

# 人工智能时代的数据安全挑战及应对

## ——《人工智能安全:原理剖析与实践》

数据作为驱动人工智能技术发展的核心要素,安全问题不容忽视。在人工智能时代,数据的收集、存储、处理和传输都面临着前所未有的安全挑战。从个人隐私泄露到企业商业机密失窃,再到国家安全威胁,数据安全已成为全球共同关注的焦点。基于此,人工智能时代的数据安全,对于保护个人隐私、维护企业利益、保障国家安全具有重要意义,亦可为推动人工智能技术的健康发展和应用提供有力支撑。

随着人工智能技术的飞速发展,《人工智能安全:原理剖析与实践》一书应运而生,为读者提供了全面、深入的人工智能安全知识和实践指导。该书不仅涵盖信息安全和人工智能的基础知识,还从多个角度深入剖析了人工智能安全,展现了人工智能安全研究的丰富内涵和广阔前景。编者首先介绍了信息安全和人工智能的基本概念,为读者后续的学习打下坚实基础。然后从人工智能信息系统安全和算法安全2个层面,详细阐述了人工智能安全的研究方法和手段。信息系统安全部分涵盖了环境安全、基础架构安全、依赖库安全以及云服务安全等方面,通过丰富的实例和程序代码,展示了如何在实际环境中发现和应对人工智能系统的安全漏洞;算法安全部分则深入探讨了对抗攻击和防御策略,包括白盒和黑盒场景下的对抗攻击方法以及多种对抗防御手段,为读者提供了全面的算法安全知识和实践指导。最后,该书还关注了人工智能辅助攻防新场景,如自动化漏洞挖掘、生成对抗网络在安全领域的应用等,展现了人工智能技术在安全领域的创新应用和巨大潜力。通过实验环节,如为视频中的人物换脸等,让读者亲身体验了人工智能技术的魅力和风险,增强了读者的实践能力和安全意识。该书内容全面且注重实践,通过大量实例和程序代码让读者掌握了人工智能安全的知识和技能,同时,前瞻性强,关注最新动态,语言通俗易懂,结构清晰,为读者提供了前沿且易于理解的学习体验。通过学习本书,读者能够对人工智能安全有更深入的了解和认识,并为人工智能安全的研究和实践做出更大贡献。

编者指出,合规性与监管缺失是人工智能发展面临的现实问题,引导人工智能企业在遵守法律法规的前提下开展业务,对社会稳定和人工智能技术可持续发展有重要意义。合规性是企业运营的基础,也是企业信誉的保障。在人工智能领域,合规性要求企业在收集、存储、处理和共享数据时必须遵守相关法律法规,尊重用户的隐私和数据安全。然而,由于人工智能技术的复杂性和创新性,相关的法律法规往往难以覆盖所有情况,给企业带来了不确定性和风险。尽管各国政府都在加强对人工智能技术的监管,但由于技术发展迅速,监管机构和手段往往难以更进。监管的滞后和缺失,使得一些不法分子有机可乘,利用人工智能技术从事非法活动,给企业和社会带来了严重损失。为应对这一挑战,企业需要采取一系列措施来加强合规性和监管。首先,企业需要密切关注相关法律法规的动态变化,及时了解并遵守最新的法律法规要求,人工智能技术开发与应用企业可建立专门的法律团队,负责跟踪和研究相关法律法规,为企业提供法律咨询和指导;其次,企业需要建立完善的合规管理制度和流程,包括制定详细的数据收集、存储、处理和共享规范,明确各部门和人员的职责和权限,确保企业在运营过程中始终遵守相关法律法规,并加强对员工的合规培训和教育,提高员工的合规意识和能力;然后,企业需要与监管机构建立良好的沟通机制,及时汇报和反馈相关问题,共同推动人工智能技术的健康发展和应用;最后,企业需积极参与行业标准的制定和推广,通过参与行业标准的制定,企业可以推动行业向更加规范、安全的方向发展,通过推广行业标准,促进整个行业的合规性和监管水平的提升。

笔者认为,在人工智能时代,数据泄露与隐私保护的安全事故时有发生,敲醒了人工智能的安全警钟。人



书名:人工智能安全:原理剖析与实践

编者:王琦,朱军,王海兵

出版社:电子工业出版社

ISBN:9787121445033

出版时间:2022年12月

定价:128元

人工智能系统为了提升准确性和效率,往往需要训练和学习大量的数据。然而,近年来数据泄露事件频频发生,这些数据涵盖了用户的个人信息、企业的敏感数据,甚至包括国家安全的机密信息。不仅给个人和企业带来了严重威胁,也对国家安全造成了潜在风险。为了应对这一挑战,人工智能企业需要采取一系列措施加强数据加密和访问控制,确保数据的安全性和隐私性。措施1:采用先进的加密技术保护数据,通过利用强加密算法和密钥管理技术来加密处理数据,使得未经授权的人员无法获取敏感数据的明文内容,如此即使数据在传输或存储过程中被截获或窃取,攻击者也无法轻易解密和获取数据的真实内容。措施2:建立完善的访问控制机制也是防止数据泄露的重要手段,通过身份认证和权限管理,确保只有授权人员才能够访问敏感数据,包括验证用户的身份,以及对不同用户设置不同的访问权限。措施3:除了加密和访问控制外,人工智能企业还需要建立完善的数据管理体系来确保数据的安全性和隐私性,包括严格监管数据的采集、存储、使用和共享。在数据采集阶段,需要明确告知用户数据的收集目的和使用范围,并获得用户的明确授权;在数据存储阶段,需要采用可靠的存储设备和备份策略,确保数据的完整性和可用性;在数据使用阶段,需要遵循最小必要原则,只收集和使用与业务相关的必要数据,避免过度收集和滥用用户信息;在数据共享阶段,需要与合作伙伴签订严格的保密协议,确保数据在共享过程中不被泄露或滥用。措施4:为进一步提高数据的安全性,还可采用一些先进的技术手段来增强保护效果,如利用差分隐私技术,在保护用户隐私的同时提供有用的统计信息;利用同态加密技术在不暴露原始数据情况下计算和处理数据等。

笔者强调,随着人工智能技术的广泛应用,高度智能化的系统也成为了黑客和恶意攻击者的新目标,黑客利用各种手段,如注入恶意代码、篡改数据等,试图破坏系统的正常运行,窃取敏感信息,甚至制造严重的安全事故。恶意攻击和数据篡改对人工智能系统构成了巨大威胁,攻击者可能会通过精心设计的恶意代码,利用系统漏洞或弱密码等途径,侵入系统内部,获取敏感数据或篡改关键信息。如在自动驾驶汽车中,攻击者可能会篡改传感器数据,导致车辆做出误判,从而引发交通事故;在智能家居系统中,攻击者可能会控制智能设备,窃取用户的隐私信息或制造混乱。对人工智能系统的攻击不仅对个人隐私和企业安全造成严重威胁,也给社会稳定和国家安全带来了潜在风险。为了应对这一挑战,需要构建强大的安全防护体系来实时监测和防御外部攻击。第一,防火墙是保护系统安全的第一道防线,通过配置合理的防火墙规则,可以过滤掉大部分恶意流量和攻击行为,确保只有合法的请求能够访问系统。入侵检测系统(IDS)可以实时监控网络流量和系统日志,发现异常行为并及时报警。通过结合防火墙和IDS等安全设备,可构建一个多层次、立体化的安全防护体系,有效抵御外部攻击。第二,除实时监测和防御外部攻击外,企业还需要加强对人工智能系统的安全审计和漏洞扫描。定期对系统进行全面的安全审计和漏洞扫描是发现潜在安全隐患的重要手段,通过模拟攻击者的行为测试系统,可发现系统存在的漏洞和弱点,并及时修复问题。同时,建立完善的安全漏洞管理制度和应急响应机制必不可少,一旦发现安全漏洞或遭受攻击,需要迅速响应并采取修复和处置措施,确保系统的安全稳定。第三,提高系统的容错性和鲁棒性也是应对恶意攻击和数据篡改的重要手段。在设计人工智能系统时,需要充分考虑系统的容错性和鲁棒性,确保在遭受攻击或出现异常时能够迅速恢复正常运行。例如,企业可采用分布式架构和冗余设计来提高系统的可用性和可靠性。此外,加强对系统日志的收集和分析也有助于及时发现并应对潜在的安全威胁。

综上,在人工智能时代,数据安全的挑战日益严峻,从合规性与监管缺失,到数据泄露与隐私保护,再到恶意攻击与数据篡改,每一个问题都关乎着个人、企业和国家的核心利益。为了人工智能技术的可持续健康发展,人工智能企业与国家主管部门必须时刻保持警惕,不断加强技术研发和制度建设,以应对这些挑战。从加强数据加密和访问控制,到构建强大的安全防护体系,再到密切关注法律法规的动态变化并建立完善的合规管理制度,每项措施都是对数据安全的有力保障。然而,数据安全并非一劳永逸,而是需要持续关注和努力。随着技术的不断进步和攻击手段的不断翻新,人工智能行业必须时刻保持敏锐的洞察力和前瞻性,不断完善和提升安全防护能力。

(张子挺/淄博师范高等专科学校/讲师;马欣/淄博师范高等专科学校/讲师)