第一课：推荐系统与机器学习overview（概览与思维维度）

ZB EB TB GB MB

数据——机器学习——需求

金融的风控应用机器学习模型预测

收集数据——数据分析——数据预处理——

回归：预测的数据是具体数值 分类：预测的数值是离散的标签

线性回归： 逻辑回归：

训练与测试的关系：训练与测试数据的交集应当为空

参数是模型自己调节的，超参数是后期手动调节的

分类算法：C4.5，朴素贝叶斯，SVM，KNN，Adaboost，CART

聚类算法：K-means，EM

关联分析：Apriori（在推荐中的关联规则推荐）

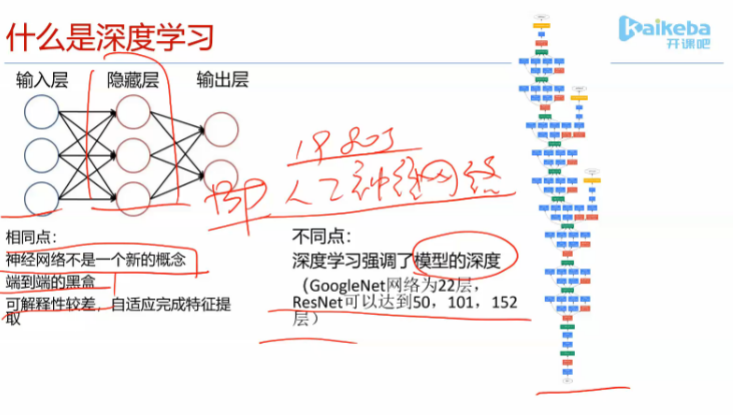
连接分析：PageRank，textrank，edgerank

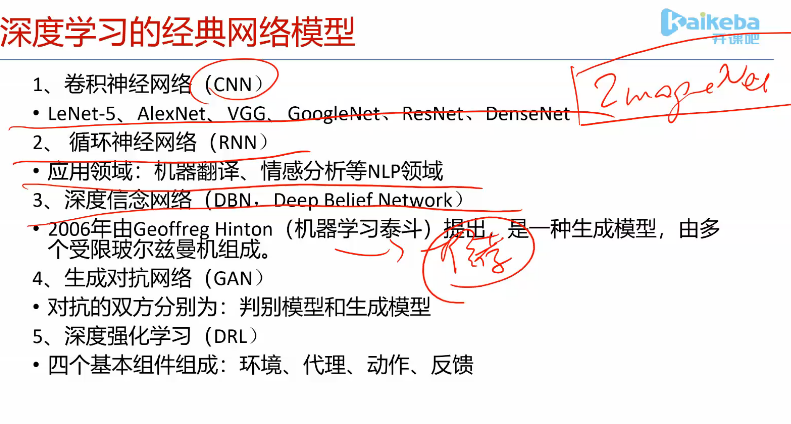


作业：使用cart算法对minist进行训练



什么是深度学习





深度学习框架：

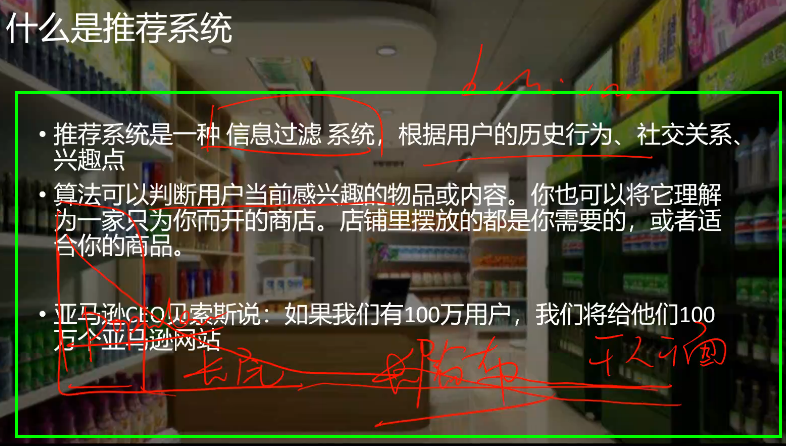
Tensorflow：tensorboard（提供tensorflow的训练进度可视化），tensorserving

Keras：Keras封装的API应用简单

Pytorch：Facebook的深度学习模型

推荐系统：

长尾分布：

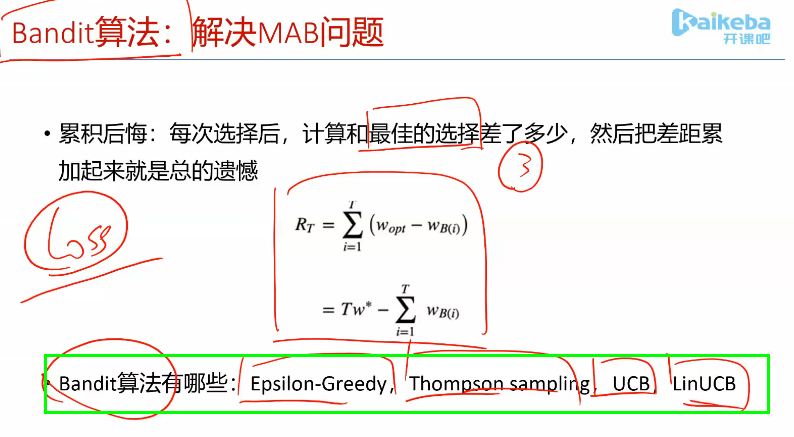


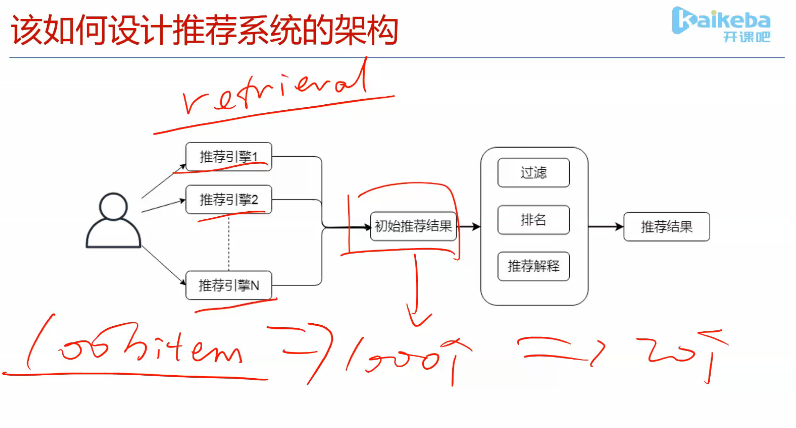
基于内容的推荐：

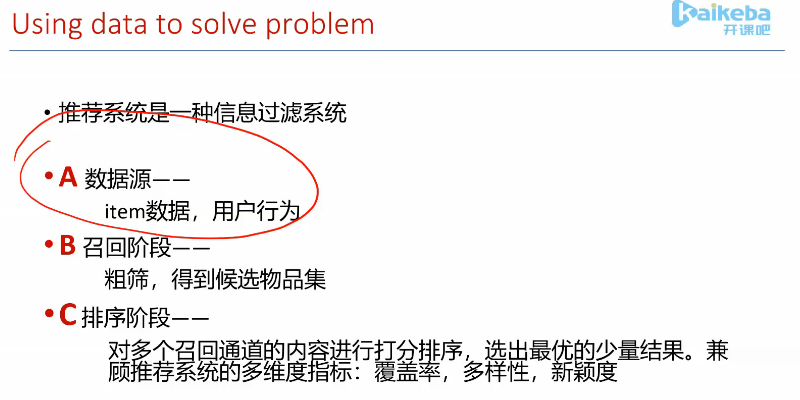
信息流>人数：usercf，新浪微博

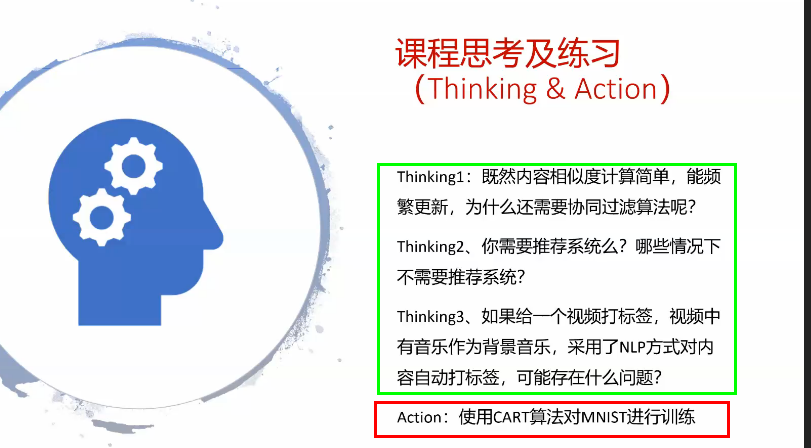
冷启动：MAB（）

Bandit算法：









矩阵分解

Bandit解决冷启动问题

监督学习：给出真实结果y，自行预测

无监督学习：不给出真实结果y，自行预测

半监督学习：给出部分真实结果y，另一部分未知的，自行预测