

机器学习大礼包

机器学习与人工智能 带您领略世界上最前沿的智能黑科技，帮助您在机器学习与人工智能领域更好的发展。

扫描二维码，加入我们吧！

我们专注机器学习和人工智能，关注前沿技术和业界实践，旨在提供一线资源和消息。这里有最热门的新闻，这里有最专业的文章，这里有最具有价值的干货。



欢迎关注 机器学习与人工智能 公众号，获取更多咨询！

不知道各位小伙伴是否苦于缺乏项目经验，而在求职过程中屡屡碰壁。本次大礼包将带领大家了解一些常见的竞赛平台，并告诉大家如何在这些比赛中取得好的成绩，丰富自己的简历。

有哪些竞赛平台？

[Kaggle](#)

[datacastle](#)

[阿里天池](#)

在这些平台里，笔者更偏好于Kaggle，参与者更加多样化，参赛者可以在Discussion页面讨论，社区文化浓厚。

如何入门？

动手是最好的学习。笔者推荐通过Kaggle提供的入门比赛[泰坦尼克号生还者预测](#)来入门。下面给出该问题的一些解决方案，帮助大家更好的完成该比赛。

[Titanic Data Science Solutions](#)

[逻辑回归应用之Kaggle泰坦尼克之灾](#)

[kaggle数据挖掘——以Titanic为例介绍处理数据大致步骤](#)

该学什么？

常用Python库

科学计算：

[pandas](#)

[numpy](#)

机器学习：

[Sklearn](#)：可使用该网页上的搜索功能获取相关算法的具体介绍及demo.

绘图：

下面两个绘图库推荐参看gallery中自己所需的图，然后在代码上进行改进，得到自己所需的图。

[seaborn gallery](#)

[matplotlib gallery](#)

数据处理

数据可视化

[可视化方法对机器学习至关重要（系列）之一](#)

[A visual introduction to machine learning](#)

缺失值处理

[kaggle数据挖掘竞赛初步--Titanic<原始数据分析&缺失值处理>](#)

[wiki-缺失数据](#)

[机器学习中如何处理缺失数据？](#)

数据变换

[kaggle数据挖掘竞赛初步--Titanic<数据变换>](#)

[为什么一些机器学习模型需要对数据进行归一化？](#)

[再谈机器学习中的归一化方法（Normalization Method）](#)

特征选择

[机器学习中特征选择概述](#)

[机器学习中的特征——特征选择的方法以及注意点](#)

[【斯坦福---机器学习】复习笔记之特征选择](#)

[干货：结合Scikit-learn介绍几种常用的特征选择方法](#)

[Scikit-learn Feature selection](#)

模型选择

Kaggle 上获得较好成绩的模型往往都是基于树的有：

Gradient Boosting

[模型组合之 Boosting 与 Gradient Boosting](#)

[Gradient Boosting算法简介](#)

[Gradient boosting](#)

[GBDT\(Gradient Boosting Decision Tree\) 没有实现只有原理](#)

Random Forest

[RandomForest随机森林总结](#)

[Sklearn-RandomForestClassifier](#)

[机器学习中的算法——决策树模型组合之随机森林与GBDT](#)

[随机森林入门介绍](#)

交叉验证

当我们完成训练后，如何确定我们的模型能够对于测试集也有很好的拟合效果呢。交叉验证是取得好成绩的保证。

[留一法交叉验证和普通交叉验证有什么区别？](#)

[交叉验证\(Cross Validation\)方法思想简介](#)

[交叉验证](#)

集成学习

想在Kaggle众多参赛者中脱颖而出，集成学习是不二法门。上文模型选择中提到的效果好的基于树的算法，究其原因，便是他们算法本身便结合了集成学习的思想。

[集成学习基础通俗入门](#)

[集成学习原理小结](#)

[集成学习：机器学习兵器谱的“屠龙刀”](#)

常用的方法有：Bagging、Boosting、Bootstrap、Stacking

[【机器学习】模型融合方法概述](#)

[bootstrap, boosting, bagging 几种方法的联系](#)

[Boosting算法简介](#)

[总结：Bootstrap\(自助法\), Bagging, Boosting\(提升\)](#)

[集成学习——bagging and boosting](#)

[数据比赛大杀器----模型融合\(stacking&blending\)](#)

[kaggle-ensembling-guide](#)

[kaggle比赛集成指南](#)

[Kaggle机器学习之模型融合 \(stacking\) 心得](#)

