

案例:自动信用建模系统

背景

在消费贷款中,贷方使用信用建模方法来评估借款人的违约风险,并决定是否接受贷款申请人。机器学习技术的最新进展提高了风险评估模型的有效性,但也带来了实际挑战。对输入数据的透彻了解有助于增强对模型的信任,但此类分析任务通常需要繁重的编码工作并且非常耗时。复杂的机器学习模型难以

解释模型是如何工作的,以及为什么申请人会从模型中获得高或低的信用评分。

应进一步努力在财务决策过程中有效地建立可信模型。

在本次比赛中,参赛者将设计一个自动分析和建模系统,为消费贷款的信用评估提供可靠的机器学习模型。系统应对输入数据进行分析,以帮助用户全面了解样本和特征。该系统有望对模型和某些预测进行解释,而不是简单地提供黑盒模型。鼓励参与者使用各种可视化

提供见解的技术。

数据

1. 标签数据:包含贷款申请日期 (“APPLICATION_DATE”) 和默认标签 (“DEFAULT_LABEL”)。 “DEFAULT_LABEL” 中的值 1 代表默认值,0 代表正常值。
2. 特征数据:它包含在申请日期 (“APPLICATION_DATE”) 收集的申请人的匿名特征 (“V000” - “V999”) , 可以作为自变量来预测默认标签。它可以匹配到带有 “APPLICATION_ID” 的标签数据。

要求

原型的最低要求应满足以下标准:

1. 应提供分析功能,让用户在建模型前了解训练数据。
2. 应提供建模功能,使用机器学习算法训练至少一种分类模型并评估模型的有效性。为性能评估提供了一个测试数据集。
3. 它应该提供模型和预测的解释。