

【基金业数字化转型专题】易方达基金：Bond. AI 助力债券风险分析

【编者按】为深入贯彻落实党的二十大精神，引导基金行业机构践行《证券期货业科技发展“十四五”规划》，共促基金行业数字化转型，按照中国证监会总体工作部署，于2022年11月开展“证券期货业数字化转型主题宣传月”活动。通过开展“证券期货业数字化转型主题宣传月”活动，搭建交流平台，展现数字化转型成果案例，激发金融科技创新活力，营造金融科技长效发展新生态。该篇为“证券期货业数字化转型主题宣传月”系列宣传之十六。

Bond. AI 助力债券风险分析 ——易方达基金

在资产管理领域，债券信用风险和交易风险研究是非常核心和关键的步骤。随着我国债券存量规模不断增加和债券发行数量持续增长，债券违约事件不时发生，债券投资研究的价值愈发凸显，资产管理公司对相关利用科技手段解决研究问题的需求也愈加迫切。

近几年在债券投资分析领域，已经出现一些与债券研究相关的系统，但主要提供的功能为债券信用评级，侧重于简单的信息收集、聚合与整理，缺乏对信息的深度加工及分析。这类债券研究系统通常采用专家评价方法，但随着市场上债券数量的快速增长，专家评价方法需要越来越多的人力物力才能完成，且一般分析的颗粒度较粗，难以精准挖掘隐藏在表层之下的深度信息。因此，需要引入新的金融科技方法和手段，研发出一套高效实用的债券风险分析系统，提升对债券风险度量的精准性及合理度。

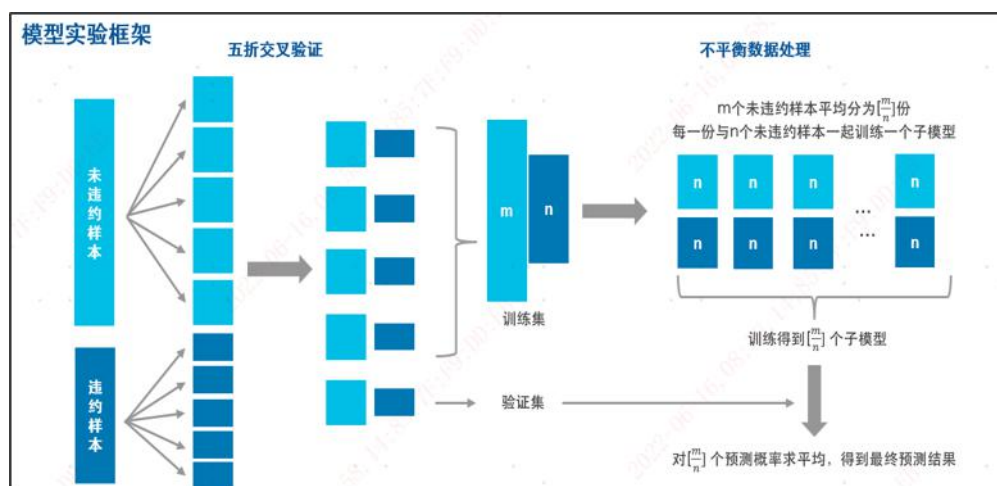
为解决长期以来债券风险分析所面临的痛点，易方达基金依靠多年的IT技术积累及业务实际需求，从债券分析的多个维度出发，自主研发了集信用评价系统、财务造假识别系统、债券流动性分析系统、债券舆情分析系统于一体的综合债券风险分析平台（以下简称Bond. AI）。

一、探索大数据及人工智能技术的应用

区别于传统的债券分析系统，Bond. AI充分运用了学术界及产业界前沿的人工智能技术。运用到的新兴技术主要包含以下三个领域：

（一）、基于小样本数据的机器学习债券投资分析技术

债券投资分析是非常重视知识和经验积累的领域，机器学习算法也是人工智能领域中非常依赖大量数据进行训练的一种方法。对于训练人工智



能模型的数据量要求来说，由于发生债券违约事件的发行人数量非常少，如何有效利用小样本数据进行有效的债券投资分析，对于Bond. AI来说是至关重要的技术问题。Bond. AI基于自主开发的机器学习算法，利用市场客观数据和投资目标之间的关系，构建多种人工智能模型协助债券投资人员进行业务开展。Bond. AI研发团队根据实际情况设计新的模型训练框架，对小样本数据进行衍生，同时对数据进度滚动切分以达到模型训练的数据要求。通过研发Easy Ensemble五折交叉实验框架，有效权衡了小样本及样本的平衡。

图1：机器学习债券投资分析技术模型实验框架

（二）、机器学习模型的可解释性技术

机器学习模型使用大量数据作为决策依据，通过模型的复杂计算，得出的结论通常难以用常规的业务逻辑去理解和解释，在实际业务应用中经常会大打折扣。因此，对机器学习模型产生的结论能够进行解释，让用户能够理解机器学习模型得出结果的依据和逻辑，是必要性、重要性和技术难度共存的一个发展方向。Bond. AI借鉴LIME, SHAP等市场主流人工智能算法的思想，独立研发针对XGBoost的模型可解释性系统。现成的算法（如SHAP）只针对单个模型做解释，但单个算法模型难以达到高准确率的应用要求，应用效果难以满足也无需求。Bond. AI研发团队针对集成模型作解释，通过研究针对机器学习模型的可解释算法并采用多模型融合方案，将一个大业务划分成多个小业务，并使用单个模型处理一个小业务，最后通过模型集成的方式，使用集成模型将各个小模型的结果集成起来，解决一个复杂的实际业务分析问题。

（三）、异构数据结构化及整合技术

债券投资分析业务需要大量的数据和信息作为决策的依据，而使用的数据来源、类型、格式、存储方式各有不一。如何高效地将不同类型的数据进行统一的结构化存储并使用，是重要的研究方向之一。Bond. AI采用异构数据结构化整合技术，可以避免使用单一来源数据，将各种结构的数据进行统一操作，形成较大类型、较多维度的数据，提供给机器学习模型进行训练和预测。通过这种技术，Bond. AI能够融合大量的不同来源的信息，提升模型的效果。系统以债券发行人为核心，构建财务的结构化数据库，同时以发行人为关键字段，构建非结构化文本数据。通过自然语言处理技术对文本数据进行分析提取，与结构化的工商、财务、风险指标等数据进行关联，综合性地构成机器学习模型的特征数据输入。Bond. AI针对不同结构数据构建子模型，并将子模型的结果数据进行集成组合，以有效支持复杂集成模型的训练和决策。利用此技术，Bond. AI得以全面分析近5000家债券发行公司的交易、财务、工商及各类网络公开数据，实时关注各类大小媒体网站的相关新闻，如果有对应的新闻和网页信息，将进行信息结构化识别和提取，并推送到舆情分析和订阅推送模型进行下一步分析处理。

从研发过程来看，基于小数据样本的人工智能模型训练及应用难度很大，

这样导致团队投入了大量的精力和时间，对模型的研发和打磨精益求精。目前大多数AI应用都是基于大数据场景的，但是对于Bond. AI系统面临着模型训练数据较小的难题。虽然债券违约事件在成熟的资本市场属于正常现象，但在国内可供参考的案例与样本不多，难以检测模型有效性，验证基准亦较难得出，针对小样本数据的机器学习技术难度非常大。本系统专门针对小数据的机器学习技术难题进行了攻关，取得了较好的成果。

二、建设综合债券风险分析平台

为加强对公司在债券投资中的信息分析处理及违约预测领域的系统化支持，使得投资研究人员能够利用最新的人工智能技术辅助决策，Bond. AI搭建了四个子系统，可实现对债券风险的全面分析。

（一）、债券舆情分析子系统

利用大数据技术对债券相关的舆情进行分析，判断舆情类事件对债券发行人的偿债能力是否有影响，同时通过对大量新闻舆情信息的深度挖掘，及时地从细微信息中提取对违约预测有价值的信息。

（二）、财务造假识别子系统

利用人工智能技术对债券发行人的财务质量进行评估，通过行业横向对比、自身历史数据纵向分析、财务造假坏样本学习分析等方式，准确快速地判断债券发行人的财务健康程度，并判断财务报表的真实可信度，用于捕捉筛选涉嫌财务造假的企业。

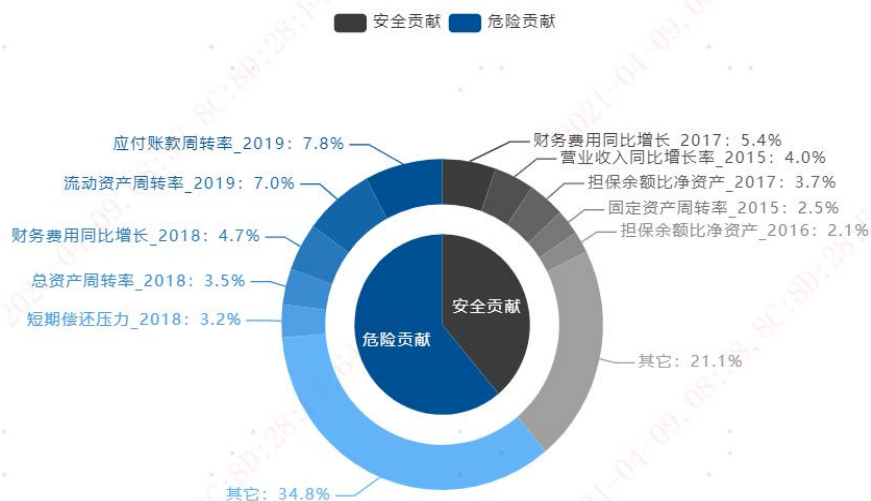


图2：财务造假识别子系统分析样例

（三）、债券流动性分析子系统

利用机器学习技术，收集相关债券数据并判断其交易的风险。通过构建针对债券交易风险的流动性预测模型，对债券流动性进行分析排名，使得交易员在交易过程中能够提前了解当前债券的潜在交易风险和价格风险，给予债券合理的定价，保障债券市场的流动性和有效性。值得一提的是，Bond. AI能够将流动性作为价格影响因素单独拆分出来，并分析其对价格的影响，让债券定价和估值更加准确。

（四）、信用评价子系统

使用债券市场上已经违约的债券发行人数据进行人工智能模型训练，并预测未来存在违约风险的企业。Bond. AI收集了上百个财务指标作为训练特征，经过机器学习模型训练之后，能对全市场的信用债发行人进行精准前置的债券信用风险分析及违约概率预测，对市场上每一家债券的发行人都能得出潜在的违约风险分析判断。

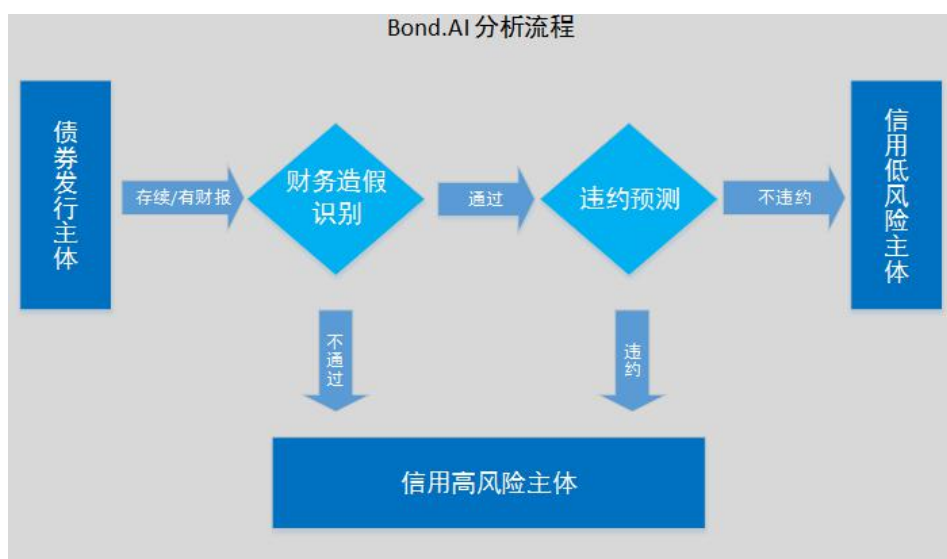


图3: Bond, AI信用风险分析流程

三、积极推广应用，助力业务发展

Bond. AI从2018年开始进行模型研究和系统研发，上线后仍持续深入迭代优化，目前功能已经基本完备，并且在公司的各项资产管理业务中推广应用，覆盖了固收交易、固收投资研究、债券投资风控、量化投资等业务领域。

以信用评价为例，Bond. AI可以协助投资人员构建债券投资备选池，利用人工智能技术针对债券发行人的违约概率给出风险评价和提示，为信用研究人员的债券发行人初筛工作提供支持，研究人员可以根据初筛结果重点研究被系统进行风险警示的债券发行人。另外，利用决策性智能系统构建黑名单业务体系，为信用研究人员提供可靠的黑名单结果。Bond. AI上线后，能够预判债券的违约概率，并且对违约债券的预测准确率长期保持在90%以上，有效减少了因债券违约造成的损失，具备较好的参考价值。此外，Bond. AI还定期提供潜在的违约发行人关联的上市公司，在股票量化投资中作为参考风险因子，可以对债券违约导致股票大跌的情况进行有效预警。

以债券流动性为例，Bond. AI可以协助交易员在债券交易流程中更加准确地判断债券交易风险、衡量债券估值、并进行风险防范。据统计，Bond. AI对没有成交信息的债券流动性预测准确度提升了10%。

从应用结果整体上来看，高质量的研究过程也输出了高准确性的分析

预测模型。随着违约债券的增加，Bond. AI也在不断的学习进化，违约预测能力持续提升，模型的风险发现能力甚至在一定程度上超过了相关从业人员。

目前，Bond. AI已取得一项软件著作权及三项发明专利，并且获得了中国人民银行颁发的“2020年度金融科技发展奖”二等奖。

四、小结

人工智能作为新兴技术，经常面临在在具体场景中难以落地应用的情况，特别是在资管领域，综合性地应用丰富的人工智能技术，并没有太多综合性应用人工智能技术的先例可供参考，工作挑战性非常大。

Bond. AI结合了易方达基金多年丰富的业务实践及技术研发经验，将人工智能技术引入债券投资中，有效地帮助了投研工作人员提高投资效率、控制投资风险。