



纺织业大模型测试床

引言/导读

中国联通研究院作为中国联通集团的直属二级机构，承担着国家战略、行业发展和企业生产的战略决策咨询、技术发展引领和产业发展推动的重任。不仅是原创技术发展的主力军，也是数字经济和实体经济融合创新的排头兵。在工业大模型领域，联通研究院依托中国联通元景大模型底座，面向工业领域对工业大模型的迫切需求，致力于关键技术的研究与开发，旨在创建并推广适用于垂直行业的工业大模型。

当前，新一轮科技革命和产业变革加速演进，以大模型为代表的新一代人工智能技术在能力通用性上大幅提升，大大扩展了人工智能落地的场景/行业范围、降低了人工智能落地到多种行业/场景的再定制成本。但是目前工业领域对大模型的应用主要局限在研发设计和经营管理等环节，作为工业生产的核心场景，由于生产制造环节对安全性和稳定性要求较高，大模型在该环节的渗透率不高。作为赋能新型工业化的重要支撑，大模型要与制造业深度融合，就必须将从其单一环节赋能向垂直行业多场景、全流程覆盖拓展。

在纺织业的智能化浪潮中，我们基于联通元景大模型底座，深度融合纺织业的特定数据和行业知识，进行微调与优化构建纺织业大模型。这一行业大模型将探索实现纺织业全链条智能管理，为纺织产品的全生命周期管理提供强大的技术支持。通过纺织业大模型，切实解决大模型在生产制造领域的应用难题，打破技术与实际生产之间的壁垒，完成行业大模型落地的“最后一公里”。这一探索将不仅提升纺织企业的生产效率和产品质量，更为纺织业的数字化转型和智能化升级注入新的活力。未来可逐步在更多纺织企业中推广应用，帮助更多的企业实现智能化转型，提升整体竞争力，利用大模型为纺织业迎来更加广阔的发展前景。

一、关键词

纺织业全链条智能 纺织业大模型 元景大模型

二、测试床项目承接主体

2.1. 发起公司和主要联系人联系方式

中国联通研究院

周晓龙 手机：18601103858 电子邮箱：zhouxl50@chinaunicom.cn

邸青玥 手机：18601106379 电子邮箱：diqy@chinaunicom.cn

李金博 手机：18610172852 电子邮箱：lijb7213@chinaunicom.cn

李想 手机：16619899166 电子邮箱：lix682@chinaunicom.cn

通信地址：北京市大兴区亦庄经济开发区北环东路1号 100176

2.2. 合作公司

无

三、测试床项目目标

纺织业以其产业链长且复杂著称，从原料生产至最终产品，涵盖纤维、纺纱、织造、印染等多个环节，远超一般制造业的复杂度，数据互通困难，需高度协调与管理。该行业为劳动密集型产业，减少劳动成本支出对提升竞争力至关重要，而纺织品的疵点检测、质量控制等方面较多依赖传统的人工方式，效率低下且质量难以保证，需要采用计算机视觉检测技术快速识别瑕疵和问题，提高质检效率和准确性。同时，纺织业高度市场化和国际化，出口导向显著，受国际市场波动和贸易政策影响大，需要及时对市场政策进行分析给出相应对策。在技术层面，纺织业展现技术创新与传统工艺并存的独特风貌，既采用先进自动化控制和新材料，又保留传统技艺。然而，纺织业也是高污染行业，面临水污染、高能耗及固体废弃物等环境问题，环保转型压力大，需要采用人工智能技术优化生产参数和工艺流程，降低能耗和排放。此外，纺织业产品极具多样性和时尚性，快速响应市场需求，这要求企业不断适应市场变化，进行柔性生产。综上所述，纺织业以其独特的产业链结构、劳动密集性、国际化程度、技术创新与传统融合、环境压力及市场时尚性，与其他行业形成鲜明对比，亟需引入工业智能的新技术来解决这些问题。

纺织业大模型测试床是专为推动纺织业智能化、高效化发展而精心打造的解决方案。与通用大模型不同，该测试床基于联通大模型底座，充分利用纺织业的专业数进行微调，以确保其在纺织领域的知识具备高度的专业性和针对性。通过集成大模型技术与纺织生产流程，实现了纺织品全生命周期的智能化管理，例如，该测试床可在原材料生产阶段，使用大模型收集并分析各种纤维材料的市场趋势、性能和成本；在织布阶段，根据纤维特性和市场需求，利用大模型推荐最优的织法；在染整阶段，利用大模型进行颜色匹配和预测，确保染色效果符合设计要求；在销售与分销阶段，利用大模型分析市场趋势和消费者需求，为销售策略提供支持；在市场与服务阶段，基于大模型制定个性化的营销策略和促销活动，利用大模型进行客户服务预测和自动化应答，提升客户满意度等。总之，通过在通用大模型的基础上构建纺织业大模型测试床，可以实现对纺织业全生产流程的智能化管理，提高生产效率、降低成本、优化产品质量，并推动纺织业的可持续发展。这一举措不仅能够对纺织业产生技术升级，也为纺织业与人工智能的融合发展奠定了坚实的基础。

四、测试床方案架构

4.1. 测试床应用场景

本测试床应用领域为纺织业，涉及原材料生产、纺纱、织布、染整、后整理与精加工、辅料生产、服装设计与制衣、检测与质控、包装与物流、销售与分销、市场与服务、回收与再利用等产品全生命周期的应用场景。

4.2. 测试床架构

本测试床以联通元景大模型底座为基础，利用纺织业特定数据，针对纺织业原材料生产、纺纱、织布、染整、后整理与精加工、辅料生产、服装设计与制衣、检测与质控、包装与物流、销售与分销、市场与服务、回收与再利用等步骤，进行纺织品全生命周期的智能化管理。本测试床架构如图 1 所示：



图 1. 测试床总体架构

元景大模型底座：包括元景语言大模型、视觉大模型、多模态大模型等基础模型。其中语言大模型具备文本提取、文本召回、意图识别、参数解析、安全合规等能力；视觉大模型完成开放式目标检测和多任务视觉模型，具备场景自动适配和基于少量数据微调能力；多模态大模型具备图文理解、文生图（扩展对中文专有名词支持）、视频生成能力；语音大模型具备普通话及部分方言识别能力，可用 5 句话克隆人声等。

行业数据：使用场景数据集，融合专家规则、商品手册、操作流程、故障诊断书内容进行微调。

模型部署：具备模型量化、模型裁剪、模型编译、模型部署能力。

模型训练：基于不同的训练策略、优化策略和提示库对模型进行训练。

大模型赋能产品全生命周期：纺织业大模型赋能原材料生产、纺纱、织布、染整、后整理与精加工、辅料生产、服装设计与制衣、检测与质控、包装与物流、销售与分销、市场与服务、回收与再利用等纺织品全生命周期的步骤。

模型测试：利用基础功能验证测试、多样性和覆盖性测试、鲁棒性测试、效率和可扩展性测试、用户反馈和改进测试和安全性测试评估测试床性能。

4.3. 测试床方案

1、基础功能验证

工序验证：模拟纺纱、织布、染整、制衣等不同纺织工序，测试模型在这些环节中的表现，确保模型能够准确回答各个环节的操作问题。

行业标准验证：收集纺织业的相关标准和规范，将模型输出与行业标准和规范进行对比，确保模型的回答符合行业标准和规范。

生产设备验证：收集不同类型和品牌的纺织生产设备信息，测试模型是否能够提供准确的设备信息，包括设备性能、参数、使用说明等。

2、多样性和覆盖性验证

任务多样性测试：设计不同的任务场景，如质量检测、生产优化、库存管理等，测试模型在不同任务中的表现，评估其适应性和泛化性。

数据多样性测试：使用不同类型的数据，如原材料数据、生产过程数据、市场反馈数据等，测试模型在处理这些数据时的表现，评估其数据兼容性和处理能力。

3、鲁棒性验证

设计包含错误信息的测试用例，如输入错误的的数据、不完整的指令等，测试模型对这些错误信息的识别能力和处理能力，评估其容错性和稳定性。

4、效率和可扩展性验证

效率测试：通过测量运行模型的内存占用和推理速度，评估模型的运行效率。

可扩展性测试：测试模型在新任务、新数据时的性能表现，评估其可扩展性和灵活性。同时，测试模型在不同硬件和软件环境下的表现，确保其在各种环境下都能稳定运行。

5、用户反馈和改进验证

收集用户在使用模型过程中的反馈信息，包括模型表现、用户满意度、改进建议等。针对用户反馈的问题和建议，对模型进行改进并重新测试，确保模型能够不断优化并满足用户需求。

6、安全性验证

设计包含违法、违规等不安全信息的测试用例，测试模型对这些信息的识别能力。同时，评估模型在面临恶意攻击和数据泄露等情况下的安全性表现，确保模型能够保障用户数据和隐私的安全。

4.4. 方案重点技术

1、常见性能评估指标

本测试床在设计时充分考虑了如何准确评估模型的性能，因此采用了多种业内广泛认可的指标来衡量模型的准确性和可靠性。其中，使用准确率（Accuracy）表示模型预测正确的样本占总样本的比例。然而，在某些不平衡数据集上，准确率可能不够全面，作为补充引入了召回率（Recall），在纺织业中表示模型在检测质量问题时，能够准确识别出存在缺陷产品的能力。为了进一步综合评估模型的性能，采用了 F1 分数（F1 Score）权衡准确率和召回率，给出更全面的评估。F1 分数越高，说明模型在准确率和召回率上表现越均衡，性能越好。使用 ROC-AUC 评估模型的分​​类能力。ROC-AUC 的值越高，表明模型在各种阈值下都具有较好的分类性能，对于纺织业中需要灵活调整阈值的场景尤为重要。

在实际使用过程中，首先运行待评估的模型，收集模型在测试数据集上的预测结果。然后，根据真实标签和预测结果计算上述指标的值。通过对比不同模型在同一测试数据集上的指标值，可以对模型的性能进行客观的比较和评估，从而选择出性能最优的模型。

2、基准测试

本测试床为了确保模型的性能达到行业领先水平，采取了对比评估的方法。具体而言，将待评估的模型在标准数据集上进行测试，以获取模型在该数据集上的表现数据。这些标准数据集不仅代表了行业的普遍情况，还涵盖了各种复杂场景和变化因素，确保测试的全面性和准确性。获得模型在标准数据集上的表现数据后，将其与已有的多种方法进行对比。这些已有方法可能包括传统的经验法则，以及其他先进的机器学习或深度学习模型。通过对比不同方法在相同数据集上的性能表现，评估模型的相对优势。通过对常见性能评估指标的对比，可以发掘模型的优势和不足，进而为模型的优化和改进提供有力的指导。

3、根据用例评估检索增强生成文本质量

本测试床利用具体的用例，对模型在寻找和提取给定文档内目标信息的能力进行评测。具体来说，首先准备一系列包含不同难度和复杂度的文档样本，并在其中标注出目标信息

的位置和类型。然后将这些文档输入到待评估的模型中，观察模型是否能够准确地找到并提取出目标信息。评估过程中不仅关注模型是否能找到目标信息，还会进一步评估模型找到的信息的准确性和完整性，以全面评估模型在检索和生成方面的性能。

此外，还会考虑模型在处理不同文档类型时的表现，例如，行业标准、设备信息、维保手册等，以检验模型在不同场景下的适应能力。通过这样全面而细致的评估，能够更准确地了解待评估模型的质量，并为后续的模式优化和改进提供有力的依据。

4.5. 方案自主研发性、创新性先进性

1、创新性

纺织业大模型测试床经过微调，融合了纺织业的专业数据，确保了其在纺织领域的知识和应用上具备高度的专业性和针对性。这种针对特定行业的深度优化，使得模型能够更精准地理解和解决纺织产业中的具体问题。通过集成大模型技术与纺织生产流程，探索实现从原材料检测到生产过程控制，再到成品质量检测的全面智能化管理。

2、国产自主研发性

本测试床采用国产大模型联通元景大模型底座，该基础大模型底座在 15 个主流榜单中 7 个维度能力超 GPT3.5，通过网信办备案、安全测试通过工信部软件测评中心 A 级认证，具备先进性。该大模型底座自主研发，不依赖于外部技术保障了国家技术安全，也能够更好地保护纺织业数据的安全，防止敏感信息泄露。

五、测试床实施部署

5.1. 测试床实施规划

时间	任务	具体工作
2024 年 6 月-2024 年 9 月	需求分析，架构设计	确认需求和设计架构方案
2024 年 9 月-2024 年 12 月	搭建纺织业大模型测试床环境	开展测试床实施方案设计、数据采集、清洗与大模型微调训练验证
2024 年 12 月-2025 年 3 月	测试验证	开展测试床部署、典型应用场景验证

2025年3月-2025年5月	解决方案验证	针对纺织行业形成特定解决方案并验证，梳理并总结测试床
-----------------	--------	----------------------------

5.2. 测试床实施的技术支撑及保障措施

中国联通研究院提供技术支撑，负责测试床的设计、部署和验证方案，并联合行业客户进行应用场景落地及验证。

5.3. 测试床实施的自主可控性

本测试床所涉及的软硬件设施均由参与单位自主生产和提供，具有自主可控性。

六、测试床预期成果

6.1. 测试床的预期可量化实施结果

完成基于联通元景大模型+纺织业数据微调的纺织业大模型测试床建设，并在多个纺织企业落地应用。

6.2. 测试床的商业价值、经济效益

本测试床在商业价值方面，通过整合大模型技术与纺织生产流程实现从原材料到成品的全链条智能管理，为企业提供了实时的数据分析和预测能力，帮助管理层做出更加精准和快速的决策；通过大数据的收集和分析，企业可以深入了解市场趋势、消费者需求，从满足市场需求；通过监控生产过程中的各项参数和指标，测试床能够确保产品质量符合标准，还可以分析产品质量问题的根源，帮助企业制定改进措施，提高产品质量和稳定性。

6.3. 测试床的社会价值

1、基于国家政策，推动纺织行业智能化发展

深入贯彻总书记关于网络强国的重要思想和数字中国、新型工业化等重要论述，落实全国新型工业化推进大会会议精神，大力推进人工智能全方位、深层次赋能新型工业化。持续探索人工智能等新科技在工业领域应用的新模式、新路径，推动数字技术与实体经济深度融合，强化科技创新主体地位。纺织业大模型测试床为纺织业带来了全新的智能化生

产和管理模式，不仅提高了生产效率和产品质量，还有助于推动整个纺织业向智能化、高效化方向发展。

2、促进产业链协同与区域经济发展

纺织业大模型测试床的建设将促进纺织产业链上下游企业之间的协同合作，形成紧密的产业链生态圈。同时，大模型技术的应用将带动相关产业的发展，如大数据、人工智能等，促进区域经济的繁荣和发展。

3、基于国产联通元景大模型底座，实现工业大模型自主化供给

通过采用国产大模型底座，纺织业大模型在技术上实现了自主性，不依赖于外部技术，从而保障了国家技术安全。国产大模型底座能够更好地保护纺织业数据的安全，防止敏感信息泄露，对维护国家信息安全具有重要意义。

6.4. 测试床初步推广应用案例

选择具有代表性的纺织企业作为试点单位，进行纺织业大模型测试床的示范应用。收集试点单位的反馈意见和数据，对测试床进行持续优化和改进。通过行业会议、展览等渠道，向纺织企业展示纺织业大模型测试床的优势和应用效果。与行业协会、科研机构等建立合作关系，共同推动纺织业的智能化转型。

七、测试床成果验证

7.1. 测试床成果验证计划

本测试床对基于联通元景大模型+纺织业数据微调的纺织业产品全生命周期智能化管控大模型进行测试验证。

7.2. 测试床成果验证方案

编号	应用	验证内容
1	功能验证测试	验证模型在纺织工序、行业标准文档检索、生产设备信息输出等基本能力
2	多样性和覆盖性测试	验证模型在不同任务（如质量检测、生产优化、库存管

		理等) 和处理不同类型数据 (如原材料数据、生产过程数据、市场反馈数据) 时的能力
3	鲁棒性测试	验证模型对错误信息的识别能力
4	效率和可扩展性测试	验证模型的评估内存占用, 推理速度, 扩展性
5	用户反馈和改进测试	针对用户的反馈信息, 模型改进的效果
6	安全性测试	模型针对违法等不安全的信息识别能力

八、与已存在 AII 测试床的关系

本测试床和之前已经审批的测试床无关联。

九、测试床成果交付

9.1. 测试床成果交付件

最终交付成果应包含:

- 1、1 套基于联通元景大模型的纺织业大模型, 验证成功的标准是可以在该模型上实现纺织业全流程智能化管理并通过测试来验证各项指标;
- 2、1 份测试报告, 包括整体及各种关键技术的测试报告;
- 3、1 项纺织业大模型实际应用相关专利;
- 4、1 项以上纺织行业解决方案。

9.2. 测试床可复制性

本测试床可以应用于纺织业生产全流程, 能够显著提升生产的稳定性和可靠性, 提升生产效率。本测试床的解决方案和相关技术能够适用于纺织业不同的生产工序和需求, 是大模型赋能纺织业商品全流程的案例, 为纺织业相关应用场景提供参考依据, 可广泛用于纺织企业的商品全流程品质把控。

9.3. 测试床开放性

本测试床是对纺织业赋能全流程大模型的一种测试验证，与现有的工业生产独立。后续将继续开放给更多纺织企业共同参与测试床的验证和推广工作。

十、其他信息

10.1. 测试床使用者

中国联通研究院负责具体测试床方案设计与部署、后续运维等技术支持工作。欢迎 AII 成员企业参与和使用本测试床项目，参加单位需要对测试床的建设提供必要的验证设备或必要的验证场景。在初始阶段，本测试床将仅限于现有的成员单位。之后，逐步开放给更多的 AII 成员企业使用。

10.2. 测试床知识产权说明

中国联通研究院对测试床的建设、运营以及使用拥有产权。

10.3. 测试床运营及访问使用

本测试床将部署于中国联通研究院内，由中国联通研究院运营，并提供相关技术支持。对测试床的访问使用需经过申请使用，并通过付费、知识产权共享等多种方式有条件使用。

10.4. 测试床资金

测试床的资金来源于测试床参与单位。

10.5. 测试床时间轴

任务描述	时间进度	详情
沟通交流、行业调研	2024 年 6 月-2024 年 7 月	调研当前纺织行业的市场需求、技术趋势和竞争态势，掌握具体需求，评估大模型技术在纺织业的应用潜力和优势，形成调研文档。
需求分析、可行性分析、	2024 年 7 月-2024 年 9 月	根据调研结果，设计纺织业大模

架构设计		型测试床的系统架构，进行可行性分析。
模型搭建、企业应用	2024 年 9 月-2025 年 2 月	选择元景大模型底座作为技术基础，根据纺织业的特点和需求，对大模型进行定制化设计，面向示范企业进行测试。
应用效果评估	2025 年 2 月-2025 年 4 月	针对纺织业形成特定解决方案并验证，梳理并总结应用效果。
宣传推广	2025 年 4 月-2025 年 5 月	宣传推广，扩大应用范围，形成国际领先、可推广的应用。



工业互联网产业联盟
Alliance of Industrial Internet