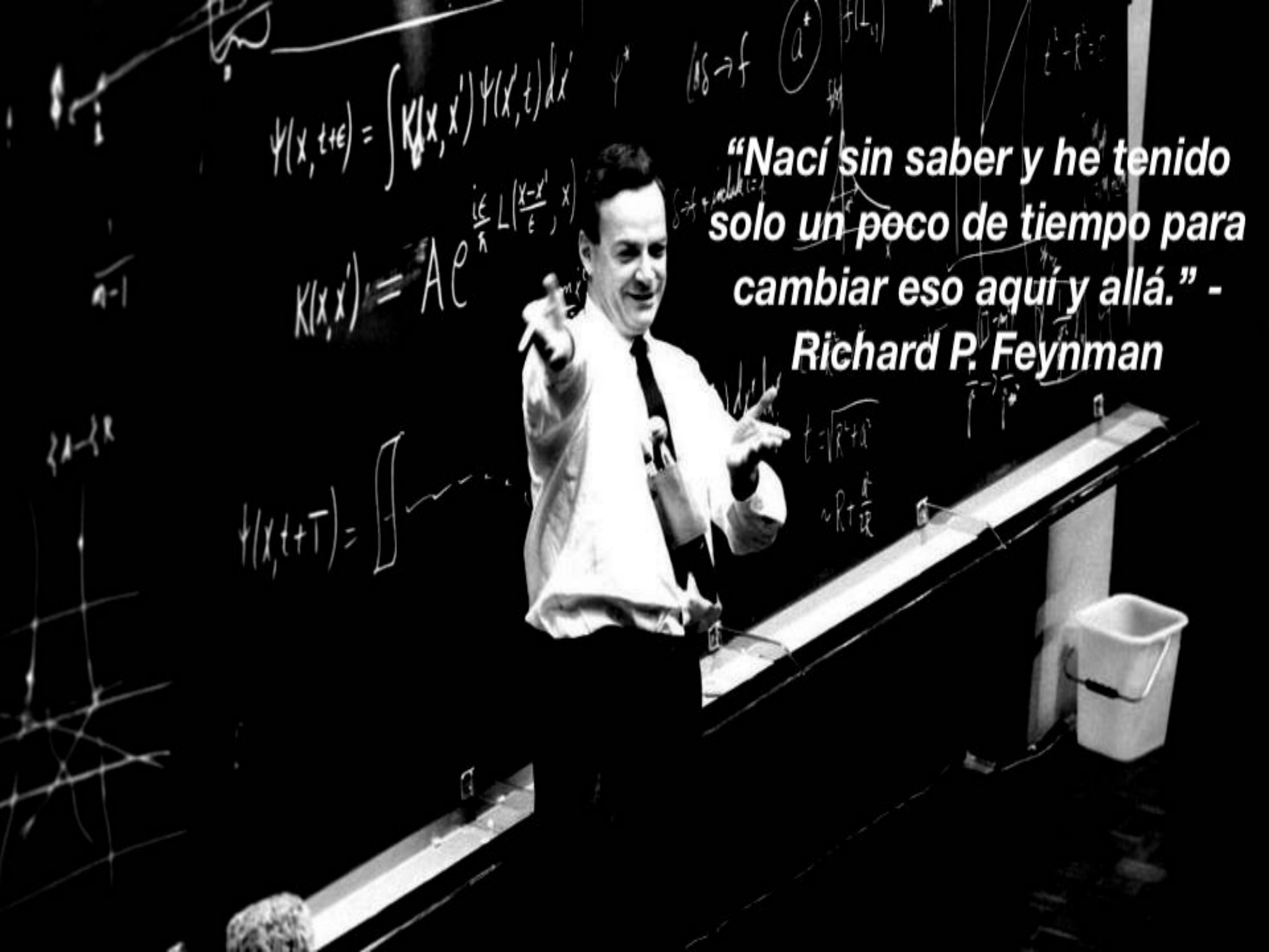

"Nací sin saber y he tenido solo un poco de tiempo para cambiar eso aquí y allá." - Richard P. Feynman

# Estrategias de *Secure Learning* para detección de Android Malware

**Jhoan Steven Delgado Villarreal**

Tutores: Christian Urcuqui, Msc.[1], Javier Díaz, Ph.D.[2], Andrés Navarro, Ph.D.[3]

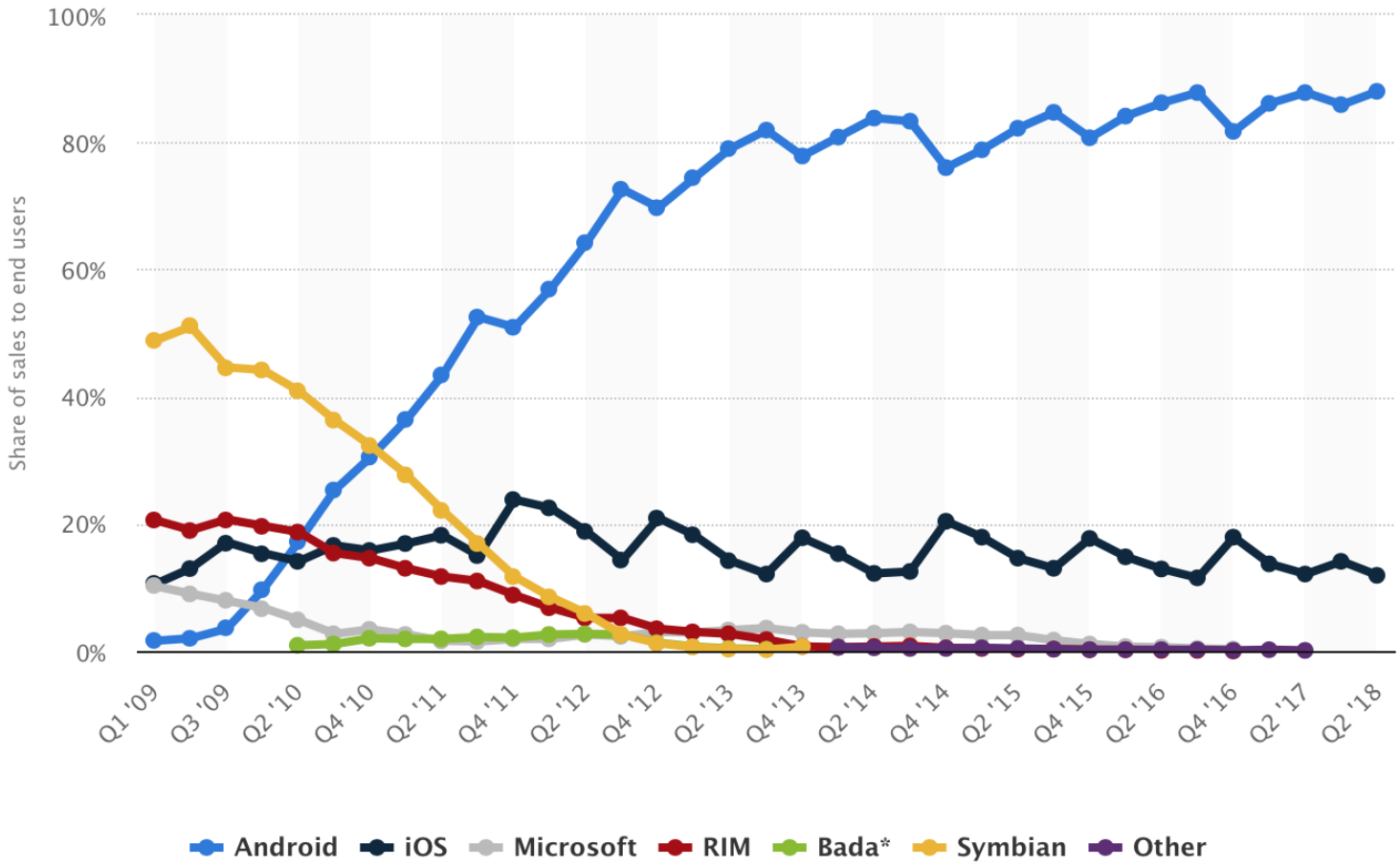
Universidad Icesi

Ingeniería de sistemas, facultad de Ingeniería

29 de noviembre de 2018

# Pregunta...

# Cuota de mercado en Android




*Global mobile OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2018 [Statista]*

# Android



- Sistema operativo para dispositivos móviles
- Código abierto (Open Source) basado en el kérnel de Linux
- Arquitectura de 5 componentes

# Android



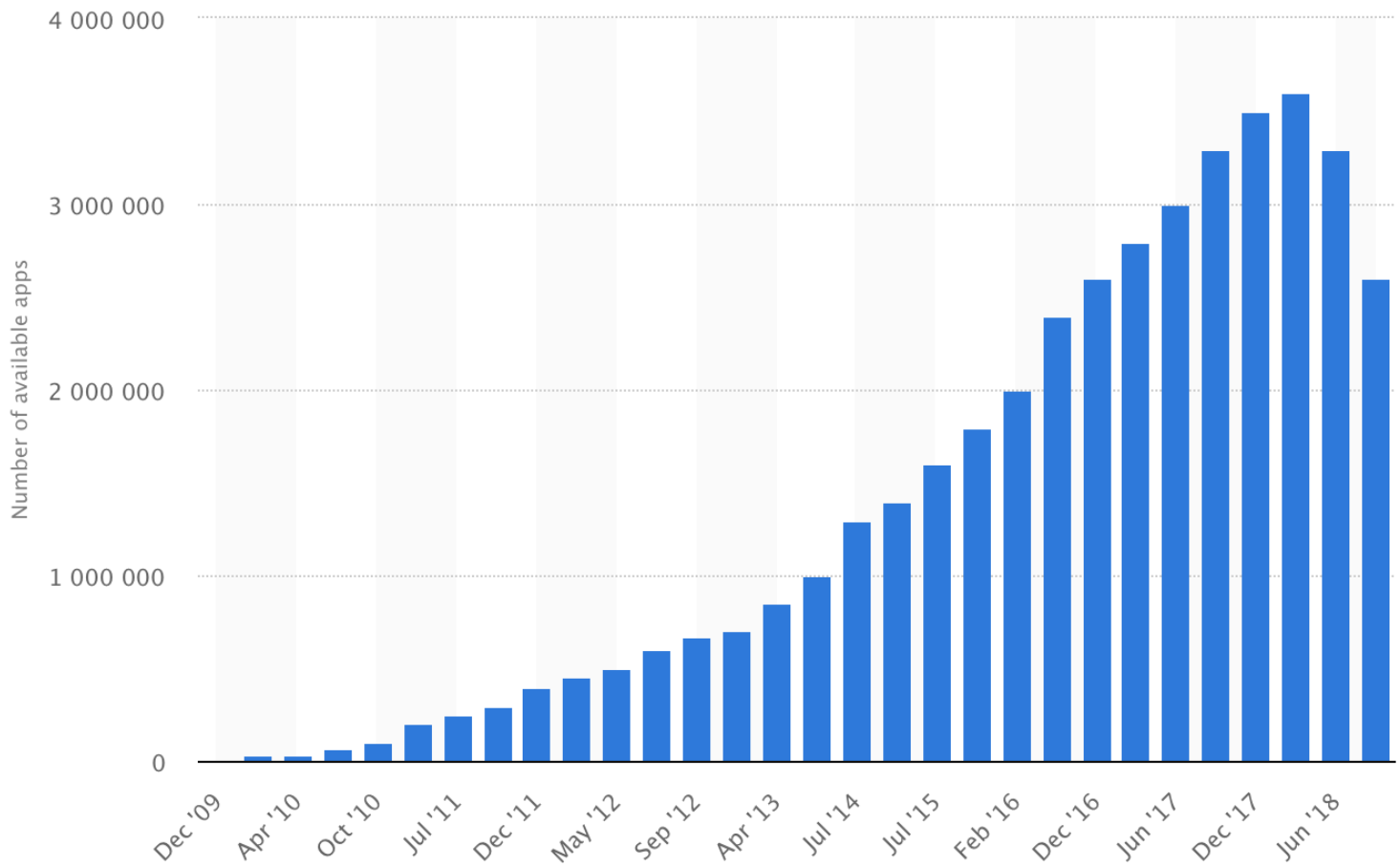
- Sistema operativo para dispositivos móviles
- Código abierto (Open Source) basado en el kérnel de Linux
- Arquitectura de 5 componentes

# Android

- Sistema operativo para dispositivos móviles
- Código abierto (Open Source) basado en el kérnel de Linux
- Arquitectura de 5 componentes

# Apps






*Number of available applications in the Google Play Store from December 2009 to September 2018 [Statista]*

# Malware




FayerWayer.
Uber fue hackeado, datos de 57 millones de usuarios comprometidos
La compañía pagó USD $100.000 por ocultar el hackeo.
Major DDoS attack on Dyn DNS knocks Spotify, Twitter, Github, PayPal, and more offline
The sound of silence.
Critical Flaws in Intel Processors Leave Millions of PCs Vulnerable
intel
In past few months, several research groups have uncovered vulnerabilities in the Intel remote administration feature known as the Management Engine (ME) which could allow remote attackers to gain full control of a targeted computer.
Now, Intel has admitted that these security vulnerabilities could "potentially place impacted platforms at risk."
WannaCry
MORE LIKE THIS
Friday's DDoS attack came from 100,000 infected devices
STUXnet
w Brokers, Who Leaked WannaCry SMB Exploit, Are With More 0-Days
Warning! Over 900 Million Android Pho
New 'QuadRooter' Attack
QuadRooter
New Flaws Affect Over 900 Million Devices
The Hacker News
Security in a serious way
Wanna Cry Again? NSA's Windows 'EsteemAudit' RDP Exploit Remains Unpatched
EsteemAudit (No Patch)
Windows RDP Hacking Tool
w Brokers
th more 0-Day Exploits

# Malware



- Software malicioso que busca perjudicar a los usuarios.
- Obtiene información sensible de los usuarios.
- Los *hackers* lo desarrollan principalmente con ánimo de lucro o activismo político.
- Primer troyano para Android en 2010.

# Malware



- Software malicioso que busca perjudicar a los usuarios.
- Obtiene información sensible de los usuarios.
- Los *hackers* lo desarrollan principalmente con ánimo de lucro o activismo político.
- Primer troyano para Android en 2010.

# Malware

- Software malicioso que busca perjudicar a los usuarios.
- Obtiene información sensible de los usuarios.
- Los *hackers* lo desarrollan principalmente con ánimo de lucro o activismo político.
- Primer troyano para Android en 2010.

# Malware

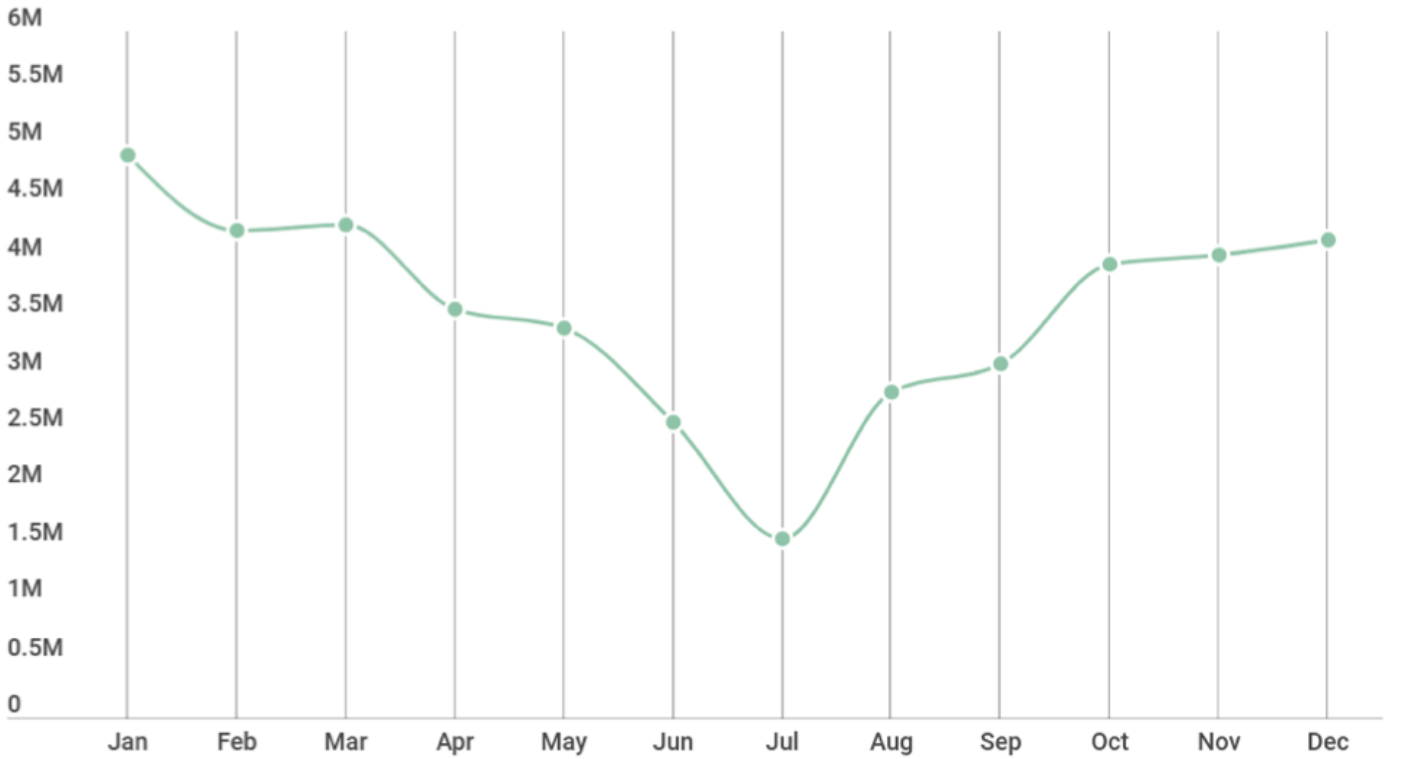


- Software malicioso que busca perjudicar a los usuarios.
- Obtiene información sensible de los usuarios.
- Los *hackers* lo desarrollan principalmente con ánimo de lucro o activismo político.
- Primer troyano para Android en 2010.

# Ataques a dispositivos móviles año 2017



Registrando un número creciente de ataques malware a móviles – 42.7 millones vs. 40 millones en 2016." [SecureList]




KASPERSKY

# Features to Detect Android Malware

Christian Camilo Urcuqui López
*Grupo de investigación i2t*
*Universidad Icesi*
Cali, Colombia
ccurcuqui@icesi.edu.co

Jhoan Steven Delgado Villarreal
*Universidad Icesi*
Cali, Colombia
jhoan.delgado@correo.icesi.edu.co

Andres Felipe Perez Belalcazar
*Universidad Icesi*
Cali, Colombia
andres.perez2@correo.icesi.edu.co

Andres Navarro Cadavid
*Grupo de investigación i2t*
*Unverisidad Icesi*
Cali, Colombia
anavarro@icesi.edu.co

Javier Gustavo Diaz Cely
*Grupo de investigación i2t*
*Universidad Icesi*
Cali, Colombia
jgdiaz@icesi.edu.co

# Características de capa de red



- (R1): Paquetes TCP
- (R2): Paquetes distintos TCP
- (R3): IP externas
- (R4): Volumen de bytes
- (R5) Paquetes UDP
- (R6) Paquetes de la aplicación fuente
- (R7) Paquetes de la aplicación remota
- (R8) Bytes de la aplicación origen
- (R9) Bytes de la aplicación remota
- (R10) Consultas DNS, número de consultas DNS.

Urcuqui, C., Navarro, A., Osorio, J., & Garcıa, M. (2017). Machine Learning Classifiers to Detect Malicious Websites. CEUR Workshop Proceedings. Vol 1950, 14-17

# Conclusiones del paper...



La esencia de Machine Learning:

- Debe existir un patrón.
- No se puede describir con exactiud matemáticamente.
- **Tenemos datos.**

*The Learning Problem - Introduction. Professor Yaser Abu-Mostafa, Caltech.*

Entonces...

- Sí es posible entrenar clasificadores de machine learning para la detección de software malicioso en Andorid con tráfico de red.
- Existen algunas tendencias en los flujos de información. EJ: Paquetes TCP: 197 (Apps benignas), y 72 (Apps malignas)

# Conclusiones del paper...



La esencia de Machine Learning:

- Debe existir un patrón.
- No se puede describir con exactiud matemáticamente.
- **Tenemos datos.**

*The Learning Problem - Introduction. Professor Yaser Abu-Mostafa, Caltech.*

Entonces...

- Sí es posible entrenar clasificadores de machine learning para la detección de software malicioso en Andorid con tráfico de red.
- Existen algunas tendencias en los flujos de información. EJ: Paquetes TCP: 197 (Apps benignas), y 72 (Apps malignas)

# Conclusiones del paper...



La esencia de Machine Learning:

- Debe existir un patrón.
- No se puede describir con exactiud matemáticamente.
- **Tenemos datos.**

*The Learning Problem - Introduction. Professor Yaser Abu-Mostafa, Caltech.*

Entonces...

- Sí es posible entrenar clasificadores de machine learning para la detección de software malicioso en Andorid con tráfico de red.
- Existen algunas tendencias en los flujos de información. EJ: Paquetes TCP: 197 (Apps benignas), y 72 (Apps malignas)

# Conclusiones del paper...



La esencia de Machine Learning:

- Debe existir un patrón.
- No se puede describir con exactiud matemáticamente.
- **Tenemos datos.**

*The Learning Problem - Introduction. Professor Yaser Abu-Mostafa, Caltech.*

Entonces...

- Sí es posible entrenar clasificadores de machine learning para la detección de software malicioso en Andorid con tráfico de red.
- Existen algunas tendencias en los flujos de información. EJ: Paquetes TCP: 197 (Apps benignas), y 72 (Apps malignas)

# Conclusiones del paper...



La esencia de Machine Learning:

- Debe existir un patrón.
- No se puede describir con exactiud matemáticamente.
- **Tenemos datos.**

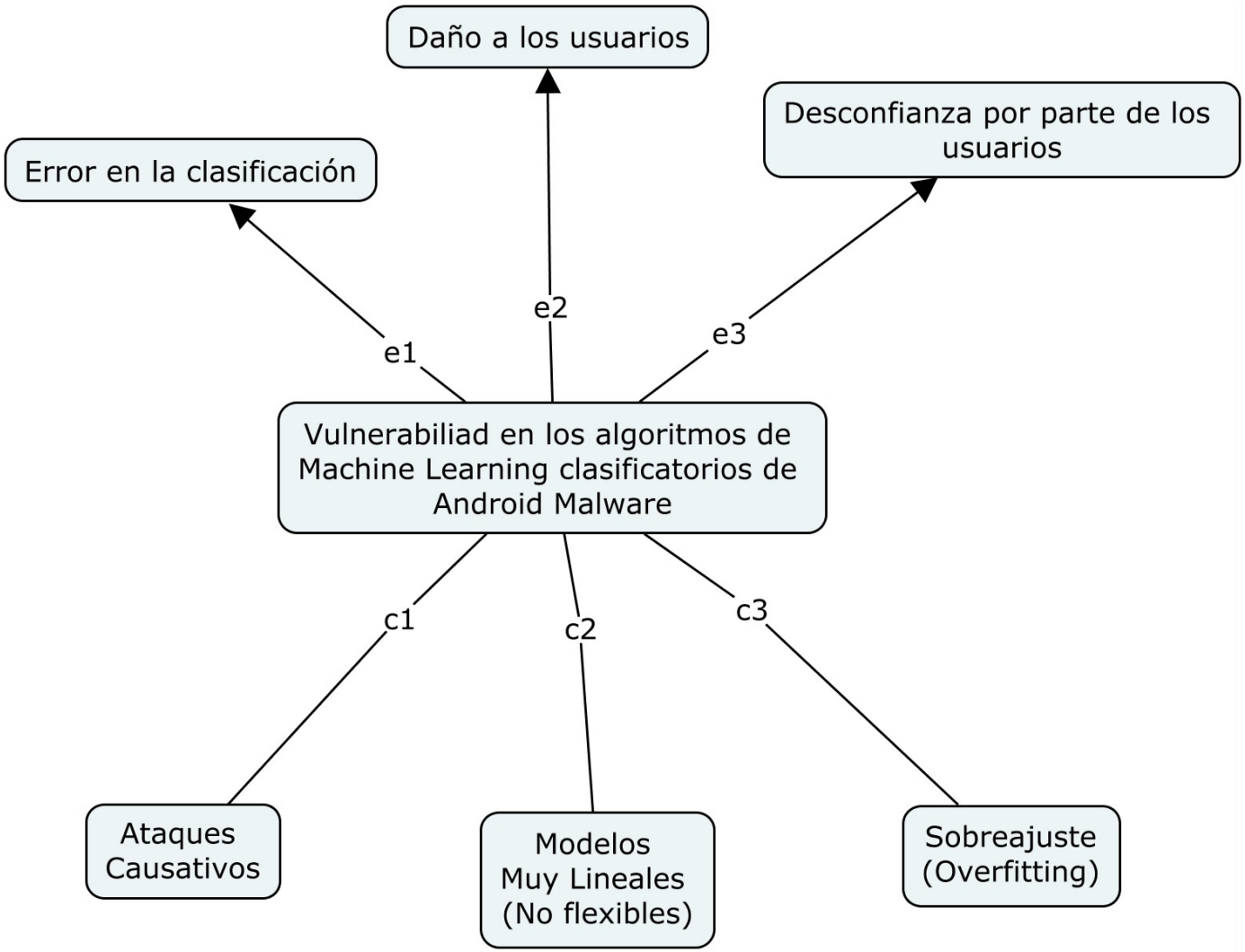
*The Learning Problem - Introduction. Professor Yaser Abu-Mostafa, Caltech.*

Entonces...

- Sí es posible entrenar clasificadores de machine learning para la detección de software malicioso en Andorid con tráfico de red.
- Existen algunas tendencias en los flujos de información. EJ: Paquetes TCP: 197 (Apps benignas), y 72 (Apps malignas)

# El problema




Daño a los usuarios
Error en la clasificación
Desconfianza por parte de los usuarios
e1
e2
e3
Vulnerabiliad en los algoritmos de Machine Learning clasificatorios de Android Malware
c1
c2
c3
Ataques Causativos
Modelos Muy Lineales (No flexibles)
Sobreajuste (Overfitting)

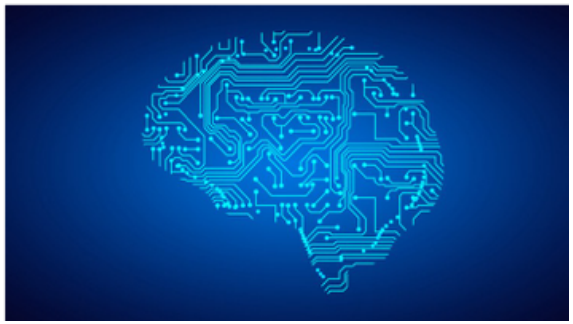
# Marco teórico



- Inteligencia Artificial (IA)
- Ciberseguridad

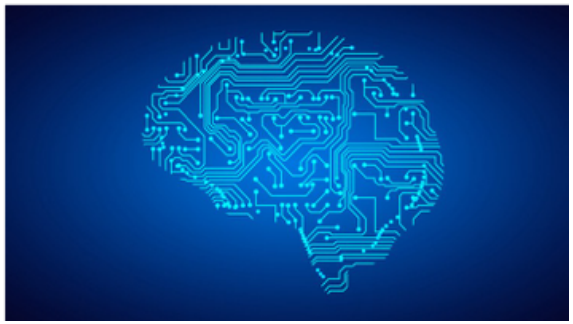
# Machine Learning

- (Samuel,1959) Se refiere al termino de Machine Learning como el campo de estudio que le brinda a los computadores la habilidad de aprender sin necesidad de estar explícitamente programados.
- Se realiza el aprendizaje a través de los datos.



# Machine Learning

- (Samuel,1959) Se refiere al termino de Machine Learning como el campo de estudio que le brinda a los computadores la habilidad de aprender sin necesidad de estar explícitamente programados.
- Se realiza el aprendizaje a través de los datos.



# Aprendizaje supervisado



- Aprender a partir de un "experto"
- Datos de entrenamiento etiquetados con una clase o valor:

$$(x_1, x_2, ..., x_n, y) \tag{1}$$

donde $y$ es la etiqueta.

- Meta: predecir una clase o valor.

*Análitica de datos, Prof. Javier Díaz, 2016*

# Aprendizaje supervisado

- Aprender a partir de un "experto"
- Datos de entrenamiento etiquetados con una clase o valor:

$$(x_1, x_2, ..., x_n, y) \tag{1}$$

donde $y$ es la etiqueta.

- Meta: predecir una clase o valor.

*Análitica de datos, Prof. Javier Díaz, 2016*

# Aprendizaje supervisado

- Aprender a partir de un "experto"
- Datos de entrenamiento etiquetados con una clase o valor:

$$(x_1, x_2, ..., x_n, y) \tag{1}$$

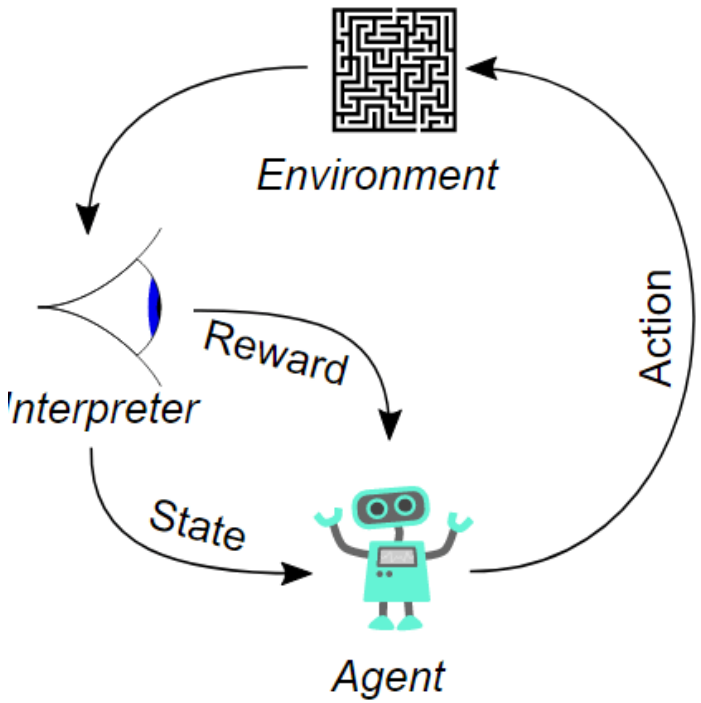
  donde $y$ es la etiqueta.
- Meta: predecir una clase o valor.

*Análitica de datos, Prof. Javier Díaz, 2016*

# Aprendizaje por refuerzo



- El aprendizaje por refuerzo es el problema de lograr que un agente actúe en el mundo para maximizar sus recompensas.
- Por ejemplo, considere enseñarle a un perro un nuevo truco: no puede decirle qué hacer, pero puede recompensarlo/castigarlo si hace lo correcto/incorrecto

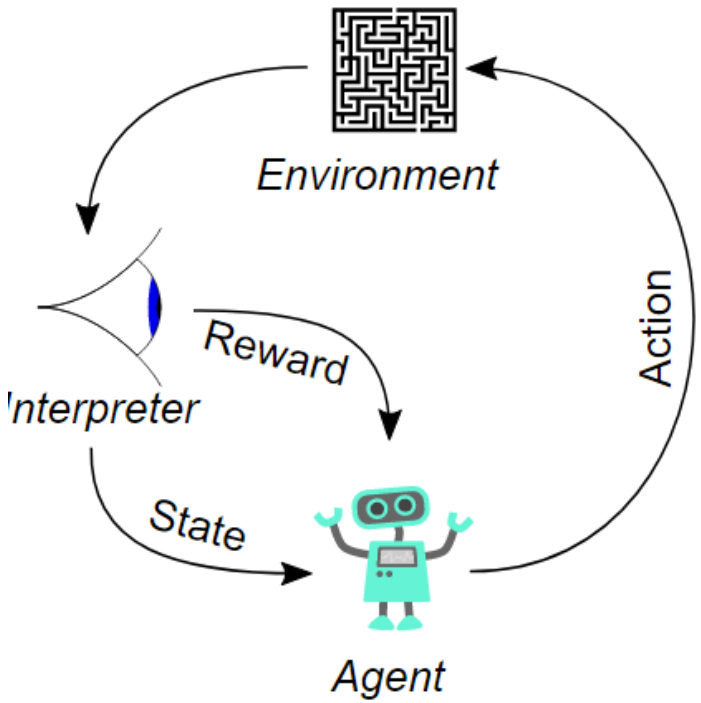



https://www.cs.ubc.ca/ murphyk/Bayes/pomdp.html

# Aprendizaje por refuerzo



- El aprendizaje por refuerzo es el problema de lograr que un agente actúe en el mundo para maximizar sus recompensas.
- Por ejemplo, considere enseñarle a un perro un nuevo truco: no puede decirle qué hacer, pero puede recompensarlo/castigarlo si hace lo correcto/incorrecto

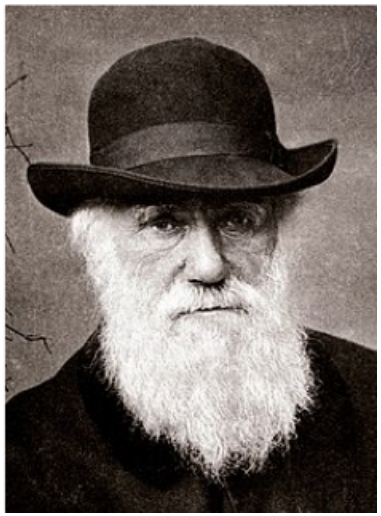



*https://www.cs.ubc.ca/ murphyk/Bayes/pomdp.html*

# Algoritmos genéticos



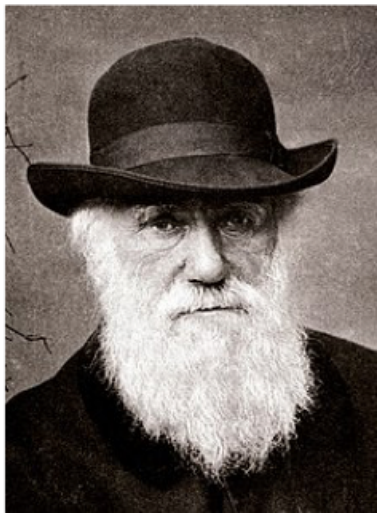
- (Goldberg, 1989) Define algoritmo genético como algoritmos de búsqueda basados en la selección natural y la genética (Charles Darwin).
- En cada generación, un conjunto de individuos (cadenas) son creados usando bits y partes de los antigüos más ajustados.
- El pionero de estos algoritmos fue el Profesor John Holland



# Algoritmos genéticos



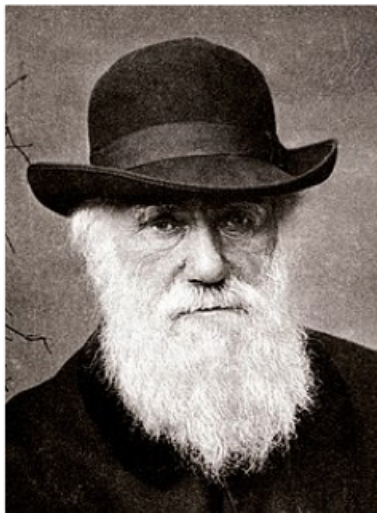
- (Goldberg, 1989) Define algoritmo genético como algoritmos de búsqueda basados en la selección natural y la genética (Charles Darwin).
- En cada generación, un conjunto de individuos (cadenas) son creados usando bits y partes de los antigüos más ajustados.
- El pionero de estos algoritmos fue el Profesor John Holland



# Algoritmos genéticos

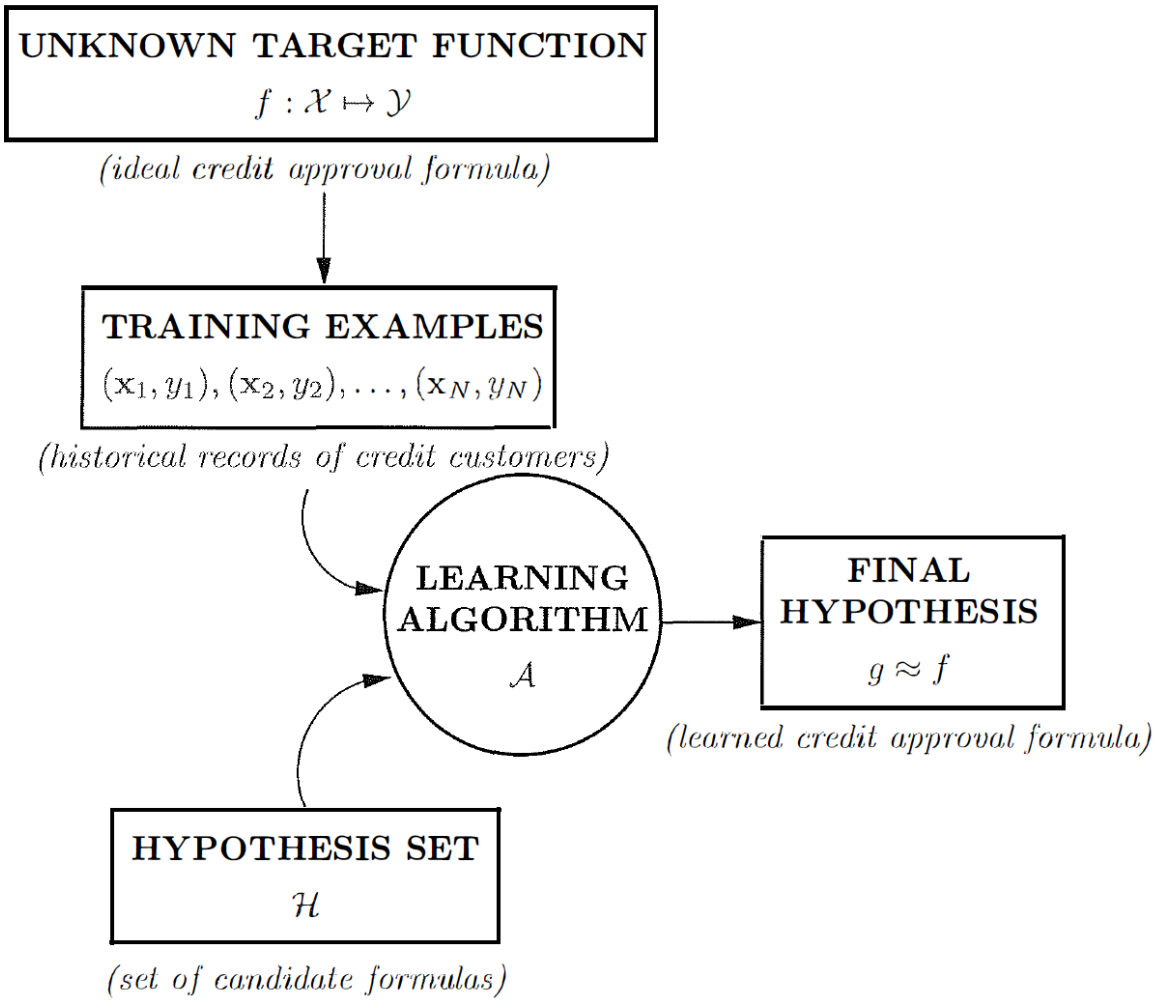


- (Goldberg, 1989) Define algoritmo genético como algoritmos de búsqueda basados en la selección natural y la genética (Charles Darwin).
- En cada generación, un conjunto de individuos (cadenas) son creados usando bits y partes de los antigüos más ajustados.
- El pionero de estos algoritmos fue el Profesor John Holland



# El problema de aprendizaje






*The Learning Problem - Introduction. , Caltech.*

# Técnicas para el análisis de amenazas



- Análsis estático
- Análisis dinámico

# Técnicas para el análisis de amenazas

- Análsis estático
- Análisis dinámico

# Análisis estático



- Técnica que evalúa los comportamientos maliciosos del codigo fuente, datos, o archivos binarios, sin ejecutar directamente la App
- Es posible evitarlo a partir de técnicas de ofuscación

*Batyuk, L., Herpich, M., Camtepe, S. A., Raddatz, K., Schmidt, A., & Albayrak, S. Using static analysis for automatic assessment and mitigation of unwanted and malicious activities within Android applications. Malicious and Unwanted Software (MALWARE), 2011 6th International Conference on. IEEE, Piscataway. 2011.*

# Análisis estático



- Técnica que evalúa los comportamientos maliciosos del codigo fuente, datos, o archivos binarios, sin ejecutar directamente la App
- Es posible evitarlo a partir de técnicas de ofuscación

*Batyuk, L., Herpich, M., Camtepe, S. A., Raddatz, K., Schmidt, A., & Albayrak, S. Using static analysis for automatic assessment and mitigation of unwanted and malicious activities within Android applications. Malicious and Unwanted Software (MALWARE), 2011 6th International Conference on. IEEE, Piscataway. 2011.*

# Análisis dinámico

- Estudia el comportamiento del malware en ejecución mediante simulación de gestos.
- Se analizan los procesos en ejecución, la interfaz de usuario, conexiones de red, entre otros.
- Existen técnicas que permiten evadirlo. El malware tiene la capacidad de detectar ambientes sandbox y detener su comportamiento malicioso

*Petsas, T., Voyatzis, G., Athanasopoulos, E., Polychronakis, M., & Ioannidis, S. Rage against the virtual machine: hindering dynamic analysis of android malware. In Proceedings of the Seventh European Workshop on System Security (p. 5). ACM. April 2014.*

# Análisis dinámico



- Estudia el comportamiento del malware en ejecución mediante simulación de gestos.
- Se analizan los procesos en ejecución, la interfaz de usuario, conexiones de red, entre otros.
- Existen técnicas que permiten evadirlo. El malware tiene la capacidad de detectar ambientes sandbox y detener su comportamiento malicioso

*Petsas, T., Voyatzis, G., Athanasopoulos, E., Polychronakis, M., & Ioannidis, S. Rage against the virtual machine: hindering dynamic analysis of android malware. In Proceedings of the Seventh European Workshop on System Security (p. 5). ACM. April 2014.*

# Análisis dinámico

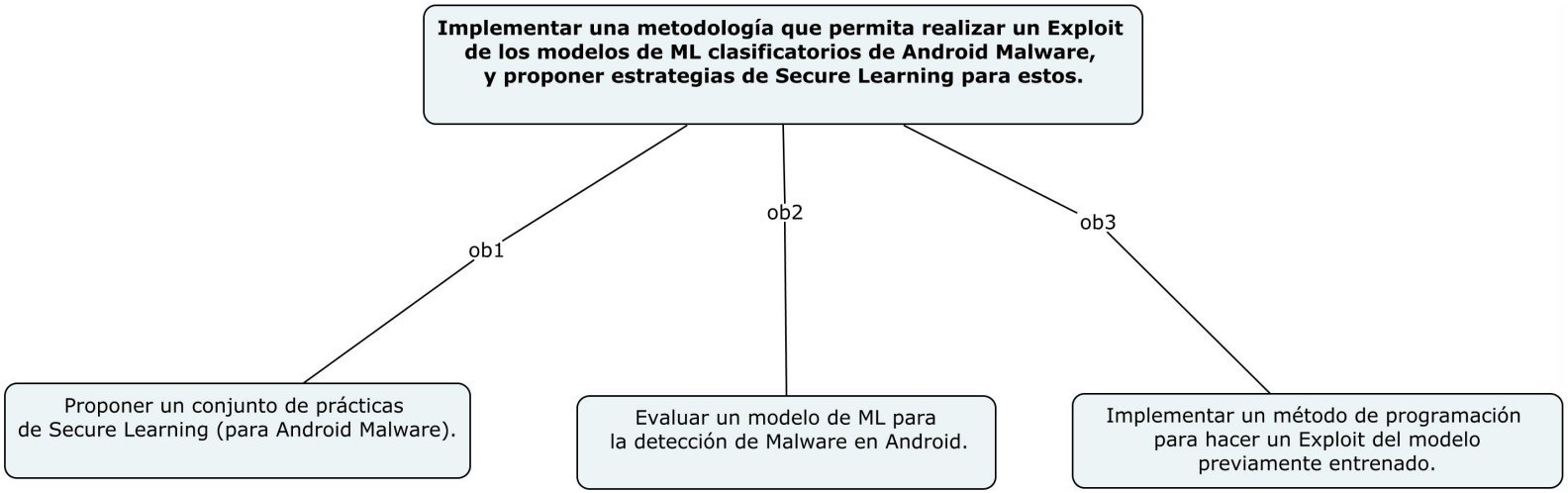


- Estudia el comportamiento del malware en ejecución mediante simulación de gestos.
- Se analizan los procesos en ejecución, la interfaz de usuario, conexiones de red, entre otros.
- Existen técnicas que permiten evadirlo. El malware tiene la capacidad de detectar ambientes sandbox y detener su comportamiento malicioso

*Petsas, T., Voyatzis, G., Athanasopoulos, E., Polychronakis, M., & Ioannidis, S. Rage against the virtual machine: hindering dynamic analysis of android malware. In Proceedings of the Seventh European Workshop on System Security (p. 5). ACM. April 2014.*

# Objetivos




Implementar una metodología que permita realizar un Exploit de los modelos de ML clasificatorios de Android Malware, y proponer estrategias de Secure Learning para estos.
ob1
ob2
ob3
Proponer un conjunto de prácticas de Secure Learning (para Android Malware).
Evaluar un modelo de ML para la detección de Malware en Android.
Implementar un método de programación para hacer un Exploit del modelo previamente entrenado.

# Proyectos



| Papers | APK perturbations | Network features | Attack framework | Estrategias de aprendizaje seguro |
|---|---|---|---|---|
| Android HIV: A Study of Repackaging Malware for Evading Machine-Learning Detection | SI | NO | SI | NO |
| Poster: Towards Adversarial Detection of Mobile Malware | NO | NO | NO | SI |
| Yes, Machine Learning Can Be More Secure! A Case Study on Android Malware Detection | NO | NO | SI | NO |
| Estrategias de Secure Learning para detección de Android Malware. | NO | SI | NO | SI |

"Minería de datos no es algo que haces solo una vez y luego olvidas, es un proceso continuo"

*D*ata Mining For Dummies®, John Wiley & Sons, Inc.

# CRISP-DM



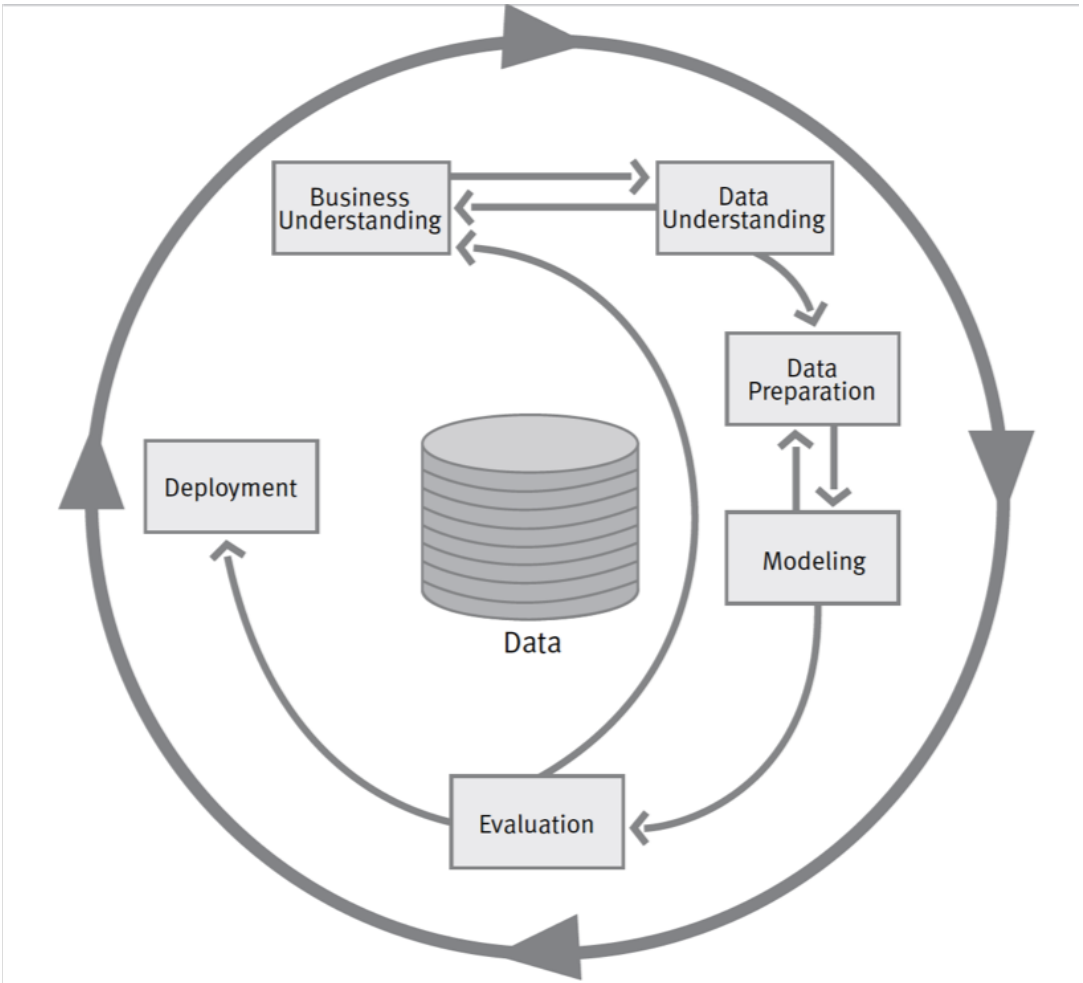

Business Understanding
Data Understanding
Data Preparation
Modeling
Evaluation
Deployment
Data

DIAGRAMA DE GANTT

| | Modo de | Nombre de tarea | Duración |
|---|---|---|---|
| 1 | | **Proyecto Malware** | **38 días?** |
| 2 | | **Estrategia de aprendizaje seguro** | **22 días?** |
| 3 | | Identificar prácticas de aprendizaje seguro | 15 días |
| 4 | | Evaluar prácticas de aprendizaje seguro | 7 días |
| 5 | | Escribir documento final del Anteproyecto | 1 día? |
| 6 | | **Modelo de detección entrenado** | **30 días?** |
| 7 | | Implementar el sistema propuesto por Andrés para la captura de tráfico de red de aplicaciones Android | 14 días |
| 8 | | Generar un conjunto de tráfico de red (Apps Benignas y Maliciosas) | 1 día |
| 9 | | Crear un dataset para training y testing | 7 días |
| 10 | | Entrenar distintos algoritmos de ML | 1 día? |
| 11 | | Evaluar algoritmos | 7 días |
| 12 | | **Método de programación para realizar exploit** | **38 días?** |
| 13 | | Analizar distintos tipos de métodos para realizar un exploit en los algoritmos de ML | 15 días |
| 14 | | Evaluar los métodos | 15 días |
| 15 | | Seleccionar un método | 7 días |
| 16 | | Implementar método para realizar exploit del algorimo de ML | 1 día? |
| 17 | | Documento final proyecto | 1 día? |

# Referencias

# ¡Muchas gracias!
# ¿Preguntas?