

人工智能技术详解

【课程内容】

本课程包含大数据、机器学习、深度学习、知识图谱、强化学习与深度强化学习的相关知识。



【课程时长】

7天 (7小时/天)

【课程对象】

理工科本科及以上，且至少了解一门编程语言。

【课程大纲】 (培训内容可根据客户需求调整)

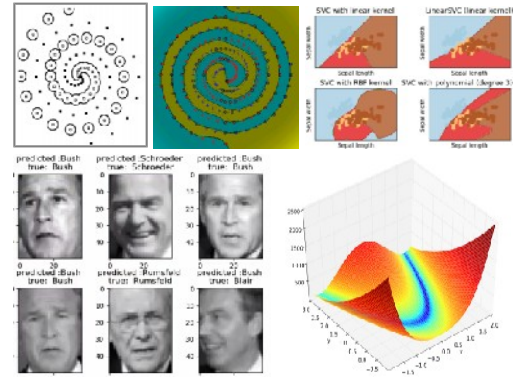
时间	内容	案例实践与练习
Day1 上午 准备工作	准备工作 1.概念与术语 2.Python (Anaconda) 的安装 3.Pycharm 的安装与使用 4.Jupyter Notebook 的安装与使用 5.Tensorflow 与 pytorch 的安装 6.Opencv、Sklearn 工具包的运用 贪婪决策过程 1.熵与熵减过程 2.贪婪法 3.ID3 与 C4.5 4.其他改进方法 5.决策树剪枝	案例实践： 1.Anaconda 安装 2.Pip install 的技巧 3.Tensorflow-GPU 的安装 4.pytorch 的安装 5.Jupyter Notebook 的使用 6.Opencv 的基本例子    

<p>Day1 下午 经典模型</p>	<p>经典无监督学习</p> <ol style="list-style-type: none"> 1.监督学习与无监督学习 2.K-means 3.k-medoids 4.判断最优分组个数的调参方法 5.基于层次、密度、网格的方法 <p>BP 神经网络</p> <ol style="list-style-type: none"> 1.人工神经元及感知机模型 2.Sigmoid 激活函数 3.前向神经网络的架构 4.梯度下降 5.误差反向传播详解 <p>性能评价指标</p> <ol style="list-style-type: none"> 1.精确率； 2.P、R 与 F1 3.ROC 与 AUC 4.对数损失 5.泛化性能评价：k 折验证验证 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.验证：K-means 是不稳定的 2.手肘法分析 NBA 球队的档次 3.绘制 ROC 并计算 AUC 4.手算神经网络 BP 算法 5.只用 numpy，手推 BPNN 
<p>Day2 上午 进阶模型</p>	<p>支持向量机</p> <ol style="list-style-type: none"> 1.“双螺旋”问题 2.基本模型与惩罚项 3.求解对偶问题 4.核函数：映射到高维 5.从二分类到多分类 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.SVM 实现人脸识别应用 2.通过深度 BP 网络实现手写数字的识别 3.皮马印第安人糖尿病风险 4.Xgboost、GBDT 等方法的比较

6.用于连续值预测的支持向量机

集成学习

- 1.集成学习的思路
- 2.bagging 与 boosting
- 3.随机森林
- 4.GBDT
- 5.XGboost
- 6.最新的集成学习算法



Day2 下午
深度学习初步

深度学习基础知识

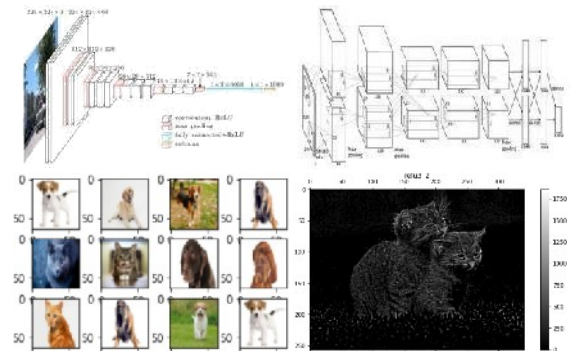
- 1.连接主义的兴衰
- 2.深度学习与 NN 的区别与联系
- 3.目标函数与激励函数
- 4.学习步长
- 5.权重初始化
- 6.权重衰减 (Weight Decay)
- 7.各种梯度下降的方法
- 8.避免过适应

卷积神经网络 CNN

- 1.CNN 概述
- 2.AlexNet 与 ZF-Net
- 3.VGG (5 层变为 5 组)
- 4.迁移学习
- 5.GoogLenet 和 Inception 模块

案例实践：

- 1.各种梯度下降方法的实战效果
- 2.VGG 各层的可视化展现
- 3.迁移学习：猫狗大战
- 4.Resnet 用于分类



	<p>6.模型退化与 ResNet</p>	
<p>Day3 上午 目标检测</p>	<p>二阶段目标检测</p> <ol style="list-style-type: none"> 1.目标检测任务介绍 2.R-CNN 3.SPPNET (全图卷积、SPP 层) 4.Fast-RCNN (多任务) 5.Faster-RCNN (RPN) <p>一阶段目标检测</p> <ol style="list-style-type: none"> 1.YOLO-v1 (一切都是回归) 2.YOLO-v2 (9000 个分类) 3.YOLO-v3 (多尺度) 4.YOLO-v4 5.YOLO-v5 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.基于 Faster-RCNN 的通用目标检测 2.改造成“血细胞识别”系统 3.基于 YOLO v3 的通用目标快速检测 
<p>Day3 下午 图谱技术准备-I</p>	<p>循环神经网络</p> <ol style="list-style-type: none"> 1、RNN 基本原理 2、LSTM、GRU 3、双向循环神经网络 4、编码器与解码器结构 5、seq2seq 模型 6、Attention <p>词向量初步</p> <ol style="list-style-type: none"> 1.Word2Vec:CBOw 2.Word2Vec:skip-gram 3.Hierachical Softmax 4.Negative Sampling 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.RNN 的基础实践 2.股票交易数据的预测 3.《绝代双骄》人物关系分析 

5.其他词向量技术

隐马尔科夫模型

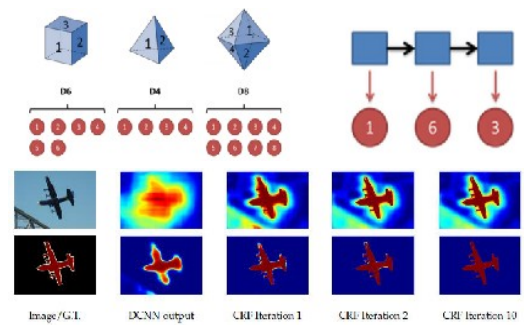
- 1.HMM 形式化定义
- 2.向前向后算法解评估问题
- 3.Viterbi 算法处理解码问题
- 4.鲍姆韦尔奇算法解学习问题

条件随机场

- 1、产生式模型与判别式模型
- 2、最大熵原理
- 3、MRF (马尔科夫随机场)
- 4、最大团与势函数
- 5、线性链条件随机场

案例实践：

- 1.出现这种情况可能吗？
- 2.今天身体怎么样？
- 3.DenseCRF 获得精细的轮廓



Day4 上午
图谱技术准备-II

Day4 下午
知识图谱基础

知识图谱基础

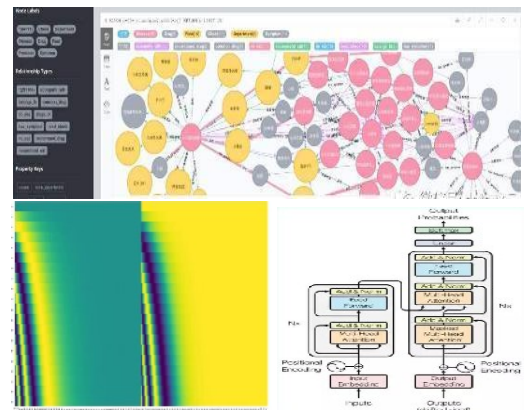
- 1.知识图谱基本概念
- 2.知识图谱的发展史
- 3.实体的属性与关系
- 4.知识图谱的相关技术介绍
- 5.典型应用案例

词向量进阶 (Transformer)

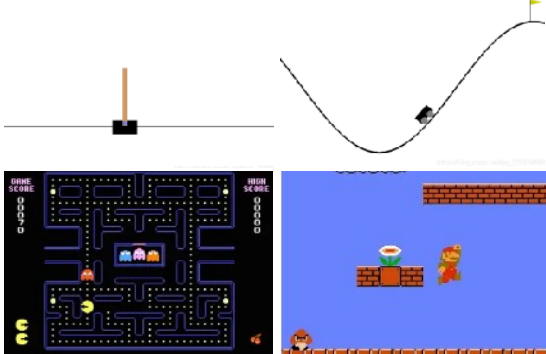
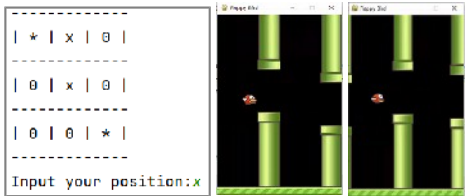
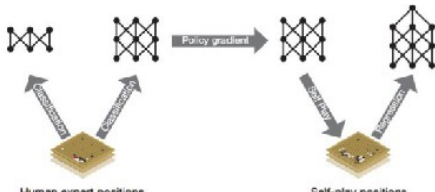
- 1.所有你需要的仅仅是“注意力”
- 2.Transformer 中的 block
- 3.自注意力机制

案例实践：

- 1.Neo4j 工具的使用
- 2.知识图谱项目展示
- 3.Transformer 实现的机器翻译



	<p>4.多头注意力</p> <p>5.位置编码 (抛弃 RNN)</p> <p>6.Batch Norm 与 Layer Norm</p> <p>7.解码器的构造</p>	
<p>Day5 上午 知识图谱进阶</p>	<p>知识抽取</p> <p>1.实体抽取</p> <p>2.关系抽取</p> <p>3.事件抽取</p> <p>4.知识抽取及其相关工具</p> <p>5.实体消歧与链接</p> <p>6.知识规则挖掘</p> <p>7.知识图谱表示学习</p> <p>基于知识图谱的问答</p> <p>1.知识问答技术概述</p> <p>2.知识问答系统简史</p> <p>3.知识问答的评测数据集</p> <p>4.KBQA 基本概念及挑战</p> <p>5.知识问答主流方法介绍</p>	<p>案例实践：</p> <p>1.实体关系抽取案例</p> <p>2.基于图谱的领域问答</p>  
<p>Day5 下午 强化学习</p>	<p>强化学习初步</p> <p>1.agent 的属性</p> <p>2.马尔科夫奖励/决策过程</p> <p>3.exploration and exploitation</p> <p>4.状态行为值函数</p> <p>5.Bellman 期望方程</p>	<p>案例实践：</p> <p>1.财宝在右</p> <p>2.格子世界</p> <p>3.谷底的小车</p> <p>4.倒立摆</p>

	<p>6.最优策略</p> <p>强化学习经典方法</p> <ol style="list-style-type: none"> 1.策略迭代与价值迭代 2.动态规划法：DP 3.蒙特卡洛法：MC 4.时序差分法：TD 5.DP、MC、TD 的关系 	
<p>Day6 上午 深度强化学习</p>	<p>深度强化学习 (DQN)</p> <ol style="list-style-type: none"> 1.值函数的参数化表示 2.值函数的估计过程 3.基础的 DQN 方法 4.Double DQN 5.Prioritized Replay 6.Dueling Network <p>深度强化学习 (PG)</p> <ol style="list-style-type: none"> 1.策略梯度方法介绍 2.常见的策略表示 3.减小方差的方法 4.引入基函数与修改估计值函数 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.自我进化的井字棋 2.笨鸟先飞：DQN 
<p>Day6 下午 深度强化学习</p>	<p>AlphaGo</p> <ol style="list-style-type: none"> 1.围棋 AI 的难点 2.MCTS 3.策略网络 4.价值网络 	<p>案例实践：</p> <ol style="list-style-type: none"> 1.山寨版 alpha go 2.Actor-Critic 的应用场景 

	<p>5.Alpha Go 的完整架构</p> <p>6.Alpha Go zero</p> <p>Actor-Critic 算法</p> <p>1.随机策略与确定性策略比较</p> <p>2.随机策略 AC 的方法</p> <p>3.确定性策略梯度方法</p> <p>4.DDPG 方法及实现</p> <p>5.A3C 方法</p>	
<p>Day7 上午 大数据技术</p>	<p>大数据技术综述</p> <p>1.从数据库，数据仓库到大数据</p> <p>2.大数据的 6V 特征</p> <p>3.Hadoop 生态圈简介</p> <p>4.Spark 生态圈简介</p> <p>5.搜索引擎：ES</p> <p>hadoop</p> <p>1.文件系统：HDFS</p> <p>2.计算框架 Map-Reduce</p> <p>3.数据库 Hbase</p> <p>4.数据仓库 Hive</p> <p>5.数据迁移：sqoop</p>	<p>案例实践：</p> <p>1.Hadoop 集群的操作</p> <p>2.Hbase 的操作</p> 
<p>Day7 下午 大数据及其他</p>	<p>Spark</p> <p>1.scala</p> <p>2.RDD</p> <p>3.数据库操作：spark-SQL</p>	<p>案例实践：</p> <p>1.spark 集群的演示</p> <p>2.计算机想象的数字</p>

4.流操作 : spark-Streaming

GAN

1.生成对抗网络 (GAN)

2.KL 散度与JS 散度

3.改进的 GAN : DCGAN

4.Wasserstein GAN

3.特朗普的孩子？

