ISSN:2705-1323（O）2737-4149（P）

基于 Python 的机器学习算法在网络安全威胁检测中的应用研究

郑海鹏

淮南联合大学　安徽淮南　232038

摘　要：随着网络攻击方式的不断演化，传统安全防护手段难以有效应对多样化的网络威胁。本文围绕 Python 语言下的机器学习算法在网络安全威胁检测中的应用展开探讨，重点分析了异常流量识别、恶意软件检测、钓鱼攻击防控与误报率控制等关键问题。通过梳理主流算法在安全检测中的技术路径，结合 Python 的编程优势，提出构建智能化、自动化防护体系的可行方案。研究表明，基于机器学习的安全检测手段具有良好的适应性与扩展性，为提高网络系统的防御能力提供了有力支持。

关键词：Python；机器学习算法；网络安全；网络异常检测；应用

引言

在信息技术快速发展的背景下，网络系统广泛应用于各行各业，网络安全问题也日益突出。传统基于规则的安全检测方法已难以应对日益复杂、多变的威胁场景。机器学习凭借其强大的数据分析和模式识别能力，为网络异常检测提供了更智能、高效的解决方案。Python 语言作为主流开发工具，拥有丰富的机器学习库和良好的可扩展性，广泛应用于安全领域。本文围绕基于 Python 的机器学习算法在网络安全中的应用展开探讨，重点分析其在异常检测、恶意软件识别、钓鱼网站防控与误报率优化等方面的实践路径，旨在为构建智能化网络安全防护体系提供理论参考和技术支持。

1 当前网络安全威胁状况

随着信息技术的迅速发展和网络应用的广泛普及，网络安全威胁呈现出类型多样化、手段智能化和攻击频繁化的趋势，严重影响着企业、政府机构及普通用户的数据信息安全与系统稳定性。目前，常见的网络安全威胁主要包括以下几类：

首先是跨站脚本攻击（XSS），攻击者通过在网页中嵌入恶意脚本，使用户在无感知的情况下执行攻击代码，从而泄露个人信息或被远程控制。该攻击方式分为反射型、存储型和基于 DOM 的三种形式。其次是 SQL 注入攻击，攻击者通过在输入字段中植入恶意 SQL 语句，绕过身份验证访问数据库，实现数据篡改、窃取甚至控制系统，常见于安全验证薄弱的 Web 应用中。第三是跨站请求伪造攻击（CSRF），黑客利用用户已登录状态诱导其发起恶意请求，在未经授权的情况下执行如修改密码、转账等敏感操作，造成严重后果。最后是拒绝服务攻击（DoS），通过发送大量无效请求，消耗目标服务器资源，致使正常用户无法访问，甚至造成系统崩溃。

2 基于 Python 的机器学习算法对网络威胁的检测

2.1 机器学习算法的网络异常检测

在网络安全领域，机器学习算法通过构建模型对海量网络数据进行分析和挖掘，能够有效识别潜在的异常行为，从而提前预警可能的安全威胁。异常检测的核心在于对系统中偏离常态的数据点进行识别，这些异常数据常常与诸如欺诈行为、系统故障、恶意入侵等事件密切相关。

在基本方法上，变量异常检测是最常见的思路之一。当研究对象仅为两个变量时，可通过数据可视化手段识别分布之外的异常点；但随着变量维度增加至十个甚至上百个，仅靠视觉手段就变得力不从心。因此，实际应用中通常会采用状态监控方法，即设定各变量的阈值区间，若某项指标超出正常范围即触发警报。此外，还可采用概率统计方法对数据集进行建模，依据每个数据点出现的概率 $p(x)$ 与设定阈值 r 进行比对，若 $p(x) < r$，即可判定该行为为异常。

更进一步的异常检测方式是借助自动编码器神经网络（Autoencoder）。自动编码器是一类无监督学习模型，其基本结构包括输入层、多个隐藏层和输出层，通过将高维输入数据压缩重构，实现特征降维并保留关键信息。在网络安全中，自动编码器可用于学习正常网络行为的特征分布。一旦网络状态偏离正常模式，其重构误差将显著增大，从而有效

地识别出潜在的异常行为。例如，在设备监控场景中，温度、压力、振动等传感器信号在运行异常时的联动变化会反映在编码器重构失败的误差中，从而作为预警信号。

2.2 机器学习算法的恶意软件检测

与传统基于特征库或规则匹配的检测方式不同，机器学习算法在恶意软件检测中展现出更强的泛化能力与智能识别能力。它能够基于样本行为模式识别出新型、变异或此前未知的恶意软件。

其中，图像检测算法是一种新兴且高效的方法。该方法的核心是将恶意软件的二进制数据转化为灰度图像，再借助图像分类的深度学习模型（如 CNN）进行恶意与良性文件的二分类处理。具体操作过程包括：首先读取恶意软件的二进制内容，每 8 位作为一个像素值转换为 0 至 255 之间的整数；然后将像素序列重构为固定尺寸的图像（如 256 × 256），并在图像不足 64KB 时以 0 进行补齐。通过图像分类模型训练后，能够快速、准确地识别恶意软件图像的特征模式。该方法不仅效率高，而且适用于大规模样本自动化检测。

另一种常用方法是 Logistic 回归算法（逻辑回归），其本质是一种二分类模型，广泛应用于判断文件是否为恶意样本。逻辑回归通过对特征向量进行加权求和，并将结果输入到 Sigmoid 函数中，将线性输出映射至 0 到 1 之间的概率值。模型训练过程中，利用最大似然估计法对模型参数进行优化，最终实现最优分类边界的构建。Logistic 回归虽然模型结构简单，但具有良好的可解释性与稳定性，适合对恶意软件文件特征向量进行快速筛查，尤其适合与其他算法共同构建混合检测体系。

2.3 机器学习算法的网络钓鱼检测

网络钓鱼是一种伪装成可信网站或平台以骗取用户敏感信息的攻击方式，具有高度隐蔽性和迷惑性。传统防护手段常依赖黑名单或规则库更新，但面对快速变化的钓鱼手段，这些方法响应滞后、效果有限。而机器学习算法通过对大量历史钓鱼样本的学习，能够从数据中提取深层特征，有效提升识别的准确性与灵活性。

网络钓鱼检测的第一步是特征提取。模型通过分析 URL 结构、域名信息、网页内容、跳转行为等因素，挖掘钓鱼网站与正常网站之间的细微差异。例如，钓鱼链接常包含混淆域名、过长参数或与真实网站结构不符的路径，这些都可作为重要判别依据。

在建模阶段，监督学习是应用最广泛的方法之一。通过对已标注的钓鱼与非钓鱼样本训练模型，系统能逐步学会区分两类网页。在此基础上，集成学习方法进一步提高了检测准确率，它将多个分类器的判断结果整合起来，以投票或加权方式得出更稳健的判断。

为了实现更加全面的防护，除了模型构建，还可结合模型训练、用户教育与实时监测三位一体的策略。一方面，通过持续收集钓鱼数据训练模型，可不断提升系统对新型钓鱼手段的感知能力；另一方面，通过机器学习辅助设计个性化安全提示与培训内容，有助于增强用户识别钓鱼攻击的意识和能力；同时，借助实时监测系统分析用户网络行为，一旦发现可疑操作即刻预警，最大限度降低风险发生率。

2.4 机器学习算法的威胁应对

在面对日益严峻的网络安全挑战，机器学习技术不仅用于检测威胁，更逐渐成为构建主动防御体系的关键力量。它可以通过分析历史数据和实时网络流量，识别异常行为、预测潜在风险，并实现自适应调整与自动化响应。

具体而言，机器学习模型在检测恶意代码、识别非法入侵、过滤垃圾邮件等方面表现出色。通过构建入侵检测系统，模型能够实时分析网络传输的数据，识别并分类正常与异常行为，并在检测到可疑行为时及时发出预警。同时，在电子邮件过滤中，机器学习算法通过学习垃圾邮件的语言模式、发送频率和结构特征，可以高效识别并拦截大量未知类型的垃圾信息。此外，模型训练过程中可选择监督学习、无监督学习或强化学习等不同方式，以适应不同的数据结构和应用场景。在系统实际部署中，还可结合特征提取、数据清洗和结果反馈，形成动态更新和持续优化的安全响应机制。最终目标是构建一个自学习、自适应、自防御的智能安全体系，实现从“被动防守”向“主动预防”的战略转型。

2.5 机器学习算法减少误报

在网络安全实践中，误报（即将正常行为误判为威胁）问题始终是检测系统优化的重点之一。误报率高不仅会增加管理员负担，也可能导致真正的威胁被忽视。为此，机器学习算法通过多种策略来降低误报风险，提升检测系统的精准度和实用性。

一是通过调整分类阈值，使模型在不同场景下保持合适的灵敏度。较低的阈值有助于发现更多潜在威胁，但也可

能带来更多误报；较高的阈值则可减少误报，但可能漏掉部分攻击。因此，需结合实际风险容忍度灵活设定。

二是处理数据不平衡问题。在网络安全中，正常数据远多于异常数据，模型易偏向多数类，导致异常行为被忽视。通过欠采样、过采样或合成少数类样本等技术手段，可以改善类别比例，增强模型对异常样本的识别能力。

三是加强特征工程，去除冗余或无效特征，并提炼出与安全威胁高度相关的关键属性。良好的特征选择不仅提高模型训练效率，也显著降低误判风险。

四是根据实际应用需求优化算法选择和模型参数。不同模型对误报率的敏感度不同，某些算法如支持向量机、随机森林在安全检测中的表现更加稳健。通过实验和调优，可找到性能最优的组合方式。

2.6 机器学习算法的可扩展性

随着数据规模的迅猛增长与攻击类型的日益复杂，网络安全系统面临性能和扩展能力的双重考验。相比传统安全模型，机器学习算法在处理大规模、多样化数据时展现出更强的可扩展性。这种可扩展性体现在：①模型可通过分布式计算框架进行训练与部署，将海量数据分散到多个节点处理，从而显著缩短训练时间并提升响应速度；②模型在面对持续新增的训练样本或新类型攻击时，可通过增量学习方式动态更新，保持检测能力的实时性和前瞻性。此外，通过模型压缩等技术手段，如参数剪枝、量化和蒸馏等，可以大幅减少模型体积和运算资源占用，提升在边缘设备或嵌入式系统中的应用能力。这使得机器学习算法不仅适用于大型服务器平台，也适用于云环境、终端设备等多种复杂运行场景。

随着企业对云计算平台的依赖程度不断加深，保障云环境中的数据安全已成为网络安全的重中之重。传统防护手段难以全面应对分布式架构下多变的安全威胁，而机器学习技术凭借强大的数据处理和自适应能力，为构建智能化云安全体系提供了有效路径。

在实践中，可通过部署数据采集代理和日志监控系统，实时获取云平台中的访问行为、程序操作和网络流量等关键数据。随后，借助机器学习算法对数据进行清洗、特征提取和模型训练，快速识别潜在风险。例如常用的无监督学习方法如 K 均值聚类，能够有效划分异常用户行为和正常访问模式，从而提升入侵检测能力。此外，通过参数优化与模型自学习机制，机器学习还能适应多租户云环境的动态变化，实现跨平台、跨节点的统一安全管控。未来，云安全将更加依赖智能算法驱动，提升防护效率与策略响应的智能化水平。

3 结束语

随着网络攻击方式日益复杂，传统安全防御手段已难以满足现代网络环境的防护需求。本文探讨了基于 Python 的多种机器学习算法在网络威胁检测中的应用路径，涵盖异常检测、恶意软件识别、钓鱼网站防控以及威胁响应机制等多个方面。研究表明，机器学习不仅提升了威胁识别的准确性与实时性，还具备良好的可扩展性和适应性，尤其在云计算等新型架构中表现出强大优势。未来，结合深度学习与大数据分析，构建更加智能、自适应的网络安全体系，将成为保障信息安全的重要发展方向。

参考文献：

[1] 谷洪彬，郑黎明，魏孔鹏 . 基于机器学习的网络安全态势感知关键技术研究与应用 [J]. 电脑与信息技术，2022，030(1):40–42，49.

[2] 李泽龙 . 基于机器学习算法的网络安全防御策略探讨 [J]. 信息记录材料，2022，000(6):23–24.

[3] 曹光忠 . 机器学习中用 Python 模拟 K 近邻算法的实现与应用 [J]. 电脑知识与技术，2024，000(21):78–80.

[4] 李九州 . 基于 Python 的网络流量特征提取和图形化编程 [J]. 电脑编程技巧与维护，2024，000(6):69–70.

[5] 张婷婷，陈云云 . 基于机器学习的网络安全态势感知关键技术研究 [J]. 网络安全技术与应用，2024，000(1)65–68:.

[6] 葛云峰 . 基于机器学习算法的计算机网络安全体系部署研究 [J]. 网络安全和信息化，2023，000(7):54–56.

[7] 赵志俊 . 基于机器学习的通信网络安全漏洞监测系统设计 [J]. 长江信息通信，2022，035(12):175–177.

基金项目：

1、淮南联合大学校级科研项目《机器学习算法的研究及应用》（LYB2002）；

2、淮南联合大学校级重点科研项目《知通信息技术创新团队》（LCX2303）。