

RESSET 文本分析客户端-小锐助手

使用说明

RESSET 技术部

2024 年

目录

1	平台介	∽绍	4
2	平台首	页	
3	使用す	7 .	5
		- b	_
4	登求人	五	
5	文本教	7据	
5	5.1	中国上市公司财经文本	
	5.1.1	股票代码选择	
	5.1.2	时间设置	
	5.1.3	关键字设置	
	5.1.4	查询结果展示	
5	5.2	政府工作文本数据	
	5.2.1	政府地区选择	
	5.2.2	时间设置	
	5.2.3	关键字设置	
	5.2.4	查询结果	
5	5.3	美国上市公司财经文本	
	5.3.1	股票代码选择	
	5.3.2	时间设置	
	5.3.3	关键字设置	
	5.3.4	查询结果展示	
5	5.4	财经资讯文本	
	5.4.1	输入查询条件	
	5.4.2	关键字设置	
	5.4.3	查询结果展示	
6	词频搏	家	
e	5.1	输入关键词检索	
6	5.2	导入模板检索	
7	文本物	; 征	
-	7.1	文本信息	24
-	7 2	信感分析	- 25
-	73	可运性分析	
-	7 /	白宝义文本特征	29
-	7 5	ロルヘ入学習世	
/	7.5 7 E 1	17 □1日仰	
	7.5.1 7 E 2	油建口的树	
	7.5.2 7 E 2	/∿™四」	
	7.J.J 7 5 1	元于山口」	סס דכ
	7.5.4 7 F F	^{則暗} ഥ囚」	
	1.5.5	应为键%%心凶↓	

RESSET 文本分析客户端─小锐助手

	7.5.6	人口老龄化指标	
	7.5.7	员工待遇因子	
	7.5.8	贸易战潜在风险	
	7.5.9	管理会计因子	
8	文本挖	图掘	
	8 1	文本预处理	46
	811	◇+ />◇- ····································	
	812	◇+1、八····································	
	Q 1 2	大平方时	50
	8 1 <i>/</i>	ムに加る。	50
	0.1.4 8 1 5	₩ ^四 /+ 文太向畳化	
	0.1.J g J	大平内里心	
	0.2	2月11日本119	
	8.2.1	相似度订昇	53
	8.2.2	文本分类	
	8.2.3	文本聚类	59
	8.2.4	主题模型	60

小锐助手

1 平台介绍

RESSET 文本分析客户端(简称小锐助手)是一种基于计算机技术的自动化文本分析工具,它可以帮助用 户从文本中提取有价值的信息,并对其进行深入的分析和处理。客户端基于 NLP(自然语言处理)技术,使 用先进的算法和模型来对文本进行文本转换、识别、分类、情感分析、实体建模等任务。

文本预处理:对文本清洗、分词、词性标注等预处理操作。

特征提取:通过词频统计、词性标注、命名实体识别等技术实现。

模型训练:以识别文本中的模式和趋势。

情感分析:以了解用户对某个主题或产品的态度和情感。

实体建模: 识别出其中的实体关系和语义关系。

标签和分类:以方便用户进行组织和查询。

可视化分析:帮助用户直观地展示和分析文本数据。

2 平台首页



3 使用方式

下载位置,访问 <u>RESSET 财经文本智能分析平台</u>,通过在系统首页左下角点击小锐助手,点击下载。下载完成 后,将其解压,然后双击小锐助手.exe 进行安装,安装过程中 360 安全防卫会出现弹窗,选择允许程序所有 操作即可,安装完成后,系统桌面会自动生成小锐助手快捷方式,双击即可登录。



治称			修改口	用	奕型	大小
📲 小锐助	〕手.zi	р	2023/	12/13 15:16	压缩(zipped)文件	248,131 KB
	6	打开(O) 在新窗口中打开(E) 使用 Skype 共享				
		全部解压缩(T)				
		7-Zip	>	打开压缩包		
		固定到"开始"屏幕(P)		打开压缩包		>
	8	通过QQ发送到	>	提取文件		
	ai	TortoiseSVN	>		宝 助手\"	
	2	Edit with Notepad++		测试压缩包		
	6	使用 360解除占用		添加到压缩包		
		使用 360强力删除		压缩并邮寄…		
	Ð	使用 360进行木马云查杀		添加到 "小锐!	助手_2.7z"	
	Ð	使用 360管理右键菜单		压缩 "小锐助:	手_2.7z" 并邮寄	
	Ŕ	共享		添加到 "小锐!	助手_2.zip" デーロード # ######	

RESSET 锐思数据 RESSET 文本分析客户端-小锐助手 1.111 R 小锐助手.exe 192 KX, L1 740 1.00 1217 大王 R 小锐助手.exe 2023/12/13 15:01 应用程序 249,333 KB R WinRAR 自解压文件 \times _ • 点击 解压 按钮开始解压。 • 使用 浏览 按钮从目录树中选择目标文件 夹。它也可以手动输入。 • 如果指定的目标文件夹不存在,在文件解压前它将被自动创建。已在提取之前自动创 建。 目标文件夹(D) C:\Program Files 浏览**(**W)... \sim 解压进度 解压 取消 LUESTIETIS ISTOT ינ כברנו כבני 240,000 KD ІΝΤΕ∟-ΥΤ 误报反馈 Х 360安全大脑提醒您 文件防护 有程序试图修改关键程序DLL文件 文件防护 通过韵 后,系 风险程序: Range F:\05、RESSET 产品安装文件 最新 - 2019 \9、文本分析客户端\小锐助手\小锐助手.EXE 目标: F:\Program Files\小锐助手\Firefox64 \d3dcompiler_47.dll 拦截时间: 2023.12.13 16:55 通常木马会修改这些DLL文件,以此来启动自己入侵您的电 脑。如果不是您主动修改,请阻止。 极智守护 不再提醒 阻止本次操作 (29) 0 敗据文 允许本次操作 解压进度 5文本教展時和19 阻止程序所有操作 据等。文本分析 的色制标件、会计 允许程序所有操作 利数据库等。 同時 著唱

4 登录方式

用户通过双击桌面小锐助手快捷方式 ## 打开小锐助手,在登录窗口中输入用户信息。小锐助手账号同 BS 模式账号一致。





备注: 校内使用小锐助手,需要学校已购买或开通财经文本智能分析平台 B/S 模式账号,通过在登录窗口输入本校账号及密码后点击登录,助手自动识别 IP 进入小锐助手;也可通过 CARSI 登录方式登录小锐助手。

5 文本数据

5.1 中国上市公司财经文本

文本数据起止时间: 2001-至今

数据库子库包含:定期报告(季报和年报)、问询函及回复说明、IPO 招股说明书、内部控制评价报告、业 绩说明会全文、社会责任报告、上市公司重大事项公告、审计报告、风险管理业务公告以及上市公司典型案 例。

以定期报告-年度报告分类为例进行操作。

		首页	文本数据	词频搜索	文本特征	文本挖掘	设置			
中国上市公司财经文本	→ 文本数据起止时间: → 文本数据包含 定期報	2001-至今 告 (季报和年报)	间询函及	回复说明。 IPO将	2股说明书.	审计报告 自动	≥⊻文档 等11	大分类		折叠≪
政府工作文本数据	→ 文本分析: 基础字段		主题分析、	特色指标等						
美国上市公司财经文本	→ 如使用本库数据,请	添加引用:数据来)	E 信仰/T友 夏于RESSET) 「文本数据库(htt	p://rtas.re	sset.com/)				
财经资讯文本	数据库子库	定期报告		问询函及回复说	明 🔤	IPO招股说明书	ŝ 💌	内部控制评价报告		
		业绩说明会全文	NEW	社会责任报告	NEW	上市公司重大	事项公告	审计报告		
		风险管理业务公告	i	上市公司典型案	例					

▶ 定期报告 ^	😂 财经文本数据库
●年度报告	
● 第一季度报告	第一步: 输入查询条件
● 第二季度报告	↓ 股票代码选择
● 第三季度报告	 ◆ 全市场股票 ○ 从股票池选择 ○ 股票代码导入
▶ 问询函及回复说明	
▶ IPO招股说明书	
▶ 内部控制评价报告	☐ 关键字设置
▶ 业绩说明会全文	◆ 标题 ○ 全文
→社会责任报告	标题 包含任关键字(-分隔) 选择股票
▶ 上市公司重大事项公告	标题 包含所有关键字(-分隔)
▶ 审计报告	标题 不包含任关键字(-分隔)
▶ 风险管理业务公告	查询文档

5.1.1 股票代码选择

代码选择分为三种方式:全市场股票、从股票池选择、股票代码导入。

全市场股票:用户可选择全市场股票即选择全部的股票代码,不需要单独选择股票代码;

 於 股票代码选择 ◆ 全市场股票 O 从股票池选择 	≩ ○股	票代码导入			
从股票池选择: 选择左侧行业分 选区代码/单个选择代码将代码自	类,点击 动移至已选	选 ,股票代码: 区,如图:	进入中间待选区,	用户可以通过点击	金选一待
第一步:输入查询条件	○ 股票代码导入				
行业选择 筛选 ① 股票市场分类 □ ① 上市状态 □ ① 上市状态 □ ① 正监会行业分类2012版 □ ① 中信行业分类 □ ① 中证指数行业分类 □ ① 中证指数_证监会行业分类2012版 □ ① 中证指数_证监会行业分类2012版 □ ① 申证指数_证监会行业分类2012版 □	待选区 请输入股3 000001 000002 000003 000004- 000005 000006 000006	5150个 全选 栗名称/代码/拼音 Q 平安银行 へ 万 科 A 「 町金田A 「 国华网安 「 ST 星 源 「 深振业A 「 *ST 全新 「	已进0个 ◆	全部删除	

注:点击待选字段选定,点击已选字段删除

注:点击待选字段选定,点击已选字段删除。

股票代码导入:用户需要先下载导入模板,在模板文件加入需要的股票代码信息,编辑完成后,点击批量导入股票代码,导入的代码会在股票代码框中显示,如图:

第一步: 輸入查询条件		
◊♀ 股票代码选择		
○ 全市场股票 ○ 从股票池选择 ◇ 股票代码导入		
股票代码	批量导入股票代码	下载导入模板

正在打开 fileDownLoad_code.txt		×		
 您选择了打开: fileDownLoad_code.txt 文件类型: TXT 文件 来源: http://192.168.0.14 	40:58081			
您想要 Firefox 如何处理此文件 ? ① 打开 , 通过(②) 浏览(B). ④ 保存文件(S) ⑦ 保存到百度网盘 □ 以后自动采用相同的动作划			批量导入股票代码	下载导入模板
	備定 取消 Ⅲ ▼ 凵 Ø			
名称 fileDownLoad_cod1.txt fileDownLoad_cod2.txt fileDownLoad_code.txt	修改日期 类型 2022/1/25 14:29 文本 2022/2/8 14:26 文本 1920/42/30 15:50 文本			
	01,000002	代码导入 000004,000005,00000	6,00007,000008,00000	此量导入股票代码下载导入模板
< m 0:	e ('.txt,''.text) (O) 取消 20	22		

5.1.2 时间设置

时间范围设置:系统后台内设置时间范围,如果设置年份超出设定范围,在下一步点击查询时则会给出相应的提示,如图:

📋 时间设置	
2017	2019
◊ 股票代码选择	① rtas.resset.com
🔿 全市场股票	请在规定年限内选择时间,当前年限为1年
	确定
🖹 时间设置	
2017	2019

5.1.3 关键字设置

关键字设置可对标题/全文进行筛选,设置关键字:包含任一关键字(-分隔)、包含所有关键字(-分隔)以及 不包含任一关键字(-分隔)。 标题:包含任一关键字(-分隔)

=	关键字设置	
⊘标	题 🔷 全文	
标题	包含任一关键字(-分隔)	选择股票
标题	包含所有关键字(-分隔)	
标题	不包含任一关键字(-分隔)	

点击选取股票,用户可以在弹出框中选择关键字信息,如图:

行业选择 筛选 / <th<< th=""><th></th><th>101</th><th></th><th></th><th></th><th></th><th></th></th<<>		101					
行业选择 筛选 ● 股票市场分类 ● 上市状态 ● 上市状态 ● 上市状态 ● 中信行业分类 ● 中证指数行业分类 ● 中证指数行业分类 ● 中面指数 证监会行业分类2012版 ● 中证指数 证监会行业分类2012版 ● 中证指数 证监会行业分类2012版 ● 中面指数 近 监会行业分类2012版 ● 中面指数 近 监会行业分类2012版 ● 中面指数 近 监会行业分类2012版 ● 中面行业分类2014版 ● 中面方行业分类2014版 ● 中面方行业分类2014版			-检索-				
 ① 股票市场分类 □ 上市状态 □ 上市状态 □ 近监会行业分类2012版 □ 中信行业分类 □ 中证指数行业分类 □ 中证指数 证监会行业分类2012版 □ 印证指数 证监会行业分类2012版 □ 000021-深 科 技 □ 000022-深赤湾 A □ 000023-深天地 A □ 000024-招商地产 □ 000026 体 + 1 	行业选择 筛选		待选4761个	全选		已选4个	全部删除
 ● 上市状态 ● し市状态 ● し市状态 ● 000018-神城A退 ● 000019-深粮控股 ● 000019-深粮控及 ● 000020-深华发A ● 000020-深华发A ● 000021-深科技 ● 000021-深科技 ● 000022-深赤湾A ● 000023-深天地A ● 000024-招商地产 ● 000026 体 + A 	田 股票市场分类		请输入股票名称/代码/拼音	۹		000014-沙河股份	
 ① 证监会行业分类2012版 ● 中信行业分类 ● 中证指数行业分类 ● 中证指数行业分类 ● 中证指数 证监会行业分类2012版 ● 中证指数 证监会行业分类2012版 ● の00023-深天地A ● 000024-招商地产 ● 000026 体 + 1 			000018-神城A退	^		000015-PT中浩A	
 ① 中信行业分类 ① 中证指数行业分类 ① 中证指数证监会行业分类2012版 ① 申万行业分类2014版 ② 00021-深科技 ③ 000022-深赤湾A ③ 000023-深天地A ○ 000024-招商地产 ○ 000026 株 + A 	田 证监会行业分类2012版 ■		000019-深粮控股		-	000016-深康佳A	
 ① 中证指数行业分类 ① 中证指数,证监会行业分类2012版 ① 申万行业分类2014版 ② 000021-深科技 ○ 000022-深赤湾A ○ 000023-深天地A ○ 000024-招商地产 ○ 000025 体 + A 		4	000020-深华发A	-	Φ	000017-深中华A	
① 中证指数_证监会行业分类2012版 000022-深赤湾A ① 申万行业分类2014版 000023-深天地A ① 000024-招商地产 000025 体 + A			000021-深科技		AP.	Λ	
① 申万行业分类2014版 ■ 000023-深天地A 000024-招商地产 000025 体 + A V V 1000025 体 + A V V V V 1000025 体 + A V V V V V V V V V V V V V V V V V V	田 中证指数_证监会行业分类2012版 ■		000022-深赤湾A				
000024-招商地产 000035 ## +> A	□		000023-深天地A				
			000024-招商地产				
			00001E #± + A	~			
	■ 标题 包含任一关键字(-分隔)						

📋 关键字设置

♦ 标题 () 全文						
标题	沙河股份-PT中浩A-深康佳A-深中华A-	查询				
标题	包含所有关键字(-分隔)					
标题	不包含任一关键字(-分隔)					

全文:

📋 关键字设置

〇标	题 🔹 📀 全文	
全文	沙河股份-PT中浩A-深康佳A-深中华A-	查询
全文	包含所有关键字(-分隔)	
全文	不包含任一关键字(-分隔)	

RESSET 锐思	数据				RESSET 文:	本分析客户端−小锐助ヲ	F
5.1.4 查询	结果展示						
综上操作选持	译完成后,点击	^{查询文档} 按钮,	结果即可显示,	用户ī	可以点击 <mark> [•]数</mark>	^{据下载} 按钮,进行查询约	古江
果下载,如图	ह :						
	查询文档						
	已全选				找到 56229条结果 显示 10 🗸		
	标题				披露时间		
	600530_*ST交昂: 2022年年	度报告			2023-08-31		
	001211_双枪科技: 2022年年	F度报告			2023-07-01		
	002656_ST摩登: 2022年年月	度报告			2023-06-30		
	603959_百利科技: 2022年年	F度报告			2023-06-30		
	300937_药易购: 2022年年度	詞指			2023-06-29		
	002564_天沃科技: 2022年年	F度报告			2023-06-29		
	600570_恒生电子: 2022年年	F度报告			2023-06-28		
	601212_白银有色: 2022年年	F度报告			2023-06-27		
	688578_艾力斯: 2022年年度	試报告			2023-06-22		
	688737_中自科技: 2022年年	F度报告			2023-06-22		
		首页 上一页 1 2	3 4 5 6 7 下一	页 尾页			
						(

5.2 政府工作文本数据

政府工作文本数据库目前包括国务院、省级以及地级市政府工作报告三大数据源。政府工作报告是外界 了解政府的主要信息来源,包含着宏观政策、发展规划、资源配置等重要信息。文本的主要内容有对过去一 年工作业绩与问题的回顾性总结,对未来一年政府工作的预期目标,以及具体实施的方案与安排,是一个具 有施政纲领性质的政策性文本,是记录政府履职的重要政策性工具。

±数据下载

文本数据起止时间: 1954-至今

文本数据包含:国务院、省级以及地级市政府工作报告。其是外界了解政府的主要信息来源,包含着宏观政 策、发展规划、资源配置等重要信息。

文本分析:数据库子库。

±数据下载

• Q ~ ¢ Ĥ • → 文本数据起止时间: **1954-至今** 中国上市公司财经文本 折叠 « → 文本数据包含国务院、省级以及地级市政府工作报告。其是外界了解政府的主要信息来源,包含着宏观政策、发展规划、资源配置等重要信息。 → 文本分析: 词频、文本特征、主题分析、特色指标等 → 支持**定制开发**(文本批量导入导出、特定指标开发) → 如使用本库数据, 请添加引用: 数据来源于RESSET文本数据库(http://rtas.resset.com/) 数据库子库 政府工作报告(国务院) 政府工作报告(省级行政区) 政府工作报告(地级行政区)

5.2.1 政府地区选择

RESSET 锐思数据

选择政府工作报告(国务院),如图:

RESSET		▲ 2010 Q 2010 000 000 000 000 000 000 000 000 00	
	□ 国务院	😌 政府工作文本数据库	
	田 省级单位		
	北京市	第一步:输入直询条件	
	□ 天津市	■ 时间设置 1954 ■ 2023 ■	
	□ 和平区		
	□ 河东区	□ 关键字设置	
	□ 河西区		
	□ 南开区		
		\$2422 已高阶份大陆子(-778)	
	□ 红桥区	你道 不包含住一关键字(-分隔)	
	□ 东丽区	查询文档	
	□ 西青区		
	□ 津南区		0
	□ 北辰区		
	□ 武清区	100102001 加平氏2023年時月工作時合 2023-12-51 100102001 約1072020年時月工作時合 2023-12-51	
	□ 宝坻区	100102001 101020001 10102001 10102001	
	□ 滨海新区	100102001 和平区2021平段府工作探告 2021-12-51	
		□ 100102001_和平区2020年版府工作版合 2020-12-31	
		□ 100102001_和平区2019年政府工作股告 2019-12-31	
		□ 100102001_机学区2018年政府工作股告 2018-12-31	
		1047 43 34 3047 43 34	

5.2.2 时间设置

时间设置:系统内设置时间范围为1年,如果设置年份超出1年,下一步点击查询时则会给出相应的提示,如图:

田 辽宁省	I	
		⊕ rtas.resset.com
📋 时间设置		请在规定年限内选择,当前年限为2020~2021
2018	2021	确定

📋 时间设置

2020	2021	
------	------	--

5.2.3 关键字设置

关键字设置可分:

标题:包含任一关键字(-分隔)、包含所有关键字(-分隔)、不包含任一关键字(-分隔); **全文:**包含任一关键字(-分隔)、包含所有关键字(-分隔)、不包含任一关键字(-分隔)。

三 关键字设置					
♦ 标题 🔿 全文					
标题 包含任关键字(-分隔)					
标题 包含所有关键字(-分隔)					
标题 不包含任一关键字(-分隔)					

5.2.4 查询结果

点击查询按钮,用户可以对查询结果进行下载。如图:

查	间文档	
Ê	黄育已选 已选 10 条 清空	找到 18条结果 显示 10 -
	标题	操作
	100102001_和平区2023年政府工作报告	2023-12-31
	100102001_和平区2022年政府工作报告	2022-12-31
	100102001_和平区2021年政府工作报告	2021-12-31
\checkmark	100102001_和平区2020年政府工作报告	2020-12-31
\checkmark	100102001_和平区2019年政府工作报告	2019-12-31
	100102001_和平区2018年政府工作报告	2018-12-31
	100102001_和平区2017年政府工作报告	2017-12-31
	100102001_和平区2016年政府工作报告	2016-12-31
	100102001_和平区2015年政府工作报告	2015-12-31
\checkmark	100102001_和平区2014年政府工作报告	2014-12-31
	首页 上一页 1 2 下一页 尾页	



5.3 美国上市公司财经文本

文本数据起止时间: 1954-至今;

文本数据包含**年报、季报、招股说明书;**

文本分析:数据库子库。

市公司财经文本	→ 文本数据起止时间:	1954-至今	ra ola ola site alcolule						折
工作文本数据	 → 又本数据包含年报、 → 文本分析: 词频、文 → 支持会制平台(文本) 	李掖、招股说明书林 本特征、主题分析、 ^{业 — — — — — — — —}	11目定义又档 特色指标等						
市公司财经文本	→ 如使用本库数据, 请	添加引用:数据来源	唐于RESSET文2	本数据库(http	o://rtas.res	set.com/)			
经资讯文本	数据库子库	年度财务报告	季)	度财务报告		招股说明书			
	工作文本数据 书公司财经文本 圣资讯文本	□ 文本分析: 词姨、文 □ 支持定制开发(文本 計公司财经文本 登窃讯文本 数据库子库	□ 文本分析: 词频,文本特征、主题分析、 □ 文本分析: 词频,文本特征、主题分析、 □ 支持定制开发(文本批量导入导出、特定 □ 如使用本库数据,请添加引用:数据来】 ○ 数据库子库	□ 文本分析: 词频、文本特征、主题分析、特色指标等 □ 支持定制开发(文本批量导入导出、特定指标开发) □ 女持定制开发(文本批量导入导出、特定指标开发) □ 如使用本库数据, 请添加引用:数据来源于RESSET文 圣资讯文本 数据库子库 年度财务报告	□ 文本分析: 词频、文本特征、主题分析、特色指标等 □ 支持定制开发(文本批量导入导出、特定指标开发) □ 立持定制开发(文本批量导入号出、特定指标开发) □ 如使用本库数据, 请添加引用:数据来源于RESSET文本数据库(http://www.setable/set	□ 文本分析: 词频,文本特征、主题分析,特色指标等 □ 支持定制开发(文本批量导入导出,特定指标开发) □ 立時定制开发(文本批量导入导出,特定指标开发) □ 如使用本库数据,请添加引用:数据来源于RESSET文本数据库(http://rtas.ress 登资用文本 数据库子库 年度财务报告	□ 文本分析: 词频,文本特征、主题分析、特色指标等 □ 支持定制开发(文本批量导入导出、特定指标开发) □ 立時定制开发(文本批量导入导出、特定指标开发) □ 如使用本库数据,请添加引用:数据来源于RESSET文本数据库(http://rtas.resset.com/) 经资讯文本 数据库子库 年度财务报告 择度财务报告 招股说明书	□ 文本分析: 词姨,文本特征、主题分析、特色指标等 □ 支持定制开发(文本批量导入导出、特定指标开发) □ 立肯定制开发(文本批量导入导出、特定指标开发) □ 如使用本库数据, 请添加引用: 数据来源于RESSET文本数据库(http://rtas.resset.com/) 经消讯文本 数据库子库 年度划务报告 年度划务报告	□ 文 立 今 分 折: 词频、文本特征、主题分析、特色指标等 □ 支 方 定 制开发 (文本 批量导入导出、特定指标开发) □ 立 方 定 制开发 (文本 批量导入导出、特定指标开发) □ 如使用本库数据, 请添加引用: 数据来源于RESSET文本数据库(http://rtas.resset.com/) 圣资讯文本 数据库子库 年度财务报告 于 成 的 新 日

5.3.1 股票代码选择

代码选择分为两种方式:从股票池选择、股票代码导入。

♦ 股票代码选择

从股票池选择:用户可以在左侧代码待选区,点击单个/全选,代码进入已选区,如图:

从股票池选择 〇 股票	戰代码导入		
待选区11423个 全	选	已选5个	全部删除
请输入股票名称/代码/拼音	Q	ABB	
ABCM	^	ABCB	
ABEO		ABC	
ABEQ		ABBV	
ABEV	4	ABCL	
ABG			
ABM			
ABMD			
	\sim		

注:点击待选字段选定,点击已选字段删除。

10 股重代码选择

股票代码导入:用户可以通过先下载导入模板,在模板文件加入需要的代码信息,编辑完成后,点击批量导 入股票代码,导入的代码会在股票代码框中显示,如图:

1			
○ 从股票池选择	🕏 股票代码导入		
股票代码		批量导入股票代码	下载导入模板



5.3.2 时间设置

时间范围设置:系统内设置时间范围为1年,如果设置年份超出1年,在下一步点击查询时则会给出相应的提示,如图:

E	时间设置		
	2017	2019	
Ē	关键字设置		
۲	●标题 ○ 全文	⊕ rtas.resset.com	
标	题 包含任一关键字(请在规定年限内选择时间,当前年限为1年	
标	题 句会所有关键字(── 不允许 rtas.resset.com 再次向您提示	
		确定	
标	题 不包含任一关键字		

5.3.3 关键字设置

关键字设置可对标题/全文进行筛选,设置关键字:包含所有关键字、包含任一关键字以及不包含任一关键 字。

标题:

±数据

▶ 关键字设置
 ▶ 休题 ○ 全文
 ▶ 标题 包含所有关键字,(空格分隔)
 ▶ 标题 包含任一关键字,(空格分隔)
 ▶ 标题 不包含任一关键字,(空格分隔)

全文:

🖹 关键字设置

○标	题 🛛 😒 全文
全文	包含所有关键字,(空格分隔)
-	
全文	包含任一关键字,(空格分隔)
全文	不包含任一关键字,(空格分隔)

5.3.4 查询结果展示

	标题	披露时间
	ABR_ABR_20211231Annual report	2021-12-31
\checkmark	ABBV_ABBV_20211231Annual report	2021-12-31
	AB_AB_20211231Annual report	2021-12-31
	ABM_ABM_20211031Annual report	2021-10-31
	ABC_ABC_20210930Annual report	2021-09-30
	ABMD_ABMD_20210331Annual report	2021-03-31
	ABCB_ABCB_20201231Annual report	2020-12-31
	ABCL_ABCL_20201231Annual report	2020-12-31
	ABG_ABG_20201231Annual report	2020-12-31
	ABBV_ABBV_20201231Annual report	2020-12-31

5.4 财经资讯文本

文本数据起止时间: 2000-至今;

财经资讯包含:证券日报、证券时报、智通财经、中财网、中国经济网、中国经济报、中国证券报、上海证券报、财经网、新华社、凤凰新闻、今日头条、腾讯财经等 30 多个具有影响力的财经资讯;

文本分析:**数据库子库。**

中国上市公司财经文本	 → 文本数据起止时间: 2000-至今 → 文本数据包含证券日报、证券时报、智通财经、中财网、中国经济网、中国经济报、中国证券报、上海证券报、财经网、新华社、凤凰新闻、今日头条、『 	折叠 新聞好
政府工作文本数据	30多个具有影响力的财经资讯。 → 文本分析: 词频、文本特征、主题分析、特色指标等	
美国上市公司财经文本	→ 支持定制开发【文本批量导入导出(十大官媒、道琼斯全球资讯、股吧评论等)、特定指标开发】 → 如停用本库教报、请添加引用:教报来源于RESSET文本教报库(http://ttas.resset.com/)	
财经资讯文本		
	数据库子库 财经资讯	

5.4.1 输入查询条件

时间范围设置:系统后台内设置时间范围,如果设置年份超出设定范围,在下一步点击查询时则会给出相应的提示,如图:

📋 时间设置	
2017	2019
📋 关键字设置	
♦ 标题 ○ 全文	() rtas.resset.com
标题 包含任一关键字(请在规定年限内选择时间,当前年限为1年
标题 包含所有关键字(不允许 rtas.resset.com 再次向您提示
标题 不包含任一关键号	确定

5.4.2 关键字设置

关键字设置可对标题/全文进行筛选,设置关键字:包含所有关键字、包含任一关键字以及不包含任一关键 字。

标题:

= ‡	关键字设置
📀 标题	题 〇 全文
标题	包含所有关键字,(空格分隔)
标题	包含任一关键字,(空格分隔)
标题	不包含任一关键字,(空格分隔)

RESSET 復	说思数据	RESSET 文本分析客户端-小锐助手
	关键字设置	
〇标	题 🛛 📀 全文	
全文	包含所有关键字,(空格分隔)	
全文	包含任—关键字,(空格分隔)	
全文	不包含任一关键字,(空格分隔)	

5.4.3 查询结果展示

综上操作选择完成后,点击 查询文档 按钮,查询结果即可显示,用户可以对查询结果进行下载,如图:

查询文档	
全选 查看已选 已选 10 条 清空	找到 49115条结果 显示 10 ~
☑ 标题	披露时间
☑ *ST庞大被证监会立案调查 遗留问题远不止信息披露违规	2023-05-28
☑ 奥贝胆酸NASH适应症上市申请被推迟相关赛道药物"上市第一名"之争横生变数	2023-05-28
☑ 上市6周登顶畅销榜TOP1 海信电视U7K为"爆款王"打样	2023-05-28
☑ 信息披露不准确沈阳化工收警示函 2022年净利润现较大额度亏损	2023-05-28
☑ 森源电气回复深交所问询函新实控人能够对公司实施有效、稳定的控制	2023-05-28
☑ 具备开发800Gbps及以上铌酸锂调制器芯片能力光库科技助推芯片国产化	2023-05-28
☑ <股东来了>(2023)甘肃片区燃高校 走进校园百余场	2023-05-28
☑ 2023数博会"东数西算"高端对话在贵阳成功召开	2023-05-28
☑ 2023数博会"数据二十条"安全产业发展论坛成功举办	2023-05-28
☑ <股东来了>(2023)"投教万里星聚天山"系列投教活动走进伊力特	2023-05-28





6 词频搜索



6.1 输入关键词检索

用户可以通过在输入关键词框内输入单个、多个关键词进行检索,多个关键词用逗号分开,本助手支持中 文、英文文档解析。

使用方法: 1、输入关键词; 2、选择自定义文件夹; 3、点击检索; 4、显示分析结果; 5、下载分析结果文件。

	RESS 秒传海	SET 词频搭 ^{受支本 本地分析接}	搜索	
财务,经济	关键词模板下载	自定义文件夹	▲ 中文	
	请选择目定义文件夹			

RESSET 炒砖海量支本	词频 *地分析	叟索					
财务,经济							
◆ 关键词模板下载	定义文件夹 全测试器下载\1111	× e	中文		1	到 出TXT S	导出表格
文件	时间	大小	字数	总词数	总句数	关键词	词频
	2024-05-14	67.07KB	27322	14500	905	财务	31
ST全新・关于对深交所(2021) 第225号年报问询函问复的公告6f036587fb9e4bc9847c623f60883630 p	2024-05-14	67.07KB	27322	14500	905	经济	8
ST全新: 独立董事关于对深交所公司部年报问询函 [2022] 第 162 号回复之独立意见08c6a88b6dc842e0	2024-05-14	5.10KB	1971	1096	70	财务	
ST全新:独立董事关于对深交所公司部年损问询阅 (2022) 第 162 号回复之独立意见08c6a88b6dc842e0 ST全新:独立董事关于对深交所公司部年损问询阅 (2022) 第 162 号回复之独立意见08c6a88b6dc842e0	2024-05-14 2024-05-14	5.10KB 5.10KB	1971 1971	1096 1096	70 70	财务 经济	2 2

6.2 导入模板检索

点击关键词模板下载,设置关键词,关键词整理完成后,双击输入框上传关键词文件,选择自定义文件夹, 点击检索,查看分析结果并支持下载。

刘览文件夹		×			
> 🏪 本地磁曲 (C:)		RES	SET 词频	叟索	
 ▲ 新加巻 (D:) > 360Downloads > 360安全浏览請下載 BaiduNetdiskDownload 		桫鴣	每量支本 本地分析扌	究姫	
> develop > G		建词用逗号分开,双击导	入关键词		
HadoopUpLoad PHP Program Files Program Files (x86) res		→ 关键词模板下载	自定义文件夹	▲ 中文	
新建文件夹(M)	确定 取消				

	🗐 关键	词模版.txt	- 记事本				
	文件(F)	编辑(E)	格式(O)	查看(V)	帮助(H)		
	财经数行基数 行基数						
R 打开				:	×		
$\leftrightarrow \rightarrow \cdot \cdot$	▶ 🔜 ≪ 360安 > :	关键词文件 🗸 🗸	ひ 搜索"关键》	1文件* ク			



	RES	SET 词频	要索	
	妙传语	・量叉本 本地分析オ	死始	
财务,经济				
	📚 关键词模板下载	自定义文件夹	≤■ 中文	

桫传海量支本	本地分析	究掘					
财务,经济							
◆ 关键词模板下载 — □ 自	□□ 自定义文件夹 D\\$60安全浏览器下载\1111		▲ 中文		🚺 导出TXT 🔄 导出表格		
文件	时间	大小	字数	总词数	总句数	关键词	词频
ST全新:关于对深交所(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.p	2024-05-14	67.07KB	27322	14500	905	财务	31
ST全新:关于对深交所(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.p	2024-05-14	67.07KB	27322	14500	905	经济	8
ST全新:独立董事关于对深交所公司部年报问询函(2022)第 162 号回复之独立意见08c6a88b6dc842e0…	2024-05-14	5.10KB	1971	1096	70	财务	2
	2024-05-14	5.10KB	1971	1096	70	经济	2
							_

7 文本特征



7.1 文本信息

▲ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○	Q 词频搜索		€ ₩=		- ×
特征分析	自动化约	充计文本中的]文本指标		
文本分析方法 🔹 支持任	E意文本文件	• 自定义词汇和字典 •	一键导出分析结果		
情感分析		可读性分析		自定义文本特征	
提供常用情感词库		提供常用可读性词库		指标构建	
支持自定义情感词库		支持自定义的词库		词典定制	
提供百分比的语调分析		支持多种格式自定义	文本	特征计算	
支持多种格式自定义文本		一键导出分析结果		分析结果导出	
			特色指标		
创新 风险因子	竞争性剧	财子 前增性因子	区块链概念因子		
人口老龄化指标	江待遇因子	贸易战潜在风险	会计因子		ſ
					_
					説思
					寄服
	 ・ 文本数据 ・ 文本数 ・ 文本数 ・ 文本数 ・ 文 ・ 文 ・ 文	 ・ ・ ・		★ ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●	Image: State State Image: State State State Image: State St

文本信息功能: 1、解析文件大小; 2、分析文件页数; 3、准确统计文本句子数; 4、提供文本字数和词汇数。

操作:上传 PDF、DOC、DOCX、TXT 文件,系统自动生成文件基本信息。用户可以对解析结果进行导出。

浏览文件夹	
 ✓ 新加巻 (D:) ▲ 360Downloads ▲ 360支全浏览器下載 ▲ doc ▲ pf > ■ RTAS 	点击打开文件夹 可直接上传PDF、DUC、DOCX、TXT格式的文件夹
 > 财经文本 - 客户端 关键词搜索结果_20231211095828 > 立信 - 财经文本 模板文件 	
新建文件夹(M) 确定 取消	

RESSET 4	第助手	マンクロン (本特征) 文本授短	€ tem					<u>1</u> - ×
 文本特征 文本信息 	文本信息 统计文件的一些基本信息							功能简介 1、上传PDF、DOC、DOCX
 ♥ 情感分析 □ 可读性分析 ■ 自定义文本特征 	点击打开文件 可直接上体PDF、DOC、DOC、 已选择数据文件:DA360安全测示器下数\1111	夹 TXT格式的文件夹						、IXI文件,轰腾目动生成交 件基本信息 2、统计TXT文件时页数为1,
	E94	ff: 7/7						
	文件名	上传时间	文件大小	总字数	总词数	是句数	总页数	
	ST全新:关于对深交所(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-04-12	67.87KB	26835	15206	1871	30	
	ST全新: 独立董事关于对深交所公司部年报问询函 (2022) 第 162 号回复之独立意见08c6a88b6dc842e0be1d0c8261743a8c	2024-04-12	5.13KB	1923	1101	122	3	
	ST全新:碱公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf.txt	2024-04-12	36.34KB	13936	8286	969	18	
	ST国华:致同会计师事务所(特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询通的回复d0c057d68d4d447	2024-04-12	94.37KB	35953	20772	2498	42	
	昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.bxt	2024-04-12	42.84KB	15306	8589	955	1	
	 	下—页 尾页						R R R

7.2 情感分析

RESSET小锐助手	▲ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○	文本特征 文本挖掘 设置		- ×
			_	
	文本特征分析 自动化约	统计文本中的文本指标	7	
	 常用文本分析方法 支持任意文本文件 	 自定义词汇和字典 一键导出分析结果 		
全本信息	信感分析	可读性分析	自定义文本特征	
解析文件大小	提供常用情感词库	提供常用可读性词库	指标构建	
分析文件贝数	支持目定义情感词库	支持目定义的词库	词典定制 	
/在1995年7年9月38 提供文本字数和词汇数	支持多种格式自定义文本	一键导出分析结果	分析结果导出	
	描述性创新 网络因子 音分性	因子 前始性因子 反体路缆金因子		
	人口老龄化指标员工待遇因子	贸易战潜在风险 管理会计因子		ĺ
				思 客
				ER .

功能简介:根据情感词典计算文本的情感极性,检索文本中所包含的积极情感词和消极情感词的个数。

最终文本的情感得分为 TONE=(积极情感词个数-消极情感词个数)/(总情感词个数)。

情感得分取值范围为[-1,1],情感得分越接近1,文本的情感极性越偏向于积极;反之,情感的分越接近-1, 文本的情感极性越偏向于消极。

功能操作:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,上传完成后,点击 查询按钮进行文件解析,用户可对解析结果进行预览、导出。

 文本 文本信 	特征 息	情感分析	分析得出文本数据的正负面语调		
❤ 情感分	祈				
■ 可论 测 测 ●● 自知	次 遼文件夹)) 音乐 → 章面		×	こ。 点击打开文件夹 可直接上传PDF、COC、DOCX、TXT格式的文件夹	
	 美田 美田 美田 新加速(1) 新加速(1) 360bcc 	ii) j河览器下载 j j j j j j j j 文件 得 大件 大 政府报告		宣询	
	新建文件夹(M)		确定 取消		



 ☑ 文本特征 ☐ 文本信息 	情感分析。分析得出文本数据的正负面语调								功能简介 根据情感词典计算文本的 极性,检索文本中所包含		
❤ 情感分析											极情感词和消极情感词的 教。
🛄 可读性分析										最终文本的情感得分为 TONF=(积极情感词个参	
👫 自定义文本特征	<u>高击打开文件実</u> 可 <u>前線上作PDF、DOC、DOCX、TXT修式的文件</u> 実 已造得数篇文件DA360完全波动器下载\1111										情感词个数/(总情感词个 情感得分取值范围为[-1, 情感得分越接近1,文本 感极性越偏向于积极; 新
Вэ й : 7/7											情感的分越接近-1,文本 感极性越偏向于消极。
		查道	Ð							2 日州事故	
	文件名	上传时间	文件大小	总字数	总词数	总句数	总页数	积极词频	消极词频	Tone	
	ST全新:关于对深交所(2021)第225号年版问询函回复的公告6f036587fb9e4bc9847c623f60883	2024-04-12	67.87KB	26835	15206	1871	30	4136	1356	50.619%	
	ST全新:独立董事关于对深交所公司部年报问询函(2022)第 162 号回复之独立意见08c6a88b6dc	2024-04-12	5.13KB	1923	1101	122	3	366	101	56.745%	
	ST全新: 碱公律师事务所对全新好2021年报问询查的法律意见书62c8a29e02e84169aae679089e5	2024-04-12	36.34KB	13936	8286	969	18	2387	703	54.498%	
	ST国华: 致同会计师事务所 (特殊普通合伙) 关于深圳国华网安科技股份有限公司2021年年报问询函	2024-04-12	94.37KB	35953	20772	2498	42	5736	2083	46.720%	
	昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-04-12	42.84KB	15306	8589	955	1	4333	818	68.239%	
	首页	〔 <u>上</u> —页 1	2 下一页	尾页							

7.3 可读性分析

RESSET小说助手	☆ この	☆本将征 文本挖掘 设置	-	- ×
	文本特征分析 自动化	统计文本中的文本指标		
	 常用文本分析方法 支持任意文本文件 		•	
文本信息	情感分析	可读性分析	自定义文本特征	
解析文件大小	提供常用情感词库	提供常用可读性词库	指标均量	
分析文件页数	支持自定义情感词库	支持自定义的词库	词典定制	
准确统计文本句子数 提供文本字数和词汇数	提供百分比的语调分析 支持多种格式自定义文本	支持多种格式自定义文本 一键导出分析结果	特征计算 分析结果导出	
_			的指标	
	描述性创新 风险囚子 竞争性	因子 前喻性因子 区块链概念因子		
	人口老龄化指标 员工待遇因子	贸易战潜在风险管理会计因子		ſ
				锐思客
				服

功能简介:基于文本中所包含的生僻字数量、专业术语数量和语法结构(逆转词汇,如"但是"、"然而"等)进行文本的可读性分析。基于此,构建可读性词库统计文本中生僻词数量、专业术语数量和逆转词汇数量,计算文本复杂度: COMPLEXITY=(生僻字数量+专业术语数量+逆转词数量)/文本总词汇数量。

功能操作:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,上传完成后,点击 查询按钮进行文件解析,用户可对解析结果进行预览、导出。

可读性分析 提供可读性通用	指标	
浏览文件夹	× 「 可直接上传PUF、DO	<mark>三</mark> , 打开文件夹 DC、DOCX、TXT樁式的文件夹
 ✓ 360安全浏览器下载 1111 20230322 pdf, txt, word txt txt1 关键词文件 		
> 解析后的文件 空文件夹 天津政府报告 新建文件夹(M)		直询

RESSET 文本分析客户端─小锐助手

RESSET 锐思数据

 交本特征 う本信息 	可读性分析 提供可读性通用指标											功能简介
 情感分析 可读性分析 請定义文本特征 	<mark>度。</mark> 点击打开文件夹 可 <u>自後上</u> 修PDF、DOC、DOCX、TXT储式的文件夹 已选择数据文件:DA360安全测数器下数\\111									量、专业木器数量和耐活的 (近转调定、如"倡量"、" 然而"等,进行文本的可读性 分析。 基于此、传递可读性调味统计 文本中生音调数量、专业术制 数量和近转调订数量、 计算文末差型表示。 COMPLEXITY=(生得学数量+ 专业术器数量+世标间汇数量) (文本态调定数量)。		
		()										
🖪 文本特征	可法进入扩 相供可法供该回报性											
	可读性分析 提供可读性通用指标											功能简介 基于文本中所包含的生態字数
 情感分析 可读性分析 自定义文本特征 	已选择数据文件:D\360安全测览器下载\1111	可直接上传PD	<mark>言</mark> 点击打开: F、DOC、DC	〕 文件夹 ɔcx、тxтŧ	<u> </u> 者式的文件	史						至, 夺业术语数量和语法结构 (送转间汇, 如" 伺是"、 然而" 夸) 进行文本的可读性 分析。 基于此, 构建可读性词座统计 文本中生都问数量, 夺业术语 数量和记述将前款 数量, 计算文本复杂度; COMPLEXITY~(生像子数量、 专业术语数量-送转前门数量
												/文本总词汇数量。
			查询]								
										T 导出TXT	5 导出表格	
		上传时间	文件大小 (7.07//2	总字数 20025	总词数	总句数	总页数	逆接成分词频	生僻词词颊	专业术语词频	复杂性	
	51至新: 夫丁刈床交析 (2021) 第225号年报回询图回复的公告61036587代994bc9847 51 全新: 法丁ガ床交析 (2021) 第255号年报回询图 (2022) 第 162 号回复之社改章回の8	2024-04-12	67.87KB	26835	15206	18/1	30	597	94	5860	43.082%	
	51主新:现立重单天于对决交所公司起中报问询图 [2022] 第 162 亏回复之独立意见08 51全新:现立重单天于对决交所公司起中报问询函约注注音印书62 r8+29+02+284169+3	2024-04-12	36 34KB	13936	8286	969	- 3	308	81	2471	44.959% 34.516%	
	31主新,藏公中和中药所以至新贝221年很问题因的法律急处中5628829e02e84169aa 	2024-04-12	94 37KB	35953	20772	2498	42	991	4	8135	43 953%	
	日本で2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-04-12	42.84KB	15306	8589	955	1	177	2	2357	29.526%	
		道页上	一页 1	2 下—页	瓦 尾页							

7.4 自定义文本特征

自定义文本特征提供"词典定制"和"特征计算"两大模块。其中,"特征词典定制"模块用于实际自定义 指标构建,通过"种子词集输入"完成用户自定义文本指标的词典构建;"特征计算"模块为用户提供文本 数据针对已有自定义文本指标的计算和下载功能。

功能步骤:

① 选中点击打开文件夹,上传需要解析的文档;



 ☑ 文本特征 ○ 文本信息 ♥ 信感分析 ☑ 可读性分析 ■ 自定义文本特征 	自定义文本特征 提供自定义文本指标的计算 点击 已选择数额文件:0入360绘全测定器下数\1111	<mark>テ,</mark> 打开文件夹 C、DOCX、TXTI 版式的文件来	功能简介 田定义文本特征提供"间晚定 新 和"特征计算"两大楼 快,其中"特征词典定朝 提供开主张周启之训练内 整:通过"称于理秘分、完 成将于自定之文本并描标的词确 户道供文本故题针对已有自定 义文本描标的计算和下载功 能,
	选择指标 操作 指标名称 操作 新增新版画法 下載調整 经济 下載調整 金融 下載調整	关键词列表 操作 大数据 新聞新版測试 開始 信息 新聞前版測试 開始 集白 新聞前版測试 開始	
	添加指标 请输入关键词,多个关键词用逗号分开 移动的	添加关键词 <u>请输入关键词,多个关键词用语号分开</u>	ĺ

② 创建指标:用户可以通过在输入框内输入指标名称;

③ 点击添加指标按钮进行添加;

 文本特征 文本信息 	自定义文本特征 提供自定义文本指标的计算
❤ 情感分析	
些 可读性分析	E,
計 自定义文本特征	<u>点击打开文件夹</u> 可直接上传PDF、DOC、DOCX、TXT格式的文件夹 已选择数据文件:D1/360安全测觉器下载1111
	计经指标 学辑初列事
	新増損防剤は 3 下戦時期 大数据 新増損防剤は 開始
	经济 下致調集 删除 信息 新增指标则试 <mark>删除</mark>
	金融 下戰國集 删除 集合 新增指示规试 删除
	添加指标 请输入关键词,多个关键词用逗号分开 添加新 添加关键词 清输入关键词,多个关键词用逗号分开 使温导入关键词
	已度新闻集!

④ 选择新添加指标,用户可以通过在右侧输入框内输入单个或多个(用英文逗号隔开)关键词,点击●,关键词则显示在关键词列表区域内。

自定义文本特征 提供自定义 3	文本指标的计算			
	点击打	<mark>王</mark> , I开文件夹		
	可直接上传PDF、DOC、	DOCX、TXT格式的文件夹		
选择指标		关键词列表		
指标名称	操作	关键词	来源	操作
新增指标测试 下载词集				
经济 下载词集	删除			
金融				
添加指标 请输入关键词,多个关键词用逗号分开	添加指标	添加关键词 大数据, 信!	急,集合	+ 批量导入关键词
			已添加指标!	
		查询		

自定义文本特征 提供自定义文本指标的计算 E, 点击打开文件夹 可直接上传PDF、DOC、DOCX、TXT格式的文件夹 选择指标 关键词列表 指标名称 关键词 操作 来源 操作 新增指标测试 大数据 新增指标测试 下载词集 下载 信息 新增指标测试 金融 下载词集 集合 新增指标测试 添加指标 请输入关键词,多个关键词用逗号分开 添加关键词 + 批量导入关键词

以上操作完成后,点击查询按钮进行文档解析,支持结果导出下载功能。

 文本特征 文本信息 (前成分析) 可读性分析 (前定义文本特征) 	自定义文本特征	提供自定义文本指标的	片算 点击 _{可国体上作PDF、DOC}	<mark>三,</mark> 17开文件夹 2. DOCK、TXT 能 式的文件4	ι			功能简介 目記文文本移起其"调味走 制"和"特征订算"两大堪 块。其中,"特征订调性型"两大堪 块。其中,"特征订调性型" 成界中行运输论"资 成界中行运输论"资 成界中自定文文本指标论词则, 内望、本证"算"也称之词 中述文本指标论计算和下载功 能。
	选择指标 指标名称 新聞編編時期試 母ぶ 金融	操作 下致调集		关键词列表 关键词 左数端 信息 集合	来源 新增抽标则试 新增抽标则试	提作 删除 删除		
	添加指标 遺輸入关键词 多	5个关键词用逗号分开	ISactors	添加关键词 遭输入	关 键词,多个关键词用退号分开 文档转换中。	•	批重导入关键词	

文档解析中:

 三 文本特征 章 文本信息 	自定义文本特征 提供自定义文本指标的计算									
♥ 情感分析										
😐 可读性分析										
計 自定义文本特征	<u>点击打开文件夹</u> 可直接上体PDF、DOC、DOCX、TXT描式的文件夹 已选择数据文件:D\\360突全测觉器下载\1111									
	选择指标									
	经济 下载词集 船轮 Reconting the OCR project									
	全融 下载调集 副除 34 of 42 pages recognized 指标测试 删除									
	Cancel 3									
	添加指标 请输入关键词,多个关键词用逗号分开									

结果展示:

🛃 文本特征												功能简介
目 文本信息	选择指标			关键	司列表							475月8日月7日 自会义文本特征提供"词典会
	指标名称		操作	关键	词	来	源		操作			制"和"特征计算"两大模
ॐ 情感分析	新增指标测试	下载词集	删除	大数	居	新	增指标测试		删除			块。其中, "特征词典定制" 横块用于实现自完义指标构
👛 可读性分析	经济	下载词集	删除	信息		新	增指标测试		删除			建,通过"种子词集输入"完
■ 自定义文本特征	金融	下抵词类		集 合		आ :	增指标测试		翻除			1007 日本、文中市的500年 特征主义,中位江首、银大力用 户提供文本发展针为已寿自进 义文本者能称约计算和下载功 能。
	添加指标 请输入关键词,多	*个关键词用逗号分开	F Koortava	添加	关键词 <mark>请输入:</mark> ff:7/7	关键词,多个 文件	▶ 关键词用逗 分析完成	号分开	-	•	批量导入关键词	
г				查询						(P Put		
		文件名		上传时间	文件大小	总字数	总句数	总页数	总词数	词接合计		
	ST全新:关于对深交所 (2021) 第	225号年报问询函回复的纪	公告6f036587fb9e4bc9847c623f60883630.p	2024-04-16	67.87KB	26835	1871	30	15206	27357	179.909%	
	ST全新:独立董事关于对深交所公司	同部年报问询函(2022)	第 162 号回复之独立意见08c6a88b6dc842e0	2024-04-16	5.13KB	1923	122	3	1101	1960	178.020%	
	ST全新: 減公律师事务所对全新好2	021年报问询函的法律意	见书62c8a29e02e84169aae679089e5732d7	2024-04-16	36.34KB	13936	969	18	8286	14375	173.485%	
	ST国华: 致同会计师事务所 (特殊部	普通合伙) 关于深圳国华网	列安科技股份有限公司2021年年报问词圈的回复	2024-04-16	94.37KB	35953	2498		20772	36818	177.248%	
	昌平区2022年政	府工作报告c1e521a1526	64ff5bb462f83b3de4d17.txt	2024-04-16	42.84KB	15306	955	1	8589	15310	178.251%	
			首页。	上一页 1 2	下一页 尾页							6 18 18

7.5 特色指标

RESSET小説助手	▲ ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●	文本特征 文本挖掘 设置		- ×
	文本特征分析 自动化	统计文本中的文本指	标	
	■ 常用文本分析方法 🛛 ■ 支持任意文本文件	 自定义词汇和字典 一键导出分析结果 	R	
文本信息	情感分析	可读性分析	自定义文本特征	
解析文件大小	提供常用情感词库	提供常用可读性词库	指标构建	
分析文件页数	支持自定义情感词库	支持自定义的词库	词典定制	
准确统计文本句子数	提供百分比的语调分析	支持多种格式自定义文本	特征计算	
提供文本字数和词汇数	支持多种格式自定义文本	一键导出分析结果	分析结果导出	
			特色指标	
	加速吐动和 风险出于 元十			
	人口老龄化指标员工传递因子	贸易战潜在风险管理会计因子		ſ
				10
				思客
				B.

7.5.1 描述性创新

创新指标又称为描述性创新,是衡量企业创新水平的文本指标。描述性创新指标刻画了与技术创新有关的投入和产出描述性创新信息。创新作为企业的核心竞争力,是衡量企业价值和可持续发展能力的重要,指标创新性指标根据"创新"相关词汇总词频在公告中文本总词频的占比计算得出。

RESSET 4		有新版本
文本特征 五本特征 五本特征 五本特征		功能简介 创新指标又称为描述性创新,
 ・ 風险因子 ・ ・ ・	 	是國重正並创新水干的文举指 标。 描述性创新指标刻画了与技术 创新有关的投入和产出描述性 创新信息。
♥ 前瞻性因子● 区块链概念因子	点击打开文件夹 可直接上体PDF、DOC、DOCX、TXT格式的文件夹	创新作为企业的核心竞争力, 是衡量企业价值和可持续发展 能力的重要指标 创新性指标根 据"创新"相关词汇总词颎在
🖍 人口老龄化指标 🤰 员工待遇因子		公告中文本总词频的占比计算 得出。
♥ 贸易战潜在风险중 管理会计因子		

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。

	描述性创新	统计分析得出创新因子指标	
『 閲文(Ч ж	×	<mark>≡</mark> ,
	 ✓ 360安全浏览器下载 ↓ 1111 ↓ 20230322 ↓ 20240419 	Î	点击打开文件夹 可直接上传PDF、DOC、DOCX、TXT格式的文件夹
	pdf, txt, word txt > txt1 关键词文件		
	 解析后的文件 空文件夹 天津政府报告 问询密0409 英文文件 		
新	建文件夹(M)	施定 取消 .:	

描述性创新统计分析得出创新因子指标											
「 「 「 「 二 志打开文件夹 可 直接上传PDF、DOC、DOCX、TXT格式的文件夹 已选择数据文件:D\360安全浏览器下载\1111											
	已分析:										
l	直询						T 导出TXT	5 导出表格			
文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	词颊合计	指标值			
ST全新:关于对深交所(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-05-14	67.87KB	26835	1871	30	15206	2	0.013%			
ST全新:独立董事关于对深交所公司部年报问询函(2022)第 162 号回复之独立意见08c6a88b6dc842e0be1	2024-05-14	5.13KB	1923	122	3	1101	0	0.000%			
ST全新: 诚公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf	2024-05-14	36.34KB	13936	969	18	8286	0	0.000%			
ST国华:致同会计师事务所(特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询函的回复d0c	2024-05-14	94.37KB	35953	2498	42	20772	140	0.674%			
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	32	0.373%			
· · · · · · · · · · · · · · · · · · ·	_页 1 2 下	页尾页									

7.5.2 风险因子

风险信息具有异质性与不确定性,证监会要求公司对风险进行"有意义"的披露,以帮助信息使用者更加准确的预测公司未来现金流,从而做出正确的投资决策。风险指标根据"风险"相关词汇总词频在公告中文本总词频的占比计算得出。

RESSET 文本分析客户端-小锐助手

뢷 文本特征	风险因子。分析得出文本数据的风险因子指标	功能简介
▲ 描述性创新		风险信息具有异质性与不确定
🚊 风险因子		(住, 业监受要求公司对风险进行"有意义"的披露,以帮助 信息使用素面加进确地预测公
🔏 竞争性因子	, ≓,	司未来现金流,从而做出正确 的投资决策。风险指标根据"
🛠 前瞻性因子	点击打开文件夹 可直接上体的F_DC_DCC_TXT接式做文件来	风脸"相关词汇总词频在公告 中文本总词频的占比计算得
② 区块链概念因子		ш.
👗 人口老龄化指标		
🕹 员工待遇因子		
🕑 贸易战潜在风险		
📀 管理会计因子		

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。



风险因子 分析得出文本数据的风险因子指标												
この 「「「「」」」 「「」」」」 「」」 「」」」 「」」 「」」」 「」」」												
	查询]					T 导出TXT	5 导出表格				
文件名	上传时间	文件大小	总宁数	总句数	总页数	总词数	词颊合计	指标值				
ST全新:关于对深交所 (2021) 第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-05-14	67.87KB	26835	1871	30	15206	254	1.670%				
ST全新:独立董事关于对深交所公司部年报问询函(2022)第 162 号回复之独立意见08c6a88b6dc842e0be1	2024-05-14	5.13KB	1923	122	3	1101	23	2.089%				
ST全新:碱公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf	2024-05-14	36.34KB	13936	969	18	8286	102	1.231%				
ST国华: 致同会计师事务所 (特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询函的回复d0c	2024-05-14	94.37KB	35953	2498	42	20772	196	0.944%				
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	32	0.373%				
	-页 1 2 下	页 尾页										

7.5.3 竞争性因子

企业竞争战略指的是在企业总体战略的制约下,指导和管理具体战略经营单位的计划和行动。它是企业进行 经营管理和内部资源配置的指南针,竞争战略的选择决定了企业最终表现结果的质量。 竞争性指标根据"竞 争性"相关词汇总词频在公告中文本总词频的占比计算得出。

文本特征	竞争性因子 分析得出文本数据的竞争性因子数据指标	功能简介
抽述性创新 通 风险因子		企业党争战略指的是在企业息 体战略的制约下,指导和管理 具体战略经营单位的计划和行 动。它是企业进行经营管理和
🔀 竞争性因子	E. E	内部资源配置的指南针, 竞争 战略的选择决定了企业最终表
🛠 前瞻性因子	点击打开文件夹 可直接上传PDF、DOC、DOCX、TXT格式的文件夹	现结果的质量。 竞争性指标相 据 "竞争性"相关词汇总词频
② 区块链概念因子		在公告中文本总词频的占比计 算得出。
👗 人口老龄化指标		
🥹 员工待遇因子		
🕑 贸易战潜在风险		
📀 管理会计因子		

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。



竞争性因子	分析得出文本数据的竞争性因子数据指标										
「 「 「 「 」 」 」 」 」 二 二 二 二 二 二 二 二 二 二 二 二 二											
		已分析:		文件分析							
	l	查询						T 导出TXT	5 导出表格		
	文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	词颊合计	指标值		
ST全新:关于对深交所	(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-05-14	67.87KB	26835	1871	30	15206	32	0.210%		
ST全新:独立董事关于对	深交所公司部年报问询函〔2022〕第 162 号回复之独立意见08c6a88b6dc842e0be1	2024-05-14	5.13KB	1923	122	3	1101	1	0.091%		
ST全新: 诚公律师事务所	对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf	2024-05-14	36.34KB	13936	969	18	8286	3	0.036%		
ST国华: 致同会计师事务	所(特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询函的回复doc	2024-05-14	94.37KB	35953	2498	42	20772	193	0.929%		
昌平	区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	188	2.189%		
	首页 上-	-页 1 2 下	页 尾页								

7.5.4 前瞻性因子

前瞻性,指文本中包含与公司未来发展和业绩相关的信息含量。对投资者了解公司前景、预测发展趋势以及 评估未来公司价值具有极大帮助。前瞻性指标根据"前瞻性"相关词汇总词频在公告中文本总词频的占比计 算得出。

RESSET 文本分析客户端-小锐助手

RESSET 锐思数据

뢷 文本特征	前瞻性因子 分析得出文本数据的前瞻性因子数据指标	功能简介
△ 描述性创新		前瞻性, 指文本中包含与公司 未来发展和业结相关的信息含
道 风险因子		量。对投资者了解公司前景、 预测发展趋势以及评估表来公
🔀 竞争性因子		司价值具有极大帮助。前瞻性 指标根据"前瞻性"相关词汇
分前 前 着	点击打开文件夹 可直接上作PDF、DOC、DOCX、TXT格式的文件夹	总词频在公告中文本总词频的 占比计算得出。
● 区块链概念因子		
👗 人口老龄化指标		
🕹 员工待遇因子		
🕑 贸易战潜在风险		
🗟 管理会计因子		

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。





7.5.5 区块链概念因子

区块链研究数据库是基于对中国上市公司公告进行整理、解析和分析,从而建立的专业数据库。区块链技术 的集成应用在新的技术革新和产业变革中起着重要作用,是企业自主创新的重要突破口。本数据库通过分析 上市公司在公告中涉及区块链及区块链相似词集的程度。

뢷 文本特征	区块链概念因子 分析得出文本数据的区块链概念因子数据指标	功能简介
▲ 描述性创新		区块链研究数据库是基于对中 国上市公司公告进行整理 創
🚊 风险因子		析和分析,从而建立的专业数 据库 区块链技术的集成应用
🔀 竞争性因子		在新的技术革新和产业变革中 記差重要作用 易企业自主
📢 前瞻性因子	点击打开文件夹	新的重要突破口。本数据库通过分析上市公司在公告中涉及
② 区块链概念因子	り <u>当後上</u> 行PDF、DUC、UUCA、IATMAABX代来	区块链及区块链相似词集的稍 度。
👗 人口老龄化指标		
🕹 员工待遇因子		
🕑 贸易战潜在风险		
🙈 管理会计因子		

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。

区块链概念因子 分析得出文本数据的区块链概念因子数据指标



区块链概念因子 分析得出文本数据的区块链概念因子数据指标

可直接上传PDF 已选择数据文件:D\360安全浏览器下载\1111	<mark>三,</mark> 点击打开文件夹 ;, DOC, DOCX, TX	「格式的文件夹						
	已分析:		文件分析					
[查询						T 导出TXT	5 导出表格
文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	词颊合计	指标值
ST全新:关于对深交所(2021)第225号年报问询图回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-05-14	67.87KB	26835	1871	30	15206	0	0.000%
ST全新:独立董事关于对深交所公司部年报问询函〔2022〕第 162 号回复之独立意见08c6a88b6dc842e0be1…	2024-05-14	5.13KB	1923	122	3	1101	0	0.000%
ST全新: 诚公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf	2024-05-14	36.34KB	13936	969	18	8286	0	0.000%
ST国华:致同会计师事务所(特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询函的回复d0c	2024-05-14	94.37KB	35953	2498	42	20772	0	0.000%
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	1	0.012%
	-页 1 2 下	页 尾页						

7.5.6 人口老龄化指标

	人口老龄化指标 分析得出文本数据的人口老龄化指标数据指标	功能简介
· 加速性创新		老龄化指标根据"人山老龄 化"相关词汇总词频在年度财 务报告中文本总词频的占比计
🔀 竞争性因子	E.	算侍田, 结百财经义本语喝闷 建出能够有效测度上市公司对 老龄化相关信息披露的指标
🛠 前瞻性因子		管在帮助信息使用者更为高效 地分析上市公司老龄化信息波
● 区块链概念因子	SHRTML DOCK DOCK INTHEMPISIH	露情况。
👗 人口老龄化指标		
👂 员工待遇因子		
🕑 贸易战潜在风险		
📀 管理会计因子		

老龄化指标根据"人口老龄化"相关词汇总词频在年度财务报告中文本总词频的占比计算得出,结合财经文本语境构建出能够有效测度上市公司对老龄化相关信息披露的指标,旨在帮助信息使用者更为高效地分析上市公司老龄化信息披露情况。

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。



人口老龄化指标 分析得出文本数据的人口老龄化指标数据指标 E, 点击打开文件 可直接上传PDF、DOC、DOCX、TXT格式的文件夹 👖 导出TXT 🗧 导出表格 上传时间 总句数 总页数 总词数 词颊合计 文件名 文件大小 总字数 指标值 2024-05-14 ST全新:关于对深交所 (2021) 第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt 0.000% 67.07KB 14500 0 ST全新:独立董事关于对深交所公司部年报问询函 [2022] 第 162 号回复之独立意见08c6a88b6dc842e0be1... 2024-05-14 5.10KB 1096 0.000% ST全新: 诚公律师事务所对全新好2021年报问词函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf... 2024-05-14 35.96KB 0.000% ST国华: 致同会计师事务所 (特殊普通合伙) 关于深圳国华网安科技股份有限公司2021年年报问询函的回复d0c 2024-05-14 92.57KB 35803 20089 0.000% 昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt 2024-05-14 42.84KB 15306 955 8589 0.012% 首页 上一页 1 2 下一页 尾页

7.5.7 员工待遇因子

RESSET 4	\锐助手	▲ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○	Q 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一 一		- ×
 文本特征 描述性创新 	员工待遇因子 分析得出文本数据的员工	_待遇因子数据	指标		功能简介 员工待遇因子是衡量企业对员 工友好程度的文本指标。本教
▲ 风险因子			_		据库基于马斯洛需求理论选取 出7个员工友好维度进行员工
			<mark>三</mark> , 点击打开文件夹		友好度指标的构建。具体来 讲,7个员工友好维度包括: 基本权益、员工福利、健康与
★ 前端住凶子 ● 区块链概念因子		可直接上传	PDF、DOC、DOCX、TXT框	试的文件夹	安全、人文关怀、沟通、晋 升、自我实现。
ALL					
🕹 员工待遇因子					
🔮 贸易战潜在风险			the Ma		
🖻 管理会计因子			宣闻		

员工待遇因子是衡量企业对员工友好程度的文本指标。本数据库基于马斯洛需求理论选取出 7 个员工友好维 度进行员工友好度指标的构建。具体来讲,7 个员工友好维度包括:基本权益、员工福利、健康与安全、人文 关怀、沟通、晋升、自我实现。

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。



员工待遇因子分析得出文本数据的员工	待遇因子	·数据指	眎										
この 「 二 二 二 二 二 二 二 二 二 二 二 二 打 开 文 件 夹 可 直接上体PDF、DOC、DOCX、TXT格式的文件夹 已选择数据文件:D\360安全浏览器下载\1111													
							文件分析						
			럩	銄							1 5	≩出TXT [s	,导出表格
文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	基本权益	员工福利	健康与安全	人文关怀	沟通	晋升	自我实现
ST全新:关于对深交所 (2021) 第225号年报问询函回复的公告6f036587fb9e	2024-05-14	67.07KB	27322	905	30	14500	0.007%	0.041%	0.000%	0.000%	0.090%	0.000%	0.000%
ST全新:独立董事关于对深交所公司部年报问询图 [2022] 第 162 号回复之独	2024-05-14	5.10KB	1971	70	3	1096	0.000%	0.000%	0.000%	0.000%	0.091%	0.000%	0.000%
ST全新:诚公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e8	2024-05-14	35.96KB	14251	515	18	8115	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
ST国华: 致同会计师事务所 (特殊普通合伙) 关于深圳国华网安科技股份有限公	2024-05-14	92.57KB	35803	1457	42	20089	0.025%	0.030%	0.000%	0.000%	0.055%	0.000%	0.000%
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	0.000%	0.012%	0.012%	0.000%	0.116%	0.000%	0.000%
	[首页 上-	_页 1	2 下	一页属	页							

7.5.8 贸易战潜在风险



企业微观层面的、随时间变化的、用于衡量中国上市公司感知到的贸易战不确定性指标。

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。

贸易战潜在风险分析得出文本数据的贸易战潜在风险因	日子数据指标	ŧ							
三 - - - - - - - - - - - - -									
	查询						T 导出TXT	5 导出表格	
文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	词颊合计	指标值	
ST全新:关于对深交所(2021)第225号年报问询函回复的公告6f036587fb9e4bc9847c623f60883630.pdf.txt	2024-05-14	67.07KB	27322	905	30	14500	0	0.000%	
ST全新:独立董事关于对深交所公司部年报问询图 [2022] 第 162 号回复之独立意见08c6a88b6dc842e0be1	2024-05-14	5.10KB	1971	70	3	1096	0	0.000%	
ST全新: 诚公律师事务所对全新好2021年报问询函的法律意见书62c8a29e02e84169aae679089e5732d7.pdf	2024-05-14	35.96KB	14251	515	18	8115	0	0.000%	
ST国华: 致同会计师事务所(特殊普通合伙)关于深圳国华网安科技股份有限公司2021年年报问询函的回复d0c	2024-05-14	92.57KB	35803	1457	42	20089	0	0.000%	
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	0	0.000%	
首页上一	-页 1 2 下	页 尾页							

7.5.9 管理会计因子

文本特征 描述性创新	管理会计因子分析得出文本数据的管理会计因子数据指标	功能简介 管理会计指标是衡量企业管理
道 风险因子		会计活动的指标。根据指标相 关词汇总词频在年度财务报告
🔀 竞争性因子		中文本念向過的口比口算得出。
🛠 前瞻性因子	点击打开文件夹 可自然 计维护区 DOC DOCK TXT器式做文件来	
● 区块链概念因子		
👗 人口老龄化指标		
👂 员工待遇因子		
👽 贸易战潜在风险		
🗠 管理会计因子		
		and a second

管理会计指标是衡量企业管理会计活动的指标。根据指标相关词汇总词频在年度财务报告中文本总词频的占比计算得出。

功能简介:点击打开文件夹位置,可直接上传含 PDF、DOC、DOCX、TXT 格式的文件夹,小锐助手可将上传的文件转换成方便计算机处理的 TXT 格式文件进行解析。

管理会计因子 分析得出文本数据的管理会计因子数据指标



管理会计因子 分析得出文本数据的管	理会计区	因子数	据指	示										
こ。 点击打开文件夹 可直接上传PDF、DOC、DOCX、TXT格式的文件夹 已选择数据文件:D\\360安全浏览器下载\1111														
								文件分析学						
				查道]							.)出TXT E	导出表格
文件名	上传时间	文件大小	总字数	总句数	总页数	总词数	词颊合计	战略管理	预算管理	成本管理	营运管理	投融资管理	绩效管理	风险管理
ST全新:关于对深交所 (2021) 第225号年报问询函回复的公告6f036587f	2024-05-14	67.07KB	27322	905	30	14500	0	0.000%	0.000%	0.000%	0.021%	0.062%	0.000%	0.000%
ST全新:独立董事关于对深交所公司部年报问询函 (2022) 第 162 号回复	2024-05-14	5.10KB	1971	70	3	1096	0	0.000%	0.000%	0.000%	0.000%	0.091%	0.000%	0.000%
ST全新: 诚公律师事务所对全新好2021年报问询函的法律意见书62c8a29e	2024-05-14	35.96KB	14251	515	18	8115	0	0.000%	0.000%	0.000%	0.025%	0.074%	0.000%	0.000%
ST国华:致同会计师事务所(特殊普通合伙)关于深圳国华网安科技股份有	2024-05-14	92.57KB	35803	1457	42	20089	0	0.000%	0.040%	0.000%	0.065%	0.005%	0.010%	0.000%
昌平区2022年政府工作报告c1e521a152664ff5bb462f83b3de4d17.txt	2024-05-14	42.84KB	15306	955	1	8589	1	0.000%	0.070%	0.000%	0.023%	0.000%	0.035%	0.012%
		首页	上一页	ī 1	2 下-	页尾	页							

8 文本挖掘

RESSET	▲ Q 前页 通频提案 文	✓ ◆ ◆ ★特征 文本/2版 投資	- ×
	文本挖掘 自动化多	维度多文本的文本挖掘	
	步骤1:文本预处理 本步骤是文本挖掘的预处理过程,每一个小步骤都可以得到对应的结果文件,后 一步可以相關前一步的分析结果文件进行设定,也可以自定义文件进行设定,既 支持多指标流程化也支持独立分析单个指标。	步骤2:机器学习 本步骤包含文本相似度、文本分类、文本聚类和主题模型等功能,支持步骤一的无缝 结合,也支持自定义文本进行一键式源程化分析,	
	 ⇒ 文本转换 ↓ ▲ 	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	
	文即休验 ↓ ↓ ↓	立 即体 強 立 政 集 立 政 集 立 原 体 验 立 原 体 验	

8.1 文本预处理

RESSET小锐助手		前页	● ● 文本数据	Q 词频搜索	~~ 文本特征	人 文本挖掘	(後期)		- ×
	文本	挖掘	自动	北多	维度多	文本的	的文本排	挖掘	
	步骤1:文本预处理 本步骤最文本挖掘的预处理过程,每一个小步骤都可 一步可以根据前一步的分析结果文件进行设定,也可 支持多指标流程化也支持拉立分析单个指标。	T以得到对应的 T以自定义文件	结果文件,后 进行设定,既	ξ E	步骤 本步 结合	12:机器学习 骤包含文本相似 ,也支持自定义	度、文本分类、文 文本进行一键式游	文本聚类和主题模型等功能,支持步骤一的无缝 施理化分析,	
	 ⇒ 文本转換 ↓ ▲ ▲ 	立即体验	∂ ≙			相位		立 即 体 验 文本分类	
		立即体到 本向量化 1	☆ ZIII)(43会			ا ¢		立 即 中 始 全 三 題 模型 全 節 序 体 验 章	
					I				
									袋 思 服

文本预处理:本步骤是文本挖掘的预处理过程,每一个小步骤都可以得到对应的结果文件,后一步可以根据 前一步的分析结果文件进行设定,也可以自定义进行设定,既支持多指标流程化也支持独立分析单个指标。

8.1.1 文本转换

RESSET		- ×
x本預处費 ② 文本转換 ③ 文本分词 〒 古奈用詞 『 特征词库 ※ 文本の词里化 明潔字习 ● 相似度计算 ◎ 文本分淡 ◎ 文本分淡	政策 文井祉 文井祉 文井祉 政策 文本转换 將PDF、DOC、DOCX转为TXT文件	- × 功能简介 1、通过选择文件夹直接读取PDF、DOCX、 DOC的内容,转换成方便计算机处理的TXT 格式文件。
 ● 文本築奏 ◇ 主题模型 		

功能简介:通过选择文件夹直接读取 PDF、DOCX、DOC 的内容,转换成方便计算机处理的 TXT 格式文件。



文本预处理	文本转换 将PDF、DOC、DOCX转为TXT文件
 文本转換 文本行词 文本分词 法停用词 特征词库 文本向量化 	この 「 こ 二 二 二 二 二 二 二 二 二 二 二 二 二
机器学习	已选择数据文件:D\\360安全浏览器下载\pdf、doc、txt
 	
■ 文本聚类	
♀ 主题模型	

文本预处理	文本转换 將PDF、DOC、DOCX转为TXT文件								
 ⇒ 文本转換 ▲ 文本分词 ↓ 法停用词 ⑦ 特征词库 ☑ 文本向量化 机器学习 	この 点击打开文 可直接上传PDF、DOC、DOC 已选择数据文件D\360安全浏览器下载\pdf、doc、bt	<mark>7件夹</mark> CX、TXT格式的文件夹							
参 文本分类		已分析: 5/5	文件分析完	成					
 — → → → → → → → → → → → → → → → → → → → → → → → → → → → → → → → → → → 									
						日本	5 导出表格		
	文件名	上传时间	文件大小	总字数	总词数	总句数	总页数		
	PBL联合CBS教学模式对实习生批判性思维能力的影响,pdf.txt	2023-12-18	23.01KB	11357	5163	459	5		
	会计教学中批判性思维教学法运用策略分析.pdf.txt	2023-12-18	11.47KB	4147	2568	271	1		
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	2023-12-18	2.36KB	832	832	243	1		
	技术参数.docx.txt	2023-12-18	4.81KB	1815	1083	109	1		
	本科生批判性思维能力及高影双一流"建设高校的实证研究.pdf.txt	2023-12-18	39.12KB	17730	8860	467	8		
	首页 上一页 1	下一页 尾页							

8.1.2 文本分词

功能简介:上传需要的文本文件,自动输出分词结果。

分词模型分文三种:精准模式分词、全模式分词和搜索引擎模式。

》**精准模式分词**:将文本精确地切分成词语,适用于文本分析和挖掘等任务。在精确模式下,会对待分析的 句子进行全词匹配,找出所有可能的词语,然后依据词典中的词语频率计算出最可能切分结果,分词结果不 存在冗余。

文本预处理	文本分词 提供可视化分词工具,自动得出分词结果						
 ○ 文本转換 ▲ 文本分词 环 大使用词同 図 特征词ゅ 必 文本向量 • 机器学习 ● 相似度计算 ◆ 文本聚業 ④ 主题模型 	こ。 点击打开文件夹 可直接上传TXT格式的文件夹 已选择数据文件:F\Program Files\小税助手\customdata\6d8996fa1c5349b9969f18244408b88d						
	▲ 全模式分词 ■ 接索引擎分词						

文件为上一步上传解析后的结果文件。

文本预处理 	文本分词 提供可视化分词工具, 自动得出分词结果										
 ⇒ 文本转換 ▲ 文本分词 征 去停用词 ⑦ 特征词库 ☑ 文本向量化 机器学习 	点击打开文件夹 可直接上作TXT格式的文件夹 已选择数据文件F\Program Files\小税助手\customdata\6d8996fa1c5349b9969f18244408b88d										
 □ 相似度计算 ◆ 文本分类 □ 文本聚类 ▲ 大本聚类 	☑ 藉產模式分词 ■ 全模式分词 ■ 搬索引擎分词		已分析: 5/5 文件分析完成								
① 主题模型				<mark>て</mark> 导出	TXT 5	导出表格					
	文件名	上传时间	文件路径	文件大小	总字数	总词数					
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	2023-12-18	customdata/29b2f6861a3a492b8a60df4622326226/PBL联合CBS教学模式对实	23.01KB	11357	5163					
	会计教学中批判性思维教学法运用策略分析.pdf.txt	2023-12-18	customdata/29b2f6861a3a492b8a60df4622326226/会计教学中批判性思维教学	11.47KB	4147	2568					
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	2023-12-18	customdata/29b2f6861a3a492b8a60df4622326226/初中物理科学探究中培养学	2.36KB	832	832					
	技术参数.docx.txt	2023-12-18	customdata/29b2f6861a3a492b8a60df4622326226/技术参数.docx.txt	4.81KB	1815	1083					
	本科生批判性思维能力及高影…双一流"建设高校的实证研究.pdf.txt	2023-12-18	customdata/29b2f6861a3a492b8a60df4622326226/本科生批判性思维能力及高	39.12KB	17730	8860					
		首页	上-页 1 下-页 尾页								

》**全模式分词:**将文本中所有可能词语都分解出来,适用于一些对速度要求比较高的场景,分词结果存在冗余。例如:"中华人民中和国",分词结果为:"中华","中华人民","中华人民共和国","华 人","人民"等。

文本预处理	文本分词 提供可视化分词工具,自动得出分词结果									
 	こ。 点击打开文件夹 可直接上作TXT格式的文件夹 已选择数缆文件乐\Program Files\小铁助手\customdata\6d8996fa1c5349b9969f18244408b88d									
 • • •	■ 精准模式分词 🛛 全模式分词 🛛 ■ 搜索引擎分词		已分析: 5/5 文件分析完成							
■ 文本聚类 ◇ 土販増型										
				1 导出	TXT S	导出表格				
	文件名	上传时间	文件路径	文件大小	总字数	总词数				
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	2023-12-18	customdata/1a145fcc6da3439d884c7e7b99db8b18/PBL联合CBS教学模式对	23.01KB	11357	5826				
	会计教学中批判性思维教学法运用策略分析.pdf.txt	2023-12-18	customdata/1a145fcc6da3439d884c7e7b99db8b18/会计教学中批判性思维教学	11.47KB	4147	3082				
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	2023-12-18	customdata/1a145fcc6da3439d884c7e7b99db8b18/初中物理科学探究中培养学	2.36KB	832	791				
	技术参数.docx.txt	2023-12-18	customdata/1a145fcc6da3439d884c7e7b99db8b18/技术参数.docx.txt	4.81KB	1815	1237				
	本科生批判性思维能力及高影双一流"建设高校的实证研究.pdf.txt	2023-12-18	customdata/1a145fcc6da3439d884c7e7b99db8b18/本科生批判性思维能力及高	39.12KB	17730	9873				
		首页	上一页 1 下一页 尾页							

》**搜索引擎模式**:在精确模式的基础上,对长词再次使用全模式进行切分,最终得到最可能的切分结果,适用于搜索引擎等场景。

RESSEI 況	思数据	RESSET 文本分析客F	⊐端−ノ	小锐国	助手					
文本预处理	文本分词 提供可视化分词工具,自动得出分词结果									
 ○ 文本转換 ◎ 文本有過 ○ 大本分词 □ 大停用词 □ 特征:同庫 ☑ 文本向量化 机器学习 □ 相似度计算 ◎ 文本分类 □ 本の当 	こ こ こ こ こ た ま お 市 ガ 子 大 中 た は 大 大 中 の は 上 低 下 大 中 大 中 た は 大 大 中 大 中 の さ は 上 低 下 大 作 来 の う は 上 低 下 大 下 も の 文 件 大 い 戦 助 手 し に よ い 本 い 本 の 本 し 本 本 4 4 4 8 8 8 8 6 4 8 9 9 9 1 1 5 5 4 4 4 8 8 8 8 6 4 8 9 9 9 1 8 1 4 4 4 8 8 8 8 8 8 8 8 8 8 8 8 8									
 ■ 文本聚类 ⑦ 主题模型 			查询							
		- E4	чтут 🛛 🖂	日中主权						
	文件名	上传时间		□ 导出 文件大小	∃TXT S	导出表格 总词数				
	文件名 PBL联合CBS数学模式对实习生批判性思维能力的影响.pdf.txt	上传时间 2023-12-18	文件路径 customdata/31a462f731eb422ab283ea22c91c8c20/PBL联合CBS教学模式对实	文件大小 23.01KB	±TXT 国 总字数 11357	导出表格 总词数 5758				
	文件名 PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt 会计教学中批判性思维教学法运用策略分析.pdf.txt	上传时间 2023-12-18 2023-12-18	文件器经 customdata/31a462f731eb422ab283ea22c91c8c20/PBL联合CBS数学模式对实 customdata/31a462f731eb422ab283ea22c91c8c20/会计数学中批判性思维数学	文件大小 23.01KB 11.47KB	3TXT S 总字数 11357 4147	导出表格 总词数 5758 2795				
	文件名 PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt 会计教学中批判性思维教学法运用策略分析.pdf.txt 初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	上传时间 2023-12-18 2023-12-18 2023-12-18	文件器经 customdata/31a462f731eb422ab283ea22c91c8c20/PBL联合CBS数学模式对实 customdata/31a462f731eb422ab283ea22c91c8c20/会计教学中批判性思维教学 customdata/31a462f731eb422ab283ea22c91c8c20/知中物理科学探究中培养学	文件大小 23.01KB 11.47KB 2.36KB	TXT 写	导出表格 总词数 5758 2795 832				
	文件名 PBL联合CBS教学模式对实习生批判性思维能力的影响,pdf.txt 会计教学中批判性思维教学法运用策略分析,pdf.txt 初中物理科学探究中培养学生科版8年级"浮力"教学为例,pdf.txt 技术参数,docx.txt	上传时间 2023-12-18 2023-12-18 2023-12-18 2023-12-18	文件器径 customdata/31a462f731eb422ab283ea22c91c8c20/PBL联合CBS数学模式对实 customdata/31a462f731eb422ab283ea22c91c8c20/会计数学中批判性思维教学 customdata/31a462f731eb422ab283ea22c91c8c20/初中物理科学探究中培养学 customdata/31a462f731eb422ab283ea22c91c8c20/技术参数.docx.txt	 文件大小 23.01KB 11.47KB 2.36KB 4.81KB 	 TXT 国 記字数 11357 4147 832 1815 	导出表格 总词数 5758 2795 832 1119				
	文件名 PBL联合CBS教学模式对实习生批涉性思维能力的影响,pdf.txt 会计教学中批涉性思维教学法运用策略分析,pdf.txt 初中物理科学探究中培养学生科版8年级 "浮力"教学为例,pdf.txt 技术参数,docx.txt 本科生批判性思维能力及高影双一流 建设高校的实证研究,pdf.txt	上作时间 2023-12-18 2023-12-18 2023-12-18 2023-12-18 2023-12-18	文件器径 customdata/31a462f731eb422ab283ea22c91c8c20/PBL联合CBS教学模式对实 customdata/31a462f731eb422ab283ea22c91c8c20/会计教学中批判性思维教学 customdata/31a462f731eb422ab283ea22c91c8c20/切中物理科学探究中培养学 customdata/31a462f731eb422ab283ea22c91c8c20/技术参数.docx.txt customdata/31a462f731eb422ab283ea22c91c8c20/技术参数.docx.txt	 (1) 号 文件大小 23.01KB 11.47KB 2.36KB 4.81KB 39.12KB 	Image: Total and the second	导出表格 总词数 5758 2795 832 1119 9906				

8.1.3 去停用词

功能简介:在文本分析和文本挖掘的任务中,只关注有价值、有意义的文本信息,去停用词是为了将一些无价值意义的词语(也被称为停用词)剔除,停用词不仅对挖掘文本信息没有帮助,而且会降低文本挖掘的效率,增加数据中无用的噪声。

上传文件使用文本转换后的文件,用户也可以重新上传新的文件进行解析。

文本预处理	去停用词 智能剔除分词后的停用词,得到全新分词结果										
 ⇒ 文本转換 △ 文本 (小) △ 文本分词 ○ 法 (小) ○ 特征词库 ② 文本向量化 机器学习 	こ。 点击打开文件夹 可直接上传TXT格式的文件夹 已选择数据文件F\Program Files\小税助手\customdata\31a462f731eb422ab283ea22c91c8c20										
 • 田似度计算 ◇ 文本分类 ● 文本聚类 ⑦ 主题模型 											
	文件名	上传时间	文件諸径	文件大小	总字数	总词数					
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	2023-12-18	customdata/716a7980ae9f4038bcdac506472a3ce0/PBL联合CBS教学模式对实	32.56KB	12542	2975					
	会计教学中批判性思维教学法运用策略分析.pdf.txt	2023-12-18	customdata/716a7980ae9f4038bcdac506472a3ce0/会计教学中批判性思维教学	15.70KB	4597	1412					
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	2023-12-18	customdata/716a7980ae9f4038bcdac506472a3ce0/初中物理科学探究中培养学	3.21KB	832	42					
	技术参数.docx.txt	2023-12-18	customdata/716a7980ae9f4038bcdac506472a3ce0/技术参数.docx.txt	6.12KB	1885	602					
	本科生批判性思维能力及高影双一流"建设高校的实证研究.pdf.txt	2023-12-18	customdata/716a7980ae9f4038bcdac506472a3ce0/本科生批判性思维能力及高	55.97KB	19733	5832					
		首页	上页 1 下页 尾页								

8.1.4 特征词库

功能简介:特征词指的是在文本中具有一定信息量或表征文本特征的词语,这些词语通常被用来描述文本的 主题、情感或其他重要的属性。根据用户上传文件及内嵌特征词库提取文本中的特征词。



Ť		닏		1/41	· _\$	-1	2=	Υ, ι	9	С,	ů,		(44×											
	张薇 周兰 火红	中学 理论	翻转 课堂	と 设计 成人	人 护理 理	学教学	评价 解放	杂志 红梅	≢ 涉外 ₱	€平 号书	枕 改革 刻	改果 研究 中[国 管理 周	家 混合 学	习 理念 胡	雁 我国 博士	と 课程 设置	閏 模式 构建	1 用于 妇产	产科 实践 都	参考 文献 基	础 技能 分	斤 农村 卫生	生事 事业 🕴

☐ 5 ~ ♂ ऱ	د	文本挖掘结果集_1.csv -	Excel		録 🖬	- 0 ×
文件 开始 插入 页面布局 公式 数据 审阅	视图 帮助 Acroba	at 更多工具 📿	操作说明搜索			\Box
● ● ● ● ● ● ● ● ● ■	き 三 ● = ◆ ・ 。 き 三 = ● = ● = ○ ジャー ・ がた 式	 常规 □ □ 数字	 ・ ・ ・	 	。 选 查找和选择 章	加 載项 加載项 へ
A1 • : × ✓ fx 文件名						~
A 1 文件名 2 716a7980ae9f4038bcdac506472a3ce0 3 4 5 6 6 7 8 9 10 11 12 13 14 15 16	BC 	D 里论 翻转 课堂 设计,	F G 或人 护理 理学 教学 评价	H J 解放 杂志 红梅 涉外 专业 f	6式改革效果	L ◆ 研究中国管理 /
17 18 19 20 21 22						

8.1.5 文本向量化

功能简介:用户上传文件合集,根据用户选择的算法计算出文本中词的向量值,进一步可以得到整篇文档的向量表示,可用于后续的文本挖掘计算。





文本预处理	文本向量化 数值向量来表示文本的语义										
🖨 文本转换											
실 文本分词											
℃ 去停用词											
🕞 特征词库											
🗾 文本向量化	点击打开文件夹										
机器学习 	可直接上作TXT格式的文件夹 已选择数据文件F:\Program Files\小税助手\customdata\f33f49d17d3d47cf947b0735db9191eb										
中 相似度计算											
会 文本分类	TF-IDF TFC V LTC		文件分析完成								
🗟 文本聚类											
⑦ 主题模型											
	文件名	词	向量值								
	Feature.txt	一幕	0.03162277660168379								
	Feature.txt	一步	0.03162277660168379								
	Feature.txt	一段	0.03162277660168379								
	Feature.txt	一流	0.03162277660168379								
	Feature.txt	# –	0.03162277660168379								
		首页 上一页 1	2 3 4 5 6 7 下一页 尾页								

8.2 机器学习

RESSET小锐助手	★ この ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●		- ×
	文本挖掘 自动化多约	佳度多文本的文本挖掘	
	步骤1:文本预处理 本步要意文本短期的预处理过程。每一个小步骤都可以得到时应的结果文件,后 一步可以根据前一步的分析结果文件进行设定,也可以自定义文件进行设定,既 支持多指标流程化也支持独立分析单个指标。	步骤2:机器学习 本步骤包含文本相似意、文本分类、文本聚类和主题模型等功能,支持步骤一的无缝 结合,也支持自定义文本进行一键式流程化分析。	
	 ⇒ 文本转換 立即体验 ↓ ↓ ↓ 文本分词 立即体验 		
	↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	正正 正正 正正 正正 作 生 証頼理 主 証頼理 生 証頼理	
			ļ
			思 著 服

机器学习:本步骤包含文本相似度、文本分类、文本聚类和主题模型等功能,支持步骤一的无缝结合,也支持自定义文本进行一键式流程化分析。

8.2.1 相似度计算

功能简介:上传文本文件合集,利用 TF-IDF 算法计算每篇文档中的向量表示,选择不同的距离公式计算得出 所有文档之间的两两相似度。

RESSET 文本分析客户端-小锐助手





余弦距离:

文本预处理 	相似度计算·算得出文件夹中每个文件之间的相似度		
 文本转换 文本转换 文本分词 云本停用词 标征词库 文本向量化 机器学习 	可」 已选择数据文件:DA360安全浏览器下载ATXT	<mark>こし、</mark> 点击打开文件夹 li接上作TXT版式的文件夹	
 □ 相似度计算 シ本分类 	▼ 余弦距离 ■ 欧式距离 ■ Jaroar4距离	已分析: 18/18 文件分析完成	
 ■ 文本聚类 ⑦ 主题模型 			
	文件1	文件2	相似度
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	100.000
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	49.376
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	会计教学中批判性思维教学法运用策略分析.pdf.txt	54.259
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	0
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	基于思辨能力培养的中学语文学语文批判性思维教学室例》.pdf.txt	41.620
	首页 上一页 1	2 3 4 5 6 7 下一页 尾页	

欧氏距离:

文本預处理	相似度计算·算得出文件夹中	每个文件之间的相似度			
 ○ 文本转換 △ 文本分词 ↓ 法停用词 ⑦ 特征词库 ☑ 文本向量化 机器学习 	已选择数额文件:D\360安全浏览	8 F\$\$\TXT	<mark>三</mark> 点击打开文件夹 可直接上传TXT格式的文件夹		
日 相似度计算					
📚 文本分类	🗌 余弦距离 🔛 欧式距离	Jarcard距离	已分析: 6/18	文件分析中	
島 文本聚类					
⑦ 主题模型			進行		5 导出麦格

文本预处理 	相似度计算 ·算得出文件夹中每个文件之间的相似度			
 ⇒ 文本转換 ▲ 文本分词 ① 文本分词 □ 示 告停用词 □ 行征词库 ☑ 文本向量化 机器学习 	可」 已选择数据文件D\360安全浏览器下载\TXT	<mark>テ</mark> , 点击打开文件夹 目接上修TXT描式的文件夹		
 ● 相似度计算 ◎ 文本分类 	■ 余弦距离 🖌 歐式距离 📕 Jarcar 御藤			
■ 文本聚类				
		畫狗	5 导出表格	
	文件1	文件2	相似度	
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	100	
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	中学语文教学培养学生思辨能学语文批判性思维教学室例》.pdf.txt	9.235	
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	会计教学中批判性思维教学法运用策略分析.pdf.txt	9.662	
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	7.966	
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	基于思辨能力培养的中学语文学语文批判性思维教学案例》.pdf.txt	10.414	
	首页 上一页 1	2 3 4 5 6 7 下一页 尾页		

Jarcard 距离:

文本预处理	相似度计算·算得出文件夹中	每个文件之间的相似度		
 文本转换 文本分词 文本分词 去停用词 特征词库 文本向量化 机器学习 	快 词 词 库 量化 已选择数 <u>原文</u> 件:D:\360安全浏览器下载\TXT		<mark>こ</mark> , 点击打开文件夹 可直接上作TXT格式的文件夹	
 □ 相似度计算 ⇒ 文本分类 ■ 文本聚类 		🖌 Jaroard距离	文件	上传完成!
⑦ 主题模型				6 导出表格)



	首页	上一页		2	3	4	5	6	7	下一页	尾页
--	----	-----	--	---	---	---	---	---	---	-----	----

A	A1 \checkmark : $\times \checkmark f_{*}$ \Diamond /t+1					
	Α	В	С	D		
1	文件1	文件2	相似度			
2	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	100			
3	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	2.5			
4	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	会计教学中批判性思维教学法运用策略分析.pdf.txt	2			
5	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	0			
6	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	基于思辨能力培养的中学语文学语文批判性思维教学案例》.pdf.txt	2.222			
7	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	审辩式思维在创造力发展中的重要性.pdf.txt	0			
8	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	审辩式思维培养水平的评估指标研究.pdf.txt	1.37			
9	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	审辩式思维教育具有的现实意义.pdf.txt	0			
10	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	审辩式思维能力及其测量.pdf.txt	1.887			
11	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	审辩式思维观照下的初中文言文教学.pdf.txt	1.818			
12	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	循本源抓关键 提升审辩式思维能力.pdf.txt	0			
13	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	思维的批判性与大学的批判性思维教育.pdf.txt	1.562			
14	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	批判性思维教学:意蕴、特征及建构路径.pdf.txt	1.786			
15	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	批判性思维教育的一个议题:究? 罗素、杜威和图尔敏.pdf.txt	1.493			
16	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	本科生批判性思维能力及高影双一流"建设高校的实证研究.pdf.txt	1.37			
17	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	浅析审辩式思维的培养策略.pdf.txt	1.923			
18	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	美国斯坦福大学批判性思维教育及启示.pdf.txt	1.818			
19	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	认知偏差与批判性思维教育.pdf.txt	0			
20	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	2.5			
21	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	100			
22	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	会计教学中批判性思维教学法运用策略分析.pdf.txt	3.226			
23	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	0			
24	于账当小学账款体帐底干田售货 能分子等些军用条件服务的(M	甘工田塾46-1位美学中学组合、银行学生生生生生产的16-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-	2046			

8.2.2 文本分类

功能简介:

÷

》用户上传训练集文件和测试文件夹。

训练集文件夹:包含若干个文件夹,这些若干文件夹名称为分类名称,每个分类文件夹中包含若干 txt 文件。

测试集文件夹:包含多个带预测的 txt 文件。

》提取文本特征得到文件特征矩阵,根据用户选择的分类算法建立相应模型,同时根据训练集文件夹文本数据训练模型,模型训练完成后输出测试集文件夹下文件的分类结果。





KNN:

RESSET 文本分析客户端─小锐助手

RESSET 锐思数据

文本预处理 	文本分类 训练文本并预测文本结果				
 文本转换 文本分词 证 去停用词 序 特征词库 文本向量化 机器学习 	<mark>こ。</mark> <u>点击打开文件夹</u> 可自接上作TXT格式的文件夹(JIM线集文件夹) 训练集:D\\360安全浏览器下载\财经文本 - 客户端\\80d38f0fefe548c8ad5f108b8b53f9c1	<mark> 三</mark> 点击打开文件夹 可自接上传TXT语元的文件夹(例试组 激试集:D\\360安全刘览器下载\财经文本 - 客户幽\\353c1e32	: 臣文(十史) 5f1b41ad85d5b9bb2a077d4a		
▶ 相似度计算		فليرجل الإرار الإرار			
◎ 又本分交 ◎ 文本聚类	VRN Decizionizes 25M	又作为作完成			
 主题模型 		_			
			5 导出表格		
	文件		类別		
	万科A: 2021年第二季度报告.pdf.txt		n文件夹1 - 副本		
	乐普医疗: 300003乐普医疗业绩说明会、路演活动信息20210201.pdf.txt n文件夹1 - 副本				
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预测 (6) - 副本.pdf.txt				
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预测 (7) - 副本,pdf.txt n文件夹1 - 副本				
	首页 上一页 1	下页 尾页			

DecisionTree:

文本预处理	文本分类 训练文本并预测文本结果			
 ⇒ 文本转換 ▲ 文本分词 译 去停用词 ☆ 特征词库 业 文本向量化 机器学习 	<mark>こ。</mark> <u>点击打开文件夹</u> 可直接上传TXT格式的文件夹(训练集文件夹) 训练集:D:\360安全浏览篇下载\财经文本 - 寄户端\80d38f0fefe548c8ad5f108b8b53f9c1	こ。 点击打开文件3 可直接上作TXT格式的文件夹(例成) 週試集:D\360安全浏览器下载\财经文本 - 音户幽\353c1e32	2 重文件实) 5/11b41ad85d5b9bb2a077d4a	
 ① 相似度计算 ◇ 文本分类 圖 文本聚类 ◇ ± 55倍刑 	I 1501 ✓ DecisionTree SVM	文件分析完成		
			□ 导出表格	
	文件		类別	
	万科A: 2021年第二季度报告.pdf.txt		n文件夹1	
	乐普医疗: 300003乐普医疗业绩说明会、路演活动信息202	10201.pdf.txt	n文件夹1 - 副本	
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预	迥) (6) - 副本.pdf.txt	n文件夹1 - 副本	
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预	测 (7) - 副本.pdf.txt	n文件夹1 - 副本	
	首页 上一页	1 下一页 尾页		

SVM:

文本预处理	文本分类 训练文本并预测文本结果		
 ➡ 文本转換 ▲ 文本分词 征 去停用词 ☑ 特征词库 ☑ 文本向量化 	こ 点击打开文件夹 可直接上传TXT格式的文件夹(训练集文件夹)	<mark>三,</mark> 点击打开3 可直接上传TXT格式的文件夹	【件夹 (测试重文件夹)
机器学习 	训练集:D:\360安全浏览器下载\财经文本 -	测试集:D:\360安全浏览器下载\财经文本 - 客户端\353	c1e325f1b41ad85d5b9bb2a077d4a
 ① 相似度计算 文本分类 ▲ 文本聚类 	L ISAN DecisionTree SMM	文件分析完成	
⑦ 主题模型		询	5 导出表格
	文件		类別
	万科A: 2021年第二季度报告.pdf.txt		n文件夹1
	乐普医疗: 300003乐普医疗业绩说明会、路演活动信息20210	201.pdf.txt	n文件夹1 - 副本 (2)
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预测	l (6) - 副本.pdf.txt	n文件夹1 - 副本 (2)
	疫情对一带一路沿线国家主要经济指标影响的灰色测度模型构建及预测		n文件夹1 - 副本 (2)
	首页上一页	1 下—页 尾页	

8.2.3 文本聚类

功能简介:对用户上传的数据进行分词处理,然后利用 tf-idf 算法获取 tf-idf 矩阵,利用 PCA 选择重要的特征 信息,最后根据用户选择的聚类算法对所有文档进行聚类。

顾处]	¥	文本聚类	计算出文件夹中的	文本聚类类别			
	划览文件夹					Х	
本本停征本 彩一化本本题	→ ♪ → → → → → → → → → →	音乐 桌面 本地磁盘 (C:) 新加卷 (D:) 360Download 360安全浏览器 doc pdf pdf, doc pdf gdf ydf, doc , RTAS TXT 关键词搜想	ds 音下载 . bxt 客户端 段结果_20231211095828	确定	取消	Line and the second sec	

文本预处理	文本聚类 计算出文件夹中的文本聚类类别
 文本转換 文本转換 文本分词 法停用词 特征词库 文本向量化 机器学习 	こ 点击打开文件夹 可直接上传TXT樁式的文件夹
 • 田 相似度计算 	✓ K-Means算法 ■ DESCAN算法 查询

K-Means 算法:

文本预处理	文本聚类 计算出文件夹中的文本聚类类别			
 文本转换 文本转换 文本分词 还 去停用词 牙征词库 父 文本向量化 机器学习 	三 点击打开文件夹 可直接上传TXT格式的文件夹 已选择数据文件:D\\360安全浏览器下载\TXT			
• 相似度计算				
会 文本分类	v K-Heans算法 ■ DBSCAN算法 文件分析完成			
■ 文本聚类				
♥ 主题模型		5 导出表格		
	文件	类別		
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	类别1		
	中学语文数学培养学生思辨能学语文批判性思维教学案例).pdf.bt	类别1		
	会计教学中批判性思维教学法运用策略分析。pdf.txt 类别1			
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt 类别1			
	基于思辨能力培养的中学语文学语文批判性思维教学案例).pdf.bxt	类别1		

DBSCAN 算法:

文本预处理	文本聚类 计算出文件夹中的文本聚类类别	
 ● 文本转換 ▲ 文本转换 ▲ 文本方词 ↓ 法停用词 ◎ 特征词库 ▲ 文本向量化 机器学习 	こ。 点击打开文件夹 可直接上作TXT格式的文件夹 已选择数版文件:D\\360安全浏览器下载\TXT	
 □ 相似度计算 ◆ 文本分类 □ 文本聚类 ○ 、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、、	■ K-Meana算法 I DESCAN算法 文件分析完成	
Y 王题模型		5 导出表格
	文件	类別
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	类别0
	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	类别0
	会计教学中批判性思维教学法运用策略分析,pdf.txt	类别0
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	类别0
	基于思辨能力培养的中学语文学语文批判性思维教学案例》.pdf.txt	类别0
	首页 上一页 1 2 3 4 下一页 尾页	

8.2.4 主题模型

功能简介:根据上传的文档,完成分词、去停用词等预处理操作,根据文本预处理的结果,利用 TextBank 算 法挖掘文本中主题词,返回每个文档中前是个主题词。

RESSET 文本分析客户端-小锐助手





文本预处理	主题模型 计算出文件夹中的文本LDA主题	
 文本转换 文本分词 文本分词 去停用词 特征词库 文本向量化 	こ 点击打开 可自接上传TXT#) 文件夹 (红的文件夹
机器学习 	已选择数据文件:D\360安全浏览器下载\TXT	
 田似度计算		已分析: 18/18 文件分析完成
		■ 日本
	文件	主題词
	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	护理 教学模式 批判性 教学 思维能力 思维 学生 课堂 实习生 能力
	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	批判性 教学 语文 思维 案例 阅读教学 学生 教育工作者 课堂 语言
	会计教学中批判性思维教学法运用策略分析,pdf.txt	批判性 思维 学生 教学 问题 教师 基础 内容 教学法 过程
	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	
	基于思辨能力培养的中学语文学语文批判性思维教学案例》.pdf.txt	教学 学生 语文 案例 能力 素养 批判性 学校 教师 学科
	首页 上一页 1 2 3	3 4 下一页 尾页

A1 \checkmark f_x χ f_x χ		
		2
1		L 子師海
1		
2	PBL联合CBS教学模式对实习生批判性思维能力的影响.pdf.txt	护理 教学模式 批判性 教学 思维能力 思维 学生 课室 实习生 能力
3	中学语文教学培养学生思辨能学语文批判性思维教学案例》.pdf.txt	批判性 教学 语文 思维 案例 阅读教学 学生 教育工作者 课堂 语言
4	会计教学中批判性思维教学法运用策略分析.pdf.txt	批判性 思维 学生 教学 问题 教师 基础 内容 教学法 过程
5	初中物理科学探究中培养学生科版8年级"浮力"教学为例.pdf.txt	
6	基于思辨能力培养的中学语文学语文批判性思维教学案例》.pdf.txt	教学 学生 语文 案例 能力 素养 批判性 学校 教师 学科
7	审辩式思维在创造力发展中的重要性.pdf.txt	
8	审辩式思维培养水平的评估指标研究.pdf.txt	学生 思维 教师 问题 水平 论证 教学 理由 结论 观点
9	审辩式思维教育具有的现实意义.pdf.txt	
10	审辩式思维能力及其测量.pdf.txt	思维 论证 评价 思维能力 参议员 批判性 药剂师 政府 医生 民主
11	审辩式思维观照下的初中文言文教学.pdf.txt	文言文 教学 思维 学生 观照 文本 教师 文化 流程 目标
12	循本源抓关键 提升审辩式思维能力.pdf.txt	
13	思维的批判性与大学的批判性思维教育.pdf.txt	思维 批判性 品质 逻辑 思想 理性 大学 前提 领域 准则
14	批判性思维教学:意蕴、特征及建构路径.pdf.txt	思维 批判性 学生 教学 教师 评价 建构 特征 师生关系 品质
15	批判性思维教育的一个议题:究? 罗素、杜威和图尔敏.pdf.txt	论证 情境 真理 批判性 观念 证据 思维 问题 理论 船主
16	本科生批判性思维能力及高影双一流"建设高校的实证研究.pdf.txt	批判性 思维能力 学生 思维 影响力 课程 基线 科研 高校 能力
17	浅析审辩式思维的培养策略.pdf.txt	学生 思维 教学 课堂 灯泡 小组 科学 小学 课程标准 教师
18	美国斯坦福大学批判性思维教育及启示.pdf.txt	批判性 思维 课程 高校 学生 思维能力 主题 内容 名声 通识
19	认知偏差与批判性思维教育.pdf.txt	
20		
21		
22		
23		