

问题：

为了更有效地服务客户，一家机构想建立一个机器学习模型来预测哪些客户可能会流失。他们已经收集了一些历史数据（存储在customer_churn.csv文件中），请利用这些数据以及Spark创建一个分类模型，帮助他们决定哪些新客户（记录在new_customers.csv中）可能会发生流失。并将这个模型应用到new_customers.csv的文件中所记录的客户中，来预测哪些客户可能会发生流失。

customer_churn.csv 文件中包含以下特征：

- 1.Names:** 客户的姓名。
- 2.Age:** 客户的年龄。
- 3.Total_Purchase:** 客户的总消费金额。
- 4.Account_Manager:** 是否有客户经理分配（0表示没有，1表示有）。
- 5.Years:** 客户使用服务的年限。
- 6.Num_Sites:** 客户使用的站点数量。
- 7.Onboard_date:** 客户开始使用服务的日期。
- 8.Location:** 客户的地址。
- 9.Company:** 客户所在的公司。
- 10.Churn:** 是否流失（1表示客户已经流失，0表示客户仍在使用服务）。