

数据科学家笔试题目

附件 Excel 中是一张表名为 LoanData 的数据表及相关的数据字典 DataDictionary ,

请利用这张表完成以下问题：

1. 请以借款金额 (loan_amt) 为目标变量建立一个线性模型。

* 注：请将 Python/R/SAS 程序和结果单独列出，并简单解释模型结果

2. 以借款利息 (int_rate) 高低区分不同信用等级的借款人，以此为目标变量建立逻辑回归模型。

* 注：请将 Python/R/SAS 程序和结果单独列出，并简单解释模型结果

3. 以借款利息 (int_rate) 高低区分不同信用等级的借款人，以此为目标变量建立一至两个机器学习模型 (GBM、Random Forrest、Neural Network、SVM 等) 。

* 注：请将 Python/R/SAS 程序和结果单独列出，并简单解释模型结果