



论道自动机器学习与可解释机器学习

AI Time 第二期

AI Time rallies researchers specializing in the field of artificial intelligence. The goal of AI Time is to explore the essence of AI foundations, theories, algorithms, and applications, with an emphasis to encourage debate on AI-related topics.



■ 嘉宾



Philip S. Yu
美国伊利诺伊大学
芝加哥分校(UIC)特
聘主任教授



Qiaozhu Mei
密歇根大学信息学
院与计算机系终身
教授



王立威
北京大学教授



李兴建
百度高级研究员

■ 主持人



李文珏
中科创星投资总监
CCF YOCSEF学术秘书



刘譞哲
北京大学副教授
CCF青年科学家

■ 今日辩题



AI Time



自动机器学习
AutoML

VS

可解释机器学习
Explainable ML

Philip S. Yu (俞士纶) Qiaozhu Mei 王立威 李兴建

时间：2019年6月19日15:00-17:00

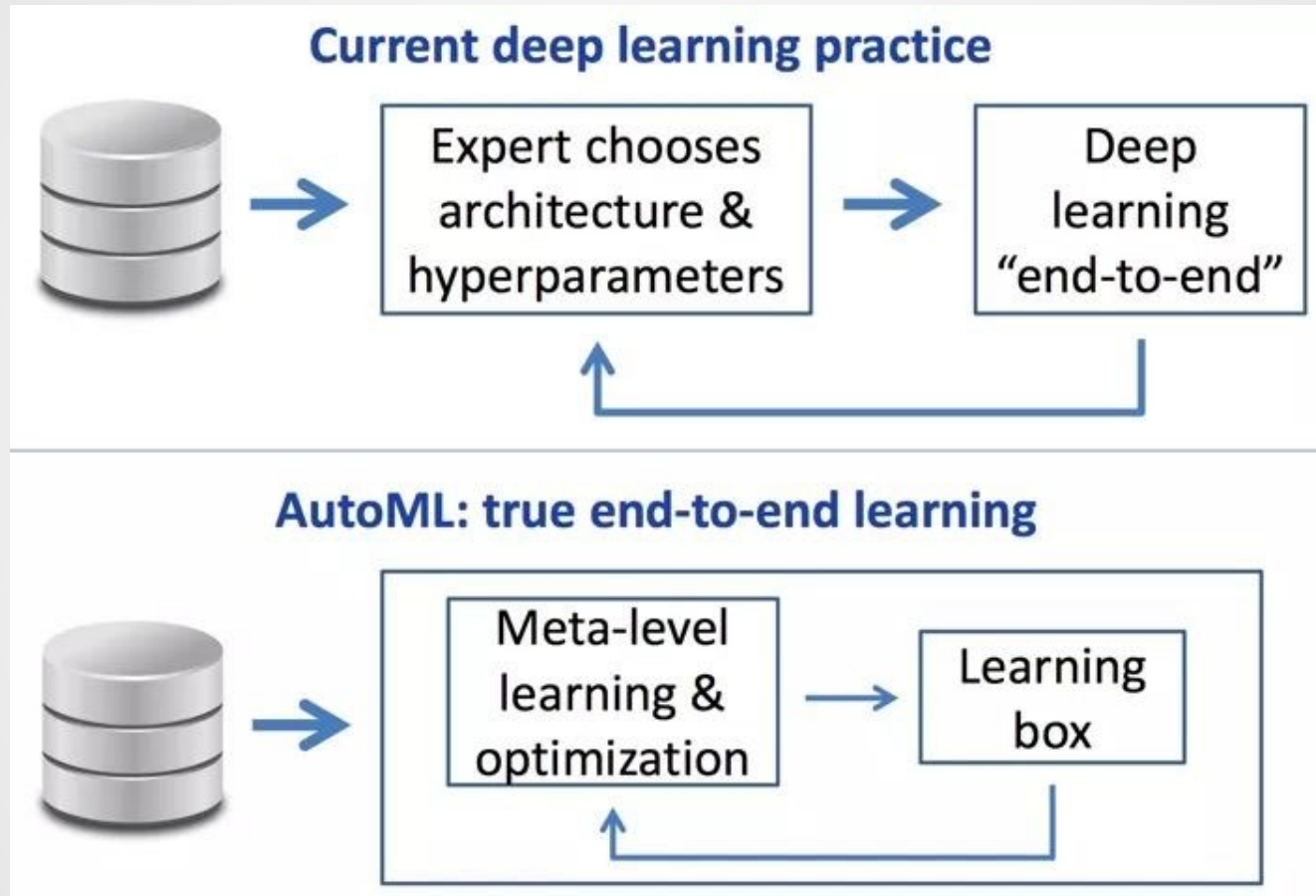
地点：海淀区中关村东路搜狐大厦二楼1911（清华科技园）



欢迎进群讨论，加好友注明：AI-Time



■ 自动机器学习 (AutoML)



企业界的最爱



■ AutoML形式化定义

Definition: Combined Algorithm Selection and Hyperparameter Optimization (CASH)

Let

- $\mathcal{A} = \{A^{(1)}, \dots, A^{(n)}\}$ be a set of algorithms
- $\Lambda^{(i)}$ denote the hyperparameter space of $A^{(i)}$, for $i = 1, \dots, n$
- $\mathcal{L}(A_{\lambda}^{(i)}, D_{train}, D_{valid})$ denote the loss of $A^{(i)}$, using $\lambda \in \Lambda^{(i)}$ trained on D_{train} and evaluated on D_{valid} .

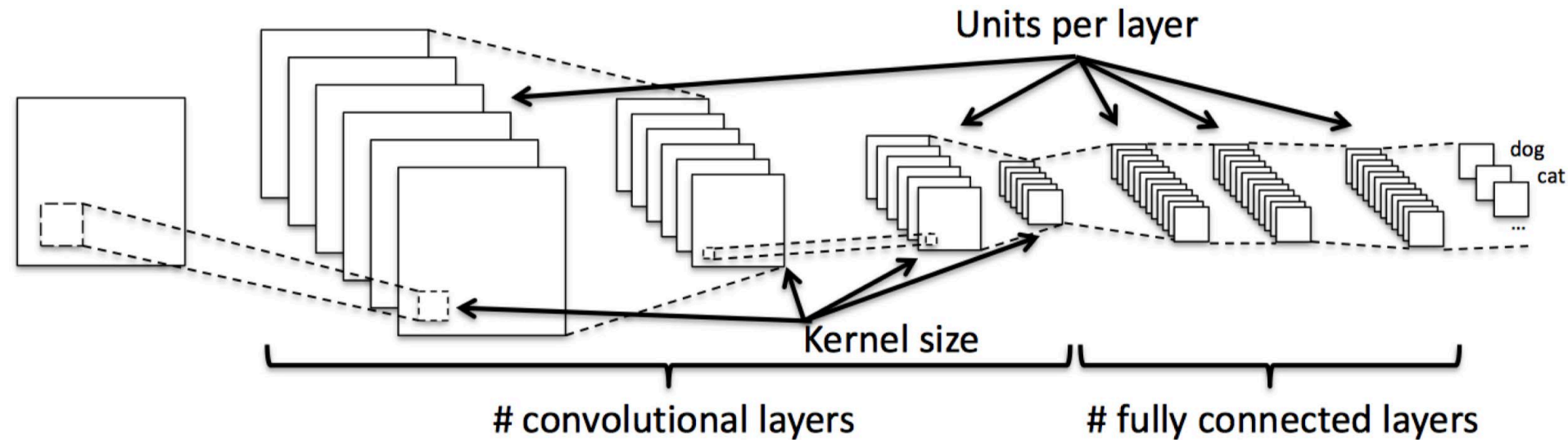
The Combined Algorithm Selection and Hyperparameter Optimization (CASH) problem is to find a combination of algorithm $A^* = A^{(i)}$ and hyperparameter configuration $\lambda^* \in \Lambda^{(i)}$ that minimizes this loss:

$$A_{\lambda^*}^* \in \arg \min_{A^{(i)} \in \mathcal{A}, \lambda \in \Lambda^{(i)}} \mathcal{L}(A_{\lambda}^{(i)}, D_{train}, D_{valid})$$



深度学习问题——超参数

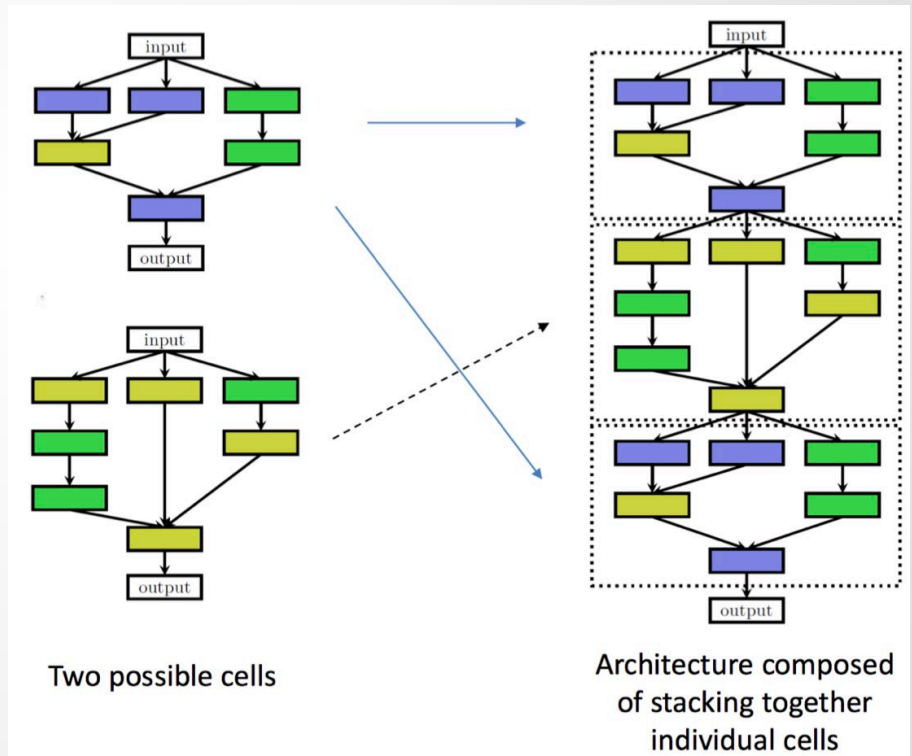
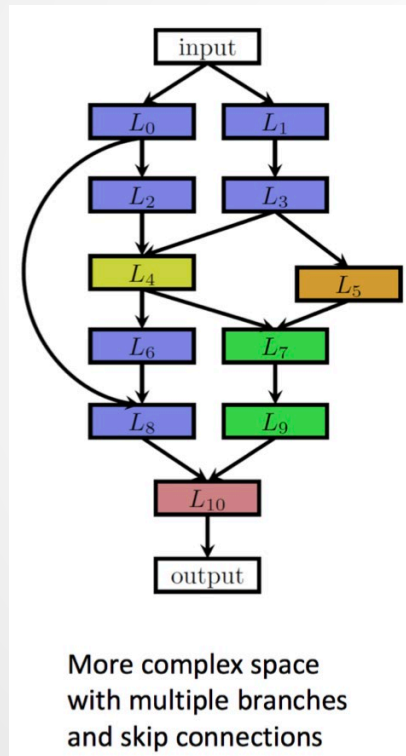
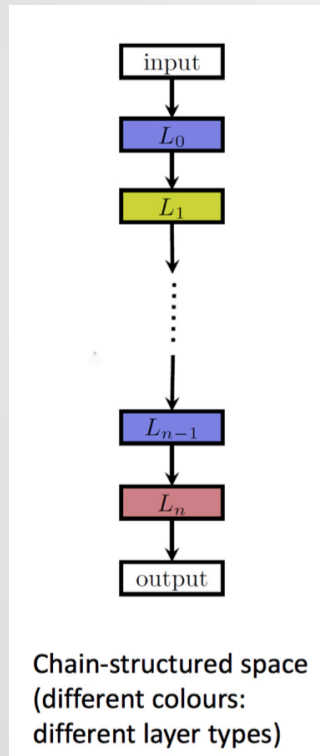
深度学习性能对超参数敏感，需要进行大量调优工作





深度学习问题—模型框架

深度学习框架灵活，需要自动寻找最优框架



Neural Architecture Search

■ 可解释机器学习 (XML)



x

“panda”

57.7% confidence

+ .007 ×



$\text{sign}(\nabla_x J(\theta, x, y))$

“nematode”

8.2% confidence

=



$x +$

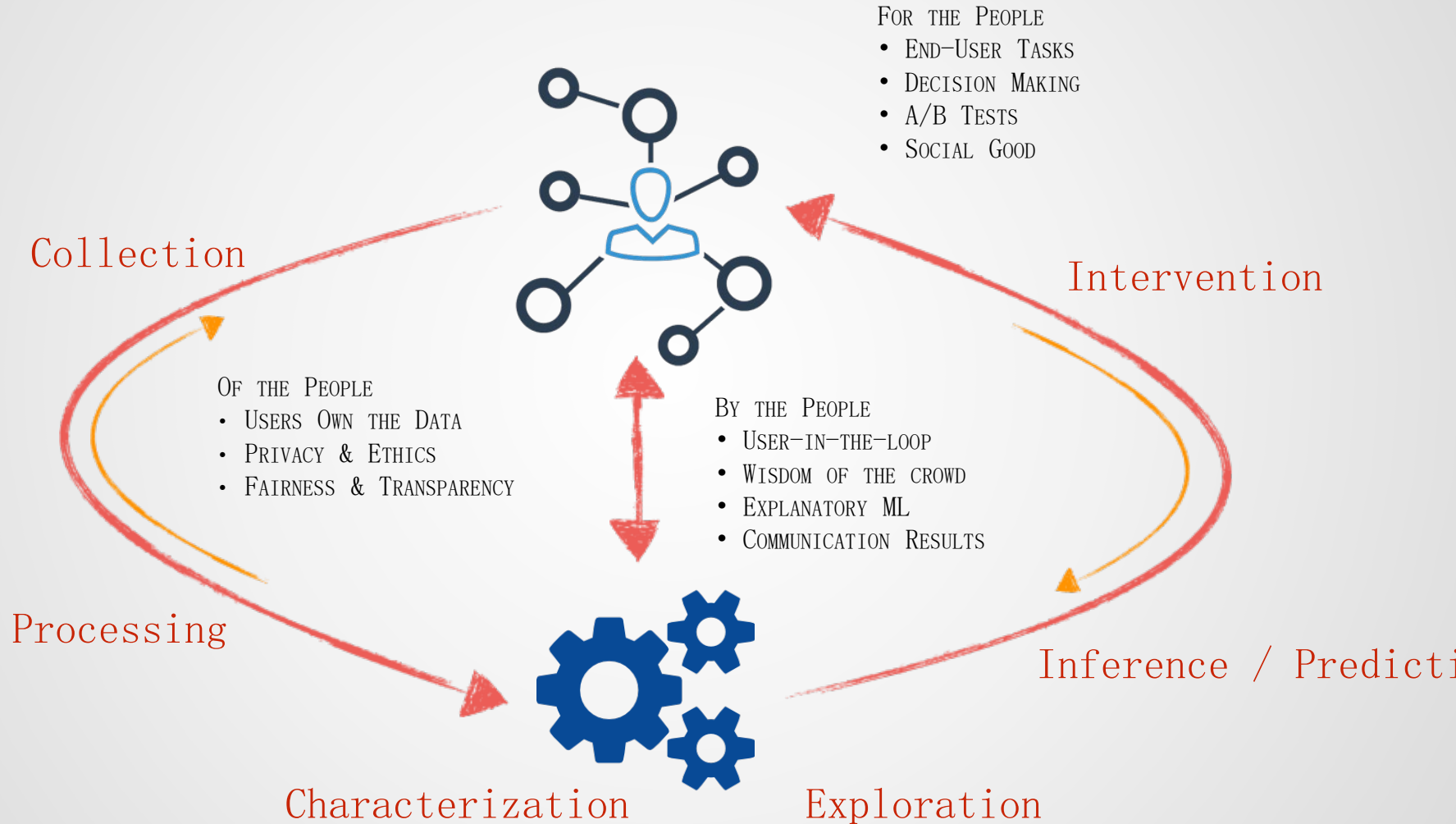
$\epsilon \text{sign}(\nabla_x J(\theta, x, y))$

“gibbon”

99.3 % confidence

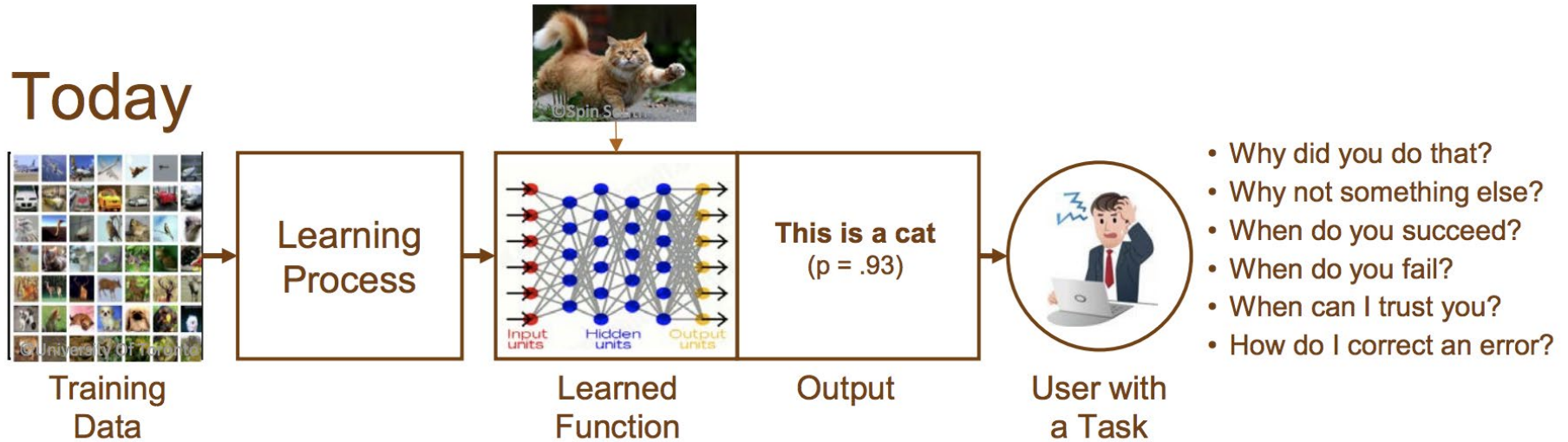
不可解释的危险性

End-to-end Data Science (Data-driven decision making)

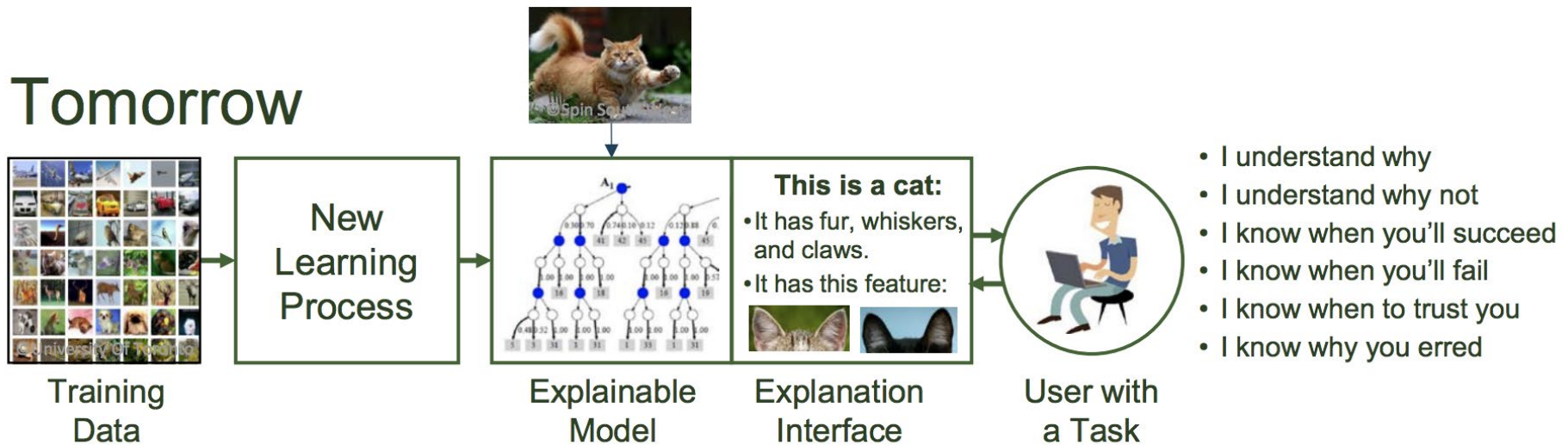


XML的愿景

Today



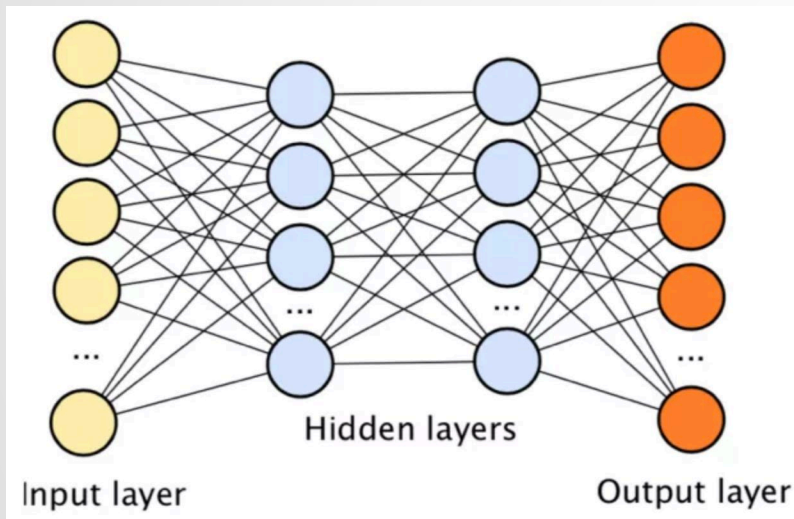
Tomorrow



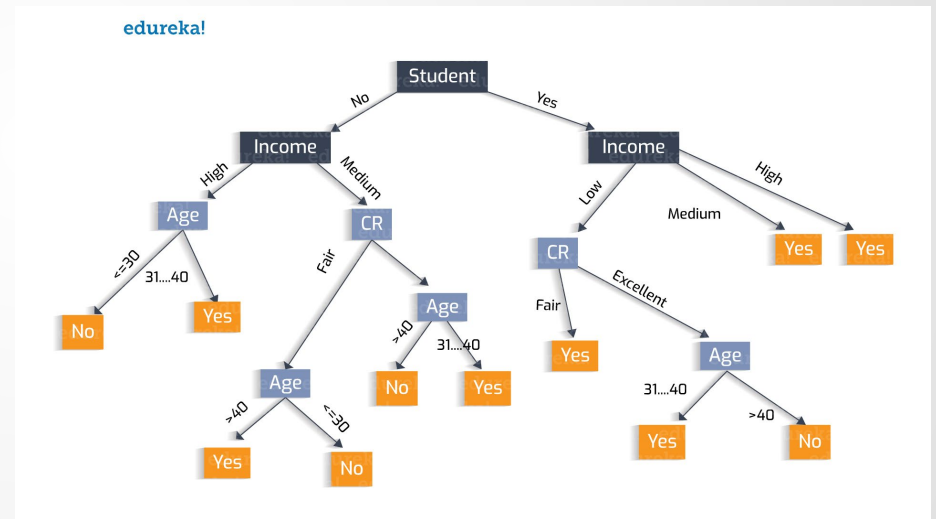


AutoML vs XML

电话诈骗自动检测



85%



78%



■ XML面临的问题

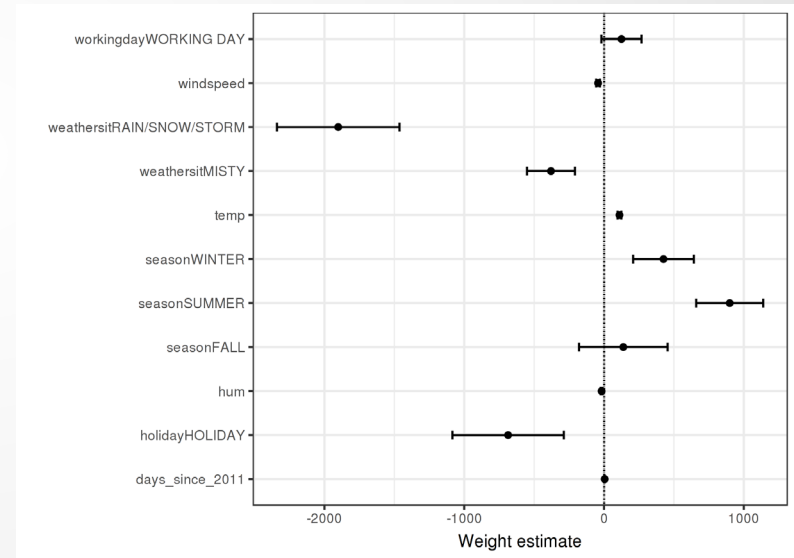
全局可解释性 = 诠释 (神经网络(输入空间, 输出空间)) $\cap$ 人类理解能力

Who 谁向谁解释?

What 解释什么?

How 如何解释?

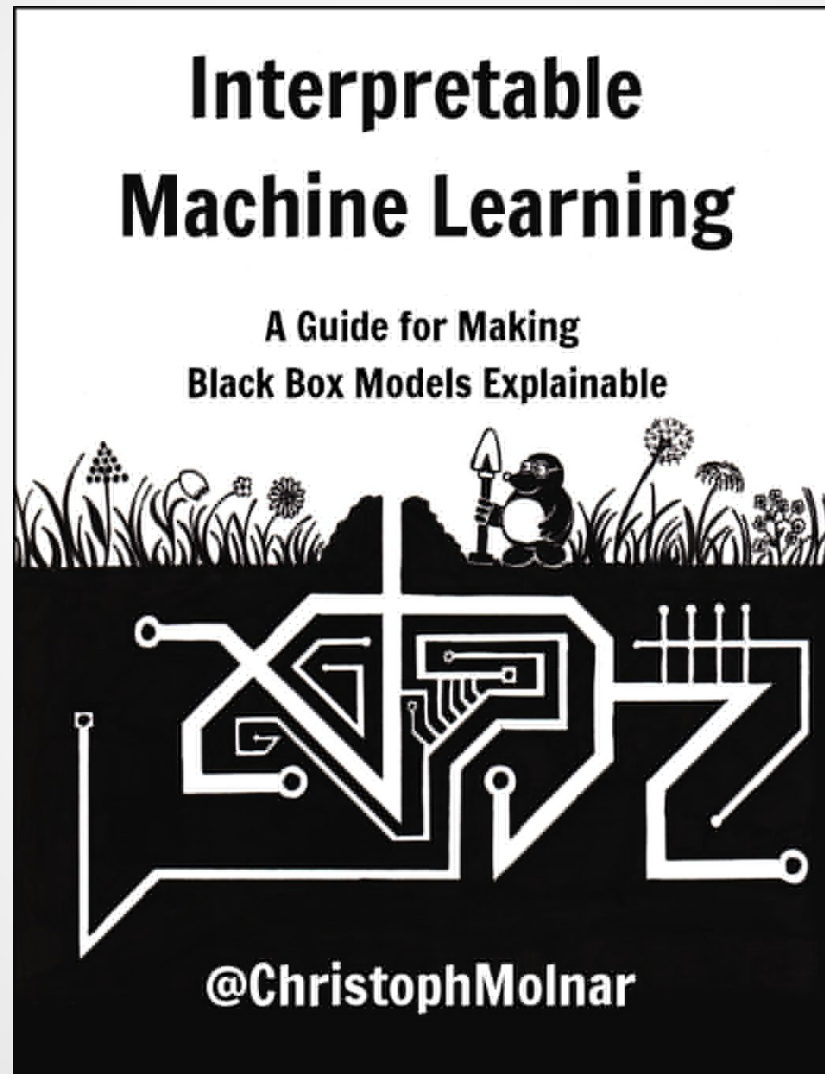
When 何时解释?



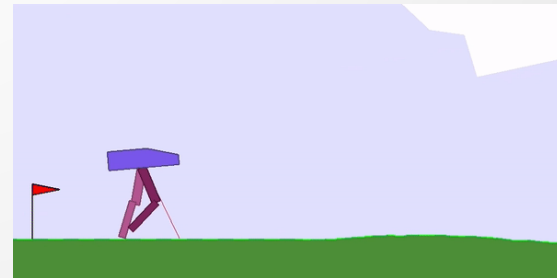
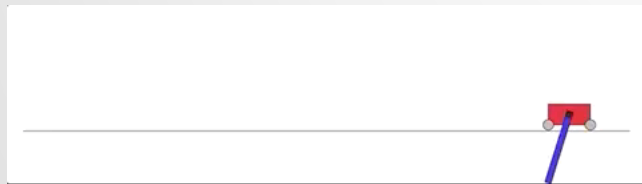
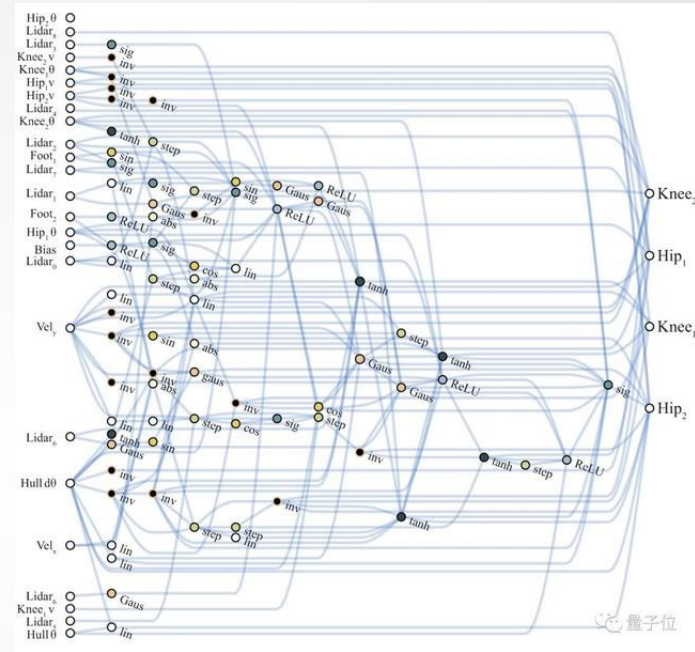
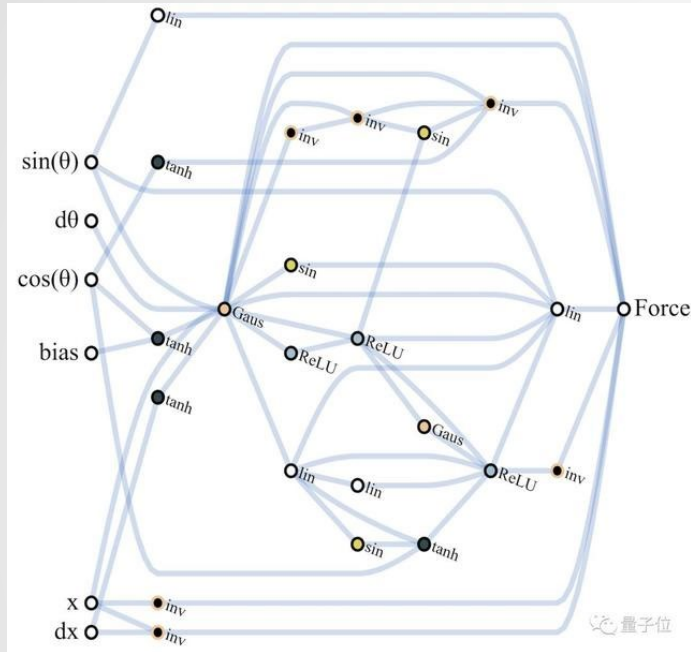
Model-Specific Methods vs. Model-Agnostic Methods
vs. Example-based Explanation



■ XML推荐书目



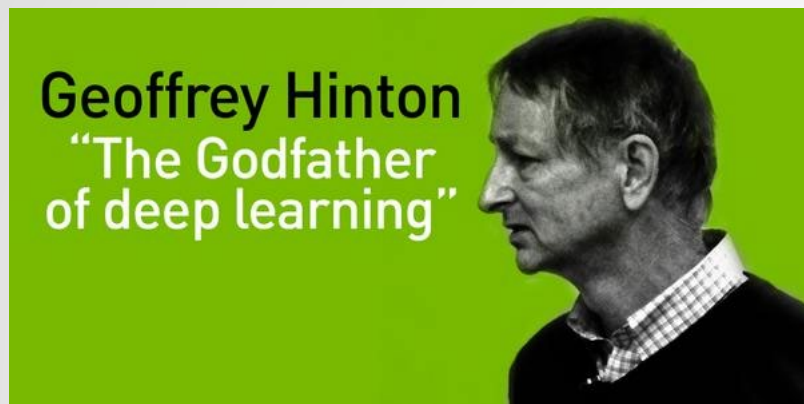
如何实现AutoML



Hyperparameter Optimization vs. Neural Architecture Search
vs. Meta Learning

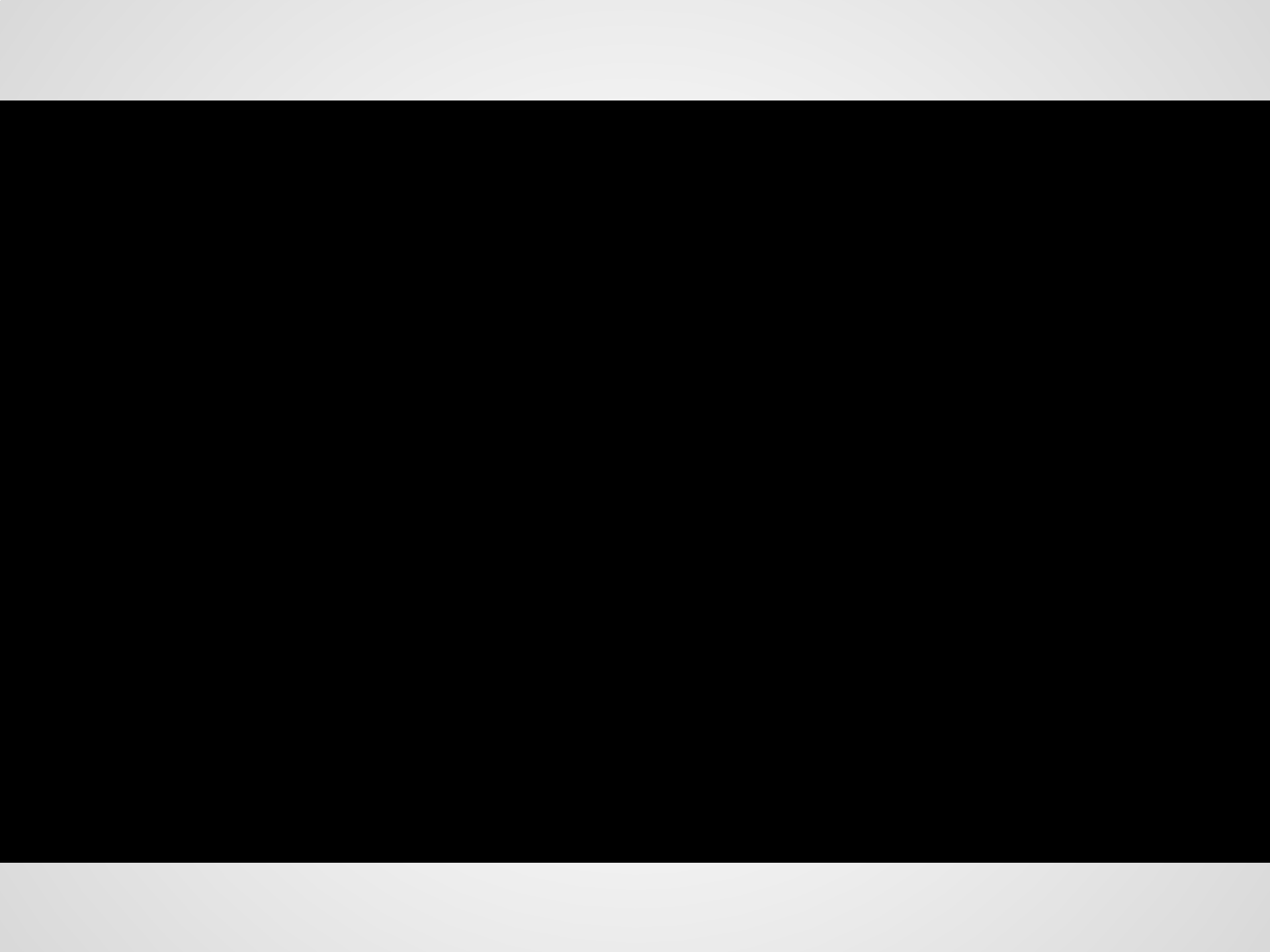


■ 来自Hinton的声音



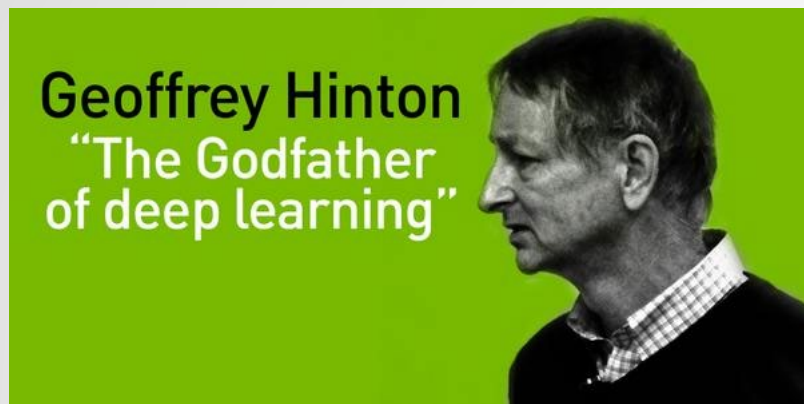
纠结深度学习（可与不可）解释性问题是一个伪命题。为什么一定要存在识别数字的理论才能证明我们擅长识别数字？难道非要通透骑车每一个细节的物理力学，才能证明自己会骑车？

不是神经网络需要理论解释，而是人类克制不住自己讲故事的冲动；理论再合理也只是主观判断，并不能帮助我们理解为什么。





■ 来自Hinton的声音



纠结深度学习（可与不可）解释性问题是一个伪命题。为什么一定要存在识别数字的理论才能证明我们擅长识别数字？难道非要通透骑车每一个细节的物理力学，才能证明自己会骑车？

不是神经网络需要理论解释，而是人类克制不住自己讲故事的冲动；理论再合理也只是主观判断，并不能帮助我们理解为什么。



■ 各方观点

ACL2019新论文，痛批“不计成本追逐丁点提升”的深度学习研究方法

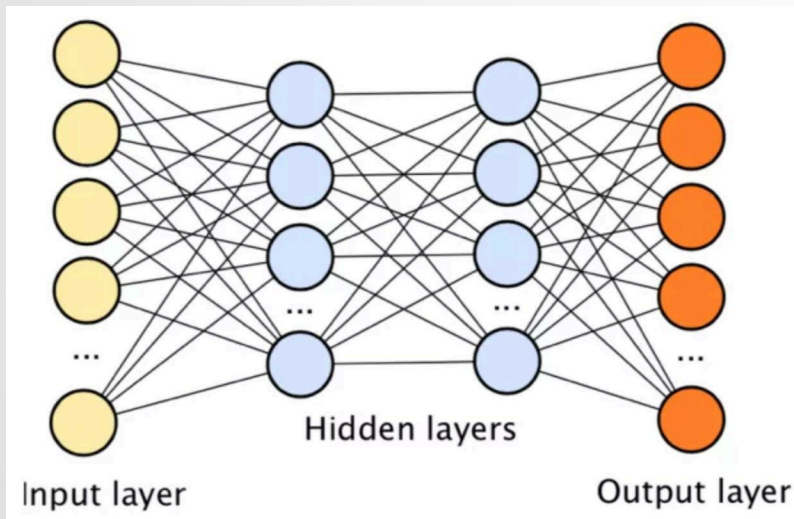
某些新的模型/训练方式取得了最优结果，但是进展微小，副作用却是不成比例的计算量/碳排放增加。

Model	Hardware	Power (W)	Hours	kWh·PUE	CO ₂ e	Cloud compute cost
Transformer _{base}	P100x8	1415.78	12	27	26	\$41–\$140
Transformer _{big}	P100x8	1515.43	84	201	192	\$289–\$981
ELMo	P100x3	517.66	336	275	262	\$433–\$1472
BERT _{base}	V100x64	12,041.51	79	1507	1438	\$3751–\$12,571
BERT _{base}	TPUv2x16	—	96	—	—	\$2074–\$6912
NAS	P100x8	1515.43	274,120	656,347	626,155	\$942,973–\$3,201,722
NAS	TPUv2x1	—	32,623	—	—	\$44,055–\$146,848
GPT-2	TPUv3x32	—	168	—	—	\$12,902–\$43,008

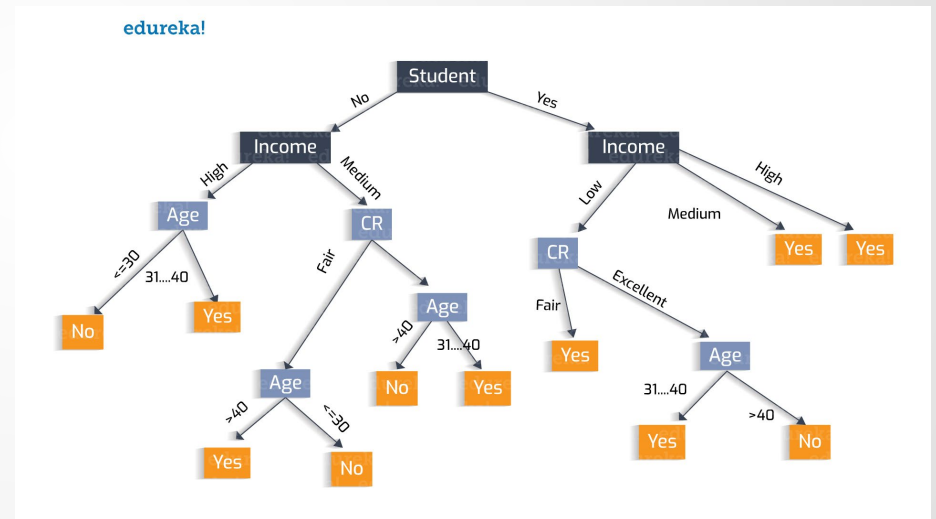


■ Back to AutoML vs XML

电话诈骗自动检测



85%



78%

■ It Is Your Turn



AI Time



自动机器学习
AutoML

VS

可解释机器学习
Explainable ML

Philip S. Yu (俞士纶) Qiaozhu Mei 王立威 李兴建

时间：2019年6月19日15:00-17:00

地点：海淀区中关村东路搜狐大厦二楼1911（清华科技园）



■ AutoML未来之路是什么？

NAS without defining the search space? True “AutoML”

■ XML未来之路是什么？

Analyzing the model without any knowledge about the model

■ 两者是否可以结合？

Can machine learning be automated and, with interpretability?



谢谢大家

