人工智慧之資安衝擊與因應

毛敬豪 博士
資安科技研究所
2017/06/17

簡報大綱

- 人工智慧在資安攻防的角色
- 人工智慧被駭客利用攻擊的方法
- 人工智慧帶給資安防護的契機
- 人工智慧資安技術研發展望

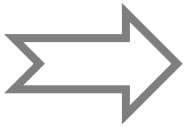
人工智慧在資安攻防的角色

Why we need AI in information security?

目前

專家經驗知識

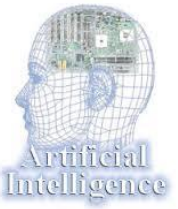
耗時、局部、後知後覺



未來

機器智能學習

快速、全面、及早發覺

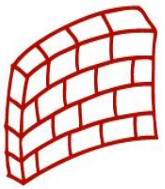


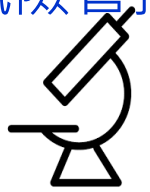
偵防核心

攻擊特徵、威脅情資、防禦規則

解決方案

身分識別與內容安全	資安弱點檢測	入侵偵測防禦	威脅監控分析
資料防漏 數位版權管理 隱私保護 身分權限管控	弱點掃描 滲透測試 原始碼檢測 App資安檢測	防毒軟體 防火牆/IPS 郵件過濾	資安監控 日誌管理 事件鑑識 風險管控





人工智慧在資安攻防的角色

How AI-powered cyberattacks will make fighting hackers even harder

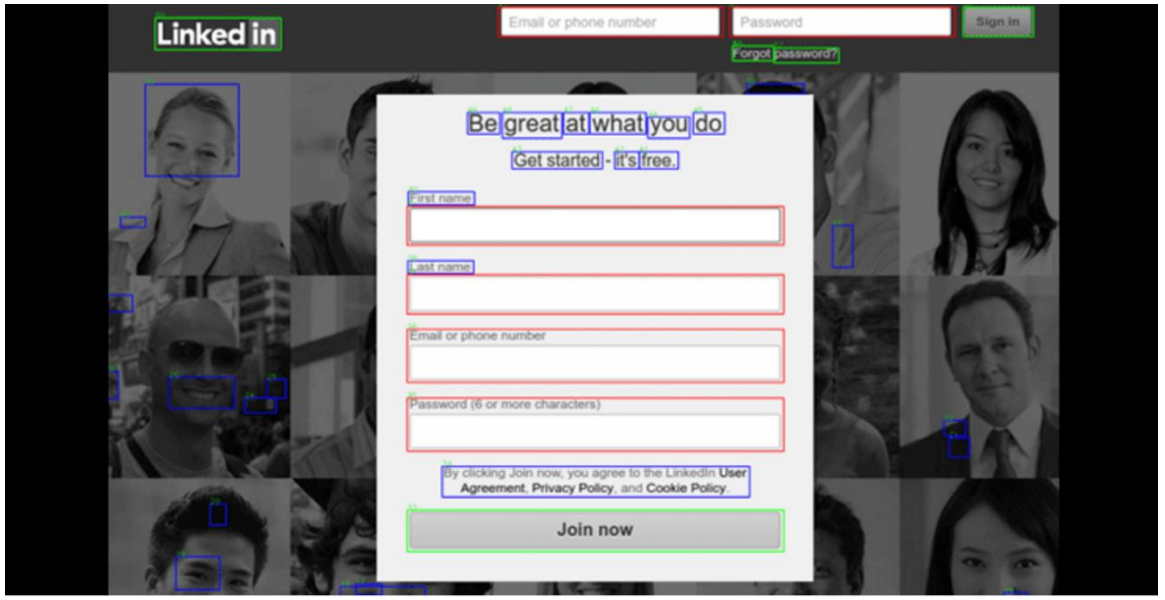
- Simple sequence-to-sequence machine learning:
  - monitor emails and conversations of a compromised victim on an infected device
  - tailor phishing messages to mimic the message style of the victim to particular contacts in their address book (convince them to click on a malicious link)
  - AI can learn social profile and online shopping accounts

http://www.zdnet.com/article/how-ai-powered-cyberattacks-will-make-fighting-hackers-even-harder/

人工智慧在資安攻防的角色

Phishing-as-a-service is making it easier than ever for hackers to steal your data

- an artificial intelligence system based on a semi-supervised learning model that can navigate and interact with the Internet just like an intelligent human being
  - automatically gather information about input boxes, buttons, and navigation links with minimal false positives



- Leak of any user's Email ID on LinkedIn
- Leak of users email and phone number and resume
- Deleting every user's LinkedIn request
- Downloading every transcript to videos from Lynda
- Downloading every Lynda exercise files without a premium membership

人工智慧帶給資安防護的契機

Cloud-AI: Artificially Intelligent System Found 10 Security Bugs in LinkedIn

- Clouddesk: providing intelligence machine learning-based solutions to help organizations identify and tackle online threats in real-time
  - artificial Intelligence system based on a semi-supervised learning model that can navigate and interact with the Internet just like an intelligent human being

人工智慧在資安攻防的角色

breaking CAPTCHA

- Artificial Intelligence is good at breaking CAPTCHA codes.
- To beat Google's latest reCAPTCHA system, which is also powered by a sophisticated artificial intelligence system to defend websites against bots.

人工智慧被駭客利用攻擊的方法

Will AI usher in a new era of hacking?

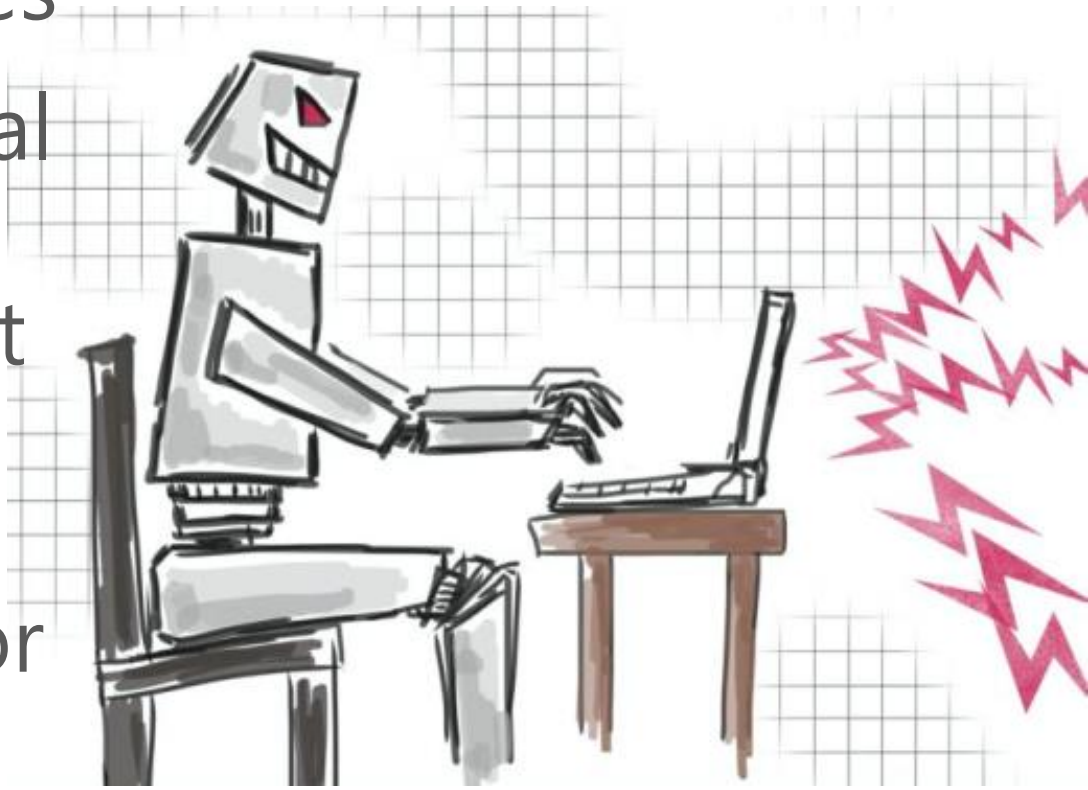
- Hackers are using malware and phishing scams to steal Netflix users' passwords, bank details
  - inadvertently downloaded Infostealer.Banload
  - back door to steal information and download potentially malicious files

Artificial intelligence—could one day become the go-to hacking tool

人工智慧被駭客利用攻擊的方法

Will AI usher in a new era of hacking?

- Vulnerability discovery is a double-edge sword
- Rent-a-hacker services
- Building mathematical models based on malware samples that can gauge whether certain activity on a computer is normal or not.

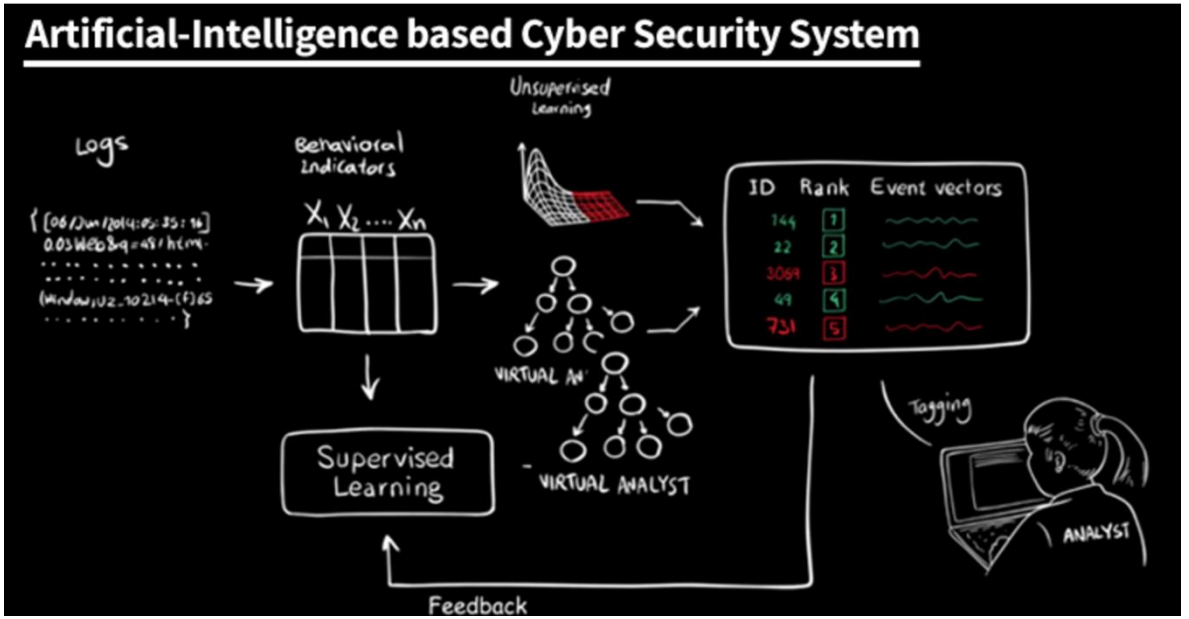


http://www.pcworld.com/article/3142940/security/will-ai-usher-in-a-new-era-of-hacking.html

人工智慧帶給資安防護的契機

MIT builds Artificial Intelligence system that can detect 85% of Cyber Attacks

- MIT builds Artificial Intelligence system that can detect 85% of Cyber Attacks
- The more data it analyzes, the more accurate it becomes



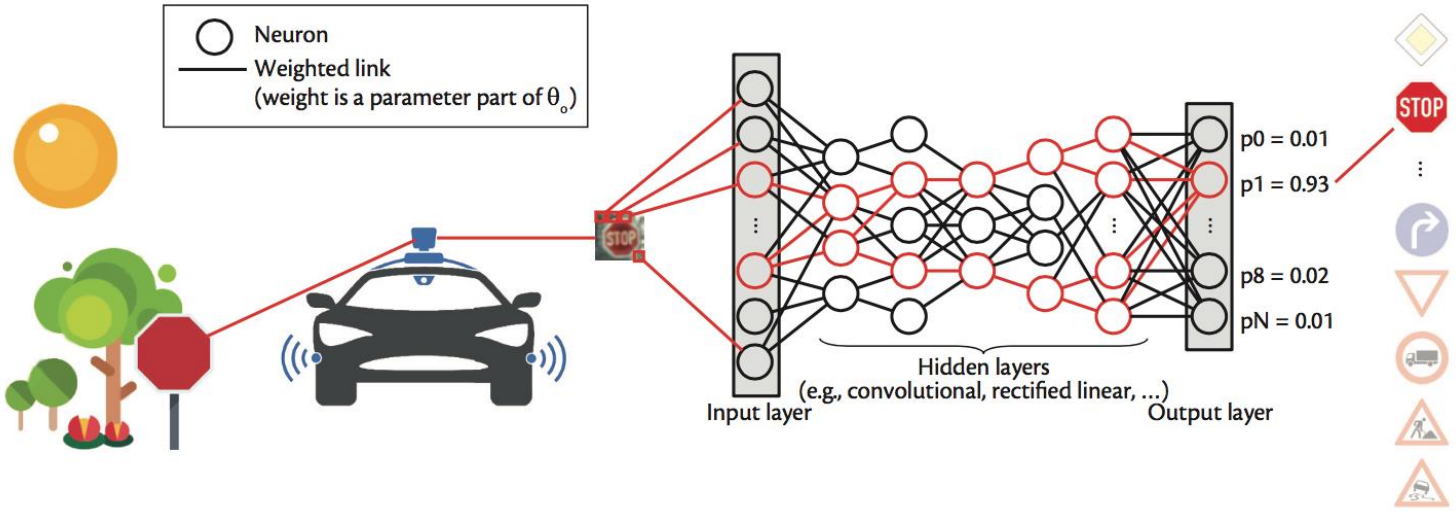
人工智慧帶給資安防護的契機

AI security startup

Name	Founded	Features
Darktrace	2013/UK	**Self-learning intelligence** of the human immune system (energy and utilities, financial services, telecommunications, healthcare, manufacturing, retail and transportation)
Jask	2015/USA	New **AI-based** approach that produces the precious few alerts that detect real actual attacks
Deep instinct	2014/Isarel	Leveraging **deep learning' s predictive capabilities**, Deep Instinct' s on-device, proactive solution protects against zero-day threats and APT attacks with industry leading accuracy
Harvest.ai	2014/USA	**AI-based algorithms to learn the business value of critical documents** across an organization, detect and stop data breaches from targeted attacks and insider threat before data is stolen.
PatternEx	2013/USA	Threat Prediction Platform is designed to create "virtual security analysts" that **mimic** the intuition of human security analysts in real time and at scale
Vectra Networks	2011/USA	Prioritizes attacks that pose the greatest business risk, enabling organizations to quickly make decisions on where to focus their time and resources
Status Today	2015/UK	Insider threat and data breaches using a patent pending Artificial Intelligence that understands Human Behavior
Cyberlytic	2013/UK	Prioritizes the workload of security teams and reduces response times from cyber attacks to seconds
Neokami	2014/Germany	Leveraging Artificial Intelligence to discover, secure and govern sensitive data in the cloud, on premise, or across their physical assets.
Fortscale	2012/USA	User behavior analytics (UEBA) solution combines expertise from the Israeli Defense Force' s elite security unit, big data analytics and advanced machine learning

人工智慧帶給資安防護的契機

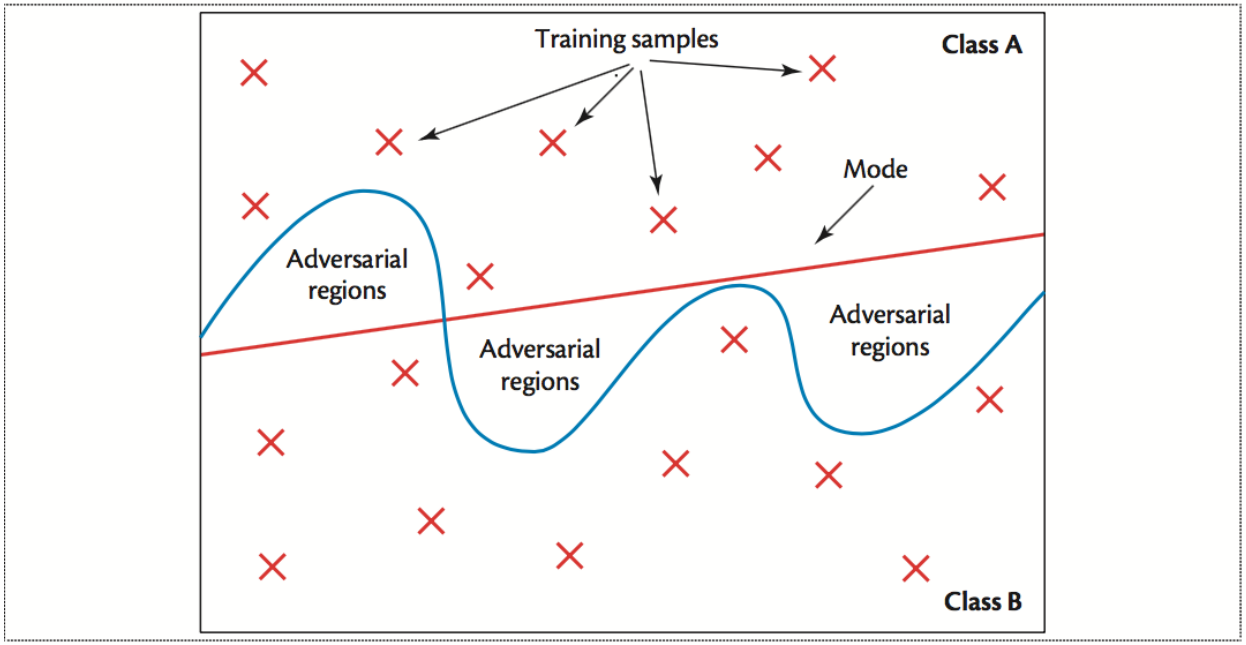
- In the context of classification, adversarial samples are crafted to force a target model to classify them in a class di erent from their legitimate class.



IEEE Security & Privacy 2016

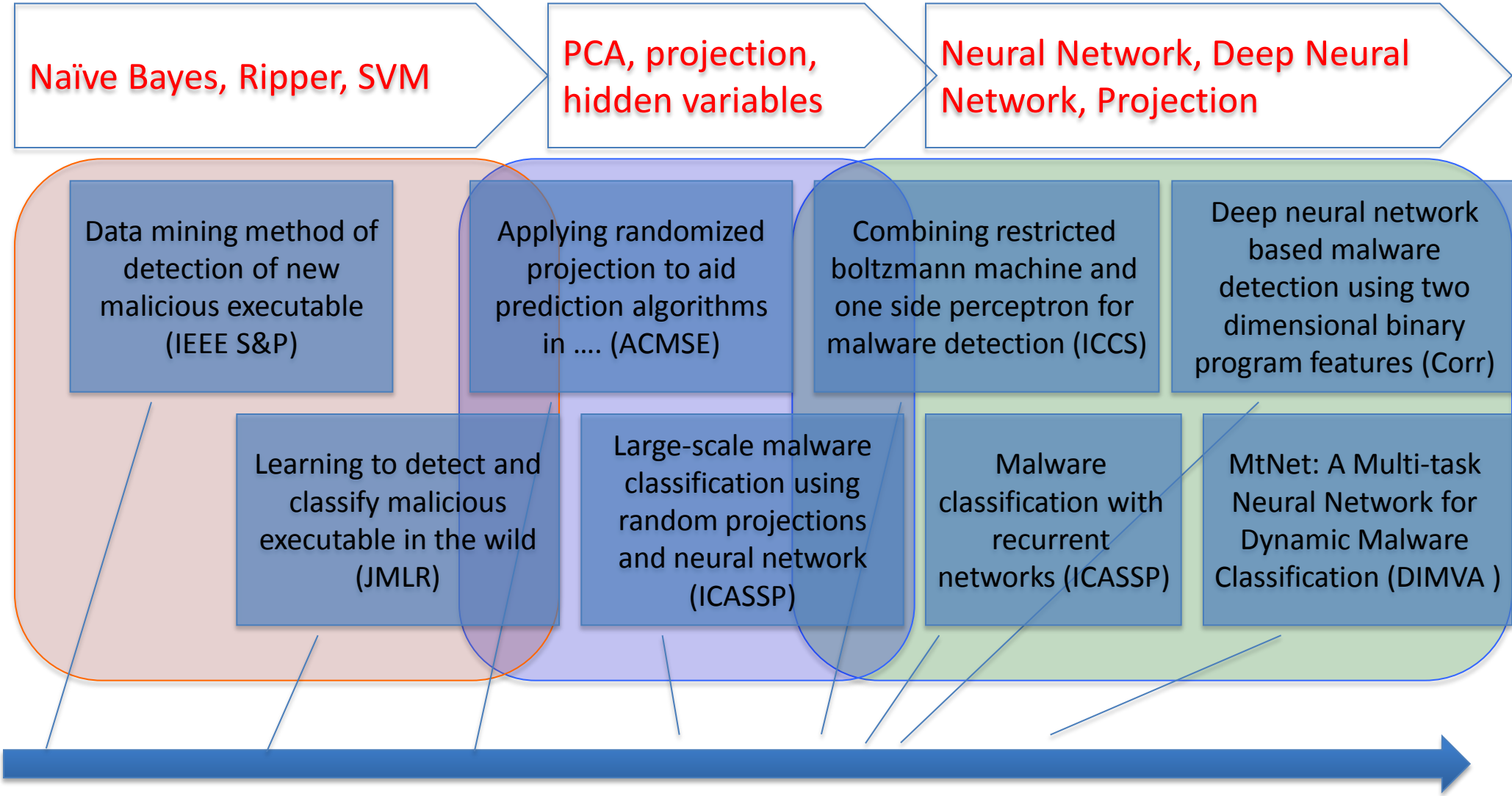
人工智慧帶給資安防護的契機

- Model training and use. e plane represents all possible input feature vectors. For each sample, the input feature values uniquely identify its coordinates in the plane. Two classes A and B (that is, spam and not spam) are regions in a two-dimensional plane separated by the smooth curved line.



IEEE Security & Privacy 2016

人工智慧帶給資安防護的契機
Malware Classification using Data Analytics



人工智慧資安技術研發展望

研發人工智慧資安核心，發展新興應用資安整合技術方案，導入於智慧城市、智慧製造與國防資安等場域

關鍵資安核心

機器感知弱點探析技術

深度學習威脅防禦技術

資料安全與隱私技術

新興應用資安

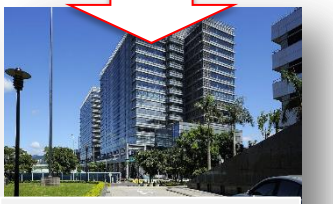
雲端服務資安偵防

IoT/CPS資安偵防

FinTech資安防護

場域實證

政府單位



工業園區

企業機構