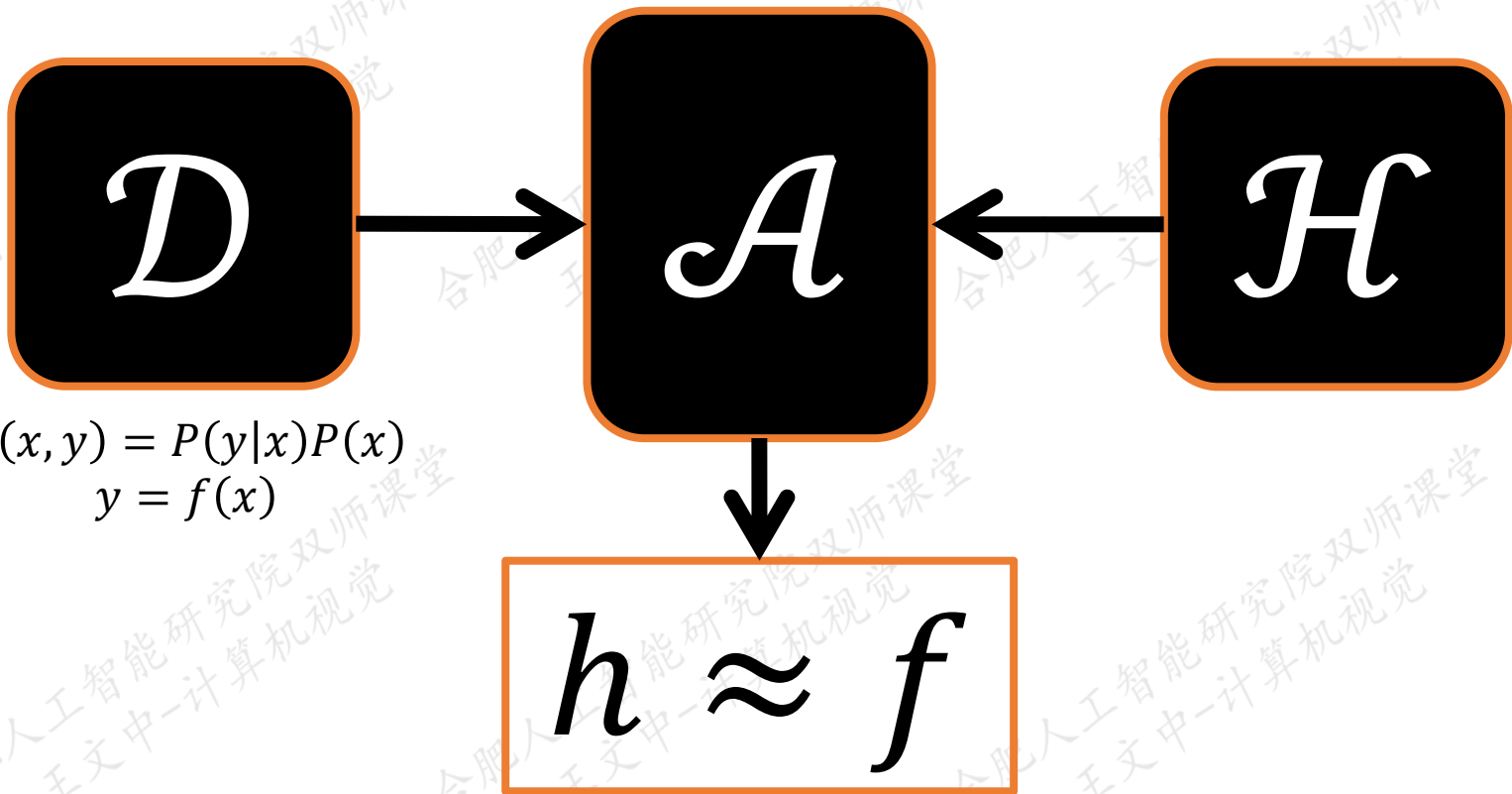
Image Classification

(from theory to practice)

王文中

安徽大学计算机学院

机器学习基本概念



为什么学习是可行的？

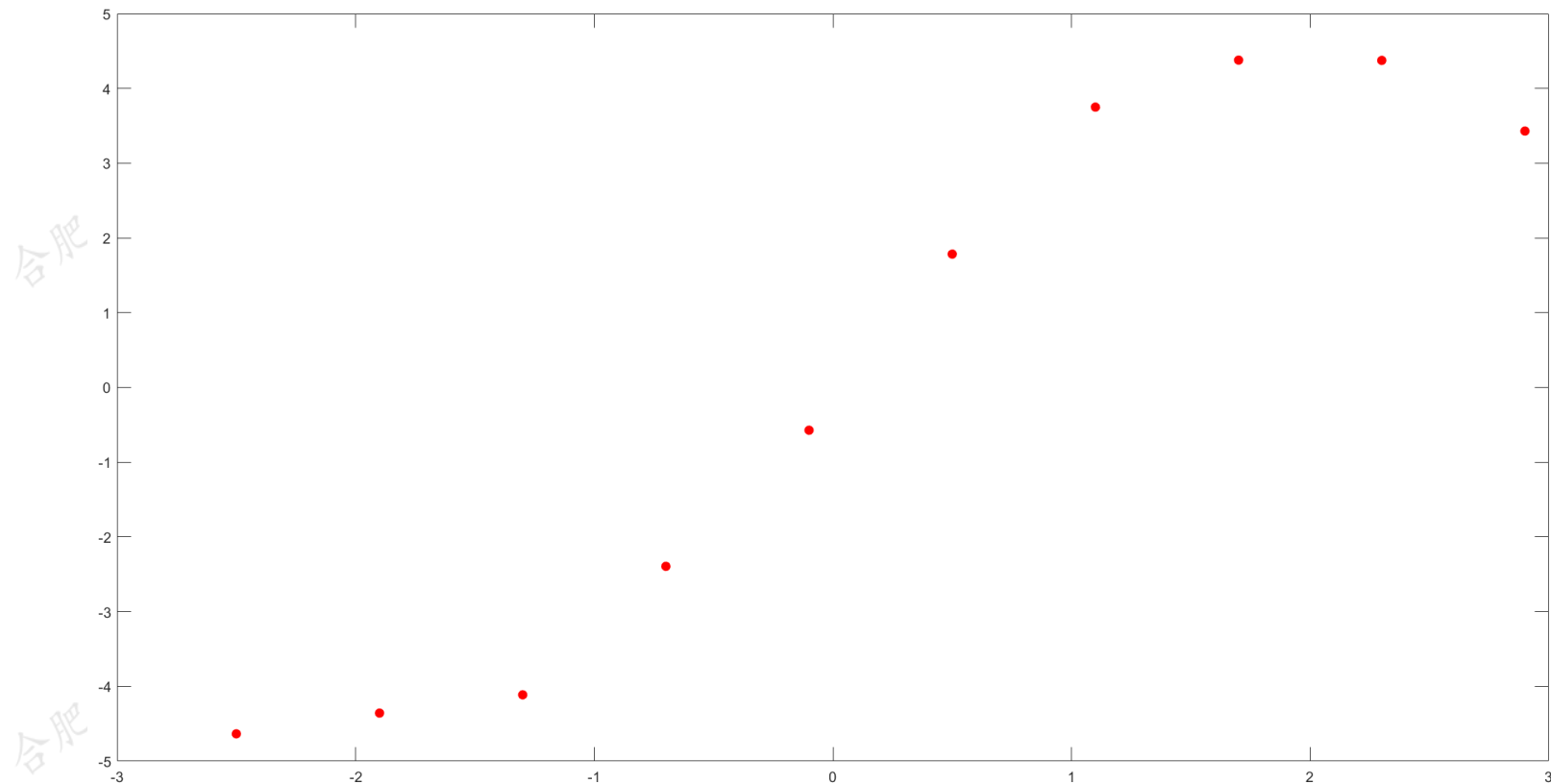
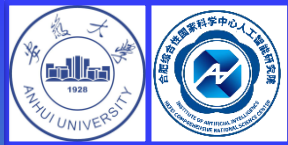


观测样本

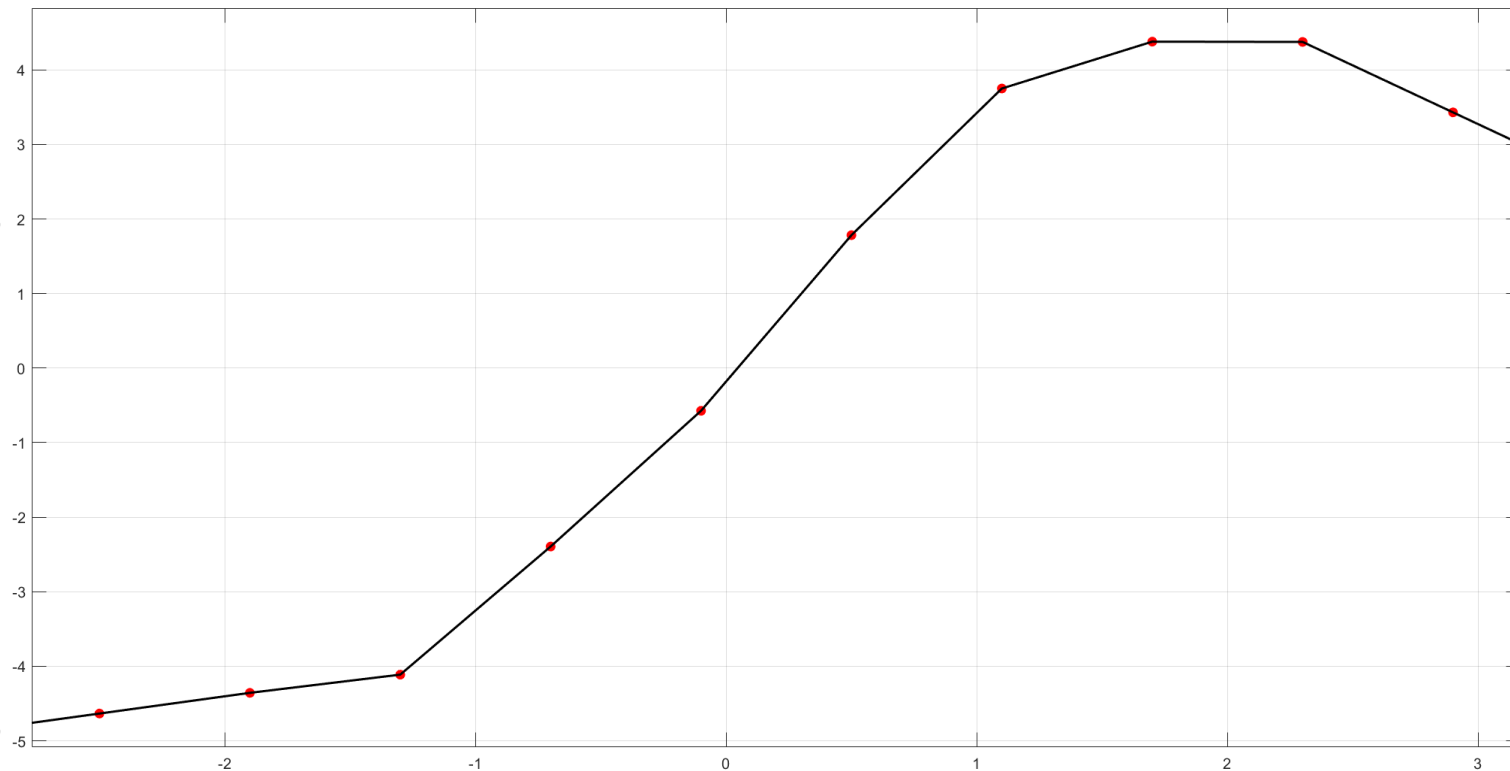
测试样本

x			y	h	f1	f2	f3	f4	f5	f6	f7	f8
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	1	1	1	1
0	1	0	1	1	1	1	1	1	1	1	1	1
0	1	1	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1		?	0	0	0	0	1	1	1	1
1	1	0		?	0	0	1	1	0	0	1	1
1	1	1		?	0	1	0	1	0	1	0	1

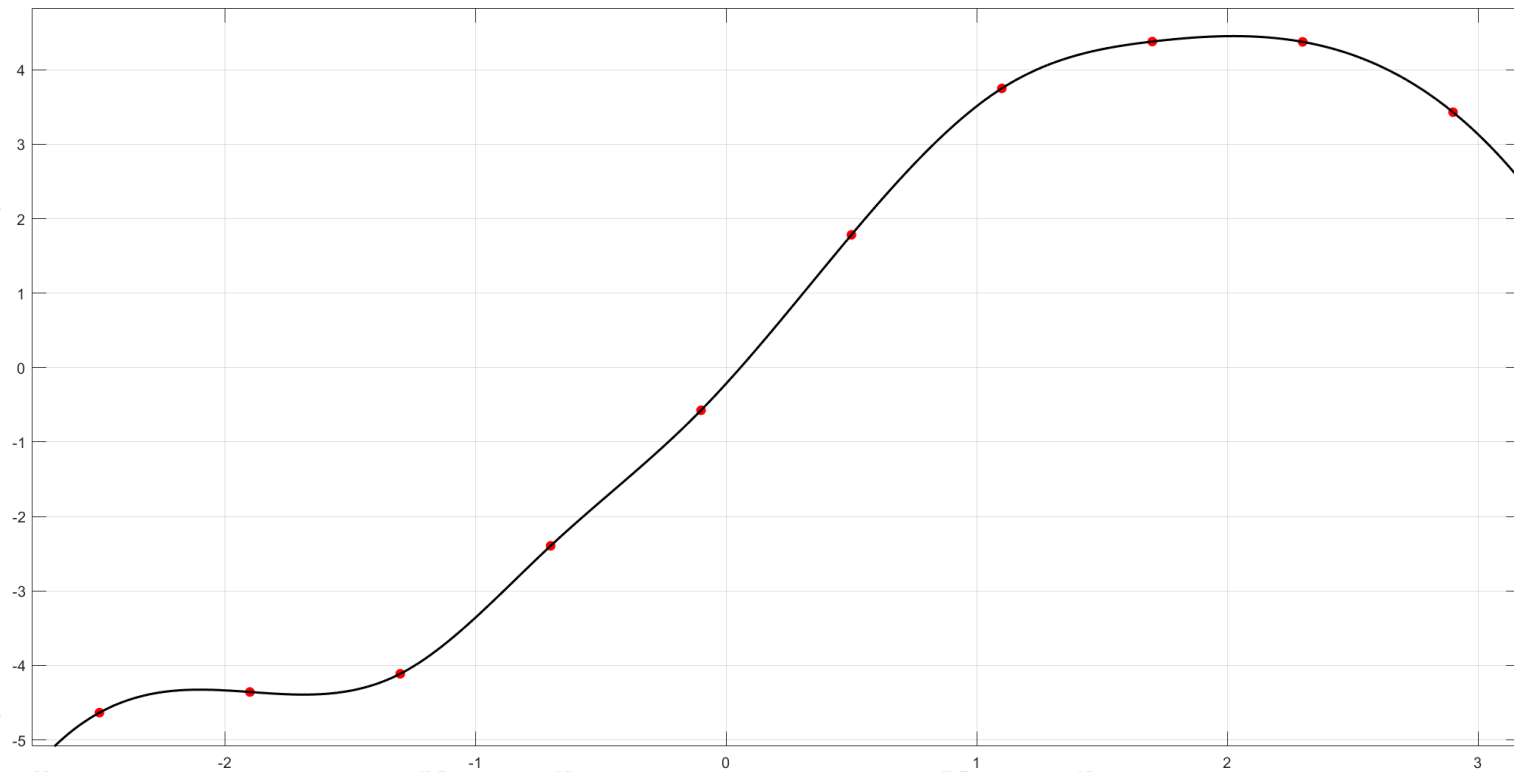
为什么学习是可行的？



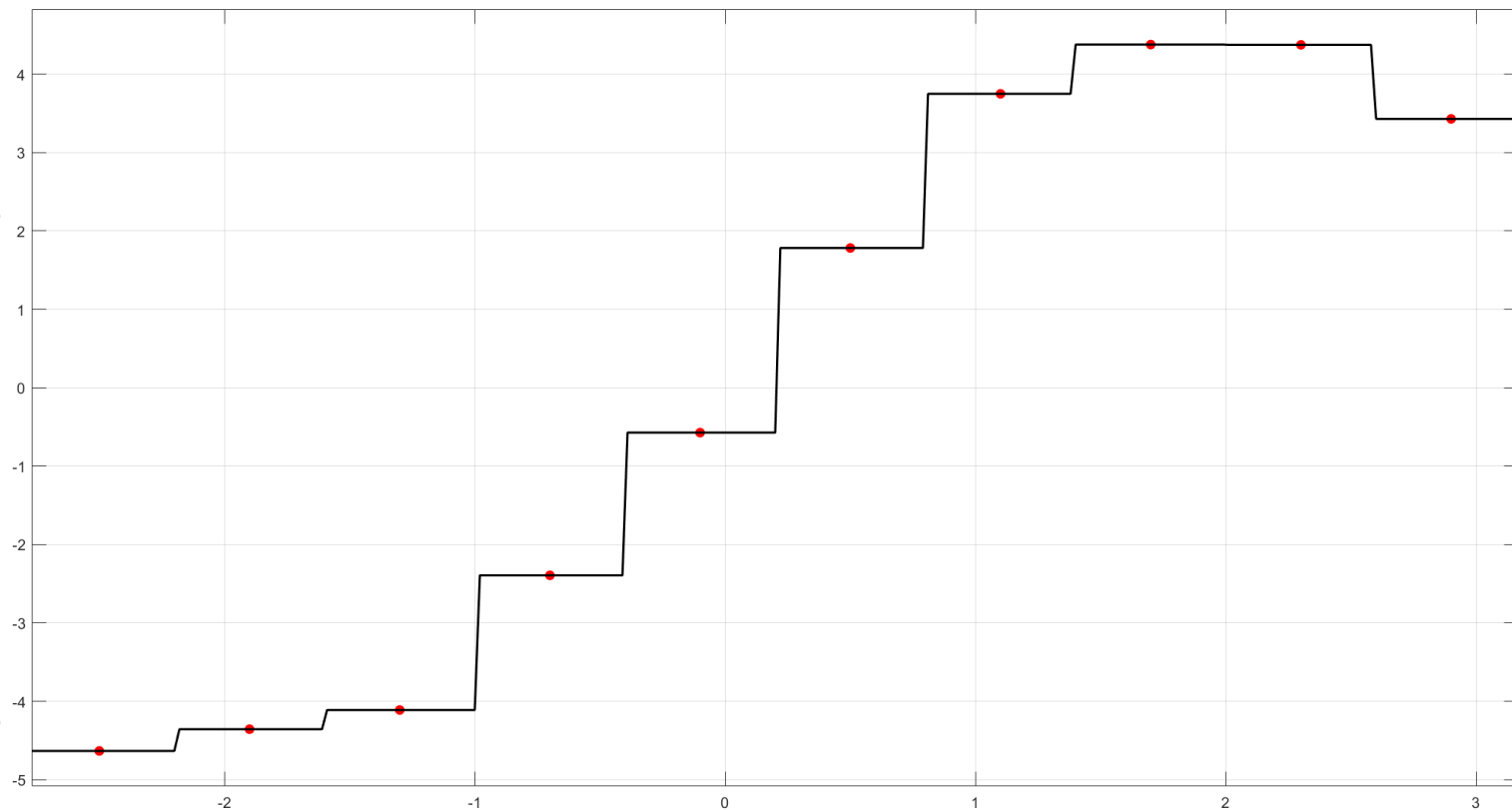
为什么学习是可行的？



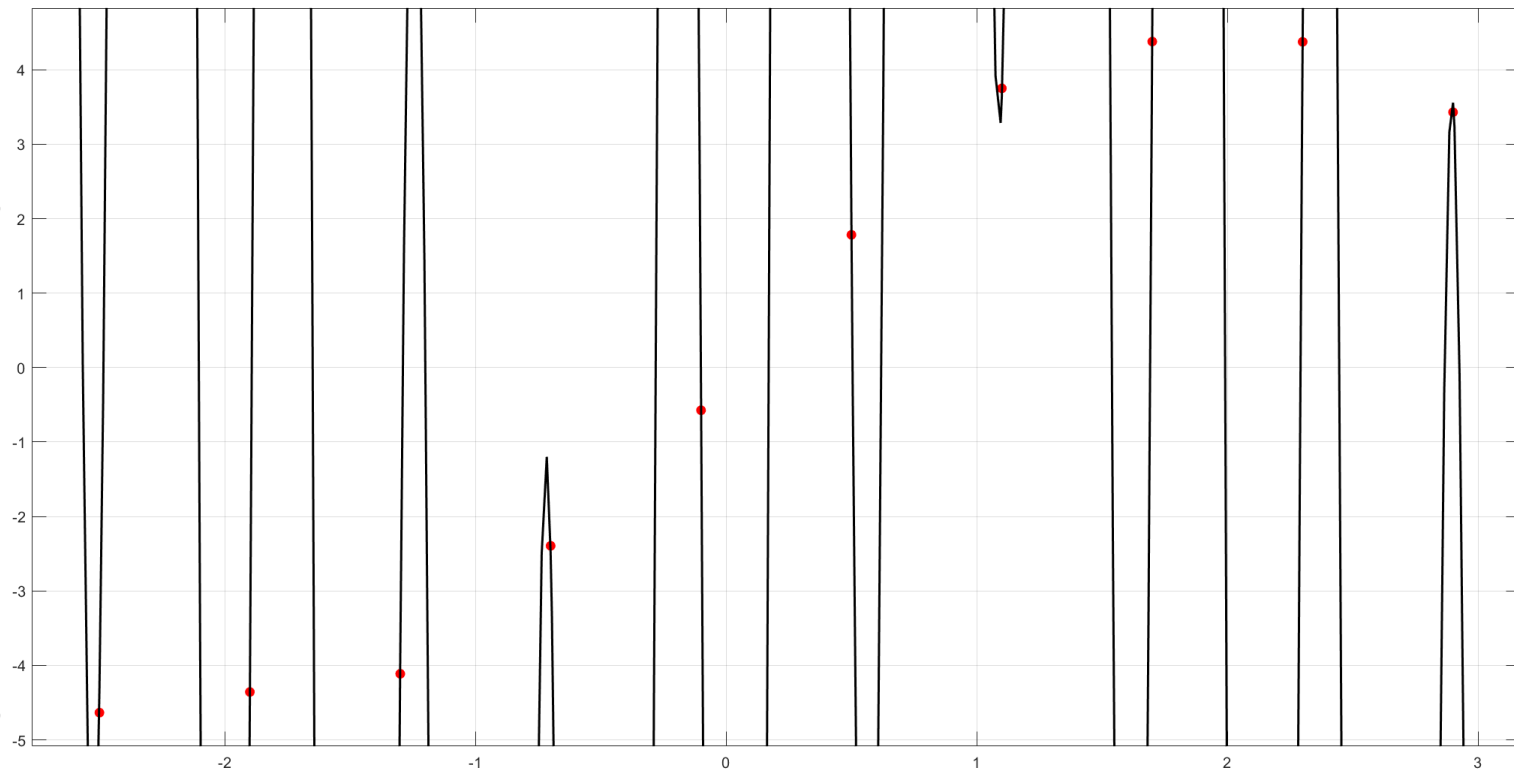
为什么学习是可行的？



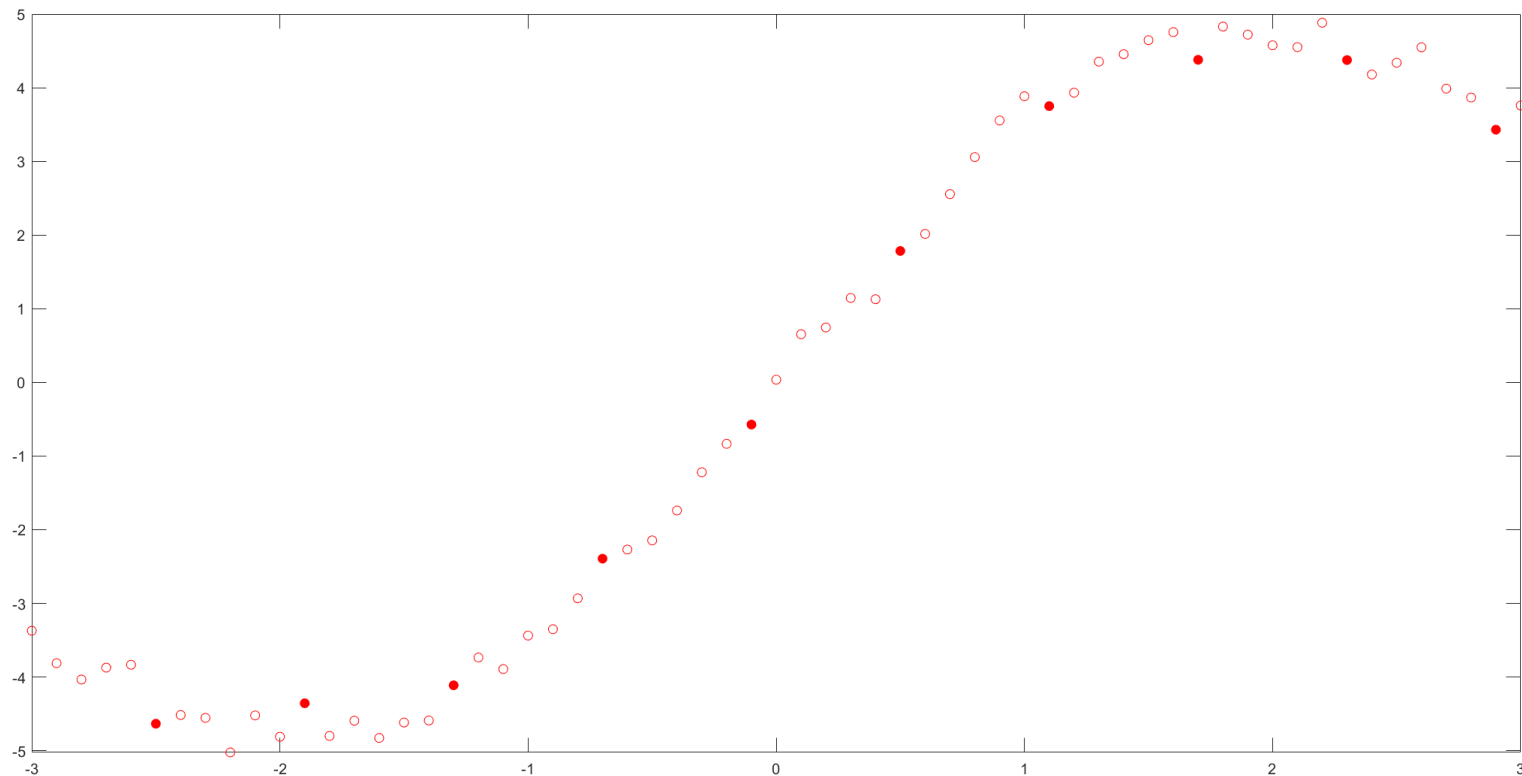
为什么学习是可行的？



为什么学习是可行的？



为什么学习是可行的？



未来观测到的数据与训练数据差不多

独立同分布假设(Independent Identical Distributed, I.I.D)



$$P(x, y) = P(y|x)P(x)$$
$$y = f(x)$$

训练/测试样本是从同一个概率分布独立采样的

训练样本



2021/8/16

Non-I.I.D

测试样本



12

训练样本



Non-I.I.D



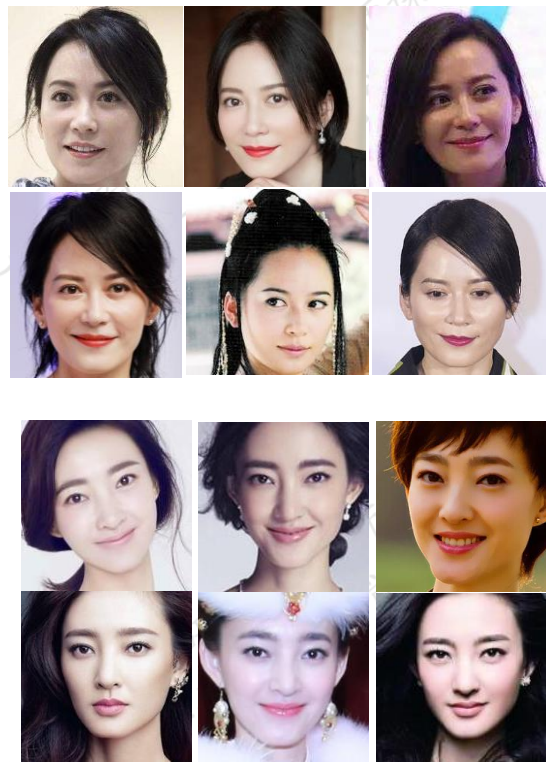
测试样本

训练样本



2021/8/16

Non-I.I.D



测试样本

训练样本



Non-I.I.D

测试样本





医院A

Non-I.I.D



医院B

独立同分布假设：分布漂移



	Training	Serving
Joint	$P_{\text{train}}(y, x)$	$P_{\text{serve}}(y, x)$
Conditional	$P_{\text{train}}(y x)$	$P_{\text{serve}}(y x)$
Marginal	$P_{\text{train}}(x)$	$P_{\text{serve}}(x)$

Dataset shift

$$P_{\text{train}}(y, x) \neq P_{\text{serve}}(y, x)$$

Covariate shift

$$P_{\text{train}}(y|x) = P_{\text{serve}}(y|x)$$

$$P_{\text{train}}(x) \neq P_{\text{serve}}(x)$$

Concept shift

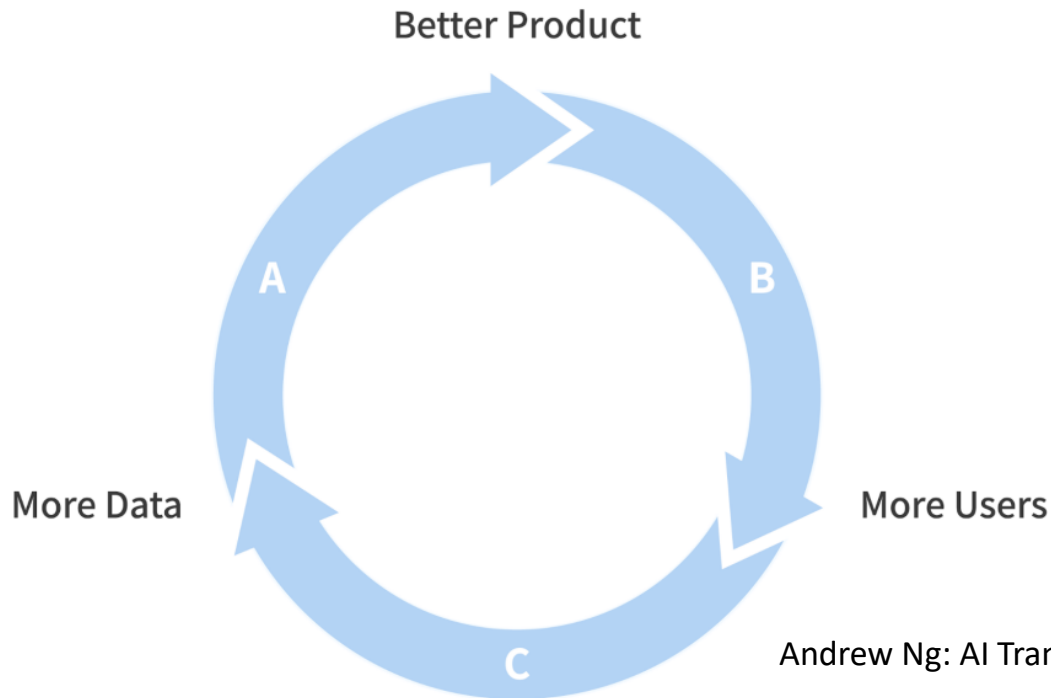
$$P_{\text{train}}(y|x) \neq P_{\text{serve}}(y|x)$$

$$P_{\text{train}}(x) = P_{\text{serve}}(x)$$

The Virtuous Cycle of AI

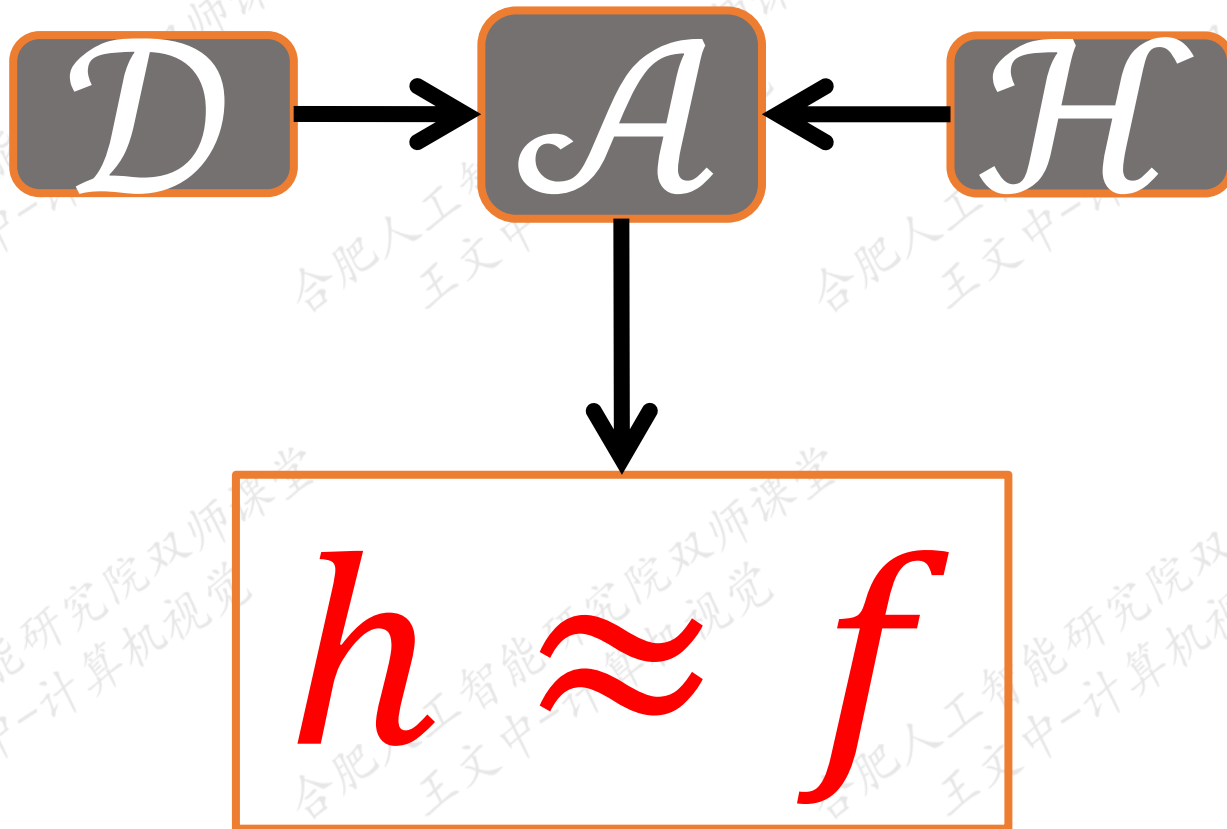


The Virtuous Cycle of AI

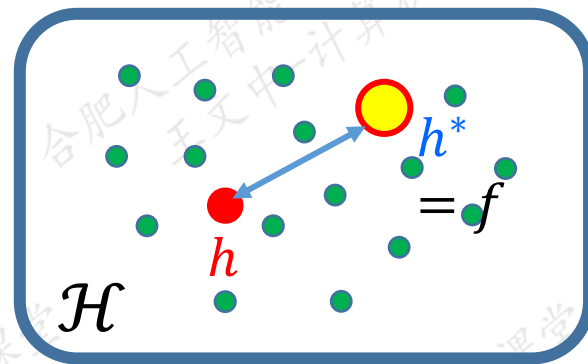
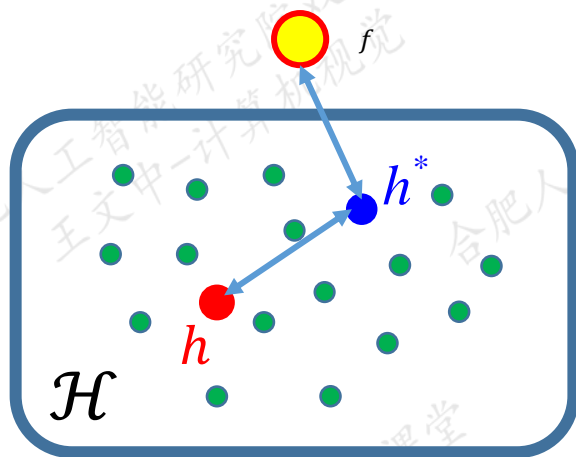


Andrew Ng: AI Transformation Playbook

泛化性能(Generalization)



泛化性能(Generalization)



h^* : \mathcal{H} 中最好的假设(真实误差最小的假设)



$$P(|h \text{ 的真实误差} - h^* \text{ 的真实误差}| > \epsilon) \leq \delta$$

$$\epsilon = O\left(\sqrt{\frac{d \times \log m - \log \delta}{m}}\right)$$

m: 训练样本的数量

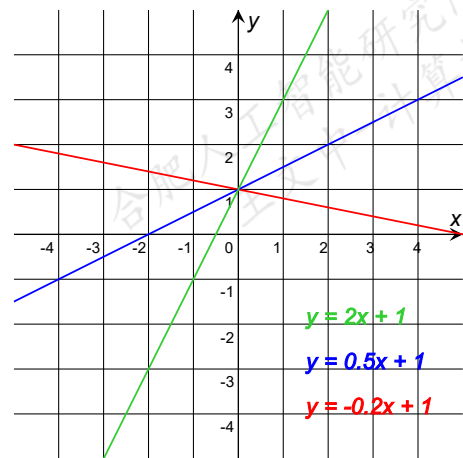
d: 假设空间的复杂程度

模型越复杂（d很大），需要的训练样本越多（m很大）

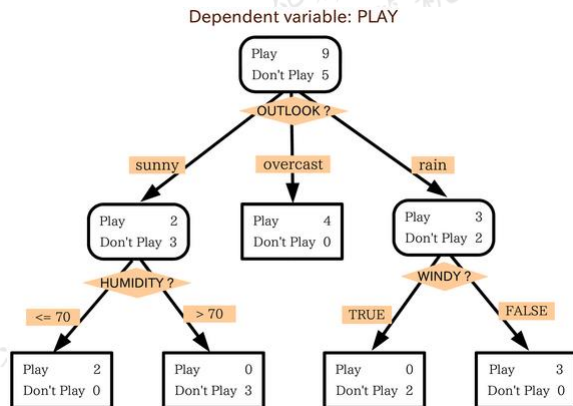
A stylized, white, cursive letter 'H' is centered within a black rounded square. The square is framed by a thick orange border. The background of the entire image is a dark, textured surface with some green and yellow elements visible at the bottom left.

$$y = h(x; \theta)$$

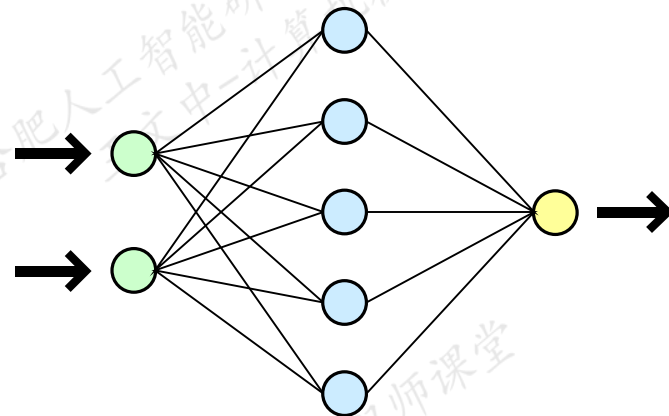
假设空间



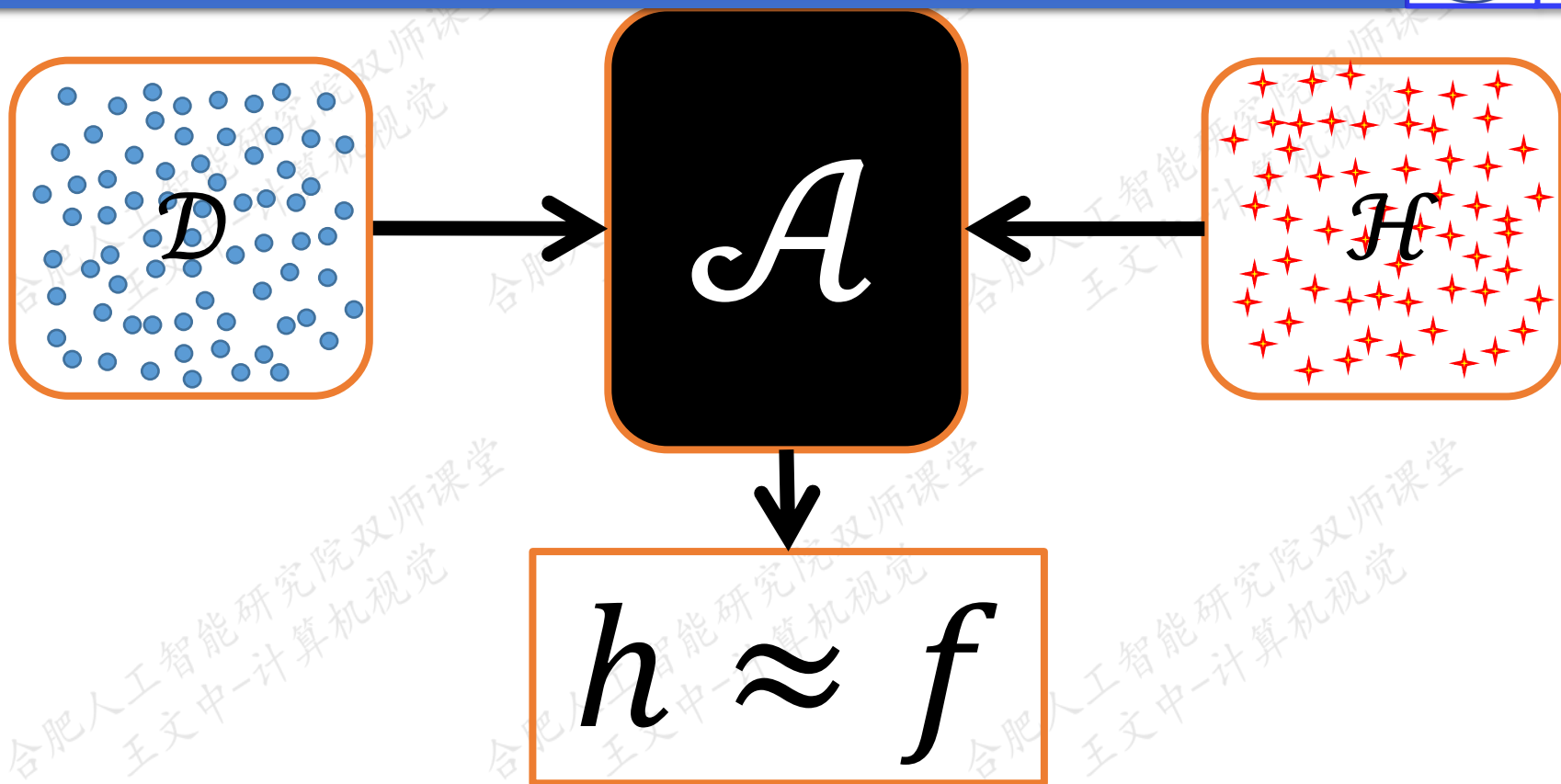
线性函数



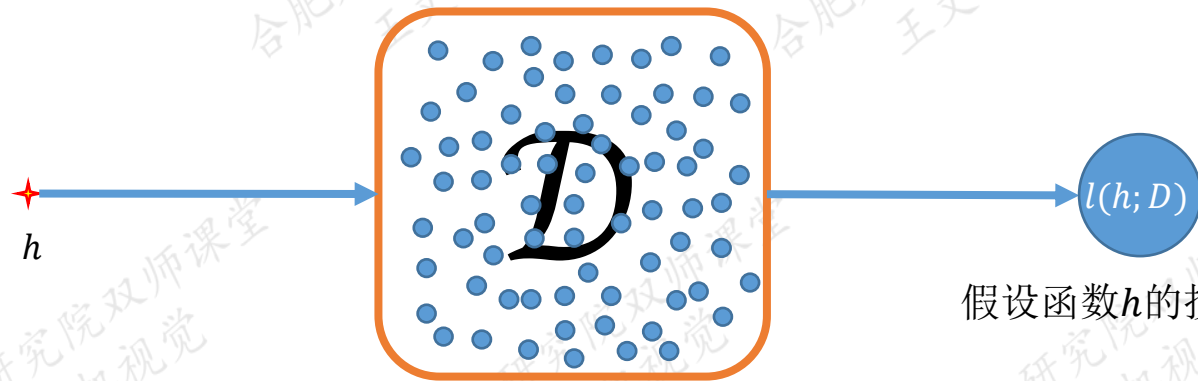
树



神经网络

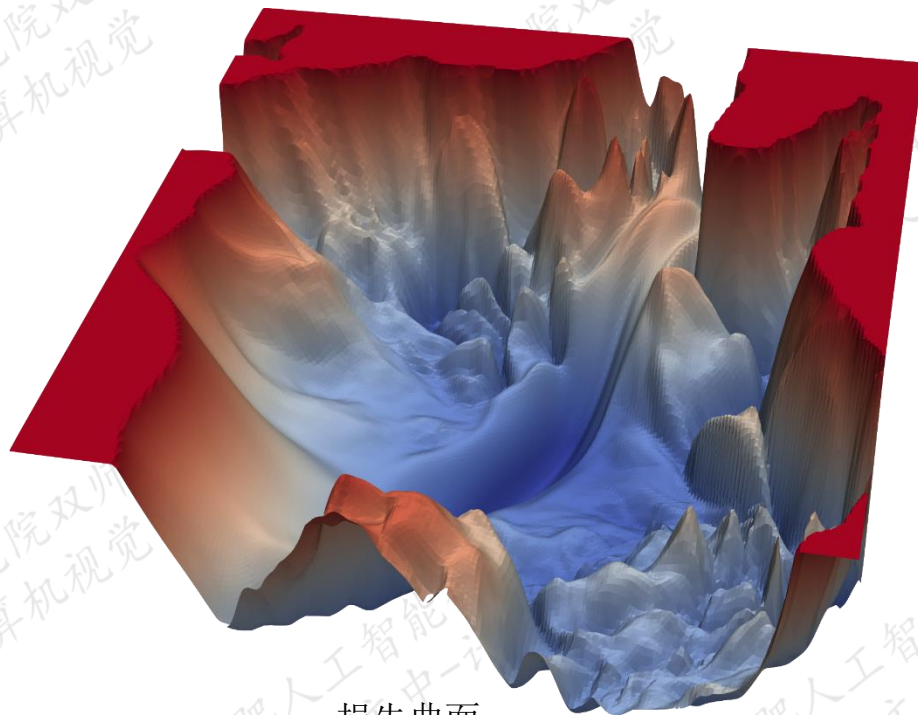


算法 \mathcal{A} : 1)利用样本集 D 评价假设 h ;



假设函数 h 的损失(风险)

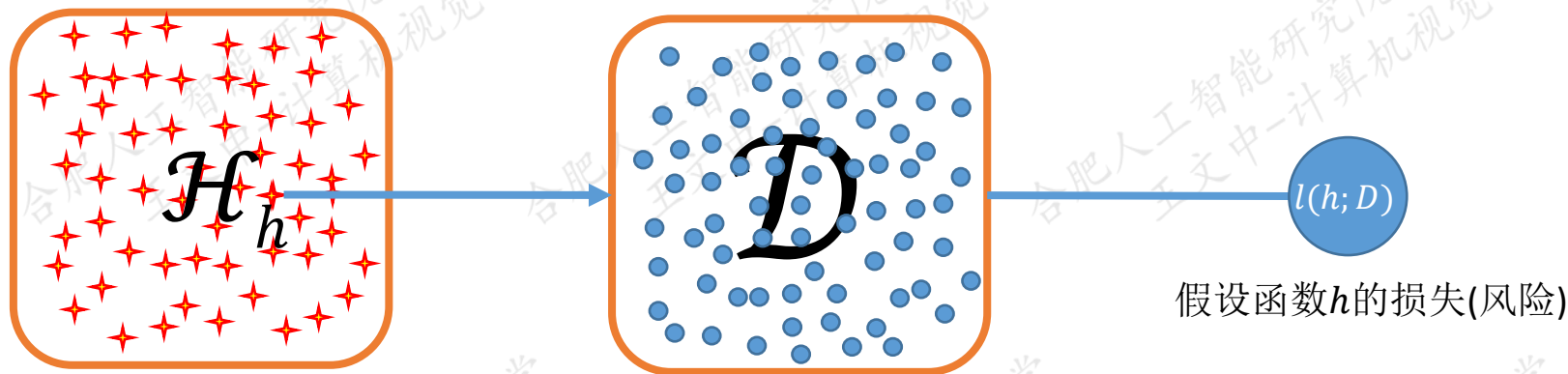
算法 \mathcal{A} : 2)搜索预测损失最小的 h , 并输出;



损失曲面

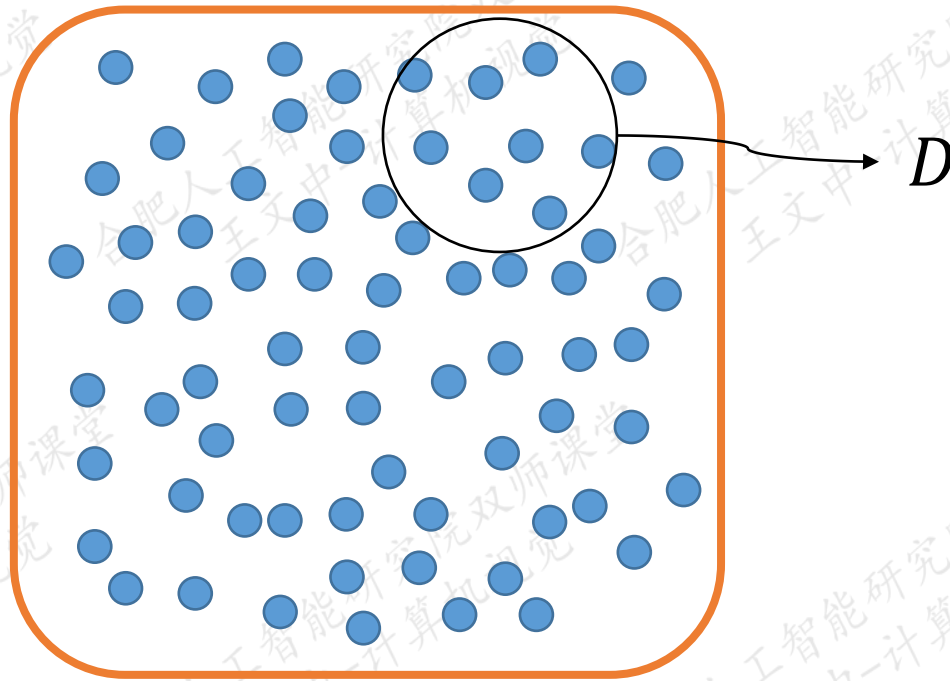
<https://www.cs.umd.edu/~tomg/projects/landscapes/>

过拟合与欠拟合



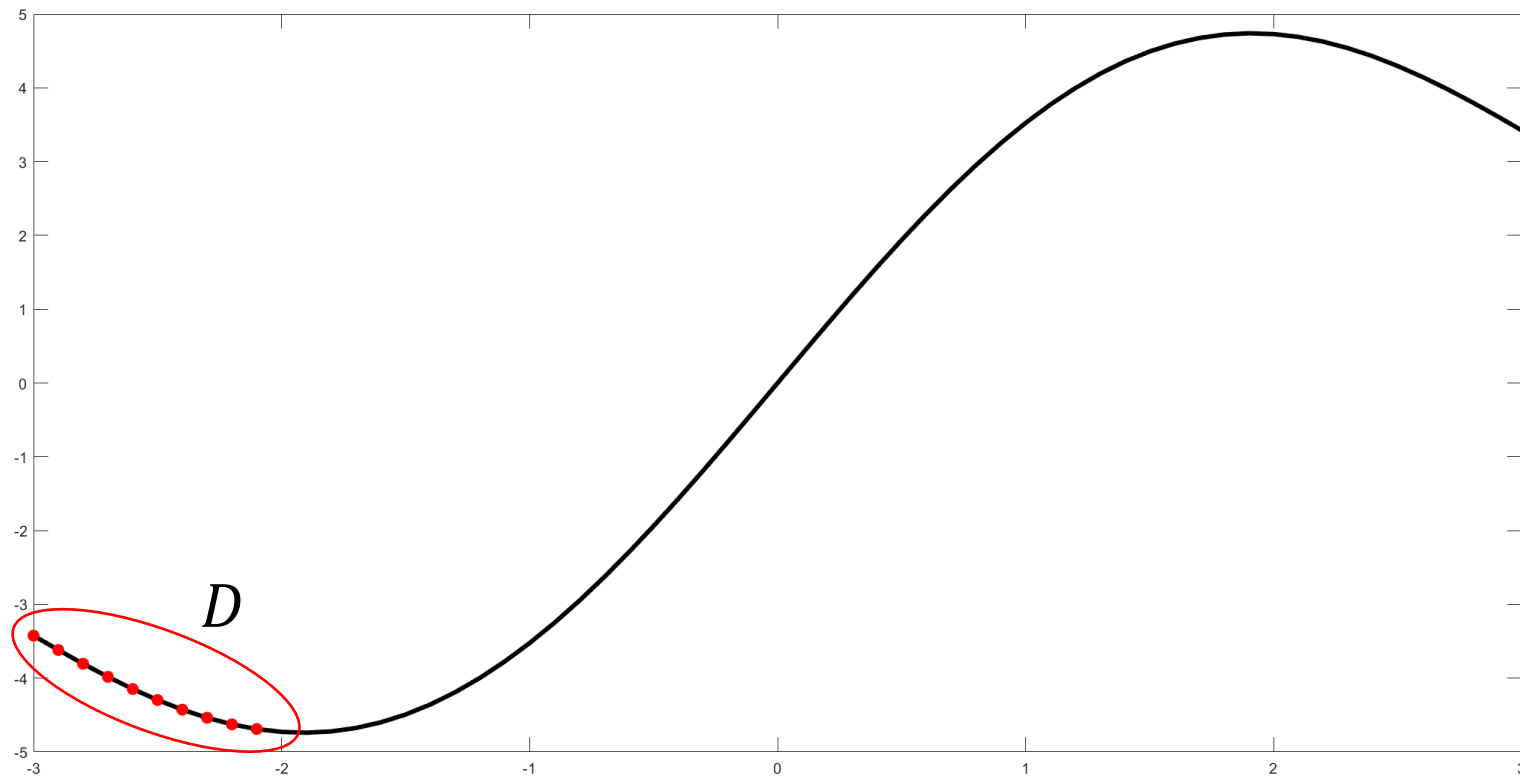
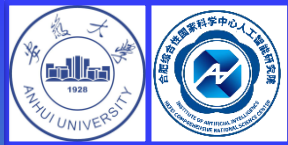
算法输出的最好的假设 h 受制于训练样本 D 和假设空间 H 。

训练样本 D 可能不具有代表性

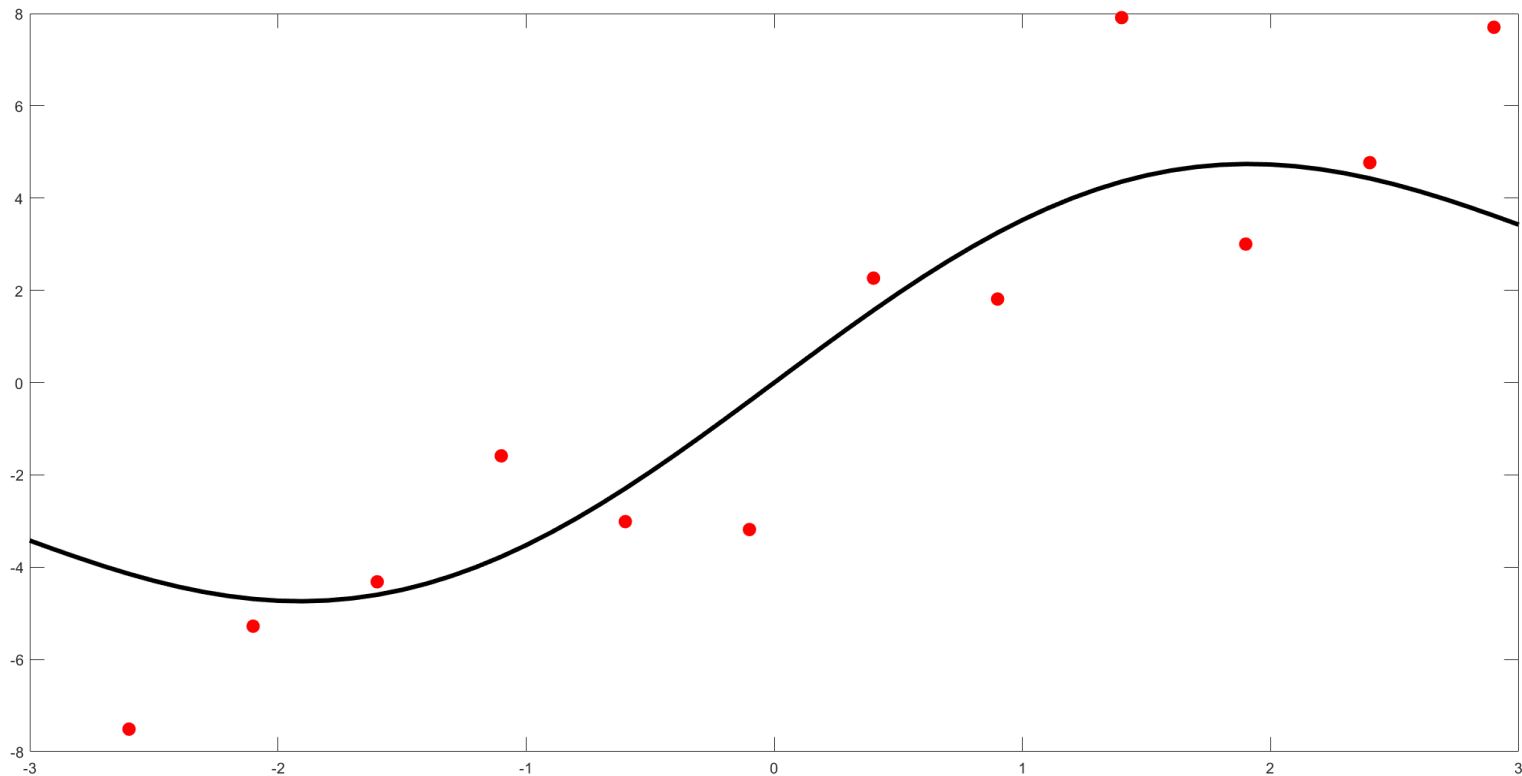


样本空间

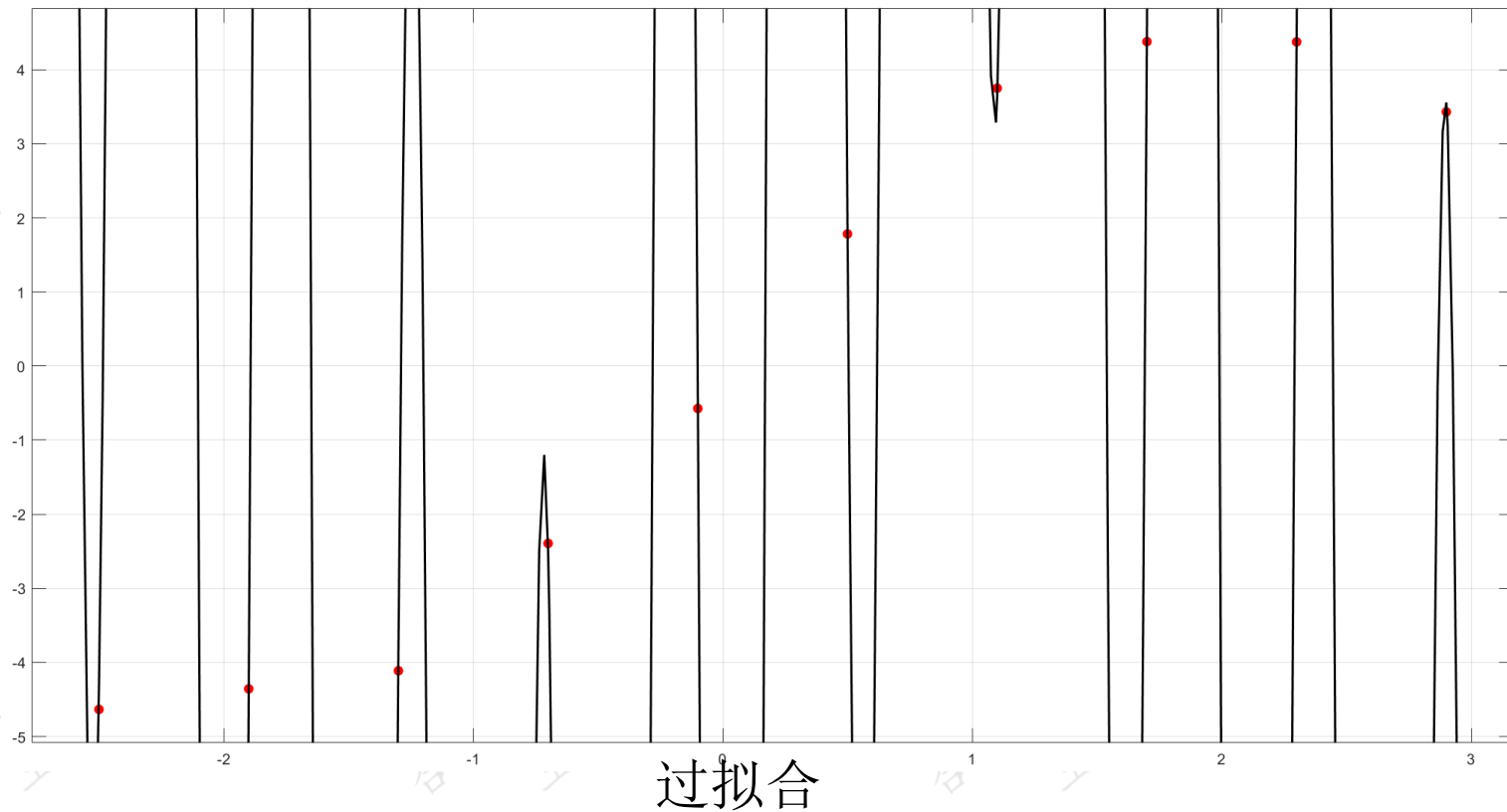
训练样本 D 可能不具有代表性



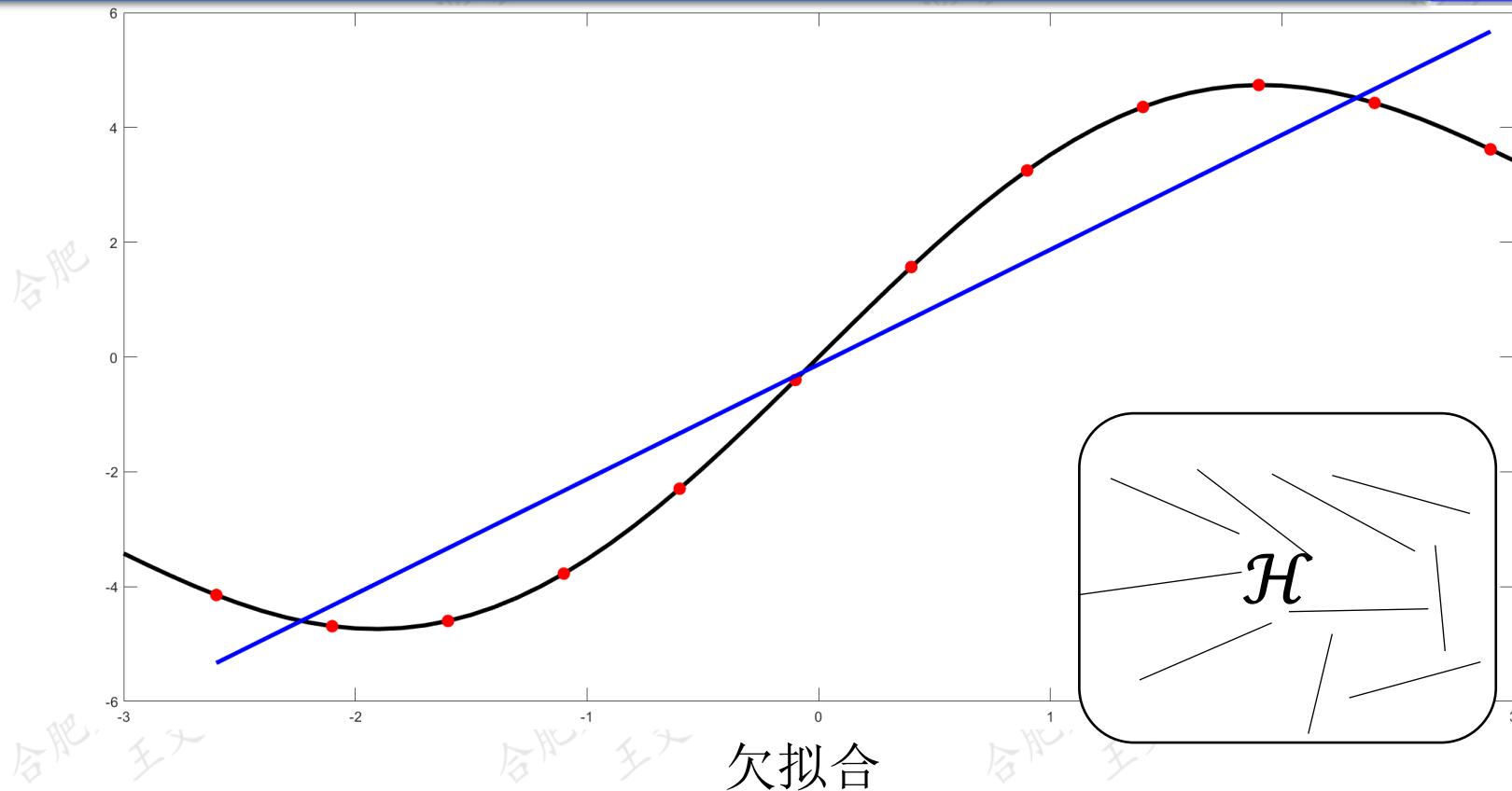
训练样本 D 可能噪声太大



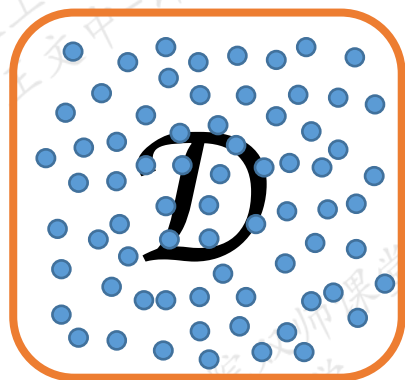
假设空间 H 可能太复杂



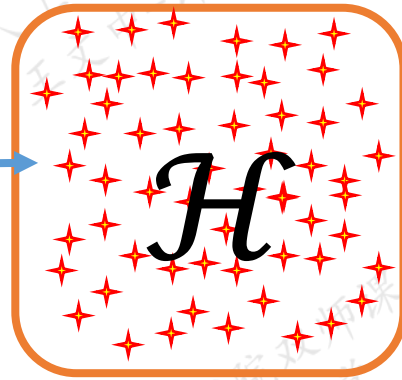
假设空间 H 可能太简单



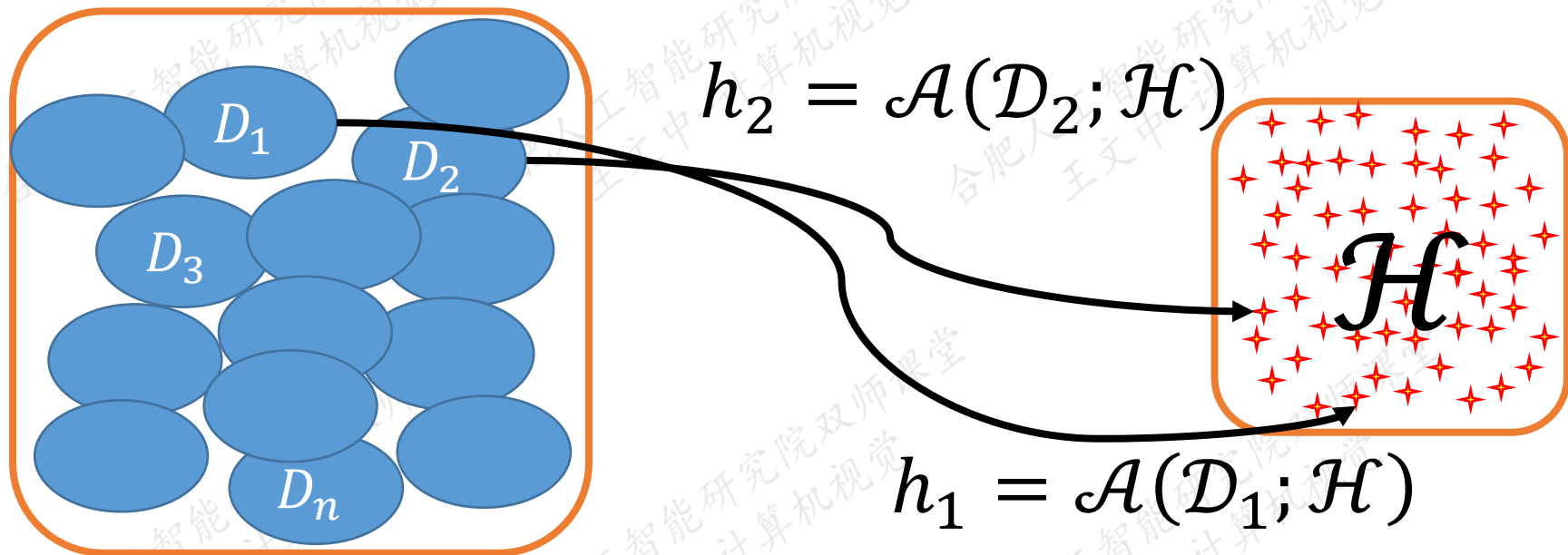
方差与偏差



$$h = \mathcal{A}(\mathcal{D}; \mathcal{H})$$



方差与偏差



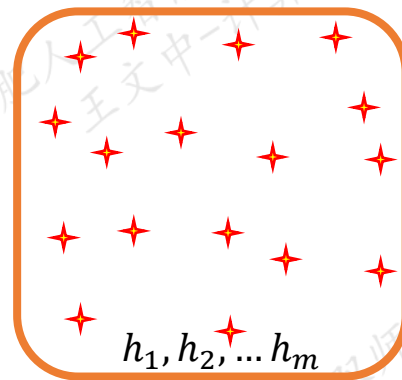
D_i : 大小相同的样本集

$$D_1, D_2, \dots, D_m \Rightarrow h_1, h_2, \dots, h_m$$

方差与偏差

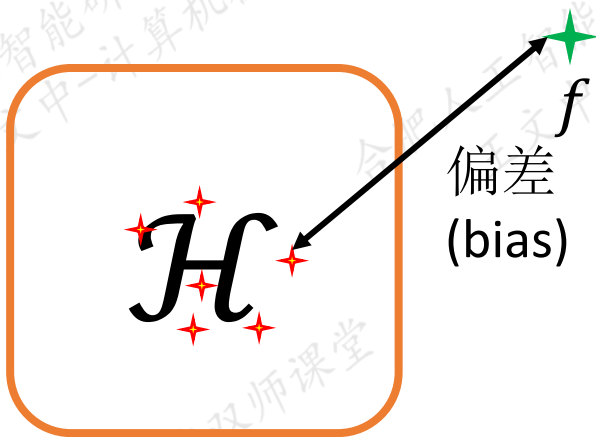


方差低

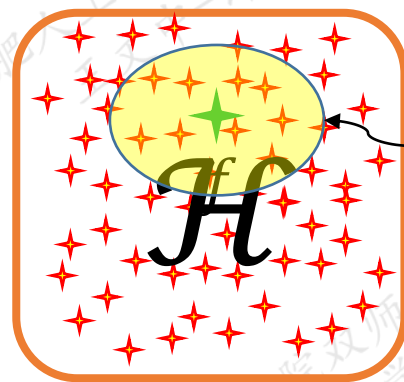


方差高

方差与偏差

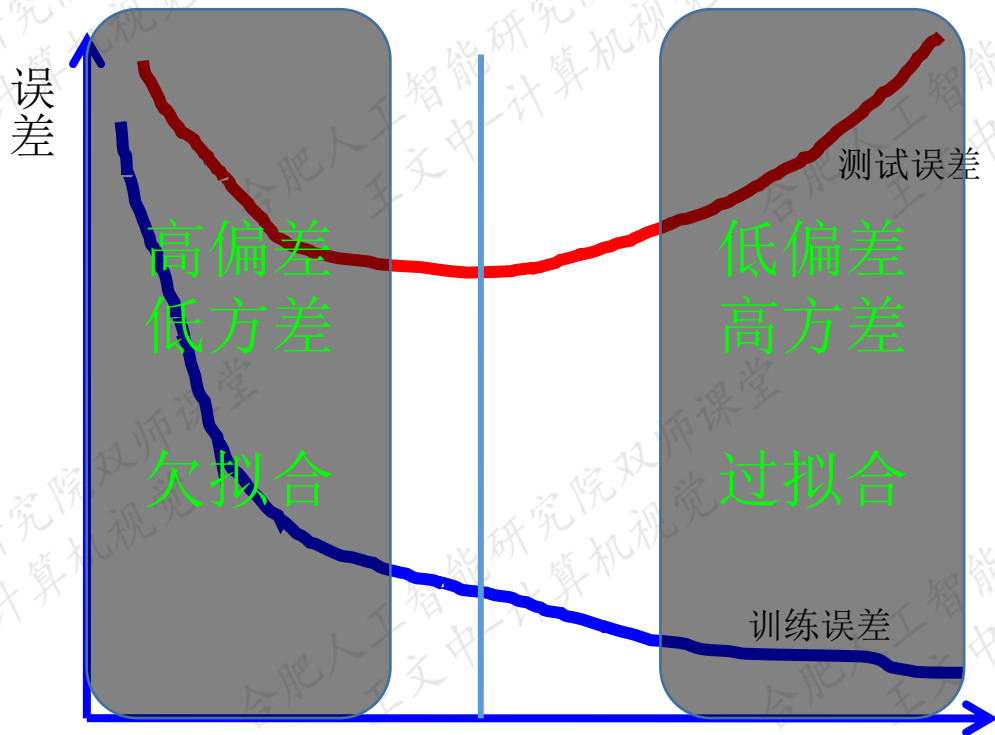


假设空间 \mathcal{H} 比较简单



假设空间 \mathcal{H} 比较复杂

方差与偏差



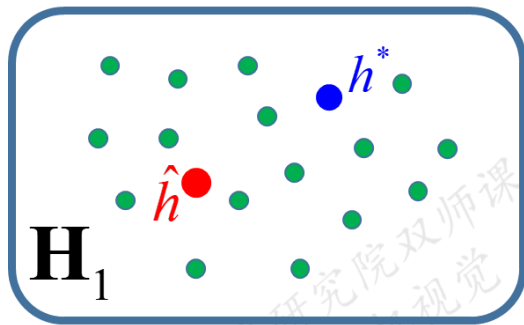
偏差与方差



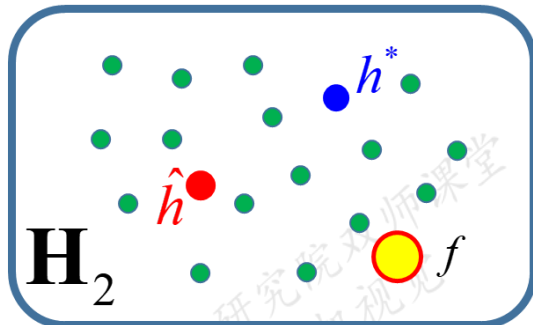
$$f(x) = x^2$$



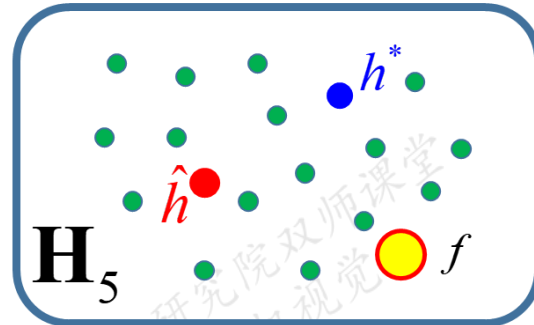
f



$$H_1 = \{h(x) = a_1x + a_0\}$$

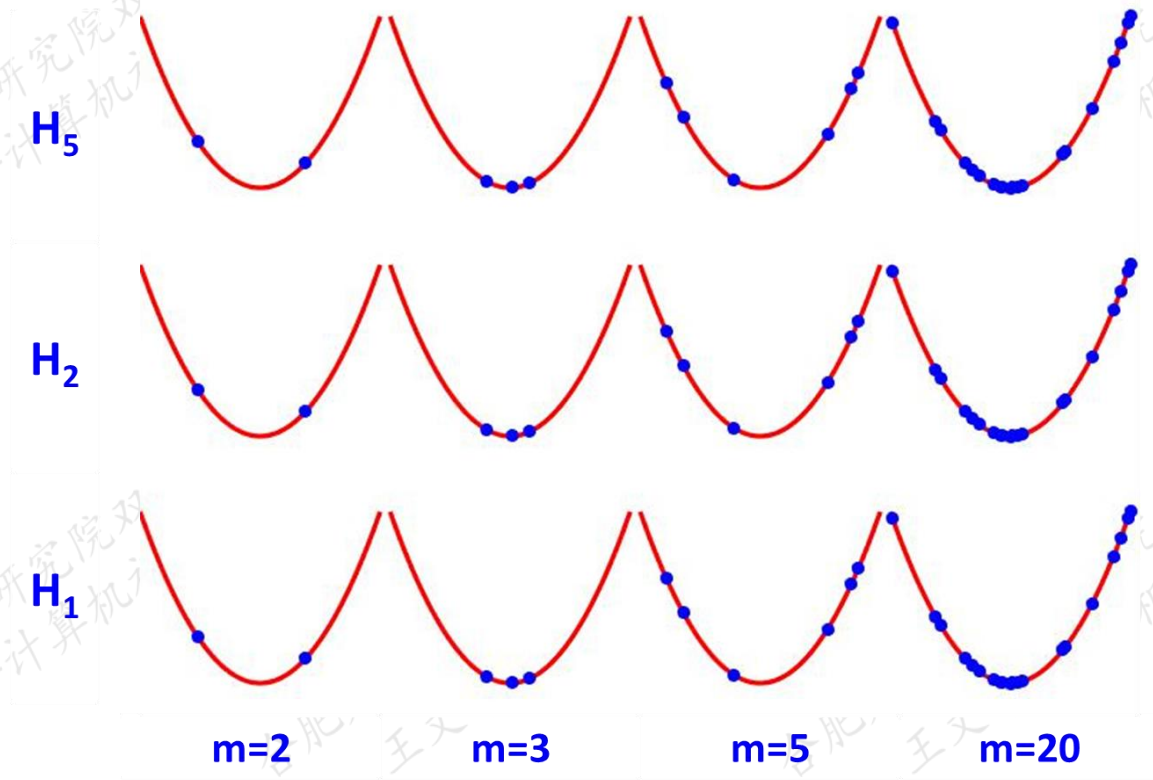


$$H_2 = \{h(x) = a_2x^2 + a_1x + a_0\}$$

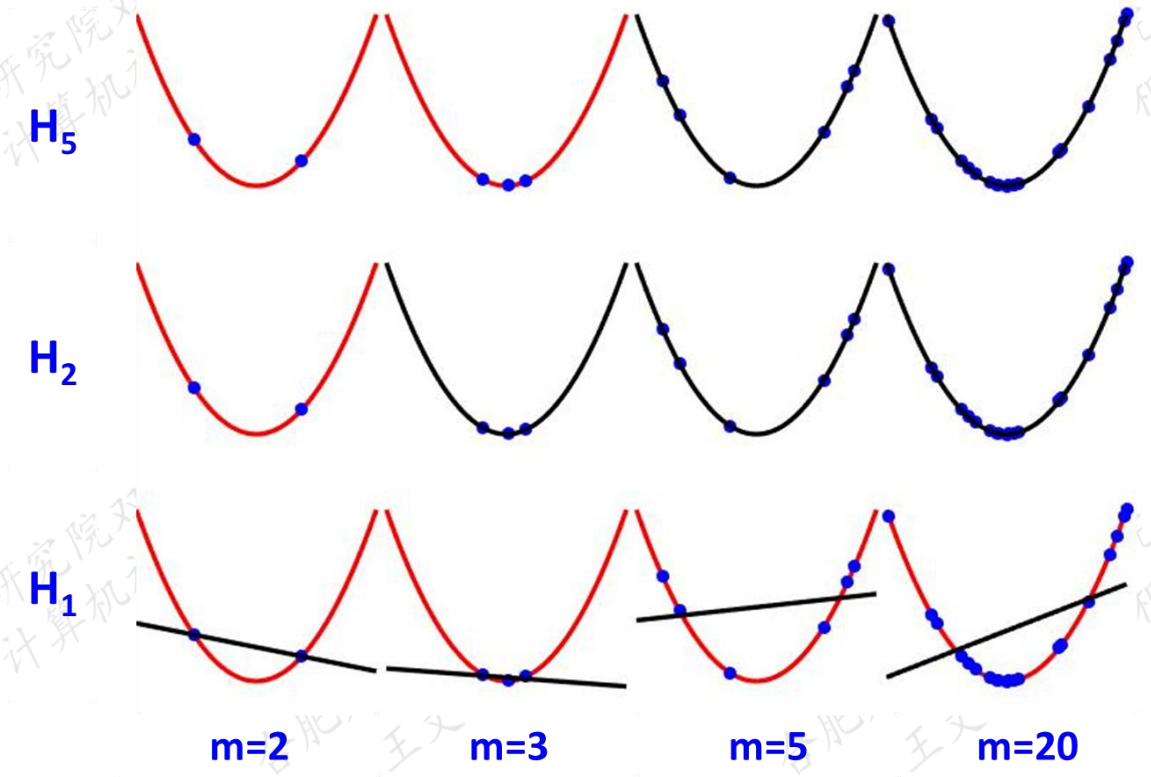


$$H_5 = \{h(x) = a_5x^5 + a_4x^4 + \dots + a_0\}$$

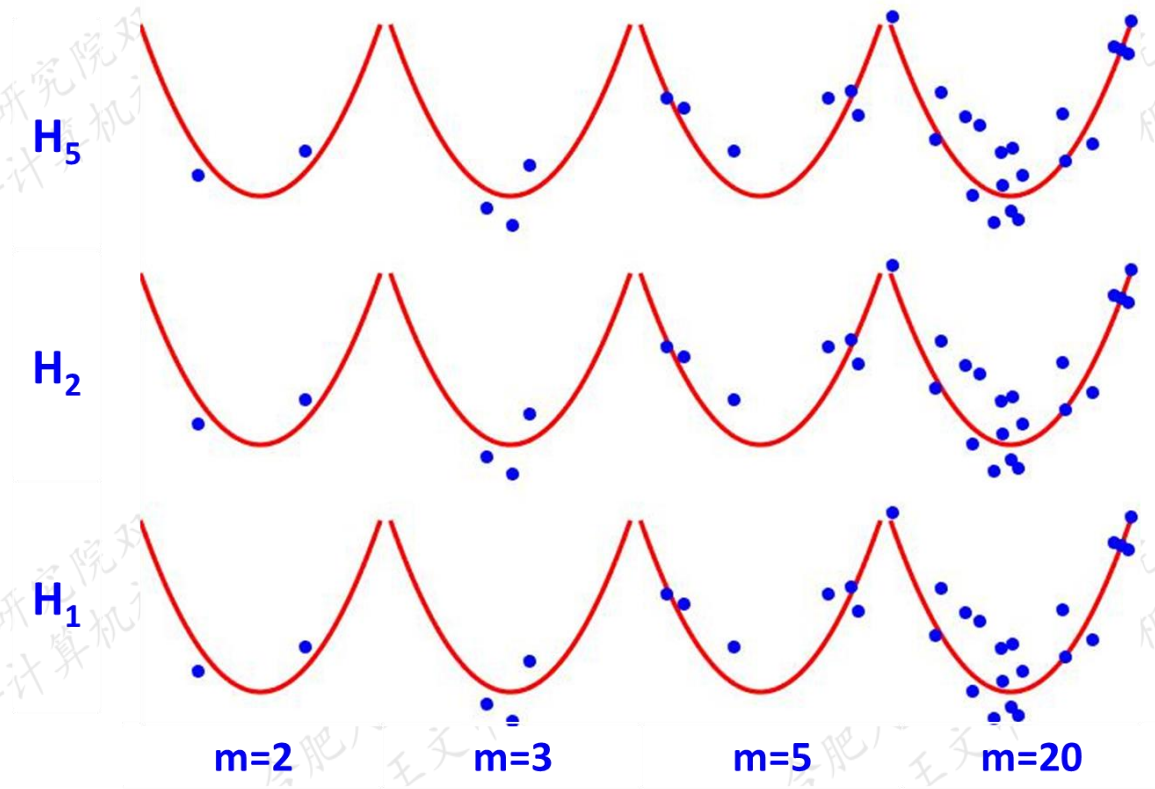
偏差与方差



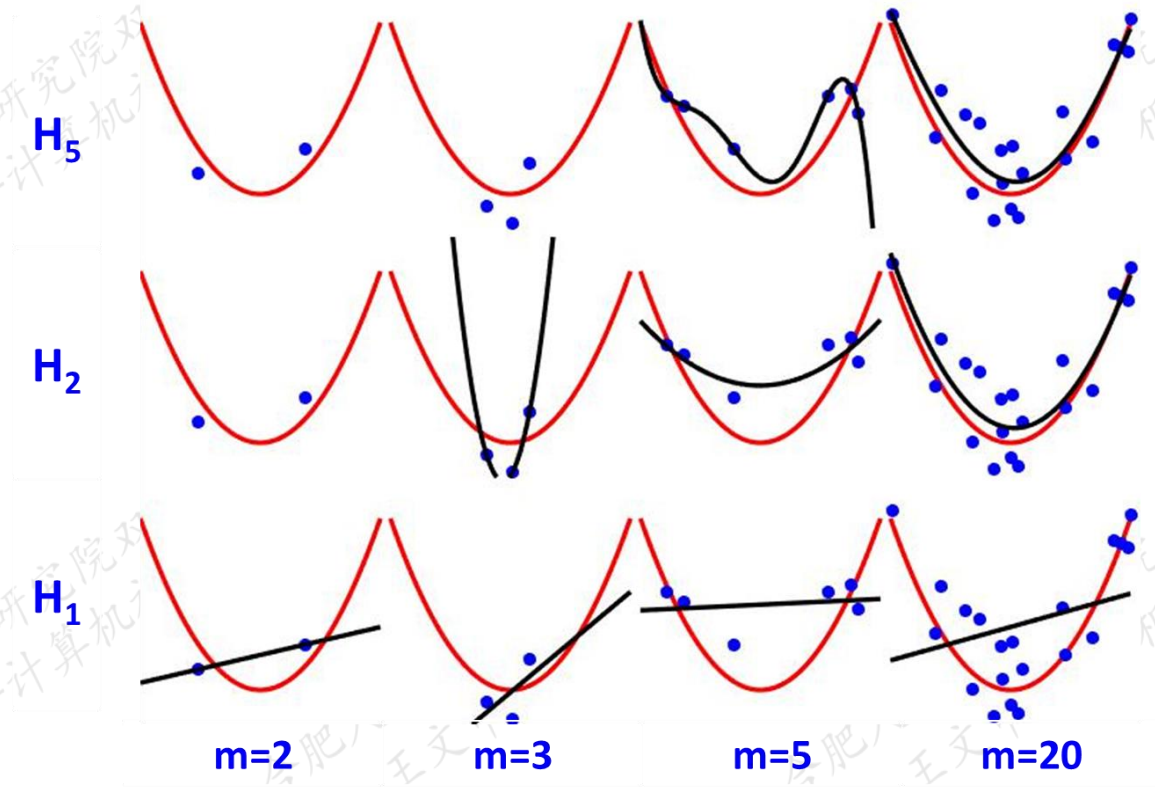
偏差与方差



偏差与方差



偏差与方差



什么情况下可以使用机器学习方法



- 1. 没有已知、高效的算法

- 2. 有数据

- 3. 输入与输出之间有关联

思考：哪些问题可以用机器学习解决？



- 从无人机航拍图像，识别农作物的种类：玉米、烟草、高粱。

思考：哪些问题可以用机器学习解决？



- 从十万张明星人脸照片，以及十万张猫脸照片，训练一个从明星人脸到猫脸的转换程序。

思考：哪些问题可以用机器学习解决？



- 根据人脸判别该人是否是罪犯？

思考：哪些问题可以用机器学习解决？

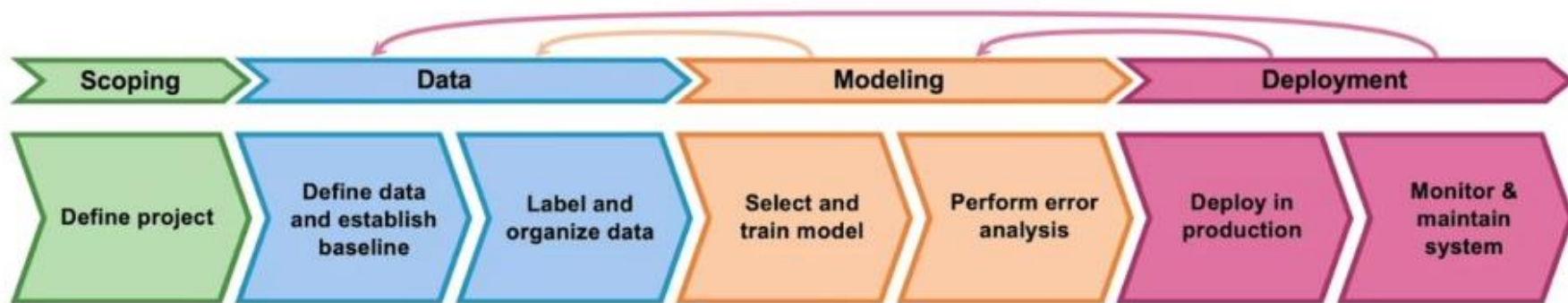


- 从监控视频图像中提取的行人图像，判别他的性别？种族？

图像分类实践

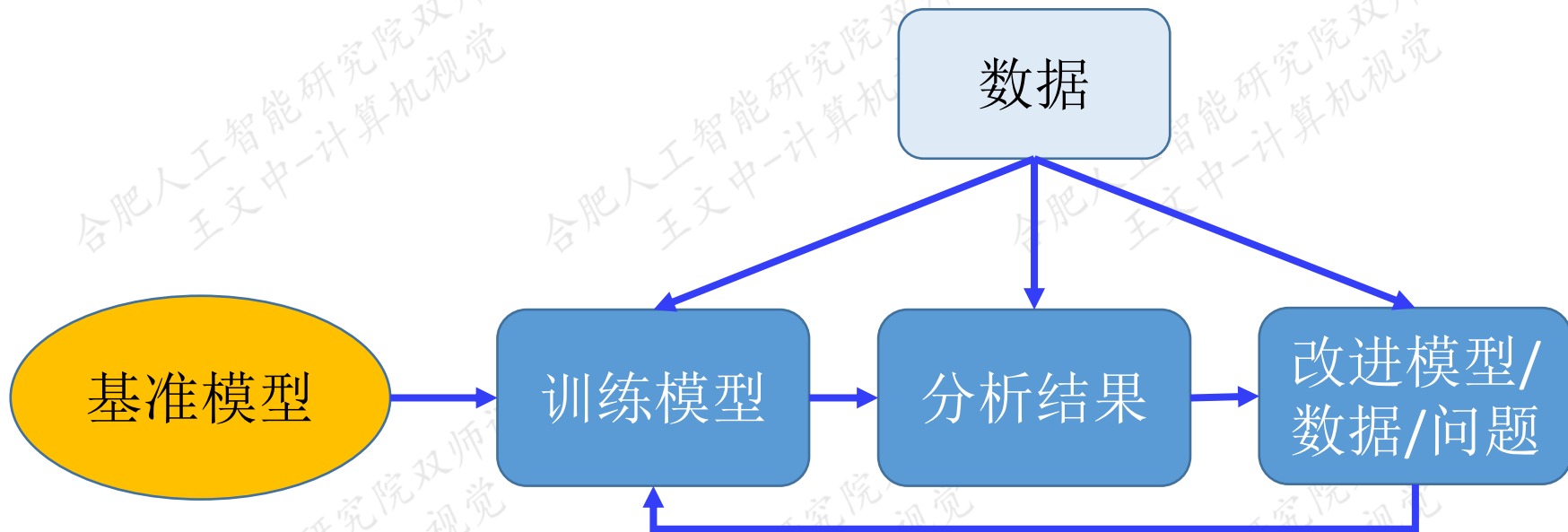
Andrew Ng: Machine Learning Yearning

The ML project lifecycle



Andrew Ng

快速迭代

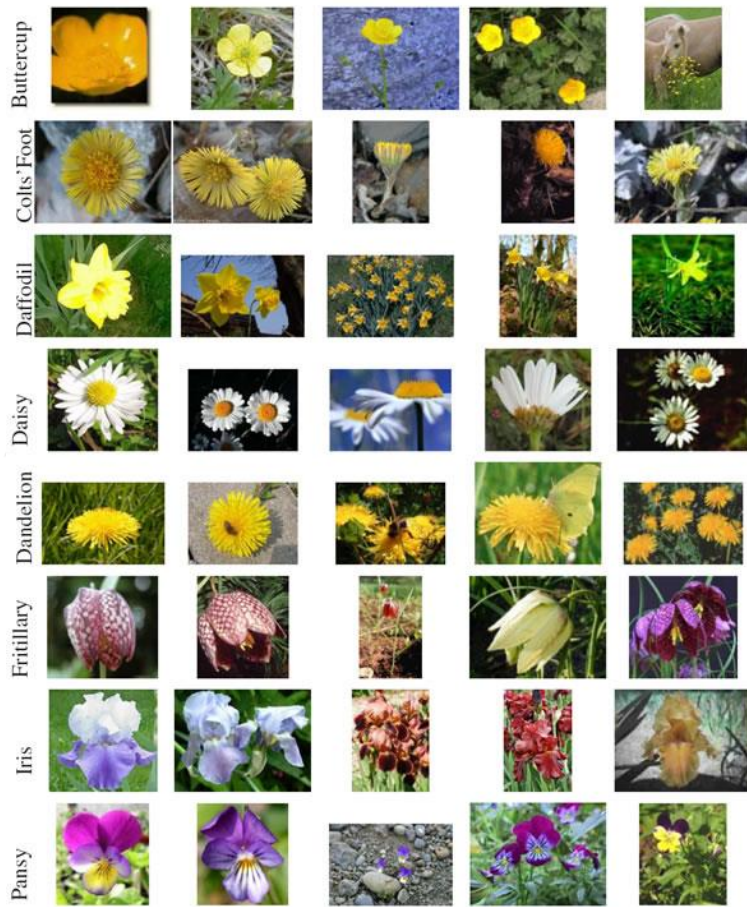


一般流程



- 1. 确定项目任务和应用场景
 - 定义输入和输出
 - 确定其它约束
- 2. 收集样本
 - 1) 采集图像（网络爬虫、自己拍摄）
 - 2) 标注图像
 - 3) 数据清洗（删除错误的图像，比如非花卉图像；删除低质量图像，比如分辨率非常低、成像非常模糊的图像；删除重复图像；修正标注错误）

- 3. 确定并复现基准：
 - 样本比较少：手工分类，得到Human Performance
 - SOTA（State-Of-The-Art）,站在巨人的肩膀上
- 4. 诊断并改进模型
 - 错误分析(Error Analysis)



确定基准模型(BaseLine)



- 图像分类：
 - ImageNet图像识别竞赛的模型
 - ResNet, GoogLeNet...
 - 例外：
 - MNIST手写体识别，不适合用ResNet这类预训练模型，为什么？
 - 交通信号灯颜色（红、绿、黄）识别，不适合用ResNet？
- 本例中数据量非常少
 - 使用ResNet-50微调

保证模型正常训练



- 使用少量样本训练模型

- 确保代码可以正常运行，快速排除代码bug。
- 确定合适的优化算法、学习速率等超参数。
- 合适的大模型在小样本上应该过拟合。

保证模型正常训练



- 损失函数不下降:

- 调节学习率
- 使用 `torch.optim.lr_scheduler`
- 更换优化器
- 检查数据是否有问题

- 损失函数出现NaN:

- 可能发生了梯度爆炸。
- 梯度裁剪:`nn.utils.clip_grad_value,`
`torch.nn.utils.clip_grad_norm.`
- 可能数据中包含了异常值

- 1. 诊断是否欠拟合：
 - 训练集精度（0.8）远低于期望值（Human Performance, 0.98）
 - 欠拟合解决对策：增加模型复杂度（比如减少正则化、增加网络层数）

- 2. 诊断是否过拟合：如果训练集精度比较高（0.95），但是验证集精度比较低（0.8），那么模型过拟合。
 - 检查验证集数与训练集数据是否满足独立同分布假设
 - 如果不满足：更新数据集。
 - 增加训练样本数量：收集更多样本、样本增广
 - 减少模型复杂度：增加正则化因子、减少可训练网络层数
 - 采用Bagging集成学习

错误分析



- 对分类结果进行分析
- 找出错误样本的规律
 - 模糊图像容易出错？
 - 低照度图像容易出错？
 - 目标较小的图像容易出错？
- 根据分析结果，确定下一步改进方向

错误分析



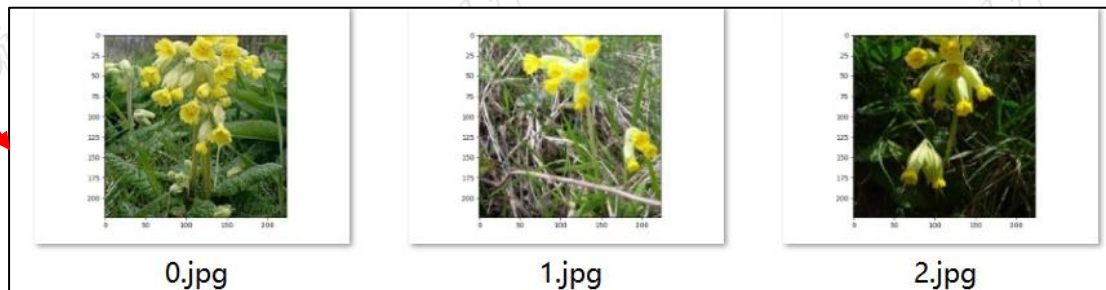
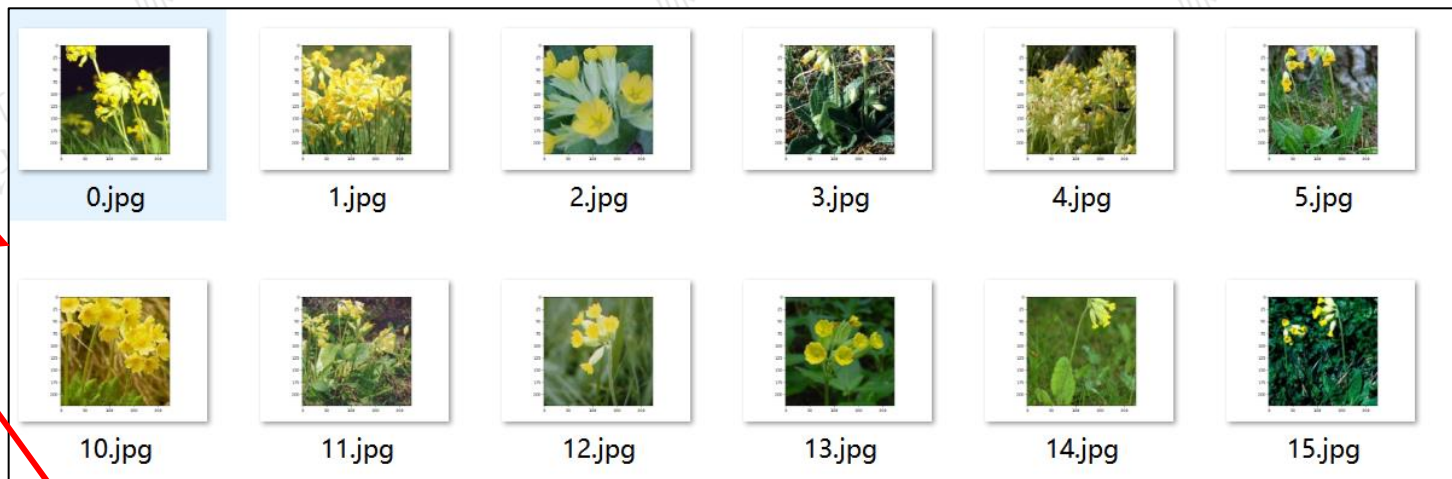
Image	Dog	Great cat	Blurry	Comments
1	✓			Usual pitbull color
2			✓	
3		✓	✓	Lion; picture taken at zoo on rainy day
4		✓		Panther behind tree
...
% of total	8%	43%	61%	

Andrew Ng

错误分析



- Bluebell
- Buttercup
- ColtsFoot
- Cowslip
- Cowslip
- Tulip**
- Crocus
- Daffodil
- Daisy
- Dandelion
- Fritillary
- Iris
- LilyValley
- Pansy
- Snowdrop
- Sunflower
- TigerLily
- Tulip
- Windflower



[54	0	0	0	1	0	0	0	0	1	1	2	1	0	0	0	0]	Bluebell : 0.9
[0	47	1	6	0	0	1	0	0	0	0	0	0	3	2	0]	Buttercup : 0.78
[0	10	24	5	0	10	3	0	0	0	0	0	0	1	7	0]	ColtsFoot : 0.4
[0	0	0	57	0	0	0	0	0	0	0	0	0	0	3	0]	Cowslip : 0.95
[0	0	0	0	50	0	7	0	0	0	1	2	0	0	0	0]	Crocus : 0.83
[0	11	0	11	0	27	2	0	0	0	0	1	0	1	6	1]	Daffodil : 0.45
[0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0]	Daisy : 1.0
[0	15	4	3	0	5	0	29	0	0	0	0	0	1	3	0]	Dandelion : 0.48
[0	0	0	0	0	0	0	0	54	0	0	6	0	0	0	0]	Fritillary : 0.9
[1	2	0	0	1	2	1	0	0	48	0	3	0	1	1	0]	Iris : 0.8
[0	0	0	3	0	0	1	0	0	0	43	0	11	0	1	0]	LilyValley : 0.72
[0	0	0	1	0	0	0	0	1	0	0	56	0	0	0	2]	Pansy : 0.93
[0	0	0	1	2	0	1	0	0	0	1	0	53	0	0	2]	Snowdrop : 0.88
[0	1	0	1	0	1	0	0	0	0	0	0	55	2	0	0]	Sunflower : 0.92
[0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0]	TigerLily : 1.0
[0	7	0	13	1	0	2	0	0	0	1	0	0	1	35	0]	Tulip : 0.58
[0	0	0	0	0	0	4	0	0	1	0	0	1	0	0	54]	Windflower : 0.9

ColtsFoot-> Buttercup: 17%



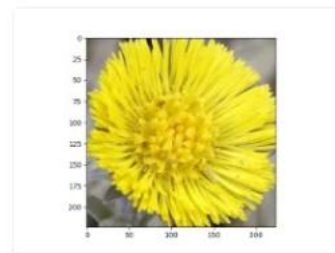
0.jpg



1.jpg



2.jpg



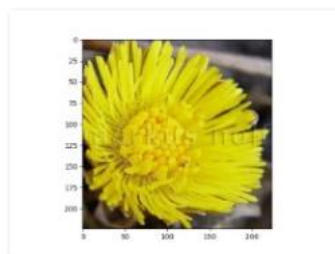
3.jpg



4.jpg



5.jpg



6.jpg



7.jpg

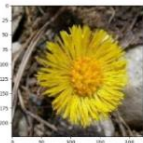


8.jpg



9.jpg

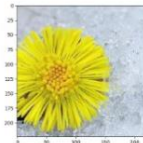
ColtsFoot-> ColtsFoot: 40%



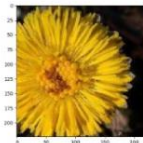
0.jpg



1.jpg



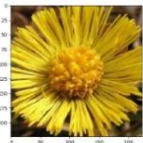
2.jpg



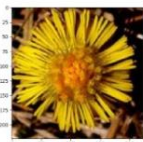
3.jpg



4.jpg



5.jpg



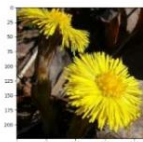
6.jpg



7.jpg



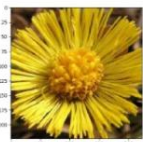
8.jpg



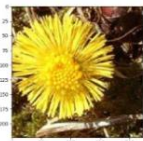
9.jpg



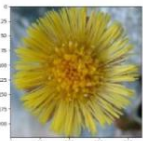
10.jpg



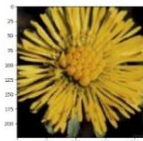
11.jpg



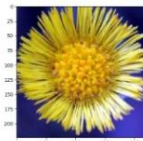
12.jpg



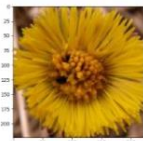
13.jpg



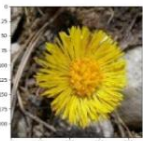
14.jpg



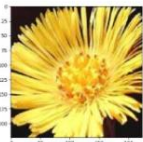
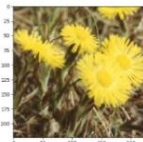
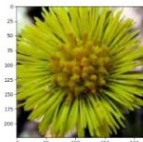
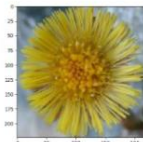
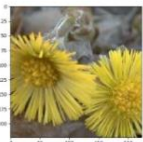
15.jpg



16.jpg



17.jpg



真实label: ColtsFoot



改成暖色调



预测结果: Buttercup ☹

预测结果: ColtsFoot ☺

Dandelion->Cowslip



检查输入网络的图片，而不是原始图像



```
train_set = datasets.ImageFolder(  
    train_dir,  
    transforms.Compose([  
        transforms.RandomResizedCrop(224),  
        transforms.RandomHorizontalFlip(),  
        transforms.ToTensor(),  
        normalize,  
    ]))
```


检查输入网络的图片，而不是原始图像



合肥人工智能研究院双师课
王文中-计算机视觉

合肥人工智能研究院双师课
王文中-计算机视觉

合肥人工智能研究院双师课
王文中-计算机视觉

了解图像增广可能带来的问题



了解图像增广可能带来的问题



```
train_set = datasets.ImageFolder(  
    train_dir,  
    transforms.Compose([  
        transforms.RandomResizedCrop(224,  
                                       scale=(0.9,1), ratio=(0.8,1.2)),  
        transforms.RandomHorizontalFlip(),  
        transforms.ToTensor(),  
        normalize,  
    ]))
```

调整输入x



调整输入x



- Resnet50, Layer4微调
- 原始图像, RandomResizedCrop(224):
 - 97.94%, 93.82%
- 原始图像, RandomResizedCrop(224, scale=(0.9,1), ratio=(0.8,1.2)):
 - 99%, 95%
- 裁剪图像, RandomResizedCrop(224, scale=(0.9,1), ratio=(0.8,1.2)):
 - 100%, 95.59%

调整输出 y



Crocus

- 1.机器学习的基本概念
 - 独立同分布假设
 - 过拟合与欠拟合的原因及其对策
 - 机器学习的适用范围
- 2.开发图像识别模型
 - 快速迭代，从错误中找方向
 - 模型诊断
 - 错误分析