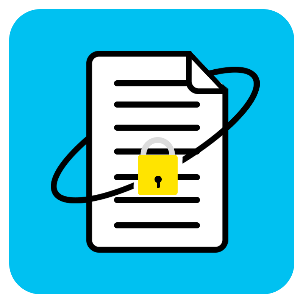
遇到困难不摆烂

2022.9.25 lyz

**软件使用说明书**

深度学习数据污染检测系统



# 目录

[目录 1](#_Toc115042202)

[1 概览 4](#_Toc115042203)

[2 界面功能 4](#_Toc115042204)

[2.1 主界面 4](#_Toc115042205)

[2.1.1界面功能 4](#_Toc115042206)

[2.1.2 快速检测 5](#_Toc115042207)

[2.1.3 文件类型选择 5](#_Toc115042208)

[2.1.4 数据类型选择 5](#_Toc115042209)

[2.1.5 对比数据集 6](#_Toc115042210)

[2.1.6 选择方式 6](#_Toc115042211)

[2.1.7 确认开始检测 7](#_Toc115042212)

[2.1.8 检测进度 7](#_Toc115042213)

[2.1.9 结果展示 8](#_Toc115042214)

[2.1.10 生成报告 8](#_Toc115042215)

[2.1.11 pdf报告详情 9](#_Toc115042216)

[2.1.12 慢速检测 9](#_Toc115042217)

[2.1.13 被污染数据 10](#_Toc115042218)

[2.1.14本地检测 10](#_Toc115042219)

[2.1.15 被污染概率 11](#_Toc115042220)

[2.2 数据管理 11](#_Toc115042221)

[2.2.1 数据管理功能 11](#_Toc115042222)

[2.2.2 修改数据类型 12](#_Toc115042223)

[2.2.3 界面变化 12](#_Toc115042224)

[2.2.4 添加数据 13](#_Toc115042225)

[2.2.5 删除数据 13](#_Toc115042226)

[2.2.6 数据被污染警告 14](#_Toc115042227)

[2.2.7 更新信息 15](#_Toc115042228)

[2.2.8 被管理文件夹 15](#_Toc115042229)

[2.3 内置浏览器 16](#_Toc115042230)

[2.3.1内置浏览器功能 16](#_Toc115042231)

[2.3.2 网页收藏 16](#_Toc115042232)

[2.3.3 网页黑名单 17](#_Toc115042233)

[2.3.4 数据下载 18](#_Toc115042234)

[2.3.5 下载提示 18](#_Toc115042235)

[2.4 日志记录 19](#_Toc115042236)

[2.4.1 日志记录功能 19](#_Toc115042237)

[2.4.2 日志备份 19](#_Toc115042238)

[2.4.3 日志重置 20](#_Toc115042239)

[2.4.4 操作类型筛选 21](#_Toc115042240)

[2.5 设置 21](#_Toc115042241)

[2.5.1 设置界面功能 21](#_Toc115042242)

[2.5.2 软件更新 22](#_Toc115042243)

[2.5.3 软件反馈 22](#_Toc115042244)

[3 定义 23](#_Toc115042245)

[4 警告与错误提示信息 23](#_Toc115042246)

[4.1 警告 23](#_Toc115042247)

[4.1.1 数据未选择警告 23](#_Toc115042248)

[4.1.2 密码输入错误警告 23](#_Toc115042249)

[4.2 错误 24](#_Toc115042250)

[4.2.1 数据路径错误 24](#_Toc115042251)

[4.2.2 找不到必要的wkhtmltopdf组件的错误 24](#_Toc115042252)

[4.2.3 配置文件查找失败错误 24](#_Toc115042253)

[4.2.4 数据库连接失败错误 24](#_Toc115042254)

[4.2.5 输出文件夹无法访问错误 24](#_Toc115042255)

[4.2.6 对比数据集错误 24](#_Toc115042256)

[4.2.7 被管理文件夹不存在错误 24](#_Toc115042257)

[5 附录一：程序附带文件规范说明 24](#_Toc115042258)

[5.1.1 依赖存储各类信息的json文件格式 24](#_Toc115042259)

[5.1.2 样式文件 25](#_Toc115042260)

[6 附录二：软件信息安全说明 26](#_Toc115042261)

[参考资料 26](#_Toc115042262)

# 概览

深度学习数据安全问题日益严重，在互联网中传播的数据很容易被污染，看似和正确数据没有什么区别，训练的模型却会受到严重影响，导致训练结果与预期差距巨大。

针对这几个问题，我们基于一个文件只有唯一的MD5码，将其正确的MD5码保存在数据库中保存，使用MD5与数据库结合，确保数据正确的MD5被保存，通过与MD5比对，鉴别数据是否被污染。将结果展示给用户，并且可以通过pdfkit生成pdf报告。

本系统使用pyQt技术构建界面，除了对数据使用前的安全检测，还为用户提供数据管理的能力，通过内置浏览器确保用户下载数据的安全，可以通过日志系统查看各类操作记录，通过设置对软件各功能进行管理。

# 界面功能

软件启动后的界面，标题栏下方为菜单栏，左侧部分为界面切换，右侧部分为界面显示位置，底部为状态栏。

## 主界面

2.1.1界面功能

默认显示的是主界面，主界面的功能为数据污染检测，选择不同的检测模式。图示

描述已自动生成

2.1.2 快速检测

点击快速检测模式，进入快速检测模式界面图形用户界面

中度可信度描述已自动生成

2.1.3 文件类型选择

选择上传文件类型，本地数据集“选择”按钮弹出对应类型文件选择框。

图形用户界面

低可信度描述已自动生成

2.1.4 数据类型选择

选择数据类型，会连接到对应的数据库，实现数据的快速查询。

图形用户界面, 应用程序

描述已自动生成

2.1.5 对比数据集

选择对比数据集时，可以直接点击“选择”按钮，进行选择。

图形用户界面, 应用程序

描述已自动生成

2.1.6 选择方式

也可以直接在输入框汇中输入，会提供近似匹配，方便选择

图形用户界面

描述已自动生成

2.1.7 确认开始检测

选择好本地数据集和对比数据集点击“确认”按钮，信息无误后即可开始检测。

图形用户界面, 应用程序

描述已自动生成

2.1.8 检测进度

开始检测后，会有进度条显示检测进度。

文本

低可信度描述已自动生成

2.1.9 结果展示

检测完成会将各类信息展示在页面上供用户查看。

图形用户界面

描述已自动生成

2.1.10 生成报告

可以点击“生成报告”按钮，弹出生成完成提示框，表示pdf报告生成完成。

图形用户界面

描述已自动生成

2.1.11 pdf报告详情

生成后的pdf报告，各类信息都会保存。

文本

描述已自动生成

2.1.12 慢速检测

点击慢速检测按钮，进入慢速检测模式界面，操作同上。



2.1.13 被污染数据

选择慢速模式，数据集详细信息会展示被污染的数据集具体信息。

图形用户界面

描述已自动生成

2.1.14本地检测

选择本地检测模式，进入本地检测模式界面，只需选择本地数据集即可。

图形用户界面, 应用程序, 网站

描述已自动生成

2.1.15 被污染概率

检测结果不再是是否被污染，而是可能被污染的概率。

图形用户界面, 应用程序

描述已自动生成

## 数据管理

2.2.1 数据管理功能

点击数据管理按钮，界面进入数据管理界面，会显示被管理文件夹下的文件信息，包括名称，最后修改时间，数据集类型以及文件大小。

图形用户界面

中度可信度描述已自动生成

2.2.2 修改数据类型

点击“管理”按钮，进入管理模式，在管理模式，可以修改数据类型。

图形用户界面

描述已自动生成

2.2.3 界面变化

进入管理模式，按钮栏会发生改变，可以对数据集进行增加删除。

图形用户界面, 应用程序

描述已自动生成

2.2.4 添加数据

点击“添加”按钮弹出文件选择窗口，选择文件后，该数据会被移动到被管理文件夹下。

图形用户界面, 应用程序, 表格, Excel

描述已自动生成

2.2.5 删除数据

选中文件点击“删除”按钮，该数据就从被管理文件夹下删除。

图形用户界面

描述已自动生成

2.2.6 数据被污染警告

当检测到文件被修改后，在该界面会增加一个警告按钮，点开会弹出警告提示

表格

描述已自动生成

图形用户界面

描述已自动生成

2.2.7 更新信息

如果不在管理模式下点击更新按钮，更新记录，提示消息在软件每次重启后都会弹出，并保存在日志中。

图形用户界面

描述已自动生成

2.2.8 被管理文件夹

点击“文件夹中打开”按钮，会打开被管理文件夹

图形用户界面, 应用程序

描述已自动生成

## 内置浏览器

2.3.1内置浏览器功能

点击左侧网页按钮，即可打开内置浏览器，



2.3.2 网页收藏

可以点击收藏将常用网站收藏，黑色实心五角星表示已收藏，

图形用户界面, 应用程序, Word

描述已自动生成

2.3.3 网页黑名单

程序默认开启钓鱼网站黑名单，当访问到这些网站后，系统会弹出警告，可以在设置中关闭。

图形用户界面, 文本, 应用程序

描述已自动生成

图形用户界面, 应用程序

描述已自动生成

2.3.4 数据下载

在该浏览器中下载数据会自动下载到被管理文件夹下，当点击下载链接会提示开始下载

文本

描述已自动生成

2.3.5 下载提示

文加下载完成后，会进行提示文件下载完成。

图形用户界面, 文本, 应用程序

描述已自动生成

## 日志记录

2.4.1 日志记录功能

点击左侧“日志记录”按钮，进入日志管理系统，在这个界面可以查看各种对软件的操作，如数据下载，数据被修改的记录等。



2.4.2 日志备份

可以将日志文件导出到其他地方用于备份， 图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

文本, 图标

描述已自动生成

2.4.3 日志重置

可以点击重置日志，但需要输入密码才能重置。

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

2.4.4 操作类型筛选

可以通过选择操作类型筛选单选按钮，选择查看不同的操作。

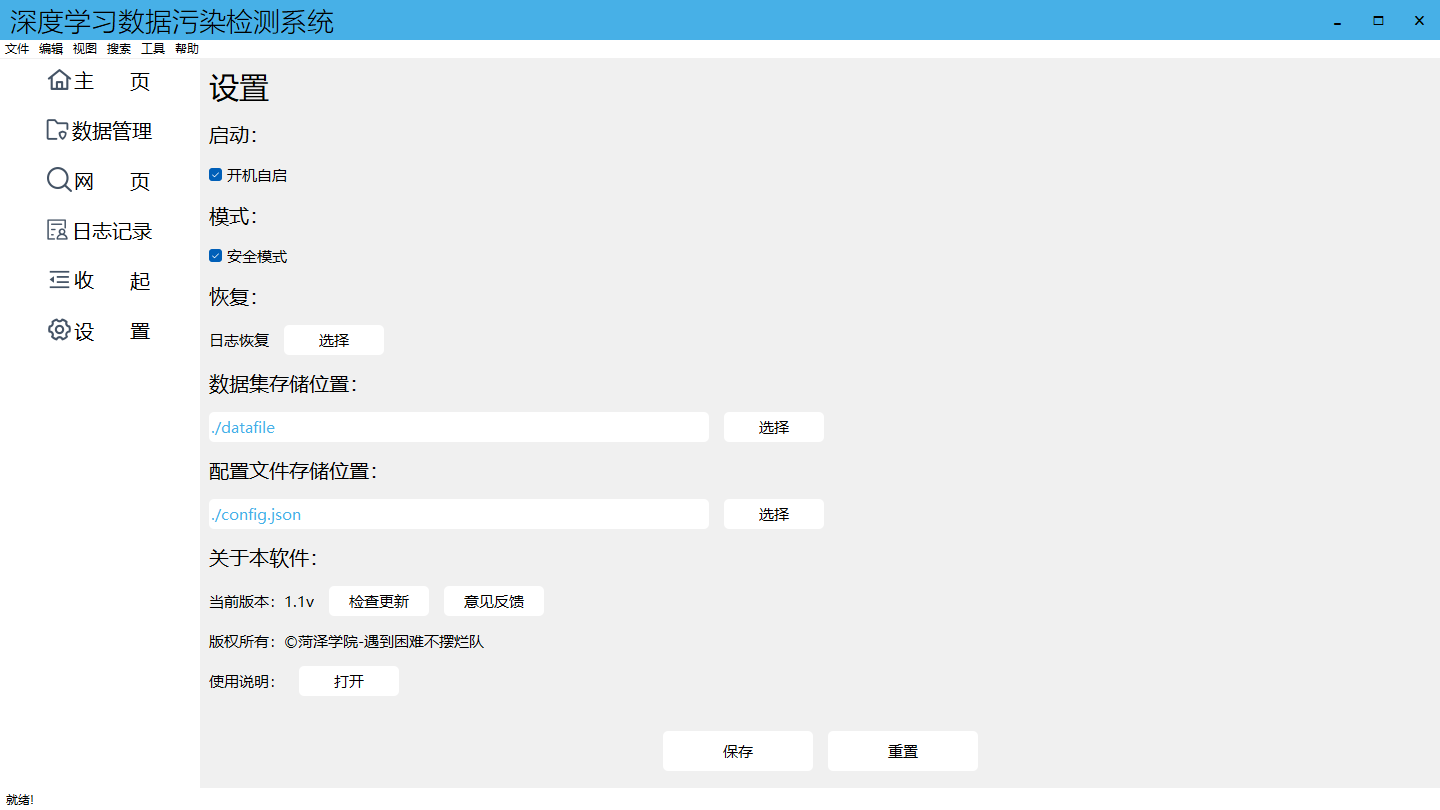
图形用户界面, 文本, 应用程序

描述已自动生成

## 设置

2.5.1 设置界面功能

在设置界面，可以对软件信息进行设置，可以选择是否开机自动启动，安全模式开启后，将会时时对被管理文件夹进行检测。当不小心将日志删除后，可以选择日志恢复，即可将备份的日志恢复。数据集存储位置就是被管理数据集的位置，配置文件的位置即软件配置信息文件的存储的位置。



2.5.2 软件更新

可以点击检查更新按钮，检查软件是否有更新，

图形用户界面

描述已自动生成

2.5.3 软件反馈

可以点击意见反馈，反馈意见。当遇到软件操作问题时，点击使用说明，即会打开软件说明书。如果修改了软件设置，需要点击保存才会保存，点击重置就会将设置恢复为默认。

图形用户界面, 应用程序

描述已自动生成

# 定义

**配置文件(\*.json) ：**是按照特定格式存储了一系列网络文本的文件。

**MD5：**信息摘要算法（英语：MD5 Message-Digest Algorithm），一种被广泛使用的密码散列函数，可以产生出一个128位（16字节）的散列值（hash value），用于确保信息传输完整一致。

# 警告与错误提示信息

由于本系统操作众多，用户不可避免地会产生若干逻辑问题。针对这些问题，我们制定了一系列警告和提示，确保用户所运行的程序在相对正确的轨道上运行。

## 警告

当用户做出错误的操作，或有违规行为时，会做出警告。

### 数据未选择警告

数据选择的提示警告

### 密码输入错误警告

重置日志时需要输入密码，密码错误会弹出警告

## 错误

### 数据路径错误

系统会对所输入的每一个路径进行有效性检验，如果发现路径错误或者访问权限不足，就会导致产生路径不存在或无法访问的错误。

### 找不到必要的wkhtmltopdf组件的错误

生成pdf需要wkhtmltopdf组件，才能正确生成pdf报告，检测不到该组件就会报错。

图形用户界面, 应用程序

描述已自动生成

### 配置文件查找失败错误

如果配置文件匹配失败或者无权访问将导致此错误。

### 数据库连接失败错误

如果无法连接到数据库，将导致此错误。

### 输出文件夹无法访问错误

如果输出文件夹匹配失败或者无权访问，将导致此错误。

### 对比数据集错误

如果无法在数据库中找到填写的对比数据集，将导致此错误。

### 被管理文件夹不存在错误

如果被管理文件夹不存在，将导致此错误。

# 附录一：程序附带文件规范说明

### 依赖存储各类信息的json文件格式

日志记录每条信息由时间，名称，类型组成

文本

描述已自动生成

数据类型由数据名和类型组成

文本

描述已自动生成

配置文件由开机自启，安全模式，下载路径，配置文件组成

文本

描述已自动生成

### 样式文件

程序样式依赖UnFrameStyle.qss文件，缺少该文件会导致程序显示错位。

# 附录二：软件相关信息说明

本程序不会上传本地数据信息，只从数据库中检索信息。本程序计算的结果仅供参考，无法保证100%正确。

程序图标来自：<https://www.iconfont.cn/collections/detail?spm=a313x.7781069.1998910419.d9df05512&cid=41678>

程序壁纸来自：

https://link.zhihu.com/?target=https%3A//tigdig.com/0314942

# 参考资料

1. Fanchao Qi et al. “Hidden Killer: Invisible Textual Backdoor Attacks with Syntactic Trigger” meeting of the association for computational linguistics (2021): n. pag.
2. Rivest. “The MD5 Message-Digest Algorithm” (1992).
3. Min Du et al. “Robust anomaly detection and backdoor attack detection via differential privacy” Learning (2020): n. pag.
4. Soheil Kolouri et al. “Universal Litmus Patterns: Revealing Backdoor Attacks in CNNs” computer vision and pattern recognition (2019): n. pag.