

北京中科资环信息技术 研究院文件

中科资环发(2018)79 号

MATLAB 最新干货课程-融合编程、图像处理、 机器学习、深度学习实践技术内容培训班

(全程 4 天实操授课，核心编程技巧，MATLAB 领域实用干货，现场答疑)

各企事业单位：

近年来，随着 AlphaGo、无人驾驶汽车、医学影像智慧辅助诊疗、ImageNet 竞赛等热点事件的发生，人工智能迎来了新一轮的发展浪潮。尤其是深度学习技术，在许多行业都取得了颠覆性的成果。因此，为了帮助广大科研人员更加系统地学习深度学习的基础理论知识及对应的 MATLAB 代码实现方法，北京中科资环信息技术研究院特举办“MATLAB 最新干货课程-融合编程、图像处理、机器学习、深度学习实践技术内容培训班”旨在帮助学员掌握深度学习的基础知识，与经典机器学习算法的区别与联系，以及最新的迁移学习、强化学习、对抗生成网络等算法的基本原理及其 MATLAB 编程实现方法。本次培训采用“理论讲解+案例实战+动手实操+讨论互动”相结合的方式，抽丝剥茧、深入浅出分析深度学习在应用时需要掌握的经验及编程技巧。此外，本次培训还将通过实际案例的形式，介绍如何提炼创新点，以及如何发表高水平论文等相关经验。

一、主办单位：北京中科资环信息技术研究院

会议指定协办单位：北京中科硕博研计算技术中心

二、时间地点：2018 年 12 月 6 日—12 月 10 日 江苏*南京

(时间安排：第一天报到、授课四天)

三、培训目标：

1. 掌握 MATLAB 基础编程及进阶提升方法
2. 通过实操培训掌握各种编程技巧
3. 熟练掌握 MATLAB 图像处理方法
4. 掌握 BP 神经网络、极限学习机、支持向量机、决策树与随机森林、遗传算法等的基本原理及其 MATLAB 代码实现方法
5. 掌握最新的卷积神经网络、长短时记忆网络 LSTM、迁移学习、强化学习等算法的基本原理及其 MATLAB 代码实现方法

主办单位：北京中科资环信息技术研究院

二零一八年十一月一日

协办单位：北京中科硕博研计算技术中心

二零一八年十一月一日

四、培训对象：

各省市、自治区从事各行各业的机器学习、数据挖掘、图像处理等方向相关的企事业单位技术骨干、科研院所研究人员和大专院校相关专业教学人员及在校本科生、硕士和博士等相关人员，以及对机器学习、深度学习和 MATLAB 编程感兴趣的广大爱好者

五、培训费用

4300 元（报名费、培训费、资料费、午餐费）住宿可统一安排，费用自理。
（发票可开具：培训费、会议费、资料费等，会场发放红头文件，用于参会人员报销使用）

六、颁发证书：

参加会议的学员可以获得《深度学习技术》专业技术培训证书。此证书作为个人学习和知识更新、专业技能提升、单位人才聘用的参考依据。
注：请学员准备电子版：姓名+身份证号+2 寸蓝底证件照片发至会务组。

七、报名方式：

请各有关部门统一组织本地区行政、企事业单位报名参加培训，各单位也可直接报名参加。
报名回执表请传真至会务组 010-53853500 或发 E_mail: 1561702198@qq.com 会务组收到回执后通知报到相关事项。


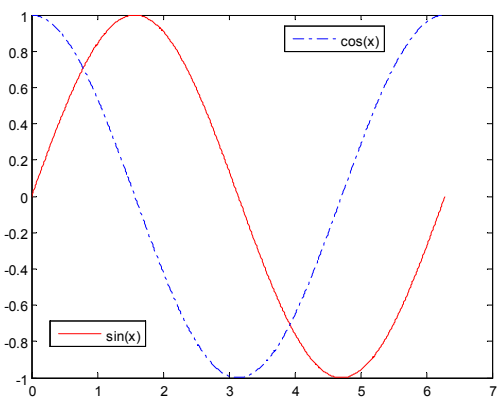
八、联系方式：

联系人：王进 18510371091（微信同步） QQ 咨询：1561702198 电话/传真：010-53853500
MATLAB 编程与机器学习交流群：540840947 加群请备注：王进邀请

报名回执表（此表可复制）

单位名称						
发票抬头				发票类型	<input type="checkbox"/> 增值普票 <input type="checkbox"/> 增值专票	
纳税人识别号				发票内容	<input type="checkbox"/> 培训费 <input type="checkbox"/> 会议费	
通讯地址						
学员姓名	性别	部门	工作（研究）方向	职务	联系方式	邮箱
希望通过会议解决哪些问题（请认真填写）		问题一： 问题二： 问题三：				
培训费		万 仟 佰 拾 元			电汇日期	
培训费用： <input type="checkbox"/> 电汇 <input type="checkbox"/> 现场办理费用 （请用“√”标注）						
汇 款 账 户	账户：北京中科硕博研计算技术中心 开户行：中国工商银行股份有限公司北京学院路支行 账号：0200 0255 0920 0106 214 注：请将银行汇款凭证传真至会务组。			账户：麦豆信息技术（保定）有限公司 开户行：中国工商银行股份有限公司保定新华支行 账号：0409 0030 0930 0163 555 会议指定服务单位		
是否需要住宿： <input type="checkbox"/> 是 <input type="checkbox"/> 否 入住日期__、结束日期__、住宿天数__天 标准间__间、大床__间。				参会单位盖章： 参会学员签名： 会议负责人：王进 <div>二零一八年 月 日</div>		

课程大纲

时间		授课方式	课程	主要内容
Day 1	08:30-09:00	课程简介 经验分享	第一章 MATLAB 入门 基础	<p>1、到底应该如何学习编程？应该学习哪种编程语言？MATLAB 过时了吗？</p>  <p>2、简单介绍 MATLAB 的安装、版本历史与编程环境（应该安装哪个版本的 MATLAB？在哪些情况下需要时刻关注最新版本的 MATLAB？）</p> <div><p>Neural Network Toolbox Release Notes</p><p>Bug Reports Bug Fixes</p><p>Found 32 notes Release Range: R2014b to R2017b</p><p>Sort by:</p><p>▼ R2017b</p><p>New Features, Bug Fixes, Compatibility Considerations</p><ul style="list-style-type: none">> Directed Acyclic Graph (DAG) Networks: Create deep learning networks with more complex architecture to improve accuracy and use many popular pretrained models> Long Short-Term Memory (LSTM) Networks: Create deep learning networks with the LSTM recurrent neural network topology for time-series classification and prediction> Deep Learning Validation: Automatically validate network and stop training when validation metrics stop improving> Deep Learning Layer Definition: Define new layers with learnable parameters, and specify loss functions for classification and regression output layers> Deep Learning Training Plots: Monitor training progress with plots of accuracy, loss, validation metrics, and more> Deep Learning Image Preprocessing: Efficiently resize and augment image data for training> Bayesian Optimization of Deep Learning: Find optimal settings for training deep networks (Requires Statistics and Machine Learning Toolbox)> GoogLeNet Pretrained Network: Transfer learning with pretrained GoogLeNet convolutional neural network> Batch Normalization Layer: Speed up network training and reduce sensitivity to network initialization> Deep Learning: New network layers> Functionality Being Removed or Changed</div> <p>3、MATLAB 基础操作：包括矩阵操作、逻辑与流程控制、函数与脚本文件、基本绘图等（浩如烟海的函数应该怎样记忆？矩阵最常用的五种索引操作是什么？MATLAB 绘图功能真的比较弱吗？如何导出高质量的图像，以供满足 SCI 论文的要求？）</p>  <p>4、文件导入：mat、txt、xls、csv、jpg、wav、avi 等格式（mat 格式的文件是怎么生成的呢？为什么 mat 文件导入到 workspace 中会有各种各样名字的变量？）</p>
	09:00-12:00	相关知识点 复习与巩固 理论讲解与 案例演示 实操练习		

13:30-15:30

理论讲解与
案例演示
实操练习

第二章
MATLAB 进阶
与提高

1、MATLAB 编程习惯与风格（Cell 模式和程序发布功能是什么？为什么代码中需要加入一些空格和空白行？MATLAB 新版本中包含的 Live Script 是什么？）

Contents

- I. 清空环境变量
- II. 训练集/测试集产生
- III. 数据归一化
- IV. BP神经网络创建、训练及仿真测试
- V. 性能评价
- VI. 绘图

I. 清空环境变量

```
clear all  
clc
```

II. 训练集/测试集产生

1. 导入数据

```
load spectra_data.mat
```

```
%% I. 清空环境变量  
clear all  
clc  
  
%% II. 训练集/测试集产生  
%%  
% 1. 导入数据  
load spectra_data.mat  
  
%%  
% 2. 随机产生训练集和测试集  
temp = randperm(size(NIR, 1));  
% 训练集——50个样本  
P_train = NIR(temp(1:50), :)' ;  
T_train = octane(temp(1:50), :)' ;  
% 测试集——10个样本  
P_test = NIR(temp(51:end), :)' ;  
T_test = octane(temp(51:end), :)' ;  
N = size(P_test, 2);
```

2、MATLAB 调试技巧（MATLAB 为什么会给出各种各样的错误信息？常见的错误信息有哪些？面对错误信息，应该是失落还是开心？如何使用断点调试工具？应该去哪些网站寻找答案？怎样能够让别人乐意帮助你解决问题？）

■ 错误信息的阅读

- Index must be a positive integer or logical.
- Undefined function or variable "B".
- Inner matrix dimensions must agree.
- Function definitions are not permitted at the prompt or in scripts.
- Index out of bounds because numel(A)=5.
- In an assignment A(I) = B, the number of elements in B and I must be the same.
- Expression or statement is incorrect--possibly unbalanced (, {, or [.
- Too many input arguments.
-

3、向量化编程与内存优化（怎样提升你的代码效率？MATLAB 的内存管理机制是什么？为什么你的代码中会出现许多红色的下划波浪线？）

- 及时清除不用的变量
- 使用变量前，预分配内存空间
- 选择恰当的数据类型
- 循环与向量化
 - 按列优先循环
 - 循环次数多的变量安排在内层
- ◆ 给一些函数“瘦身”
- ◆

4、MATLAB 深度学习工具箱介绍及所需的开发环境配置，Release Notes 解读

Name	Type	Author	Install Date
MATLAB Support Package for USB Webcams version 17.2.0	Hardware Support Package		1 November 2017
IMPIC Neural Network Toolbox Importer for ResNet-50 Network version 17.2.0	Feature		1 November 2017
IMPIC Neural Network Toolbox Importer for Tensorflow-Keras Models version 17.2.0	Feature		1 November 2017
Caffe Neural Network Toolbox Importer for Caffe Models version 17.2.0	Feature		1 November 2017
googLeNet Neural Network Toolbox Model for GoogLeNet Network version 17.2.0	Feature		1 November 2017
NIxNet Neural Network Toolbox Model for AlexNet Network version 17.2.0	Feature		1 November 2017
VGG-16 Neural Network Toolbox Model for VGG-16 Network version 17.2.0	Feature		1 November 2017
VGG-19 Neural Network Toolbox Model for VGG-19 Network version 17.2.0	Feature		1 November 2017

15:30-18:00

理论讲解与
案例演示
实操练习

第三章
MATLAB 图像
处理技术方法

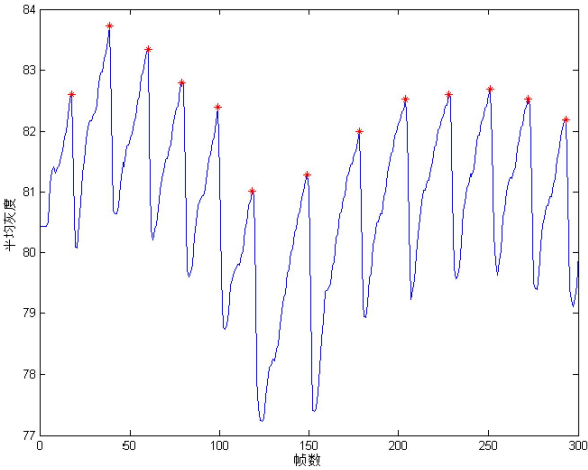
- 1、图像的常见格式及读写（彩色图像、灰度图像、二值图像等）
- 2、图像的基本操作（几何变换、时频域变换等）
- 3、图像直方图（图像的对比度、亮度等改变）
- 4、案例实践：基于手机摄像头的心率计算

一、工作原理

将手指覆盖在手机摄像头上，用摄像头照亮指尖皮下毛细血管，当心脏将新鲜的血液压入毛细血管时，亮度（红色的深度）会有轻微变化，通过摄像头监测这一有规律变化的间隔，利用图像处理分析技术，即可计算出心率。

二、实验方法

- 手机型号：小米（Android 4.2）
- 视频采集帧频：15.17帧/秒
- 数据采集方法：受试者呈坐立姿势，保持安静状态，将右手手指轻轻覆盖在手机摄像头上，左手操作手机进入摄像模式，录制30秒左右的视频并保存为3gp格式。
- 利用格式转换工具将视频文件转换为avi格式视频，以便于MATLAB软件的读取与分析。

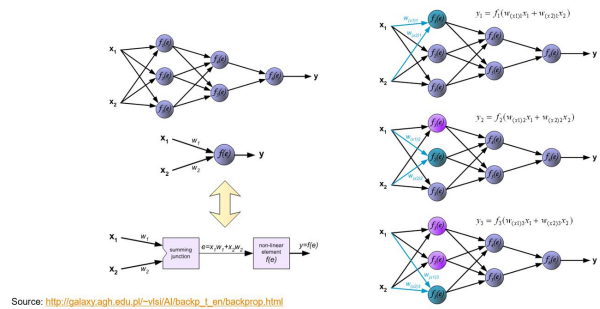


Day 2

08:30-11:00

理论讲解与
案例演示
实操练习第四章
BP 神经网络

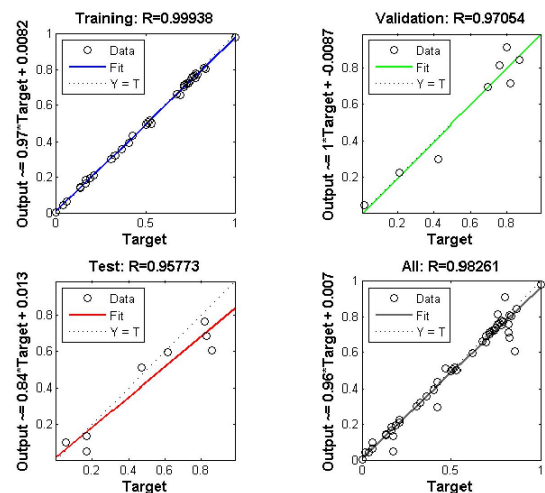
1、BP 神经网络的基本原理（人工智能发展过程经历了哪些曲折？人工神经网络的分类有哪些？BP 神经网络的拓扑结构和训练过程是怎样的？什么是梯度下降法？BP 神经网络建模的本质是什么？）



2、BP 神经网络的 MATLAB 实现（怎样划分训练集和测试集？为什么需要归一化？归一化是必须的吗？什么是梯度爆炸与梯度消失？MATLAB 中 BP 神经网络的常用函数有哪些？如何使用？）

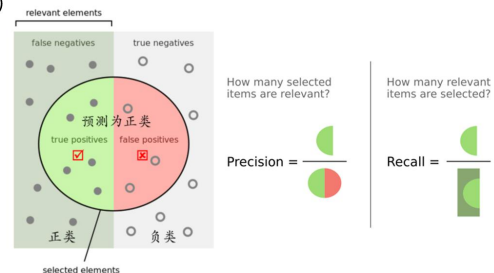
- 什么是归一化？
 - 将数据映射到[0, 1]或[-1, 1]区间或其他的区间。
- 为什么要归一化？
 - 输入数据的单位不一样，有些数据的范围可能特别大，导致的结果是神经网络收敛慢，训练时间长。
 - 数据范围大的输入在模式分类中的作用可能会较大，而数据范围小的输入作用可能会减小。
 - 由于神经网络输出层的激活函数的取值是有范围的，因此需要将网络训练的自标数据映射到激活函数的取值域。例如神经网络的输出层若采用S形激活函数，由于S形函数的取值域限制在(0, 1)，也就迫使神经网络的输出只能限制在(0, 1)，所以训练数据的输出就要归一化到(0, 1)区间。
 - S形激活函数在(0, 1)区间以外区域很平缓，区分度太小。例如S形函数在参数 $x=1$ 时，有100)与5)只相差0.0067。
- 归一化算法
 - $y = (x - \min) / (\max - \min)$
 - $y = 2 * (x - \min) / (\max - \min) - 1$

3、BP 神经网络参数的优化（隐含层神经元个数、学习率、初始权值和阈值等如何设置？什么是交叉验证？）



4、值得研究的若干问题（欠拟合与过拟合、泛化性能评价指标的设计、样本不平衡问题等）

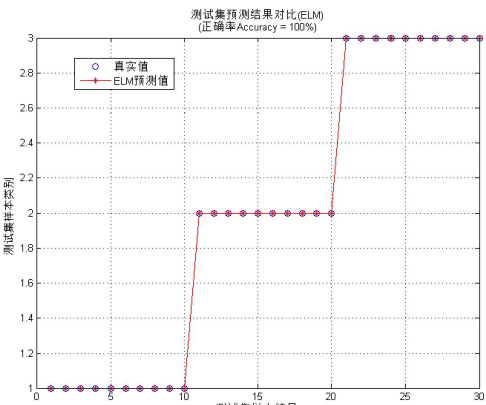
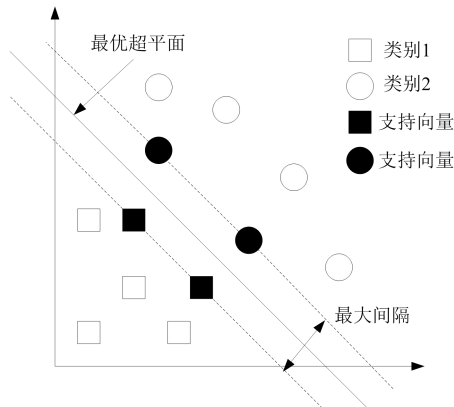
- 模型泛化性能评价（绝对误差、相对误差、均方根误差、准确率、精确率、召回率等）



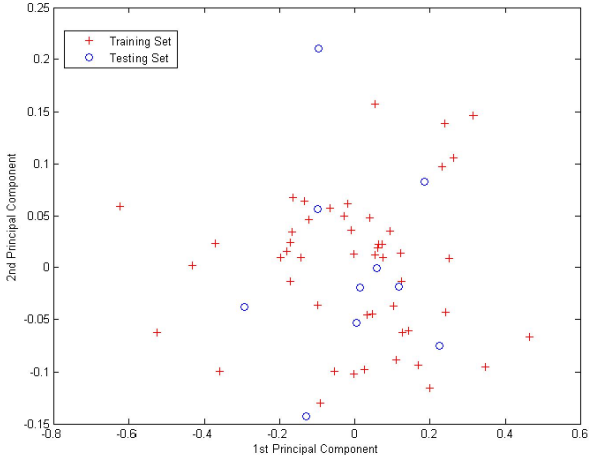
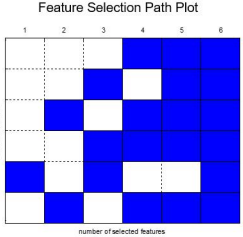
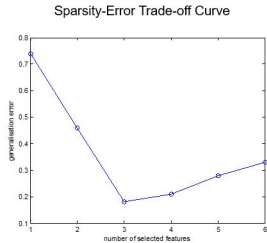
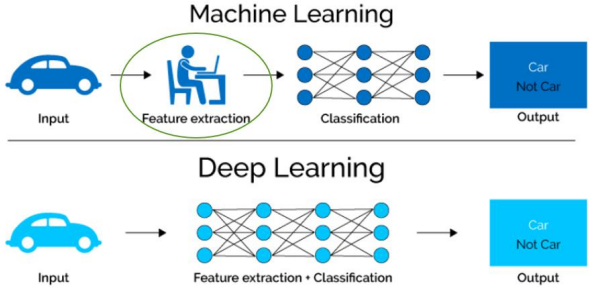
5、案例演示一：近红外光谱汽油辛烷值预测（回归拟合）

6、案例演示二：人脸朝向识别（分类识别）

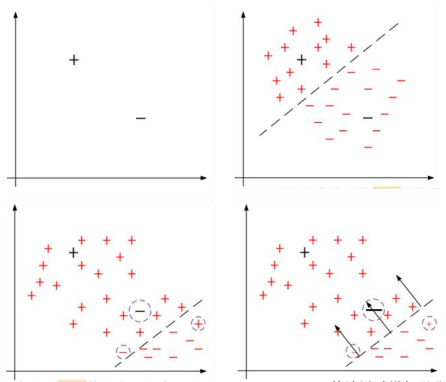
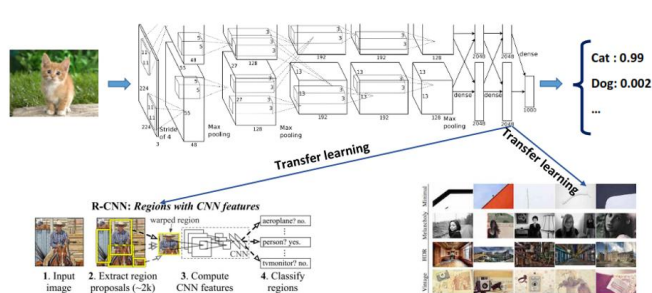
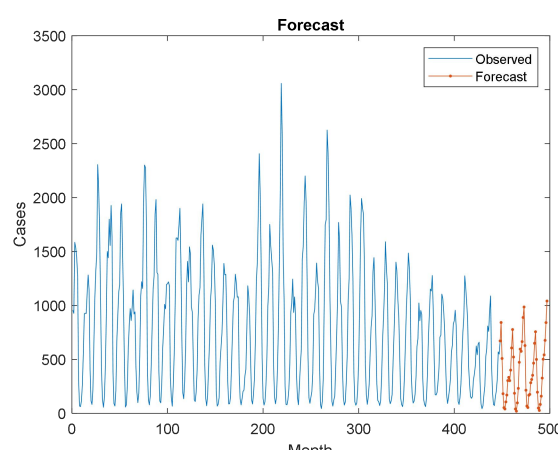
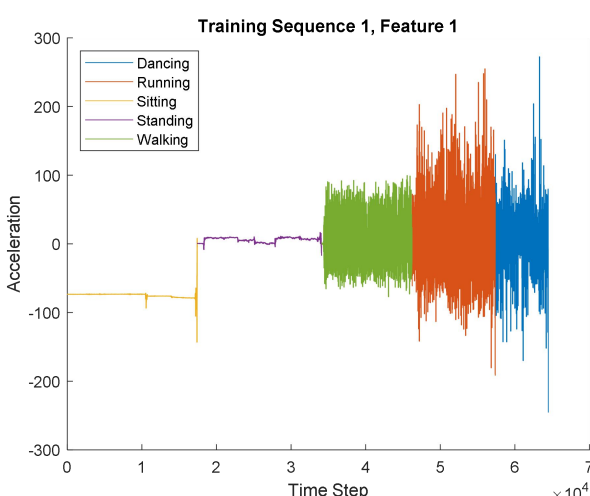
7、实操练习

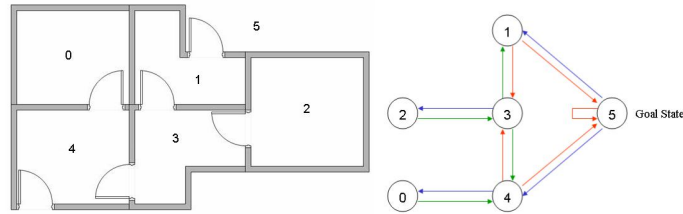
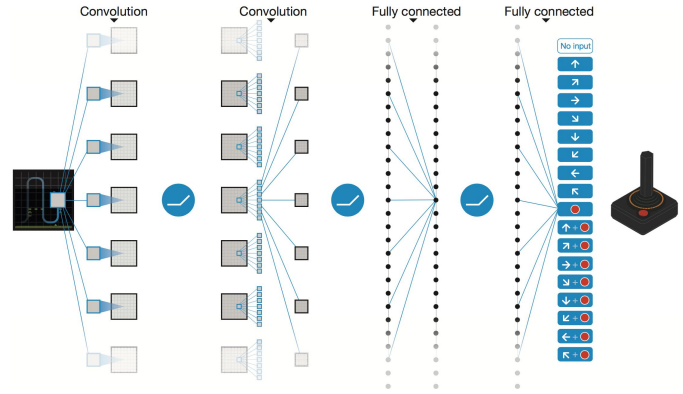

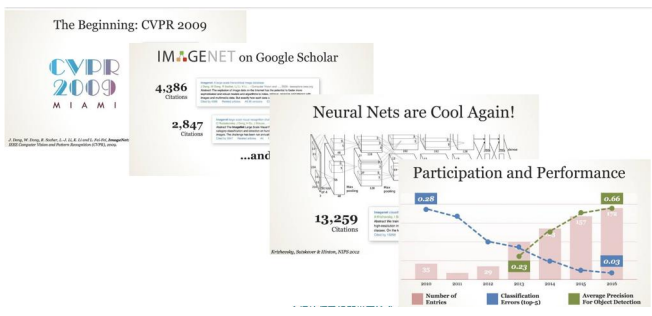
	11:00-12:00	理论讲解与 案例演示 实操练习	第五章：极限学习机(Extreme Learning Machine, ELM)	<p>1、ELM 的基本原理（ELM 的基本算法，“极限”体现在哪些地方？）</p> <p>Algorithm ELM: Given a training set $\mathcal{N} = \{(x_i, t_i) x_i \in \mathbf{R}^n, t_i \in \mathbf{R}^m, i = 1, \dots, N\}$, activation function $g(x)$, and hidden node number \tilde{N},</p> <p>Step 1: Randomly assign input weight w_i and bias b_i, $i = 1, \dots, \tilde{N}$.</p> <p>Step 2: Calculate the hidden layer output matrix \mathbf{H}.</p> <p>Step 3: Calculate the output weight β</p> $\beta = \mathbf{H}^\dagger \mathbf{T}, \quad (16)$ <p>where $\mathbf{T} = [t_1, \dots, t_N]^T$.</p> <p>2、ELM 与 BP 神经网络的区别与联系</p> <p>3、案例实践：鸢尾花种类识别</p> 
	13:30 – 15:30	理论讲解与 案例演示 实操练习 知识点总结、 延拓	第六章：支持向量机(Support Vector Machine, SVM)	<p>1、SVM 的基本原理（什么是经验误差最小和结构误差最小？SVM 的本质是解决什么问题？SVM 的四种典型结构是什么？核函数的作用是什么？什么是支持向量？）</p>  <p>2、SVM 扩展知识（如何解决多分类问题？SVM 除了建模之外，还可以帮助我们做些什么事情？）</p> <p>3、LibSVM 工具箱的安装与使用</p> <p>4、案例实践一：乳腺癌肿瘤诊断</p> <p>5、案例实践二：混凝土强度预测</p>
	15:30 – 18:00	理论讲解与 案例演示 实操练习 知识点总结、 延拓 课堂讨论	第七章：决策树与随机森林	<p>1、决策树的基本原理（微软小冰读心术的启示；什么是信息熵和信息增益？ID3 算法和 C4.5 算法的区别与联系）</p>

			<div data-bbox="858 107 1423 421"> </div> <div data-bbox="791 432 1500 551"> <p>2、随机森林的基本原理（为什么需要随机森林算法？广义与狭义意义下的“随机森林”分别指的是什么？“随机”体现在哪些地方？随机森林的本质是什么？）</p> </div> <div data-bbox="837 555 1449 840"> </div> <div data-bbox="791 846 1500 965"> <p>3、案例实践：乳腺癌肿瘤诊断 4、知识扩展：决策树除了建模型之外，还可以帮我们做什么事情？怎样解读随机森林的结果？</p> </div> <div data-bbox="837 972 1449 1458"> </div>	<div data-bbox="100 1787 359 1816"> <p>Day 3 08:30 – 10:00</p> </div> <div data-bbox="387 1680 539 1921"> <p>理论讲解与 案例演示 实操练习 知识点总结、 延拓 课堂讨论</p> </div> <div data-bbox="576 1742 762 1861"> <p>第八章：遗传算法(Genetic Algorithm, GA)</p> </div>	<div data-bbox="791 1469 1500 1588"> <p>1、遗传算法的基本原理（以遗传算法为代表的群优化算法的基本思想是什么？目前国内外的研究热点在哪些方面？）</p> </div> <div data-bbox="837 1594 1449 1973"> </div> <div data-bbox="791 2011 1500 2130"> <p>2、常见遗传算法工具箱介绍（GAOT 工具箱的添加与使用方法） 3、案例实践一：一元函数的寻优计算（极大值与极小值）</p> </div>
--	--	--	--	---	---

				<p>4、案例实践二：多元函数的寻优计算（遗传算法优化 BP 神经网络的初始权值与阈值）</p> <p>5、案例实践三：离散变量的寻优计算（基于遗传算法的特征变量筛选）</p>
10:00 – 12:00	理论讲解与 案例演示 实操练习 知识点总结、 延拓 课堂讨论	第九章：变量降 维与特征选择	<p>1、变量降维与特征选择在概念上的区分</p> <p>2、主成分分析（PCA）、偏最小二乘（PLS）的基本原理（PCA 与 PLS 的区别与联系；PCA 除了降维之外，还可以帮助我们做什么？）</p>  <p>2、常见的特征选择方法（优化搜索、Filter 和 Wrapper 等；前向与后向选择法；区间法；无信息变量消除法；正则稀疏优化方法等）</p>   <p>3、案例实践：近红外光谱的降维与波长筛选</p>	
13:30 – 18:00	理论讲解与 案例演示 实操练习 知识点总结、 延拓 课堂讨论	第十章 深度学习入门 基础与卷积神 经网络	<p>1、深度学习与传统机器学习的区别与联系（神经网络的隐含层数越多越好吗？深度学习与传统机器学习的本质区别是什么？）</p>  <p>2、深度学习开源工具箱简介（Python & MATLAB）</p>	

				<div data-bbox="858 120 1442 407" data-label="Image"> </div> <p>3、卷积神经网络的基本原理（什么是卷积核？CNN 的典型拓扑结构是怎样的？CNN 的权值共享机制是什么？CNN 提取的特征是怎样的？）</p> <div data-bbox="794 568 1401 788" data-label="Diagram"> </div> <p>4、LeNet、AlexNet、Vgg-16/19、GoogLeNet、ResNet 等典型深度神经网络的区别与联系解读</p> <p>5、卷积神经网络参数调试技巧</p> <p>6、案例演示一：11 行代码实现深度学习物体识别</p> <div data-bbox="842 967 1442 1299" data-label="Image"> </div> <p>7、案例演示二：利用卷积神经网络抽取抽象特征</p> <p>8、案例演示三：自定义卷积神经网络拓扑结构</p> <p>9、实操练习</p>
Day 4	08:30-10:00	理论讲解与 案例演示 实操练习	第十一章 迁移学习算法	<p>1、迁移学习算法的基本原理（为什么需要迁移学习？为什么可以迁移学习？迁移学习的基本思想是什么？）</p> <div data-bbox="820 1514 1471 1711" data-label="Diagram"> </div> <p>2、常用的迁移学习算法简介（基于实例、特征和模型，譬如：TrAdaboost 算法）</p>

				 <p>3、基于深度神经网络的迁移学习算法</p>  <p>4、案例演示</p> <p>1、LSTM 神经网络的基本工作原理及 MATLAB 实现方法</p> <p>2、案例演示一：时间序列预测（Time Series Forecasting）</p>  <p>3、案例演示二：序列-序列分类（Sequence-to-Sequence Classification）</p> 
10:00-12:00	理论讲解与案例演示 实操练习	第十二章 长短时记忆网 络 LSTM		

	13:30-14:30	理论讲解与 案例演示 实操练习	第十三章 强化学习 (增强学习)	<p>1、强化学习的基本思想与原理介绍</p>  <p>2、Q-Learning 算法详解</p> <ol style="list-style-type: none"> 1. Initialize Q-values ($Q(s, a)$) arbitrarily for all state-action pairs. 2. For life or until learning is stopped... 3. Choose an action (a) in the current world state (s) based on current Q-value estimates ($Q(s, \cdot)$). 4. Take the action (a) and observe the outcome state (s') and reward (r). 5. Update $Q(s, a) := Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$ <p>3、深度强化学习介绍</p>  <p>4、案例演示</p>
	14:30-16:30	理论讲解与 课题讨论	第十四章 深度学习热门 研究方向	<p>1、对抗生成网络（什么是对抗生成网络？为什么需要对抗生成网络？对抗生成网络可以帮我们做什么？）</p>  <p>2、专题讨论：数据 VS.模型，孰更重要？（模型泛化性能不好，究竟是孰之过？）</p> 

	16:30-17:30	理论讲解与 案例演示	第十五章 科研与创新方 法概述	1、如何查阅文献资料？（你会使用 Google Scholar、Sci-Hub、ResearchGate 吗？应该去哪些地方查找与论文配套的数据和代码？） 2、如何高效率撰写专业论文？（SCI 不同分区的论文差别在哪些地方？你知道你的论文为什么显得很单薄吗？） 3、从审稿人的角度看，SCI 期刊论文需要具备哪些要素？（审稿人关注的点有哪些？如何回应审稿人提出的意见？） 4、如何提炼与挖掘创新点？（如果在算法层面上难以做出原创性的工作，如何结合自己的实际问题提炼与挖掘创新点？）
	17:30-18:00	课堂讨论 复习与答疑	第十六章 讨论与答疑	1、建立微信群，便于后期的讨论与答疑 2、讨论与答疑，解答学员的实际问题 3、相关学习资料分享与拷贝（图书推荐、在线课程推荐等）