



Shanghai Advanced
Institute of Finance
上海高级金融学院

量化俱乐部-机器学习-Fundamentals

2019-08-04

1. 自我介绍
2. 课程介绍
3. 机器学习简介与应用
4. 深度学习介绍
5. 基本概念

自我介绍



SAIF

Shanghai Advanced
Institute of Finance
上海高级金融学院

张骁喆，2010年大连理工大学软件学院本科毕业，高金FMBA 2017PTE。9年计算机工作经验，熟悉开发，测试，运维，项目管理，产品设计各环节。

曾就职eBay，唯品会，2016加入SAP，现担任SAP Cloud部门开发团队主管。

2017开始接触量化投资和python，目前毕业论文研究方向使用机器学习在A股进行量化投资。

量化投资/机器学习咨询培训/项目管理产品管理。



Professionalism · Ownership · Innovation · Excellence



1. 对象：对机器学习和人工智能相关算法，应用感兴趣；对机器学习，人工智能领域已经有所了解，想深入理解一些重要算法及应用。
2. 课程前提要求 (请务必在上课前阅读)
3. 目标：
 - a) 理解机器学习基本概念
 - b) 理解机器学习中的原理以及常用算法
 - c) 使用python及相关package(numpy, sk-learn)等进行模型训练
 - d) 会使用python对原始数据进行机器学习前的预处理
 - e) 能够根据实际问题性质采取正确的算法及应用对问题分析

4. 课程设置

机器学习	基本概念	训练集，测试集，特征值，监督学习，非监督学习，半监督学习，分类，回归
	监督学习-分类	临近取样 KNN
		决策树
		向量机
	监督学习-回归	线性回归
		多项式回归
		逻辑回归

1. 概念：多领域交叉学科，设计概率论，统计学，逼近论，凸分析，算法复杂度理论等多门学科。专门研究计算机怎样模拟人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。
2. 学科定位：人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能各个领域，它主要使用归纳，综合而不是演绎。
3. 定义：探究和开发一系列算法来如何使计算机不需要通过外部明显的指示，而可以自己通过数据来学习，建模，并且利用建好的模型和新的输入来进行预测的学科。
 - a) Arthur Samuel(1959): 一门不需要通过外部程序指示而让计算机有能力自我学习的学科
 - b) Langley(1996): 机器学习是一门人工智能的学科，该领域的主要研究对象是人工智能，特别是如何在经验学习中改善具体算法的性能。
 - c) Tom Michell(1997): 机器学习是对能通过经验自动改进的计算机算法的研究。

机器学习的应用：

语音识别
自动驾驶
语言翻译
计算机视觉
推荐系统
识别垃圾邮件

Demos:

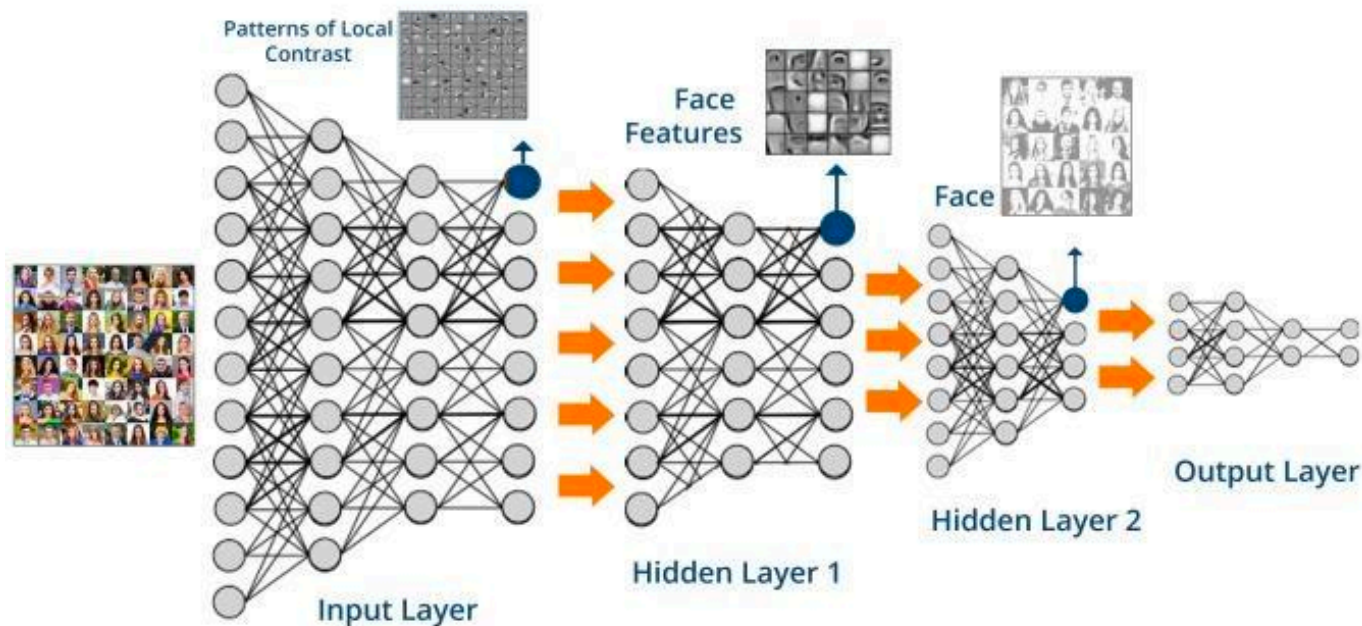
1. 全球首例！中国无人公交车上路 7个定位摄像头 精度达厘米级
2. Research at NVIDIA: AI Reconstructs Photos with Realistic Results

市场需求：

LinkedIn所有职业技能需求量第一

什么是深度学习？

深度学习是基于机器学习延伸出来的一个新的领域，由以人大脑结构为启发的神经网络算法为起源加之模型结构深度的增加发展，并伴随大数据和计算能力的提高而产生的一系列新的算法。



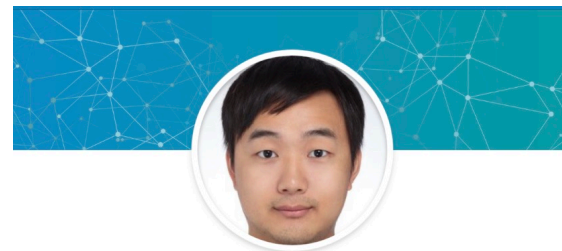
深度学习介绍

- 深度学习什么时间发展起来？
- 其概念由著名科学家Geoffrey Hinton等人在2006年和2007年在《Sciences》等上发表的文章被提出和兴起。被应用在图像处理与计算机视觉，自然语言处理以及语言识别上。自2006至今，学术界和工业界合作在深度学习方面的研究与应用在以上领域取得了突破性进展。以ImageNet为数据库的经典图像中的物体识别竞赛为例，击败了所有传统算法，取得了前所未有的精确度。



深度学习有哪些代表性的学术机构和公司？人才需求如何？

学校以多伦多大学，纽约大学，斯坦福大学为代表，工业界以Google， Facebook， Baidu 代表为深度学习研究与应用前沿。Google挖走了Hinton， Facebook挖走了LeCun， 百度硅谷实验室挖走了Andrew Ng。深度学习因技术的发展和人才的稀有造成人才抢夺战达到了前所未有的激烈程度。诸多的公司(阿里巴巴， 雅虎)也都在跟进，开始涉足深度学习领域，深度学习人才需求量会持续快速增长。



zhao zhong • 3 度

Ph.D. candidate

Institute of Automation Chinese Academy of
Sciences

北京 海淀区 • 53 人

计算机软件

加为好友

...

深度学习如今和未来将对我们生活造成怎样的影响？

目前我们使用的Android手机中的google 语音识别，百度识图，google的图片搜索，都已经使用到了深度学习技术。Facebook在名为DeepFace的项目中对人脸识别的准确率第一次接近人类肉眼(97.25%VS97.5%)。大数据时代，结合深度学习的发展在未来对我们的生活影响无法估量。保守而言，很多目前人类从事的活动都将因深度学习和相关技术的发展被机器取代，如自动汽车驾驶，无人飞机，以及更加称职的机器人，深度学习的发展让我们第一次看到并接近人工智能的终极目标。

深度学习应用展示

概念学习：人类学习概念：鸟，车，计算机

定义：概念学习是指从有关某个布尔函数的输入输出训练样例中推出该布尔函数。

例子：学习“享受运动”这一概念，小明是否享受运动取决于很多因素

样例	天气	温度	湿度	风力	水温	预报	享受运动
1	晴	暖	普通	强	暖	一样	是
2	晴	暖	大	强	暖	一样	是
3	雨	冷	大	强	暖	出入	否
4	晴	暖	大	强	冷	出入	是

概念定义在实例的集合之上，这个集合表示成 X 。待学习的概念或目标函数成为目标概念，记做 C 。

学习目标 $f: X \rightarrow y$

训练集/训练样例: 用来进行训练, 也就是用来产生模型和训练其参数的训练集。

测试集/测试样例: 用来专门进行测试已经训练好的模型。

特征向量(feature vector): 属性的集合, 通常会用一个向量来表示, 附属于一个实例。

回归算法: 目标标记为连续型数值

分类算法: 目标标记为类别性数值

监督学习: 训练集有类别标记

无监督学习: 无类别标记

半监督学习: 有类别标记训练集+无标记训练集

[SK-learn](#) [TensorFlow](#)



SAIF

Shanghai Advanced
Institute of Finance
上海高级金融学院