

R 机器学习初探_互评

班马

- 1、 **玉子：**涵盖了 caret 包的作业全流程，对机器学习的应用域、研究对象、以及相关算法做了简要的阐述，决策树和随机森林算法的原理也比较做了比较精简的说明，但缺少对于随机森林的优缺点部分的认识论述。
- 2、 **逍遥子晴：**对机器学习的认识部分做了较详细的阐述。包括建模过程中的数据处理步骤（数据获取，数据理解，数据处理，数据探索）的说明，相关的有无监督学习算法的分类。实例应用涵盖了 caret 包的作业全流程，在学习教材应用的过程中加入不少自己的注释和见解更有利于理解整个建模过程。作业缺少对随机森林算法的理解。
- 3、 **想：**对机器学习的定义，工作流程（业务理解，数据获取，数据预处理，特征选择，数据降维，显著性检验，强相关性检验，多重共线性检验，建木及模型训练，模型评估），算法（分类，选择流程图，优化）等多方面做了精要的阐述，看完对机器学习的认识有进一步的理解，选择流程图还没细看，要收藏多看几遍！应用实例包括涵盖了 caret 包的作业全流程也有自己的注释。作业也缺少对随机森林算法的理解。
- 4、 **HAN：**应用实例包括涵盖了 caret 包的作业全流程，代码注释较简练。对这个实例的变量关系，模型选择有较详细的阐述。对决策树算法的说明结合了应用实例，更加明了。随机森林算法的原理概要也做了说明。缺少对机器学习的认识说明。
（可能是我的网络被英文作业吓到了，这份作业下了好久囧~。）