基于《个人信息保护法》的 App 隐私政策合规性检测

孙雯倩，徐天辰，余佩厚，陈云芳，张伟*

（南京邮电大学计算机学院，江苏 南京 210023）

摘　要： 数据隐私保护已成为社会关注的焦点，各国和地区正在陆续制定相关的法律法规，但是由于 App 产品发布的隐私政策存在篇幅长、专业性强等问题，利用自动化手段检测隐私政策的合规性成为亟待解决的技术难题。作为主流解决方法的机器学习模型需要标签注释的数据集进行支撑，而国内目前缺少这样的 App 隐私政策数据集。在分析欧盟《通用数据保护条例》（GDPR）合规性分析相关工作的基础上，设计适合我国《个人信息保护法》的标签方案，具体包括 15 个要求标签，然后使用网络爬虫获取 10 个类别、363 个 App 的中文隐私政策，并对这些隐私政策进行语句级划分和标注，构建包括 104 134 个隐私政策语句及标签组成的中文隐私政策语料库。采用百度最新开源的预训练语言模型 ERNIE 对语料库进行训练与测试，实验结果表明，该方案检测准确率达到 85.75%。

关键词： 隐私政策；《个人信息保护法》；合规性分析；语料库；自然语言处理

中图分类号： TP391　　**文献标志码：** A　　**DOI：** 10.19678/j.issn.1000-3428.0069804

Compliance Detection of App Privacy Policies Based on Personal Information Protection Law

SUN Wenqian, XU Tianchen, YU Peihou, CHEN Yunfang, ZHANG Wei*

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, China)

【Abstract】 Data privacy protection has become the focus of social attention, and countries and regions are gradually formulating relevant laws and regulations in this regard. However, because of the long and professional privacy policies released by App products, the use of automated methods to detect compliance with privacy policies has become an urgent technical challenge. Machine learning models, the widely popular solutions for this challenge, require labeled annotated datasets for support; however, a lack of such App privacy policy datasets currently exists in China. Based on the EU General Data Protection Regulation (GDPR) compliance analysis, a labeling scheme suitable for China's Personal Information Protection Law is designed, which includes 15 required labels. Subsequently, Chinese privacy policies for 363 Apps in 10 categories are obtained using Web crawlers, and these privacy policies are classified and annotated at the sentence level. A Chinese privacy policy corpus consisting of 104 134 privacy policy statements and labels is constructed. The corpus is trained and tested using the latest open-source pretraining language model from Baidu, ERNIE, with a detection accuracy of 85.75%.

【Key words】 privacy policy; Personal Information Protection Law; compliance analysis; corpus; Natural Language Processing (NLP)

0　引言

数据共享改变了信息孤岛现象，促进了信息社会的发展，同时也带来了信息泄露、数据滥用的风险。为了应对数字时代个人数据隐私面临的挑战，并为个人提供更多的隐私权和控制权，各个国家和地区相继发布了个人信息保护法案。2018 年 5 月 25 日，欧盟正式实施了个人信息保护法规，即《通用数据保护条例》（GDPR）。GDPR 的实施激发了全球范围内加强个人数据保护法规的趋势，2020 年，加利福尼亚州实施了美国历史上最严格的个人信息隐私法规之一——《加州消费者隐私法案》。我国广泛吸取各国经验，结合本国国情，2021 年 11 月 1 日正式实施《中华人民共和国个人信息保护法》（以下简称《个保法》），其规定了如何收集、处理、存储和保护个人信息，以及在哪些情况下需要用户的明确同意等方面的要求。随着《个保法》等法规的制定并生效，应用程序和服务提供商必须遵守这些法规。

然而，相关应用软件的隐私政策违反《个保法》

收稿日期： 2024-04-29　　**修回日期：** 2024-06-25

基金项目： 国家自然科学基金（62202406）。

通信作者 E-mail： * zhangw@njupt.edu.cn

的例子屡见不鲜。2019 年,一款名为“ZAO”的手机 App 的 AI 换脸功能吸引了众多用户,然而“ZAO”的隐私政策相关条款引发了网友的争议,其中“不可撤销、永久”等表述严重侵犯了《个保法》赋予的用户对个人数据的修正权、删除权、限制处理权以及撤回同意的权利。

对于用户来说,如果不熟知《个保法》等法规规定并逐条仔细阅读隐私政策,很难判断 App 到底有没有侵犯自己的各项权利,隐私政策成为一种形式化的告知。文献[1]指出隐私政策的形式化告知导致用户做出的同意并非是自己的真实意图表示,做出的同意决定并非知情且自愿,这种缺乏实际意义的告知同意只会给个人信息处理者披上合法的外衣,无法有效保护用户的个人信息权益。另一方面,对于 App 隐私政策制定方,有时候隐私政策可能会遗漏法规规定的部分条款和权利的声明,即使这不是出于其本意,但导致的隐私政策内容不完整,被发现后也势必受到各方质疑。因此,利用自动化的手段来检测 App 隐私政策与个人信息保护法规之间的合规性问题是非常必要的。

欧盟 GDPR 发布时间早、影响范围广,针对欧盟 GDPR 的合规性自动化检测得到了广泛的研究,但是各个国家的法规在内容和执行方面存在差异,处理个人信息需要了解特定国家或地区的法规要求,还要考虑文化背景和语言差异,数据保护法规的重点也不尽相同。例如,在数据主体的知情权方面,欧盟 GDPR 对于信息处理者设置了严格的义务要求:必须在处理个人信息前告知数据主体其享有的诸多权利;而我国的立法模式采取的是数据主体“主动型”,除非数据主体主动申请要求获悉个人信息的处理情况,数据处理者可以免除该项义务,这两种立法模式差异的出发点是基于不同国情的考虑。因此,不能直接套用关于 GDPR 合规性的自动化研究,需要设计出适合我国个人信息保护法案的合规性自动化检测方案。

纵观国内研究现状,已经有一些对隐私政策合规性的相关理论和实践探索,但是现有的研究方法主要是内容分析或案例分析,采用抽样调查的方式针对有限数量的 App 隐私政策进行,在研究结果的精准度、时效性等方面存在局限性。随着自然语言处理(NLP)技术、深度学习模型的广泛应用,利用自动化手段进行自动分类等方法逐渐引入,但 App 隐私政策合规性自动化检测面临下面的关键问题:

Q1:如何汲取 GDPR 相关方案经验,设计一套针对我国法规的标签机制?

Q2:如何设计隐私政策的遴选规则和爬取范围,使语料库具有更优代表性?

Q3:如何进行文本划分和标注对象的粒度设置,使得分类模型表现更佳?

Q4:如何设置合规性检测评价指标,检测实际运行效果?

对于我国的个人信息保护相关法案,2016 年颁布《中华人民共和国网络安全法》、2021 年先后颁布《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法》。其中,《个保法》是我国第一部专门规定个人信息保护的法律,明确规定了个人信息所有者的权益,包括知情权、访问权、更正权、删除权等;明确了个人信息处理者的责任,要求其采取合适的措施来保护个人信息的安全,包括建立数据保护管理制度、进行风险评估、采取技术措施等,明确了企业和组织在隐私政策方面承担的责任。《个保法》提供了一套全面的法律框架,可以有效监管隐私政策,保护个人信息的安全和隐私权益,确保企业和组织合规运营,也是最新的法律法规。因此本文选择将《个保法》作为本文法规的研究对象。

本文提出了针对隐私政策句子级别的多标签分类方法和合规性检测方法,具体提出了一套针对《个保法》的标签方案,合理选取隐私政策类别,建立了包含 104 134 个隐私政策句子的语料库,通过多标签文本分类模型对语料库进行训练和测试,并提出有效的合规性分析提取规则,以此来检测 App 隐私政策是否符合《个保法》的标签要求。本文工作主要贡献如下:

1)以 App 隐私政策自动化合规分析的角度,提出了《个保法》的 15 个要求标签,并对标签进行分级处理,据此设计了适用于我国《个保法》的隐私政策合规性检测方法。

2)收集、筛选并构建包含 10 个类别、363 条隐私政策的语料库(包含 104 134 个隐私政策语句),根据标签方案人工标注了其中的约 20%,即 21 539 个隐私政策语句。

3)将标注后的 21 539 个隐私政策语句以 7∶3的比例划分为训练集和验证集,使用百度 ERNIE 模型进行多分类训练,模型达到了 85.75%的准确率,平均 F1 值为 82%。使用该分类器模型训练后的分类结果预测了语料库中剩余的 82 595 个隐私政策语句的标签情况,并按篇使用命题逻辑命中的方式进行合规性检测。最终发现 363 条隐私政策中,只有 14.8%的隐私政策完全遵守《个保法》的相关规定。

1 相关工作

1.1 隐私政策语料库的建立

建立高质量的隐私政策语料库是实施模型训练的前提，而语料库的标签设计和标注直接决定了模型的输出质量。文献[2]提出了一种自动分析隐私政策内容和报告违反 GDPR 第 13 条的方法，该条内容规定了数据控制者在收集个人数据时必须向数据主体提供的信息，和隐私政策具有强关联，以此制定了 10 项 GDPR 要求标签，并人工标注了一个包含 304 条隐私政策(36 610 个标注句子)的语料库。文献[3]针对 GDPR 提出了 18 项 GDPR 要求标签，其覆盖了文献[2]中所有的标签，除此之外增加考虑数据来源、数据保护、自动化决策等要求，并创建了一个专家标注覆盖 1 080 个网站隐私政策的数据集。文献[4]研究《通用数据处理协议》对 GDPR 法规的遵从性和合规性问题，提出 45 项 GDPR 要求标签，其中有 26 项强制性要求和 19 项可选要求，定义了其中法律概念的词汇表，利用 NLP 自动生成《通用数据处理协议》文本内容的短句级表示，通过对标这两种表述来判断合规性。文献[5]与法律专家合作制定了 10 个隐私政策标签，使用众包的方法创建了包含 115 条隐私政策的语料库(OPP-115)。文献[6]构建了名为 PrivaSeer 的自动生成语料库，其中包含了超过 100 万条英文网站的隐私政策。文献[7]采用识别物联网设备的隐私政策统一资源定位符(URL)技术，构建了涵盖 592 条隐私政策的语料库。文献[8]提出了一个名为 ATPChecker 的自动化系统来分析 Android 第三方库(TPLs)是否符合隐私相关法规，构建了一个数据集，其中包含 458 个 TPLs、247 条 TPLs 的隐私政策、187 个 TPLs 的二进制文件和 641 个主机应用程序及其隐私政策的列表。文献[9]从国内各种隐私相关法规中总结了 7 个组成部分作为标签，比如控制者、收集、共享等，从 483 条中文 Android 应用程序的隐私政策中标注了 11 565 个句子中所有组成部分的出现情况，构建了国内首个中文隐私政策数据集 CA4P-483。由此可见，语料库的建立和所针对的法规内容和标签方案息息相关，因此，本文将在提出 15 个要求标签的基础上构建适用于《个保法》的隐私政策语料库。

1.2 自动化分类模型和合规性检测方法

大部分隐私政策的合规性检测研究针对 GDPR。文献[10]构建了一个机器学习和规则相结合的系统来检测隐私政策中是否存在 GDPR 要求的信息，这里的规则运用 OPP-115 的标签方法。然而，该研究仅在 30 条隐私政策上对模型进行了评估。文献[11-12]建立基于深度学习的隐私政策分类模型，使用 OPP-115 和其他语料库进行训练，这些语料库进一步被用作提供具有问答功能的隐私策略摘要的服务的解决方案，称为 Polisis-Pribot。文献[3]构建了一个基于卷积神经网络的分类器来判断隐私政策是否完全符合 GDPR，利用主动学习方法提高模型的准确性，解决不同类别的隐私政策片段之间因重叠性特征导致的模型误分类问题，实现了 89.2%的准确率。该研究分析了满足 GDPR 隐私政策要求标签的网站数量，发现只有 32%的网站完全符合 18 项 GDPR 要求标签。文本分类一直是 NLP 领域中一个被广泛研究的问题[13]。经典的 NLP 技术侧重于从文本中提取特征和训练模型，如逻辑回归或支持向量机(SVM)。随着深度学习的推进，NLP 领域的前期工作集中于使用词向量进行文本分类。文献[4]利用 NLP 技术自动检查符合性，基于 BERT(Bidirectional Encoder Representations from Transformers)语言模型，利用 NLP 工具为句子中的每个动词生成一个“谓词-变元结构”，用生成器为隐私政策中的每个语句生成多个基于语义框架的表示，然后通过检查隐私政策文本和 GDPR 文本跨度之间的相似度等条件判断合规性。文献[14]根据 GDPR 定义了 33 种元数据类型，使用基于人工规则和机器学习的方法对隐私政策进行歧义检测。文献[15]在其基础上更进一步，提出了 55 种元数据类型，并借助机器学习和 NLP 从基金领域的 234 条隐私政策中自动提取元数据以检查合规性，其局限性在于隐私政策只来自特定领域，并且评估数据集数量有限。文献[16]使用 GDPR 知识图谱和分类器来改进隐私政策合规性检测系统。文献[17]使用统一建模语言(UML)和对象约束语言(OCL)来构建 GDPR 的 UML 表示。文献[18]在数据供应链中实现了基于文档的合规性检查，并开发了自动检测隐私政策是否符合 GDPR 的具体方法。文献[19]以搭建的 GDPR 知识图谱为依据，检测隐私政策是否缺少部分 GDPR 要求披露的概念。文献[20]构建了包括所有欧洲市场的有关技能的隐私政策数据集，使用基于 ChatGPT 的动态测试工具来检查该隐私政策是否符合 GDPR。

国内隐私政策的合规性研究工作中，文献[21]采用机器学习集成方法对我国医疗健康 App 隐私政策的合规性进行测评，依据国家相关政策法规构建医疗健康 App 隐私政策合规性测评指标体系，基于硬投

票分类器，综合应用卷积神经网络、循环神经网络、长短期记忆人工神经网络 3 种机器学习算法建立合规性检测模型。文献[22]根据《个保法》创建了一套的隐私政策标签方法，包含 8 个一级标签及其分别对应的 30 个二级标签，并设计了隐私政策合规性得分：满足一个标签得 1 分，否则得 0 分，通过满足的标签数量计算合规性得分，以此判断隐私政策的合规性，但其模型的训练是利用传统的机器学习模型。利用更高效的深度模型自动化检测 App 隐私政策与隐私保护法律法规的合规性问题已经成为主要趋势。

2 方法

2.1 方法概述

本文提出了一种自动检测隐私政策内容是否违反《个保法》的方法。该方法包括 3 个过程：建立语料库、训练多标签分类器和分析隐私政策的合规性，整体流程如图 1 所示。

Step1 语料库的建立。首先是标签设计，通过深入的法律研究，参考欧盟的 GDPR 以及相关法律法规，经过多次的团队讨论和筛选，针对《个保法》提取出了 15 个关键的隐私政策合规要求标签，这些标签涵盖了数据收集、存储、共享、访问和用户权利等方面的关键要求；其次是隐私政策样本的获取，为了建立一个有代表性的数据集，对安卓 App 应用市场的排名进行了综合分析，识别出了最常见的 App 应用类别，使用网络爬虫技术，采集了数百个不同应用的隐私政策样本，将采集到的隐私政策样本按照语句进行划分，创建了一个丰富的语料库，以便后续的分析和评估；最后是数据标注，为了进行合规性分析，借助 doccano 工具，培训了一组专业人工注释员，以多标签的方式标注了隐私政策中的句子，将其与提前定义的合规要求标签进行关联。

Step2 多标签分类模型。运用百度 ERNIE 的多标签分类大模型对语料库进行分类训练，使用标注的数据对提出的方法进行了评估，以验证其在隐私政策合规性分析中的有效性；并使用了一系列性能指标来评估模型的性能，包括准确率、召回率和 F1 值等。

Step3 合规性分析。将 15 个标签要求划分成 4 类，运用多标签分类模型的分类结果判断每类是否满足条件，通过命题逻辑公式分析隐私政策的合规性。

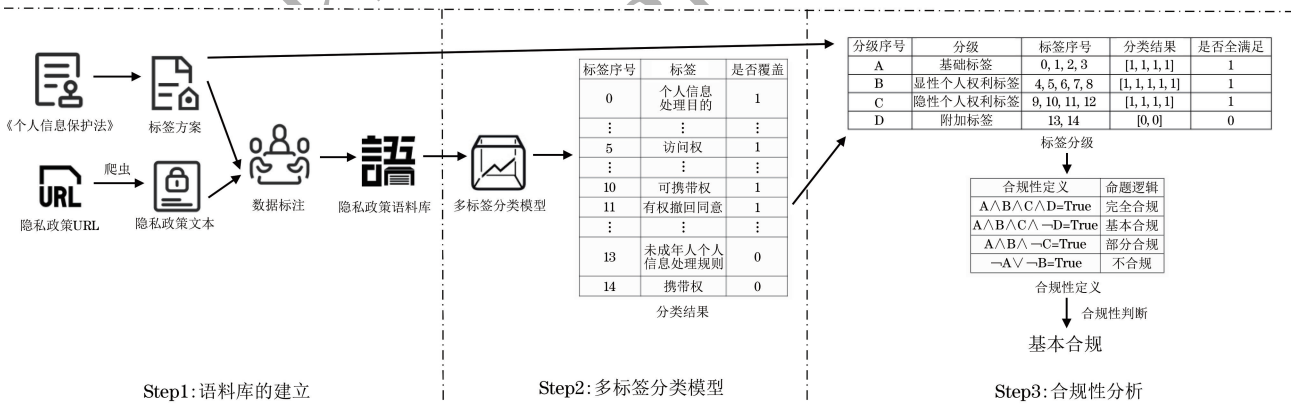

图 1 方法流程

Fig. 1 Method process

2.2 语料库的建立

2.2.1 标签方案

在本节中，提出了一套针对我国法规《个保法》的标签机制来回答 Q1 问题。欧盟颁布的 GDPR 影响范围广，是全球数据保护和隐私保护的标杆法规。参考上文提出的几种 GDPR 标签方案，并将其与《个保法》条款进行分析比对，寻找其共性和个性。笔者发现，GDPR 的标签方案中共同强调的要求在《个保法》中都可以一一对应，区别在于有的以粗粒度的形式提出要求，而有的则是细粒度形式。先以细粒度的形式提出《个保法》的标签方案，并尝试将标签方案应用到隐私政策文档中，人工阅读选取的隐私政策，进行手动贴标签。

检查对应标签发现，粒度过细会导致标签重复性和不合理性。例如，“有权悉知数据如何被收集”和“有权悉知数据如何被处理”通常在相似段落同时出现，因此合并成一个标签；再如，“有权就其损失获得赔偿”和“个人信息保护负责人的联系方式”两个标签几乎完全没有展现出来。分析其原因，这里的“损失”涉及太多方方面面，难以在隐私政策中体现出来，而个人信息保护负责人一般是该公司的负责人，这与个人信息处理者联系方式是一样的；还有，“有权向监管机构发起申诉”和“有权对控制者或处理者提起诉讼”在隐私政策中通常是出现在同一句子中，因

此在该研究项目中可将其合并成一个标签;“近亲可对死者信息行使权利”在隐私政策中也极少提到,涉及相关法律法规,容易引起冲突,对此本文不予涉及。

除此之外,未成年人个人信息越来越受到重视和保护。《个保法》第三十一条提出,“个人信息处理者处理不满十四周岁未成年人个人信息的,应当制定专门的个人信息处理规则”,即要求个人信息处理者必须采取额外的保护措施来保护未成年人的个人信息,隐私政策必须明确规定如何处理未成年人的信息以符合法律要求。因此,将“未成年人个人信息处理规则”加入标签。

通过多轮贴标签实践,最终确定了 15 个《个保法》标签,并对标签名称进行优化,使其更贴近中文阐述理解的习惯。标签按照重要性排序,详细说明如下所示(每个标签后的数字表示其在《个保法》中的具体条款编号):

1)个人信息处理目的。个人信息处理者处理个人信息的目的。[17.2]

2)收集个人信息的种类。收集的个人信息的类别。[17.2]

3)存储期限。个人信息处理者处理个人数据之前,应当公布个人信息保存期限,个人信息的保存期限应当为实现处理目的所必要的最短时间。[17.2]

4)个人信息处理者的联系方式。处理个人信息前,个人信息处理者应当提供名称或者姓名和联系方式。[17]

5)访问权。个人有权向个人信息处理者查阅个人信息。[45]

6)修正权。个人有权请求更正、补充个人信息,个人信息处理者核实后应予以更正、补充。[46]

7)删除权。遇到几种情况,个人信息处理者应当删除信息,若未删除,个人有权请求删除。[47]

8)限制处理权。个人有权限制他人对其个人信息进行处理。[44]

9)反对权。个人有权拒绝他人对其个人信息进行处理。[44]

10)有权撤回同意。基于个人同意处理个人信息的,个人有权撤回其同意。[15]

11)有权悉知数据如何被处理。知情权内容范畴,个人有权了解个人信息会被如何应用、处理目的、是否会被分享给第三方。[44]

12)有权免受自动化决策的限制。个人有权拒绝个人信息处理者仅通过自动化决策的方式作出决定。[24]

13)有权向监管机构发起申诉。任何组织、个人有权针对违法个人信息处理活动向履行个人信息保护职责的部门进行投诉、举报。[65]

14)未成年人个人信息处理规则。个人信息处理者处理不满十四周岁未成年人个人信息的,应当制定专门的个人信息处理规则。[31]

15)可携带权。个人有权要求个人信息处理者提供其个人信息副本。个人如若合法要求转移个人信息至其他个人信息处理者,原处理者应提供转移路径。[45]

2.2.2 隐私政策收集

为了回答 Q2 问题,需要收集一组高质量的、具有不同应用类别的隐私政策。App 隐私政策在不同的时间段会更新不同的版本,目前没有最新且现成的隐私政策数据集公开可用,人工收集会花费大量的时间,因此需要通过爬虫等技术手段,创建隐私政策语料库。

使用以下的步骤收集隐私政策:1)通过应用市场排名来选择应用程序及其隐私政策;2)隐私政策应该来自不同的类别,不同的类别可能对访问用户信息有不同的要求。收集的隐私政策涵盖了浏览器、视频影音、学习教育、娱乐、天气预报、新闻、购物、医疗健康、阅读、出行 10 个应用类别,每个类别的隐私政策数量占比几乎相同。在安卓应用市场浏览并统计涵盖范围最广的应用类别,分析并总结出以上 10 个类别。

通过下载量和排名来收集各类隐私政策的 URL,涵盖了下载量较高、中等、较低的隐私政策。根据下载量分级收集隐私政策是一种逐渐精细化管理和规范隐私政策的方法,其目的是更好地满足不同规模和影响力的应用程序的需求,因为无论应用程序的规模如何,都应该遵循适用的隐私法规和最佳实践,以保护用户的隐私权。以《个保法》生效时间为起止日期,收集 2021 年 11 月 1 日之后更新的隐私政策,共收集了 524 个隐私政策的 URL,并记录了该版本更新的最新日期和收集日期。

通过爬虫程序爬取隐私政策文本后发现,由于爬虫在爬取网页内容时会遇到不同的技术和文件类型,不同的网站可能使用不同的技术和框架来构建其页面。大部分网站使用传统的超文本标记语言(HTML)来呈现内容,而有一些可能依赖 JavaScript 来动态生成页面元素,有一些利用 JSP(Java Server Pages)服务器端技术生成动态网页内容。为了保证收集到的隐私政策文档的质量,需要对爬取到的隐私政策进行筛选,筛选标准如下:1)保证隐私政策完整性;2)控制隐私政策文档大小,具体

设定 40 个句子长度为隐私政策文档大小的下界；3)去除重复的隐私政策。

本文共爬取了 521 条隐私政策，经过以上的筛选标准筛选后，保留了 363 条有效的隐私政策，其中包含了 104 134 个句子。每条隐私政策的平均句子数量为 287 个句子，其中 48.2%的隐私政策长度在 200～300 个句子之间，最短的隐私政策包含 40 个句子。

2.2.3 数据标注

针对 Q3 问题，隐私政策文本的标注采用划分句子贴标签的方法来进行，而不是更为常见的文本分段标注。首先，考虑精确性和细粒度，将标签分配给句子而不是段落可以更加精确地标识隐私政策中的具体内容和规定，有助于了解隐私政策的各个方面是否符合法规要求。其次，在隐私政策中有些标签内容只在句子中出现，甚至一个句子中涵盖几个标签内容，而分段难以标注处理，直接会影响分类器结果。合规性分析通常需要对具体的法规规定和具体的隐私政策内容进行匹配，如果将标签分配到语句级，可以更容易地将法规要求与政策中的相应语句进行比较。尽管对语句级进行标注可能会比对段落级标注工作量更大，但这种方法能够更精确地确定哪些部分需要进一步修改以满足法规要求。

本项目招募了具有法律和计算机科学专业的本科生和研究生背景的 9 位志愿者，以便对隐私政策进行标注。为了确保标注的准确性和一致性，首先为志愿者提供标注任务的培训。在培训过程中，提供了一个简明的教程，并提供了带有标签的示例句子，以清晰阐明每个标签的含义。完成培训后，要求志愿者标注一组隐私政策文本，并仔细审查他们的标注结果，以便及时解决任何可能存在的误解。通过这一流程，确保每位志愿者都具备对标签含义的清晰理解，从而保证了标注质量的一致性。每个句子由 3 名志愿者独立进行标注，以进一步提高标注结果的可靠性。

每个志愿者被指派标注一组隐私政策，各自独立完成标注任务，平均需要 30 min。在所有志愿者完成各自任务后，要求 3 名志愿者对相同的隐私政策进行标注合并。按照标准程序，如果 3 名志愿者一致标记相同的标签，那么该标签被视为该句子的最终标记。如果存在差异，他们将进行讨论，直到达成共识。

最终从 363 条隐私政策中详细标注了 100 条，每个类别选择了 10 条隐私政策。根据预定的标签方案，人工标注了 21 539 个隐私政策语句，这些标注语句约占整个语料库的 20%。

2.3 多标签分类模型

分类工作使用在中文领域内模型效果和模型计算效率突出的 ERNIE 轻量级系列模型作为训练基座。ERNIE 是百度提出的一种预训练语言模型，基于深度学习框架 PaddlePaddle 实现，拥有 2 600 亿个参数，超过了 OpenAI 的 GPT-3 的 1 750 亿。ERNIE 旨在对各种不同类型的知识(如语法知识、语义知识和实体知识等)进行建模，从而提高 NLP 任务的性能[23]。然而，ERNIE 引入了一种全新的预训练任务设计，对各种类型的知识进行建模，从而改善 NLP 任务的性能。ERNIE 在多个 NLP 任务(如阅读理解、命名实体识别和情感分类等)上取得了显著的性能提升[24]。ERNIE 3.0 轻量级系列提供多种尺寸的预训练模型满足不同需求，具有广泛成熟的实践应用性。

2.3.1 文本多标签分类

文本多标签分类是 NLP 中常见的文本分类任务，文本标签分类在各种现实场景中具有广泛的适用性，例如商品分类、网页标签、新闻标注、语义场景分类等。多标签数据集中样本用来自 n_{classes} 个可能类别的 m 个标签类别标记，其中 m 的取值在 0～n_{classes} 之间，这些类别具有不相互排斥的属性。多标签数据集的标签集含有两个或两个以上的类别，输入句子/文本具有一个或多个标签，如图 2 所示，即一个多标签分类和处理例子，该隐私政策文本具有“访问权”“修正权”和“删除权”这 3 个标签。

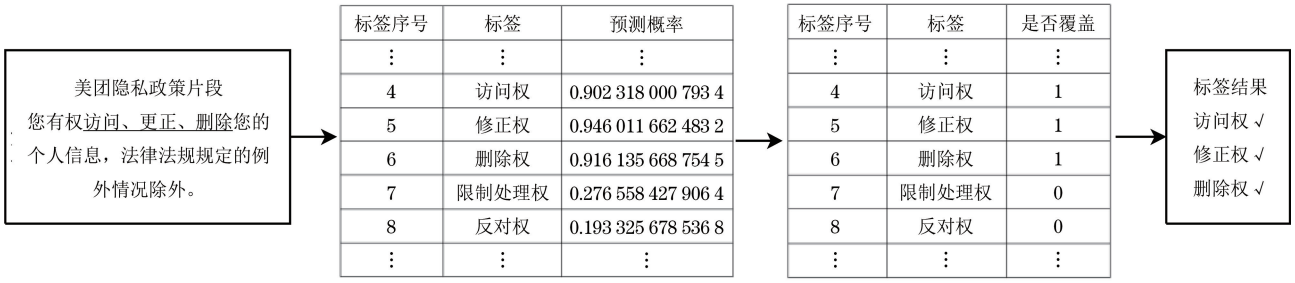

图 2 多标签示例

Fig. 2 Example of multiple labels

2.3.2 ERNIE 模型结构

ERNIE 模型的核心算法与 BERT 模型相似，它基于 Transformer 结构并采用 Masked Language Model 作为预训练模型。ERNIE 3.0 首次在百亿级预训练模型中引入大规模知识图谱，提出了海量无监督文本与大规模知识图谱的平行预训练方法(Universal Knowledge-Text Prediction)，通过将知识图谱挖掘算法得到的 5 000 万个知识图谱三元组与 4 TB 大规模语料同时输入到预训练模型中进行联合掩码训练，促进了结构化知识和无结构文本之间的信息共享，大幅提升了模型对于知识的记忆和推理能力。

图 3 为 ERNIR 模型结构，其输入为以下 4 个部分：

1) Token embedding：embedding 是将大型稀疏向量转换为保留语义关系的低维空间，这一部分为把源代码经过词法分析过后生成的 token 序列进行 embedding，即词向量本身的 embedding。

2) Sentence embedding：句子类型的 embedding。

3) Position embedding：位置信息的 embedding。

4) Task embedding：embedding 建模不同的任务。

将 4 个部分的 embedding 连接，最终的结果作为 Transformer 的输入，训练不同的子任务。模型输入(Input)为 segment embedding、token embedding、position embedding 与 task embedding 之和。

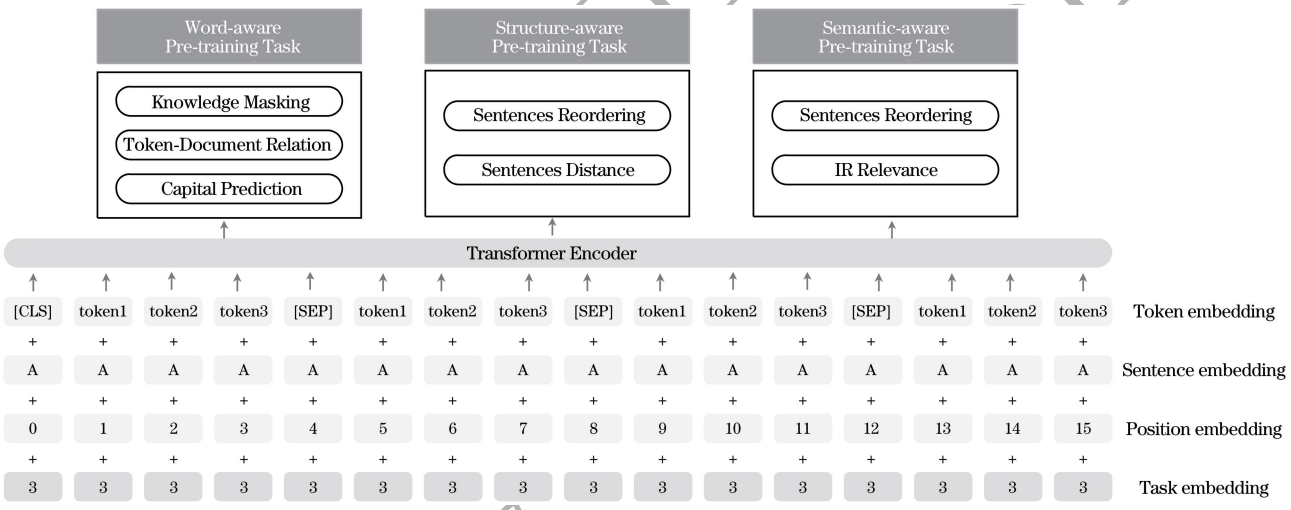

图 3 ERNIE 模型结构

Fig. 3 ERNIE model structure

2.3.3 ERNIE 模型算法

ERNIE 在预训练阶段引入了新的任务和知识，以便更好地捕获多粒度知识和结构化知识。以下是 ERNIE 模型的目标函数和损失函数。

假设输入序列为 X，经过 Mask 操作后的序列为 $\widetilde{X}$，实体识别和语义关系挖掘任务的标签为 Y_e 和 Y_r，那么 ERNIE 的目标函数可以表示为：

$$L(\theta)=L_{\mathrm{MLM}}(X,\widetilde{X};\theta)+\alpha L_{\mathrm{entity}}(X,Y_e;\theta)+\beta L_{\mathrm{relation}}(X,Y_r;\theta) \tag{1}$$

式中：L_{MLM} 是 Masked Language Model 的损失函数，用于建模句子中的词汇知识；L_{entity} 和 L_{relation} 分别是实体识别和语义关系挖掘任务的损失函数，用于捕获句子的结构化知识；α 和 β 是超参数，控制着不同任务损失函数的权重。

Masked Language Model 损失函数 L_{MLM} 表示为：

$$L_{\mathrm{MLM}}(X,\widetilde{X};\theta)=-\sum_{t=1}^{T} l_{X_t\neq\widetilde{X}_t}\ln P(X_t\mid\widetilde{X}_{-t};\theta) \tag{2}$$

式中：$\widetilde{X}_{-t}$ 是除第 t 个位置的其他所有位置的元素；l 是指示函数；T 是求和的序列长度。

实体识别损失函数 L_{entity} 表示为：

$$L_{\mathrm{entity}}(X,Y_e;\theta)=\frac{1}{T}\sum_{t=1}^{T}\ln P(y_{e,t}\mid X;\theta) \tag{3}$$

语义关系挖掘损失函数 L_{relation} 表示为：

$$L_{\mathrm{relation}}(X,Y_r;\theta)=\sum_{(i,j)\in P}\ln P(Y_{r,ij}\mid X;\theta) \tag{4}$$

ERNIE 模型的关键算法包括多粒度知识建模、连续语义融入策略和结构化知识融合，这些算法在预训练阶段通过不同的任务损失函数相互作用，从而提高模型在 NLP 任务上的性能。

2.4 合规性分析

在合规性检测方面，GDPR 对合规性要求有明确规定，因此研究文献通常要求隐私政策完全符合所提出的标签才能判定其合规。然而，对于《个保法》提出的标签，其重要性等级有所不同。例如，“未成年人个人信息处理规则”被视为非常重要，大多数 App 会为此制定额外的规则，有时甚至单独处理。

因此，不能简单地将其纳入一般合规性标签进行统一评判。此外，像“有权向监管机构发起申诉”这样的标签，通常体现为自发性行为，并且经常嵌入在语境不明显的语句中，难以单独识别和评估。

国内对隐私政策有诸多法规约束，本文研究主要关注《个保法》对隐私政策的限制。如果简单地要求隐私政策满足全部标签才能判定合规，而不考虑标签的重要性和缺失情况，这种方法在实际操作中并不适用。因此，需要将合规性检测细化为几个等级，通过分级判定隐私政策的合规程度。这种分级评估方法不仅更符合实际情况，也能更有效地指导隐私政策的改进和完善。

GDPR 在定义个人数据权利内容时采取了 2 种方式：一种是明确定义某些个人数据权利，如访问权、修正权等，这些权利可称之为显性个人数据权利；另一种是对某些个人数据权利进行简要描述而不定义其名称，这些权利可称之为隐性个人数据权利[25]。鉴于此，我国个人信息权利也可分为显性个人信息权利与隐性个人信息权利。将上述 15 个标签分为 4 级，为基础标签（隐私政策需提供的基本信息）、显性个人信息权利标签、隐性个人信息权利标签和附加标签，其中附加标签中“可携带权”重要性程度最低，考虑“未成年人个人信息处理规则”需额外制定专门规则，将其作为附加标签，如表 1 所示。

表 1 标签分级对应表

Table 1 Label classification corresponding table

序号	分级	标签
A	基础标签	个人信息处理目的、收集个人信息的种类、存储期限、个人信息处理者的联系方式
B	显性个人权利标签	访问权、修正权、删除权、限制处理权、反对权
C	隐性个人权利标签	有权撤回同意、有权悉知数据如何被处理、有权免受自动化决策的限制、有权向监管机构发起申诉
D	附加标签	未成年人个人信息处理规则、可携带权

通过将标签分为基础标签、显性个人权利标签、隐性个人权利标签和附加标签，建立了一个系统化的合规性检测框架。在这种分层次的标签分类方法中：基础标签涵盖最基本的合规性要求，确保隐私政策符合最重要的法律规定；显性个人权利标签突出用户在隐私保护方面的显性权利，确保用户能够明确知晓并行使这些权利；隐性个人权利标签涵盖一些隐性的、更细微的隐私权利保护，进一步细化合规性要求；附加标签包括一些非强制性但仍然重要的隐私保护措施，鼓励隐私政策的更高标准。这种分类方法确保了合规性检测的全面性和细致性。

若一条隐私政策完全符合《个保法》，则应涵盖表 1 所示的 A、B、C、D 4 类标签。D 类标签中因法规提出年份比较近，很多 App 还没有设立专门的未成年人个人信息处理规则，只是在隐私政策中提到了未成年的相关规定；除此之外，“携带权”影响程度也较低，因此若隐私政策涵盖 A、B、C 3 类，则认为该隐私政策基本符合《个保法》；在 B 类和 C 类中，隐性个人权利标签在权利描述中并不能明显体现，因此，若隐私政策涵盖 A、B 2 类，则称其部分符合《个保法》；而 A、B 类若缺少或没有，则认为该隐私政策不符合《个保法》的完整性表述。

据此，定义的命题逻辑性质如下：

$$Privacy\_Policy \rightarrow Fully\_Compliant = A \wedge B \wedge C \wedge D;$$

$$Privacy\_Policy \rightarrow Basic\_Compliant = A \wedge B \wedge C \wedge \neg D;$$

$$Privacy\_Policy \rightarrow Partial\_Compliant = A \wedge B \wedge \neg C;$$

$$Privacy\_Policy \rightarrow Not\_Compliant = \neg A \vee \neg B。$$

整体的合规性分析过程为：当一条隐私政策放入模型中输出预测结果时，会输出每个句子的标签预测结果，统计标签预测结果，1 表示有该标签，0 表示没有该标签。若表示 A 类的标签 label[0]、label[1]、label[2]、label[3]都为 1，则认为该隐私政策满足 A 类，A 类为真；若表示 B 类的标签 label[4]、label[5]、label[6]、label[7]、label[8]都为 1，则认为该隐私政策满足 B 类，B 类为真；若表示 C 类的标签 label[9]、label[10]、label[11]、label[12]都为 1，则认为该隐私政策满足 C 类，C 类为真；若表示 D 类的标签 label[13]、label[14]为 1，则认为该隐私政策满足 D 类，D 类为真。将合规性分析转换成考察具体的隐私政策是否存在 A、B、C、D 4 类标签，按照如上的命题逻辑性质进行判断和分析。

合规性分析伪代码如下所示：

算法 1 合规判断算法

输入 模型训练后标签结果 bool label[15]

输出 合规性结果

```
1. begin
2.   classA := True; classB := True;
```

```
3.  classC := True; classD := True;
4.  for i := 0 to 14 do
5.    case i of
6.    0..3: classA := classA and label[i];
7.    4..8: classB := classB and label[i];
8.    9..12: classC := classC and label[i];
9.    13..14: classD := classD and label[i];
10.  end;
11.  writeln(case classA and classB and classC and classD of
12.    True: 'Fully compliant!';
13.    False: case classA and classB and classC of
14.      True: 'Basic compliant!';
15.      False: case classA and classB of
16.        True: 'Partial compliance!';
17.        False: 'Not compliant!'
18.      end
19.    end
20.  )
21. end
```

3 实验与结果分析

3.1 环境搭建

采用最新开源的百度文心 ERNIE 预训练大模型进行训练和评估，使用阿里云提供的 GPU 训练，制定 GPU 卡号为 0。编程语言至少为 Python 3.6 版本，需要的开发库为：paddlepaddle>=2.3，paddlenlp>=2.4，scikit-learn>=1.0.2。PaddleNLP 采用 AutoModelForSequenceClassification，AutoTokenizer 提供了简单易用的接口，可以用过 from_pretrained()方法来加载不同的预训练模型，在输出层上叠加一层线性层，且相应预训练模型权重下载速度快、更稳定。对预训练模型进行微调，分词器 tokenizer 使用的最大序列长度默认为 128，批处理大小为 32，训练轮次为 100，并使用早停法(EarlyStopping)，即模型在开发集经过一定 Epoch 后精度表现不再上升，训练终止。选择在开发集上表现最好的参数作为最终模型参数。

3.2 分类结果

标注的样本数据(21 539 个隐私政策语句)按照 7∶3的比例划分为训练集和验证集，所有实验运行 6 次，报告中位数结果。采用的评价指标包括标准精确率(P)、召回率(R)和 F1 值(F1)。还用了微观 F1 值(Micro F1)和宏观 F1 值(Macro F1)作为评估分类模型性能的指标，作为精确率和召回率的综合度量，它们能够全面评估模型在分类任务中的综合表现，特别适合用于多模型对比评估。因此对比实验重点考察 Micro F1 和 Macro F1 这 2 个评价指标。

在 ERNIE 的多种版本模型中进行了数据集的训练，最终挑选出评分较高的 3 个版本模型，通过 Micro F1 值和 Macro F1 值作为评价指标，如表 2 所示。

本次实验还增加了其他模型做对比实验，选择传统机器学习模型(SVM)和深度学习模型[文本卷积神经网络(TextCNN)]进行对比实验，可以验证数据集的适用性和泛化能力，还可以评估模型在不同类型任务中的表现，以及它们在处理高维稀疏数据和复杂特征提取方面的优劣。将数据集在这 2 种模型上进行训练，所得的 Micro F1 值和 Macro F1 值如表 2 所示。

表 2 模型效果对比

Table 2 Model effects comparision %

模型名称	模型结构	Micro F1	Macro F1
ERNIE 1.0 Large Cw	24-layer, 1 024-hidden, 20-head	82.29	82.37
ERNIE 3.0 Base	12-layer, 768-hidden, 12-head	82.08	81.54
ERNIE 3.0 Medium	6-layer, 768-hidden, 12-head	81.81	81.23
SVM	—	76.57	70.36
TextCNN	—	60.67	56.44

在本次实验中，对比了不同模型在文本分类任务中的表现，包括 SVM、TextCNN 以及 3 种不同配置的 ERNIE 模型(ERNIE 1.0、ERNIE 3.0 Base、ERNIE 3.0 Medium)。ERNIE 1.0 取得了最高的 Micro F1 和 Macro F1 值，分别为 82.29%和 82.37%，这表明其复杂的模型结构(24 层、1 024 个隐藏单元、20 头注意力机制)在文本分类任务中表现出色。ERNIE 3.0 Base 也表现良好，Micro F1 和 Macro F1 值分别为 82.08%和 81.54%，虽然层数和隐藏单元数减少，但依然保持了较高的性能，显示出良好的泛化能力。ERNIE 3.0 Medium 的 Micro F1 和 Macro F1 值略低于 ERNIE 3.0 Base，但依然较高，分别为 81.81%和 81.23%。这表明即使在层数进一步减少的情况下，该模型依然能够有

效地进行文本分类任务。相比于深度学习模型，SVM 作为传统机器学习模型表现略显逊色，这可能是由于其无法自动提取复杂特征。TextCNN 表现了最差的 F1 值，可能是因为 TextCNN 较难捕捉文本中的复杂模式，尤其在数据规模较大或特征较复杂的情况下，其性能不如 ERNIE 大模型。

每个标签的模型分类结果如表 3 所示，选取 ERNIE 系列模型中训练结果较好的 3 类：ERNIE 1.0 Large Cw，ERNIE 3.0 Base，ERNIE 3.0 Medium。每个标签类别的最佳标准精确率、召回率和 F1 值用粗体突出显示。“AVG”一行表示 16 个标签的宏观平均值，这里明确报告了“Other”类别的分析结果，原因是“Other”类虽然与合规性分析任务无关，即不参与隐私政策和法规之间的标签对应过程，并且不属于判断一条隐私政策是否合规的标志，但它参与分类器的训练，并且“Other”类的句子数量占所有隐私政策句子的很大部分，影响着其他 15 个标签的分类结果，因此也将该类的分类结果呈现在表中。

表 3 分类模型的各项评估指标

Table 3 The evaluation indicators of the classification model %

标签	ERNIE 1.0 Large Cw			ERNIE 3.0 Base			ERNIE 3.0 Medium		
	P	R	F1	P	R	F1	P	R	F1
个人信息处理目的	67.10	83.52	74.41	69.94	76.66	73.14	69.76	79.18	74.17
收集个人信息的种类	90.10	92.18	91.13	89.49	90.64	90.06	86.50	95.26	90.67
存储期限	95.29	94.19	94.74	92.13	95.35	93.71	82.65	94.19	88.04
个人信息处理者联系方式	89.39	86.13	87.73	84.83	89.78	87.23	85.71	91.97	88.73
访问权	89.80	86.27	88.00	84.47	85.29	84.88	91.21	81.37	86.01
修正权	86.14	86.14	86.14	87.88	86.14	87.00	90.80	78.22	84.04
删除权	90.34	90.72	90.53	90.68	90.30	90.49	89.96	90.72	90.34
限制处理权	59.43	67.74	63.32	59.14	59.14	59.14	70.49	46.24	55.84
反对权	78.46	79.27	78.87	84.15	79.79	81.91	79.47	78.24	78.85
有权撤回同意	86.24	77.05	81.39	84.17	82.79	83.47	86.32	82.79	84.52
有权悉知数据如何被处理	80.38	73.36	76.71	78.26	78.60	78.43	82.98	68.12	74.82
有权免受自动化决策限制	74.55	87.23	80.39	72.92	74.47	73.68	83.72	76.60	80.00
有权向监管机构发起申诉	95.24	80.00	86.96	88.00	88.00	88.00	100.00	72.00	83.72
未成年人个人信息处理规则	100.00	69.23	81.82	90.00	69.23	78.26	100.00	69.23	81.82
可携带权	90.91	86.96	88.89	86.96	86.96	86.96	95.45	91.30	93.33
平均	84.36	81.22	82.37	82.55	80.88	81.54	85.75	78.20	81.23
其他	76.39	59.53	66.92	77.78	60.87	68.29	76.96	55.85	64.73

从结果中观察到，ERNIE 3.0 Medium 精确率最高，达到 85.75%，而 ERNIE 1.0 Large Cw 表现出最好的召回率和 F1 值，由于本文主要关注精确率这一指标，因此重点强调 ERNIE 3.0 Medium 的精确率。其中一些类别，比如“限制处理权”，《个保法》中定义“个人有权限制他人对其个人信息进行处理”，但是在隐私政策中的表述很模糊，不会出现“限制”等明显词语，而是隐晦地出现在“改变授权范围”“允许或禁止相关权限申请”等，所以比其他类别更难获取具体特征。除此之外，因为单个语句可以涵盖多个标签，并且多标签的数量没有限制，而“个人信息处理目的”和“收集个人信息的种类”时常出现在同一句话中，而有时又作为单独句子出现，给分类器增添了一定的困难。所以在错误分类结果中观察到，有的语句可能只有“收集个人信息的种类”，却被标注为既有“个人信息处理目的”又有“收集个人信息的种类”，或者缺少标注标签等情况。“有权悉知数据如何被处理”标签的定义涵盖范围较为广泛，需要根据上下文语义，语义特征也不是很明显，从而混淆了分类模型。

综上，本文基于 ERNIE 的多标签分类器可以将给定的隐私政策句子分类为《个保法》标签要求，并具有较高的准确性，平均达到了 85.75%的准确率、81.22 %的召回率和 82%的 F1 值。在最终模型中，误分类的主要原因是隐私政策句子描述可能过于模糊或可能用非常简短的描述覆盖多个需求。

3.3 合规性分析结果

针对 Q4 问题，应用训练好的分类器来评估 363 条隐私政策的《个保法》要求的符合情况，为了更详细地检测出 App 隐私政策遵守《个保法》的标

签要求，模型预测了除标注过的 100 条隐私政策外的剩余 262 条隐私政策(共计 82 595 个隐私政策语句)的标签情况，预测的样本数据集在训练模型过程中不可见。通过合规性检测，判断整个语料库中 363 条隐私政策的合规性情况。

在合规性检测前，需要对模型预测出的标注结果人工验证正确率，以确保模型在标注隐私政策语句时的准确性和可靠性。在剩余的 262 条未标注的隐私政策中，选取 50 条隐私政策，在 10 个类别中，每个类别选取 5 条，即 12 332 个隐私政策句子。9 名志愿者按照标注标准，对 12 332 个隐私政策句子通过模型预测后的标注结果进行验证，其中有 10 062 个隐私政策句子的模型预测结果是正确的，正确率达到 81.6%。

对整个语料库进行合规性检测和分析，判断 363 条隐私政策的合规性情况。图 4 显示了每个标签满足要求的隐私政策数量占比，从中可看出，大多数隐私政策满足 A 类基础标签和 B 类显性个人权利标签，而有很少一部分隐私政策满足 C 类隐性个人权利标签中“有权免受自动化决策”“有权向监管机构发起申诉”标签和 D 类附加标签中“未成年人个人信息处理规则”“可携带权”标签。

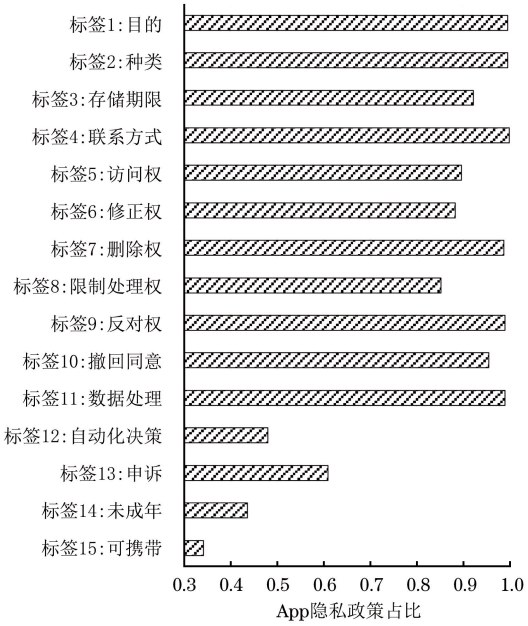

图 4 隐私政策满足《个保法》标签的占比

Fig.4 The proportion of privacy policies that meet the label requirements of the Personal Information Protection Law

图 5 显示了 363 条隐私政策中每条隐私政策符合 15 个标签的数量占比。只有 14.8%的隐私政策完全符合全部标签要求，大部分隐私政策符合 12～14 个《个保法》标签，但仍有一部分隐私政策不满足少量的《个保法》标签要求，甚至有个别隐私政策只满足几个标签要求，这将对隐私政策的合规性产生严重的影响。

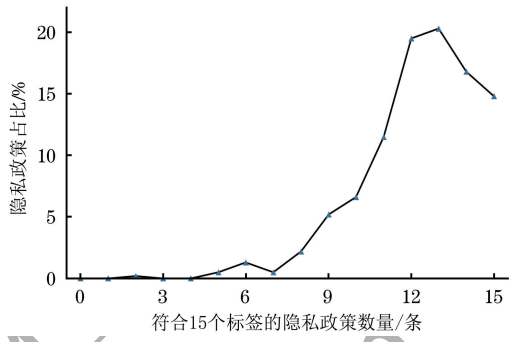

图 5 隐私政策符合标签个数

Fig.5 Number of labels that comply with privacy policy

图 6 显示了合规性结果。在 363 条隐私政策中，约 15%的隐私政策完全符合合规性要求，即满足所有要求标签；约 18% 的隐私政策基本符合合规性要求，即满足表 1 中的 A、B、C 级标签，但不满足附加标签“未成年人个人信息处理规则”“可携带权”；最多的是占比约 40% 的隐私政策，部分符合合规性要求，即满足 A 级“基础标签”和 B 级“显性个人权利标签”，但不满足 C 级“隐性个人权利标签”；而约 27%的隐私政策不符合合规性要求，即不完全满足 A、B 级标签。这些发现表明，许多 App 隐私政策仍然没有遵守《个保法》的要求。

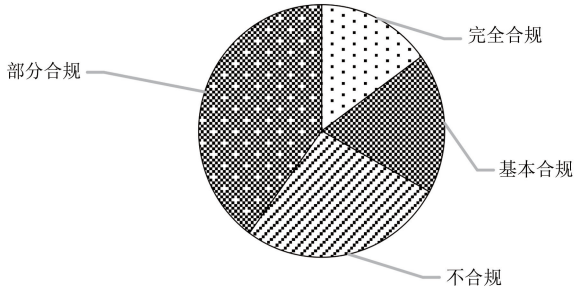

图 6 合规性结果统计

Fig.6 Compliance result statistics

4 结束语

本文面向《个人信息保护法》和 App 隐私政策之间的合规性分析任务，设计了一个基于《个人信息保护法》的标签方案，并创建了一个包含 363 条隐私政策、104 134 个多标签标注语句的语料库。用百度文心提出的 ERNIE 大语言模型对语料库进行了基准测试，隐私政策对《个保法》的合规性检测准确率达到 85.75%，并发现《个保法》的要求标签，如“有权免受自动化决策”“未成年人个人信息处理规则”“可携带权”，仅有不到 50%的覆盖率。

本文的工作为用户测试自己使用的 App 的隐私政策是否符合《个保法》要求、维护自己的权益提供了有效帮助，也为企业在制定自己的隐私政策时提供准确的建议。未来的工作和改进的方向主要是：1)本文在隐私政策收集方面主要针对安卓 App 的隐私政策，未来的工作将涉及更广泛的 App 来源，比如 iOS App；2)标签中“其他”类别标签占比过大，会对数据训练产生影响，存在数据不平衡问题，影响了分类精度，可以对其进行处理；3)需要丰富语料库的数据，提供更全面的评估，使研究者能够更全面地了解模型或算法在不同数据分布下的性能，提高模型的鲁棒性。

参考文献

[1] 刘颖，郝晓慧. 个人数据交易的法律基础[J]. 学术研究，2022(11)：85-94.
LIU Y, HAO X H. Legal basis for personal data transaction [J]. Academic Research, 2022(11): 85-94. (in Chinese)

[2] LIU S, ZHAO B Y, GUO R J, et al. Have you been properly notified? Automatic compliance analysis of privacy policy text with GDPR article 13[C]//Proceedings of the Web Conference 2021. New York, USA: ACM, 2021: 2154-2164.

[3] RAHAT T A, LONG M J, TIAN Y. Is your policy compliant? A deep learning-based empirical study of privacy policies' compliance with GDPR[C]//Proceedings of the 21st Workshop on Privacy in the Electronic Society. New York, USA: ACM, 2022: 89-102.

[4] CEJAS O A, AZEEM M I, ABUALHAIJA S, et al. NLP-based automated compliance checking of data processing agreements against GDPR [J]. IEEE Transactions on Software Engineering, 2023, 49(9): 4282-4303.

[5] WILSON S, SCHAUB F, DARA A A, et al. The creation and analysis of a website privacy policy corpus [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg, USA: ACL, 2016: 1330-1340.

[6] SRINATH M, WILSON S, GILES C L. Privacy at scale: introducing the PrivaSeer corpus of Web privacy policies[EB/OL]. [2024-02-02]. https://arxiv.org/pdf/2004.11131.

[7] KUZNETSOV M, NOVIKOVA E, KOTENKO I, et al. Privacy policies of IoT devices: collection and analysis[J]. Sensors (Basel), 2022, 22(5): 1838.

[8] ZHAO K F, ZHAN X, YU L, et al. Demystifying privacy policy of third-party libraries in mobile apps[C]//Proceedings of the IEEE/ACM 45th International Conference on Software Engineering. Melbourne, Australia: IEEE Press, 2023: 1583-1595.

[9] ZHAO K F, YU L, ZHOU S Y, et al. A fine-grained Chinese software privacy policy dataset for sequence labeling and regulation compliant identification[C]//Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Stroudsburg, USA: ACL, 2022: 10266-10277.

[10] HAMDANI R, MUSTAPHA M, AMARILES D R, et al. A combined rule-based and machine learning approach for automated GDPR compliance checking[C]//Proceedings of the 18th International Conference on Artificial Intelligence and Law. New York, USA: ACM, 2021: 40-49.

[11] HARKOUS H, FAWAZ K, LEBRET R, et al. Polisis: automated analysis and presentation of privacy policies using deep learning [C]// Proceedings of the 27th USENIX Conference on Security Symposium. [S. l.]: USENIX Association, 2018: 531-548.

[12] KIM N, OH H, CHOI J K. A privacy scoring framework: automation of privacy compliance and risk evaluation with standard indicators[J]. Journal of King Saud University-Computer and Information Sciences, 2023, 35(1): 514-525.

[13] BANNIHATTI K V, IYENGAR R, NISAL N, et al. Finding a choice in a haystack: automatic extraction of opt-out statements from privacy policy text[C]//Proceedings of the Web Conference 2020. New York, USA: ACM, 2020: 1943-1954.

[14] LIPPI M, PAŁKA P, CONTISSA G, et al. CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service[J]. Artificial Intelligence and Law, 2019, 27(2): 117-139.

[15] TORRE D, ABUALHAIJA S, SABETZADEH M, et al. An AI-assisted approach for checking the completeness of privacy policies against GDPR[C]//Proceedings of the IEEE 28th International Requirements Engineering Conference. Zurich, Switzerland: IEEE Press, 2020: 136-146.

[16] AMARAL O, ABUALHAIJA S, TORRE D, et al. AI-enabled automation for completeness checking of privacy policies[J]. IEEE Transactions on Software Engineering, 2022, 48(11): 4647-4674.

[17] TORRE D, SOLTANA G, SABETZADEH M, et al. Using models to enable compliance checking against the GDPR: an experience report[C]//Proceedings of the 2019 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems. Munich, Germany: IEEE Press, 2019: 10-20.

[18] HAMDANI R, MUSTAPHA M, AMARILES D, et al. A combined rule-based and machine learning approach for automated GDPR compliance checking[C]//Proceedings of the 18th International Conference on Artificial Intelligence and Law. New York, USA: ACM, 2021: 40-49.

[19] 李昕，唐鹏，张西珩，等. 面向 GDPR 隐私政策合规性的智能化检测方法[J]. 网络与信息安全学报，2023，9(6)：127-139.
LI X, TANG P, ZHANG X H, et al. GDPR-oriented intelligent checking method of privacy policies compliance[J]. Chinese Journal of Network and Information Security, 2023, 9(6): 127-139. (in Chinese)

[20] LIAO S, ALDEEN M, YAN J W, et al. Understanding GDPR non-compliance in privacy policies of alexa skills in European marketplaces[C]//Proceedings of the ACM Web Conference 2024. New York, USA: ACM, 2024: 1081-1091.

[21] 赵杨，严周周，沈棋琦，等. 基于机器学习的医疗健康 APP 隐私政策合规性研究[J]. 数据分析与知识发现，2022，6(5)：112-126.
ZHAO Y, YAN Z Z, SHEN Q Q, et al. Research on privacy policy compliance of medical health APP based on machine learning [J]. Data Analysis and Knowledge Discovery, 2022, 6(5): 112-126. (in Chinese)

[22] 徐奇睿. 基于《个人信息保护法》的移动互联网 APP 隐私政策合规研究[D]. 武汉：武汉大学，2022.
XU Q R. Research on privacy policy compliance of mobile

Internet APP based on Personal Information Protection Law [D]. Wuhan: Wuhan University, 2022. (in Chinese)
[23] MOTLAGH N Y, KHAJAVI M, SHARIFI A, et al. The impact of artificial intelligence on the evolution of digital education: a comparative study of OpenAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie [EB/OL]. [2024-02-02]. https://arxiv.org/abs/2309.02029.
[24] SHANG W Q, JI J L, LIU Y R. Chinese text correction system based on ernie [C]//Proceedings of the 26th ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. Washington D. C., USA: IEEE Press, 2023: 103-108.
[25] 盛小平, 唐筠杰. 我国个人信息权利与欧盟个人数据权利的比较分析: 基于《个人信息保护法》与 GDPR[J]. 图书情报工作, 2022, 66(6): 26-33.
SHENG X P, TANG J J. A comparative analysis of personal information rights in China and personal data rights in EU: based on the Personal Information Protection Law and GDPR [J]. Library and Information Service, 2022, 66(6): 26-33. (in Chinese)

编辑 金胡考