

CAMA-LAB 机器学习暑期研讨班 (2017)

- CAMA-LAB 机器学习暑期研讨班 (2017)
 - 介绍
 - 整体内容安排 (暂定)
 - 机器学习
 - 深度学习
 - 主要人员

介绍

- 面向对象：准研一学生 + 大一、大二本科生
- 场地：杭电东校区
- 时间：7月17号-9月15号（工作日：8:00-17:30）
- 内容
 - 机器学习 / 深度学习 + 计算机视觉应用
 - 课题 / 项目：选择部分学生，参与老师的研究课题 / 项目
- 形式
 - 每周
 - 老师指定学习内容，并进行部分讲解
 - 学生自学为主，鼓励自由讨论
 - 定期开会讨论
 - 提交技术报告（算法思想、编程实现、实验结果及分析）
 - 暑期结束
 - 提交大报告、项目总结
 - 选择部分学生，在 2017-18 学年持续参与老师课题
- 高强度、严要求、安全第一

整体内容安排（暂定）

机器学习

第一周：热身

- Python编程基础
- 图像表示：
 - 矩阵、颜色空间
 - 图像特征：像素值、颜色直方图、梯度直方图、LBP
 - 作业：图像读写、特征提取
- 回归：
 - 算法：线性回归，线性回归+正则项（L1, L2）
 - 优化：梯度下降法
 - 测度：欧式距离、街区距离、范数、Loss
 - 作业：波士顿房价预测

第二周：分类

- 算法：k-NN, Logistic Regression, 决策树
- 作业：
 - MNIST 手写数字识别、CIFAR-10 图像分类
 - 采用第一周的特征 + 第二周的方法（多种组合，对比结果）

第三周：经典算法

- 支撑向量机 SVM 《理解SVM的三层境界》
- 图像特征：, SIFT, Visual BoW,
- 作业：收入预测、CIFAR-10 图像分类

第四周：无监督特征学习（Unsupervised Feature Learning, Manifold Learning）

- Andrew Ng 论文
- 聚类：K-means, K-means++
- 降维：PCA, ICA, ZCA, LLE, AE
- 作业：UFL + SVM, 进行CIFAR-10图像分类
- 推荐阅读：pluskid 博客

扩展：

- 经典方法：谱聚类、朴素贝叶斯、EM算法、稀疏编码 Sparse Coding
- 集成学习：Adaboost, Random Forest, gdbt (XGBoost)
- 结构化学习 Structured Learning
- 排序学习 Learning to Rank

- 强化学习 Reinforcement Learning
- 模仿学习 Imitation Learning

深度学习

第五周：卷积神经网络 CNN

- 概念：卷积, Pooling, Stride, Padding, Data Augmentation, Learning Rate, Momentum, Softmax, ReLU, BP, SGD, Cross-Entropy Loss
- 网络：LeNet, AlexNet, VGGnet, GoogLeNet, ResNet
- 框架：Keras, Pytorch
- 作业：
 - 自己搭建浅层网络（神经网络+卷积神经网络）
 - 1) 自己实现
 - 2) 使用框架
 - 两个对比
 - Image Sentiment Classification
 - 1) Analyze the Model by Confusion Matrix
 - 2) Analyze the Model by Plotting the Saliency Map
 - 3) Analyze the Model by Visualizing Filters (1%)
 - 4) Semi-supervised Learning

第六周：递归神经网络

- 网络：RNN, LSTM
- 作业：
 - 1) 由 Cosin 预测 Sin，自己编程实现网络
 - 2) Image Captioning，使用框架实现

扩展：

- 经典应用论文：Deep Learning 推荐阅读列表（余宙）
- 生成对抗网络：GAN, CGAN, DualGAN, CircleGAN

主要人员

名单				

教师				
俞俊	高飞	谭敏	余宙	朱素果
助教				
罗宇矗	钱哲琦	朱朝阳	项晨钞	施圣洁
张海超	孟宣彤	郑光剑	吴炜晨	