**专项研究：活跃用户流失预测**

数据介绍：影响活跃用户流失的普遍判断有：在线活跃、用户账号属性（性别、好友数、等级、积分等）和玩牌情况（玩牌局数、赢牌局数、输牌局数、最高牌型等）。

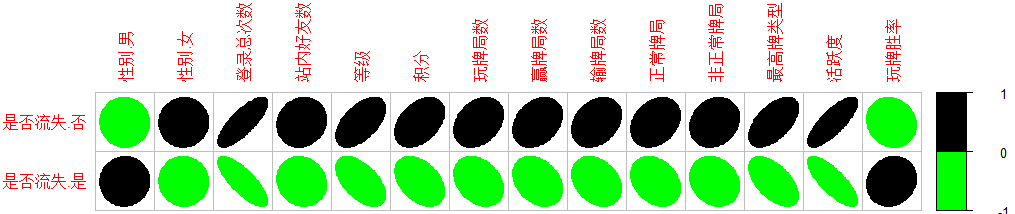
* 用户id
* 是否流失
* 性别
* 登录总次数
* 站内好友数
* 等级
* 积分
* 玩牌局数
* 赢牌局数
* 输牌局数
* 正常牌局
* 非正常牌局
* 最高牌类型

**数据样例：**



1、**数据转换：**增加两个衍生变量**：**周活跃度=登录总次数/7；玩牌胜率=赢牌局数/玩牌局数

2、**数据探索：**研究是否流失与其他变量间的关系？希望能画出以下相关系数可视化结果：



**3、数据建模：**利用十折交叉验证，利用决策树、随机森林、人工神经网络这三种分类算法建立分类模型，并通过对比结果选择最优模型最为将来的流失预测模型。