

网络安全领域五大机器学习能力之谜

非官方中文译文·安天技术公益翻译组 译注

文档信息			
原文名称	Five myths about machine learning in cybersecurity		
原文作者	Alexey Malanov	原文发布日期	2016 年 10 月 12 日
作者简介	卡巴斯基实验室是一家国际软件安全公司，在全球近 200 个国家和地区开展业务。公司总部位于俄罗斯莫斯科，在英国注册了控股公司。 https://en.wikipedia.org/wiki/Kaspersky_Lab		
原文发布单位	卡巴斯基实验室		
原文出处	https://securelist.com/blog/opinions/76351/five-myths-about-machine-learning-in-cybersecurity/		
译者	安天技术公益翻译组	校对者	安天技术公益翻译组
分享地址	请浏览创意安天论坛 bbs.antiy.cn 安天公益翻译板块		
免责声明	<ul style="list-style-type: none"> 本译文译者为安天实验室工程师，本文系出自个人兴趣在业余时间所译，本文原文来自互联网的公共方式，译者力图忠于所获得之电子版本进行翻译，但受翻译水平和技术水平所限，不能完全保证译文完全与原文含义一致，同时对所获得原文是否存在臆造、或者是否与其原始版本一致未进行可靠性验证和评价。 本译文对应原文所有观点亦不受本译文中任何打字、排版、印刷或翻译错误的影响。译者与安天实验室不对译文及原文中包含或引用的信息的真实性、准确性、可靠性、或完整性提供任何明示或暗示的保证。译者与安天实验室亦对原文和译文的任何内容不承担任何责任。翻译本文的行为不代表译者和安天实验室对原文立场持有任何立场和态度。 译者与安天实验室均与原作者与原始发布者没有联系，亦未获得相关的版权授权，鉴于译者及安天实验室出于学习参考之目的翻译本文，而无出版、发售译文等任何商业利益意图，因此亦不对任何可能因此导致的版权问题承担责任。 本文为安天内部参考文献，主要用于安天实验室内部进行外语和技术学习使用，亦向中国大陆境内的网络安全领域的研究人士进行有限分享。望尊重译者的劳动和意愿，不得以任何方式修改本译文。译者和安天实验室并未授权任何人士和第三方二次分享本译文，因此第三方对本译文的全部或者部分所做的分享、传播、报道、张贴行为，及所带来的后果与译者和安天实验室无关。本译文亦不得用于任何商业目的，基于上述问题产生的法律责任，译者与安天实验室一律不予承担。 		

网络安全领域五大机器学习能力之谜

机器学习早已渗透到人类活动的方方面面，不仅在语音识别、手势识别、手写识别和图像识别扮演了重要角色，而且难以想象机器没有学习的能力，我们该如何质量管控现代医学，银行业、生物信息学等系统。如果机器没有学习能力，甚至连天气预报都难以实现。

谜一：信息安全领域机器学习能力是新鲜玩意

讨论网络安全领域人工智能有点过时了。如果没有长期关注这个话题，你可能认为它很新奇。

背景知识：首个机器学习算法——人工神经网络，于 20 世纪 50 年代发明。有趣的是，当时人们认为该算法可以帮助我们快速建立“强大的”的人工智能，即能够思考，理解自我、解决编程之外的问题。它被人们称为弱人工智能，可以解决一些创造性的工作——识别图像，预测天气、下棋等。60 年后的现在，我们对发明真正的人工智能将花费数年的时间这一事实有了更深入的理解，而当今称为人工智能的实际上是机器学习能力。

谈到网络安全时，机器学习能力早已不是什么新鲜事了，因为在 10-12 年前我们就已在网络安全领域实施了该等级的算法。当时，新恶意软件的数量每两年翻一番，这使我们愈发明白简单的自动化分析病毒早已不够，我们急需质的飞跃。这一跨越以处理病毒群的形式使得搜索已被检查过的类似文件得以实现。最终判定发送的文件是否是恶意文件，这一功能几乎立即运用到机器人身上了。

换言之，机器学习能力在网络安全领域不是什么新鲜事。

谜二：机器学习在网络安全领域很简单——所有事情已深思熟虑过了

确实某些有现成算法的领域使用了机器学习能力，使用范围包括面部识别，表情识别或分辨猫与狗。除上述情况外，有人已对机器学习能力做了大量的思考，例如，识别必要的符号，选定合适的数学工具，留出必要的计算机资源，然后把他们的发现公之于众。现在，每个学童都可以使用这些算法了。

这也给人们造成了已经存在检测恶意软件算法的错觉。我们在卡巴斯基实验室花了 10 多年的时间开发，申请了许多专利技术。我们将继续研究，提出新的想法，因为那就是下一

个谜流行的原因。

谜三：机器学习能力——学后就忘

恶意软件检测和面部识别之间存在概念差别。脸仍是脸，这一点不会改变。在大多数使用机器学习能力的领域，物体不会随着时间而发生改变，而恶意软件则以很快的速度不断改变。那是因为网络罪犯有很强犯罪动机（例如，钱，间谍活动，恐怖主义）。他们的智力不是人工开发的，而是积极的做斗争，故意修改恶意程序，摆脱既有的培训模式。

那时因为模型必须不断地优化，有时甚至要从头开始。很明显，随着恶意软件修改的速度变快，基于无反病毒数据库模型的安全解决方案是毫无用处的。而必要的时候，网络罪犯可以创造性思考。

谜四：让安全软件从客户端学习

比如说，安全软件处理客户端文件时，大多数文件是安全的，只有少部分存在恶意。虽然后者在改变，但模型也在学习。

然而，模型的工作方式与恶意软件不同，因为恶意软件样本通过一般客户端电脑的数量远比恶意软件样本被反病毒实验系统收集的要少。而且，由于无样本可学，也就不能概括其特征了。如果把病毒作者的“创造性”因素考虑在内，检测将会失败，模型会把恶意软件识别为安全文件，“学到的内容也是错误的”。

谜五：在没有其它检测方法的情况下，仅基于机器语言模型也可能开发安全解决方案

为什么要基于不同技术使用多级防护呢？如果那个篮子是如此智能、如此高级，为什么不把所有的鸡蛋都放在一个篮子里呢？一条算法足以解决所有问题。对吗？

问题是大多数恶意软件属于不同的家族，仅一个恶意程序就由大量的修改组成。例如，勒索木马 Win32.Shade 是由 30000 个加密器组成的家族。通过给模型喂养大量的样本，可以获得发现未来威胁的能力（在一定范围内，见谜 3）。在这种情况下，机器学习能力表现良好。

然而，通常情况是一个家族仅由一些样本组成，甚至一个。或许是由于作者创作的“作品”一经发现之后就不想与安全软件开战了。转而决定攻击那些没有装安全软件或行为检测软件的电脑（例如，那些把所有的鸡蛋放在一个篮子里的人）。

这些“迷你家族”不能用于喂养模型，因为仅有一两个样本是无法概括的（机器学习的本质）。在这种情况下，使用长时间测试的方法是检测威胁最有效方式，例如基于哈希，伪装等。

另一个例子是针对性攻击。在这些攻击的背后，作者不想产生越来越多的新样本。他们为每一个受害者建立一个样本，并确认这些样本不会被防护解决方案发现（除非它是基于此目的解决方法，例如，卡巴斯基的反针对攻击平台）。再次重申，基于哈希检测效果更佳。

结论：不同的工具需使用于不同的解决方案。多级防护比单击防护更加有效——不能仅仅因为有效的工具“过时”就忽视它们。

机器警察存在的问题

最后，我想说的是，这更多是提个醒，而不是谜。当前的研究人员把注意力更多的放在复杂模型所犯的错误上：有时候，模型所做的决定不能以人类的逻辑来解释。

机器学习能力是可以被信赖的。但重要的系统（例如飞机的自动导航，汽车，医学、控制服务等）一般有非常严格的质量标准，它们使用正式的软件验证程序，而在机器学习中，我们只能展现部分思维过程，并对机器负有责任。那就是为什么质量控制模型需要德高望重的专家设计。