

海南农商银行智能客服系统人工智能算法

金融应用信息披露报告

一、重要提示

本报告是依据《人工智能算法金融应用信息披露指南》对披露人工智能算法应用情况的要求，就海南农商银行基于人工智能算法提供的金融产品和服务情况进行说明，并通过我行官网进行披露。

二、基本情况

（一）机构简介

海南农村商业银行股份有限公司（以下简称海南农商银行）是在原海南省农村信用社联合社和 19 家市县法人农信社（农商银行）基础上，采取新设合并方式组建而成，是全国首家按照全省统一法人模式改革成立的地方法人银行，其前身可溯至 1951 年琼山区美兰椰子头信用社，拥有 74 年扎根守土历史。新组建的海南农村商业银行，注册资本金 220 亿元，总行内设 19 个部门，下辖总行营业部和 18 家一级支行，440 余家营业网点，员工 7000 余人。

（二）人工智能算法应用情况

产品名称：智能客服系统

服务内容：智能客服系统通过网页、微信、手机银行 APP 等多个渠道为客户提供 7*24 小时金融咨询服务，实现个性化业务解答、自动人机对话等功能。

三、人工智能算法金融应用风险治理情况

（一）组织保障

海南农商银行按照《海南农村商业银行股份有限公司信息化项目建设实施管理办法》要求，成立智能客服系统建设项目组，项目组由项目经理和项目成员组成，其中项目成员包括科技人员和业务人员共 7 人。项目组按照科学规划、严谨设计、周密部署、精心组织的原则，扎实有效推进项目建设。项目负责落实人工智能算法相关数据治理与隐私保护、模型风险和法规遵从与行业标准等方面的组织保障措施，研究、解决项目实施遇到的问题和困难，稳步推进项目进程。

（二）实施情况

智能客服系统建设项目组组织项目组科技人员和业务人员对智能客服系统人工智能技术应用风险进行专项治理，根据法规遵从与行业标准要求，对人工智能技术在金融产品智能服务过程中，在技术应用方面、数据安全方面、系统风险方面以及科技伦理等方面进行风险控制设计与风险排查，确保系统人工智能技术应用安全可控。

四、人工智能算法技术应用信息披露

（一）算法组合信息披露情况

算法组合信息披露是对金融产品和服务所集成的全部人工智能模型算法组合使用情况的整体说明，算法组合信息披露情况见下表 1。

表 1 算法组合信息披露情况

披露分	披露项	信息披露内容
-----	-----	--------

类		
算 法 组 合 信 息 披 露	算法组合清单	算法组合清单包含 4 种人工智能算法：语音识别算法、多轮对话意图识别算法、意图识别算法、对话决策算法。
	算法组合使用的开发框架	算法组合使用 TensorFlow 开发框架，Python3 开发语言。
	算法组合使用的开源软件	算法组合使用 numpy 开源软件库。
	算法组合调度机制	语音识别算法将客户语音转成文本送入意图识别算法，意图识别算法将解析输入文本中的信息识别出文本所属的意图类别，多轮对话意图识别可通过多次与系统进行交互，以最终达到用户本次交互的目的：其中包含上下文继承，对话中的槽值修改，指代消解等细分功能。
	算法组合触发条件	客户通过网页、微信、手机银行 APP 等渠道主动向机器人提问，触发算法组合。

（二）算法逻辑信息披露情况

算法逻辑信息披露是对组合中的每个算法对象逐一说明算法机理，算法逻辑信息披露情况见表 2。

表 2 算法逻辑信息披露情况

披露分类	披露项	信息披露内容
算法逻辑信息披露	算法功能说明	语音识别算法：将自然语音转换为文字序列。 多轮对话意图识别算法：完成文字序列的词法分析、句法分析、语义分析。 意图识别算法：理解客户话语及意图。 对话决策算法：基于策略决策对话输出。
	算法推理过程说明	实体识别算法对客户表达的自然语言进行关键实体抽取。多轮对话意图识别算法和意图识别算法均为意图识别模块，对客户表达的自然语言进行语义理解。
	算法推理结果说明	多轮对话意图识别算法模块主要是基于客户对话文本进行文本分类，得到概率高的意图类别；意图识别算法主要是基于客户对话文本进行文本相似度计算，与已有意图识别算法库中的

		文本进行语义匹配，得到最相似的标准语义。对话决策算法模块将多轮对话意图识别算法和意图识别算法两个模块返回的结果进行比较判别，得到得分最高的意图类别或者标准语义，从而得到对应的返回给客户的答案。
	算法技术路线选择说明	多轮对话意图识别算法模块采用基于预训练语言模型的文本分类模型。意图识别算法采用的是粗召和精排的算法路线。文本分类是自然语言处理领域的常用成熟技术，在开源模型中广为应用。粗召采用的是文本向量召回，通过无监督向量抽取模型进行向量表征提升，从而提升粗召的精准度；精排采用文本分类模型，通过粗召缩小分类模型输入类别数，通过文本分类进一步筛选意图类别。粗召和精排是意图识别算法中较为常用的成熟技术。
	算法技术成熟度说明	所采用的各类算法均为自然语言处理和语音处理领域的常用成熟技术，在金融客服等多个领域经过实践验证，

		技术稳定性和可靠性较高。
	算法重构条件说明	当算法识别准确率持续低于预设阈值、出现新的业务场景无法有效适配或技术领域出现突破性的更优算法框架时，将对算法进行重构。
	算法假设条件说明	智能客服系统保证输出的内容符合我行各类业务的知识范围，在客户对业务的任何信息有疑问时，需给出相应解答。因此假设意图识别算法库覆盖各业务的所有信息。
	算法使用限制说明	不同的领域和行业可能存在不同的意图和语境，因此需要根据具体情况对算法进行训练和调整，以适应特定领域的需求
	算法参数及超参数说明	输入参数主要包括输入文本长度限制、输出长度限制层数等，对返回结果国值设置

（三）算法应用信息披露情况

算法应用信息披露是与人工智能算法金融应用场景相关信息的说明，避免因对算法应用的错误理解而误导客户，信息披露情况见表 3。

表 3 算法应用信息披露情况

披露分类	披露项	信息披露内容
算法应用信息披露	算法应用场景	智能客服系统主要应用于我行查询业务、信用卡业务、银行卡激活、密码问题等场景的业务交互和业务引导、聊天对话、营销活动等。
	算法应用目的	利用人工智能算法实现个性化业务解答、自动人机对话等功能，统一服务标准，提升服务效率，降低客服成本，优化客户体验。
	算法应用服务范围	服务于我行全体客户，可提供查询业务办理、信用卡业务咨询、银行卡激活指引、密码问题解决等服务，适用于采用中国标准普通话的客户，不支持方言等特殊情况。
	算法应用服务前提	智能客服通过网页、微信、手机银行 APP 等多个渠道为客户提供 7*24 小时金融服务，客户需通过上述有效渠道发起咨询请求。
	算法应用获得渠道	智能客服无需下载安装，客户可直接通过网页、微信公众号、手机银行 APP

		等渠道直接使用。
	算法应用潜在风险和防护措施	智能客服采用人工智能技术，其服务质量受客户输入文字信息顺序、语句通顺度、口音等影响，可能降低意图识别率。我行通过丰富场景化的训练语料，持续提升算法模型的适应性，提高特殊情况的识别率，并在重要交互环节通过文字确认、提示等方式降低客户使用风险。
	算法应用必要性	利用人工智能算法实现全渠道、全天候的客户服务，统一服务标准，服务质量可控，从而达到降低客服成本，提升客户满意度的业务目标。
	算法应用预期效果	日均服务量同比人工服务显著提升，客户问题识别率可达 95% 以上，有效分流人工座席压力，提升整体服务效率。

（四）算法数据信息披露情况

算法数据信息披露是对算法使用的数据来源、数据采集、数据质量控制以及数据与场景的关联性进行充分说明，信息披露情况见表 4。

表 4 算法数据信息披露情况

披露分类	披露项	信息披露内容
算法数据信息披露	算法数据与金融应用相关性说明	智能客服主要采集客户输入的文字数据和语音数据，用于业务咨询解答、问题处理、营销问答等，为精准服务客户和优化业务流程提供数据支撑。
	算法数据来源说明	算法调优、测试数据集主要来源于我行历史数据集。
	算法数据采集说明	算法模型涉及的调优数据、测试数据、推理数据均已经数据所有者（控制者）授权，采用去标识化进行脱敏处理，脱敏后无法直接关联真实客户身份。
	算法数据质控说明	建立健全严格的算法数据质量管理体系，覆盖数据采集、数据清洗、数据整理、数据标注、数据集构建各个环节，保障算法数据的完整性、一致性，数据分布的合理性、无偏性，数据样本的充足性，数据操作的规范性、合规性。
	算法组合使用的第三方软件产品	不涉及

（五）算法主体信息披露情况

算法主体信息披露是对人工智能算法金融应用服务提供者建立的算法管理相关机制（安全保障、风险防范、伦理治理等机制）的说明，信息披露情况见表 5。

表 5 算法主体信息披露情况

披露分类	披露项	信息披露内容
算法主体信息披露	算法主体建立安全保障机制说明	我行建立健全数据使用管理办法、敏感信息内控管理制度，明确各环节责任分工，为智能客服提供算法安全保障和数据安全保障。智能客服场景模型部署环境属于生产环境，且有操作日志保存，保障推理过程安全。生产环境模型有应急处理机制，遇到任何上线失败情况可立刻回退。
	算法主体建立风险防范机制说明	我行建立产品创新管理办法、数据使用管理办法，形成“事前审核准入、事中监控预警、事后处置问责”的风险防范及补偿机制，作为智能客服的责任主体确保客户权益得到有效保障。
	算法主体建立伦理	我行成立产品创新委员会和产品创新

	治理机制说明	办公室，审议本行涉及业务创新、技术创新的管理制度，审议数字化转型过程中的模式创新、机制创新等方案和制度。
	算法主体建立信息披露组织实施保障措施说明	建立健全严格的算法数据质量管理体系，覆盖数据采集、数据清洗、数据整理、数据标注、数据集构建各个环节，保障算法数据的完整性、一致性，数据分布的合理性、无偏性，数据样本的充足性，数据操作的规范性、合规性。同时建立完善的信息披露内部管理制度和流程，明确信息披露的责任部门和工作要求，确保信息披露及时、准确、完整。

（六）算法变更信息披露情况

人工智能算法金融应用采用在线实时服务，由于人工干预程度越少，服务及其策略的调整可能更加频繁快速，可能引发风险性活动，应运用与之匹配的实时监测技术并及时披露相关调整以及变更信息，信息披露情况见表 6。

表 6 算法变更信息披露情况

披露分	披露项	信息披露内容
-----	-----	--------

类		
算 法 变 更 信 息 披 露	算法变更版本号	a) 语音识别算法，版本 10.0，更新时间 2022-6-16.
	算法变更原因说明	
	算法变更影响说明	b) 语音合成算法，版本 10.0，更新时间 2022-6-16
	算法变更生效时间	e) 多轮对话算法，版本 10.0，更新时间 2022-6-16 d) 意图识别算法，版本 10.0，更新时间 2022-6-16 e) 对话决策算法，版本 10.0，更新时间 2022-6-16 f) 更新内容为语音识别算法、融合多轮对话算法和意图识别算法，提升整体意图识别准确率
	算法变更实施说明	按照我行相关实施管理办法及变更交付流程等要求，开展与算法更新、软件更新的类型、内容和程度相适宜的验证与确认活动，将风险管理、可追溯分析贯穿于更新全程，形成算法产品客户接受测试报告、算法变更审批、应急处置及回退方案和上线验证报告等记录以供核查。

	算法变更保障措施	从技术架构实现、功能完备性、数据一致性、性能稳定性等方面，对算法变更前后进行差距分析并予以记录，并针对变更前后的性能差距采取人工优化、算法回退旁路等方式进行必要补救。变更后，算法依旧限制在智能客服场景的应用范围内使用，并提供自动语音播报警示提示信息。
--	----------	---

（七）算法审计信息披露情况

人工智能算法金融应用涉及数据、算力、场景等多种要素的深度融合与交互，一定程度上增加了算法金融应用的风险。为保证金融产品和服务的业务连续性，降低技术风险、操作风险，有必要对人工智能算法金融应用进行合规性审计，并及时将算法审计活动所依据的政策、法规、标准等信息以及审计结果进行说明,信息披露情况见表 7。

表 7 算法审计信息披露情况

披露分类	披露项	信息披露内容
算法审计信息	算法审计依据	依据《人工智能算法金融应用信息披露指南》、《海南农村商业银行股份

披露		有限公司开源技术应用管理实施细则》、《海南农村商业银行股份有限公司信息化项目建设实施管理办法》等政策法规和内部制度开展算法审计工作。
	算法审计信息	1. 上线前，利用算法对抗鲁棒性测试工具进行自动化算法审计，审计活动覆盖算法安全漏洞扫描、算法逻辑漏洞扫描等方面；所有严重和一般问题已完成整改。 2. 上线前，依据相关管理制度进行算法审计，审计活动覆盖算法公平性、算法可解释性、算法鲁棒性等方面。 3. 上线后，依据相关内部制度开展算法跟踪审计，审计活动覆盖算法运行情况、算法安全风险事件、算法合规情况等方面。 通过算法审计活动，确认智能客服算法组合满足算法上线条件，在运行过程中符合国家及行业相关规范和我行内部管理制度要求，未发现重大违法违规问题和安全风险隐患。
	审计内容说明	
	算法审计结论	
	算法审计频率	

五、重大事项

报告期内，海南农商银行无人工智能算法金融应用业务相关的重大诉讼、仲裁事项。

报告期内，海南农商银行未发生人工智能算法金融应用业务相关的重大案件、重大差错等情况。