# 1、数据说明

提供的数据包含训练集和测试集。训练集是由3000 个商户构成的集合

A的交易和商户基本静态信息以及购买产品名称的标签(train.csv)。测试集是由1000 个商户构成的集合B的交易和商户基本静态信息(test.csv)。竞赛目标是提交商户集B的产品预测结果，提交样例为predict\_1.csv(需补充预测结果)。商户集A和商户集B无交集。

# 2、数据字段说明：

商户经营内容：商户的经营内容

行业大类：商户的十五大行业大类，包括零售行业、餐饮行业、服务行业等

行业小类：商户基于行业大类的进一步细分

所属区域：按商户所处地域省份的划分

收单收入：周期内，总手续费扣除发卡行服务费和转接服务费费之后，归收单机构收取的部分（已进行数据脱敏）

交易金额：周期内，商户进行交易的金额（已进行数据脱敏）

产品名称：商户购买银商服务的产品名称（共3类，一个商户可能未购买产品或者购买多个产品，多个产品之间用‘、’进行分格）。

# 3、提交样例：

对于测试集中商户产品的预测，将预测购买有对应产品的列标记为1，预测未购买的则标记为0，提交案例如predict\_1.csv所示。



# 4、评价指标：

根据每一列产品的预测准确性进行算术平均，预测准确率采用F1-score的平均值计算:



其中TP为购买产品且预测准确的商户数，FP为未购买产品但预测为购买的商户数，FN为购买产品但预测为未购买的商户数。