

更多请关注“百面机器学习”，回复 干货 即可

神经网络中怎么降维？

标准化和归一化有什么区别？

LSTM 为啥比 RNN 在梯度消失上要好？

为啥 L1 正则先验分布是 Laplace 分布，L2 正则先验分布是 Gaussian 分布？

为什么归一化能加快梯度下降法求优化速度？

简单说下 Adaboost？

SVM 和 LR 区别和联系？

机器学习中的距离和相似度度量方式有哪些？

AUC 为啥对正负样本比例不敏感？

Xgboost 哪里可做到并行？boosting 不是串行吗？

你还知道哪些降维方法？

为什么 Drop out 有效？

说说 DL 中的 Drop out？

除了 BN 你还知道哪些归一化方法？

神经网络为啥要 BN，怎么做？

batch 的大小对训练的影响？

KNN 的实现方式有哪些？

如何处理类别不均衡问题？

你知道哪些损失函数？

PCA 和 SVD 的联系和区别？

梯度法和牛顿法的区别？

机器学习有哪些优化方法？

Bootstrap 原理以及抽样到的概率是啥？

LR 如何并行化？

树模型特征重要性怎么看？

随机森林需要交叉验证吗？

当机器学习性能不是很好时，你会如何优化？

说说为什么 svm 中的某核能映射到无穷维？

说说决策树剪枝算法？

HMM 和 CRF 区别？

深度学习中防止过拟合的方法？

逻辑回归你说说能否用来做非线性分类？

为啥集成学习方法通常情况下效果就是好呢？

随机森林的泛化能力你给说说？

说说生成模型和判别模型？

Adaboost 为啥不容易过拟合？

你说说为什么 SVM 泛化能力好呢？

你说说你都知道哪些激活函数？

机器学习模型的评价指标？

CNN 为什么用 1×1 的卷积核？



KNN 中的 K 如何选择？

LR 和 SVM 的联系区别？

机器学习中一些模型为什么要对数据归一化？

聚类算法中的距离度量有哪些？

Xgboost 和 GBDT 的区别？

分类问题中如何解决正负样本不平衡问题？

K-means 聚类算法中 K 值如何选择？

请描述极大似然估计 MLE 和最大后验估计 MAP 之间的区别？

如果训练样本数量少于特征数量，怎么办？

L1 和 L2 正则的区别，如何选择 L1 和 L2 正则？L1 在 0 处不可导，怎么处理？

用 EM 算法推导解释 Kmeans 可以吗？

为什么 LR 要使用 sigmoid 函数？

有监督学习和无监督学习的区别？

百面机器学习