

“在线实时决策 + 离线机器学习” 实现高效实时反欺诈

文 || 明略数据 张旭 赵奕

科技的发展，帮助大数据、云计算、移动互联、区块链等现代信息技术能够为金融行业带来更广阔的发展空间和可能性，但各种金融欺诈形式也层出不穷。

银行线上渠道“反欺诈”挑战巨大

信用卡、借记卡等零售业务方面，随着征信体系和反欺诈体系建设的不断完善，银行在传统线下渠道的申请、受理阶段以及基于线下渠道交易阶段数据的反欺诈能力已经基本具备。但如今随着互联网，移动互联网的快速普及，银行陆续推出了在线业务，线上渠道在给银行带来更大创新空间、给用户带来更加便捷的使用体验的同时，在反欺诈层面也带来了新的挑战。不法分子利用伪基站发送钓鱼链接，利用木马、黑客技术或利用电信诈骗手段骗取、盗取用户线上渠道账户信息的情况层出不穷，致使近年来在银行线上渠道发生了大量非本人交易等欺诈情况，给银行、用户都造成了严重的损失。

区别于传统渠道，线上渠道的特点使得银行在使用、交易阶段进行用户身份真实性核实变得非常困难，同时线上渠道会产生高并发、海量、非结构化、多维度的数据，无论从业务角度还是技术角度对于银行的反欺诈能力都提出了更高的要求。

如何有效使用线上渠道产生的海量数据，如何基于大数据方法防控线上渠道的欺诈风险，如何利用新型技术架构解决海量、实时的问题，如何通过机器学习手段缩短反欺诈模型升级周期，这些都是银行迫切需要解决的问题。

“实时 + 离线”实现高效线上渠道反欺诈

大数据流式处理是一种新兴的数据处理技术，以“流”的形式处理交易产生的海量数据，并基于事件驱动。利用 Spark 分布式集群计算引擎，可以快速、高效地对数据进行协处理、流式处理、交互式分析等。实时根据反欺诈规则

库的规则，以及当前用户的特征数据，判断是否存在欺诈风险以及欺诈风险等级，向银行交易监控系统、处置系统输出决策结果。

基于大数据平台存储的历史数据和计算能力，根据实时采集、同步和存储的多维度客户行为数据，批量统计和分析客户历史特征，同样基于 Spark 架构的数据挖掘和机器学习平台，离线构建规则自学习模型，为客户画像系统和数据挖掘平台提供计算能力，帮助风控人员及时发现新型欺诈行为并产生对应的反欺诈规则，同时，提供整体反欺诈解决方案的资源管理和运维保障。

通过在线离线结合的解决方案，可以实时监测海量数据，有效避免线上业务中潜在的如信息泄露、冒名、盗刷等欺诈风险，最大限度的监测和防范线上反欺诈行为的发生。

明略数据线上渠道实时反欺诈解决方案

明略数据在利用分布式实时数据采集技术和实时决策引擎，处理海量高并发线上行为数据上有着丰富的经验。通过线上渠道实时反欺诈项目，可以为金融机构构建反作弊模型，识别恶意用户和欺诈行为，并实时预警和处置，引入机器学习框架通过对海量数据的分析、挖掘构建并周期性更新反欺诈规则和反欺诈模型。通过实时计算、实时决策系统，应对官网、网银、APP 等线上渠道产生的虚假账号、伪装账号、异常登录、频繁登录等新型风险和欺诈行为。

通过实时反欺诈解决方案的实施，可以提高金融机构的反欺诈能力，降低数据处理的复杂程度，提高数据处理的实时性能，并节约平台运营、开发和管理成本。

“在线实时决策 + 离线机器学习”结合的形式，还可以在其他业务上也有着广泛的应用前景，例如，对营销信息进行实时推送，进行终端在线营销，实时监控高危行为和异常交易，提供实施服务等。明略数据希望利用自己的技术优势，为金融行业解决最实际、最重要的问题。■