

人工智能技术应用风险规避探析

张鹏然

(广州市技师学院, 广东 广州 510000)

摘要: 人工智能技术的迅猛发展, 为推动社会、经济和科技创新带来了巨大的机遇与挑战。在人工智能技术的应用过程中, 我们必须认识到一些潜在的风险和挑战。从数据安全、隐私保护、智能算法不公的角度, 我们需要全面分析人工智能技术的安全风险, 并研究各种有效的规避策略。为了更好地应对这些挑战, 需要从多个角度出发, 探索出有效的规避策略, 如严格管理和规范数据的来源和使用, 加强数据安全保护, 建立隐私保护机制, 制定人工智能算法的公正性准则等。同时, 也需要加强政策法规的制定和完善, 规范人工智能技术的应用, 促进人工智能可信共识的形成。只有这样, 才能让人工智能技术更好地服务于人类社会, 带来更多的益处。

关键词: 人工智能; 应用风险; 规避

一、人工智能安全风险分析

(一) 网络安全风险分析

网络安全风险是指在人工智能技术应用过程中, 由于网络环境的不稳定和机器学习算法的不确定性, 可能出现的安全漏洞和风险。具体来说, 主要包括以下几个方面: 人工智能技术需要大量的数据传输和存储, 因此网络安全风险是人工智能的一个重要问题。比如, 黑客攻击、网络钓鱼以及其他网络攻击都可能导致数据泄露、机密信息泄露或者数据被篡改, 从而影响到人工智能模型的准确性和稳定性; 人工智能技术对于网络稳定性和可靠性的要求较高。例如, 如果人工智能应用程序的算法逻辑出现错误, 或者网络延迟过高, 都可能导致程序出现错误或者崩溃, 从而影响到人工智能应用的正常运行; 在人工智能应用过程中, 需要涉及到一些敏感信息。这些信息可能包含企业机密、个人隐私等, 极易受到黑客攻击和网络安全漏洞的威胁, 这也是网络安全风险的一个重要方面; 人工智能应用通常是面向大众的, 而大众对于网络安全的意识普遍较低, 这也增加了人工智能应用的网络安全风险。由于用户的安全意识不足, 可能会泄露账户信息、密码等重要信息, 这会导致用户财产受到损失, 同时也会间接影响到人工智能应用的稳定性和信誉度。

(二) 数据安全风险分析

数据安全风险分析是人工智能领域中非常重要的一部分, 因为人工智能依赖于大量的数据来进行学习和预测, 因此数据的安全性就显得尤为重要。下面将从以下三个方面来详细分析数据安全风险分析。第一, 数据采集和存储过程中的安全风险。在数据采集和存储的过程中, 数据可能会受到黑客攻击或者数据泄露的风险。为了避免这种情况, 需要采用严格的控制措施来保证数据的安全, 比如使用加密技术来保护数据传输和存储过程中的安全性。第二, 数据共享过程中的安全风险。在人工智能领域中, 数据的共享是非常重要的, 因为只有大量的数据才能让人工智能算法更加准确地进行预测和学习。但是, 数据共享的过程中也会存在安全风险, 比如数据泄露或者恶意攻击等。为了避免这种情况, 需要加强数据的权限管理和安全机制, 确保仅有授权人员才能够使用数据。第三, 人工智能算法的安全风险。在人工智能算法中, 可能会存在某些算法会对数据进行非法操作或者通过算法的方式将数据泄露出去的风险。为了避免这种情况, 需要对算法进行安全性评估, 并采取相应的防范措施来确保算法的安全性。

(三) 信息安全风险分析

信息安全风险是人工智能技术应用中一个重要的问题。随着人工智能的发展, 信息安全风险也日益凸显。在人工智能应用中,

信息安全风险可以分为以下几类: 一, 隐私泄露风险是人工智能应用中较为常见的信息安全风险之一。在人工智能算法中, 用户的个人信息可能会被收集、存储、分析和共享, 这可能会导致用户的隐私泄露。因此, 需要对用户的隐私进行有效的保护, 采取适当的措施来保证数据的安全性。二, 算法歧视风险是一个比较新的信息安全风险, 由于人工智能算法的运作依赖于大量的数据, 如果数据存在歧视性, 那么人工智能算法也会存在歧视性。例如, 在职场招聘中, 如果人工智能算法依据性别或者种族等因素来筛选候选人, 那么就会造成歧视。因此, 需要对算法进行有效的分析和控制, 尽可能避免算法歧视。三, 恶意攻击风险是人工智能应用中最严重的信息安全风险之一。恶意攻击可以采用多种方式, 例如病毒攻击、网络钓鱼、人工智能欺骗(例如deepfake视频), 这些攻击不仅会导致信息泄露, 还可能使得整个人工智能系统受到破坏。因此, 需要对整个系统进行有效的保护和防御, 采取多种技术手段和措施来防范恶意攻击。

(四) 算法安全风险分析

算法安全风险在人工智能技术应用中具有重要的意义。人工智能算法作为人工智能应用的核心, 其安全性直接关系到整个系统的可靠性和稳定性。算法安全风险主要包括以下几个方面: 一方面, 算法漏洞风险。由于复杂的算法模型和海量的数据输入, 人工智能算法很容易受到攻击者的针对性攻击。攻击者可以通过在输入数据中注入恶意代码或者进行特定的干扰, 来利用算法的漏洞实施攻击。这种算法漏洞可能导致人工智能系统输出错误的结果, 甚至造成系统崩溃。另一方面, 算法隐私泄露风险。人工智能算法在训练和运行过程中需要大量的数据支持, 而这些数据往往包含着用户的隐私信息。如果算法的设计、实现或者运行中存在隐私泄露的漏洞, 攻击者就有可能获取到用户的敏感信息。

二、人工智能技术应用风险规避策略

(一) 出台政策法规, 加强对突出问题的规制

1. 针对人工智能信息安全风险突出问题的规制

人工智能技术应用的一个重要问题是信息安全。因为人工智能技术需要处理大量的数据, 其中可能包含有隐私信息。同时, 人工智能技术也可能被黑客攻击, 导致数据泄露和其他安全问题。为了规避这些风险, 政府应该颁布相关法规, 并要求企业和机构在应用人工智能技术时遵循这些法规。在制定法规时, 应考虑以下几个方面: 安全标准, 政府应该建立与人工智能技术相关的安全标准, 以确保所有涉及到的企业和机构都能遵循这些标准。安全标准应该包括对数据和系统的保护措施, 防止黑客攻击和数据泄露; 数据安全性, 政府应该规定企业和机构必须采取措施确

保数据的安全性,比如数据加密和备份。此外,政府还应该规定企业和机构必须保护用户隐私,不得擅自使用用户数据;惩罚措施,政府应该颁布相应的法规,规定违反安全标准和数据安全性规定的企业和机构将受到何种惩罚,以保证企业和机构能够遵循这些规定。

2. 针对人工智能数据安全突出问题的规制

人工智能技术的应用需要大量的数据支持,因此数据安全成为人工智能技术应用过程中另一个重要问题。数据安全的风险主要来自于数据存储、传输和使用的过程。对于数据的存储,企业和机构应该采用安全的存储方式,比如加密存储和备份存储。对于数据的传输,企业和机构应该采用安全的传输方式,比如SSL证书、VPN等。对于数据的使用,企业和机构应该限制数据的访问权限,只有授权的人员才能处理这些数据。政府应该颁布相关法规,规定企业和机构必须遵循安全标准,并要求企业和机构采用安全的存储和传输方式。此外,政府还应该规定企业和机构必须限制数据的访问权限,只有授权的人员才能访问这些数据。

3. 针对人工智能算法安全风险突出问题的规制

人工智能技术的核心是算法。在人工智能技术的应用过程中,算法可能出现各种问题,比如算法的偏差、算法的错误、算法的歧视等。为了规避这些风险,政府应该颁布相关法规,规定企业和机构采用合理的算法,避免算法的偏差和错误。此外,政府还应该规定企业和机构必须保证算法的公正性,不得歧视某些群体或个人。此外,政府还应该要求企业和机构公示使用的算法,以便监督和评估算法的效果和公正性。

(二) 制定准则、规范和指南,促进发展可信人工智能共识形成

1. 制定人工智能管理、治理相关框架、原则和准则

随着人工智能技术应用范围的扩大,管理、治理等问题也日益凸显。为了规范人工智能技术的应用,保障人类的利益和安全,需要制定相关框架、原则和准则。人工智能管理和治理的框架建立应该从以下几个方面入手:(1)建立法律法规:需要制定专门的法律法规,明确人工智能技术的应用范围、责任和义务,保障人类的利益和安全。(2)建立监管机制:建立人工智能技术应用的监管机制,加强对人工智能技术应用的监管和管理。(3)建立伦理准则:制定人工智能伦理准则,明确人工智能技术应用的道德和伦理标准,保护用户隐私。

2. 制定人工智能设计、开发、评估、应用相关准则和规范

人工智能技术的设计、开发、评估和应用都需要遵循相关的准则和规范。为了规范人工智能技术的应用,保障人类的利益和安全,需要制定相关准则和规范。以下从四个方面展开说说:(1)人工智能设计准则:制定人工智能设计准则,规范人工智能技术的设计过程,提高人工智能技术的可控性和可靠性,减少人工智能技术的安全风险。(2)人工智能开发规范:制定人工智能开发规范,明确开发要求和标准,提高人工智能技术的可靠性和安全性。例如,要求开发人员必须保障人工智能技术的公正、公平和透明性,不得歧视任何人群和集体。(3)人工智能评估指南:制定人工智能评估指南,明确评估要求和标准,提高人工智能技术的可靠性和安全性。例如,要求评估人员必须保障人工智能技术的公正、公平和透明性,减少评估中的主观性和偏见性。

(三) 探索创新技术手段,积极应对人工智能安全威胁

1. 探索应对人工智能网络安全问题的技术

目前网络安全问题是人工智能技术面临的主要挑战之一。网

络攻击者可以通过各种手段对人工智能系统进行攻击,如恶意软件、Botnet攻击、DoS攻击等。因此,对于人工智能系统的网络安全问题需要采取相应的措施。主要包括以下方面:(1)强化体系结构设计。人工智能系统的体系结构设计需要考虑到网络安全这一要素。安全性需要在整个设计过程中得到重视,在系统的各个层面都要考虑到安全策略。例如,在系统实现初期需要考虑安全功能,防范未来的安全攻击。另外,对于人工智能系统的网络安全问题,建议对客户端和服务器两个方面进行攻击检测,以便及早发现安全漏洞。(2)加强身份验证。身份验证是保护人工智能系统不受未经授权的访问的关键。传统的身份验证技术已经无法满足人工智能系统的要求。因此,新的身份验证技术需要与智能计算相结合,并充分考虑到人工智能系统的特点。例如,可以使用机器学习算法来识别恶意访问,并分析其特征和模式,以识别安全威胁。(3)加强数据加密。加强数据加密是确保人工智能系统的安全性的关键。现有加密技术需要进一步优化,以适应人工智能技术的发展。例如,在数据传输和存储过程中,可以使用加密算法对数据进行加密,保证数据的机密性和完整性。

2. 探索应对人工智能信息安全问题的技术

信息安全问题是人工智能技术应用中另一个需要解决的重要问题。信息安全包括保护数据的保密性、完整性和可用性等方面。基于这些方面,探索应对人工智能信息安全问题的技术需要包括以下方面:(1)强化身份验证和访问控制。身份验证和访问控制是保护人工智能系统对数据的保密性和完整性的关键因素。这需要采取适当的身份验证技术和访问控制机制。在身份验证方面,可以使用生物特征识别技术和基于机器学习的身份验证技术。在访问控制方面,可以采用基于策略的访问控制和基于角色的访问控制方法,以保障系统对数据的保护。(2)采用加密技术。加密技术是保护数据机密性的重要手段。在人工智能技术应用中,需要找到一种更加适合人工智能数据的加密技术。可以利用机器学习方法进行加密,使得加密算法与人工智能技术相协同,提高加密效率和性能。(3)强化数据清洗和去重技术。数据清洗和去重技术是保障人工智能系统对数据的准确性和完整性的重要手段。数据清洗可以过滤掉不符合数据要求的信息,降低数据噪音;数据去重可以删除重复数据,提高数据质量。在人工智能技术应用中,应加强对数据清洗和去重技术的研究,以保证人工智能系统对数据的准确性和完整性。

三、结语

人工智能技术是当今最具前沿和潜力的技术之一,在各个领域都有着广泛的应用。但同时,它也孕育着一系列的安全风险,这些风险有可能对我们的社会、经济、政治等多个方面造成严重影响。因此,出台政策法规、制定准则规范、探索创新技术手段,将是我们有效规避人工智能安全风险、推动人工智能技术发展的重要举措。

参考文献:

- [1] 赵精武.生成式人工智能应用风险治理的理论误区与路径转向[J].荆楚法学,2023(03):47-58.
- [2] 张铤.人工智能嵌入社会治理的风险及其规避[J].浙江工商大学学报,2022(03):120-126.
- [3] 潘援.人工智能技术在媒体应用中的风险防范研究[J].中国记者,2022(02):113-114.