

云智未来⁹th

第九届中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2017



人工智能驱动Fintech

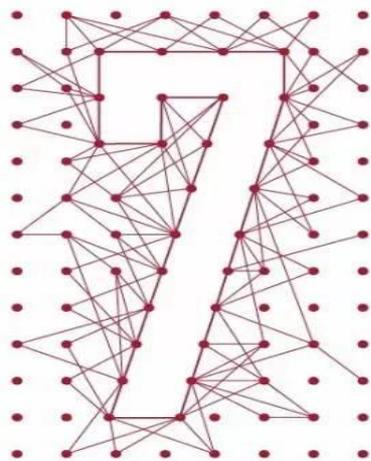
天云大数据副总裁：李从武

获取机器智能像读书一样简单

Get machine intelligence as simple as reading

人工智能对金融市场的影响

Euromoney Institutional Investor Thought Leadership 从全世界的金融机构中调查了424 名高级管理人员。其中，超过四分之一（26%）的受访者在资产管理公司工作，16% 的受访者在投资银行工作，其余的在银行、保险公司、对冲基金和经纪公司工作。大多数的受访者（57%）在他们的公司拥有 C 级职位；剩下的其他人是不同领域的高级管理人员，包括数据、技术、法律、合规等。此外，调查对象中大型、中型、小型企业都有涉及，51% 的公司拥有 100 名或更多的员工。最后，调查样本是全球范围的：三分之一的受访者来自欧洲，三分之一的受访者来自北美，16% 来自亚洲，剩下的受访者来自拉丁美洲、中东和非洲。



7 IN 10

十分之七的人相信，十五年内，人工智能会给他们自己的工作带来完全或很大程度上的变化

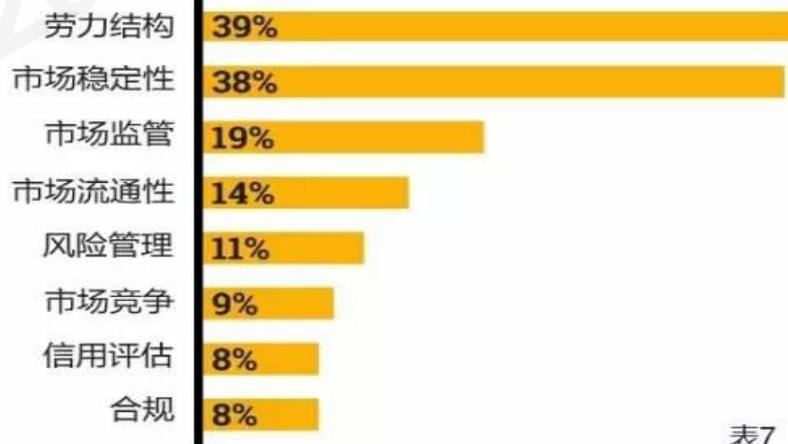


表7

FinTech

IT → DT

规则流程驱动到智能数据驱动

机器的角色，从快速思维到智能思维



Saeed Amen
The Thalesians
公司

机器学习的好处在于它能让交易者发现那些不易察觉的关系，因此不用再和其他市场参与者进行贴身肉搏去争夺这些交易机会。



Peter Havez
RavenPack 公司

这场机器学习革命，是从急剧扩大的可用数据和信息中识别复杂的模式，从而做出任何视角来看都是最优秀的决策。该市场正从做到更快转向做到更智能。

在过去的几十年里，计算机被广泛用于完成自动化任务，后者往往是被清晰的规则和算法描述的。如今，机器学习技术允许我们在难以精确描述规则的边界内完成同样的任务。

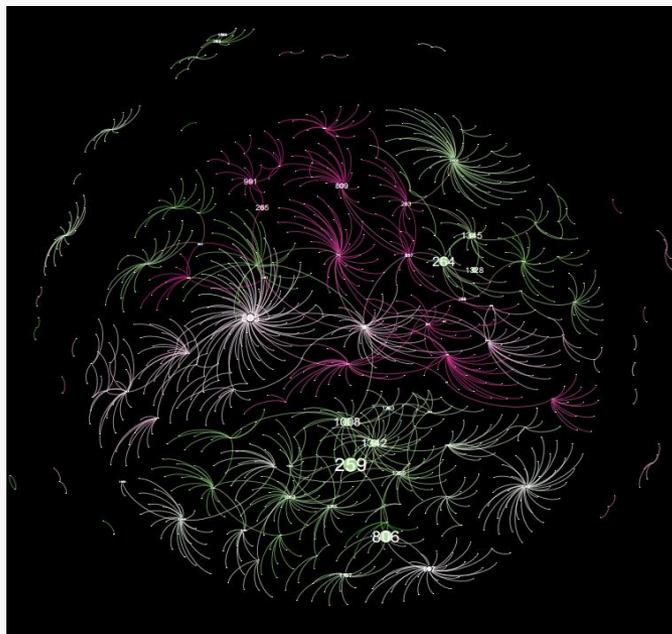
—— 来源于亚马逊创始人杰夫·贝索斯（Jeff Bezos）2017年度致股东的公开信

人类习惯抽象认知



350年前，科学巨人牛顿用三个简约表达式，揭示了自然规律，客观抽象了传统参照系中的世界。

人类用简单来抽象世界 机器用复杂来量化世界

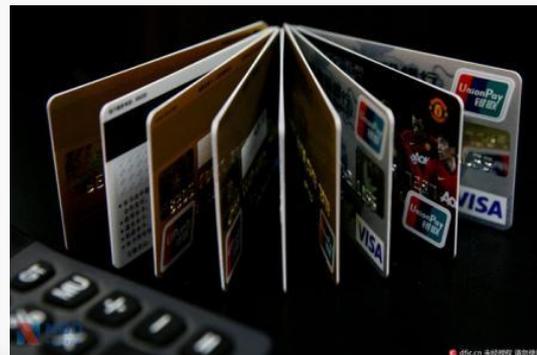
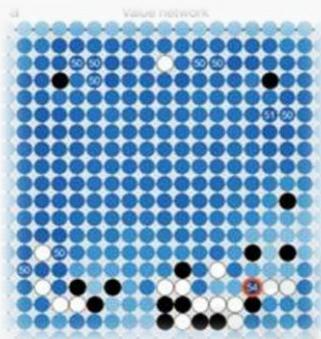
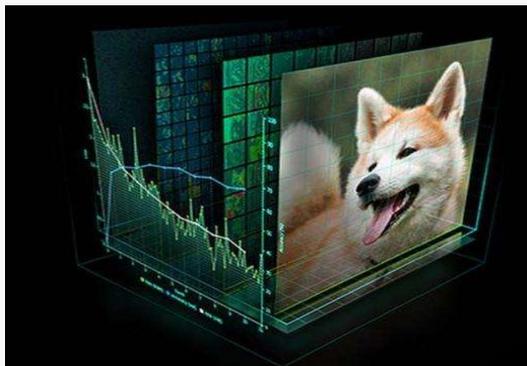


某互金理财产品的营销获客传播的复杂网络

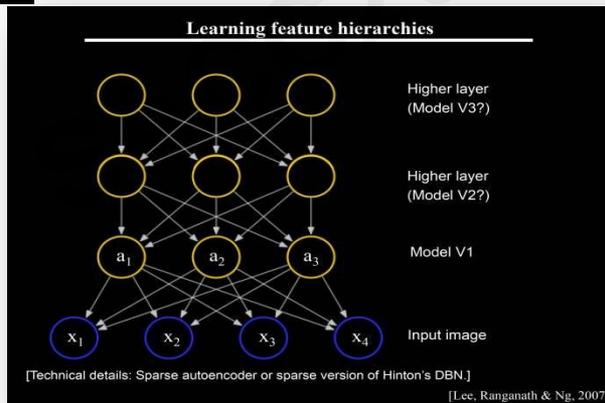


某保险公司的代理人成功销售的获客网络

人工智能依靠质朴的数学和超强的计算能力，还原了世界的复杂性。



如何用RGB像素色差等信号体系描述图片内容？



如何制订动态防范的欺诈规则？

深度学习的特征建立过程，就是协助我们对复杂问题描述的精确量化。

今天是笛卡尔的时代，等待着艾萨克牛顿的到来

笛卡尔为之后的牛顿准备了一个坐标系，使 $F=ma$ 的推演成为经典。

今天我们可能不知道目的地像什么，单我们现在知道如何绘制一张地图。



Yann LeCun: 1960 - Current
Father of Convolutional Neural Networks



René Descartes: 1596 - 1650
Father of Cartesian Geometry

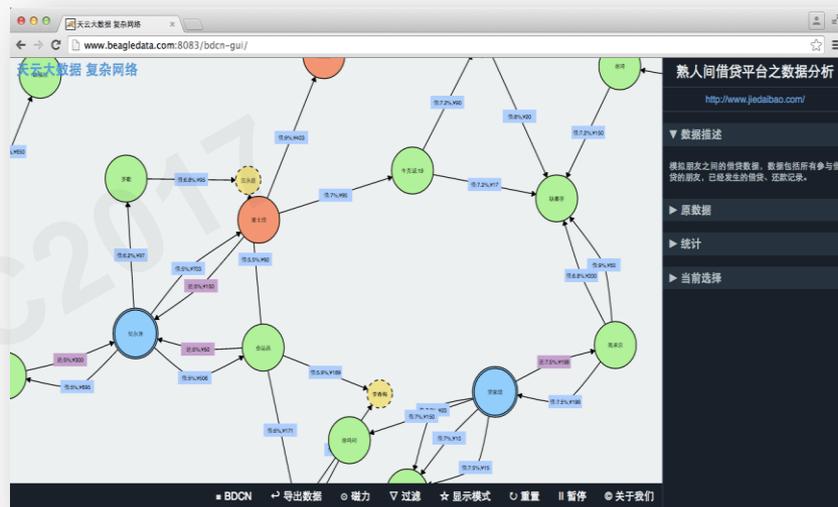
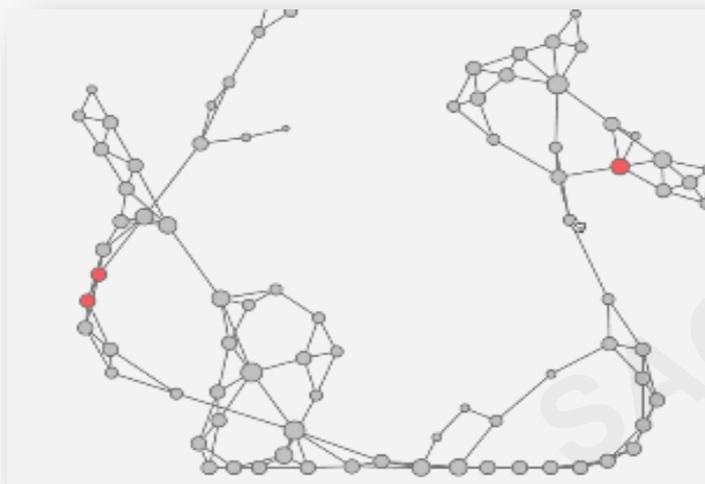
- 物理学上牛顿之前的日子。许多聪明的科学家能够使用数学来预测物体的运动，曾经聪明的笛卡尔教会我们如何将我们的物理思维考虑到坐标系统中。Yann LeCun（深度学习之父其中之一）就是一个现代的笛卡尔，开创性的工作是指日可待。ConvNets思想框架就像是一个必备的坐标系统。

开放的 B 价值远大于性感的 A

年份	人工智能的突破	年份	数据集（首次可用）	年份	算法（首次提出）
1994	人类自发的语音识别	1991	华尔街日报文章与其他文本	1984	隐藏的M arkov模型
1997	IBM 深蓝打败G any Kasparov	1991	70万国际象棋大师赛，又称“扩展的书”	1983	N egascout 规划算法
2005	谷歌阿拉伯语-中文-英语翻译软件	2005	1.8万亿符号的搜索，基于谷歌网站与新闻网页	1988	统计学机器翻译算法
2011	IBM 沃森机器人成为世界“危险”节目冠军	2010	860万文件的上传，基于维基百科，维基词典引用以及古登堡计划	1991	专家混合系统算法
2014	谷歌的危机谷歌网站分类计划	2010	150万标签的图片1000项目分类的视觉数据库	1989	卷积神经网络算法
2015	谷歌deepm ind游戏水平达到与人类同等水平	2013	超过50种超高难度雅达利游戏的学习环境数据集	1992	Q 学习算法
平均突破年限：		3 年		18 年	

复杂网络风险量化模型

强调事物“联系”与全局量化的数据结构，主要用它来解决企业大量沉淀的结构化数据，使用向量矩阵替代数据库表结构计算。

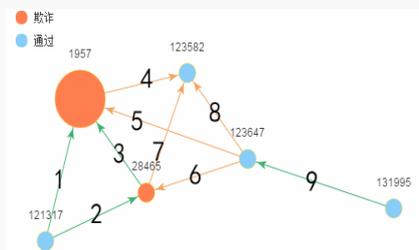


- 风险共同体：投资一致行动人，循环担保，重复抵押
- 社交P2P借贷
- 流动资本计量
- 风险种子揭示

信用申请反欺诈

通过分析银行信用卡的“通过信用卡”信息和“欺诈信用卡”信息，找到注册信息中包含的关系，同时对关系信息统计分析，计算相关指标，然后通过统计分析的结果构建社交网络，最终支撑欺诈用户发现。

申请进件的关联特征

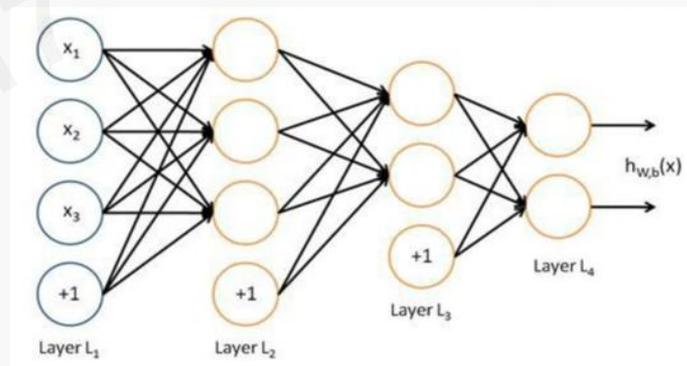


基础金融属性

年龄	
年收入	
学历	
职位	
区域	
职业	
第三方信用卡	
...	



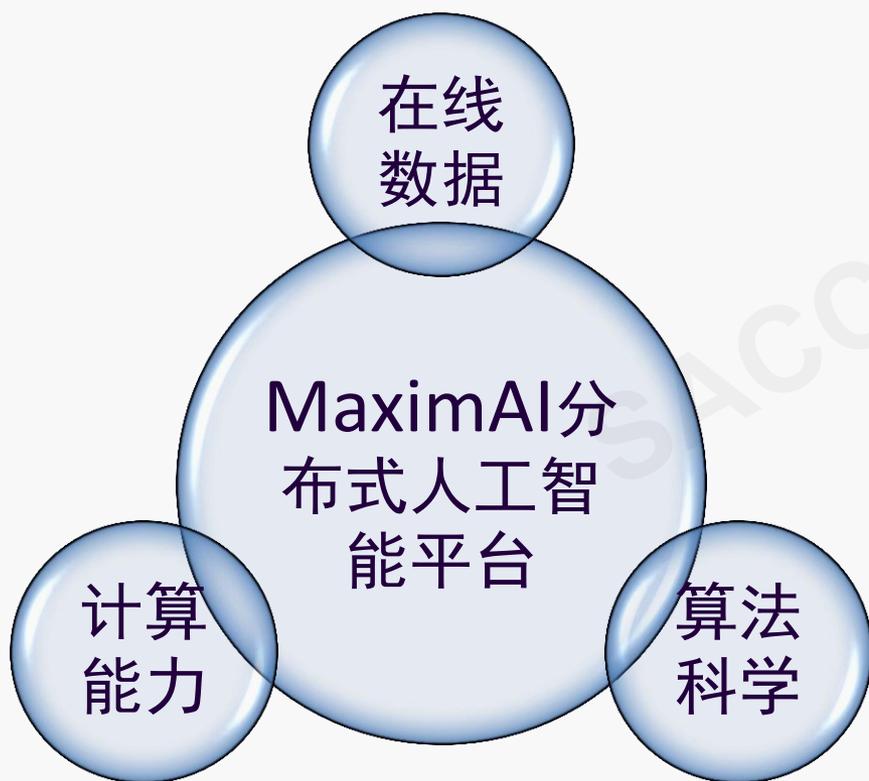
深度学习网络



Combine	AUC	Accuracy	Precision	Recall	F1-measure	Lift
LR	0.85	0.86	0.90	0.90	0.90	1.25
DL	0.90	0.87	0.90	0.92	0.91	1.32
RF	0.93	0.88	0.93	0.91	0.92	1.35

融合Algorithm Bigdata Cloud

MaximAI企业级人工智能平台产品



融合计算能力

从并行计算到分布式计算的创新

Scala分布式程序的算法代码重构，充分发挥SPARC/Alluxia内存计算能力。

融合在线数据

从流程驱动到数据驱动的创新

数据无需在生产系统和挖掘系统间抽取离线，实时的全量数据建模

融合业务价值

从零到一的创新

从业务问题定义到前沿算法模型反复迭代，最终体现商业价值化的模型，可以在平台中发布、分享和继承。业务创新可以规模化复制。

MaximAI 企业级人工智能平台产品 (续)

FreeCoding

采用完全界面化的操作
用户无需任何编程背景，
也可轻松使用数据挖掘技术

Subscription

通过REST接口整合、订阅算法
包和数据分析模版
面向高阶用户，自主编写Spark
Scala,R,Python代码

向导服务

订阅服务

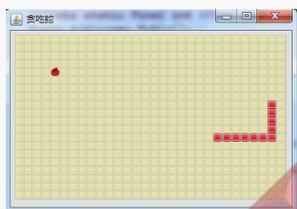


轻AI的前沿算法民主化



AI的平台化，催生Fintech

移动互联网前夜



开放平台应用开发



通讯科技巨头专利科技



Android/iOS平台化，屏蔽了技术复杂度，推动了移动互联网的繁荣。AI平台的兴起，将释放巨大AI潜能和催生更广阔人工智能市场。例：Google Tensorflow, Facebook FBLearner, MS CNTK, 腾讯Angel, 天云MaximAI.



THANKS

获取机器智能 像读书一样简单