利用LightGBM/XGboost实现Happy Customer Bank目标客户（贷款成功的客户）识别

1. 任务说明：Happy Customer Bank目标客户识别

https://discuss.analyticsvidhya.com/t/hackathon-3-x-predict-customer-worth-for-happy-customer-bank/3802

1) 文件说明

Train.csv：训练数据

Test.csv：测试数据

2) 字段说明

数据集共26个字段: 其中1-24列为输入特征，25-26列为输出特征。

1. ID - 唯一ID（不能用于预测）
2. Gender - 性别
3. City - 城市
4. Monthly\_Income - 月收入（以卢比为单位）
5. DOB - 出生日期
6. Lead\_Creation\_Date - 潜在（贷款）创建日期
7. Loan\_Amount\_Applied - 贷款申请请求金额（印度卢比，INR）
8. Loan\_Tenure\_Applied - 贷款申请期限（单位为年）
9. Existing\_EMI -现有贷款的EMI（EMI：电子货币机构许可证）
10. Employer\_Name雇主名称
11. Salary\_Account - 薪资帐户银行
12. Mobile\_Verified - 是否移动验证（Y / N）
13. VAR5 - 连续型变量
14. VAR1- 类别型变量
15. Loan\_Amount\_Submitted - 提交的贷款金额（在看到资格后修改和选择）
16. Loan\_Tenure\_Submitted - 提交的贷款期限（单位为年，在看到资格后修改和选择）
17. Interest\_Rate - 提交贷款金额的利率
18. Processing\_Fee - 提交贷款的处理费（INR）
19. EMI\_Loan\_Submitted -提交的EMI贷款金额（INR）
20. Filled\_Form - 后期报价后是否已填写申请表格
21. Device\_Type - 进行申请的设备（浏览器/移动设备）
22. Var2 - 类别型变量
23. Source - 类别型变量
24. Var4 - 类别型变量

输出：

1. LoggedIn - 是否login（只用于理解问题的变量，不能用于预测，测试集中没有）

26. Disbursed - 是否发放贷款（目标变量），1为发放贷款（目标客户）

二、作业要求：

1. 适当的特征工程（20分）
2. 用LightGBM完成任务，并用交叉验证对模型的超参数（learning\_rate、n\_estimators、num\_leaves、max\_depth、min\_data\_in\_leaf、colsample\_bytree、subsample）进行调优。（70分）

或者用XGBoost完成任务，并用交叉验证对模型的超参数（learning\_rate、n\_estimators、max\_depth、min\_child\_weight、colsample\_bytree、subsample、reg\_lambda、reg\_）进行调优。

1. 对最终模型给出特征重要性（10分）