前言

近几年，人工智能无疑成为人们口中的热点话题。先是 Google 的 AlphaGo，后有百度的度秘、无人车，乃至微软必应搜索推出的小冰也是新闻不断。这一系列人工智能产品的推陈出新，令人眼花缭乱，一时间给人的感觉是人工智能遍地开花。无论人们接受还是不接受，人工智能都在迅速渗透各行各业。

安全作为一个传统行业，基于规则以及黑白名单的检测技术已经发展到了一定的瓶颈，而利益驱动的黑产团伙，其技术的发展已经远远超乎我们的想象。如何借助人工智能的力量，提升安全行业的整体检测与防护能力，成为各大安全厂商研究的课题。

国内安全领域，随着 BAT3 以及大量新兴的创业企业进入企业安全领域，他们凭借着自身数据搜集、处理、积累以及人工智能方面优势，正在逐渐改变着整个安全行业，安全产品的形态从硬件盒子逐步走向混合模式以及云端 SaaS 服务，安全技术从重防御逐步走向数据以及智能驱动。传统安全企业也凭借其强大的安全人才储备，迅速推进人工智能在安全产品的落地。

我在安全这个行业快十年了，做了五年的超大型互联网公司的企业安全建设，从准入系统到 WAF、SIEM、IPS，基本都开发或者使用过，最近三年一直负责云安全产品，从抗 D 产品到态势感知、入侵检测，使用的技术从规则、黑白名单、模型、沙箱再到机器学习，从单机的 OSSIM 到 Hadoop、Storm、Spark、ELK，也算目睹了安全技术或者更准确的说数据分析处理技术的迅猛发展。我深深感到使用人工智能技术改变这个行业，不是我们的选择，而是必经之路。我真正意义上接触人工智能是 2015 年，当时带领了一个很小的团队尝试使用机器学习算法解决安全问题，磕磕碰碰一直走到现在，变成几十人的一个大产品团队。这本书基本也是我对机器学习应用于安全领域的一些理解。

本书的第 1 章概括介绍了机器学习的发展以及互联网目前的安全形势。第 2 章介绍了如何打造自己的机器学习工具箱。第 3 章概括介绍了机器学习的基本概念。第 4 章介绍 web 安全的基础知识。第 5 章到第 13 章介绍浅层机器学习算法，包括常见的 K 近邻、决策树、朴素贝叶斯、逻辑回归、支持向量机、K 均值、FP-growth 与 Apriori、隐式马尔可夫、有向图。第 14 章到第 17 章介绍神经网络以及深度学习中常用的递归神经网络和卷积神经网络。每章都会以生活中的例子开头，让读者有一个感性的认识，然后是简短的基础知识的介绍，最后是以安全领域的 2-3 个例子，讲解如何使用该算法解决问题，全书定位是能让更多的安全爱好者以及信息安全从业者可以了解机器学习，

动手使用简单的机器学习算法解决实际问题，所以全书尽量避免生硬的说教，能用文字描述的尽量不用冷冰冰的公式，能用图和代码说明的尽量不用多余的文字，正如霍金说言，”多写 1 个公式，少一半读者”，希望反之亦然。

本书面向信息安全从业人员、大专院校计算机相关专业学生以及信息安全爱好者，对于想了解人工智能的 CTO、运维总监、架构师同样也是一本不错的科普书籍。

这里我要感谢我的家人对我的支持，本来工作就很忙，没有太多时间处理家务，写书以后更是侵占了我大量的休息时间，我的妻子无条件承担起了全部家务尤其是照料孩子方面，我和感谢我的女儿，写书这段时间几乎没有时间陪她玩，我也想用这本书作为她的生日礼物。我还要感谢编辑吴怡对我的支持和鼓励，让我可以坚持把这本书写完。最后还要感谢各位业内好友尤其是我 boss 对我的支持，排名不分先后。

马杰@百度安全、冯景辉@百度安全、Lenx@百度安全、黄正@百度安全、程岩@百度云、云鹏@百度无人车、林晓东@百度基础架构、Tony@京东安全、赵林林@微步在线、谢忱@Freebuf、王宇@蚂蚁金服、周涛@启明星辰、姚志武@借贷宝、刘静@安天、廖威@易宝支付、尹毅@sobug，齐鲁@搜狐安全、吴圣@58 安全、康宇@新浪安全、幻泉@i 春秋、西瓜@四叶草。

我平时在 Freebuf 专栏以及 i 春秋分享企业安全建设以及人工智能相关经验与最新话题，我平时也运营我的微信公众号”兜哥带你学安全，欢迎大家关注并在线交流。