## 入学前或一年级研究生能力提升

编程 + 基础知识 (0.5-1年)

1）机器学习经典方法 + 编程实现 [准备5个UCI数据集，3-4 months]

* 各种线性分类器方法 (logistic regression, LDA等)
* 决策树与朴素贝叶斯
* SVM (liblinear、libsvm论文及开源代码阅读、使用) [3 weeks]
* Boosting和Gradient Boosting (GBDT、XGboost、Lightgbm论文及开源代码阅读、使用) [3 weeks]
* Random Forests
* 最近邻方法、降维
* 聚类(K-means, DBSCAN, GMM，t-sne)
* 概率图模型 (LDA主题模型, HMM)

2）各种深度学习方法 [3个数据集：图像、视频、声音，2-3 months]

* 原理：书籍Deep Learning或简单容易入手书籍
* 多层感知机MLP的原理，以及BP的原理
* 卷积网络：LeNet-5、AlexNet、VGG、ResNet、DenseNet、MobileNet
* 序列网络：RNN、LSTM、GRU、transformer及其组成深度模型(如Bert)
* 生成式模型：GAN、VAE

3）常用的软件 [容入以上学习内容，每种方法至少用两种编程工具实现]

* Python、Matlab、TensorFlow、Pytorch 学术届Pytorch，工业界Tensorflow
* 工具库: Scikit-learn、XGboost+Lightgbm、libsvm+liblinear、Numpy
* Linux：要能够熟练使用服务器命令行、vim等，能远程跑程序，熟练使用Git, Github等
* Latex论文编辑、尤其训练编辑数学公式