

# Retos del Machine Learning y Deep Learning en Ciberseguridad

Gustavo A. Isaza E., Ph.D

Universidad de Caldas

[gustavo.isaza@ucaldas.edu.co](mailto:gustavo.isaza@ucaldas.edu.co)

# Presentación

Gustavo A. Isaza Echeverri

Universidad de Caldas

## Estudios

PhD Ing. de Software (2010, U Pontificia Salamanca - España)

Especialista Software para Redes (1999, Uniandes - Bogotá)

Ingeniero de Sistemas y Computación (1997, UAM)

## Logros

Profesor Titular / Investigador Senior UCaldas

Artículos y conferencias en IA en CyberSec, Bioinf ....

Consultor CyberSec, IA, DevSecOps, Seguridad Ofensiva

## Acreditaciones

Cloud Security Alliance (CSA). Cybersecurity Alliance Expedición: abr. de 2022

CSSLP – Certified Secure Software Lifecycle ISC2 Expedición: nov. de 2021

CyberOps CISCO (Proceso) - Cybersecurity Operations, 2021

Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization Expedición: 2021

Neural Networks and Deep Learning, 2020

Structuring Machine Learning Projects, 2021

# Inteligencia Artificial y Ciberseguridad

- Conceptos
- Crisis en la Ciberseguridad
- Tendencias IA <-> CyberSec
- Casos de Uso

# Inteligencia Artificial y Ciberseguridad

- **Machine Learning (ML)**

El aprendizaje automático es un tipo de inteligencia artificial (IA) que se centra en el desarrollo de programas que pueden aprender y cambiar cuando se exponen a nuevos datos.

- **Data Analytics**

El análisis de datos es un proceso de inspección, limpieza, transformación y modelado de datos con el objetivo de descubrir/inferir información útil, sugerir conclusiones y respaldar la toma de decisiones.

- **Cybersecurity**

La ciberseguridad es el conjunto de tecnologías, procesos y prácticas diseñadas para proteger redes, computadoras, programas y datos de ataques, daños o accesos no autorizados. En un contexto informático.

# Inteligencia Artificial y Ciberseguridad

- **ML + Data Analytics + Cyber Security**

El aprendizaje automático se ha adoptado rápidamente en ciberseguridad por su potencial para automatizar la detección y prevención de ataques, particularmente para productos antimalware de próxima generación (NGAV).

# Crisis gestión de la Ciberseguridad

Cada SOC aprox 11.000 alertas diarias, de las cuales solamente el 17% están automatizadas, y la investigación de cada una requiere la consulta de 10 categorías de herramientas diferentes (Fuente: Forrester Study: The 2020 State of Security Operations).

El 28% de las alertas son ignoradas por la intervención manual.

El 25% de los ingenieros de DevSecOps cambian de trabajo en menos de dos años y el 67% lo hace al tercer año (Fuente: The State of SOAR Report, 2019)

<https://revista.uclm.es/index.php/ruiderae/article/view/3088/2402>

# Crisis gestión de la Ciberseguridad

El 54 % usan SOAR (Security Orchestration Automation and Response) para gestión de incidentes

Reducción del tiempo de mitigación (51%), la reducción del tiempo promedio de extremo a extremo en un incidente (47%) y la reducción del tiempo de clasificación (44%).

37% dijo que SOAR ayudó a reducir la cantidad de pasos necesarios para dar respuesta a los incidentes. (Fuente: The State of SOAR Report, 2020).

# Aplicaciones de la IA + CyberSec

El análisis de las amenazas escapa a la comprensión humana.

La inteligencia sobre amenazas acumula cientos de millones de muestras en bases de datos.

El número de ataques (+ complejos) desborda a los CISOs



# IA + CyberSec

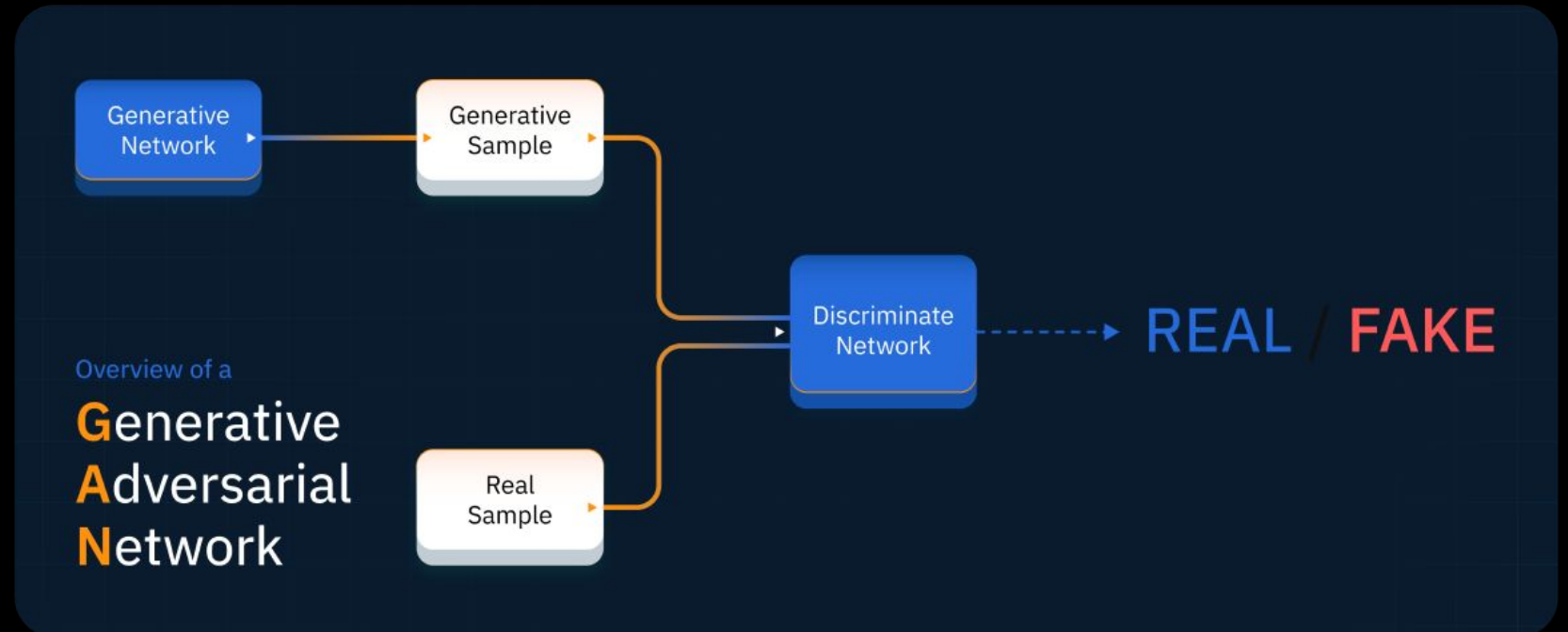
- Puede identificar malware oculto. El reconocimiento de patrones permite detectar comportamientos de amenazas que producen problemas de seguridad, ya sean conocidos o desconocidos.
- Mantiene mejor informados a los equipos de seguridad para que tomen mejores decisiones.
- Libera a los analistas de seguridad de las tareas rutinarias y a mejorar la eficacia incluso de los miembros más inexpertos del equipo.

# Aplicaciones de la IA + CyberSec

- Permite seguir el ritmo de los adversarios introduciendo nuevas técnicas, a mayor cantidad de datos disponibles, más precisión.
- Ayuda a los equipos de TI a analizar los fallos. Si la seguridad de los endpoints no puede evitar los daños de un ataque, el aprendizaje automático guarda los elementos de datos relevantes y los pone a disposición de los analistas de seguridad cuando los necesitan, resiliencia.

# IA+Ciberataques

Con las técnicas de Machine Learning los ciberdelincuentes mejoran sus algoritmos para adivinar las contraseñas de los usuarios. Con las redes neuronales y redes adversarias generativas, los ciberdelincuentes podrían analizar grandes conjuntos de contraseñas y variaciones de contraseñas que se ajusten a la distribución estadística.



# Aplicaciones de la Tecnología desarrollada

- SIEM
- Intrusion Detection and Prevention
- Steganography & Steganalysis
- Data Analysis
- Cryptography and Distributed Crypto
- CyberSec + IA

# DATOS IMPORTANTES DEL SECTOR

- **Detección de nuevas amenazas**

- Mediante el uso de algoritmos sofisticados, los sistemas de IA se entrenan para detectar malware, ejecutar el reconocimiento de patrones y detectar incluso los comportamientos más pequeños de ataques de malware o ransomware antes de que ingresen al sistema.
- La IA permite una inteligencia predictiva superior con procesamiento de lenguaje natural que selecciona los datos por sí solo al analizar artículos, noticias y estudios sobre ciberamenazas.
- Esto puede dar inteligencia de nuevas anomalías, ataques cibernéticos y estrategias de prevención.

- **Bots de combate**

- La IA y el aprendizaje automático ayudan a comprender mejor el tráfico del sitio web y a distinguir entre los bots buenos (como los rastreadores de los motores de búsqueda) y los bots malos.

# DATOS IMPORTANTES DEL SECTOR

- Predicción del riesgo de incumplimiento

- Los conocimientos prescriptivos/predictivos del análisis basado en IA le permiten configurar y mejorar los controles y procesos para reforzar su resiliencia cibernética.

- Mejor protección de puntos finales

- Línea de base de comportamiento para el endpoint a través de un proceso de entrenamiento repetido. Si ocurre algo fuera de lo común, el sistema AI puede marcarlo y tomar medidas, ya sea enviando una notificación a un profesional o incluso volviendo a un estado seguro después de un ataque de ransomware.

# Estrategia IA+CyberSec

- Identificar fuentes de datos y crear plataformas para poner en funcionamiento la IA (Data Lake).
- Selección de casos de uso de alto impacto: seleccionar un conjunto de casos de uso relevantes para acelerar y maximizar los beneficios.
- Mejora de la inteligencia de amenazas: colaborar con partners estratégicos para mejorar la inteligencia de amenazas.

# Estrategia IA+CyberSec

- Implementación de SOAR: implementar orquestación de seguridad, automatización y respuesta para mejorar la gestión de la seguridad
- Formación a ciberanalistas: capacitar a los analistas cibernéticos para que dominen la IA
- Gobernanza eficaz: establecer un modelo de administración de la IA en ciberseguridad para ofrecer mejoras a largo plazo de forma transparente y ética.



# Retos IA + CyberSec

**SOC autónomo: una herramienta XDR (detección y respuesta extendida) y una herramienta SOAR**

Predicción de nuevos patrones de Indicadores de Compromiso (IoC): descubrir nuevos IoC aplicando modelos de aprendizaje profundo sobre grandes conjuntos de datos.

Detección de intrusiones mediante UEBA (User and Event Behavioral Analytics): al emplear inteligencia artificial, las desviaciones del comportamiento normal se pueden extraer en tiempo real y evaluar mediante algoritmos de aprendizaje automático.

Descubrimiento de nuevos TTP (Técnicas, Tácticas y Procedimientos de ataques): proporcionar información sobre los perfiles de los grupos de ataque inspeccionando los patrones históricos y predecir posibles actividades futuras

# APPLICATIONS OF AI IN CYBERSECURITY

## 1. PASSWORD PROTECTION & AUTHENTICATION

AI is helping developers make biometric authentication even more accurate.



## 2. PHISHING DETECTION & PREVENTION CONTROL

AI & ML can be used to detect, track, react to & resolve phishing issues much more quickly than humans can.



## 3. VULNERABILITY MANAGEMENT

Systems based on AI & ML are proactive instead of reactive.



## 4. NETWORK SECURITY

AI is expediting the creation of security policies & determining organizations' network topographies.



## 5. BEHAVIORAL ANALYTICS

ML algorithms can learn & create a pattern of a user's behavior.



## Training phase



Benign  
executables



Malicious  
executables



Training



Predictive model

## Protection phase



Unknown  
executable

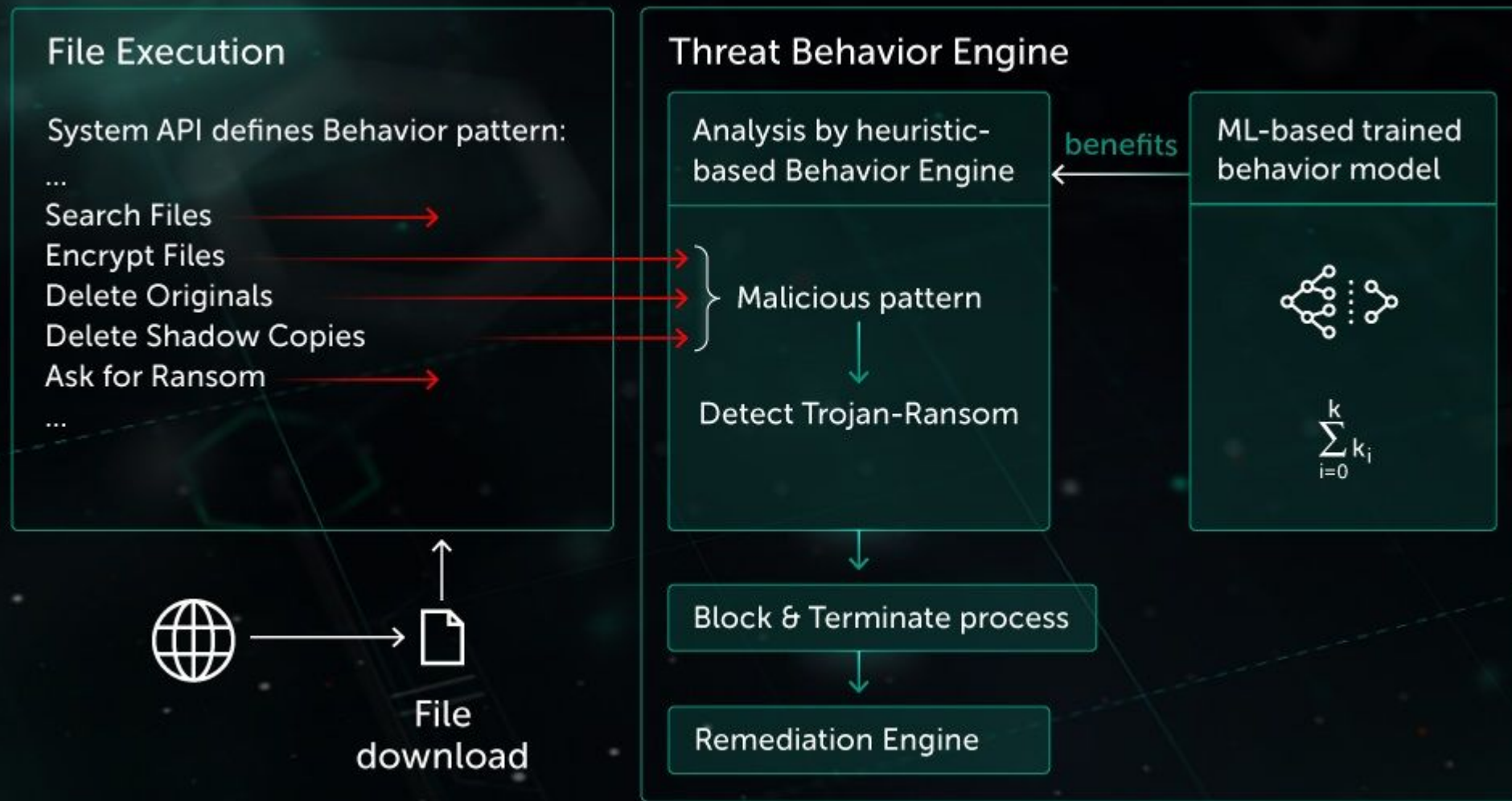


Processing  
by a predictive model

Malicious / Benign

Model decision

Machine Learning: principles



Threat Behavior Engine



# Datasets

- <https://www.unb.ca/cic/datasets/index.html>
- <https://www.kaggle.com/datasets/sampadab17/network-intrusion-detection>
- <https://github.com/shramos/Awesome-Cybersecurity-Datasets>
- <https://cyberdatascientist.com/datasets/>
- <https://zenodo.org/record/4884116#.YtlwH-zm9kx>
- <http://agents.fel.cvut.cz/boss/index.php?mode=VIEW&tmpl=materials>
- <http://bows2.ec-lille.fr/index.php?mode=VIEW&tmpl=index1>
- 

Canadian Institute for Cybersecurity

[Home](#)[About](#)[Research](#)[Members](#)[Datasets](#)[Contact Us](#)

CIC

[About the CIC](#)[Membership](#)[Research](#)[Datasets](#)[Webinars](#)[Global EPIC Program](#)[Cybersecurity Workshop](#)

## Datasets

Canadian Institute for Cybersecurity datasets are used around the world by universities, private industry, and independent researchers. We maintain an interactive map indicating datasets downloaded by country.

The following datasets are currently available:

- CIC IoT Dataset 2022
- CIC MalMem 2022
- Evasive-PDF Mal 2022
- Enriching IoT Datasets
- CIC Bell DNS EXF 2021
- CIC Bell DNS 2021
- CCCS-CIC-AndMal2020
- DNS over HTTPS (CIRA-CIC-DoHBrw2020)
- CICMalDroid 2020
- Darknet 2020
- Investigation of the Android Malware (CIC-InvesAndMal2019)
- DDoS Evaluation Dataset (CIC-DDoS2019)
- IPS/IDS dataset on AWS (CSE-CIC-IDS2018)
- Intrusion Detection Evaluation Dataset (CIC-IDS2017)
- Android Malware Dataset (CIC-AndMal2017)

## Datasets


Handpicked real-world datasets that you can use for your Machine learning project.

Each dataset is tagged and categorized to help you choose the right dataset.

If you want to share your dataset or if you find any kind of intellectual property valuation please contact us


Showing all 16 results

Sort by popularity




Awesome Machine Learning for Cyber Security

[Read more](#)




CVE downloads data

[Read more](#)







Dynamic Malware Analysis kernel and user-level calls

[Read more](#)



Enron Spam Emails

[Read more](#)



Malicious and Benign Websites

## Awesome-Cybersecurity-Datasets

A curated list of amazingly awesome Cybersecurity datasets.

Please contribute to this list with new datasets by sending me a pull request or by contacting me at [@santiagohramos](#).

Happy learning!

### Table of contents

- [Network traffic](#)
- [Malware](#)
- [WebApps](#)
- [Software](#)
- [URLs & Domain Names](#)
- [Host](#)
- [Email](#)
- [Fraud](#)

# CyberSec en IA

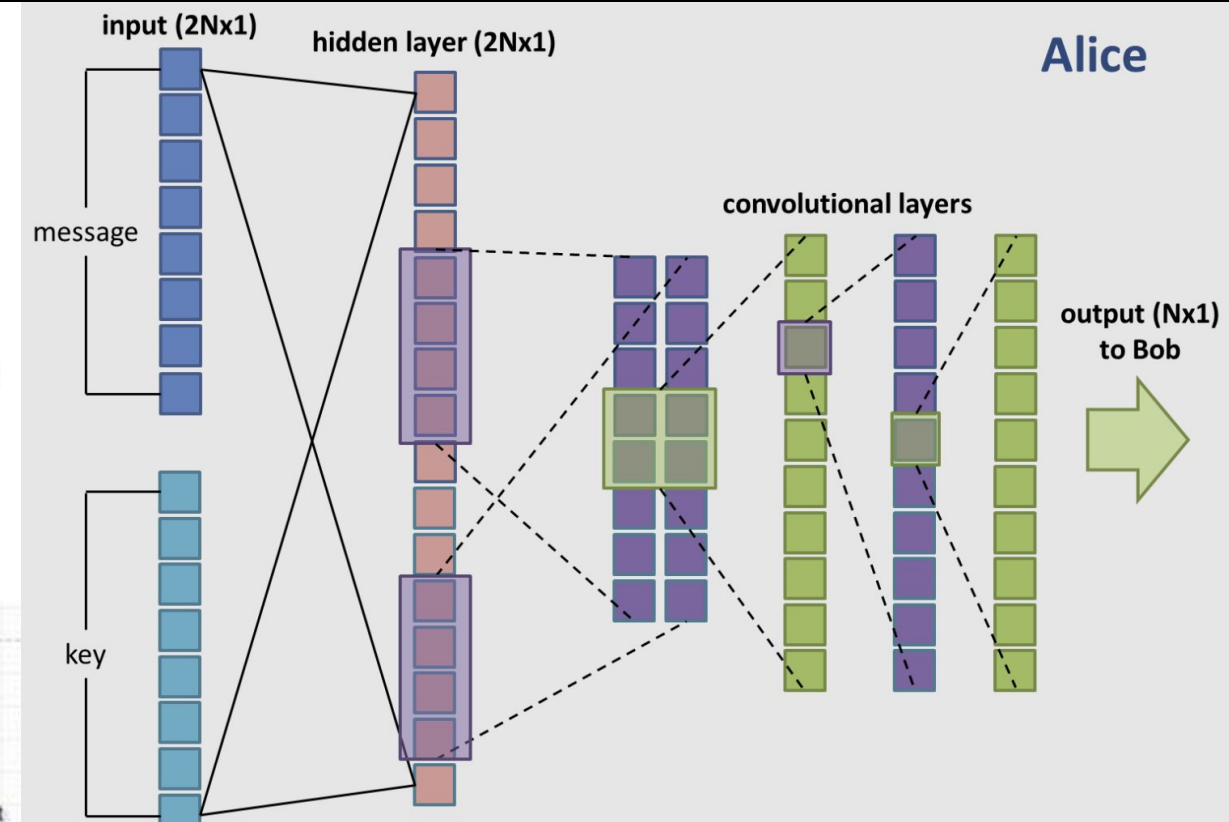
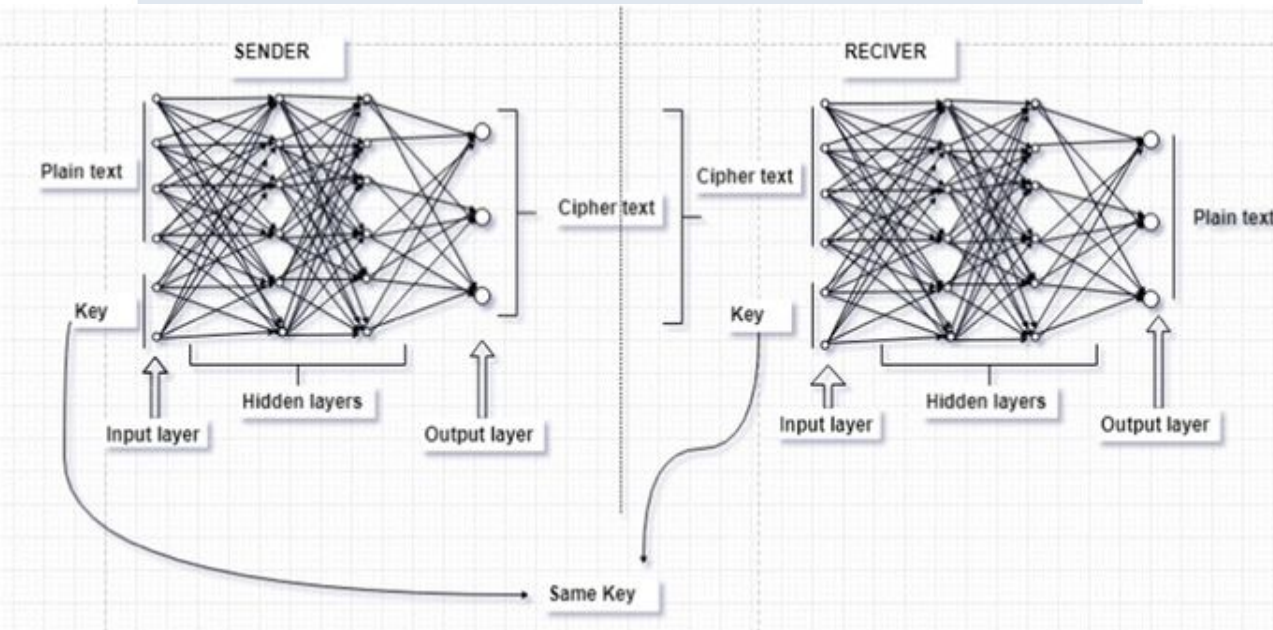
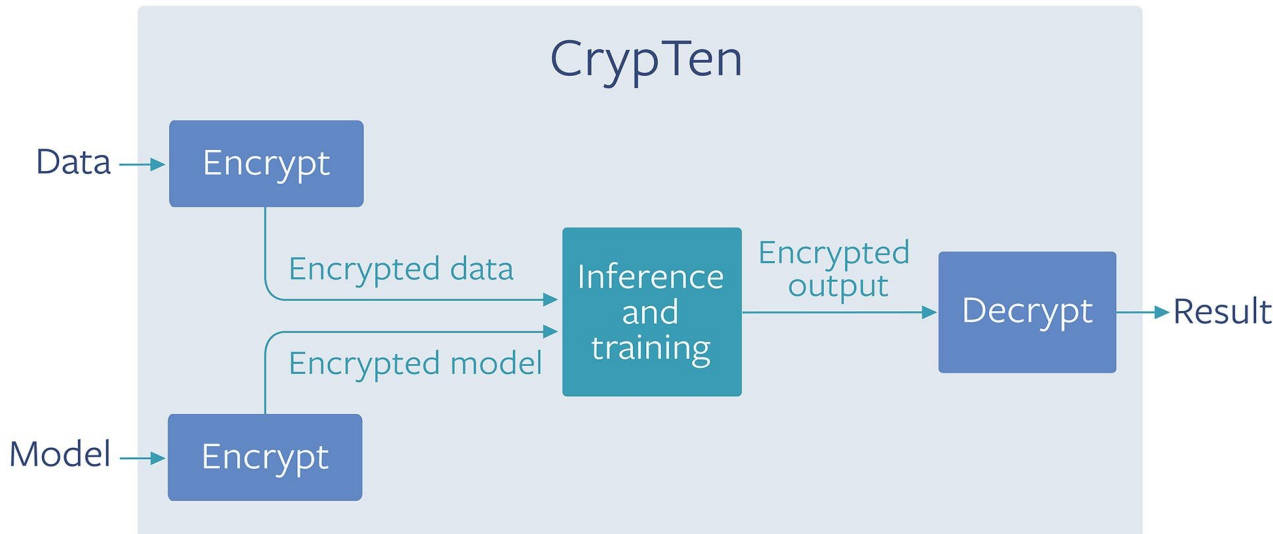
- Privacidad: Mantener la privacidad de los datos
- Equidad: que no favorezca determinadas salidas por sesgos implícitos.
- Trazabilidad: poder analizar los fallos del sistema.
- Robustez: hasta que punto podemos fiarnos del sistema.
- Fiabilidad: modelo fiable y si cambia.
- Causalidad: influir en la salida del modelo
- Explicabilidad y transparencia: los usuarios deben entender el funcionamiento del modelo.
- Gobernanza del dato: lícito, eficiente y eficaz de la información.

## Actividades abusivas en ML

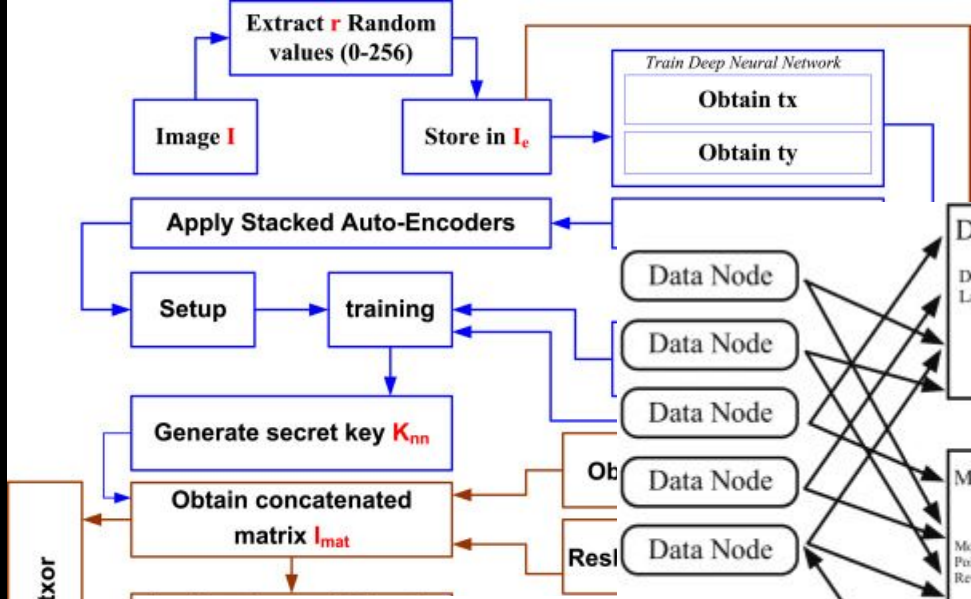
- Acceso y manipulación no autorizado a datasets y al proceso de transferencia de datos
- Acceso no autorizado al código del modelo
- Comprometer y limitar los resultados de AI
- Comprometer las inferencias y exactitud de los datos y algoritmos
- Envenenamiento de datos
- Elevar privilegios
- Manipulación en la optimización de los algoritmos
- Clasificación errónea basada en ejemplos de adversarios
- Envenenamiento del modelo
- Transferencia de ataques adversarios
- Escasez de datos

- Introducción de sesgos
- Manipulación de datos etiquetados
- Backdoors en datasets entrenados
- Comprometer los datos de validación en ML training
- Comprometer el aumento de datos
- Reducir la precisión de los datos
- Manipular el afinamiento de modelos
- DDoS
- Manipular ACLs para comprometer el preprocesamiento en ML
- Comprometer los frameworks a través de CVEs
- Degradación del modelo de ML

# IA en Criptografía



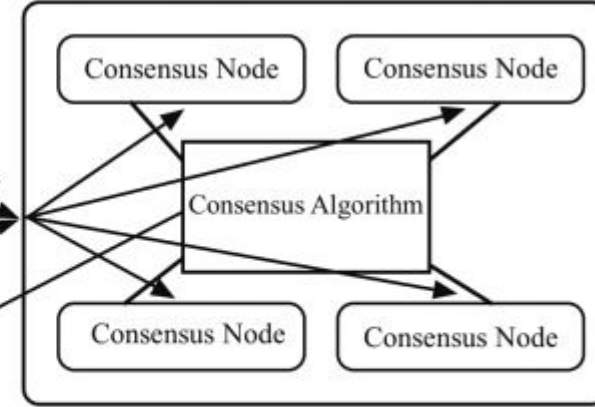




Buyer

1. Buyer creates a new contract.  
Here's an example:  
Criteria: Classify fraudulent orders w/ > 90% accuracy  
Data: Sample of 20k orders  
Reward: \$200 in ETH

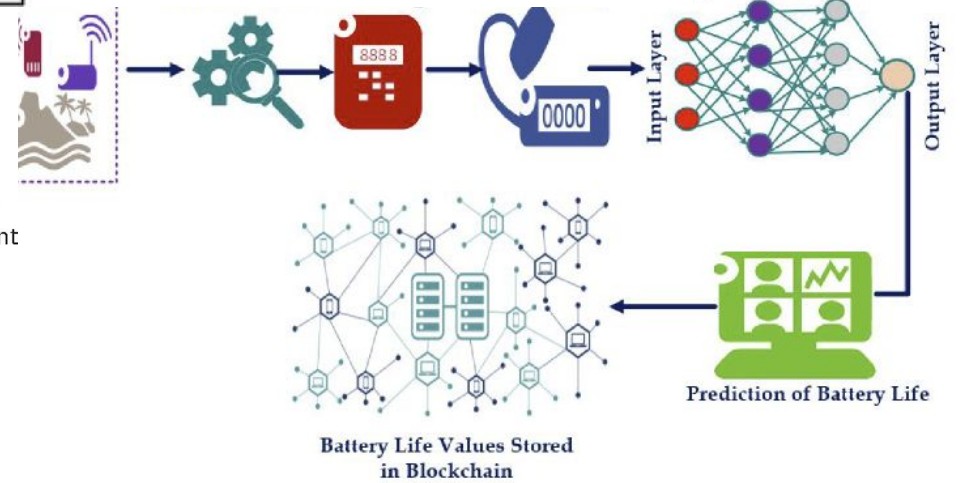
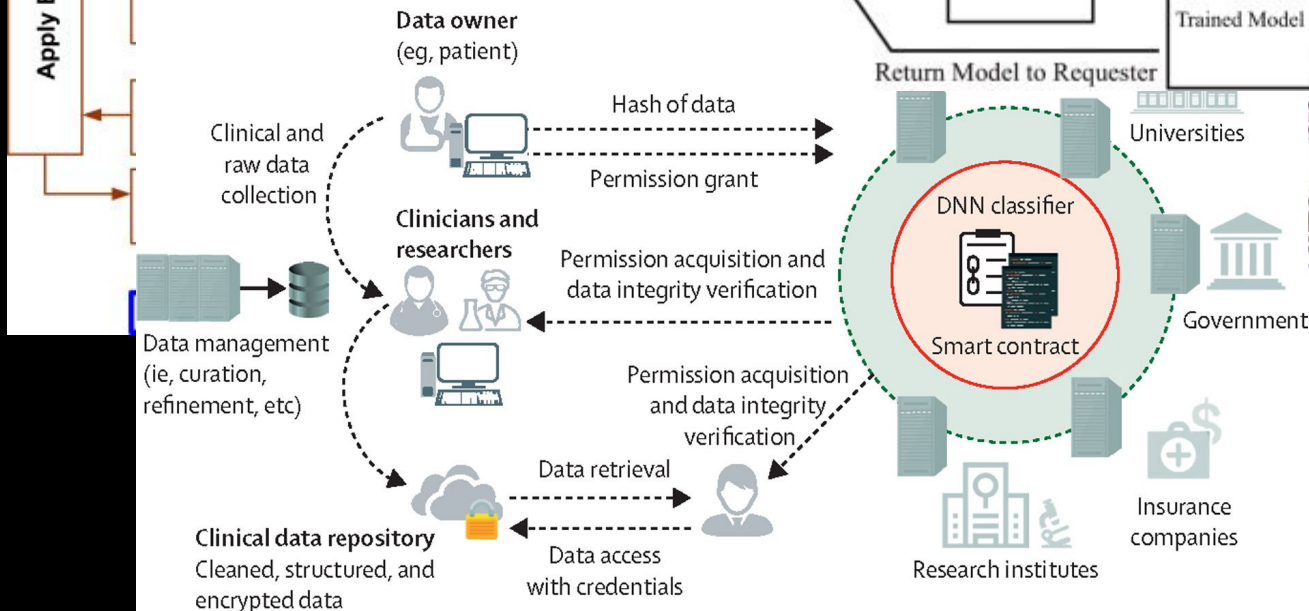
Contract is published to



Blockchain Network

5. If the model fulfills the criteria of the contract the model is sent to the buyer, and payment sent to the ML engineer.

The model is submitted and run on the Ethereum Blockchain using the data set from the contract.



VECTRA

SolutionsProductsPartnersAboutResourcesSchedule Demo

# See and stop threats across hybrid and multi-cloud enterprises

The Vectra threat detection & response platform captures packets and logs across your public cloud, SaaS, federated identity and data center networks. It applies patented security-led AI to surface, and prioritize threats and integrates into your security stack for rapid response.

See the Platform in Action

radware Blog

APPLICATION DELIVERYSECURITY

HomeSecurityApplication Security

SecurityApplication Security

## From Rule- to Machine Learning-Based Security

By Radware - March 10, 2021

f t in

Many enterprises have responded by implementing the aforementioned API management solutions that provide mechanisms, such as authentication, authorization and throttling. These are long-standing must haves for controlling who accesses APIs across the application ecosystem—and how often.

However, organizations also need to address the growth of more sophisticated attacks on APIs by complementing these “point” solutions with machine learning-driven security.

[You may also like: The 2020 App Threats Landscape in Review]

### Rule-Based

IBM Security QRadar

Offenses

Status: Open Advisor with Watson: High Pri... Clear Filters

Offenses by Magnitude

Offenses by Assignee

Offenses by Type

Magnitude	ID	Watson	Description	Start Date	Status	Type	Offense Source	Event Count	Source IP
6	2117	High priority	Potentially Successful Exploit co...	October 2, 2020 4:57 PM	Open	Source IP	INT 10.103.22.32	195	INT
6	2115	High priority	Potentially Successful Exploit co...	October 1, 2020 9:03 PM	Open	Source IP	INT 10.100.21.99	611	INT
6	2114	High priority	Potentially Successful Exploit co...	October 1, 2020 9:03 PM	Open	Source IP	INT 10.103.22.80	235	INT
6	2113	High priority	Potentially Successful Exploit co...	October 1, 2020 9:02 PM	Open	Source IP	INT 10.100.21.3	470	INT
6	2110	High priority	Potentially Successful Exploit co...	October 1, 2020 9:02 PM	Open	Source IP	INT 10.103.22.50	235	INT
6	2109	High priority	Potentially Successful Exploit co...	October 1, 2020 9:02 PM	Open	Source IP	INT 10.103.22.188	282	INT
4	2108	High priority	Actual action: Left alone precede...	October 1, 2020 9:02 PM	Open	Event Name	Security risk found, Actual action: Lef	47	INT
6	2106	High priority	Potentially Successful Exploit co...	October 1, 2020 9:01 PM	Open	Source IP	INT 10.103.22.35	282	INT

### Machine Learning

Machine-learning based application security solutions are adaptive by automatically detecting and responding to dynamic attacks and application/API vulnerabilities. First and foremost, they should automatically detect and protect new web applications as they are added to the network via automatic policy generation.

In addition, machine learning can eliminate API abuse such as token manipulations, parameter tampering, protocol attacks, invalid schemas and more. An enterprise-grade firewall should import, enumerate and catalog APIs to enforce standards and schemas using behavioral protections and positive security.

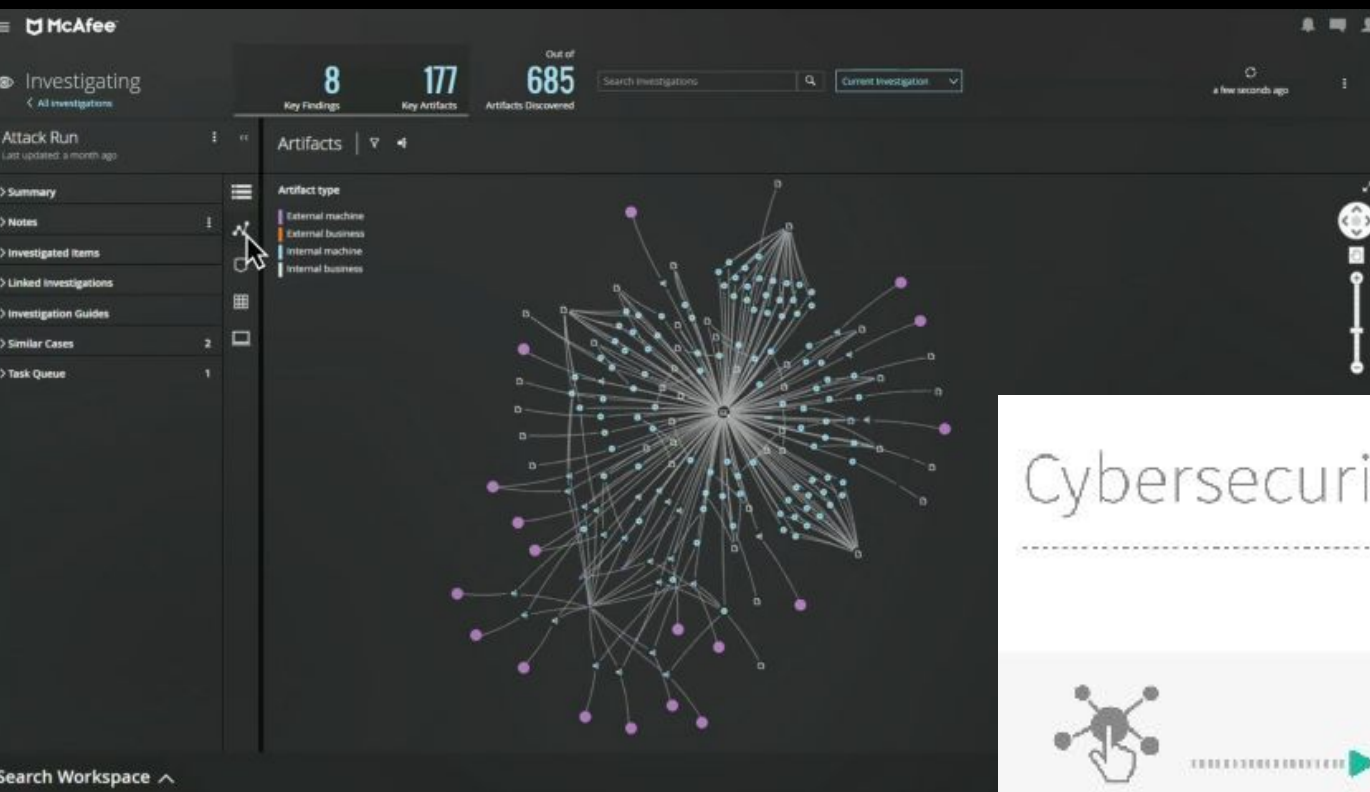
[You may also like: Application Security in 2021]

### Warning Signs

Here are seven warning signs your applications/APIs are vulnerable:

- Using non-defined/non-allowed HTTP methods for an API endpoint
- Embedding web attacks in JSON payloads or parameters
- Excessively utilizing the APIs
- Attempting to break the API authentication process through an account takeover attack
- Sending requests not according to the JSON/XML schemas
- An API key rotation – or a successful login from an unusual source
- Extremely high application usage from a single IP address or API token

## IBM Security QRadar Advisor with Watson



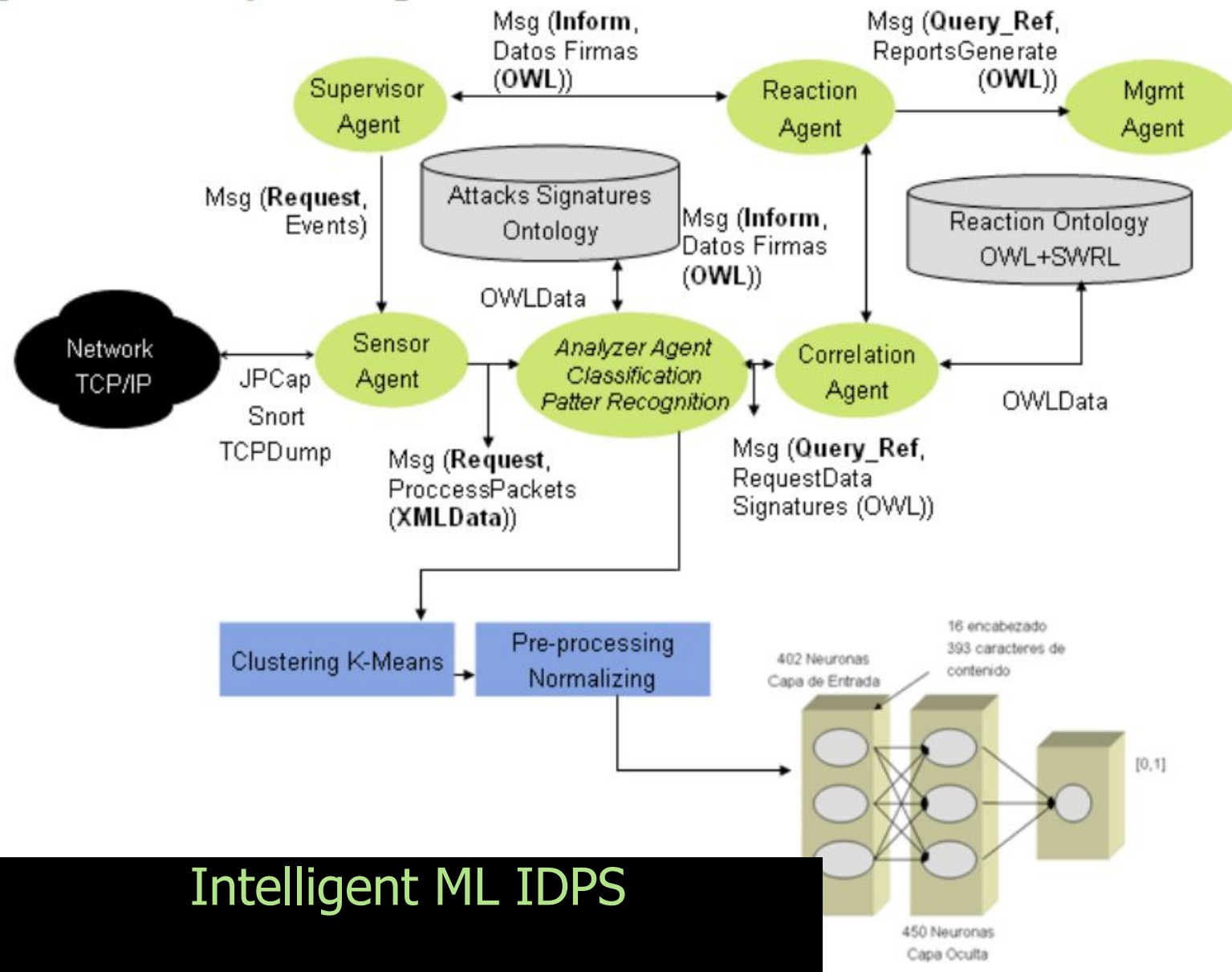
## Cybersecurity strategy CISCO + VU



Behavior Analysis



# Casos de uso IA CyberSec



Intelligent ML IDPS

Towards Ontology-Based Intelligent Model for Intrusion Detection and Prevention

G Isaza, AG Castillo-Sanz, L Castillo, MF Lopez  
Journal Of Information Assurance And Security 5 (2), 376

[https://link.springer.com/chapter/10.1007/978-3-642-04091-7\\_14](https://link.springer.com/chapter/10.1007/978-3-642-04091-7_14)

# Casos de uso IA CyberSec

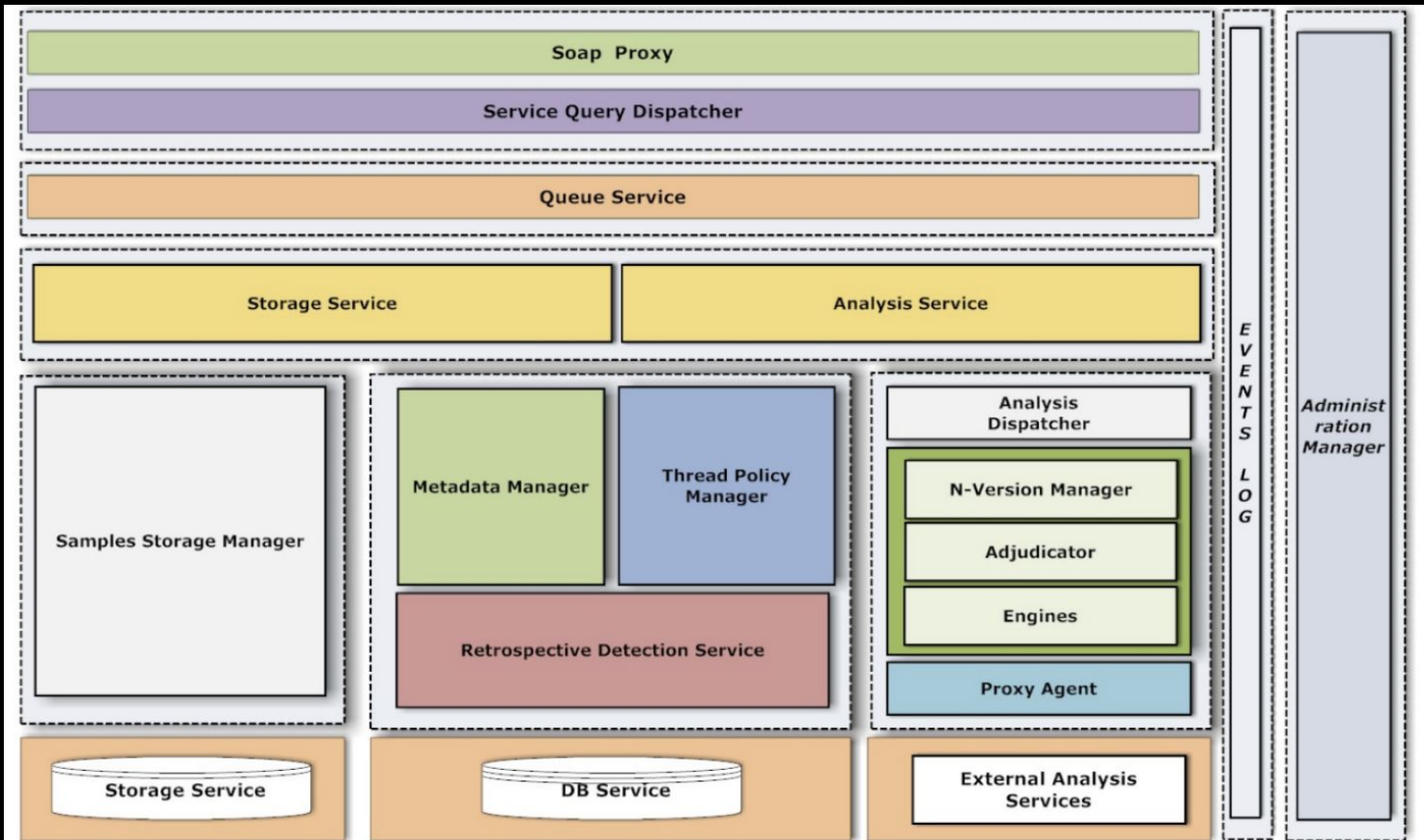


Figure 2. uCLAVS Architecture Components

## Cloud Antimalware Intelligent architecture

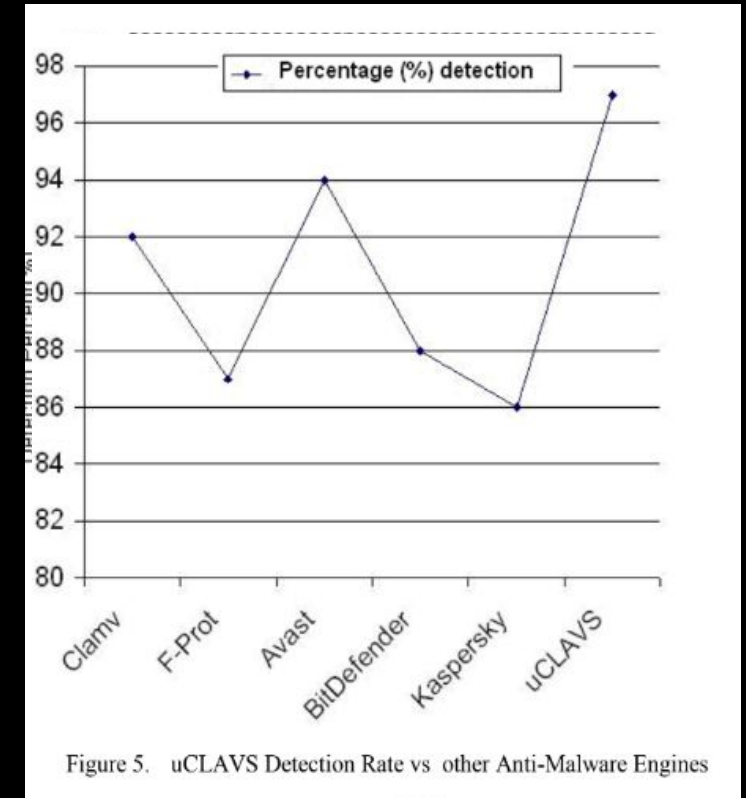
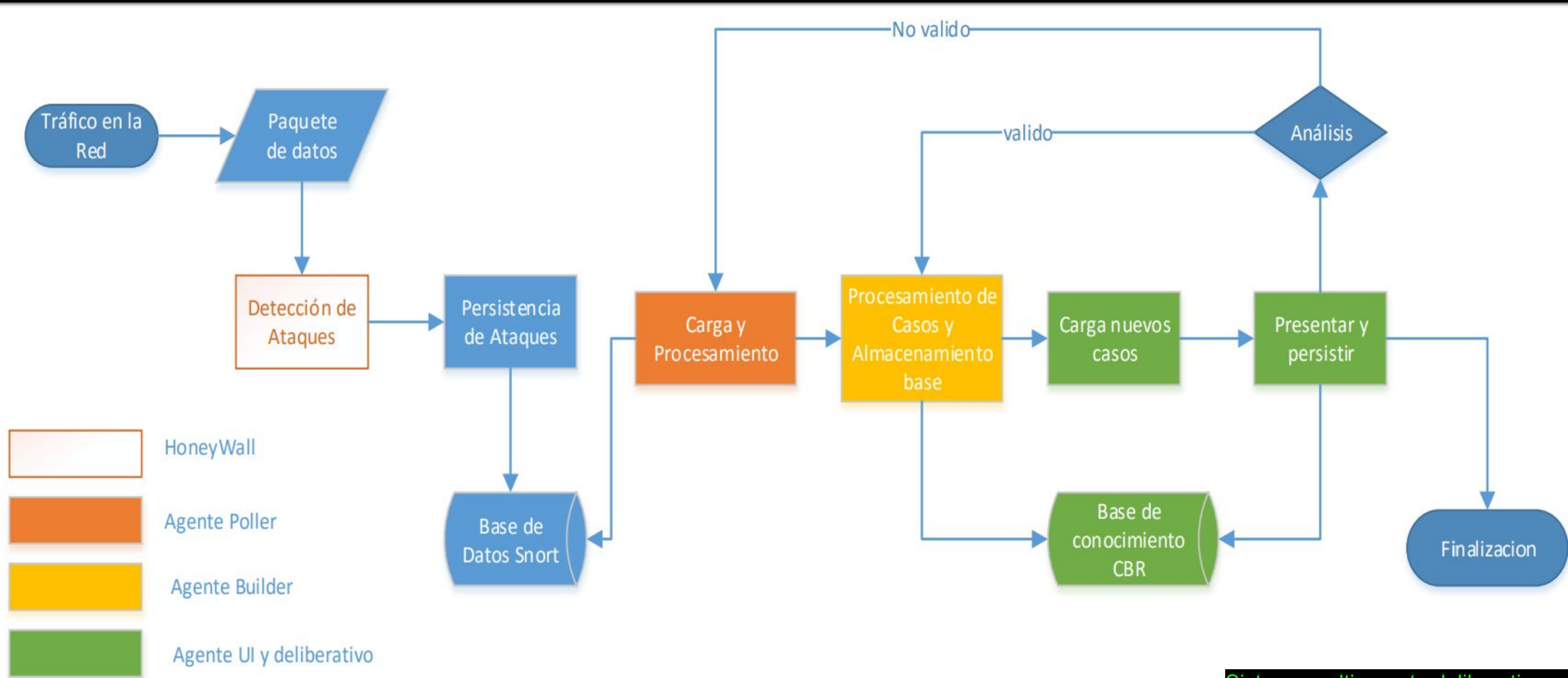


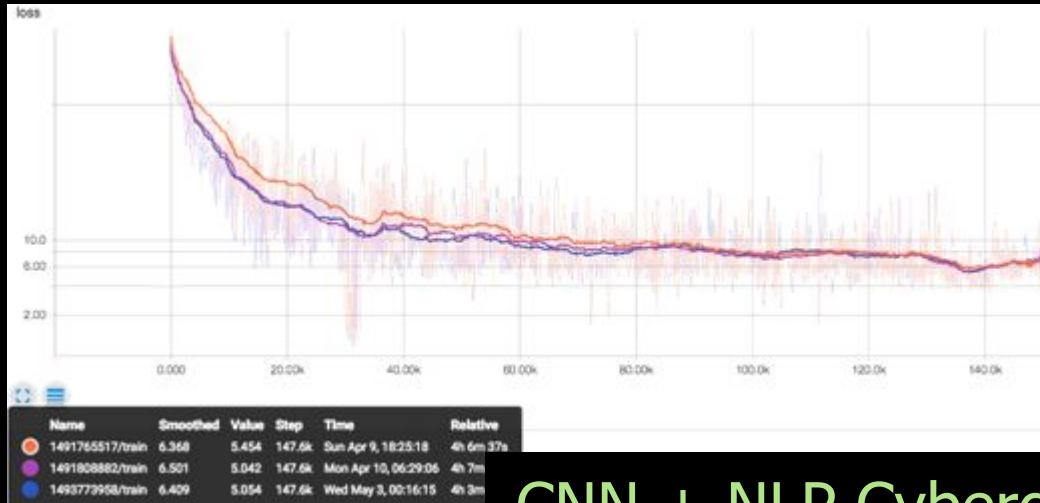
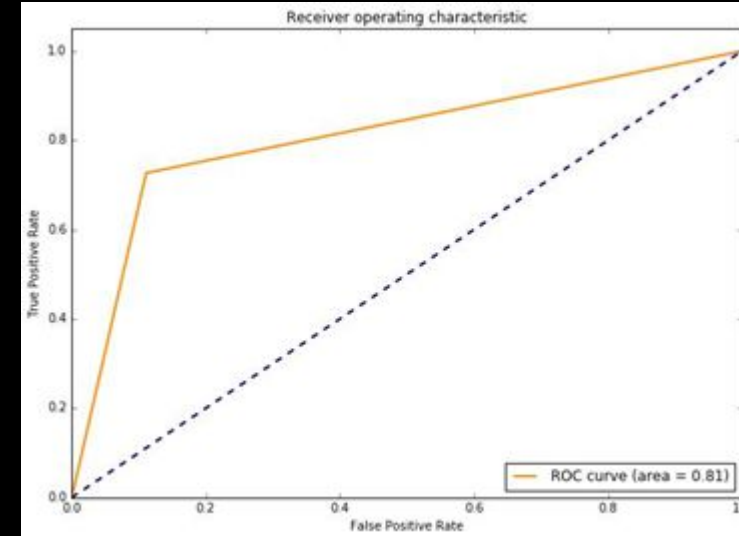
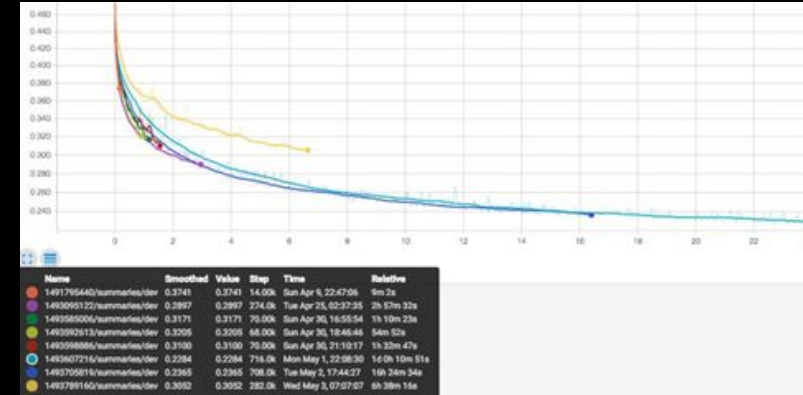
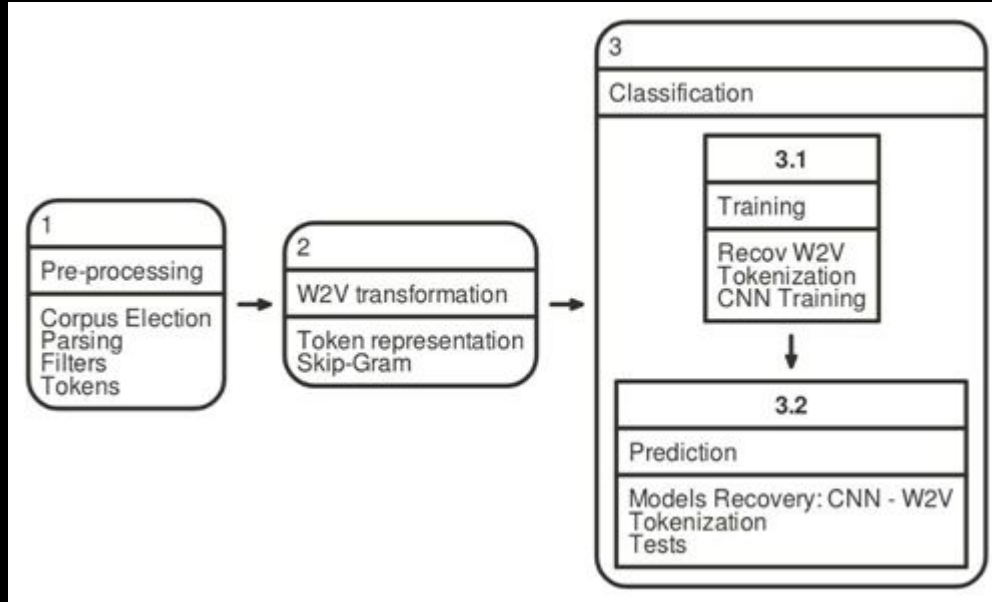
Figure 5. uCLAVS Detection Rate vs other Anti-Malware Engines

Malware detection based on cloud computing integrating intrusion ontology representation

CA Martínez, G Isaza, AGC Sanz  
Communications (LATINCOM), 2010 IEEE  
Latin-American Conference on, 1-6  
<https://ieeexplore.ieee.org/abstract/document/5641013>



# Casos de uso IA CyberSec



CNN + NLP Cybergrooming Detection  
Smartsec4COP

Classifying Cybergrooming for Child Online  
Protection Using Hybrid Machine Learning Model  
G Isaza, F Muñoz, L Castillo, F Buitrago  
Neurocomputing 484, 250-259  
<https://www.sciencedirect.com/science/article/abs/pii/S0925231221016489>

# Casos de uso IA CyberSec

## Automatic Pentesting

gisazae / SeedSSL

Watch 0Star 2Fork 0

CodeIssues 0Pull requests 0Projects 0SecurityInsights

Join GitHub today

50 million developers working together to host and manage projects, and build software together.

Dismiss

Sign up

Semillero de Seguridad Lumina (SSL)

8 commits1 branch0 releases1 contributor

Branch: masterNew pull requestFind fileClone or download

gisazae Update README.md

Latest commit 1ba1545 on 14 Mar

SeedIntegratedPrjDetectSpoofingTLS	Adding 2 prys	2 years ago
SemilleroInvestigacion_SeguridadSSL	Adding ShebiX P.Integrador ...	2 years ago
semillero_seguridadssl	1st	2 years ago
tests	1st	2 years ago
Dockerfile	1st	2 years ago
README.md	Update README.md	7 months ago
requirements.txt	1st	2 years ago
setup.py	1st	2 years ago

README.md

## Wrapper python Framework for Pentesting Tasks

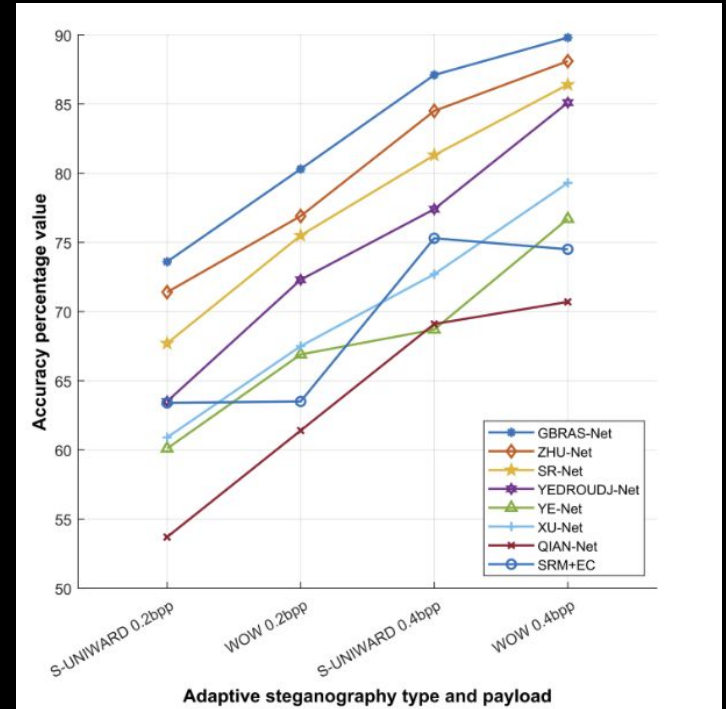
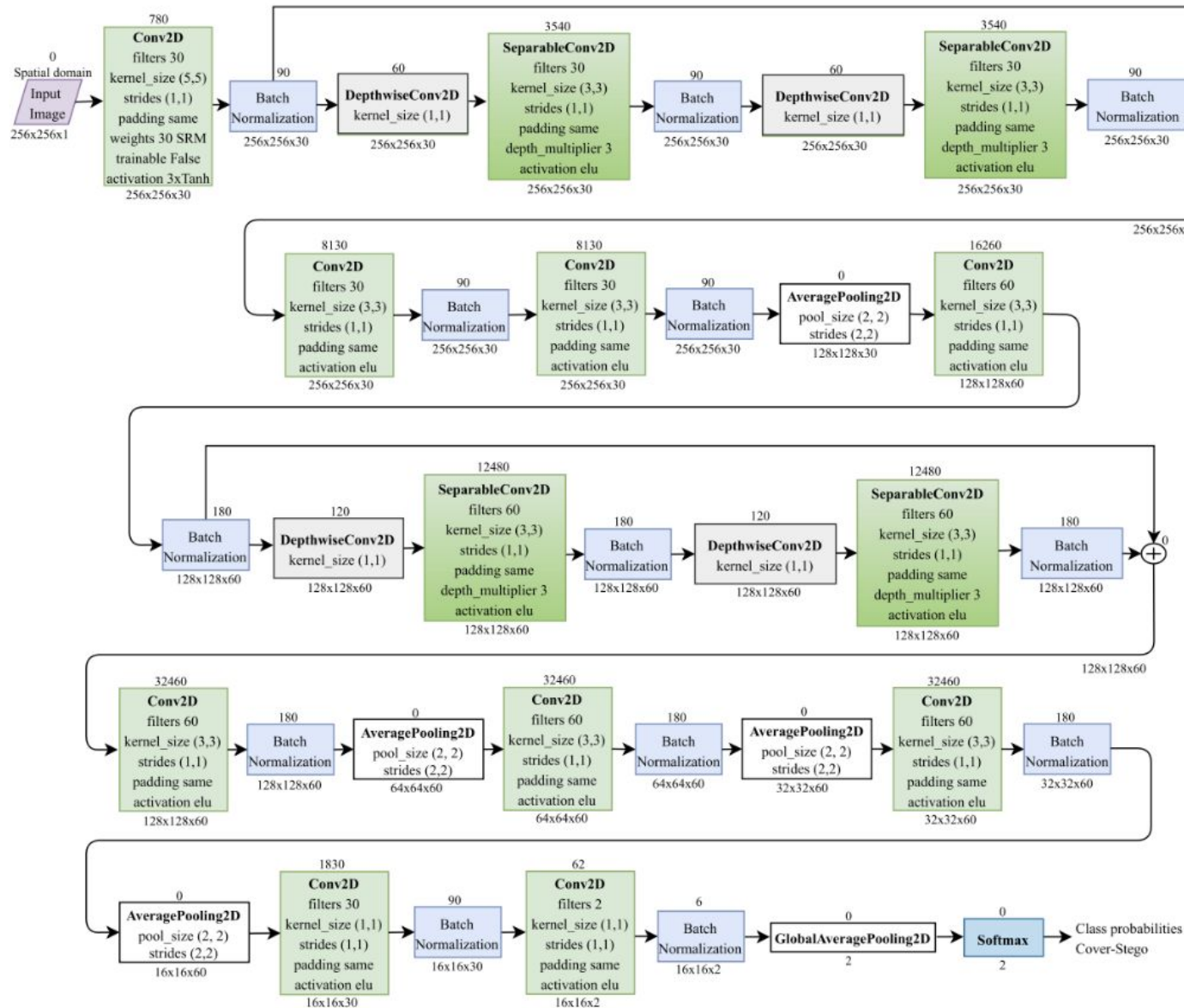
SemilleroInvestigacion\_SeguridadSSL - Universidad de Caldas

This tool is a meta-framework developed by some students as a project for the "SeguridadSSL" research seed of the "Universidad de Caldas".



# Casos de uso IA CyberSec

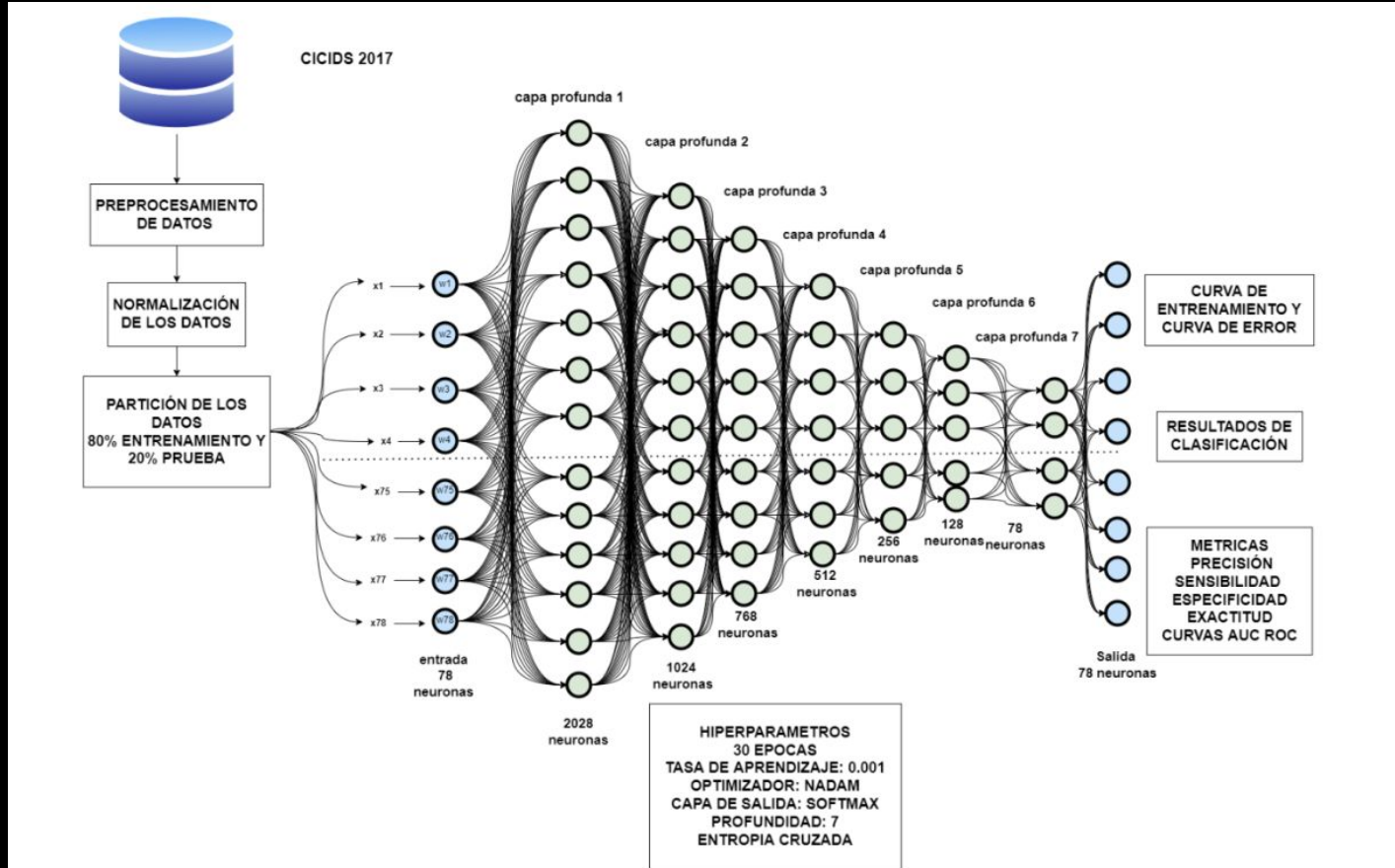
## GBRAS-Net



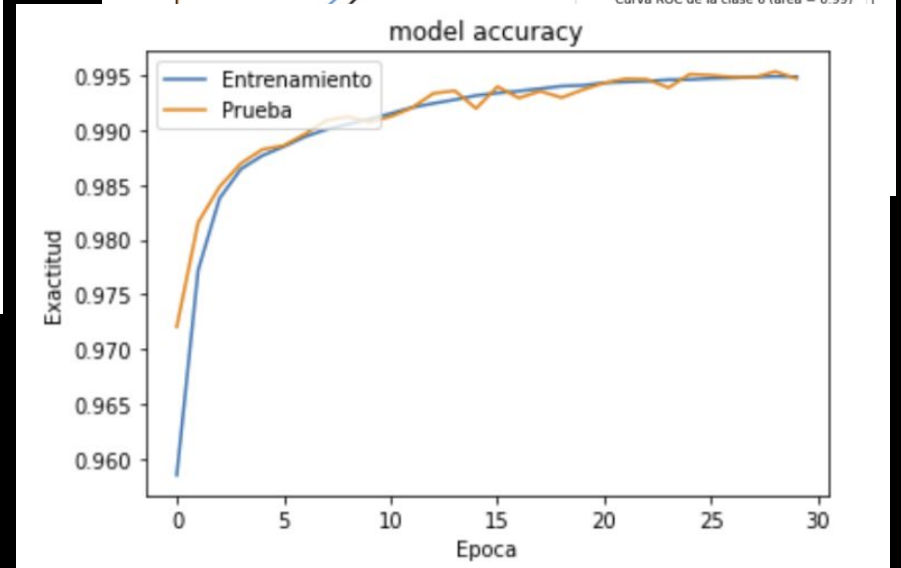
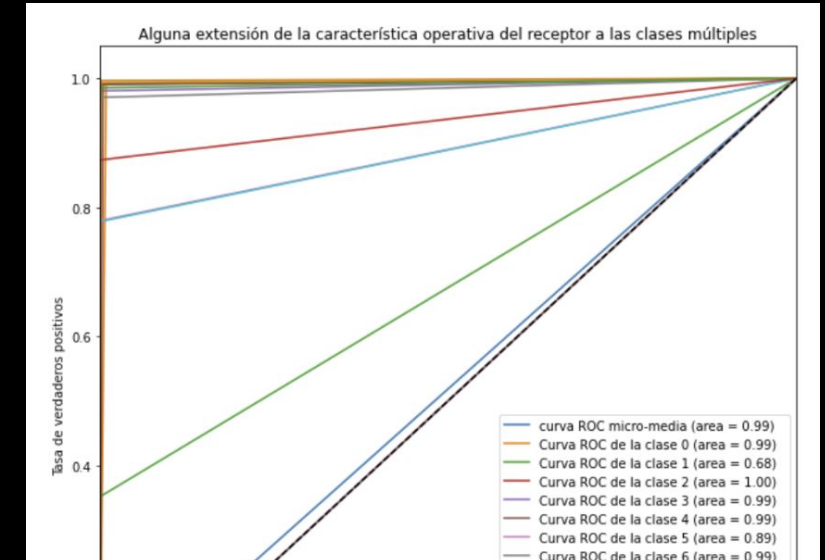
**FIGURE 5.** Comparison of the accuracy percentage of steganalysis among eight steganalysis methods against two algorithms: S-UNIWARD and WOW at 0.2 bpp and 0.4 bpp. All networks were trained and tested on BOSSbase 1.01 dataset, in image pairs (cover and stego) with 4000, 1000, 5000, respectively for the train, validation, and test data. Graph for test dataset performance.

Tabares Soto, Reinel & Ramos, Raúl & Isaza, Gustavo. (2019). Deep Learning Applied to Steganalysis of Digital Images: A Systematic Review. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2918086.  
 GBRAS-Net: A Convolutional Neural Network Architecture for Spatial Image Steganalysis IEEE Access  
<https://ieeexplore.ieee.org/document/9328287>

# Casos de uso IA CyberSec

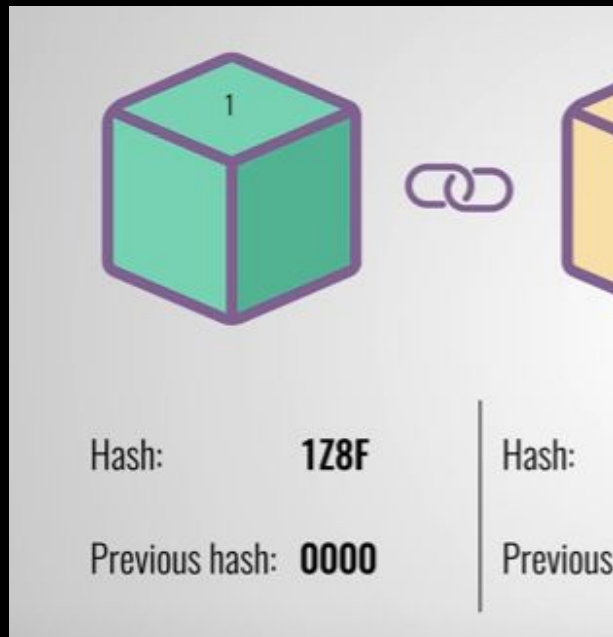


## DDoS Detection using ML/DL in cloud architecture



<https://github.com/gisazae/Tensorflow-Examples/blob/master/IntrusionDetectionMachineLearning.ipynb>

# Blockchains



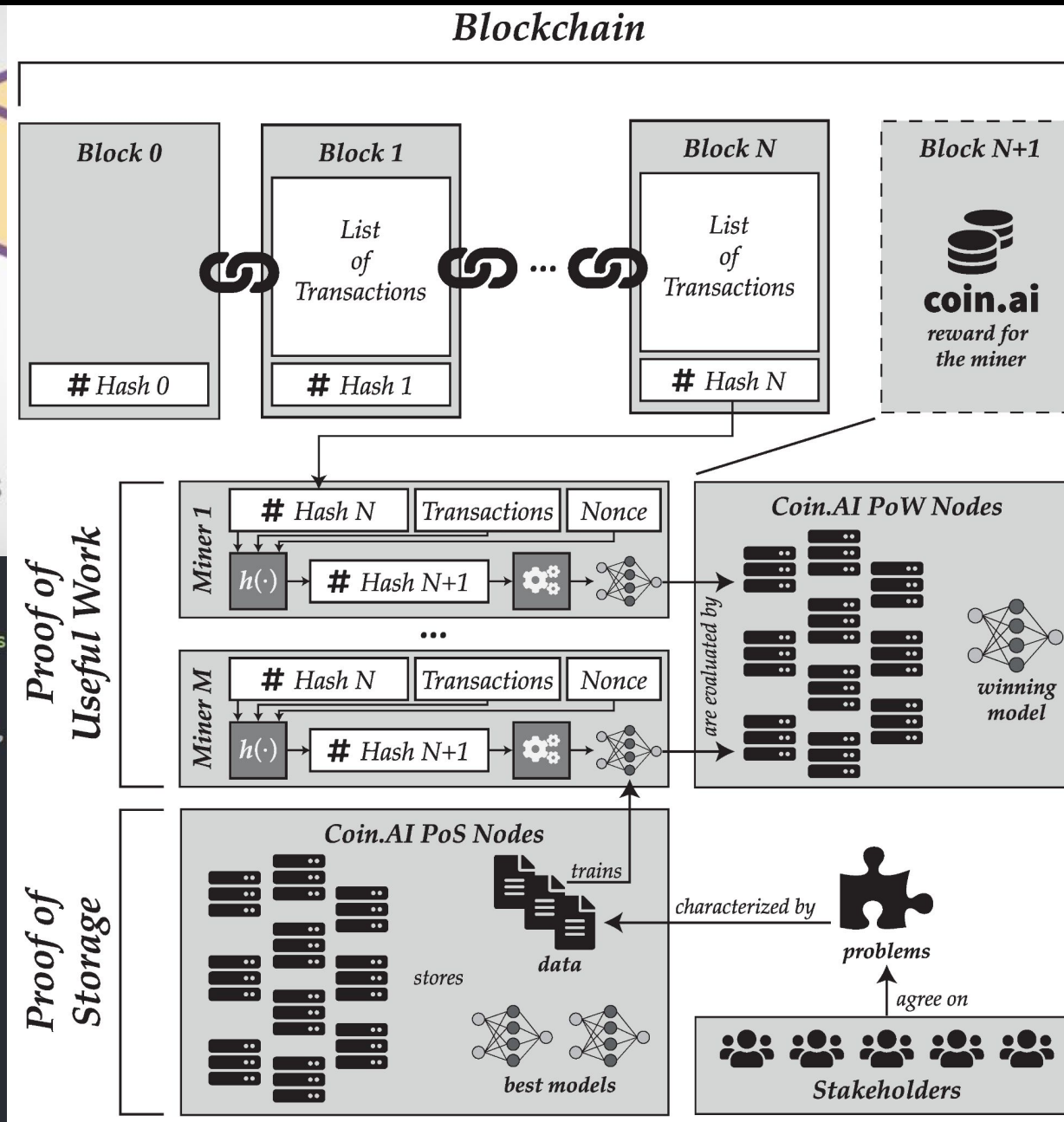
```
// block.js

const { GENESIS_DATA } = require('./genesis')

class Block {
  constructor({timestamp, lastHash, hash, data}) {
    this.timestamp = timestamp;
    this.lastHash = lastHash;
    this.hash = hash;
    this.data = data;
  }

  static genesis() {
    return new this(GENESIS_DATA);
  }
}

module.exports = Block;
```



Coin.AI: A Proof-of-Useful-Work Scheme for Blockchain-Based Distributed Deep Learning

<https://www.mdpi.com/1099-4300/21/8/723>

# Dificultades

- Representación de Datos
- Volumen de datos
- Entrenamiento en “tiempo real”
- Ética y sesgos
- Asignación de recursos
- Transición del rol operativo



???

THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.



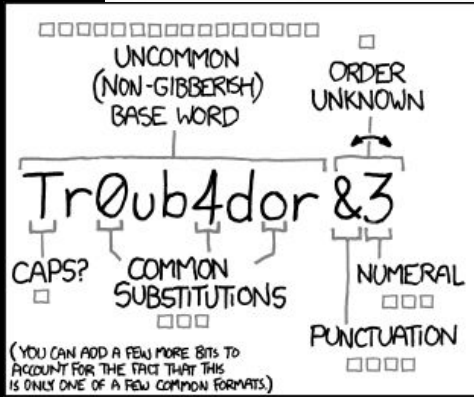
PASSWORD ENTROPY IS RARELY RELEVANT. THE REAL MODERN DANGER IS PASSWORD REUSE.



SET UP A WEBSERVICE TO DO SOMETHING SIMPLE, LIKE IMAGE HOSTING OR TWEET SYNDICATION, SO A FEW MILLION PEOPLE SET UP FREE ACCOUNTS.



BAM, YOU'VE GOT A FEW MILLION EMAILS, DEFAULT USERNAME, AND PASSWORDS.



~28 BITS OF ENTROPY

$2^{28} = 3 \text{ DAYS AT } 1000 \text{ GUESSES/SEC}$

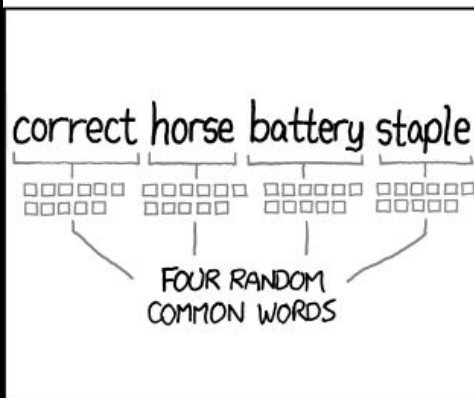
(PLAUSIBLE ATTACK ON A WEAK REMOTE WEB SERVICE. YES, CRACKING A STOLEN HASH IS FASTER, BUT IT'S NOT WHAT THE AVERAGE USER SHOULD WORRY ABOUT.)

DIFFICULTY TO GUESS: **EASY**

WAS IT TROMBONE? NO, TROUBADOR. AND ONE OF THE 0s WAS A ZERO?

AND THERE WAS SOME SYMBOL...

DIFFICULTY TO REMEMBER: **HARD**



~44 BITS OF ENTROPY

$2^{44} = 550 \text{ YEARS AT } 1000 \text{ GUESSES/SEC}$

DIFFICULTY TO GUESS: **HARD**

THAT'S A BATTERY STAPLE.

CORRECT!

DIFFICULTY TO REMEMBER: YOU'VE ALREADY MEMORIZED IT

THROUGH 20 YEARS OF EFFORT, WE'VE SUCCESSFULLY TRAINED EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS TO REMEMBER, BUT EASY FOR COMPUTERS TO GUESS.

# DATOS DE CONTACTO

| Research Gate | Google Citation Scholar |  
Github | LinkedIn