徽标, 公司名称

描述已自动生成

**2024年（第17届）**

**中国大学生计算机设计大赛**

人工智能实践赛作品报告

作品编号：　　　　　　　202405010010

作品名称：　　　　　　　“明甄谛听”

基于对抗式网络的AIGC检测系统

填写日期：　　　　　　　2024年4月30日

**目 录**

[第1章 作品概述 1](#_Toc165906872)

[1.1 作品背景 1](#_Toc165906873)

[1.2 主要功能与特色 1](#_Toc165906874)

[第2章 问题分析 3](#_Toc165906875)

[2.1 问题来源 3](#_Toc165906876)

[2.2 现有解决方案 3](#_Toc165906877)

[2.3 本作品要解决的痛点问题 4](#_Toc165906878)

[2.4 解决问题的思路 4](#_Toc165906879)

[2.4.1 作品功能和性能需求 4](#_Toc165906880)

[2.4.2 数据集概述 5](#_Toc165906881)

[第3章 技术方案 6](#_Toc165906882)

[3.1 总体介绍 6](#_Toc165906883)

[3.2 数据预处理 7](#_Toc165906884)

[3.3 模型与系统方案设计 8](#_Toc165906885)

[3.4 多模态检测 11](#_Toc165906886)

[3.5 结果输出与用户交互 12](#_Toc165906887)

[3.6 用户个性化模型 13](#_Toc165906888)

[3.7 其他技术 15](#_Toc165906889)

[第4章 系统实现 16](#_Toc165906890)

[4.1 用户界面 16](#_Toc165906891)

[4.2 数据训练 17](#_Toc165906892)

[4.3 改进过程 17](#_Toc165906893)

[4.4 系统部署方式 17](#_Toc165906894)

[第5章 测试分析 19](#_Toc165906895)

[第6章 作品总结 20](#_Toc165906896)

[6.1 作品特色与创新点 20](#_Toc165906897)

[6.2 应用推广 20](#_Toc165906898)

[6.3 作品展望 21](#_Toc165906899)

[参考文献 22](#_Toc165906900)

# 

# 作品概述

## 作品背景

随着信息技术的飞速发展，人工智能已经深度融入各个行业领域，为人们的生活带来便利。但是随着生成式AI的广泛应用，其带来的便利性和风险性并存。

近期，有关规范人工智能的各项法规陆续出台，例如由国家网信办等七部门联合公布的《生成式人工智能服务管理暂行办法》实施，这些说明当前社会需要逐步规范化生成式人工智能的使用。在**2023年12月开展了主题为《基于成都市市民群体的有关生成式人工智能发展与管控主观愿意调研》的市场调查**分析。分析发现，目前现有的AI作品监测系统发展尚不成熟，大多仅限于文案监测功能，且对中文内容的监测准确率仍有待提高。此外，系统还缺乏对图像、音频、视频等其他媒体内容的检测能力。同时，国内外检测网站的收费普遍较高，不利于中低端市场的普及和应用。

由此，致力于开发一款可以对AI参与工作程度进行检测的多平台检测系统，当前项目**已经在AI绘图检测以及文案的检测中取得了明显效果**。

## 主要功能与特色

明甄谛听系统具有文本检测、图像检测、音频检测三大模块。文本监测中包含了智能检测、学术类检测、文学类检测。图像监测包含智能检测和情感分析。音频监测具有智能检测和声音特征。分析用户上传的文件，识别其中的AI参与度，给出反馈结果。

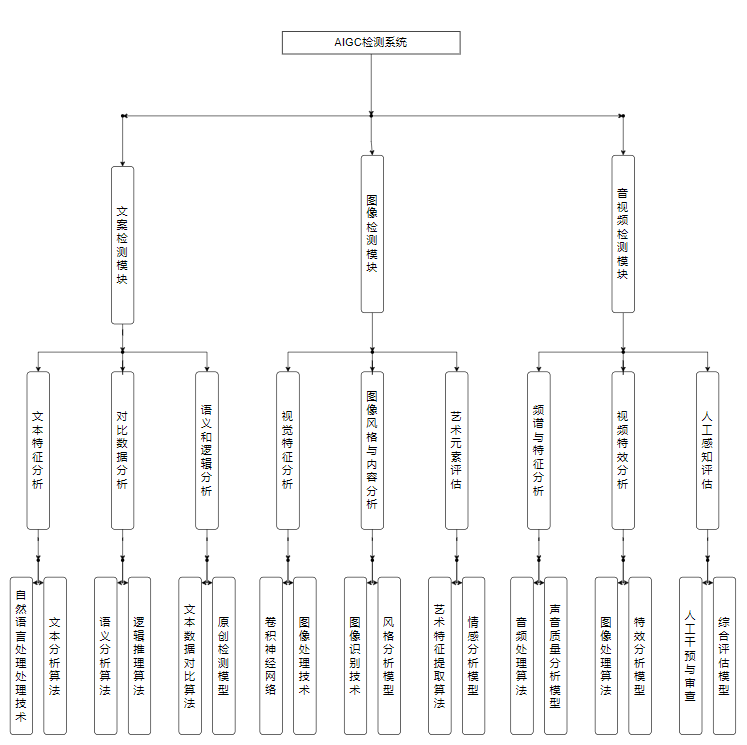


图 1-1 项目功能总体架构

眀甄谛听系统是一个多模态的AI参与度检测系统。该系统能够全面覆盖文本、图像和音频三种模态的数据，识别和分析其中的AI参与度。同时结合先进技术，让系统具有数据安全性、检测精确度高、成本可控的特点。

# 问题分析

## 问题来源

近年来，生成式AI技术的迅猛发展为人们的生活带来了诸多便利，与此同时，也伴随着一系列问题的浮现。在创作与版权领域，生成式AI生成的内容涉及原创作品，可能引发版权争议和法律纠纷，对创作者和版权所有者的权益造成潜在威胁。在信息传播领域，生成式AI被用于书写媒体内容，可能会导致虚假信息的广泛传播，对社会公众的认知产生误导。在个人信息安全方面，诈骗分子可能会利用AI技术模拟他人的声音和形象，进而实施诈骗行为，严重威胁个人信息安全。在学术领域，生成式AI自动生成文本抄袭部分论文，将造成学术诚信问题的频发，损害学术研究的可信度和声誉。

## 现有解决方案

当前，虽然市面上已涌现出若干AI检测系统，如龙源AI检测、知网学术文献检测系统、StealthWriter及ORIGINAUTY.AI等，但这些系统大多聚焦于文案监测功能，尤其以英文内容为主。对于中文内容的监测，其准确率仍有待进一步提升。此外，这些系统普遍缺乏对图像、音频、视频等多模态数据的检测能力，使得在多模态数据的AI参与度检测方面面临较大挑战。且这些系统的收费较贵，在中低端市场有较大的开拓空间。

表2-1 眀甄谛听系统与其他竞品的分析

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 龙源AI检测 | StealthWriter | ORIGINAUTY.AI | 知网学术文献检测系统 |
| 功能覆盖 | 中英文本检测 | 英文文本检测 | 英文文本检测 | 中英文本检测 |
| 用户体验 | 收费较贵 | 收费较贵 | 收费较贵 | 收费较贵 |
| 准确性 | 准确率高 | 准确率低 | 准确率高 | 准确率高 |
| 成本效益 | 维护成本高 | 维护成本高 | 维护成本高 | 维护成本高 |
| 可拓展性 | 可拓展性少 | 可拓展性少 | 可拓展性少 | 可拓展性少 |

## 本作品要解决的痛点问题

目前，市场上已有的AI生成作品检测系统进展相对缓慢，其主要功能**局限**于文案检测，并且检测准确率不高。现有的检测技术，如知网学术文献检测系统（已加入AI文本检测功能）和某些小型检测网站（主要是国外网站），在中文内容检测方面的效果并不理想。而且，现有的AIGC检测系统主要专注于学术文本，缺乏对图像、音频和视频等其他类型数据的全面检测。此外，国内检测网站的**收费普遍较高**，使得**中低端市场（如大量本科学业论文和中学作业检测）难以承受**。

眀甄谛听检测系统的设计正是基于这样的市场需求。该系统通过结合生成对抗网络模型和联邦学习技术，构建了一个多模态的AI参与度检测系统。这种创新设计使得系统能够全面处理文本、图像和音频三种模态的数据，有效识别和分析其中的AI参与度。同时，**通过自定义模型训练以及封闭式动态学习循环可以大幅提供系统的可拓展性以及成长性**。

眀甄谛听检测系统在用户数据安全和隐私保护方面也给予了高度关注。采用联邦学习技术，系统可以在不直接传输用户数据的情况下进行模型训练和优化，从而有效保障用户数据的安全和隐私。这一设计使得眀甄谛听系统在满足市场需求的同时，也符合了当前对数据安全和隐私保护的高标准。

## 解决问题的思路

### 作品功能和性能需求

眀甄谛听系统的设计旨在解决当前AI生成作品检测系统的痛点问题，实现多模态数据的全面检测，提高检测准确率，并保障用户数据的安全和隐私。为实现这些目标，系统需要具备以下功能和性能需求：

一、多模态数据支持

眀甄谛听系统应能够处理文本、图像和音频三种模态的数据，实现对这些数据中AI参与度的全面检测。系统应具备强大的数据处理能力，能够高效地识别和分析多模态数据中的关键信息。

二、高精度检测

为提高检测准确率，眀甄谛听系统需要引入中文预训练模型和针对中文检测的优化。通过利用先进的自然语言处理技术和计算机视觉技术，系统应能够准确识别AI生成的内容，避免误判和漏判的情况。

三、良好的用户体验

系统应具备简洁明了的操作界面和流畅的使用体验，方便用户快速上手。同时，系统应提供详细的检测报告和解读，帮助用户更好地理解检测结果。

四、高效性能

眀甄谛听系统需要具备高效的处理能力，能够快速完成大量数据的检测任务。此外，系统还应具备良好的稳定性和可扩展性，以适应不断增长的数据量和用户需求。

五、数据安全与隐私保护

为保障用户数据的安全和隐私，眀甄谛听系统应采用联邦学习技术。通过在不直接传输用户数据的情况下进行模型训练和优化，系统可以有效防止用户数据泄露和滥用。同时，系统还应采取其他安全措施，如数据加密、访问控制等，确保用户数据的安全性和完整性。

### 数据集概述

为了训练和验证眀甄谛听系统，使用多个来源的数据集。这些数据集包括文本、图像和音频三种模态的数据，涵盖了不同领域和场景。数据来源主要包括公开数据集的数据以及系统用户上传的数据。在数据获取方式上，遵循合法、合规的原则，确保数据的合法性和可靠性。

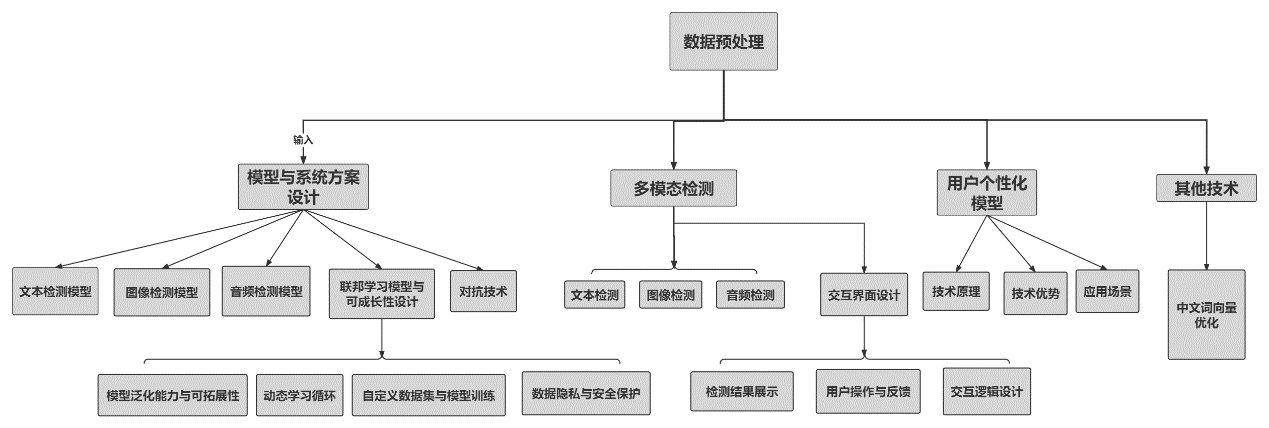
数据特点方面，这些数据集具有多样性、丰富性和真实性等特点。多样性体现在数据来源的多样性以及数据类型的多样性；丰富性体现在数据规模的庞大以及数据质量的优质；真实性体现在数据反映了真实场景下的AI参与度情况。

数据规模方面，使用了数万条数据进行训练和验证。这些数据涵盖了不同领域和场景下的AI生成作品，为系统的准确性和泛化能力提供了有力保障。

# 技术方案

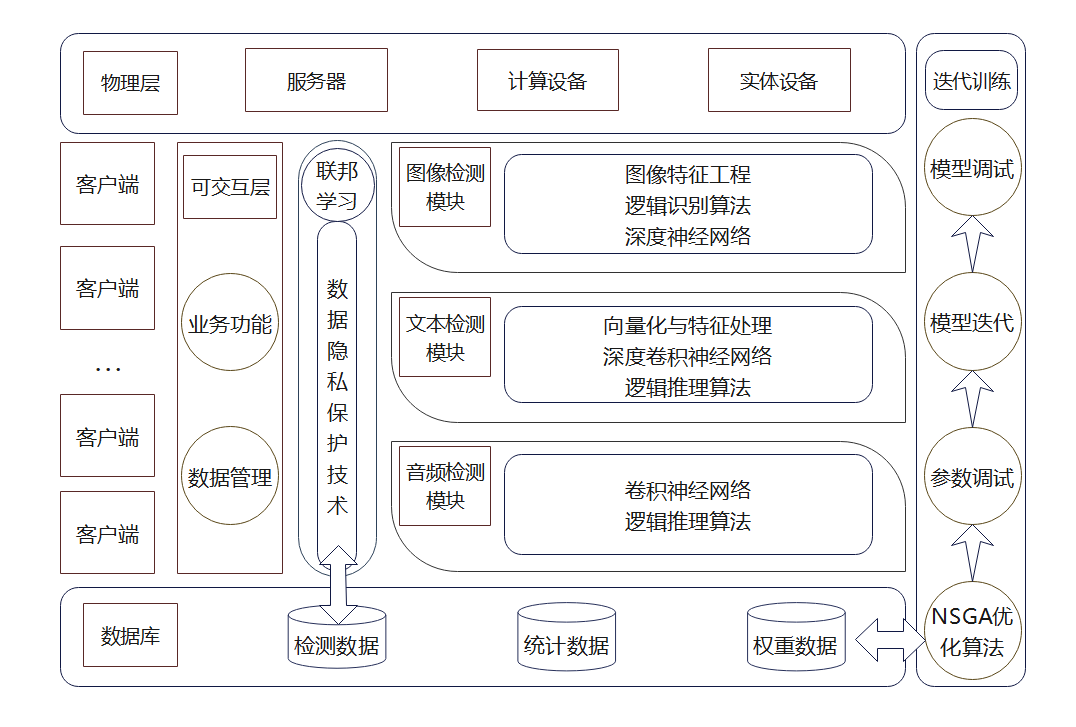
## 总体介绍

眀甄谛听系统的设计采用了先进的对抗网络模型和联邦学习技术，以实现对多模态数据的全面检测，在提高AI参与度识别精确度的同时保障了用户的个人隐私信息安全。系统的技术路线框架图如下：



**图3-1 系统的技术路线框架图**

该框架图展示了本系统的核心技术组成，包括数据预处理、模型与系统方案设计、多模态检测、结果输出与用户交互以及用户个性化模型等技术模块。



**图3-2 项目技术架构图**

以上技术架构图主要展示了本系统运行逻辑，用户在可交互层进行基础功能的使用，然后用户指令会由前端到后端，再由计算服务器处理计算信息，返回结果。用户使用后的信息会以加密的方式保存在数据库中，检测结果将结合用户反馈为模型提供下次迭代的数据。下面将详细介绍本系统的技术路线。

## 数据预处理

数据预处理是系统的重要组成部分，负责将原始数据转化为适合模型处理的格式。采用了多种数据预处理技术，包括数据清洗、格式化、归一化等，以确保数据的准确性和一致性。此外，对图像和音频数据进行了特征提取和降维处理，以提高检测效率和准确性。

文本类信息：

在模型训练之前，首先要对文本数据进行清洗，以去除无关信息、纠正错误和去除重复内容。这包括去除HTML标签、停用词过滤、纠正错别字和消除文本中的噪声。统一文本的格式，如将所有文本转换为小写，以保证模型训练的统一性。将文本分解为更小的单元，如单词或短语。中文文本通常需要分词来识别独立的语义单元。使用大量本地文本数据自行训练的词向量模型Word2Vec将文本中的单词转换为向量表示，这些向量能捕捉单词的语义和上下文信息。然后结合使用TF-IDF方法来提取文本特征，这些特征将用于后续的模型训练。接着使用PCA（主成分分析）对文本特征进行降维，以减少模型训练时的计算复杂度，同时尽可能保持数据的信息量。

图像类信息：

从原始图像中提取出所需的信息，如颜色、纹理、形状等。然后使用图像增强技术改善图像质量，如旋转、缩放、裁剪和颜色调整，以提高模型的泛化能力。使用深度学习模型（如卷积神经网络CNN）自动提取图像特征，这些特征能够捕捉到图像中的关键信息。对提取的图像特征使用PCA进行降维。

音频类信息：

从音频文件中提取出波形数据，进行频谱分析，以识别音调、节奏等特征。

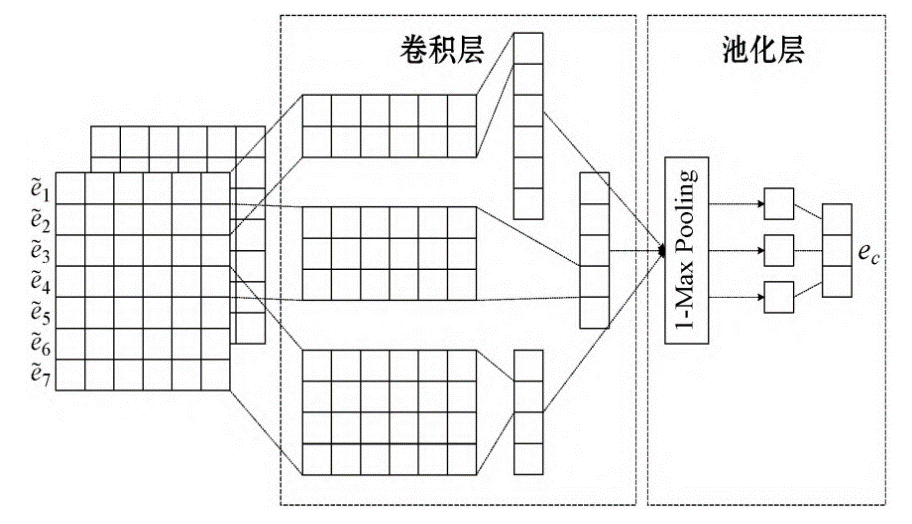
利用MFCC（梅尔频率倒谱系数）声音特征提取技术来提取音频数据的特征。对提取的特征使用PCA进行降维。

对于所有类型的数据，预处理的最后一步通常是数据归一化，这通过对数据进行标准化处理，确保数据在模型训练过程中能够更好地收敛。

## 模型与系统方案设计

在模型设计阶段，采用了深度卷积神经网络模型作为特征提取模型。该模型作为判别器模型，通过不断迭代优化，能够充分理解当前市面上已有的生成器模型的生成特征。

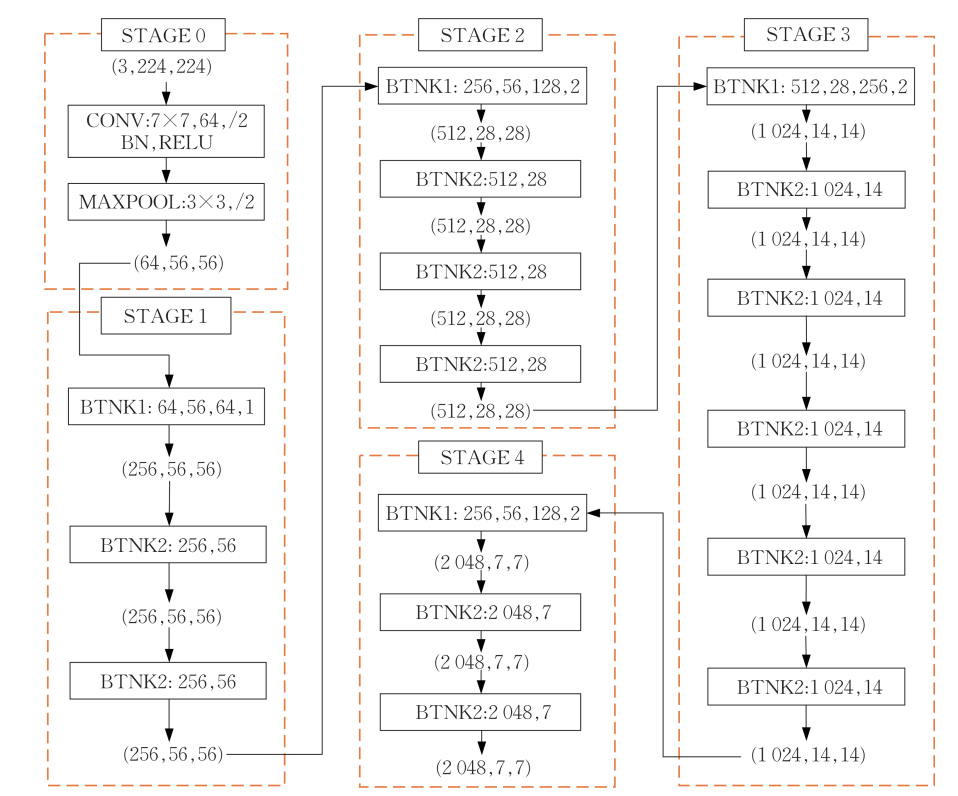
1、文本检测模型

采用门卷积神经网络（GCNN）作为预测模型。其使用多层卷积层、多层全连接层和池化层来提取和转换文本特征，并最终输出分类结果。

**图3-3 文本预测卷积神经网络图**

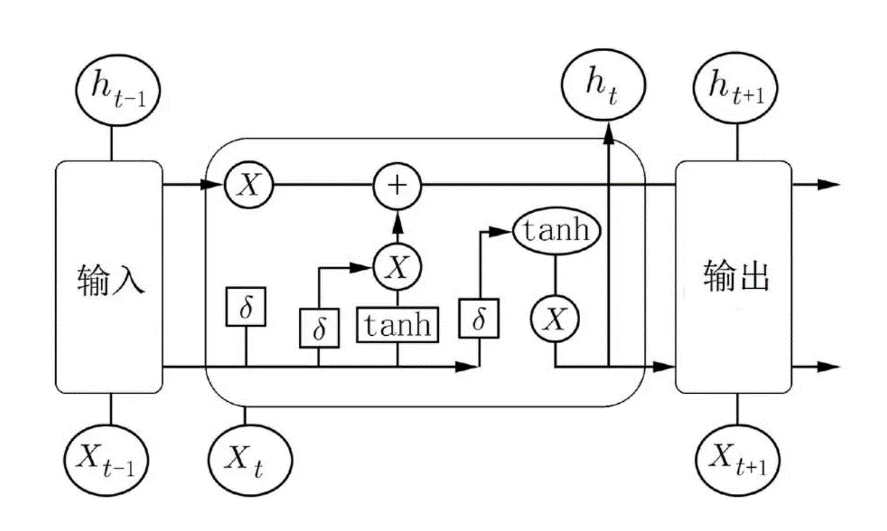
2、图像检测模型

采用残差神经网络（ResNet）作为预测模型。其架构包括多个残差块，全局平均池化层和全连接层。可以用于各种图像识别任务。



**图3-4 残差深度卷积神经网络图**

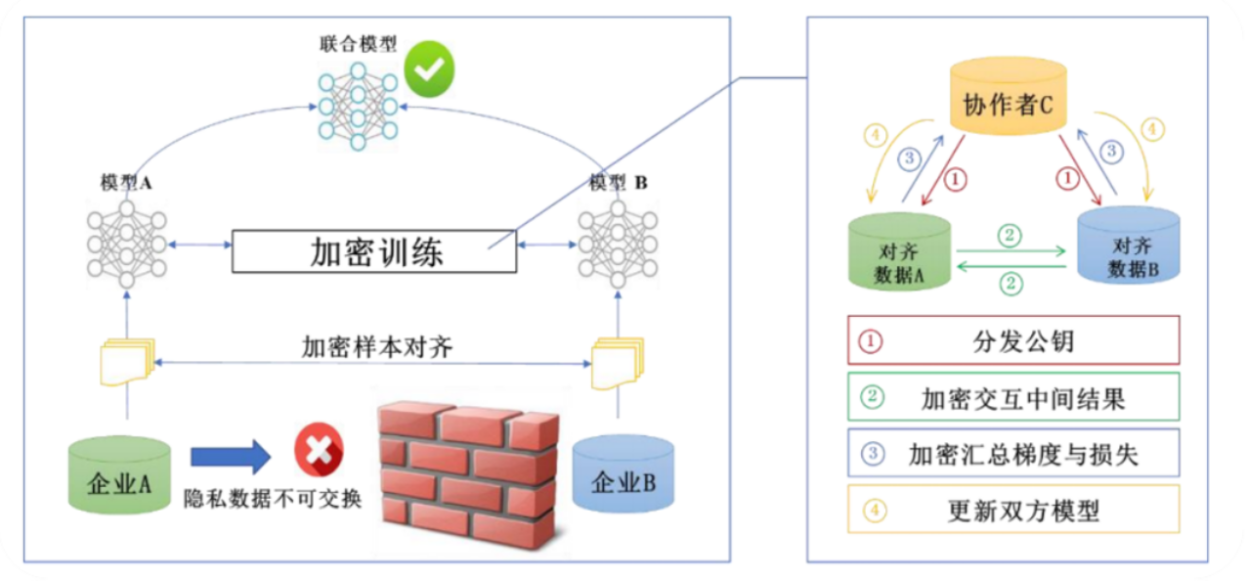
3、音频检测模型

采用长短期记忆神经网络（LSTM）作为的音频检测模型。其具备捕捉时间序列数据中的长期依赖关系的特点。模型采用LSTM层来处理音频信号，并引入注意力机制以聚焦关键信息。最后，通过全连接层和Softmax分类器输出分类结果。

**图3-5 长短期记忆神经网络图**

4、联邦学习模型与可成长性设计

联邦学习（Federated Learning）是一种机器学习技术，它允许多个设备（客户端）协作训练一个共享模型，同时保持各自的数据在本地，不离开设备。这种方法可以保护用户的隐私，同时提高模型的泛化能力。



**图3-6 联邦学习技术示意图**

在完整系统的使用页面，添加用户预测反馈选项，用户可以对自己的检测结果逐段反馈，反馈的数据将作为新的训练数据，在通过恶意判别模型自筛和人工精筛后，将用于下一轮模型的迭代。

此外，由于系统具备轻量级训练的子模型，所以为不同领域的客户提供自定义训练。客户可以按需求采集少量数据，然后在用户操作界面选择自定义模型训练的功能，这将为用户提供个性化的部署服务。

在本系统中，联邦学习技术与可成长性设计相结合，提供了以下几个关键优势：

（1）模型泛化能力与可拓展性：通过将多个客户端的数据聚合在一起，联邦学习模型可以学习到更加泛化的特征，从而提高模型的准确性和鲁棒性。随着用户数量的增加，系统可以轻松扩展，通过增加更多的客户端来提高模型的训练质量和效率。

（2） 动态学习循环：用户的预测反馈将被实时收集并用于模型的训练数据。在数据聚合之前，系统会通过恶意判别模型和少量的人工筛选来确保反馈数据的质量和安全性，从而保证了训练数据的可靠性。这种动态的学习循环允许模型根据用户的实际使用情况不断进化，提高了模型的适应性和实用性。

（3）自定义数据集与模型训练：用户可以提供自己的数据集，并参与模型的训练过程。这使用户能够根据自己的特定需求定制检测模型，提高了模型的定制化和个性化。

（4）数据隐私与安全保护：由于模型训练是在用户的设备上本地进行的，因此用户的个人数据不会离开设备，从而保护了用户的隐私。在模型训练过程中，通过加密技术确保数据在传输过程中的安全性，防止数据被未授权的用户访问。

5、对抗式技术

以生成对抗网络模型为基础核心，该网络模型一般是一个生成模型与判别模型之间进行对抗与训练得出的，当生成模型的仿真能力超过了判别模型就可以得到了AIGC模型。通过设计训练算法与对抗逻辑，训练出一个可以应对生成模型的判别模型。目前得出的模型便是Beta版预测模型——“谛听DIS-1.0”

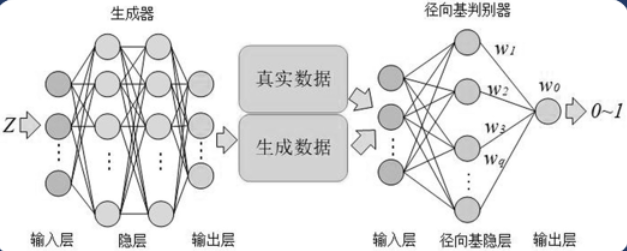


图3-7 对抗式网络技术示意图

本系统的对抗逻辑核心是大规模的多样性数据采集，鉴于各大AIGC模型企业是不可能开放模型的核心生成逻辑的，因此结合调研情况以及设计成本等多方面因素，最终设计了采集大规模多样性数据的对抗逻辑。对于数据采集任务进行了详细的精细划分，例如对于本文数据，从经典文学、近代文学、历史文学、网络文学、学术研究、新闻资讯、推文博客等多个方面进行了采集。对于数据的长短，又从软件评论、知乎问答、论坛等多个采集了不同长度的数据。由此构建了具体可行对抗逻辑。最终训练的模型预测结构也确实达到了预期。

## 多模态检测

多模态检测是本系统的核心业务模块。在实际应用场景中，信息往往以例如文本、图像和音频等多种形式存在。为了更全面、准确地检测和识别这些信息，设计了文本、图像和音频多模式的检测功能作为多模态检测模块。

1、文本检测

文本检测主要用于识别和定位图像中的文本信息。采用了门卷积神经网络（GCNN）作为预测模型，该模型通过多层卷积层、多层全连接层和池化层来提取和转换文本特征，并输出分类结果。在实际应用中，文本检测模型可以用于中低端学术文本审查、媒体平台初审等。

2、图像检测

图像检测主要用于识别和分类图像中的对象和场景。采用了残差神经网络（ResNet）作为预测模型，该模型具有多个残差块、全局平均池化层和全连接层。通过训练，图像检测模型可以识别和分类各种图像内容，如人物、动物、植物、日常事物等。在实际应用中，图像检测模型可以用于检测AI图像作品、AI修改图片以及AI伪造图像等多个场景。

3、音频检测

音频检测主要用于识别和分类音频中的声音和语调。采用了长短期记忆神经网络（LSTM）作为音频检测模型，该模型能够捕捉时间序列数据中的长期依赖关系。通过引入注意力机制，音频检测模型可以聚焦关键信息，并输出分类结果。在实际应用中，音频检测模型可以用于AI音乐检测、AI伪造音频等多场景。

在多模态检测模块中，将这三种检测功能进行了有效的集成，使得系统能够同时处理和分析文本、图像和音频等多种模态的信息。这种集成化的多模态检测方式，不仅提高了系统的检测效率，也使得系统具备了更强的信息处理能力。联邦学习

## 结果输出与用户交互

在结果输出与用户交互模块中，重点关注了用户交互体验，旨在为用户提供一个直观、易用的交互界面。以下是用户交互技术模块的详细介绍：

1、交互界面设计

采用了简洁明了的界面设计，使得用户能够快速上手并理解检测结果。界面主要包括以下几个部分：

（1）标题栏：展示系统名称、版本信息等基本信息。

（2）菜单栏：提供用户操作入口，如查看检测报告、重新检测、自定义训练等。

（3）结果展示区：以图表、列表等形式展示检测结果，包括检测准确率、AI参与度等信息。

（4）详情查看区：展示检测报告的详细内容，包括各项指标的解析和建议。

（5）状态栏：显示系统运行状态、提示信息等。

2、检测结果展示

采用了图表和列表相结合的方式展示检测结果，以满足不同用户的需求。以下是一些关键的展示内容：

（1）检测准确率：以图表形式展示文本、图像、音频三种模态的检测准确率。

（2）AI参与度：以列表形式展示各模态中AI参与的占比和具体实例。

（3）错误分类统计：以图表形式展示各分类的错误数量和占比。

（4）模型训练进度：以进度条形式展示模型训练的实时进度。

3、用户操作与反馈

为了提高用户体验，提供了以下操作和反馈功能：

（1）查看检测报告：用户可以点击查看详细的检测报告，包括各项指标的解析和优化建议。

（2）重新检测：用户可以随时进行重新检测，以获取最新的检测结果。

（3）自定义训练：用户可以参与模型训练，提供自己的数据集，定制检测模型。

（4）结果导出：用户可以将检测结果导出为PDF、CSV等格式，方便进行后续分析。

（5）在线客服：用户提供问题反馈和咨询在线客服的功能，以便解决在使用过程中遇到的问题。

## 用户个性化模型

1、使用背景

用户个性化模型训练模块是针对特定领域的数据集，利用系统所提供的轻量级模型进行用户自定义的检测模型训练的技术模块。该技术模块旨在为用户提供高度个性化的数据分析和处理能力，帮助用户更好地理解和使用本系统，从而提高业务效率和准确性。

2、技术原理

（1）数据预处理：用户需要上传特定领域的数据集，系统将对数据进行预处理，包括数据清洗、格式转换、缺失值处理等，以确保数据的质量和可用性。

（2）模型选择与配置：系统提供多种轻量级模型供用户选择，如轻量级门卷积神经网络（GCNN）等，用户可以根据自己的需求进行配置。

（3）模型训练与优化：用户可以根据自己的需求自定义训练参数，如学习率、批次大小、优化算法等。系统将根据用户提供的参数进行模型训练，并实时监控模型的训练状态和性能。

（4）模型评估与反馈：训练完成后，系统将对模型进行评估，并根据评估结果为用户提供反馈和建议。用户可以根据反馈信息对模型进行进一步优化。

3、技术优势

（1）轻量级模型：使用轻量级模型进行训练，可以减少计算资源和时间成本，提高训练效率。

（2）个性化定制：用户可以根据自己的需求自定义训练参数和模型结构，从而实现高度个性化的检测模型。

（3）高效业务：通过实时监控和评估模型的性能，可以快速发现和解决训练过程中的问题，提高模型的准确性和稳定性。

（4）易于使用：用户只需上传特定领域的数据集，无需具备复杂的编程和深度学习知识，即可轻松使用该模块。

4、应用场景

该技术模块适用于各种需要个性化数据分析和处理的场景，如电商、医疗、金融、安全等领域。用户可以根据自己的需求上传特定领域的数据集，使用系统所提供的轻量级模型训练自己的检测模型，以提高业务效率和准确性。

## 其他技术

中文词向量优化：

针对中文数据的特点，进行了多方面的优化和改进。首先，引入了中文预训练模型，以提高对中文数据的处理和理解能力。其次，针对中文文本的特点，设计了适合中文的文本特征提取和表示方法，以更好地捕捉文本中的语义信息。此外，针对中文语音和图像数据进行了特殊处理，如语音的音节分割和图像的中文文字识别等，以提高多模态检测的准确性和效率。

为了进一步提高中文检测的准确性，采用了基于迁移学习的方法。利用在大量中文数据上预训练的模型，将其知识迁移到检测任务中，从而实现对中文数据的更好理解和处理。此外，采用了多任务学习的方法，将不同的检测任务联合起来进行训练，以提高模型的泛化能力和鲁棒性。

# 系统实现

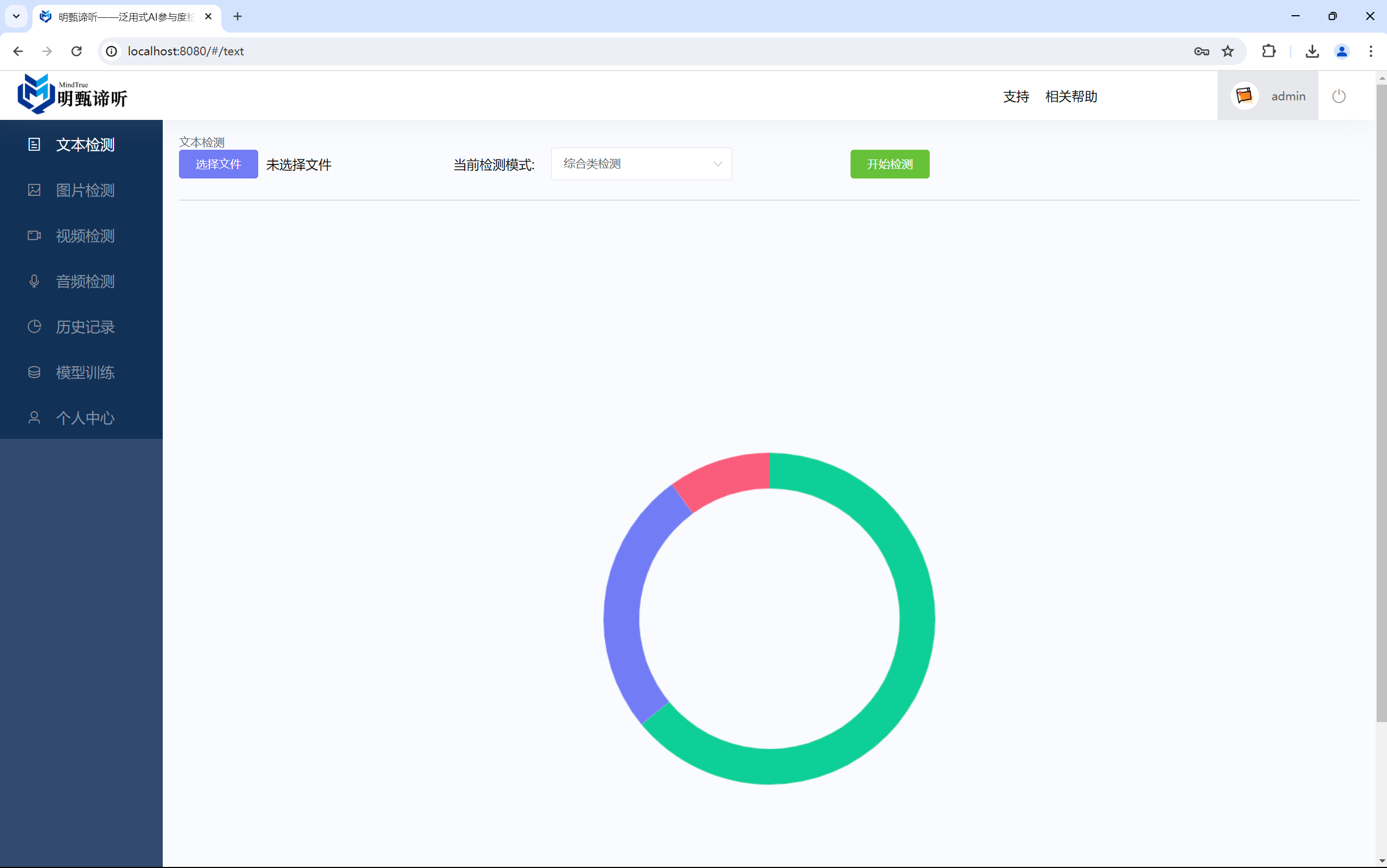
## 用户界面

用户界面是眀甄谛听系统与用户进行交互的重要桥梁。设计简洁、直观的用户界面，方便用户进行操作和查看检测结果。在界面设计上，充分考虑了用户的使用习惯和需求，提供了丰富的交互功能和可视化展示，使用户能够轻松掌握系统的使用方法和检测结果。

在界面的布局上，采用清晰的模块划分，包括数据上传、检测参数设置、检测结果显示等模块。用户可以通过简单的拖拽和点击操作，完成数据上传和参数设置，系统则会根据用户设置自动进行检测，并将结果显示在界面上。同时，提供了多种可视化展示方式，如折线图、柱状图等，帮助用户更直观地了解检测结果。

在数据上传方面，支持多种格式的数据导入，包括文本、图像、音频等。用户可以通过界面上的文件选择按钮，选择需要检测的文件进行上传。系统会对上传的数据进行自动解析和处理，以适应后续的检测任务。

在检测参数设置方面，提供了丰富的参数选项，以满足用户不同的检测需求。用户可以根据自己的需求，设置检测模型、检测范围、阈值等参数。系统会根据用户设置自动调整检测策略，以达到最佳检测效果。

在检测结果显示方面，提供了详细的检测报告和解读。报告包括检测准确率、AI参与度等信息，帮助用户更好地理解检测结果。同时，还提供了可视化的检测结果展示，如高亮显示、标记等，方便用户快速定位和分析问题。

**图4-1 用户界面**

## 数据训练

文本数据：收集大量包含AI生成和人工编写的文本数据。这些数据具有多样性，包括科普、散文、小说、学术语言等多种类型的文字。  
 图片数据：同样地，收集包含AI生成和真实拍摄的图片数据。这些数据包括写实、动漫、3D、建筑、素描、人物、动物、植物等不同类型的图片，很好的覆盖了生活中可能遇到的各种图片  
 数据分类:关于文本数据，通过词袋模型和词向量模型，将文本转换为向量表示，以便机器学习算法可以处理。词袋模型将文本视为文字单词的集合，而词向量模型则试图捕捉单词之间的语义关系。对于图片数据，基于数据做了文件分类，然后使用标签和关键词来进一步描述图片的内容，以便更快地找到所需要的图片，然后通过卷积神经网络CNN进行特征的提取，将数据转换为可供模型训练的特征表示

训练：使用训练数据集进行模型训练。在每个训练迭代中，将输入数据提供给模型进行前向传播计算，得到预测结果，然后计算损失函数的值。接下来，使用反向传播算法计算梯度，并使用优化算法更新模型参数，使损失函数逐渐减小。在模型训练过程中，可能需要调整一些超参数，如学习率、正则化项系数、隐藏层大小等。可以使用交叉验证技术，通过在验证集上评估模型性能来选择最佳的超参数组合。

## 改进过程

通过数据分析，仔细分析和理解数据集，包括数据的分布、特征之间的关系、异常值等。发现数据的特征和模式，并为后续的改进提供指导。针对数据不足或样本不平衡的情况，考虑数据增强技术，如旋转、翻转、裁剪、添加噪声等，来生成更多的训练样本，以提高模型的泛化能力。

## 系统部署方式

1.Flask模型部署：

开发模型并在本地环境中验证其准确性。使用Flask创建一个轻量级的Web服务，将模型封装为API接口。确保Flask服务可以在指定的端口上正常运行，并且能够处理前端发送的请求并返回相应的模型输出。

2.Vue前端展示：

设计并开发Vue前端应用，用于展示模型输出的结果。实现用户交互功能，如数据输入、结果展示等。使用HTTP请求（如Axios）调用SpingBoot提供的API接口，进行用户管理，权限认证等。

3.Spring Boot后端管理：

使用Spring Boot框架构建后端服务，负责用户管理、权限控制等功能。设计并实现用户注册、登录、权限认证等核心功能。

4.Spring Boot调用Flask接口：

在Spring Boot后端服务中，配置并实现与Flask API接口的通信机制。编写代码，以便在用户触发特定操作（如数据请求）时，Spring Boot能够发送HTTP请求到Flask API接口。处理Flask API返回的数据，并在后端进行适当的处理（如格式化、缓存等），然后将其发送到前端进行展示。

5.部署：

分别部署Flask模型和Vue前端应用，确保它们能够在各自的服务器上正常运行。部署Spring Boot后端服务，并配置好与SpringBoot和Vue的通信设置、Spring Boot后端访问和Flask API的通信设置。进行端到端测试，验证整个系统是否按预期工作，包括用户管理、权限控制、模型调用和数据展示等方面。

6.扩展与优化：

根据实际需求和反馈，不断优化和完善系统的功能和性能。考虑使用负载均衡、缓存等技术手段提高系统的可扩展性和响应速度。

# 测试分析

**表5-1 text detection**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| text detection | | | | |
| Epochs | MSE | RMSE | MAE | MAPE |
| 2000 | 1.20226 | 1.18417 | 7.10306 | 9.89% |
| 1000 | 1.5241 | 1.39336 | 7.77844 | 11.58% |
| 500 | 1.7658 | 1.4513 | 7.9283 | 15.65% |

**表5-2 image detection**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| image detection | | | | |
| Epochs | MSE | RMSE | MAE | MAPE |
| 2000 | 1.10226 | 1.08417 | 7.10306 | 18.49% |
| 1000 | 1.68592 | 1.61898 | 7.46277 | 21.35% |
| 500 | 2.87648 | 1.95558 | 8.1328 | 25.94% |

上述为我们已经完成部署的两个模型的测试分析，从上述数据可以看出，当前我们的模型已经具备较高的准确率，文本检测模型基本具备了市场使用的能力，但是图像检测模型由于数据量与数据类型的因素，当前预测波动较大，虽然已经具备显著的预测效果，但是投入正式使用的模型还需要持续改进。

我们将进一步收集数据，重新调整训练参数进行下一阶段的训练。

# 作品总结

## 作品特色与创新点

1. 个人隐私保护安全：眀甄谛听系统采用了成长式模型训练技术和隐私保护训练技术，确保用户数据的安全和隐私。

2. 多模态的处理识别：眀甄谛听系统具有文本检测、图像检测、音频检测三大模块，能够全面覆盖文本、图像和音频三种模态的数据，识别和分析其中的AI参与度。

3.识别准确度高：系统采用了对抗式技术，通过生成对抗网络模型进行训练和优化，得到了高准确度的预测模型——“谛听DIS-1.0”，能够更准确地识别和分析数据中的AI参与度。

## 应用推广

计划使用网页销售模式结合许可证销售模式、基于订阅的服务模式、定制化解决方案模式、授权模式等多种方式推广眀甄谛听系统。先通过市场研究了解目标客户的需求和偏好，然后制定相应的营销策略和推广计划，将眀甄谛听系统推广给更多的用户和组织，帮助他们更好地管理和使用AI技术，提高AI参与度的透明度和可解释性。

（1）通过网页销售模式，将眀甄谛听系统作为一款独立的软件产品进行销售。建立完善的销售渠道和客户服务体系，为用户提供方便快捷的购买和使用体验。同时，提供多种许可证类型，以满足不同用户的需求和预算。

（2）基于订阅的服务模式，为用户提供持续的技术支持和更新服务。用户可以通过定期支付一定的费用，获得系统的使用权和技术支持，无需担心软件版本升级和技术更新的问题。这种模式可以为用户提供更加稳定和可靠的技术支持，提高用户的满意度和忠诚度。

（3）提供定制化解决方案模式，根据用户的特殊需求和使用场景，为用户量身定制符合其需求的解决方案。派遣专业的技术团队与用户进行深入沟通，了解用户的需求和痛点，然后提供相应的技术解决方案，帮助用户解决实际问题。

（4）授权模式，与其他企业或组织合作，将眀甄谛听系统的技术授权给他们使用。这种模式可以扩大市场份额和影响力，同时也可以帮助其他企业和组织提高AI技术的使用效率和透明度。

## 作品展望

未来，将继续优化和改进眀甄谛听系统，提高其检测准确性和效率。同时，还将探索更多的应用场景，如智能客服、智能推荐等，进一步拓展系统的应用领域。眀甄谛听系统将在未来的AI领域发挥更加重要的作用，为人们的生活和工作带来更多的便利和效益。

# 参考文献

[1]梁木,李方,许怡. 推动人工智能安全高效发展[N]. 经济日报,2024-05-05(003).

[2]曲忠芳. 中国AI大模型的差异化发展之路[N]. 中国经营报,2024-04-29(D04).

[3].AIGC技术在学术出版应用的学术诚信与风险管控研讨会在京举行[J].科技与出版,2024(03):95.

[4]杨春霞,吴佳君,瞿涛,等.基于注意力机制和CNN的多标签文本分类模型[J].计算机应用与软件,2024,41(03):156-162.

[5]傅梦希,朱效宇,张良,等.基于深度残差神经网络的光场PIV粒子场重建方法研究[J/OL].光学学报:1-21[2024-05-07].http://kns.cnki.net/kcms/detail/31.1252.o4.20240424.1056.050.html.