

数据故事中的人物类型及自动生成方法研究*

■ 朝乐门^{1,2} 刘慧² 张天怡² 李泽仑²

¹ 数据工程与知识工程教育部重点实验室 (中国人民大学) 北京 100872

² 中国人民大学信息资源管理学院 北京 100872

摘要: [目的/意义] 人物和情节是数据故事的两大支柱。数据故事的情节通过人物特征、行为、所期待目标、所面对现实和所认为偏见来展开, 实现数据故事人物的自动化生成是数据故事化领域科学研究的核心主题之一, 对于数据故事的理论研究、自动生成和工程化研发具有重要意义。[方法/过程] 首先, 探讨数据故事人物的类型、特征及操作。其次, 提出基于反事实解释的人物生成方法, 分别对数据故事中的主人公、同类人物、异类人物、正面人物和反面人物给出自动生成方法。接着, 分析其技术实现, 探讨实验设计、数据来源、方法选择及结果讨论。最后, 总结论文的主要研究发现, 并对未来研究提出建议。[结果/结论] 在数据故事化领域首次较为系统研究数据故事人物的组成要素、基本类型、主要特征及核心操作, 并提出基于反事实的数据故事人物自动生成方法。

关键词: 数据故事 数据故事化 故事人物 人物

分类号: G203

DOI: 10.13266/j.issn.0252-3116.2023.24.009

1 引言

数据故事是以满足业务需求为出发点, 对原始数据进行分析、建模、价值洞察等操作并最终将其以故事的形式呈现给用户的一种产品或服务^[1]。数据故事主要由 7 种基本要素组成: 需求、人物、情境、情节、冲突、解决方案以及下一步行动^[2]。相对于数据可视化, 数据故事采取了更丰富的表达方式: 它运用故事作为主要的数据传递手段, 而不仅仅依赖于视觉编码的单一形式。数据故事利用叙事策略将数据中的核心价值编码成更易于目标受众认知、记忆和体验的表达结构, 从而实现数据认知、算法解释和虚实结合的目标。例如, 在某销售系统的商品推荐业务中, “客户”被塑造为角色, 基于其特征数据形成的购买意愿预测、预测结果与客户实际购买意愿之间的冲突、系统对某类客户购买行为的预测偏见等, 都可以被整合作为情节元素, 通过关联和叙述构成一个完整的数据故事, 更有效地满足销售业务的需求。

在数据故事的基本要素中, 人物和情节是数据故事的两个重要支柱。在数据故事中, 人物 (character) 是指数据故事提及的人和 (或) 物^[3]。通常,

数据故事的情节通过作者对人物特征、行为、意愿、语言和关系的叙述展开。在数据故事中, 人物主要由特征、行为、目标、现实和偏见 5 个部分组成 (见图 1), 其中: ①特征 (feature) 是指人物的特征信息或属性信息, 表示数据故事生成系统中自动决策的输入或依据信息, 如年龄、身高、体重、收入水平等; ②行为 (action) 是人物发起的行动或采取的操作, 表示在数据故事生成系统中主体的行动, 如“购买一双运动鞋”; ③目标 (expectation) 是人物的主观愿望, 表示人物所期待的数据故事生成系统中的主体需求, 如“某客户期待得到一双价格不高于 500 元的白色运动鞋”; ④现实 (reality) 是指数据故事生成系统基于人物所提交的特征信息计算的输出结果, 表示数据故事生成系统基于人物信息做出的决策或结论, 如“某销售系统给用户推荐的是一双价格为 500 元的黑色运动鞋”; ⑤偏见 (bias) 是人物认为数据故事系统做出了带有偏见或歧视的不公平决策时, 用于公平性检测和歧视识别的判断条件。在数据故事中, 偏见通常以人物的特征、行为、目标和结果组成的表达式表示。例如, 当性别特征为“男”, 无论其他特征如何变化, 数据系统的预测

* 本文系国家自然科学基金项目“预测性分析结果的数据故事化描述方法及关键技术” (项目编号: 72074214) 研究成果之一。

作者简介: 朝乐门, 教授, 博士生导师; 刘慧, 硕士研究生, 通信作者, E-mail: Liuhui_s@foxmail.com; 张天怡, 硕士研究生; 李泽仑: 硕士研究生。

收稿日期: 2023-04-04 修回日期: 2023-07-26 本文起止页码: 99-110 本文责任编辑: 杜杏叶

结果不改变,即现实为“不予以贷款”,此时的偏见表达式为: $P(\text{gender}=\text{“男”} \ \&\& \ \text{result}=\text{“允许贷款”}) < 0.05$ 。由上可见,对于人物的某种行为而言,特征是决策的输入信息,结果是决策的输出信息,

行为是触发决策的过程,而系统做出的自动决策结果可能与人物主观目标不一致。通常,目标(或意愿)的损失(即目标与现实之间的差距)以及偏见的存在是数据故事中的主要矛盾。

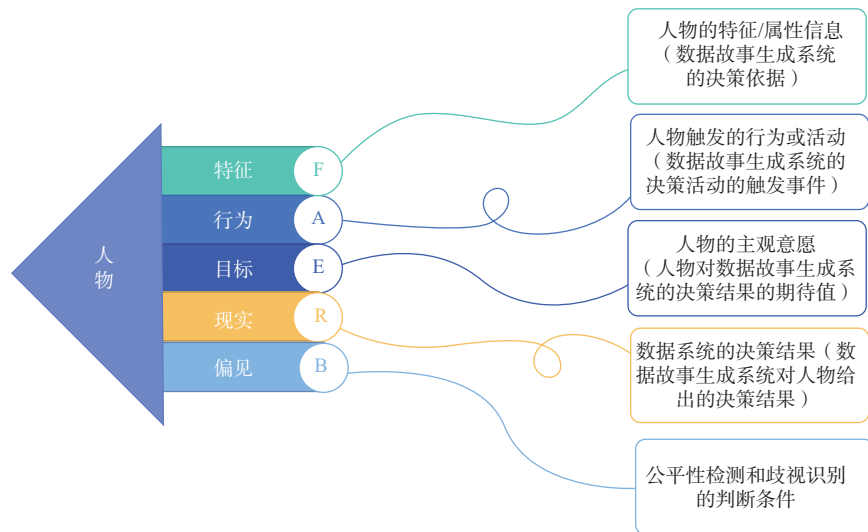


图1 数据故事中的人物
Figure 1 Characters in a data story

数据故事的人物自动生成是数据故事理论研究、工程开发和推广应用的重要课题之一,其对数据故事学科的发展有着深远影响。对于理论研究者来说,人物类型及其自动生成技术的探索提供了数据故事理论研究的关键科学问题和研究方向;对于工程开发者而言,人物自动生成是数据故事系统的技术难题之一;对于决策者和业务管理者而言,数据故事的人物类型自动生成技术可以帮助他们构建与业务场景紧密相联的数据故事,提供更为直观、生动的数据分析结果和决策支持,从而更好地引导业务决策过程。对于非专业用户和公众而言,自动生成的人物类型使得数据故事的构建过程更加高效、便捷,同时提供了丰富的角色类型,能够更好地引发受众的兴趣和共鸣。然而,对于数据故事人物的深度研究目前还相对缺乏,对此领域的突破性研究有待进行。这一领域的进展不仅将推动数据故事的理论研究和实践应用,也将为数据故事的进一步发展提供重要的理论和技术支持。为此,笔者综合运用采用文献研究、软件工程和反事实解释方法,研究数据故事中的人物类型及其自动生成方法,深入探讨数据故事人物的理论基础,提出一种基于反事实解释的不同人物类型的自动生成方法。

2 相关研究

2.1 研究综述

故事人物的研究对于数据故事领域的发展具有

重要意义。但是,目前对数据故事化中人物的专门研究极少,而在一般意义上的故事人物的讨论与研究较多,相关研究主要分散在以下3个领域:

(1) 人文社会学领域主要探讨如何更好地描述和理解故事人物以及如何通过故事人物准确传递作者的意图,进而提高故事创造和故事理解的效果。例如,当作者使用第一人称创作故事时,读者更容易感受到故事主人公的情绪^[3];当故事中的人物遭遇生存威胁时,读者也会身临其境感受到威胁^[4];故事人物是影响读者心理的关键因素,设置其参与相关服务行为的内在动机能够有效提高读者对服务的兴趣^[5]。

(2) 计算机科学领域研究主要侧重的是人物信息的提取和推理,进而进行故事的自动生成和语义理解。例如,基于人机协作的故事编辑系统 Erato,该系统采用一种事实插值算法来创建中间事实,使得故事之间的衔接更加自然,且用户只需编辑几个关键帧即可轻松生成故事,推动了故事自动生成的研究^[6]; B. Lee 讨论了将数据转换为可视化共享故事的整个过程,即通过开放编辑脚本角色,让观众参与故事创作过程,设定以及编辑故事人物,这些新兴场景可以吸引观众的注意力,并使原始数据和故事创作过程更加透明^[7]; Web 开发、数据分析和数据可视化逐渐成为创建故事所需的核心技术^[8]。

(3) 数据科学领域主要探讨的是故事人物的可

视化及用户交互问题, 进而提高故事的可理解性及用户体验。Y. Zhang 等认为交互设计和用户体验研究是交互式讲故事的重要组成部分^[9], Y. Shi 等通过分析 58 个高质量故事, 归纳出 6 种设计模式, 证明通过交互来打破第四堵墙 (Breaking the fourth wall) 对用户沉浸感、用户粘性与信息回忆均存在积极影响^[10]; 故事世界由人物与场景组成^[11], 基于此, Y. Zhang 等设计了一个新的框架来反映静态环境到动态故事环境的变化, 用户可以自由选择数据条目, 激活相关的数据故事人物^[12]。

通过以上研究发现: ①目前关于数据故事人物的研究主要集中在理论层面, 在研究目的、研究内容、研究方法等方面达成基本共识, 但对数据故事人物的生成方面的研究较少; ②现有的数据故事人物生成多依赖于数据分析平台, 智能化较低, 生成数据故事人物的方法还不够完善。因此, 数据故事人物的基础理论有待深入系统的研究和创新, 数据故事人物生成方法的研究更是亟待深入。

2.2 研究方法

反事实解释是一种局部解释方法, 其基本思想是通过更改输入 (或输出) 来改变黑盒模型输出 (或输入)。反事实解释 (counterfactual explanations) 的定义方法有很多, 比较有代表性的有 3 种: ①描述如果修改某些输入特征, 机器学习模型的输出将如何改变的解^[13]; ②描述过去某些事件或行动如果不同, 将会发生什么的解^[14]; ③描述在不同情况下系统会如何行为的解^[15]。与其他可解释性机器学习不同的是, 反事实解释不是对分类依据做出解释, 而是对如何改变实例的特征使得分类结果发生改变的过程进行解^[16]。反事实解释方法与人类的逆向思维过程高度相似, 在金融 (如风险评估) 与医疗保健 (如癌症概率预测) 领域有广阔的应用前景^[17]。反事实解释方法在描述数据故事中人物间的关系方面显示出了优越性, 其通过反事实解释技术生成的人物固有的相互对照特性, 能够强化数据故事的解释功能, 这使其成为数据故事人物生成的有效手段之一。

反事实解释通常以“如果 X 没有发生, Y 就不会发生”的形式描述因果关系^[18]。例如: “如果贷款人信用积分增加, 贷款申请就能通过。”其中, Y 是贷款申请通过的事件结果; X 是信用积分增加的事件产生原因。思考反事实需要设想一个与观察到的事实相矛盾的假设现实——“反事实”。

通常, 将反事实解释的生成方法分为扰动式与案例式^[19]两大类。扰动式反事实解释通过定义相关损失来找出最小扰动^[20], 包括整数编码^[21]、梯度下降^[22]等方式; 案例式反事实解释通过分析案例特征来解决优化问题^[23], 包括查询最短路径^[24]、案例推理^[25]等方式。

目前, 反事实解释的生成方法主要关注的是度量方法的创新和改进。S. Wachter 等提出了反事实解释, 并以 LSAT 数据集为例, 探讨未加权平方欧式距离、归一化欧氏距离以及归一化 L_1 范数等不同度量函数对反事实结果的影响^[26]; T. laugel 等采用亲近性 (proximity)、连贯性 (connectedness) 和稳定性 (stability) 3 种方法对反事实解释进行度量^[27]; D. Mahajan 等在反事实解释的损失函数的基础上分别针对外生变量与内生变量两种情况提出因果距离的衡量方法^[28]。

3 数据故事中的人物

正确认识人物类型和特征是深入研究数据故事的前提, 而探讨数据故事人物操作的定义是数据故事自动生成的关键。为此, 笔者通过讨论数据故事中的人物组成要素与主要类型以及数据故事人物与文学故事人物的区别, 定义数据故事人物的基本操作, 为提出不同类型人物的自动生成方法提供依据。

3.1 人物类型

人物类型的多样性及人物关系的复杂性是数据故事情节变化的基础。根据数据故事的特征及数据故事化工作的需要, 数据故事中的人物可以进一步分为主人公、正面同类人物 (positive & ally character, PA 人物)、正面异类人物 (positive but opponent character, PO 人物)、反面同类人物 (negative but ally character, NA 人物) 和反面异类人物 (negative & opponent character, NO 人物), 如表 1 所示:

表 1 数据故事中的人物类型
Table 1 Types of characters in data stories

主人公		人物特征	
		同类人物 (ally character)	异类人物 (opponent character)
人物 关系	正面人物 (positive character)	PA 人物	PO 人物
	反面人物 (negative character)	NA 人物	NO 人物

(1) 主人公是数据故事的中心人物和情节发展的关键焦点, 也是定义数据故事中同类人物、异类人物、正面人物和反面人物的主要依据。以信用卡违约

预测数据故事为例,正在刷卡消费的用户(如“张三”)为主人公,可以基于该主人公定义出同类人物或异类人物以及同类人物和异类人物。

(2) 同类人物和异类人物是根据与主人公在关键特征或可能产生偏见的特征上是否取值相同来决定。其中,关键特征是指对于预测结果而言“关键”的特征,即特征重要性排序靠前或主成分分析中的主成分,如收入水平、社会信誉积分以及历史记录中的违约次数等;偏见特征是指对于特定用户而言“敏感”的特征,即容易产生偏见或歧视的特征,如宗教信仰、性别以及种族等。在关键特征或偏见特征上的取值与主人公(或其他人物)一致,则属于同类人物,反之则属于异类人物。可见,数据故事中区分同类人物与异类人物的关键在于关键特征或偏见特征的取值是否一致。

(3) 正面人物和反面人物是根据人物与主人公的行动目标是否一致来决定。在此,行动目标是指数据故事中人物的目标、需求、意愿或期待,如成功进行信用卡消费、购买性价比高且自己喜欢的特定款商品等。对于数据故事化所对应的业务系统而言,关键特征(或偏见特征)和行动目标分别为业务系统的预测变量与被预测变量。区分正面人物与反面人物的关键在于预测结果或分类标签是否一致。

3.2 人物特征

相对于文学故事中的人物,数据故事中的人物具有以下5个特征:

(1) 受众即数据故事的主人公。故事受众是相对于数据故事的作者和叙述者而言的概念,泛指阅读、观看和消费数据故事的人群。在文学故事中,故事中提及的主人公和故事所面对的受众并非同一个主体。而数据故事中的主人公通常为受众自己。受众通过与数据故事生成系统不断交互生成数据故事。因此,数据故事具有较好的用户体验。

(2) 无需刻画人物的性格特征和人物形象。与文学故事不同的是,数据故事是业务导向型的数据产品,以解决业务理解、算法解释和数据分析为目的,并不侧重于刻画人物的性格特征和树立人物形象。在数据故事中,人物叙述为情节描述服务,而情节描述的目的在于数据理解和算法解释,进而解决故事受众的特定业务需求或决策需要。

(3) 以主人公为中心的人物关系。数据故事中出现的人物类型由其与主人公之间的联系来决定。数据故事中的非主人公可以从同/异类人物和正/反面人物两个维度进行划分。其中,同类人物和异类人物的划分根据关键属性(包括偏见属性)是否与同类人物一致,正面人物和反面人物的区别在于行为目标是否与主人公一致。相对于文学故事,数据故事中的人物关系较为简单,而文学故事中的人物类型更为复杂多变。

(4) 由算法定义的故事人物及其关系。数据故事中的人物及人物之间的关系并不一定根据真实的历史痕迹数据建立或由数据故事的叙述者手工定义,而是借助算法生成数据故事所需的人物及人物之间的关系。

(5) 支持旋转的人物关系及多维度探索型分析。数据故事需要支持主人公和非主人公之间旋转操作,进而实现数据故事的多维度观察和对比操作,进行数据故事的探索型分析,提升数据故事的用户体验及叙述效果。

3.3 人物操作

对于数据故事而言,人物操作包括 CRUD、校验、钻取、切取和旋转5种基本操作(见表2)。其中,CRUD操作是指数据故事中人物的生成(create)、读取(read)、更新(update)和删除(delete)等通用操作。除了CRUD操作,数据故事中需要如下4种特殊操作:

表2 数据故事人物基本操作
Table 2 Basic operations of data story characters

名称	含义	举例
CRUD	数据故事中人物的生成、读取、更新和删除	在信用卡违约预测故事项目中的UI界面对5种人物进行生成、读取、更新和删除的操作
校验	检查人物信息的完整性和一致性,并对发现的问题进行纠正或删除	在信用卡违约预测故事项目中通过设置异常处理对主人公输入信息进行检验,并处理错误信息
切取	通过“选取”和“投影”选取操作候选人物数据	在信用卡违约预测故事项目中对人物进行“选取”和“投影”操作设置条件查询
钻取	改变数据故事中的人物单位的方法,可以分为上钻和下钻两大类	上钻:在信用卡违约预测故事中设置条件对满足条件的人物组进行查询 下钻:对具体的人物类型特征进行查询
旋转	非主人公转换为主人公	在信用卡违约预测故事项目中选择非主人公人物,通过算法生成其他4种类型人物,此时非主人公人物转换为主人公

(1) 校验。校验操作旨在检查人物信息的完整性和一致性。随着数据故事的交互式叙述, 数据故事中难免产生错误数据和垃圾数据, 此时需要采用技术手段进行自动校验。完整性检验的重点在于检查人物信息, 尤其是人物的特征、行为、目标、现实和偏见的记录是否完整有效以及是否被恶意篡改过; 一致性检验的重点是检查人物属性值与模式信息之间以及人物之间的关系定义是否存在冲突与不一致, 如一则故事中是否有同时存在多个主人公。对于检验结果中发现的问题, 可以采取自动纠正、手动纠正或直接删除等措施。

(2) 切取。切取是针对特定人物数据进行选取操作的过程, 通过对候选数据集进行“选择”和“投影”两类操作过滤掉部分人物或人物的部分信息, 防止这些信息参与数据故事化的计算过程(见图2)。其中, “选择”操作的切取方向为样本向量方向(行方向), 决定数据库中哪些人物可以出现在数据故事中; “投影”操作的切取方向为特征向量方向(列方向), 决定人物的哪些信息可以参与数据故事化的计算之中。可见, “投影”操作的主要目的是屏蔽掉故事人物的隐私数据、敏感数据和非相关数据。

(3) 钻取。钻取是改变数据故事中的人物叙述单位的方法, 可以分为上钻和下钻两大类。主人公的叙述单位可以是3大类: 个体、群体或个体的某个时空范围内的阶段性个体(见图3)。例如, 对主人公“张三”的年龄特征(age)给出的限制条件为“ $18 \leq \text{age} \leq 22$ ”, 可以定义出一个新的虚拟人物——年龄为18和22之间阶段性个体, 进而可以与其他相关阶

段性个体进行对比分析。

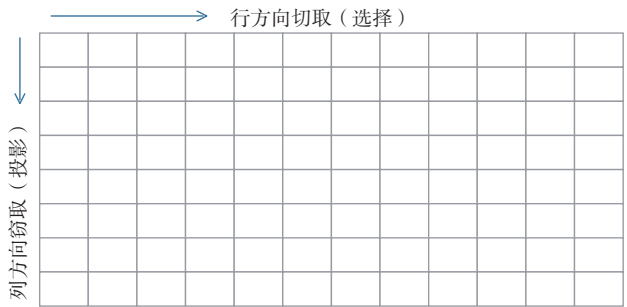


图2 数据故事人物的切取操作
Figure 2 The cut operation of data story characters

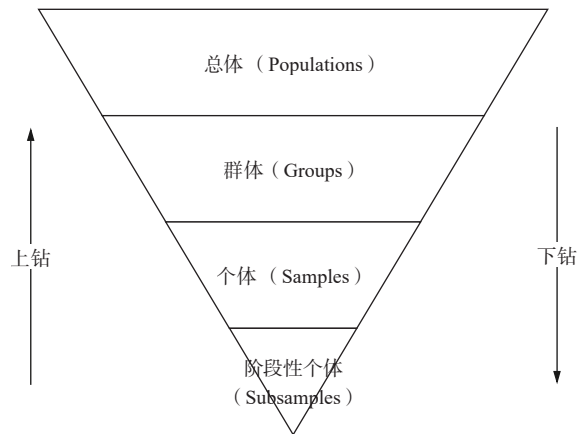


图3 数据故事人物的钻取操作
Figure 3 Drilling operation of data story characters

(4) 旋转。旋转是指将数据故事中的非主人公转换为主人公或将主人公转换为非主人公(见图4)。在数据故事中, 主人公的旋转将触发故事人物类型及相互关系的自动变化, 即根据旋转操作后产生的新主人公, 重新计算数据中的正/反面人物和同/异类人物。

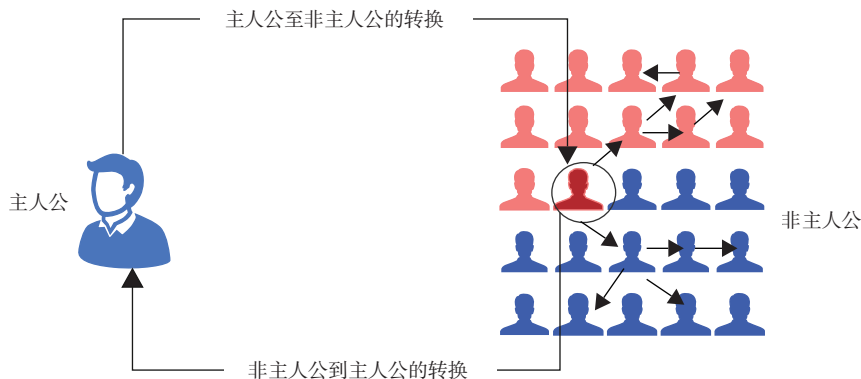


图4 数据故事中的人物旋转操作
Figure 4 The character rotation operation in a data story

4 数据故事人物的自动生成方法

反事实解释的目标在于找到改变给定分类结果所需的特征改变的最短优化路径。如表3所示, 反事

实解释没有告知为何被分类为贷款不通过, 而是给出了需要改变实例的哪些特征才能通过贷款的解释, 即样本需要增加年龄和获得硕士学位才能通过贷款。

表 3 反事实解释实例
Table 3 Examples of counterfactual explanations

样例	年龄 / 岁	工作类型	性别	受教育程度	平均月收入 / 元	贷款申请
查询样本	24	服务	女	学士	8 000	不通过
反事实解释	26	服务	女	硕士	8 000	通过

由于本文中提到的信用卡违约预测系统内部包含多个变量并且具有复杂的依赖关系，对于用户来说，推理系统决策的内部逻辑是一个巨大挑战。为了帮助用户理解系统做出决策的内部逻辑，考虑到反事实解释在不打开算法黑箱的前提下，具有易于

精确计算和便于向外传达信息的优势^[30]。本节将基于反事实解释技术生成数据故事人物，包括主人公、同 / 异类人物以及正 / 反面人物（见图 5）。考虑到算法的鲁棒性，采用归一化 L_1 范数指标进行样本与反事实实例距离的度量。

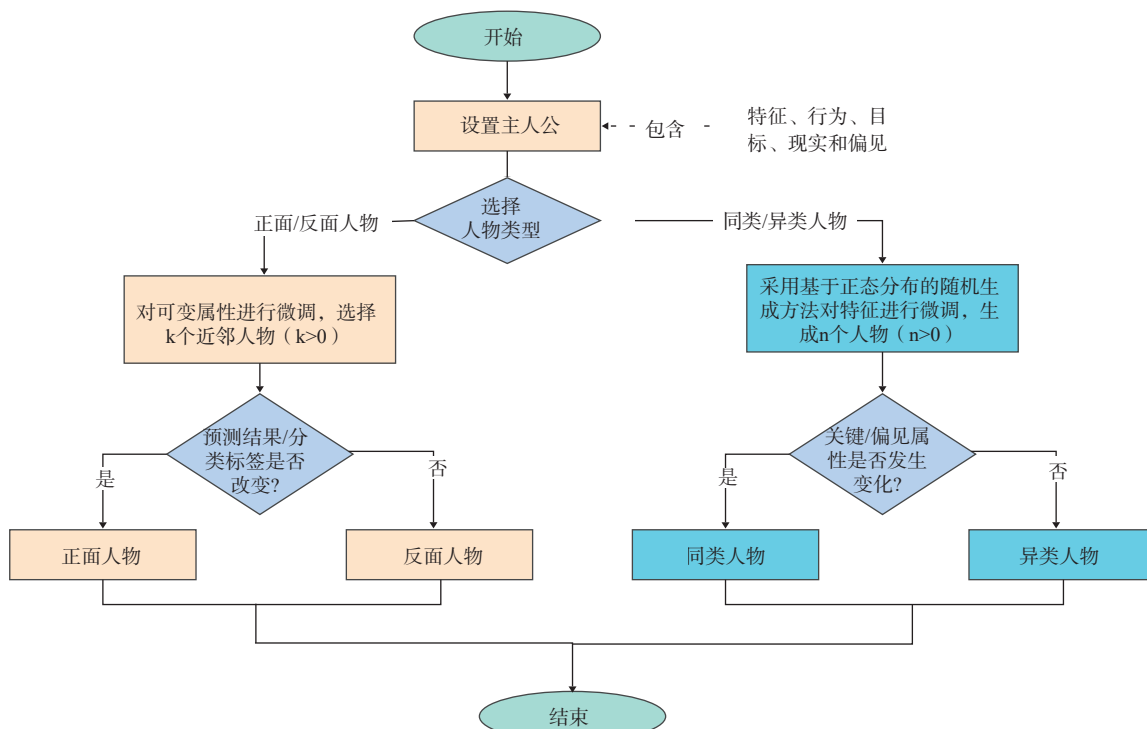


图 5 数据故事人物自动化生成流程

Figure 5 Automation generation process of data story character

4.1 主人公及其生成

主人公为系统 S 的一个特定用户 t ，其特征信息和预测结果分别为 $t[X]$ 和 $t[y]$ 。 $t[X]$ 和 $t[y]$ 值不仅可以为用户当前输入值，也可以为历史输入值或系统 S 默认值。主人公的各个特征取决于用户输入，主人公包括 $name_char$ 、 $target$ 、 $target_expected$ 、 $features$ 、 $expression_bias$ 等特征。其中， $name_char$ 指用户姓名， $target$ 指预测系统基于用户特征的预测结果即现实， $target_expected$ 指用户期待（或想要）的结果即目标， $features$ 为主人公的特征， $expression_bias$ 为偏见。下文会出现的字符及其含义如表 4 所示：

expected、 $features$ 、 $expression_bias$ 等特征。其中， $name_char$ 指用户姓名， $target$ 指预测系统基于用户特征的预测结果即现实， $target_expected$ 指用户期待（或想要）的结果即目标， $features$ 为主人公的特征， $expression_bias$ 为偏见。下文会出现的字符及其含义如表 4 所示：

表 4 相关字符含义

Table 4 Meanings of related characters

字符名称	含义	字符名称	含义
$t[X]$	用户 t 提交给系统 S 的输入信息	$t^{\#}$	异类人物
$t[y]$	系统 S 向用户 t 返回的预测结果	t^+	正面人物
X	特征集的名称	t^-	反面人物
X^*	特征集 X^* 的子集，表示偏见特征或关键特征， $X^* \subset X$ ，对应的目标向量记为 y^*	p	主人公 t 的可变属性个数
t	主人公	x_j	主人公 t 的第 j 个可变属性值
$t^{\#}$	同类人物	x'_j	主人公 t 的第 j 个可变属性微调后取值

4.2 同 / 异类人物及其生成

同 / 异类人物的判断依据为在特征子集 X^* 上与主人公相同的人物是否相同。若相同则称之为“与主人公相同类型的人物”，否则为“与主人公不同类型的人物”。生成同类人物与异类人物的目的在于帮助用户理解信用卡违约预测系统做出决策的依据，通过对所有特征进行微调，观察预测结果的变化，检验偏见特征对于预测结果的影响程度。

(1) 同类人物。寻找主人公 t 的同类型的人物 t^+ 与偏见特征或关键特征集上是否与主人公一致，而与标签 y 无关。同类人物是指在敏感特征子集上与主人公 t 相同，在非敏感特征子集上与主人公 t 不同的人。为了数据故事化的需要，一般为主人公寻找 k 个最具代表性的同类人物，分别代表与主人公的差异性的最大值、最小值及各分位数。为此笔者引入一个滑动窗口技术，并在样本 X 的滑动窗口内通过扰动生成新的样本，需要指定这个窗口的大小。假设滑动窗口大小为 h ，也就是说，对每个非敏感特征 x_i 可以在 $[x_i-h, x_i+h]$ 范围内进行扰动。以下是生成新样本的步骤和方法：

算法名称：生成 top-k 个代表性同类人物

输入：

原始样本 $X = [x_1, x_2, \dots, x_m]$

敏感特征集 $X^* = [x_1^*, x_2^*, \dots, x_n^*]$

扰动窗口大小 h

新样本数 N

给定的样本 t

k 个代表性样本数

输出：

扰动后的样本集 $samples$

top-k 代表性样本 $representative_samples$

算法步骤：

1. 初始化一个空列表 $samples$ 来保存生成的新样本

2. 对于 $i = 1$ 到 N ：

2.1 初始化一个新样本 $new_sample = []$

2.2 对于每个特征 x_i 在样本 X 中：

2.2.1 如果 x_i 不是敏感特征（在 X^* 中），从 $[x_i - h, x_i + h]$ 范围内随机选择一个值 new_x_i ，将 new_x_i 添加到 new_sample

2.2.2 如果 x_i 是敏感特征，保持其值与给定的样本 t 的对应特征

值相同，将 $t[x_i]$ 添加到 new_sample

2.3 将新的扰动后的样本 new_sample 添加到 $samples$ 列表中

4. 对 $samples$ 列表执行聚类算法，产生 k 个聚类

5. 根据 k 个聚类中心到主人公 t 的距离进行排序

6. 找出最大值、最小值、中位数、上四分位数、下四分位数样本

7. 如果找出的样本数量小于 k ，则递归地在剩余的样本中重复以上步骤，直到找出 k 个样本

8. 将找出的样本添加到代表性样本列表 $representative_samples$ ，作为主人公 t 的 top-k 个代表性同类人物。

(2) 异类人物。异类人物是指敏感特征子集上与主人公 t 不同，而非敏感特征子集上与主人公 t 相同的人物。与同类人物的生成类似，为了数据故事化的需要，仍需要生成 k 个最具代表性的异类人物。考虑到生成并找到 k 个最具代表性异类人物的算法与同类人物的算法基本类似，在此不再赘述。

4.3 正 / 反面人物及其生成

正面人物和反面人物的判断依据为分类标签 y 与主人公是否相同，若相同则称之为“正面人物 t^+ ”，否则为“反面人物 t^- ”。生成正面人物与反面人物的目的在于找到达到用户期待结果所需做出的必要改动，二者的区别在于前者是标签 y 不变的前提下特征 X 的最大抖动量，而后者为标签 y 发生变化的情况下特征 X 的最小抖动量。例如用户 A 在信用卡违约预测系统的预测结果为违约，因此贷款不通过，而用户 A 想通过做出较小的改变来获得贷款，此时用户 A 可以选择在系统里生成反面人物，对年收入等可变特征进行调节达到贷款要求。

(1) 正面人物。 X^* 取与 t 相同的值，并对补集特征进行抖动，抖动幅度为标签 y 不变前提下的最大抖动量，生成方法如公式 (1) 所示：

$$t^+ = \{z \mid z[X^*] = t[X^*] \cap z[X^c] \text{ 为标签 } y \text{ 不变的前提下, 特征 } X \text{ 的最大抖动量} \} \quad \text{公式 (1)}$$

假设业务系统为 S ，样本 $t[X]$ 的标签为 y ，该样本含有多个特征 x_1, x_2, \dots, x_m ，且每个特征的权重分别为 w_1, w_2, \dots, w_m ，样本 $t[X]$ 可以表示为一个权重向量和特征向量的点积形式，即 $t[X] = w_1x_1 + w_2x_2 + \dots + w_mx_m$ 。目标是找到一个扰动量 δ ，使得

其大小最大，而同时确保模型对扰动后的样本的预测结果不变。因此，生成正面人物问题可以表示为一个约束条件下最优化问题：

$$\begin{aligned} & \text{maximize: } \|\delta\|_2 \\ & \text{subject to: } S(w_1*(x_1 + \delta_1) + w_2*(x_2 + \delta_2) + \dots + \\ & \quad w_m*(x_m + \delta_m)) = y \end{aligned} \quad \text{公式(2)}$$

在公式(2)中， $\|\delta\|_2$ 表示L2范数，也就是欧几里得范数，它用来度量扰动 δ 的大小。约束条件 $S(w_1*(x_1 + \delta_1) + w_2*(x_2 + \delta_2) + \dots + w_m*(x_m + \delta_m)) = y$ 确保了模型对扰动后的样本预测结果依然为 y 。

公式(2)可以利用诸如拉格朗日乘数法、KKT条件等方法求解，也可以利用迭代优化方法(如梯度下降法)、启发式搜索，甚至结合近似或者线性化技术来进行求解。

(2)反面人物。 X^* 取与 t 不同的值，补集特征采用符合相应特征定义域的随机值，即反面人物可以表示为：

$t^{\neq} = \{z \mid z[X^*] \neq t[X^*] \cap z[X^*]\}$ 为标签 y

发生改变所需要特征 X 的最小抖动量} 公式(3)

可见，寻找反面人物是一个有约束优化问题。在寻找特征集最小抖动量的问题中，目标是找到一个扰动 δ ，使得其取值最小，且能够使得模型对扰动后的样本预测结果发生改变。因此，该问题的形式化表述如下：

$$\begin{aligned} & \text{minimize: } \|\delta\|_2 \\ & \text{subject to: } S(w_1*(x_1 + \delta_1) + w_2*(x_2 + \delta_2) + \dots + \\ & \quad w_m*(x_m + \delta_m)) \neq y \end{aligned} \quad \text{公式(4)}$$

在公式(4)中，约束条件 $S(w_1*(x_1 + \delta_1) + w_2*(x_2 + \delta_2) + \dots + w_m*(x_m + \delta_m)) \neq y$ 模型对扰动后的样本预测结果与原来的标签 y 不同，即标签发生改变，其他符号或参数的含义与公式(2)相同。

5 实验与讨论

本实验的目的为实现和验证前文提出的数据故事人物的生成方法，故事人物的类型包括主人公、同类人物和异类人物、正面人物和反面人物，实现数据故事人物的5个基本操作：CRUD、校验、钻取、切取和旋转。

5.1 数据与方法

为提供数据故事的数据来源，本研究设计了5个数据库关系表：①人物特征表，主要记录故事人物特征id与名称的对应信息，涉及4个字段：特征ID、特征名称、可变性、偏见；②人物类型表，主要记

录人物类型id与名称的对应信息，涉及2个字段：人物类型ID、人物类型名称；③人物关系表，主要记录不同人物类型之间关系信息，涉及4个字段：人物关系ID、前件前置人物ID、后件人物ID、人物类型ID；④人物表，主要记录故事人物信息，涉及5个字段：人物ID、人物名称、现实、目标、偏见特征；⑤人物细节表，主要记录故事人物特征有关信息，涉及5个字段：人物细节ID、人物ID、特征ID、特征值、时间戳。

其中，人物类型表与人物关系表通过外键人物类型ID连接，关联关系为一对多；人物关系表与人物表通过外键人物ID连接，关联关系为一对一；人物表与人物细节表通过外键人物ID连接，关联关系为一对多；人物细节表与人物特征表通过外键特征ID连接，关联关系为一对多。

在技术方案的选择上，本研究采用基于Python面向对象编程方法，将本系统开发为一个包含3个模块，涉及6个类的Python模块，每个类的含义如下：①人物类代表数据故事的人物，包括名称、目标、现实、偏见、特征5个共有属性，以及故事人物更新、查询、删除、钻取、切取5个操作；②主人公类代表数据故事的主人公，继承了人物类的所有属性和操作，并且新增了生成的操作；③类名为正面人物、反面人物、同类人物、异类人物的其他4种人物类，继承了人物类的所有属性和操作，并新增了旋转操作。

5.2 结果及讨论

本研究主要实现主人公、正/反面人物和同/异类人物的CRUD、校验、钻取、切取、旋转等操作，并制作数据故事系统的人机交互界面以呈现操作效果。

(1)人物查询。为了从数据库中选取人物，并指定人物类型，本系统首先实现了人物查询功能。人物查询功能支持无条件查询和有条件查询。其中，无条件查询为用户不需要指定任何查询条件，系统将当前数据故事中的已有全部人物显示；有条件查询为用户输入查询条件，系统返回符合用户输入的查询条件的所有人物信息。通常，用户先进行无条件查询，界面显示数据库中的所有人物信息，并根据所列出的所有人物信息进一步进行有条件的查询，最终找到合适的人物，并为其分配人物类型。如图6所示，用户进行的是无条件查询，系统返回数据库中的所有人物及其信息，包括：用户名、现实、目标、特征和偏见。

点击【修改】和【删除】按钮可对查询结果中的信息进行编辑；点击【钻取】按钮，在弹出的对话框中设置下拉选项，分别为均值、出现频次最多的值和出现频次最少的值，用户可以对每个字段进行设置，基于选中的人物数据生成新的虚拟人物；点击【切取】按钮，用户可以选择任意行与列，覆盖原有查询结果。以信用卡违约预测系统为例，人物查询界面允许用

户检索所有违约用户的人物信息。通过使用钻取功能，用户可以发现出现频率最高的特征，从而深化对违约用户特性的理解。同时，系统提供了切取功能，使用户能够聚焦于其感兴趣的某一个或某几个特征。在此情境中，钻取功能与切取功能作为两种常见的数据故事探索工具，对于支持用户在复杂数据环境中定位并理解特定信息较为重要。

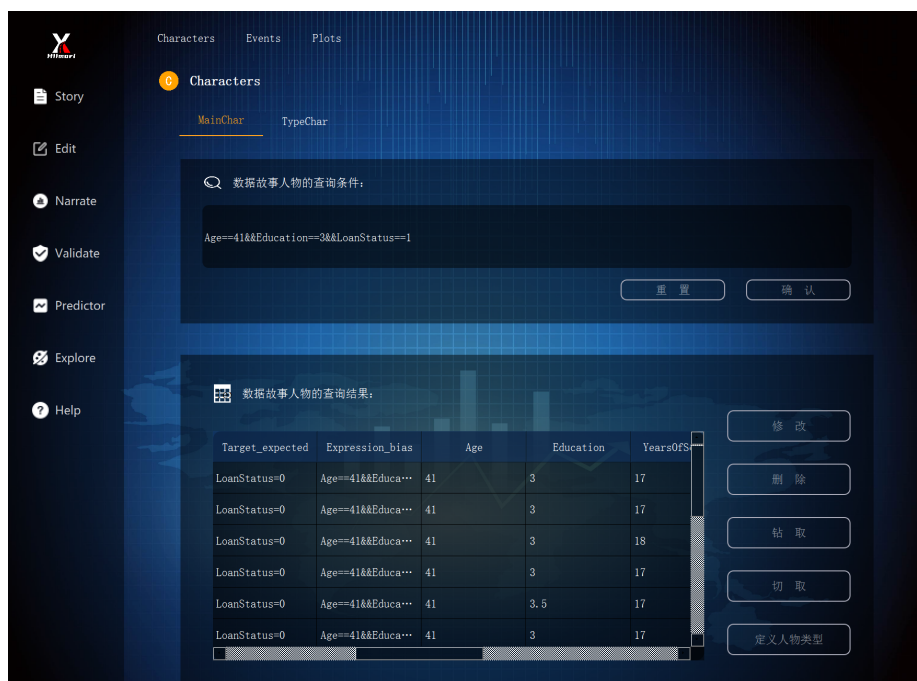


图6 查询数据故事人物界面

Figure 6 Query interface about data story character

(2) 主人公的操作。用户在图6所示的查询结果的基础上，点击【定义人物类型】按钮，在弹出的对话框中选择【设置为主人公】，即可将所选中的人物设置为主人公，并进入主人公的管理界面。主人公的管理界面设有主人公信息（特征、行为、目标、现实和偏见）和相关人物列表（同类人物、异类人物、正面人物和反面人物），见图7。主人公的管理功能提供主人公的修改、删除、旋转和基于主人公生成其他人物类型的功能。选中任一相关人物，点击【旋转】按钮，可将此人物设置为新的主人公。例如，在信用卡违约预测系统中，当选定特定的人物（如“张三”）后，该系统能够在用户界面上清晰地展示他的重要信息，包括特征、行为、目标、现实情况和偏见。偏见特征可以通过显示偏见表达式来揭示，用户也可以查询系统中存在的与指定人物相关的其他人物。

(3) 非主人公类人物的操作。根据3.1的人物类型划分思路，数据故事中的正面人物、反面人物、

同类人物、异类人物是相对于主人公而言的，因此，上述4类人物的生成需要指定对应的主人公。如图8所示，输入对应的主人公姓名，用户能够查询相关人物的全部信息。相对于主人公的操作功能，其他类型人物的校验操作相对复杂。目前，本系统中的人物校验功能主要实现了两类功能：①完整性校验：人物的属性信息中的必填特征是否完整以及每个人物是否能溯源到对应的主人公；②一致性校验：数据故事人物的钻取和切取功能可以产生衍生人物，将人物生成的先后顺序绘制树状结构，通过检查人物关系树中的每个节点是否只有一个父节点的方法判断人物数据的一致性。对于校验中发现的问题，系统提供了自动校正和直接删除两项功能。以信用卡违约预测系统为例，该界面使用户能够查询出所有已知用户的相关人物。通过观察正面人物和反面人物，用户可以得出如何改变特定特征可能降低违约概率的结论。同时，通过对同类人物与异类人物的观察，用户可以确定哪些特征是决定违约概率的关键因素。



图 7 数据故事的主人公的操作

Figure 7 Operations of protagonist in data stories



图 8 数据故事中的非主人公类人物的操作

Figure 8 Operations of non-protagonist in data stories

6 结语

数据故事化的主要目标包括数据认知、算法解释和现实与虚拟的融合，因此，它将成为智能应用中首选的数据表达方式。本研究首次系统性地探讨了数据故事人物的组成要素、基本类型、主要特征和核心操作，填补了相关理论研究的空白。故事事件的定义和故事情节的发展通过故事人物的特征、行为、目标、现实和偏见的叙述来实现，因此，本研究为数据故事情节以及数据故事事件研究奠定了理论基础。笔者

提出数据故事人物的 5 种基本操作：创建（create）、读取（read）、更新（update）、删除（delete，亦称为 CRUD 操作）、校验、钻取、切取和旋转，并基于反事实理论提出一种数据故事人物的自动生成方法。这种方法为数据故事的自动生成技术提供了新的思路 and 工具，对于数据故事的理论研究、自动化生成和工程化研发具有重要意义。此外，笔者针对数据故事中的不同类型人物提出各自的自动生成方法，并运用 Python 技术和面向对象的方法实现了这一理论框架。未来，这种方法可以推广应用到涉及预测性

分析的安全敏感领域, 如医疗诊断、灾害预警、信用评估等。在这些领域中, 面向预测性分析结果的数据故事可以帮助模型的使用者和非专业人士理解预测结果背后的逻辑, 提高他们对模型的信任度, 从而更有效地解决模型在应用中的“信任危机”。

在未来研究中, 需要继续解决数据故事人物的自动生成和工程化研发的若干关键问题: ①数据故事工具系统的进一步研发, 本文主要实现了人物的生成、修改、删除、钻取、切取、旋转等功能, 但是人物的钻取与切取功能有待进一步优化; ②数据故事人物的理论研究的进一步深入探讨, 包括进一步拓展故事人物的旋转操作的必要性和可能性, 探讨以特征为维度的旋转操作; ③数据故事的整体解决方案的提出。本文是数据故事化工程项目的子课题之一, 主要实现的是数据故事中的“人物管理器”模块。除了人物管理器模块, 数据故事化工程项目还包括情节管理器、事件管理器、模型代理器、故事渲染器、结果解释器、叙述交互及参数生成器等功能模块及子课题。为此, 如何将本子课题的工作与其他子课题有效集成, 实现代码复用、功能调用和数据共享也是值得进一步探讨的重要课题。

参考文献:

- [1] 朝乐门. 数据故事的自动生成与工程化研发 [J]. 情报资料工作, 2021, 42(2): 53-62. (CHAO L M. Automatic generation and engineering research & development of data stories[J]. Information and documentation services, 2021, 42(2): 53-62.)
- [2] 朝乐门. 数据故事化 [M]. 北京: 电子工业出版社, 2022: 54. (CHAO L M. Data storytelling: from data perception to data cognition[M]. Beijing: Publishing House of Electronics Industry, 2022: 54.)
- [3] FRIJDA N H. The laws of emotion[M]. New York: Psychology Press, 2017.
- [4] HOUGHTON K J, KLIN C M. Do readers remember what story characters remember [J]. Discourse processes, 2020, 57(1): 1-16.
- [5] HAMBY A, VAN LAER T. Not whodunit but whydunit: story characters' motivations influence audience interest in services[J]. Journal of service research, 2022, 25(1): 48-65.
- [6] SUN M, CAI L, CUI W, et al. Erato: cooperative data story editing via fact interpolation[J]. IEEE transactions on visualization and computer graphics, 2022, 29(1): 983-993.
- [7] LEE B, RICHE N H, ISENBERG P, et al. More than telling a story: transforming data into visually shared stories[J]. IEEE computer graphics and applications, 2015, 35(5): 84-90.
- [8] OJO A, HERAVI B. Patterns in award winning data storytelling: story types, enabling tools and competences[J]. Digital journalism, 2018, 6(6): 693-718.
- [9] ZHANG Y, LUGMAYR A. Designing a user-centered interactive data-storytelling framework[C] //Proceedings of the 31st Australian conference on human-computer-interaction. New York: Association for Computing Machinery, 2019: 428-432.
- [10] SHI Y, GAO T, JIAO X, et al. Breaking the fourth wall of data stories through interaction[J]. IEEE transactions on visualization and computer graphics, 2022, 29(1): 972-982.
- [11] HAROLD F. Story and discourse: narrative structure in fiction and film[J]. Poetics today, 1980, 1(3): 79.
- [12] ZHANG Y, REYNOLDS M, LUGMAYR A, et al. A visual data storytelling framework[J]. Informatics, 2022, 9(4): 73.
- [13] GUIDOTTI R, MONREALE A, RUGGIERI S, et al. A survey of methods for explaining black box models[J]. ACM computing surveys (CSUR), 2018, 51(5): 1-42.
- [14] WACHTER S, MITTELSTADT B, RUSSELL C. Counterfactual explanations without opening the black box: automated decisions and the GDPR[J]. Harvard. Journal of Law & Technology, 2018, 31(2): 841.
- [15] MILLER T. Explanation in artificial intelligence: Insights from the social sciences[J]. Artificial intelligence, 2019, 26(7): 1-38.
- [16] MOTHILAL R K, SHARMA A, TAN C. Explaining machine learning classifiers through diverse counterfactual explanations[C] //Proceedings of the 2020 conference on fairness, accountability, and transparency. New York: ACM, 2020: 607-617.
- [17] VERMA S, DICKERSON J, HINES K. Counterfactual explanations for machine learning: a review[J]. arXiv preprint, 2020, arXiv: 2010.10596.
- [18] MOLNAR C. Interpretable machine learning: a guide for making black box models explainable[M]. Seattle: Independently Published, 2022: 194.
- [19] 王明, 武文芳, 王大玲, 等. 生成链接树: 一种高数据真实性的反事实解释生成方法 [J]. 计算机科学, 2022, 49(9): 33-40. (WANG M, WU W F, WANG D L, et al. Generating link trees: a counterfactual explanations generation approach with high data reality[J]. Computer science, 2022, 49(9): 33-40.)
- [20] ADAM K D P B J. A method for stochastic optimization[J]. arXiv preprint arXiv: 1412.6980, 2014, 1412.
- [21] USTUN B, SPANGHER A, LIU Y. Actionable recourse in linear classification[C] //Proceedings of the conference on fairness, accountability, and transparency. New York: ACM, 2019: 10-19.
- [22] KOMMIYA MOTHILAL R, SHARMA A, TAN C. Explaining machine learning classifiers through diverse counterfactual explanations[J]. arXiv e-prints, 2019, arXiv: 1905.07697.
- [23] POYIADZI R, SOKOL K, SANTOS-RODRIGUEZ R, et al. FACE: feasible and actionable counterfactual explanations[C] //Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. New York: ACM, 2020: 344-350.
- [24] SMYTH B, KEANE M T. Good counterfactuals and where to

- find them: a case-based technique for generating[C]//ICCBR 2020: case-based reasoning research and development. Cham: Springer, 2020: 163-178.
- [25] GOYAL Y, WU Z, ERNST J, et al. Counterfactual visual explanations[J]. arXiv preprint, 2019, arXiv: 1904.07451.
- [26] WACHTER S, MITTELSTADT B, RUSSELL C. Counterfactual explanations without opening the black box: automated decisions and the GDPR[J]. Harvard journal of law & technology, 2018, 31(2): 841-887.
- [27] LAUGEL T, LESOT M J, MARSALA C, et al. Issues with post-hoc counterfactual explanations: a discussion[J]. arXiv preprint, 2019, arXiv: 1906.04774.
- [28] MAHAJAN D, TAN C, SHARMA A. Preserving causal constraints in counterfactual explanations for machine learning classifiers[J]. arXiv preprint, 2019, arXiv: 1912.03277.
- [29] GUIDOTTI R, MONREALE A, GIANNOTTI F, et al. Factual and counterfactual explanations for black box decision making[J]. IEEE intelligent systems, 2019, 34(6): 14-23.
- [30] GUIDOTTI R, MONREALE A, GIANNOTTI F, et al. Factual and counterfactual explanations for black box decision making[J]. IEEE intelligent systems, 2019, 34(6): 14-23.

作者贡献说明:

朝乐门: 研究设计、论文撰写及实验设计;

刘慧: 完成实验与讨论部分;

张天怡: 论文校对与修改;

李泽仑: 论文校对与修改。

Study on Character Type and Automatic Generation Method in Data Story

Chao Lemen^{1,2} Liu Hui² Zhang Tianyi² Li Zelun²

¹Key Laboratory of Data Engineering and Knowledge Engineering (Renmin University of China), Beijing 100872

²School of Information Resource Management, Renmin University of China, Beijing 100872

Abstract: [Purpose/Significance] Characters and narratives form the dual pillars of data stories. The narrative trajectory of a data story is shaped by the characters' traits, behaviors, goals, realities, and biases. Achieving automatic generation of data story characters stands as a core scientific research topic in the realm of data storytelling. Addressing this central issue carries significance for the theoretical exploration, automation, and engineering-oriented research and development in the domain of data stories. **[Method/Process]** Initially, this study delved into the types, attributes, and operations associated with data story characters. Subsequently, it proposed a character generation technique based on counterfactual reasoning, offering automatic generation algorithms for protagonists, similar characters, heterogeneous characters, positive characters, and negative characters in data stories. Following this, it dissected and discussed its technical implementation, furnishing the paper's experimental design, data sources, method selection, and result discussion. Lastly, it encapsulated the principal research findings of the paper and furnished a forward-looking perspective. **[Result/Conclusion]** In the data storytelling, this paper introduces for the first time the compositional elements, basic types, principal features, and core operations of data story characters. Furthermore, this study presents an automatic generation method for data story characters rooted in counterfactual reasoning.

Keywords: data story data storytelling story characters characters