

算法工程师

基础数学知识

- 1 概率分布
- 1 最大似然估计
- 2 最大后验估计
- 1 凸优化

机器学习

- 1 分类模型
 - 1 逻辑回归
 - 1 决策树
 - 2 感知机
 - 2 朴素贝叶斯
 - 1 SVM
 - 2 随机森林
- 1 回归模型
 - 1 线性回归
 - 2 SVR
 - 1 回归树
- 1 集成学习
 - 1 GBDT
 - 1 XGBoost
 - 1 LightGBM
- 2 聚类模型
 - 1 K-Means
 - 3 DBSCAN
 - 3 GMMs
 - 2 谱聚类
- 2 特征工程
 - 2 特征提取
 - 数值特征
 - 分桶
 - 交叉
 - 缺失值处理
 - 标准化与缩放
 - 数据平滑
 - 类别特征
 - 独热编码
 - 嵌入编码
 - 与数值特征交叉
 - 2 特征选择
 - 过滤式
 - 覆盖率
 - 方差分析
 - Pearson相关系数
 - 封装式
 - 特征子集
 - 嵌入式
 - 正则化
 - 基于树模型

深度学习

- 1 基础理论
 - 1 评估标注
 - 1 损失函数
 - 1 偏差与方差
 - 1 欠拟合与过拟合
 - 1 梯度消失与梯度爆炸
 - 2 迁移学习
 - 3 多任务学习
 - 1 优化技巧
 - 1 数据集划分
 - 1 提前终止
 - 1 学习率衰减
 - 1 正则化
 - 1 Dropout
 - 2 Normalization
- 1 参数调优
 - 1 超参数
 - 1 激活函数
 - 2 权重初始化
 - 1 优化算法
- 1 基本模型
 - 2 MLP
 - 1 RNN
 - 1 CNN
 - 1 Transformer
- 1 注意力机制
 - 3 Hard Attention
 - 1 Soft Attention
 - 1 Self Attention

统计自然语言处理

- 1 主题模型
 - 1 PLSA
 - 1 LDA
- 2 词袋模型
- 1 CRF
- 1 TF-IDF
- 1 中文分词

深度自然语言处理

- 1 词向量
 - 3 RNNLM
 - 1 Word2Vec
 - 2 GLOVE
 - 1 FastText
- 1 预训练语言模型
 - 1 BERT
 - 1 ALBERT
 - 2 XLNET
 - 1 Roberta

核心任务

- 1 文本分类
 - 1 TextCNN
 - 1 TextRNN
 - 1 FastText
 - 1 BERT
- 1 实体识别
- 2 关系抽取
- 1 语义理解
 - 2 CNN+CRF
 - 1 RNN+Attention
 - 1 RNN+CRF
 - 1 Online-RNN
 - 1 SLoTgated
- 1 文本相似度
 - 深度学习
 - 1 Word2Vec
 - 2 Doc2Vec
 - 1 RNN
 - 1 CNN
 - 1 Transformer
 - 字面匹配
 - 2 编辑距离
 - 2 Jaccard相似度
 - 1 BM25
 - 1 SimHash
 - 1 MiniHash
 - 主题模型
 - 3 LSI
 - 1 LDA
- 2 阅读理解
 - 2 Encoder-Interaction-Pointer
 - 1 +Self Attention
 - 1 +Pre-Trained Language Model
- 3 自然语言推理
- 2 对话/问答系统
- 2 文本生成
 - 1 语言模型生成
 - 3 GAN
 - 2 Encoder-Decoder
 - 2 同义词替换
- 1 搜索排序
 - 1 LambdaRank
 - 1 LambdaMART
 - 1 Wide&Deep
 - 2 LambdaDNN
- 1 关键词提取
 - 1 TF-IDF
 - 1 TextRank
- 1 推荐模型
 - 1 LR
 - 2 FM
 - 1 GBDT
 - 1 Wide&Deep
 - 2 DIN