

城乡居民养老保险风险预警模型构建及实证分析

文 | 徐玲

随着人口老龄化加速，城乡居民养老保险在基金收支、参保缴费等方面面临多重风险挑战。传统管理模式依赖经验判断，存在滞后性。本研究运用大数据、人工智能、区块链、数字孪生等技术，构建智能风险预警模型，实现从“事后补救”向“事前预防”转变，为养老保险智慧监管提供技术支撑。



(配图由 AI 生成)

风险图谱构建与数据融合框架

基于知识图谱的风险关联网络

城乡居民养老保险风险具有多维性、关联性及动态演化特征。本研究通过构建包含“风险实体-风险因素-风险事件-风险后果”的四层知识图谱架构，系统梳理出财务风险、参保行为风险、政策执行风险及系统性风险四大类别，运用自然语言处理技术，从政策文件、业务数据及舆情信息中抽取风险实体并识别风险关联关系。通过图神经网络算法挖掘风险传导路径，发现基金支付风险与人口老龄化率、缴费基数、投资收益率等因素存在显著因果关联，这为精准风险溯源和预警触发提供了知识支撑，该网络能够动态更新风险知识库，并实时捕捉新型风险特征和演化规律。

多源数据融合的智能采集体系

针对养老保险风险监测数据分散、标准不统一的现状，本研究设计了“云-边-端”三层数据融合架构。云端层整合人社部门参保缴费数据、民政部门人口数据、税务部门收入数据等政务数据，边缘层接入互联网平台就业信息、金融机构征信数据、医疗机构健康档案等社会数据，终端层部署物联网感知设备，来实时采集参保人员生存状态、活动轨迹等动态信息，通过建立统一的数据标准规范和自动化处理流程，实现异构数据的清洗、转换及加载，进而显著提升数据质量和可用性，为风险预警模型提供全面、及时、准确的数据支撑。

区块链驱动的数据治理机制

在多源数据融合过程中，数据安全和隐私保护是核心挑战，本研究引入联盟区块链技术，来构建跨部门的数据共享网络，通过智能合约实现数据访问权限的自动化管理和全

流程追溯，以此确保数据流转做到可信且可控。采用联邦学习框架，让各参与方在本地完成模型训练，仅共享加密之后的模型参数，从而实现“数据不出域、价值可流通”。引入差分隐私技术对敏感数据添加噪声扰动，在保护个人隐私的同时，还能保持模型性能，建立数据质量自动校验机制，并通过区块链存证实现数据溯源，有效保障数据安全合规及模型训练的可靠性，为跨部门协同治理提供相应的技术保障。

智能风险预警模型构建

多任务学习的风险识别模型

针对养老保险风险类型多样、任务关联性强的特点，本研究构建了基于多任务学习框架的风险识别模型。该模型采用共享底层特征提取网络和多个专用任务的架构设计，能够同时完成风险分类、风险等级评估及风险归因分析三项任务。底层网络运用 Transformer 编码器提取参保人员的时序行为特征，以此捕捉缴费模式、待遇领取规律等动态信息。针对样本不平衡的问题，采用 focal loss 损失函数和 SMOTE 过采样技术进行优化。在模型训练过程中，引入任务权重自适应调整机制来实现多任务间的协同优化，实验表明，相比单任务模型，多任务模型在风险识别准确率、召回率和 F1 值等指标上都有显著提升，且训练效率提高了大约 30%。

深度学习的动态预测算法

为实现养老保险风险的前瞻性预警，本研究设计了融合 LSTM 神经网络和注意力机制的动态预测算法。该算法通过 LSTM 层捕捉参保缴费、基金收支等时间序列数据的长期依赖关系，同时结合注意力机制自动识别关键风险的时间节点和重要特征维度，模型输入涵盖个体特征、区域经济指标、

政策变量等多维信息，能够预测未来3~12个月的风险发生概率和趋势变化。引入残差连接和批归一化技术来解决深层网络训练困难问题，采用贝叶斯优化方法进行超参数自动搜索，该算法对缴费中断风险的预测准确率达到了较高水平，平均提前预警时间能够达到数月，为管理部门采取干预措施争取了宝贵的时间窗口。

数字孪生的风险推演平台

本研究创新性地构建了养老保险系统的数字孪生体，实现物理系统与虚拟模型的实时映射和交互。该平台通过建立参保人群、基金池及政策规则的数字化镜像，能够在虚拟环境里模拟真实业务的运行状态，平台集成了风险预警模型和情景推演引擎，支持“*What-If*”分析，能模拟不同政策调整方案、经济波动状况、人口结构变化等对基金运行的影响，采用元宇宙可视化技术构建了三维沉浸式交互界面，决策者可以通过虚拟现实设备，直观感知风险态势的演变过程，平台实现了数据驱动的动态更新机制，每天自动同步真实业务数据，确保孪生体和物理系统达到高度一致，为科学决策和政策优化提供了强大的工具支撑。

模型验证与应用效果分析

模型性能对比与评估

为全面验证所构建智能风险预警模型的有效性，本研究选取某省级养老保险管理平台的真实业务数据进行实证分析，数据集涵盖数百万参保人员信息和海量交易记录。将所提模型与传统 Logistic 回归、随机森林、XGBoost 等基准方法进行对比实验，采用准确率、召回率、F1 值、AUC 等多维指标进行综合评估。如表 1 所示，本研究提出的多任务学习模型在各项指标上均显著优于对比方法，准确率和召回率均达到较高水平，在风险预测任务中，本模型的平均绝对误差和均方根误差较传统方法分别降低 35% 和 40% 以上，模型响应时间控制在毫秒级别，满足实时预警需求，跨区域泛化能力的测试表明，模型在不同经济发展水平地区均保持稳定性能。

表 1 不同模型风险识别性能对比

模型方法	准确率 (%)	召回率 (%)	F1 值 (%)	AUC 值
Logistic 回归	78.5	72.3	75.2	0.812
随机森林	84.2	79.6	81.8	0.867
XGBoost	87.6	83.4	85.4	0.891
传统精算模型	81.3	76.8	79.0	0.835
本研究模型	93.7	90.2	91.9	0.956

来源：宣城市城乡居民社会养老保险局

从表 1 可以看出，本研究提出的多任务学习模型在准确率方面比表现最好的基准方法 XGBoost 提升了 6.1%，在召回率方面提升了 6.8%，F1 值提升了 6.5%，AUC 值达到 0.956，充分证明了模型的优越性能。

典型场景的实证应用

本研究模型已在实际业务场景中得到成功应用，验证了其实用价值，在待遇冒领风险识别的场景中，模型通过关联分析生存认证数据、银行交易记录、医疗就诊信息等多源数据，成功识别出了大量疑似冒领的案例，经过人工核实，确认准确率超过了 80%，有效防范了基金的流失情况。在基金支付风险预警的场景中，模型提前数月预警了某县级统筹区基金收支失衡的风险，给管理部门争取政策调整和资金调剂的时间窗口，在缴费激励政策优化的场景中，利用数字孪生平台模拟了多种缴费补贴的方案，通过对比分析不同方案对参保率、基金收支的影响，最终选择的优化方案让参保率提升了若干个百分比，取得了比较好的政策效果。

效益分析与推广价值

智能风险预警模型的应用产生了显著的经济效益和社会效益。在直接经济效益方面，能够精准识别各类风险事件，每年能让基金损失减少达到数亿元，同时自动化的风险监测让人工审核成本降低了约 30% 以上。在间接社会效益方面，提升了养老保险基金运行的安全性和透明度，增强了参保群众对相关制度的信心，促进了参保群众缴费的积极性。模型具备良好的可推广性，在多个省份的养老保险经办机构进行推广应用，并且会根据不同地区的特点，对模型参数进行调优和功能定制，其技术架构采用模块化设计，可以灵活扩展到其他社会保险险种。

结束语

本研究构建了融合前沿技术的城乡居民养老保险智能风险预警模型，实现了风险识别、预警和治理的全流程智能化。实证分析表明，模型在识别准确率、预测精度及实时性方面显著优于传统方法，为养老保险精算管理及基金监管提供了有力技术支撑，研究创新性地将知识图谱、多任务学习、数字孪生等技术应用于养老保险领域，构建了涵盖风险监测、智能识别、动态预警、情景推演的完整技术框架。未来，需进一步加强跨险种联动监测，探索量子计算等颠覆性技术应用，强化算法伦理治理，推动养老保险风险管理持续创新发展。

作者简介：徐玲 宣城市城乡居民社会养老保险局

责任编辑：金焯 投稿邮箱：zhouhl@staff.ccidnet.com