machine learning algorithms

- 13. GAN
- 12. Ensemble
  - 种类
    - Boosting
      - 减小偏差
      - 训练残差
      - 分类器间强关联，缺乏独立性
    - Bagging Bootstrap Aggregating
      - 投票
      - 减小方差
      - 再抽样，训练结果上取平均
    - 基分类器
  - 步骤
    - Voting
    - Stacking
  - Adaboost
    - ID3作为基分类器
    - Training
    - 加权投票合并结果
  - 常用的基分类器
    - 决策树
    - 神经网络
    - Random Forest
  - 指标
    - Bias
      - 错误假设
    - Variance
      - 模型复杂度
  - Gradient Boosting Decision Tree
    - loss function 均方误差 经验损失函数
    - CART
      - 基尼指数，平方误差
    - 梯度提升训练
      - 负梯度方向进行模型更新
    - 优缺点
  - XGBoost
    - 构建决策树加入正则项
      - 控制模型复杂程度，防止过拟合
    - 对代价函数进行二级泰勒展开
    - 支持多种类型分类器，如 线性分类器
    - 支持对数据进行采样
    - 自动学习出缺失值的处理策略
- 11. Reinforcement Learning
- 10. Recurrent Neural Network
  - 处理文本数据时，循环神经和潜亏神经网络相比有什么特点
  - 梯度消失，梯度爆炸
    - back propagation through time
    - 梯度裁剪-gradient explosion
    - 深度残差网络
  - 激活函数
    - ReLU
      - W初始化为单位矩阵
  - Long Short Term Memory
    - input
    - forget
    - output
    - c记忆单元
    - 激活函数
    - 窥孔机制
      - Ct-1参与 输入门更新
  - GRU
  - Seq2Seq
    - Encode
    - Decode
      - 贪心法
      - 集束搜索
      - 注意力机制
        - 解决损失的输入序列信息
        - 对输入序列进行编码
- 9. Feed-Forward Networks
  - 深度神经网络
    - Multi-Layer Perceptron MLP
      - 最少需要几个隐含层
      - 只有一个隐层，多少隐节点能够实现n元输入的任意布尔函数
        - Disjunctive Normal Form DNF析取范式
      - 多个隐层，需要多少个网络节点和网络层
    - 激活函数和导数
      - sigmoid
      - Tanh
      - ReLU
      - 梯度消失
      - 优缺点
    - 反向传播算法
      - 平方误差
      - 交叉熵损失函数
      - 梯度计算公式
      - 适用场景
  - Training
    - 过拟合
      - Data Augmentation
      - Regularization
      - Model Ensemble
        - Dropout
      - Batch Normalization (BN)
    - Hyperparameter
      - Learning Rate
      - Decay Rate
      - Dropout Rate
    - 是否可以初始化参数为0
      - 参数取值范围
    - why dropout works
      - Bagging
    - Batch Normalization
      - 动机，原理
  - convolutional Neural Networks
    - 稀疏交互 Sparse Interaction
      - 局部特征结构
    - 和参数共享 Parameter Sharing
      - 平移等变性
    - Pooling
      - Mean Pooling
        - 背景保留效果好
      - Max Pooling
        - 纹理信息
      - 降采样
      - 空间金字塔池化
      - 平移，伸缩，旋转操作的不变性
    - 文本分类
  - Deep Residual Network, ResNet
    - 目标函数
- 8. Sampling
  - 不均衡样本集的重采样
    - 基于数据
      - undersampling
        - Informed Undersampling
          - Easy Ensemble
          - Balance Cascade
          - NearMiss
          - One-sided Selection
          - Hard Negative Mining
      - 优缺点
      - oversampling
        - SMOTE
        - borderline-SMOTE
        - ADASYN
        - Tomek Links
        - 数据扩充
    - 基于算法
      - one-class learning
      - anomaly detection
      - 代价敏感学习-使用带标签权重的模型
    - 二分类
  - bayesian Network sampling
    - 概率图模型
      - 联合概率分布
    - 采样方法
      - 祖先采样 Ancestral Sampling
    - 只考虑部分变量边缘分布，采样方法
    - 含有观测变量如何采样
      - 逻辑采样
      - 似然加权采样Likelihood Weighted Sampling
      - MCMC采样法
  - 马尔可夫蒙特卡洛采样 MCMC
    - 主要思想
    - Metropolis-Hastings
    - 吉布斯采样法
    - 如何得到相互独立的样本
      - 同时运行多个条马尔可夫链
      - 每隔若干个样本取一个，近似独立
  - 均匀分布随机数生成器
    - 线性同余法 Linear Congruential Generator
    - 高维样本
  - 优势，应用，特性
    - 适用场景
  - 采样方法
    - 均匀采样
    - 轮盘赌算法
    - 逆变换采样 Inverse Transform Sampling
    - 拒绝采样 Rejection Sampling
    - Importance Sampling
    - Adaptive Rejection Sampling
    - Sampling-Importance Re-sampling SIR
  - 高斯分布采样
    - 步骤
    - Box-Muller
    - 拒绝采样
    - 逆变换法
    - 优缺点
    - 多维高斯分布
    - 截尾高斯分布 Truncated Gaussian Distribution
- 7. Optimization algorithm
  - supervised learning Loss Functions
    - Hinge
      - Subgradient Descent Method
    - Logistic Loss
    - Cross Entropy
    - 平方损失函数
    - 绝对损失函数
    - Huber Loss Functions
  - 凸优化
    - SVM
    - Linear regression
  - 非凸优化
    - 低秩分解
    - 深度神经网络
  - 无约束优化问题
    - 直接法
      - 梯度
    - 迭代法
      - 一阶法
      - 二阶法 牛顿法
        - 鞍点 Saddle Point
  - 验证目标函数梯度功能的正确性
  - 随机梯度下降
    - Mini-Batch Gradient Descent
      - 参数m
      - 学习速率
      - 挑选m个训练数据
    - 失效的原因
      - plateau
    - Batch Gradient Descent (BGD)
    - Momentum
    - AdaGrad
    - Adam
      - first moment
      - second moment
      - exponential decay average
  - 正则化与稀疏性
    - 正则化约束
    - 解空间形状
    - 函数叠加
    - 贝叶斯先验
    - 拉普拉斯先验
  - 参数优化方法
    - 网格搜索
    - 随机搜索
    - TPE Hyperopt
    - Simulated Anneal
- 6. Probabilistic Graphical Model
  - Bayesian Network
    - 联合概率分布
    - 原理
  - Markov Network
    - 联合概率分布
  - 最大熵模型
    - 原理
  - 生成式和判别式模型
  - 马尔可夫模型
    - 马尔可夫链
    - 隐马尔可夫模型
    - 中文分词建模和训练
    - 最大熵马尔可夫模型 Maximum Entropy Markov Model (MEMM)
      - 标注偏置
    - 条件随机场 Conditional Random Field (CRF)
  - 主题模型
    - Probabilistic Latent Semantic Analysis
    - Latent Dirichlet Allocation
      - 语料库生成过程
      - 主题个数
        - 困惑度 perplexity
      - Hierarchical Dirichlet Process (HDP)狄利克雷过程
      - 推荐系统冷启动
- 5. Unsupervised Learning
  - K-Means
    - 算法步骤
    - 优缺点，调优
    - 改进模型
      - ISODATA
        - 分裂，合并
    - Expectation-Maximization algorithms
    - 收敛性
  - Gaussian Mixed Model (GMM)
    - 核心思想
    - 如何迭代计算
  - Self-Organizing Map (SOM)
    - 如何工作
    - 如何设计，训练参数
  - 聚类算法评估
- 4. Dimension reduction
  - Principal Components Analysis
    - 目标函数
    - 求解方法
    - 核主成分分析KPCA
  - Linear Discriminant Analysis
    - 目标函数
    - 求解方法
    - fisher LDA
  - LDA和PCA的区别和联系
- 3. classic Algorithms
  - SVM
  - Logistic Regression
    - 多标签分类
  - Decision Tree
    - 启发函数
    - 剪枝
- 2 Mode Evaluation
  - Limition
  - precision and recall
  - RMSE失效
  - ROC
    - Draw ROC
  - AUC
  - Cosine distance and Euclidean distance
    - cosine distance是否是一个严格定义的距离
  - A/B test
    - online
    - 实验组和对照组划分
  - 模型评估方法
  - 自助法过程中，多少数据从未被选择过
  - 超参数调优
  - under-fitting and over-fitting
    - solutions
- 1 Feature Engineer
  - 数值特征归一化 normalization
  - Categorical Feature
  - 组合特征，高维组合特征
  - 特征组合的方法
  - 文本表示模型
  - Word2Vec
  - image Insufficient data

XMind | Trial Mode