L'intelligenza artificiale come strumento "dual use" nella cybersecurity

Luca Bechelli



Clusit Education

Luca Bechelli

Clusit

- Practice Leader Information & Cyber Security Advisory Team @ PL.
- Membro del Comitato Direttivo e del Comitato Tecnico Scientifico
- Coordinatore GdL «La valutazione degli impatti del GDPR nelle clausole contrattuali dei fornitori» presso l'Osservatorio «Privacy & Security» del Politecnico di Milano
- Direttore Didattico Academy Experis per Master in Cybersecurity
- Community Clusit Oracle4Security Oracle Community For Security









TayTweets 🥥

The official account of Tay, Microsoft's A.I. tans from the internet that is got zero child The more you tak the smarter Tay gets

Q the internets

S MAY AL WASHING





96.1K

TayTweets

PDILLUNEPE

50.3K



@wowdudehahahaha I full g hate n s, I wish we could put them all in a concentration camp with k s and be done with the lot

4 63 W

Investimenti in IA in Italia







Cos'è Triton, il malware che può causare incidenti catastrofici

E' pensato per aggirare i sistemi di sicurezza di impianti sofisticati, come centrali nucleari o petrolchimici, e provocare gravi conseguenze

Triton è il nome di un <u>malware</u>, scoperto in Medio Oriente, che sarebbe in grado di **bloccare i** sistemi di sicurezza progettati per prevenire gli <u>incidenti industriali catastrofici</u>. Il malware autorizza gli aggressori a intrufolarsi nei sistemi di sicurezza.

Quando è stato scoperto da Julian Gutmanis nell'estate 2017, il malware stava infettando un impianto petrolchimico in Arabia Saudita, disattivando i controllori fisici che fungono da ultima difesa contro i possibili incidenti, che possono mettere a repentaglio la vita del personale. Questo comportamento rende Triton **uno dei malware più micidiali al mondo** poiché potrebbe, ipoteticamente, causare l'esplosione di un impianto petrolifero, facendo aumentare la pressione all'interno delle condutture piene di materiale infiammabile attraverso la chiusura da remoto delle valvole e la disattivazione dei sensori d'allarme.



Le possibilità dell'AI The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation (University of Oxford)



Recent progress in image recognition on the ImageNet benchmark. Graph from the Electronic Frontier Foundation's AI Progress Measurement project (retrieved August 25, 2017)



Increasingly realistic synthetic faces generated by variations on Generative Adversarial Networks (GANs). In order, the images are from papers by Goodfellow et al. (2014), Radford et al. (2015), Liu and Tuzel (2016), and Karras et al. (2017).

2014

Clusit

Today's AI systems suffer from a number of novel unresolved vulnerabilities.

These include **data poisoning attacks** (introducing training data that causes a learning system to make mistakes), **adversarial examples** (inputs designed to be misclassified by machine learning systems), and the **exploitation of flaws** in the design of autonomous systems' goals.

These vulnerabilities are distinct from traditional software vulnerabilities (e.g. buffer overflows) and demonstrate that **while AI systems can exceed human performance in many ways, they can also fail in ways that a human never would.**



Minacce

Espansione delle minacce esistenti

Modifica delle caratteristiche tipiche delle minacce Nuove Minacce

(pessimi) Scenari prossimi venturi

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation (Univesity of Oxford)

- Automazione degli attacchi di social engineering
- Automazione della scoperta di vulnerabilità
- Maggiore sofisticazione degli strumenti e delle tecniche di hacking
- Denial of service basato su comportamenti «human-like»
- Analisi e prioritizzazione degli obiettivi di attacco
- Violazione o contrasto delle AI «legittime»
- Alterazione o comprensione del comportamento delle AI «legittime»



Predizioni per il futuro...

Tomorrow's fears ...

Over the next five years, executives are concerned about the following threats:

43% New privacy threats

1,027 data breaches, exposing 57,667,911 records, have occurred. in 2018 through October.1

41% New cyberthreats

Malicious cyber activity cost the US economy between \$57 billion and \$109 billion in 2016.4

34% New legal liabilities and reputational risk

In 2018, just 48% of people said they trust US businesses, a decline of 10% from the prior year."

33% Too complex to understand or control

200% increase in DARPA XAI funding from 2017 to 2019.4

31% Unable to meet demand for Al skills

32% year-over-year growth of Al-related job postings.⁴

31% US falling behind other countries in Al innovation

6X, the number of deep learning patent publications by China. compared with the US.*

> Income Pure 2018 II Presidence Same 1, 2014 () Thinks of the Addressing contraction encoded As the president and a read friend in the part 2 page?

> > Fonte: PWC 2019 AI Predictions

Clusit Education

1. Martilly The's Persons Center, Chila Breach Hastory, Cettaber 21, 2018.

3. Courses of Economic Advances, The Disk of Manistran Course Actions 16 the U.S. Economic Ferrican 2018.

- 3 251d Roberture Drunt Barranteller
- 4. Separate at German Facal Test PTL2111, Bullari Damatte.
- 5. Solved Hong Lin. December 30 M Latertrain the Size. March 2018
- A CS thatpen, Articla Intelligence Dentry To Human In 2018.

Clusit

...e ambiti di sviluppo

Taking steps toward responsible AI

61%

Al models

Test for bias in data.

models, human use

of algorithms.



64%

Boost Al security with validation, monitoring, verification



52% Improve governance with At operating models, processes

Status Nur. 2014 & Paulaitant

I first pape of the represent power first in density with latits & possible for on supported by a hydrogenty, had actually "



55% Create transparent, explainable, provable

Create systems that am offical. understandable, legal



3% We currently have no plans to address these Al issues

Top five AI challenges for 2019



Ensuring that AI systems are trustworthy 37%



Training current employees to work with All

Managing the convergence of AI with other technologies 36%



Measuring Al's return on investment 31%



Moving AI initiatives from pilot to production 29%

Room Prof. 2010 of Press States

A must do blo added server effice We by provide to yet regulation to 2010.





Al...uto!

Which departments are currently using AI-powered technology?



Which departments are demanding more AI-powered technology?



(fonte: Cylance - ricerca su 100 esperti negli USA)



Al...uto!

60% +

degli esperti di cybersecurity, ritiene che l'intelligenza artificiale sarà utilizzata per attacchi informatici sempre più devastanti nei prossimi 12-18 mesi

(fonte: Cylance - ricerca su 100 esperti negli USA)



Al...uto!

39% +

degli esperti di cybersecurity preoccupati dallo stato degli aggiornamenti e delle patch sui sistemi

(fonte: Cylance - ricerca su 100 esperti negli USA)



Tecniche di attacco (rispetto al 2017)



+39% Know Vulnerabilities



Ci sono cose che non cambiano mai...

OWASP
Top ten
vulnerabilities
2017

OWASP Top 10 - 2013	→	OWASP Top 10 - 2017
A1 – Injection	•	A1:2017-Injection
A2 – Broken Authentication and Session Management	→	A2:2017-Broken Authentication
A3 – Cross-Site Scripting (XSS)	4	A3:2017-Sensitive Data Exposure
A4 – Insecure Direct Object References [Merged+A7]	U	A4:2017-XML External Entities (XXE) [NEW]
A5 – Security Misconfiguration	2	A5:2017-Broken Access Control [Merged]
A6 – Sensitive Data Exposure	7	A6:2017-Security Misconfiguration
A7 – Missing Function Level Access Contr [Merged+A4]	U	A7:2017-Cross-Site Scripting (XSS)
A8 – Cross-Site Request Forgery (CSRF)	×	A8:2017-Insecure Deserialization [NEW, Community]
A9 – Using Components with Known Vulnerabilities	→	A9:2017-Using Components with Known Vulnerabilities
A10 – Unvalidated Redirects and Forwards	×	A10:2017-Insufficient Logging&Monitoring [NEW,Comm.]





Segnali Deboli: +55% APT +140% O-day

Al per la Cybersecurity

- Adozione:
 - Il **12%** delle aziende enterprise negli USA hanno implementato soluzioni di AI per la cybersecurity in modo esteso
 - Il **27%** ha implementato solo soluzioni di security analytics
- Driver:
 - Incident Detection (29%)
 - Incident Response (27%)
 - Identificare e comunicare rischi al business 24%
 - Correlare informazioni e fornire in tempo reale lo stato della sicurezza (22%)

Fonte: ESG Research, 2018



GRAZIE Domande?

Clusit

Luca Bechelli Comitato Scientifico Clusit <u>luca@bechelli.net</u> <u>www.bechelli.net</u> https://twitter.com/luca_bechelli <u>https://www.facebook.com/bechelli.luca</u> http://www.linkedin.com/in/lucabechelli