

# Aplicaciones de Minería de Datos I

## Lectura 1: Minería de Datos, Aprendizaje Máquina y Ciber-seguridad



# Tabla de contenidos

¿Qué está sucediendo actualmente?

¿Quiénes y como nos defienden?

Requisitos del curso

Evaluación del curso

Minería de Datos el pre-ambulo de la Inteligencia Artificial y el Aprendizaje Máquina

Aprendizaje Supervisado

¿Cómo se aplican los controles inteligentes?

Retos de la Inteligencia Artificial

Sistemas defensivos de Inteligencia Artificial

Sistemas ofensivos con Inteligencia Artificial

Sistemas adversarios con Inteligencia Artificial

Cybersecurity data science: an overview from machine learning perspective

Referencias



*Los datos te hablan, si estás dispuesto a escucharlos.*



# ¿Qué está sucediendo actualmente? (1/2)

## **Nutanix amplía la protección frente a ransomware en su plataforma cloud**

Seguridad 23 FEB 2021

Cafeteras que minan Monero presas del ransomware: la pesadilla por la seguridad de la internet de las cosas nos acecha

Seguridad

**Windows Defender detecta actualizaciones de Chrome como si fuera malware**



## ¿Qué está sucediendo actualmente? (2/2)

- ▶ ¿Nuevas y sofisticadas amenazas? (métodos de ofuscación, cifrado robusto, seguridad por oscuridad y persistentes)
- ▶ ¿Ataques a cadenas de suministro?
- ▶ ¿Ataque patrocinados por gobiernos?
- ▶ ¿Gobernación de grandes contenidos de información difíciles de detectar, analizar y categorizar?
- ▶ ¿Tendencia hacia nuevos vectores de ataque **más inteligentes (automatización, deep-fakes, modelos largos de lenguaje, ...)**?



# ¿Quiénes y como nos defienden? (1/2)

**Primer ataque DDoS usando tres millones de cepillos de dientes inteligentes con malware**

**El malware se hace fuerte en la nube**

**Con esta versión maligna de ChatGPT crear malware nunca había sido tan fácil**

*Ciberdelincuentes crearon a WormGPT, un sistema basado en la tecnología de OpenAI, con fines perniciosos. Conoce cómo funciona*

**Google revela una amenaza global a la privacidad en 2024: ¿estás en riesgo?**



## ¿Quiénes y como nos defienden? (2/2)

- ▶ Políticas y reglas en dispositivos lógicos y de seguridad perimetral
- ▶ Firmas de comportamiento y de anomalías en aplicaciones y en tráfico de red
- ▶ Sandboxes de defensa
- ▶ Listas blancas/negras o filtrado
- ▶ Detección Heurística
- ▶ Sistemas de control de acceso

*¿Cuál es el resultado ante las crecientes y más sofisticadas amenazas?*

**Generación de un alto radio de FALSOS POSITIVOS**



# Requisitos del curso

- ▶ Equipo individual con Anaconda Framework instalado ([Descarga de Anaconda](#))
- ▶ Crear una cuenta en [Google Colab](#)
- ▶ Experiencia básica de Programación en Python [pueden revisar este curso introductorio](#)
- ▶ [Repositorio oficial de Lecturas y Código](#)
- ▶ Canal del equipo de [Microsoft Teams](#)





# Evaluación del curso

La evaluación se llevará a cabo mediante un proyecto final, el cual será redactado en *Jupyter Notebook*. Dicho proyecto incluirá una base de datos pertinente al ámbito de la ciberseguridad y será valorado conforme a las rúbricas que a continuación se detallan:

- ▶ Introducción: estado del arte, trabajos relacionados, entre otros
- ▶ Hipótesis del trabajo
- ▶ Resultados con código fuente: métricas de desempeño, comparación de resultados, entre otros
- ▶ Conclusiones



# Minería de Datos el pre-ambulo de la Inteligencia Artificial y el Aprendizaje Máquina

- ▶ Las soluciones escritas de manera manual, con ingeniería inversa, matching/flagging de amenazas o análisis dinámico no automatizado pueden provocar un rezago ante **nuevos riesgos más sofisticados** (*ofuscación de código, técnicas criptográficas, magecart, spear phishing, etc.*)
- ▶ La constantes actualizaciones de firmas de comportamiento y reglas de asociación **no son óptimas** en datos provenientes de diferentes orígenes (*cloud computing, IoT, SaaS, PaaS, etc.*)



# Conceptos Fundamentales (1/17)

- ▶ **Datos:** una descripción elemental de objetos, eventos, actividades y transacciones
- ▶ **Información:** datos organizados que tienen fundamento y valor
- ▶ **Conocimiento:** el concepto de entender información basada en el reconocimiento de patrones y provee intuición
- ▶ **Vulnerabilidad:** debilidad en el software, hardware o procedimiento que puede permitir un acceso o maniobra no autorizada. Se caracteriza por la ausencia o debilidad de un control, permitiendo la explotación
- ▶ **Amenaza:** cualquier daño potencial a la información o los sistemas



# Conceptos Fundamentales (2/17)

- ▶ **Riesgo:** probabilidad de que un elemento de amenaza tome ventaja de una vulnerabilidad y genere un impacto al negocio
- ▶ **Control:** contra-medida implementada para mitigar (o reducir) el riesgo potencial

En la siguiente Figura se muestran algunos tipos de amenazas, que se pueden considerar



# Conceptos Fundamentales (3/17)



# Conceptos Fundamentales (4/17)

Por lo tanto:

*La minería de datos es una técnica que busca encontrar patrones valiosos en grandes conjuntos de datos, mediante metodologías, técnicas y la aplicación de algoritmos*

**¿Cómo resolver los nuevos patrones de amenaza?**

*Debemos analizar el fenómeno de los datos en el panorama de ciber-seguridad, mediante métodos científicos, de minería de datos, inteligencia artificial y procesos predictivos, que protejan de amenazas sofisticadas, mediante controles inteligentes, reduciendo el riesgo inherente.*



## ¿Qué es la Inteligencia Artificial?

*Desarrollo y utilización de computadoras con los que se intenta reproducir los procesos de la inteligencia humana*

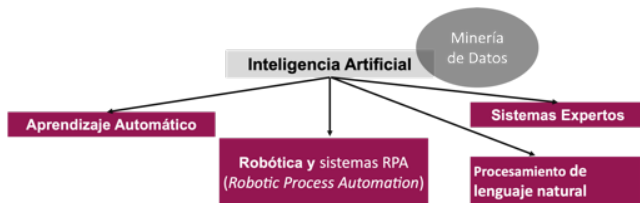


Figura 1: Relación entre la Inteligencia Artificial y la Minería de Datos

## ¿Cómo ayuda la Minería de Datos en el proceso de análisis de malware?

*Ayuda a los analistas a examinar datos, conjuntos de datos o metadatos para identificar patrones, anomalías y tendencias para identificar comportamiento malicioso y proveer valores para eventos futuros.*





## La experiencia es un punto clave en la comprensión de datos

*Un programa de una computadora aprende de la experiencia  $E$  con respecto a una clase de tareas  $T$  y una medida de rendimiento  $P$ , si su rendimiento en las tareas  $T$ , medido en base a la medida  $P$ , mejora con la experiencia  $E$*



# Conceptos Fundamentales (8/17)



## Ejemplo de Aprendizaje

*Aprender a detectar robos de tarjetas de crédito*

- **T:** detectar robos de tarjetas de crédito.
- **P:** porcentaje de robos detectados.
- **E:** base de datos de hábitos de compra con la tarjeta de crédito.

Figura 2: Ejemplo de aprendizaje

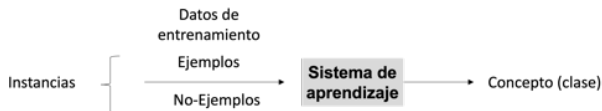


# Conceptos Fundamentales (9/17)

¿Qué conceptos se pueden aprender?



¿Qué elementos intervienen en el aprendizaje de un concepto?



¿En qué consiste aprender un concepto?



Figura 3: Aprendizaje y tareas básicas



# Conceptos Fundamentales (10/17)

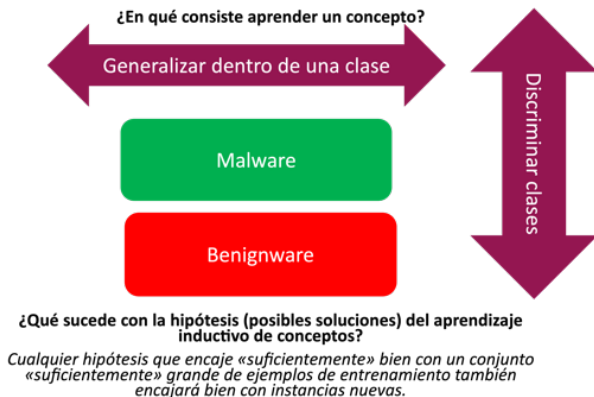


Figura 4: ¿En qué consiste aprender un concepto?

# Conceptos Fundamentales (11/17)

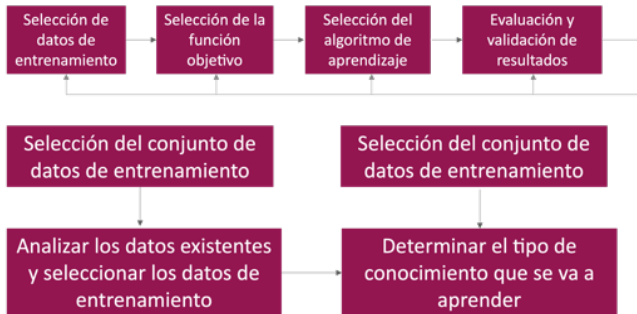


Figura 5: ¿Qué etapas comprende el aprendizaje de un concepto?

# Conceptos Fundamentales (12/17)

En la siguiente Figura se muestra los tipos de aprendizaje en términos de Inteligencia Artificial

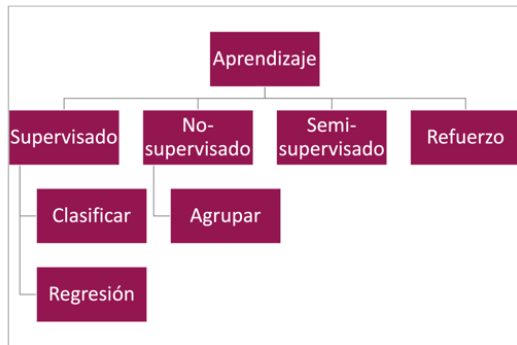


Figura 6:

# Conceptos Fundamentales (13/17)

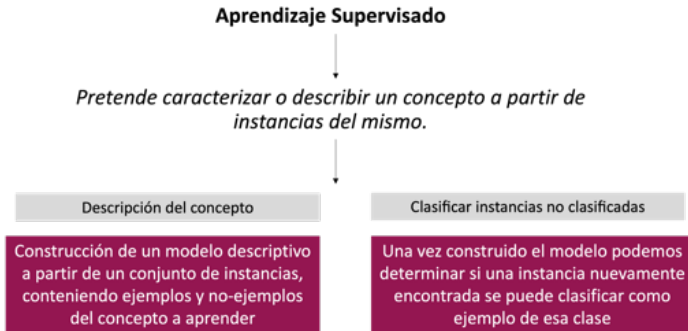


Figura 7:

# Conceptos Fundamentales (14/17)

Algunas ramas de la Inteligencia Artificial son las siguientes:

- ▶ **Supervisada:** proveer algoritmos que aprendan de muchos datos de entrenamiento etiquetados (clasificar)
- ▶ **No supervisada:** proveer algoritmos que aprendan algún problema de optimización (agrupar)
- ▶ **Deep Learning:** expande y verifica la información de manera más semejante al cerebro humano
- ▶ **Por refuerzo:** maximizar la generalización del problema, aprender por experiencia





# Conceptos Fundamentales (15/17)

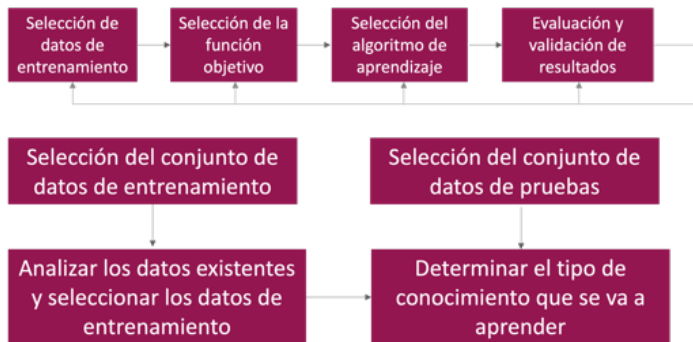


Figura 8:



# Conceptos Fundamentales (16/17)

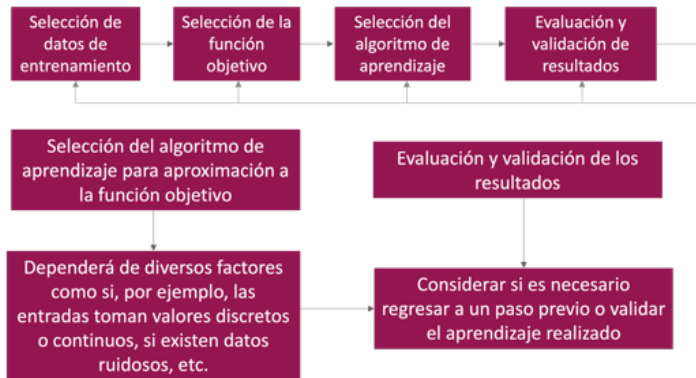


Figura 9:



## **Ingeniería de Conocimiento**



*Proceso de construir un sistema inteligente para obtener conocimiento*



## **Descubrimiento de Conocimiento en Bases de Datos**



*Procedimiento completo necesario para extraer conocimiento potencialmente útil y previamente desconocido a partir de los datos en una base de datos*



# Aprendizaje Supervisado (1/5)

En la siguiente figura se describe el flujo de trabajo del paradigma de Aprendizaje Supervisado

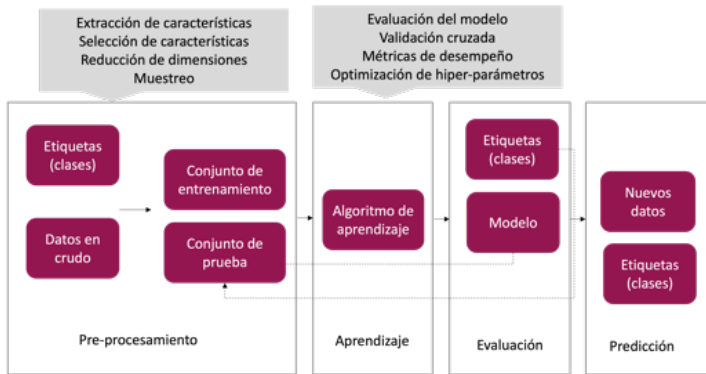


Figura 12:

# Aprendizaje Supervisado (2/5)

## Las etiquetas de clase

- ▶ La etiqueta es la salida final de la muestra, es **lo que representa** o la categoría de clase a la que pertenece
- ▶ A veces las etiquetas o clases contienen categorías, a lo que les llamamos **etiquetas categóricas**, como por ejemplo:  $\{H, M\}$ , para distinguir entre Hombre y Mujer ó  $\{B, M\}$  para distinguir entre *malware* y *benignware*. Otro tipos de etiquetas o clases son denominadas *discretas* y representan valores numéricos **bien separables**

Para poder evaluar el desempeño de un algoritmo de aprendizaje supervisado, se emplea una matriz de confusión, la cual representa un resumen de la predicción de resultados en un problema de clasificación, donde la salidas pueden ser dos o más clases. Ver Figura 36.



# Aprendizaje Supervisado (3/5)

	Malware (1) actual 	No Malware (0) actual 
Malware (1) predicho 	Verdaderos positivos	Falsos Positivos
No Malware (0) predicho 	Falsos Negativos	Verdaderos Negativos

Figura 13:

# Aprendizaje Supervisado (4/5)

Donde los valores acumulados se describen continuación:

- ▶ **Verdaderos Positivos (VP)** : las muestras etiquetadas como *fraude* y que fueron clasificadas como tal
- ▶ **Verdaderos Negativos (VN)**: las muestras etiquetadas *no fraude* como que fueron clasificadas como tal
- ▶ **Falsos Positivos (FP)**: las muestras etiquetadas como *fraude* que fueron clasificadas como *no fraude*
- ▶ **Falsos Negativos (FN)**: las muestras etiquetadas como *no fraude* que fueron clasificadas como *fraude*

De donde se desglosan las siguientes métricas de desempeño:

- ▶ **Precisión**: es la fracción de instancias relevantes dentro de las clasificadas, es decir que tan bien clasifica el modelo las muestras

$$Presición = \frac{VP}{VP + FP} \quad (1)$$



## Aprendizaje Supervisado (5/5)

- **Sensitividad:** es la fracción de instancias relevantes que se han obtenido sobre la cantidad total de instancias relevantes

$$\text{Sensitividad} = \frac{VP}{VP + FN} \quad (2)$$

- **Puntaje F1:** mide la media armónica de la precisión y la sensitividad

$$\text{Puntaje } F-1 = \frac{2 \cdot \text{sensitividad} \cdot \text{precisión}}{\text{sensitividad} + \text{precision}} \quad (3)$$

- **Macro Precisión:** calcula la métrica independientemente de cada clase y después toma el promedio (tratando todas las clases de manera igual)
- **Precisión con peso:** calcula la métrica tomando en cuenta la el valor de precisión para cada clase





# ¿Cómo se aplican los controles inteligentes?

## ► Supervisada y Deep Learning

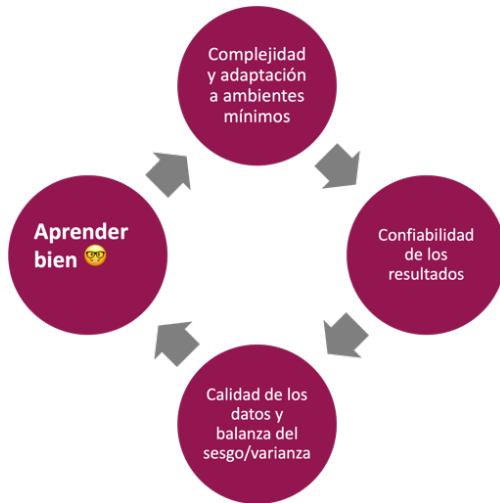
- Clasificación de malware
  - *Se pueden entrenar millones de muestras*
  - *Decremento a los falsos positivos*
- Identificación de spam
- Firewalls
  - *Analiza de manera exponencial bitácoras de firewalls para predecir y realizar puntajes de fuentes maliciosas desconocidas*

## ► No supervisada

- Análisis a los DNS
  - *Agrupación de nombres de dominio, frecuencias de lookup, etc.*
- Inteligencia de amenazas
  - *Priorización, de-duplicación*
- Análisis de comportamiento de usuarios
  - *Reglas de autenticación, estadísticas de uso, etc.*



# Retos de la Inteligencia Artificial



# Sistemas defensivos de Inteligencia Artificial



**Detección de malware:**  
Multicapa, multidefensa



**SOC, IDS/IPS y Honey Pots:**  
Auto aprendizaje y Deep Learning



**Anti-SPAM**



**Gestión de vulnerabilidades:**  
Identificar y priorizar remediación



**Clasificación de datos:**  
Rastrear datos para  
identificar, clasificar y  
proteger



**Inteligencia de amenazas:**  
Categorizar el  
comportamiento  
Analizar tráfico (¿Deep  
web?)

# Sistemas ofensivos con Inteligencia Artificial



## **Creación de malware:**

Creación en masa,  
fortaleza a capacidades  
de evasión



## **Botnets inteligentes:**

Botnets con aprendizaje  
automático  
Zombies inteligentes



## **Spear phishing:**

Ingeniería social inteligente



## **IA de adversarios:**

Redes generativas antagónicas:  
envenenar procedimientos de IA,  
producir resultados controlados o  
falsos



## **Ataques condicionales:**

Ataques inteligentes a blockchain y  
contratos



## **Clasificación de víctimas:**

Optimizar ataques dirigidos a  
organizaciones

# Sistemas adversarios con Inteligencia Artificial



## **Entradas de adversarios:**

Artefactos diseñados para engañar sistema de IA



## **Robo de modelos:**

Para aumentar las capacidades de entradas adversarias



## **Envenenamiento de datos:**

Entrenar datos envenenados en las herramientas de IA



## **Ataques de feedback:**

Envenenar los modelos de IA para generar DoS



# Lectura Recomendada

*Cybersecurity data science: an overview from machine learning perspective*

¿Cuál es la importancia de la ciencia de datos en ciber-seguridad?



# Referencias

-  Lin, C. T., Wang, N. J., Xiao, H., & Eckert, C. (2015). Feature Selection and Extraction for Malware Classification. J. Inf. Sci. Eng., 31(3), 965-992.
-  Layton, R. (2015). Learning data mining with python. Packt Publishing Ltd.

