

Practice of AI

Jim Xie

2020/7/6



Outline

1. 时序预测原理与实现
2. 图像识别原理与实现
3. 自由练习
4. 前沿技术介绍
5.

什么是时序预测？

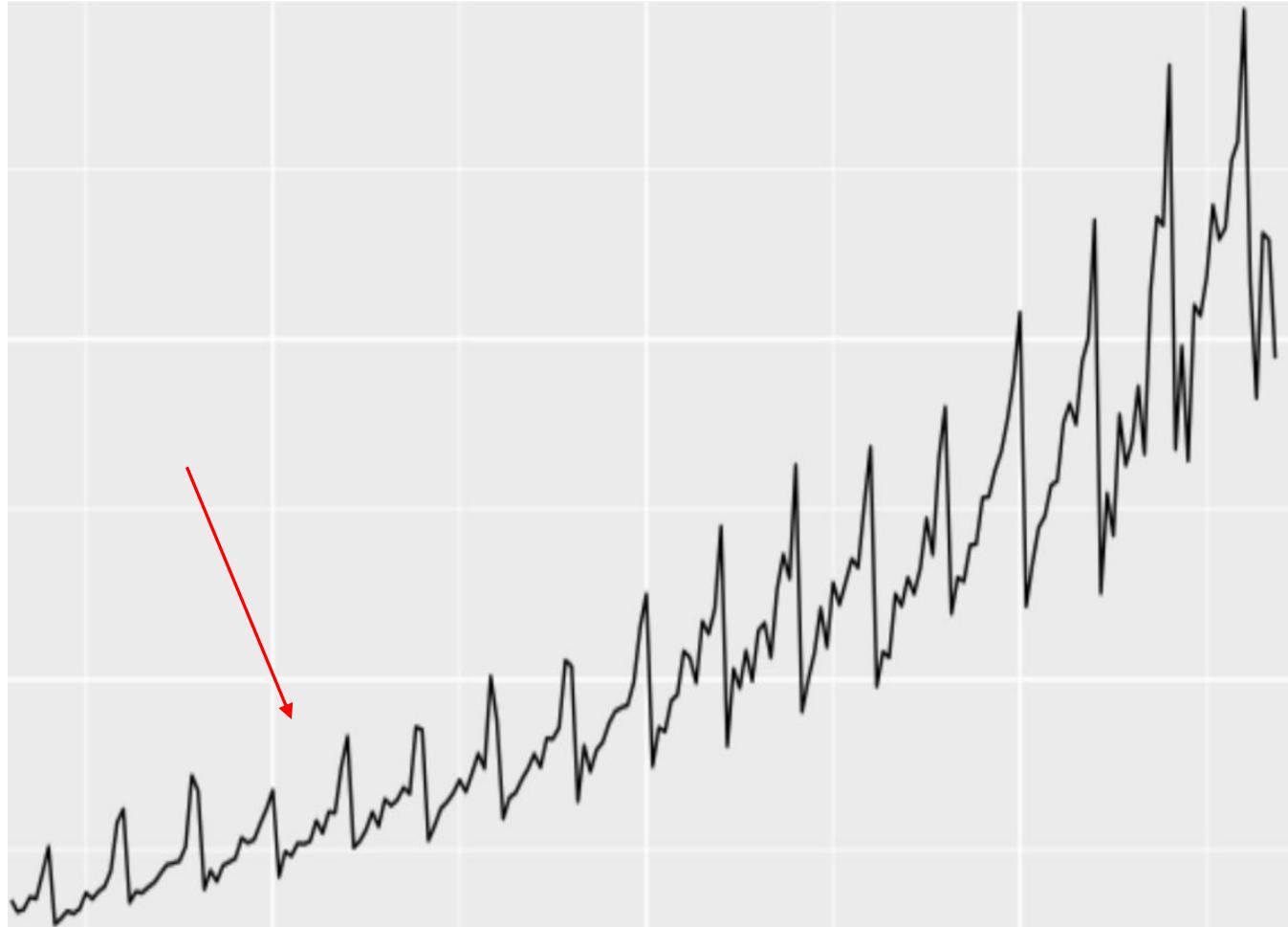
通过过去的数据，预测未来的数据

- 每小时的点击量
- 每天的气温
- 每天的股票价格
- 每天的病毒感染人数
- 每年的降水量
-



时序数据

- ❖ 趋势性
- ❖ 季节性
(间隔固定)
- ❖ 周期性
(间隔不固定)
- ❖ 随机性



时序模型 : AR

Auto Regression

日期	10-1	10-2	10-3	10-4	10-5	10-6	10-7	10-8	10-9	10-10	10-11
数据	100	102	103	104	105	106	107	108	109	110	???
x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}

目标 : 通过前几天的数据 , 得到 10-11 的数据 x_{11}

$$x_{11} = C + w_{10} \times x_{10} + w_9 \times x_9$$

$$x_{11} = C + w_{10} \times x_{10} + w_9 \times x_9 + w_8 \times x_8$$

$$x_{11} = C + w_{10} \times x_{10} + w_9 \times x_9 + w_8 \times x_8 + w_7 \times x_7$$

$$x_{11} = C + w_{10} \times x_{10} + w_9 \times x_9 + w_8 \times x_8 + w_7 \times x_7 + \dots + w_p \times x_p$$

(P 表示回溯的天数 , 训练得到 w_0, w_1, \dots, w_p)

AR模型

$$x_{11} = C + w_{10} \times x_{10} + w_9 \times x_9 + w_8 \times x_8 + w_7 \times x_7 + \dots + w_p \times x_p$$



$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t$$

\downarrow \downarrow \downarrow \downarrow

x_i C w_i 误差项

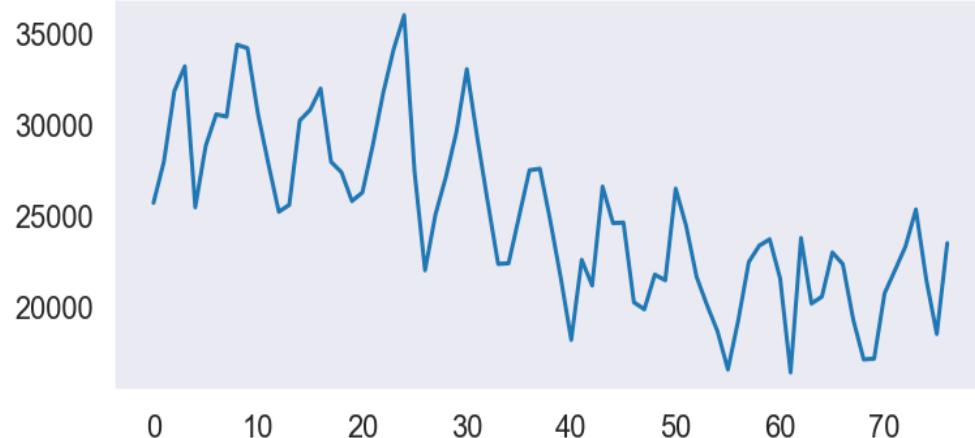
平稳序列

方差恒定，均值恒定

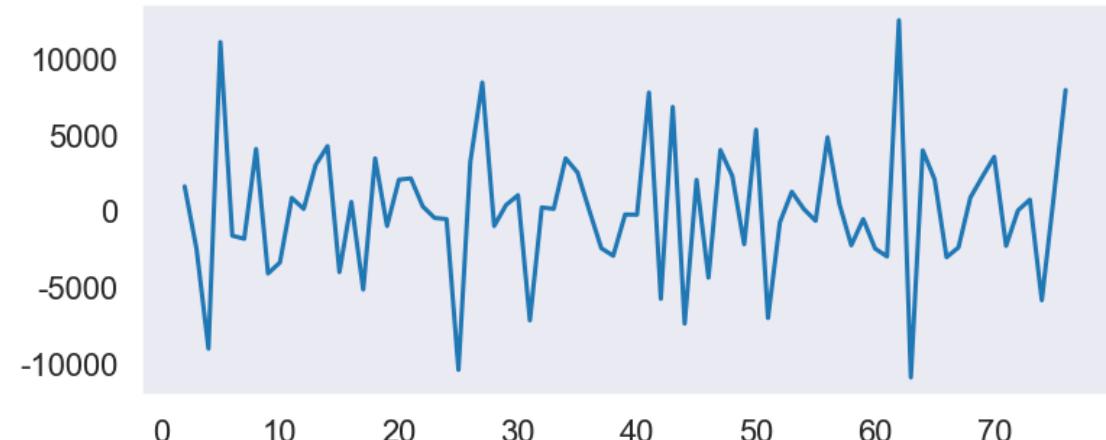


1. 如何判断序列是否平稳？
2. 怎么样将序列变得平稳？

美国新冠病毒增长量



差分
→



	日期	10-1	10-2	10-3	10-4	10-5	10-6	10-7	10-8	10-9	10-10	10-11
不平稳	数据	100	102	103	104	105	106	107	108	109	110	???
平稳	Δ	2	1	1	1	1	1	1	1	1	???	

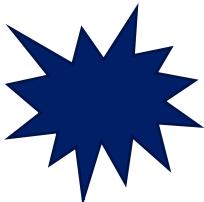
时序模型：MA

把X看成是Base + 噪音组合(ϵ)

日期	10-1	10-2	10-3	10-4	10-5	10-6	10-7	10-8	10-9	10-10	10-11
数据	100	102	103	104	105	106	107	108	109	110	???

假如Base为100

日期	10-1	10-2	10-3	10-4	10-5	10-6	10-7	10-8	10-9	10-10	10-11
数据	100	102	103	104	105	106	107	108	109	110	???
ϵ_i	0	2	3	4	5	6	7	8	9	10	???
x_i	100+0	100+2	100+3	100+4	100+5	100+6	100+7	100+8	100+9	100+10	100+???



找到噪音 ϵ_i 的变化规律，就可以计算出10-11日的值

时序模型：MA

Moving Average

日期	10-1	10-2	10-3	10-4	10-5	10-6	10-7	10-8	10-9	10-10	10-11
数据	100	102	103	104	105	106	107	108	109	110	???
ϵ_i	0	2	3	4	5	6	7	8	9	10	???
x_i	100+0	100+2	100+3	100+4	100+5	100+6	100+7	100+8	100+9	100+10	100+???

将 ϵ_i 作为变量，进行线性回归分析

$$X_t = c_0 + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

(q表示回溯的天数，训练得到 $\theta_0, \theta_1 \dots \theta_n$)

预测模型：ARMA

ARMA = AR + MA

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}$$

(P和Q是两个超参数)

新冠病毒预测

Example:

<http://10.206.67.123:8888/notebooks/JimXie/excise/example/%E7%BE%8E%E5%9B%BD%E6%96%B0%E5%86%A0%E9%A2%84%E6%B5%8B.ipynb>

练习:

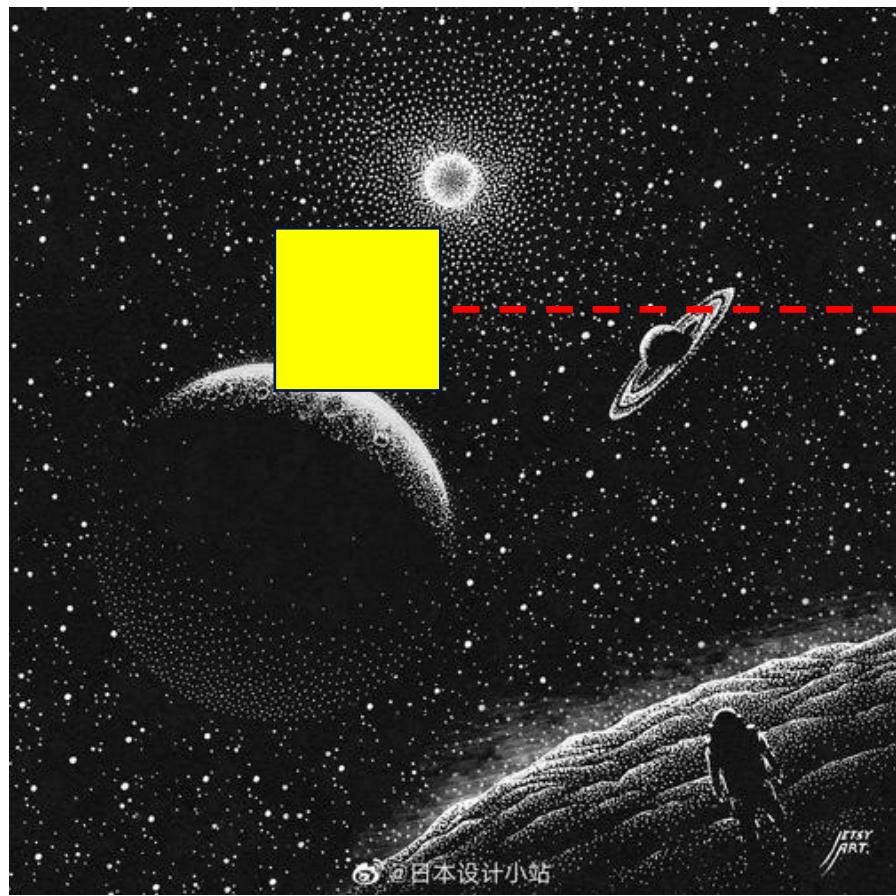
<http://10.206.67.123:8888/notebooks/JimXie/excise/%E5%8D%B0%E5%BA%A6%E6%96%B0%E5%86%A0%E9%A2%84%E6%B5%8B.ipynb>

图像识别

CV & CNN

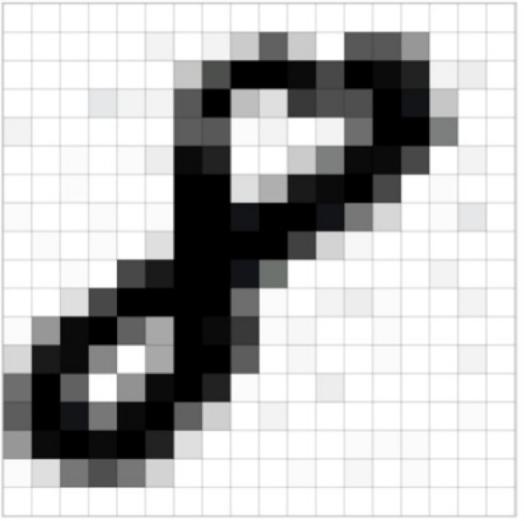
数字图像

[height,weight,channel]



08 02 22 97 38 15 00 40 00 75 06 05 07 78 52 12 50 77 91 08
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 34 65
52 70 95 25 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91
22 31 16 71 51 67 43 59 41 92 36 56 22 40 40 28 66 33 13 80
24 47 32 60 99 03 45 02 44 75 33 53 78 36 84 20 35 17 12 50
32 98 81 28 64 23 8 40 67 59 54 70 66 18 38 64 70
67 26 20 68 02 62 3 94 39 43 08 40 91 66 49 94 21
24 55 58 05 66 73 7 78 78 96 83 14 88 34 89 63 72
21 36 23 09 75 00 5 35 14 00 61 33 97 34 31 33 95
78 17 53 28 22 75 31 67 15 94 03 80 04 62 16 14 09 53 56 92
16 39 05 42 96 35 31 47 55 58 85 24 00 17 54 24 36 29 85 57
86 56 00 48 35 71 89 07 05 44 46 37 41 60 21 58 51 54 17 55
19 80 81 68 05 94 47 69 28 75 92 13 86 52 17 77 04 89 55 40
04 52 08 83 97 35 99 14 07 97 57 32 16 26 26 79 33 27 98 66
88 36 68 07 57 62 20 72 03 46 33 67 46 55 12 32 63 93 53 69
04 42 16 73 38 25 39 11 24 94 72 18 08 46 29 32 60 62 76 36
20 69 36 43 72 30 23 88 34 62 99 69 82 67 59 85 74 04 36 16
20 73 35 29 78 31 90 01 74 31 49 71 48 86 81 16 23 57 05 54
01 70 54 71 83 51 54 69 16 92 33 45 48 86 81 16 23 57 05 54

图像识别原理



8



[1]	9	1	29	70	114	76	8	8	4	5	5	0	111	162	9	8	62	62]
[3]	0	33	61	102	106	34	0	0	0	0	49	182	150	1	12	65	59]	
[1]	0	40	54	123	90	72	77	52	51	49	121	205	98	0	15	67	59]	
[3]	1	41	57	74	54	96	181	228	170	90	149	208	56	0	16	69	59]	
[6]	1	32	36	47	81	85	90	176	206	140	171	186	22	3	15	72	63]	
[4]	1	31	39	66	71	71	97	147	214	203	190	198	22	6	17	73	65]	
[2]	3	15	30	52	57	68	123	161	197	207	200	179	8	8	18	73	66]	
[2]	2	17	37	34	40	78	103	148	187	205	225	165	1	8	19	76	68]	
[2]	3	20	44	37	34	35	26	78	156	214	145	200	38	2	21	78	69]	
[2]	2	20	34	21	43	70	21	43	139	205	93	211	70	0	23	78	72]	
[3]	4	16	24	14	21	102	175	120	130	226	212	236	70	0	25	78	72]	
[6]	5	13	21	28	29	97	216	184	90	196	255	255	84	4	24	79	74]	
[6]	5	15	25	30	39	63	105	140	66	113	252	251	74	4	28	79	75]	
[5]	5	16	32	38	57	69	85	93	120	128	251	255	154	19	26	80	77]	
[6]	5	20	42	55	62	66	76	86	104	148	242	254	241	83	26	80	77]	
[2]	3	20	38	55	64	69	80	78	109	195	247	252	255	172	40	78	77]	
[10]	8	23	34	44	64	88	104	119	173	234	247	253	254	227	66	74	74]	
[32]	6	24	37	45	63	85	114	154	196	226	245	251	252	250	112	66	71]	

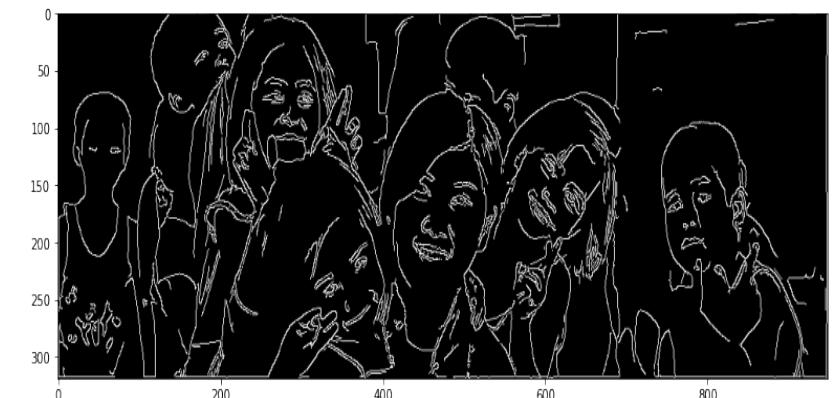
猫

图像的常见操作

❖ 灰度化

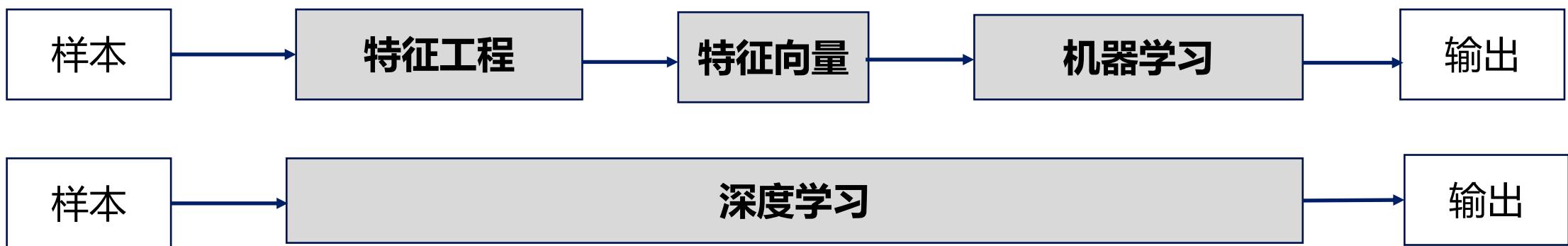


❖ 二值化



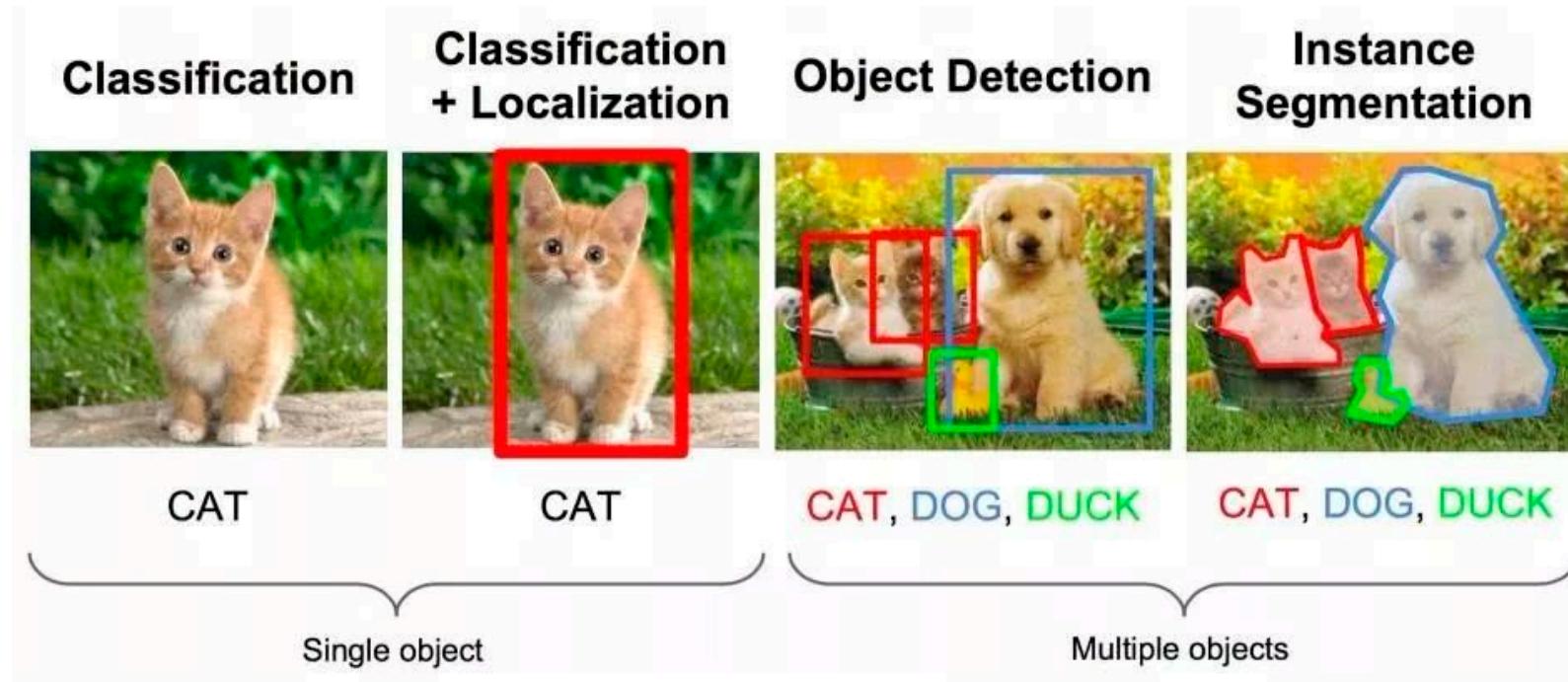
❖ 边缘检测

Deep Learning

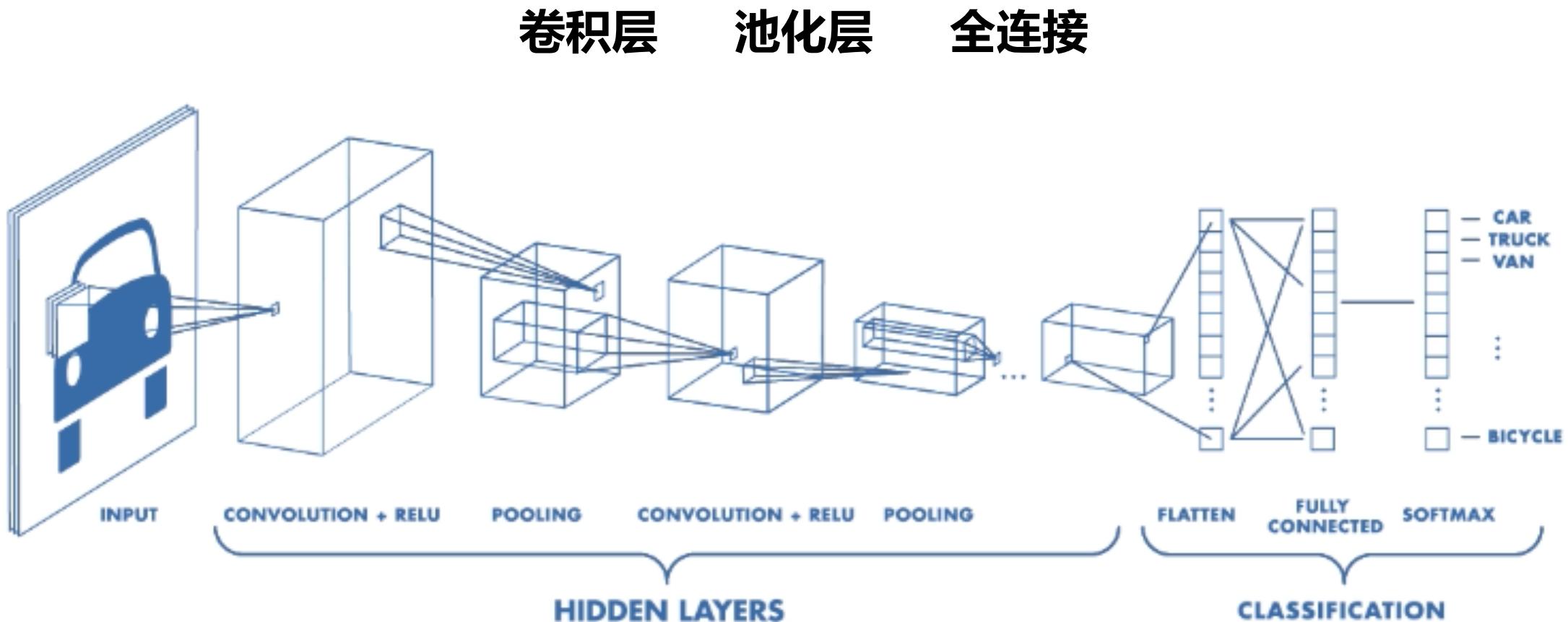


深度学习在图像领域的应用

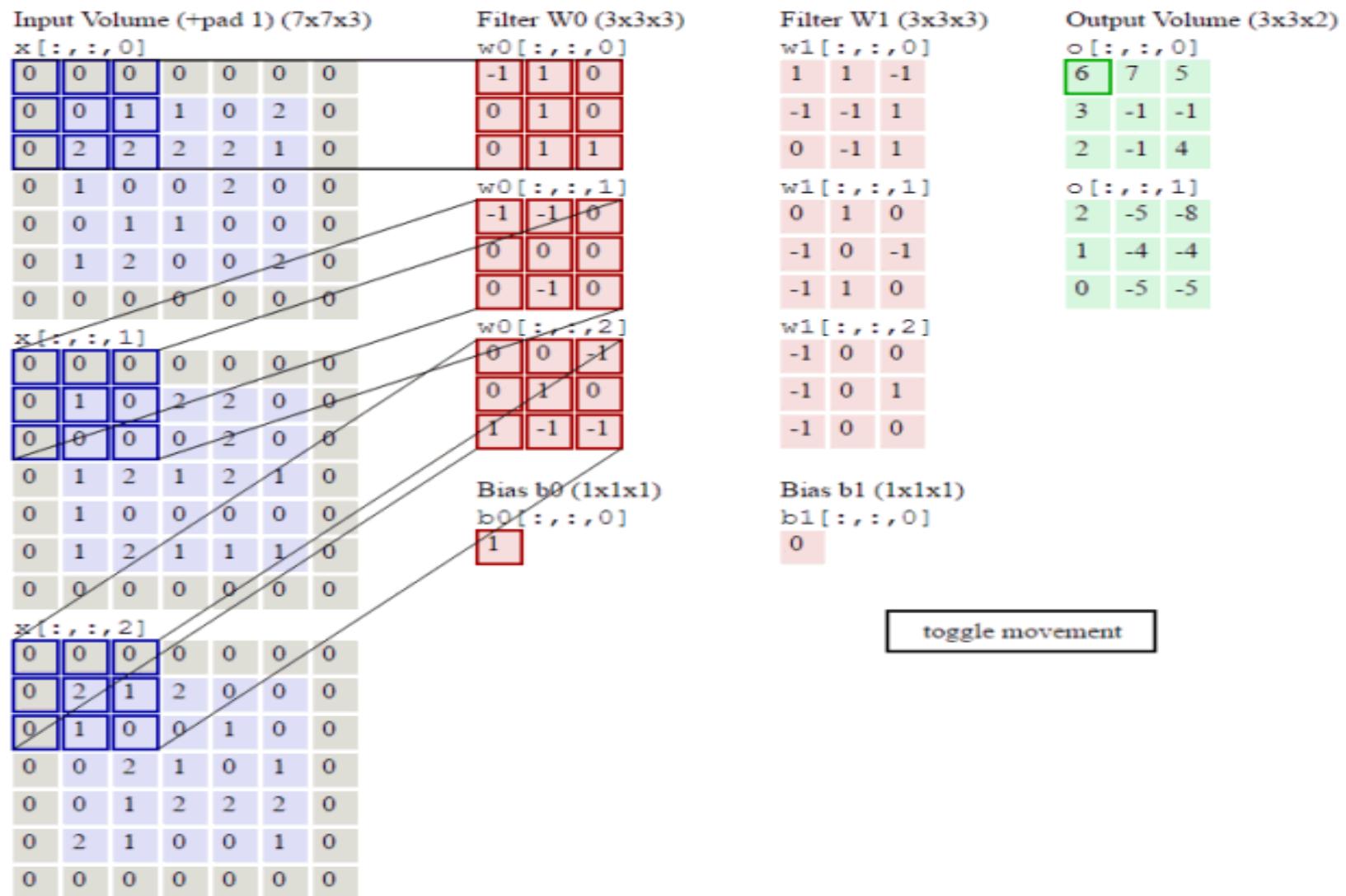
- ❖ 图像识别
- ❖ 图像定位
- ❖ 目标检测
- ❖ 实例分割



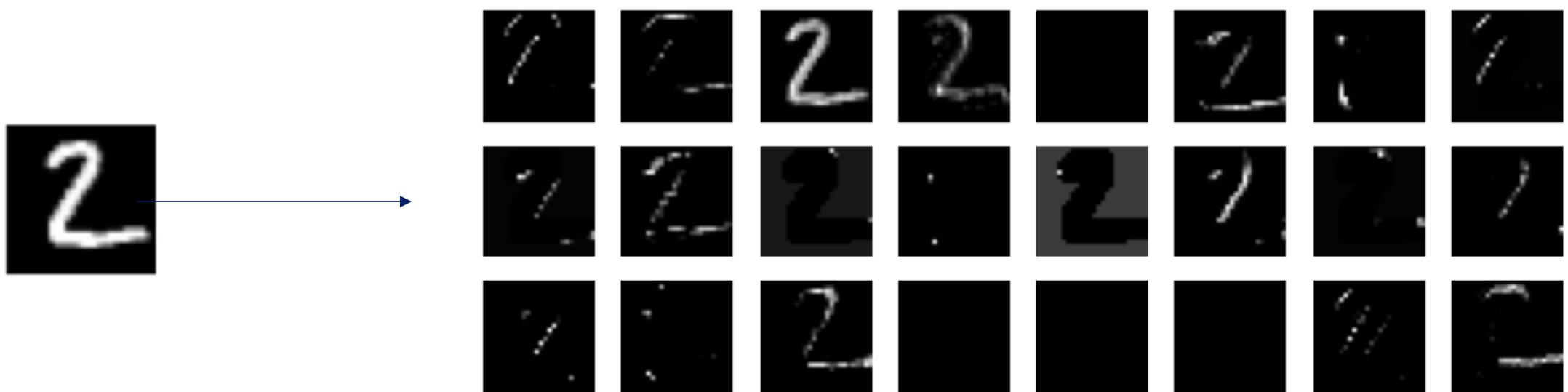
卷积神经网络(CNN)



卷积层



卷积后的特征图

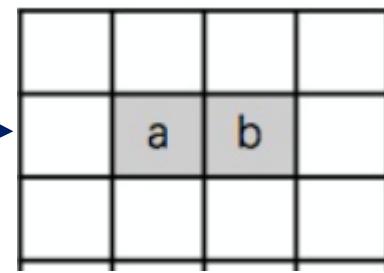
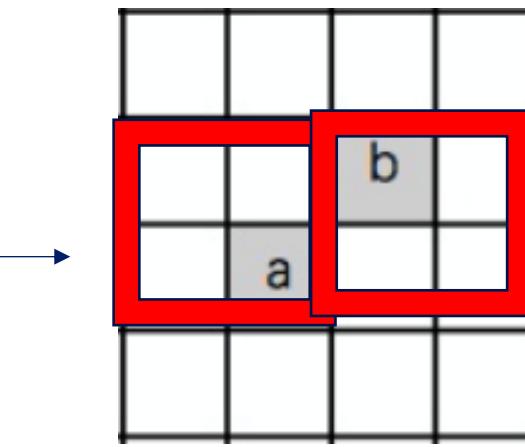
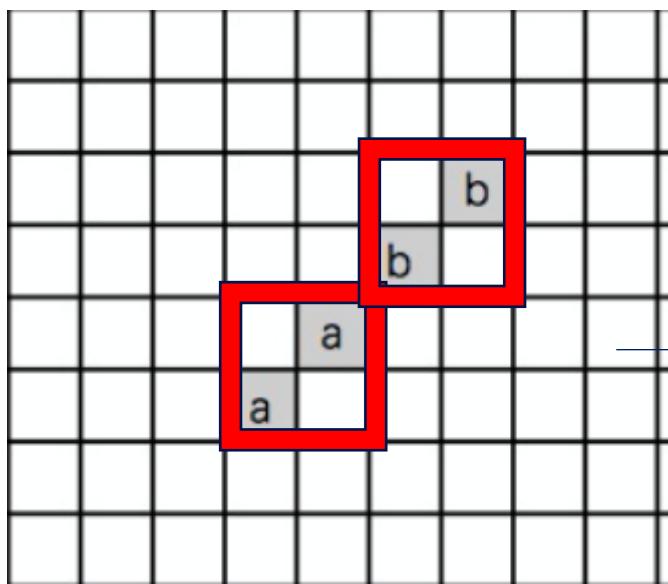


池化层

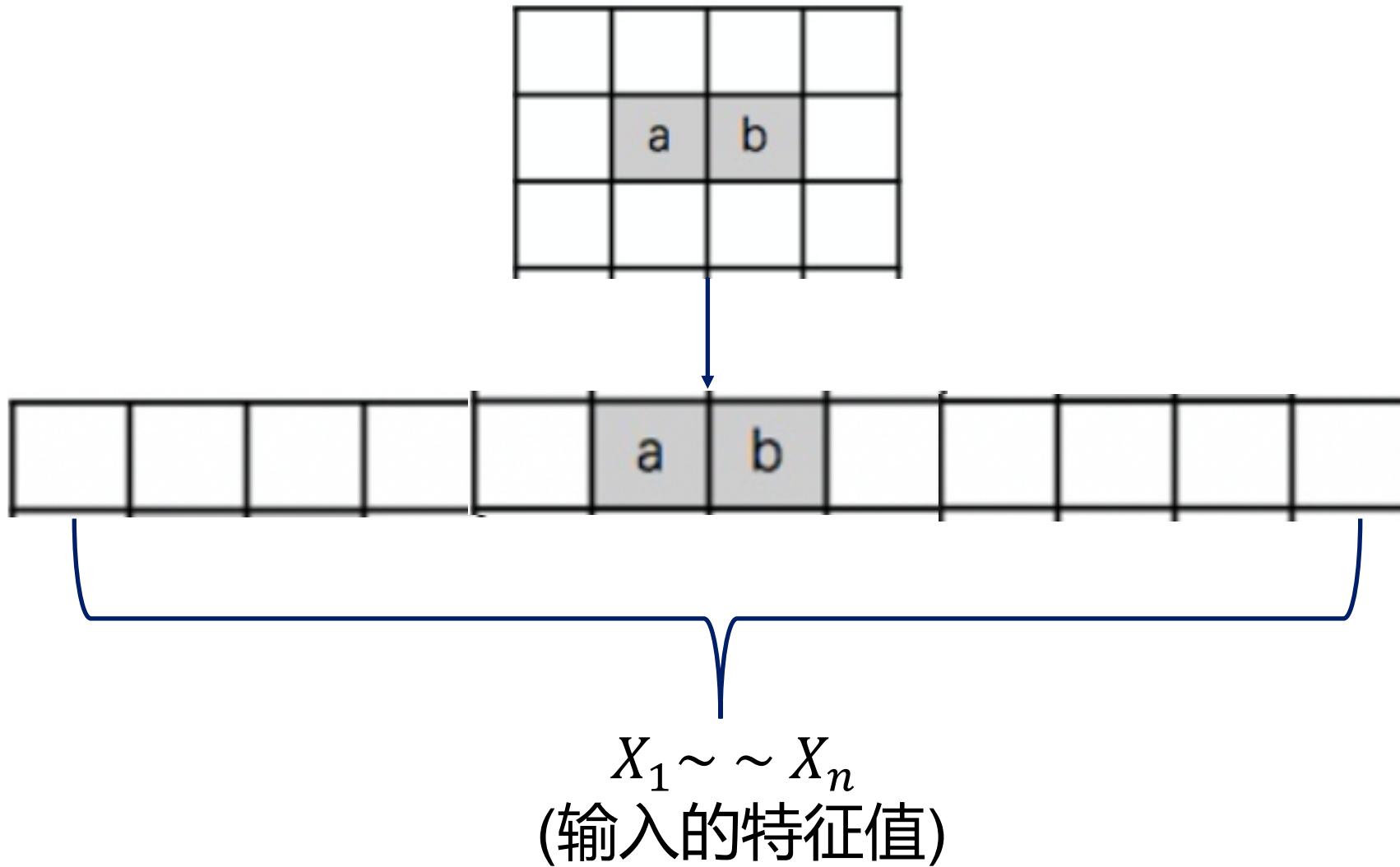
1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Max pooling

6	8
3	4



全连接层



图像识别

Example:

<http://10.206.67.123:8888/notebooks/JimXie/excise/example/%E4%BF%A1%E5%8F%B7%E7%81%AF%E6%A3%80%E6%B5%8B.ipynb>

练习:

<http://10.206.67.123:8888/notebooks/JimXie/excise/%E6%89%8B%E5%86%99%E5%AD%97%E4%BD%93%E8%AF%86%E5%88%AB.ipynb>

自由练习

① 从excise中任选一个，copy至个人目录

(<http://10.206.67.123:8888/tree/JimXie/excise>)

② 参照example code，添加模型相关的代码

(<http://10.206.67.123:8888/tree/JimXie/excise/example>)

印度新冠预测



???在这里添加你的code ???

参照美国新冠预测的模型定义部分，定义和训练模型，对感染人数进行预测

(上面的train_data和test_data分别为训练，测试集)

手写字体识别



???(在下面添加你的code)???

参照信号灯识别model定义部分，定义CNN模型m_model

注意：

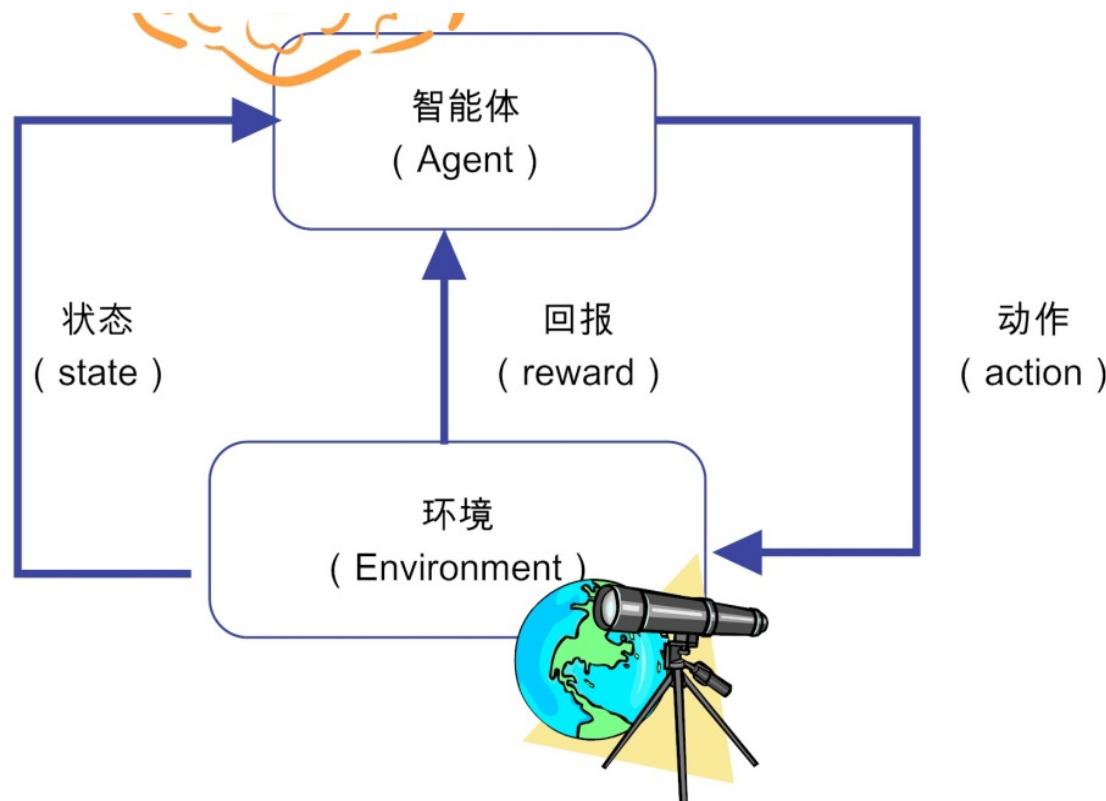
1. 图像大小input_shape为(28, 28, 1)
2. 分类为0--9，共10个类别，n_classes为10

前沿技术介绍

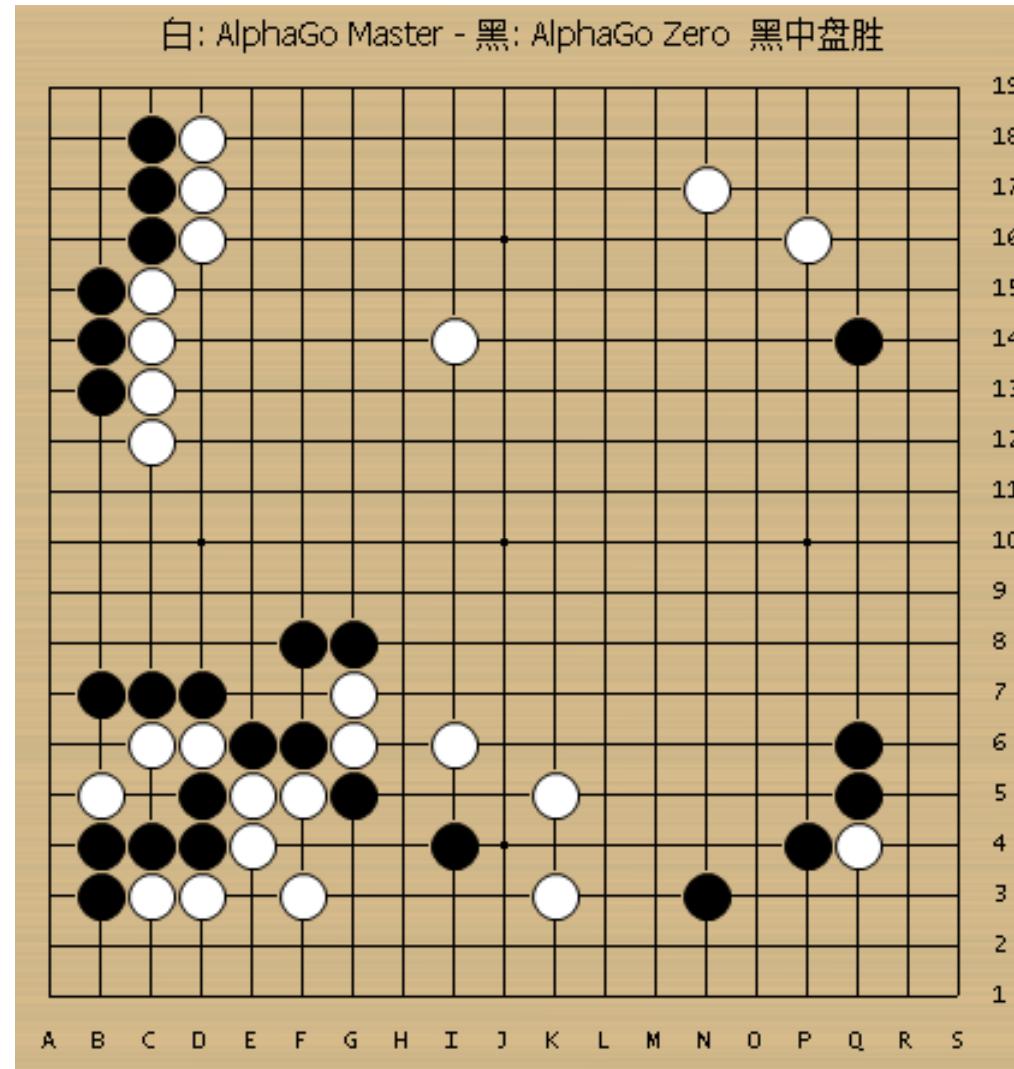
- ❖ 强化学习
- ❖ 对抗网络
- ❖ 元学习
- ❖

强化学习

学会在不同的状态下该采取什么动作 (Policy)



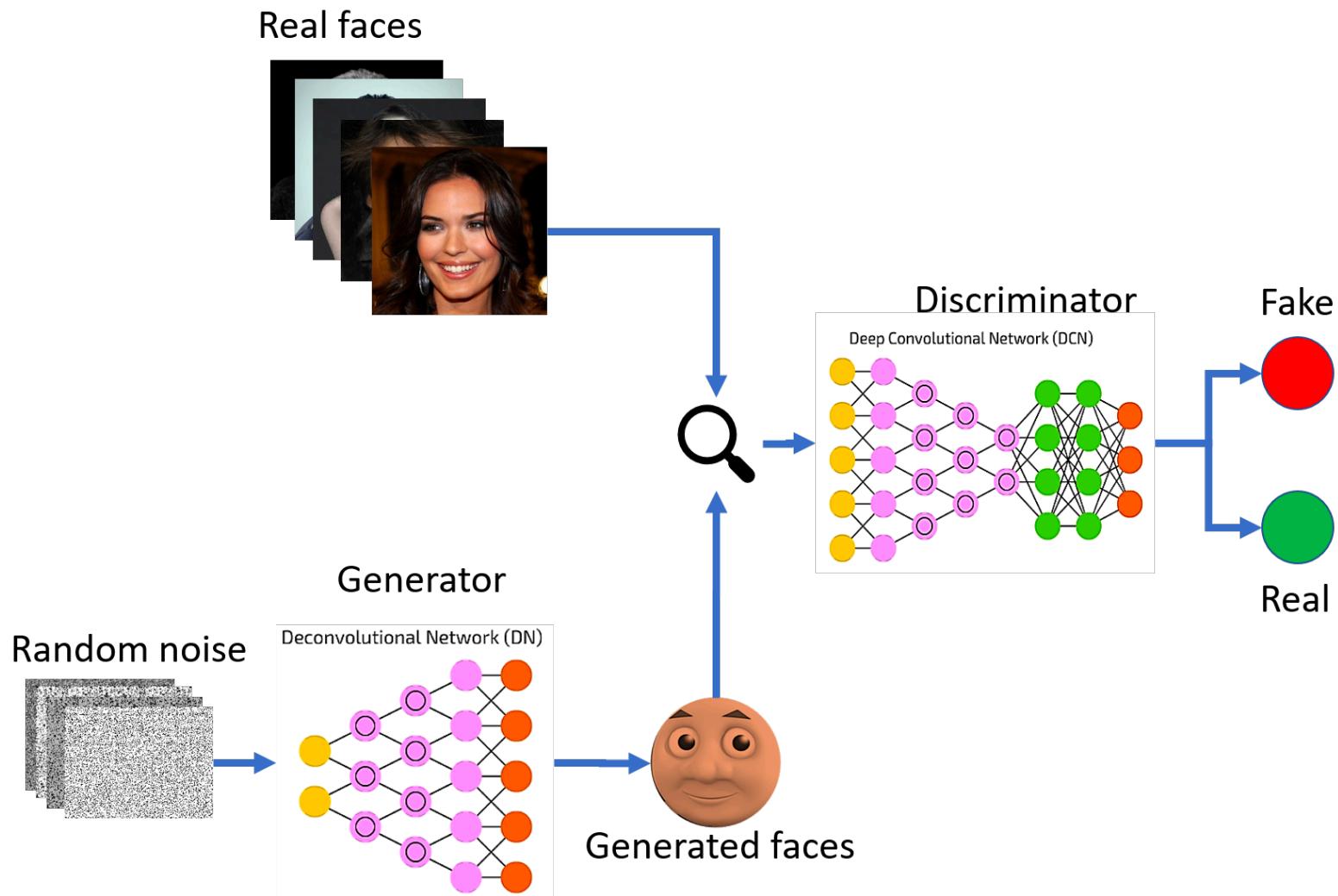
强化学习案例



GAN

https://v.douyin.com/JHGupXt/?utm_campaign=client_share&ap=p=aweme&utm_medium=ios&tt_from=copy&utm_source=copy

GAN



GAN案例



元学习

学习如何学习

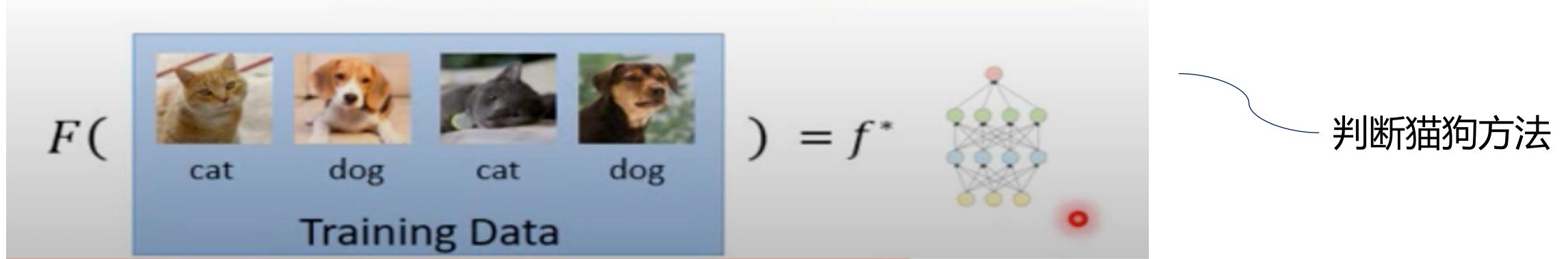


知道如何比较两个图像即可

元学习

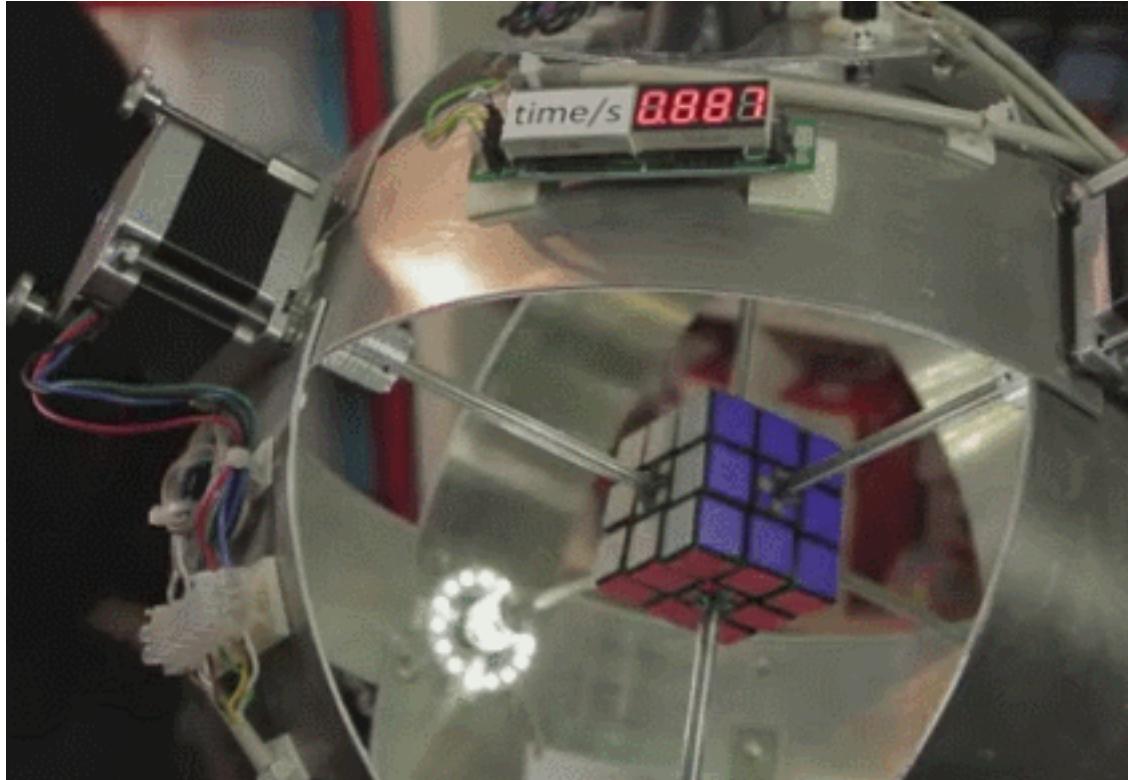


Feature学习（输出一个类别）

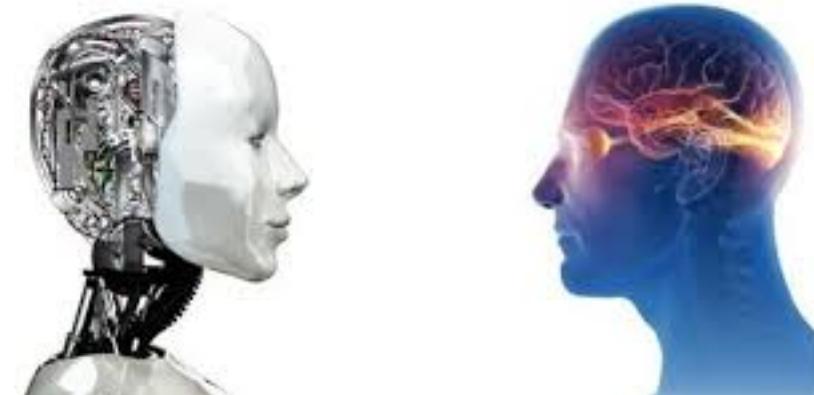
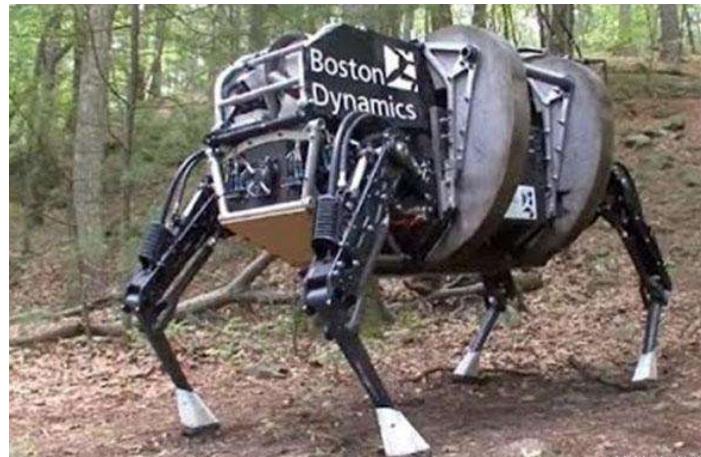


元学习（输出一个模型）

元学习案例



AI应用



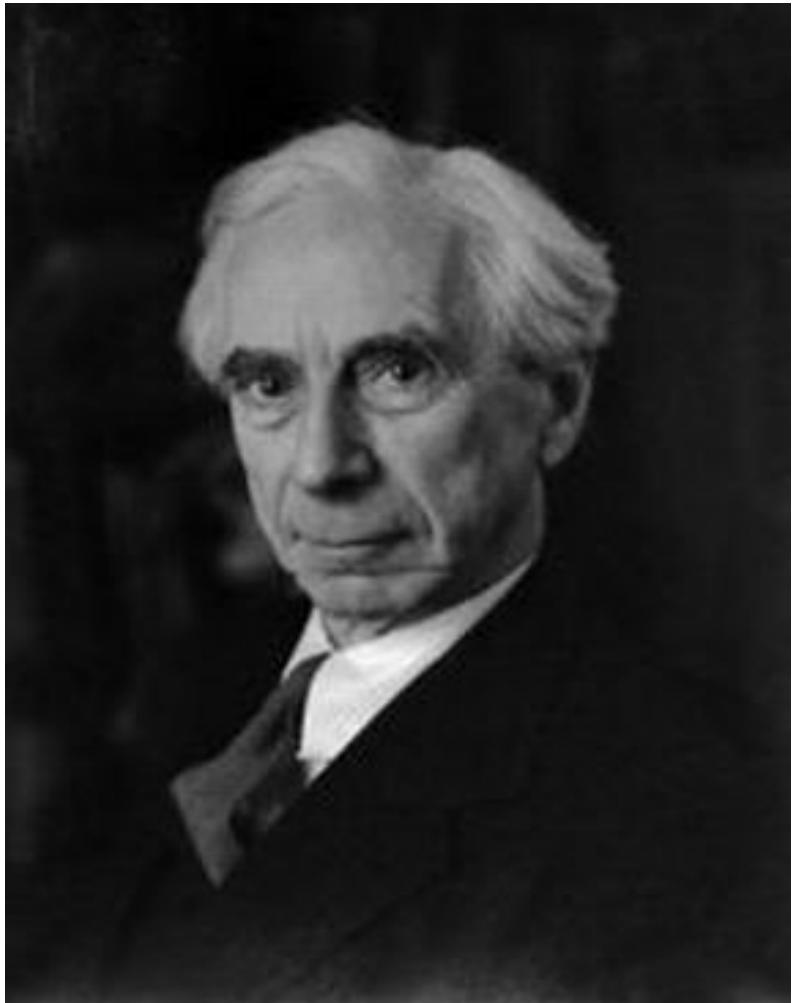
总结

- 课程回顾与建议
 - 功利驱动
 - 兴趣驱动
- 知识(智慧)到底是如何产生的？

知识(智慧)到底是如何产生的？

- 通过计算产生
- 通过回忆产生
- 通过范畴组合产生
-

❖ 结束语



也许根本就不存在智慧这种东西，
所谓的智慧，
只不过是看起来极其精炼的愚蠢。

--罗素

Thanks

