EXTRACCIÓN DE CONOCIMIENTO A PARTIR DE FUENTES DE DATOS PROCEDENTES DE LA MONITORIZACIÓN DE EVENTOS DE SEGURIDAD

Andrés Caro

Cátedra VIEWNEXT AN IBM SUBSIDIARY UEX

$whoami

Andrés Caro

- Profesor Titular de Universidad – Universidad de Extremadura
- Área Lenguajes y Sistemas Informáticos
- Doctor e Ingeniero en Informática
- Director Cátedra ViewNext-UEx Seguridad y Auditoría de Sistemas Software

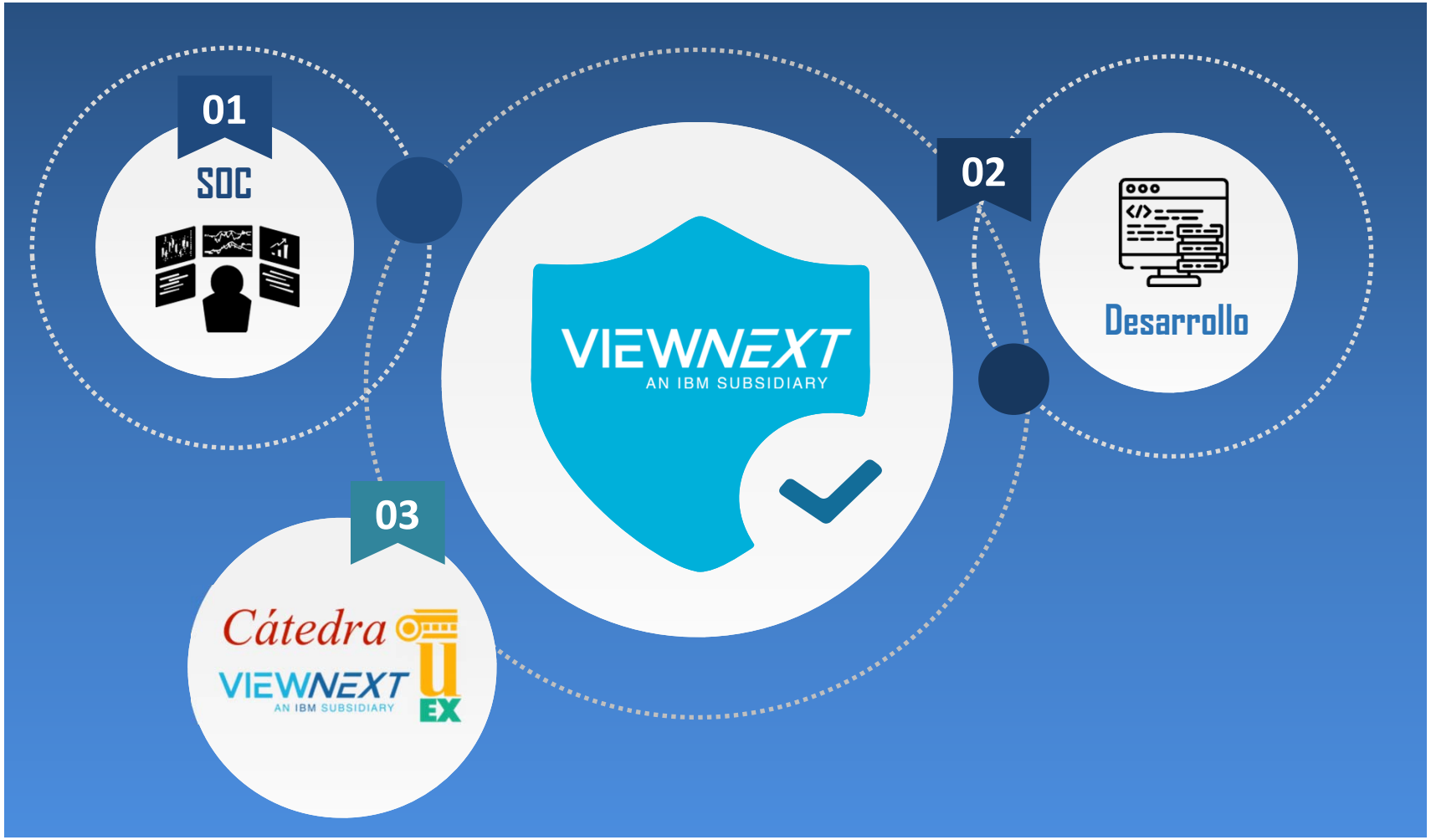
- Grados de Ingeniería en Informática
  - Ingeniería de Computadores
  - Ingeniería del Software
- Máster en Ingeniería Informática

- Organización de cursos / jornadas
  - I Foro CIBER (noviembre 2016) y II Foro CIBER (marzo 2018)
  - Curso Experto Profesional Derecho Tecnológico e Informática Forense
  - Hacking ético – Seguridad informática
  - LOPD – ENS
  - Peritaje – Informática forense

- Cibercooperante

@_AndresCaro_
andresc@unex.es

SOC (CENTRO DE OPERACIONES DE CIBERSEGURIDAD)



SIEM (*Security Information and Event Management*)

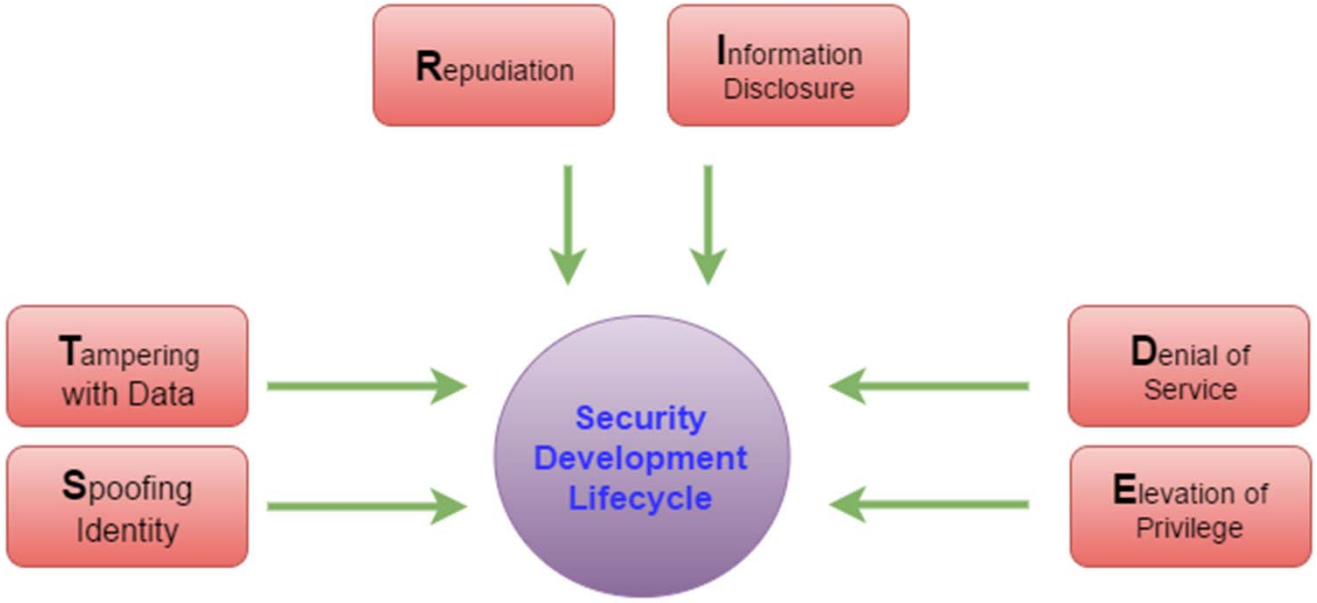
Necesidad: Gestión de Eventos e Información de Seguridad

¿Qué monitoriza un SIEM?

SIEM COMERCIALES

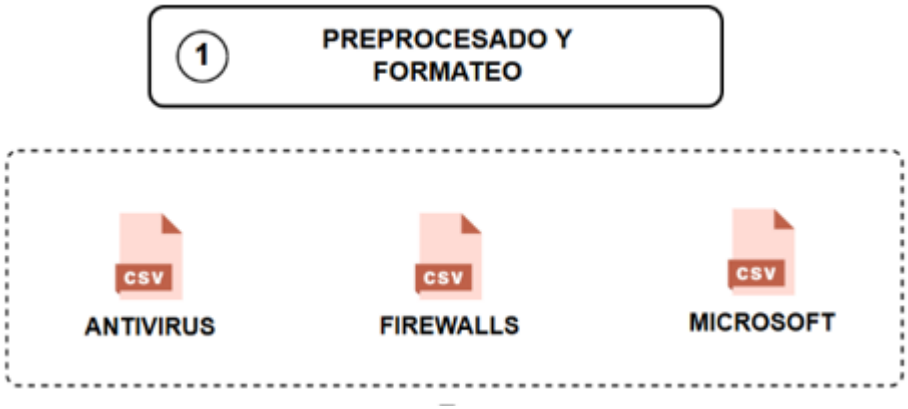
STRIDE

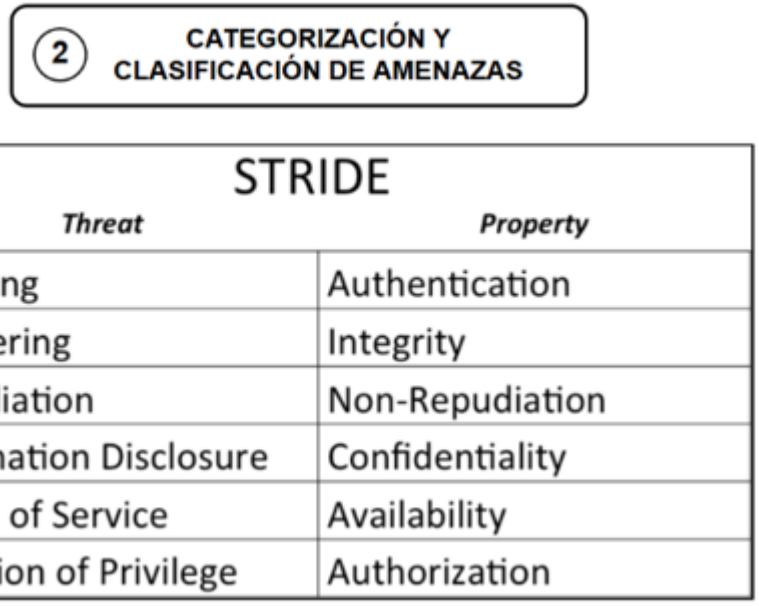
Spoofing **Suplantación de Identidad (autenticación de sistemas)**
Tampering **Manipulación de información (integridad de datos)**
Repudiation **Repudio**
Information Disclosure **Divulgación de Información**
Denial of Service **Denegación de Servicio**
Elevation of Privilege **Escalada de Privilegios**



STRIDE Threat Model

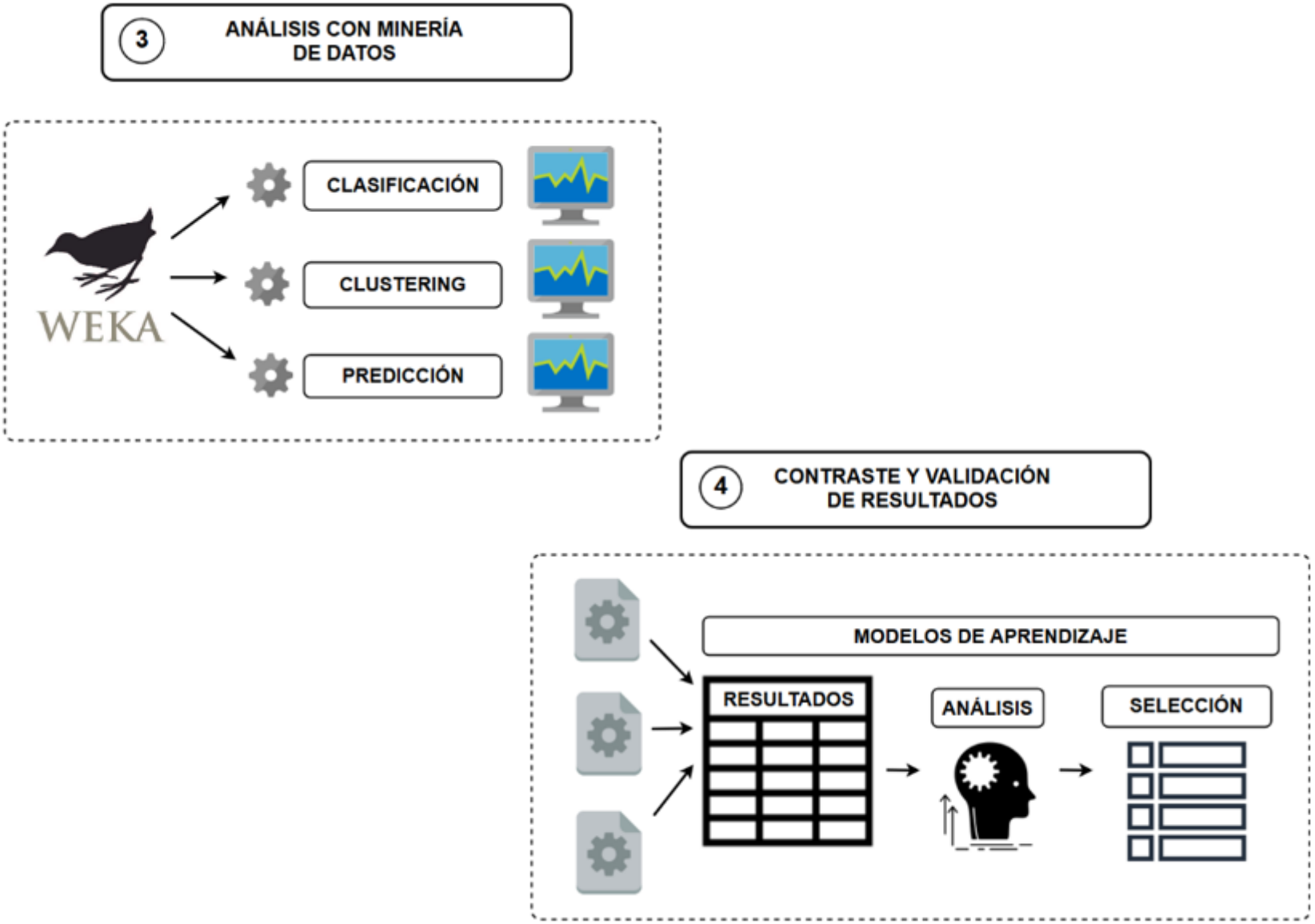
SISTEMA DE EXTRACCIÓN DE CONOCIMIENTO SIEM



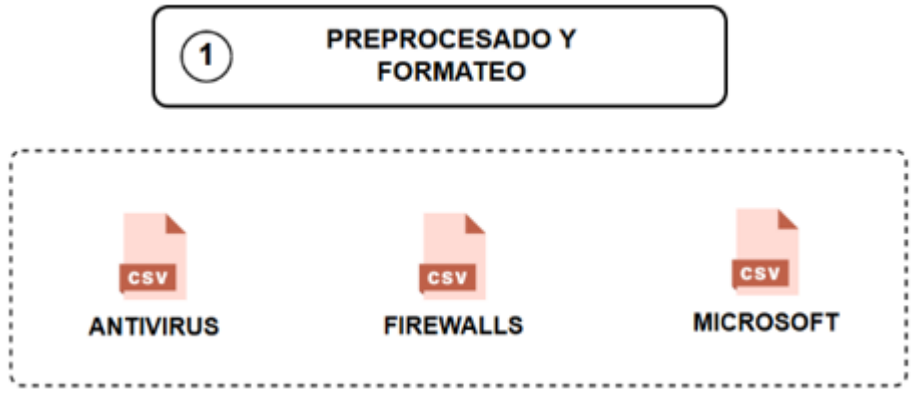


STRIDE	
Threat	*Property*
Spoofing	Authentication
Tampering	Integrity
Repudiation	Non-Repudiation
Information Disclosure	Confidentiality
Denial of Service	Availability
Elevation of Privilege	Authorization

STRIDE


3
ANÁLISIS CON MINERÍA DE DATOS
WEKA
CLASIFICACIÓN
CLUSTERING
PREDICCIÓN
4
CONTRASTE Y VALIDACIÓN DE RESULTADOS
MODELOS DE APRENDIZAJE
RESULTADOS
ANÁLISIS
SELECCIÓN

FASE 1: PREPROCESADO Y FORMATEO DE DATOS



ANONIMIZACIÓN

10.000 TUPLAS/ARCHIVO

Empresa media:

1.000 eventos/segundo

86 millones de eventos/día

Formato CSV y subestructura interna SYSLOG

FASE 2: CATEGORIZACIÓN Y CLASIFICACIÓN DE AMENAZAS

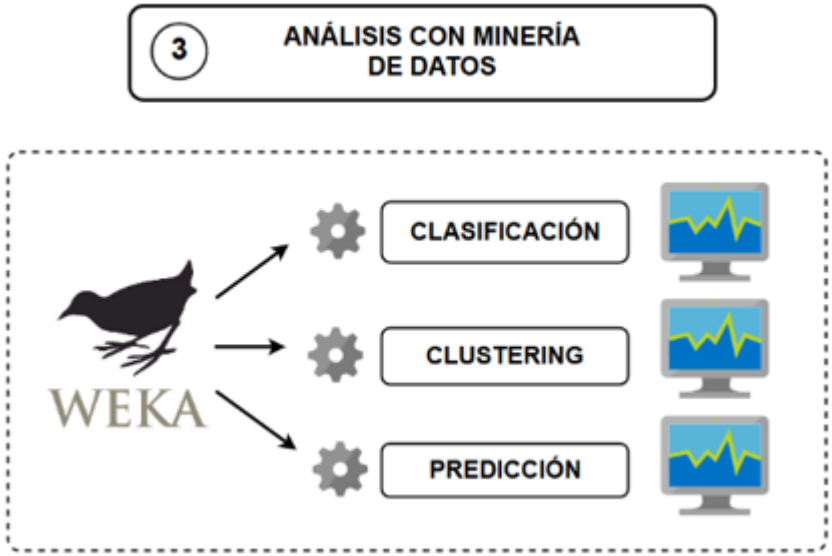
Cátedra VIEWNEXT AN IBM SUBSIDIARY U EX

② CATEGORIZACIÓN Y CLASIFICACIÓN DE AMENAZAS

STRIDE	
Threat	*Property*
Spoofing	Authentication
Tampering	Integrity
Repudiation	Non-Repudiation
Information Disclosure	Confidentiality
Denial of Service	Availability
Elevation of Privilege	Authorization

	S	T	R	I	D	E
Intentos de login fallido en PC, WiFi, correo, etc.	X	X				
Falsificar mensaje correo electrónico	X		X			
Miles de peticiones a un recurso no autorizado intentando obtener privilegios de administrador		X			X	X

FASE 3: ANÁLISIS DE DATOS



Árboles de Decisión

- J48 / C4.5
- Random Forest
- Random Tree

Modelos Bayesianos

- Naive Bayes
- Bayes Net

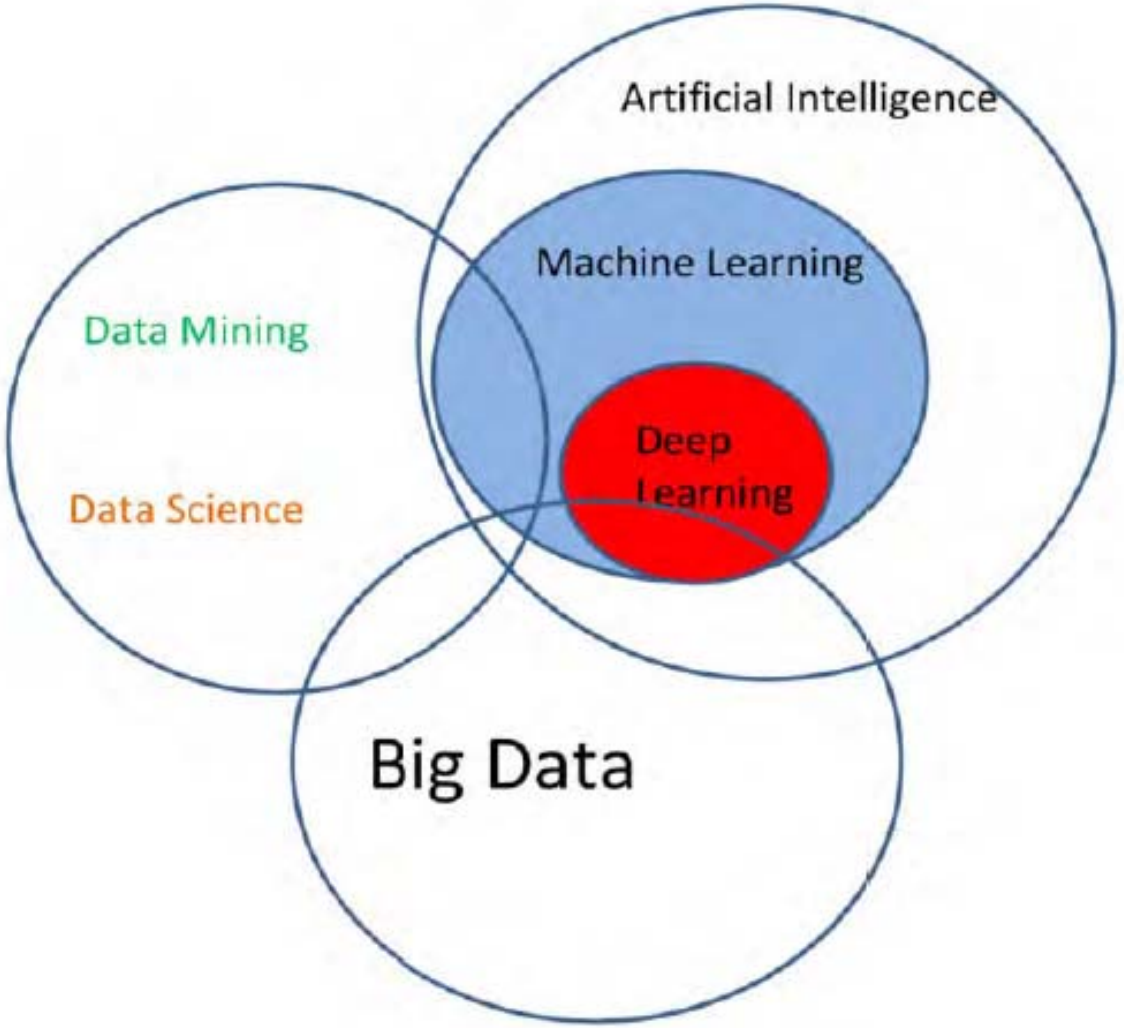
Redes Neuronales

- LibSVM

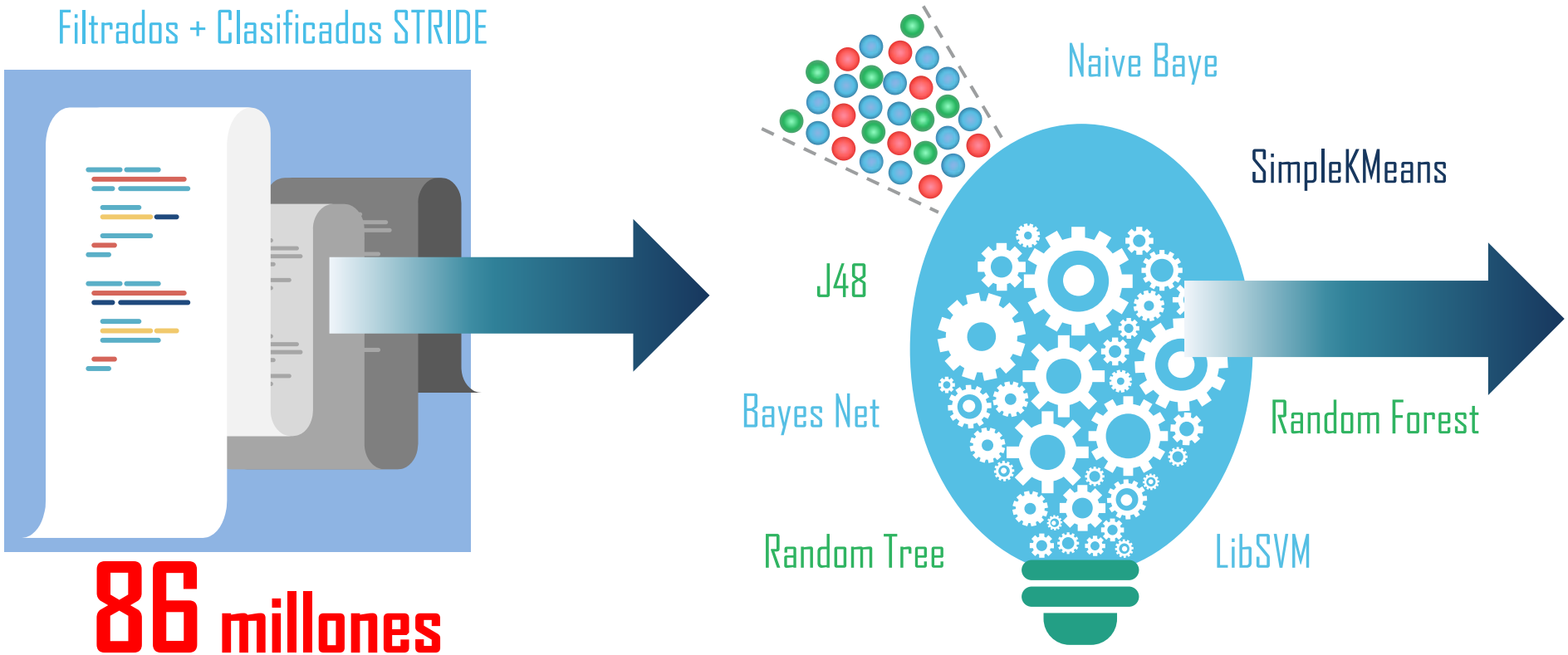
Agrupamiento (Clustering)

- SimpleKMeans

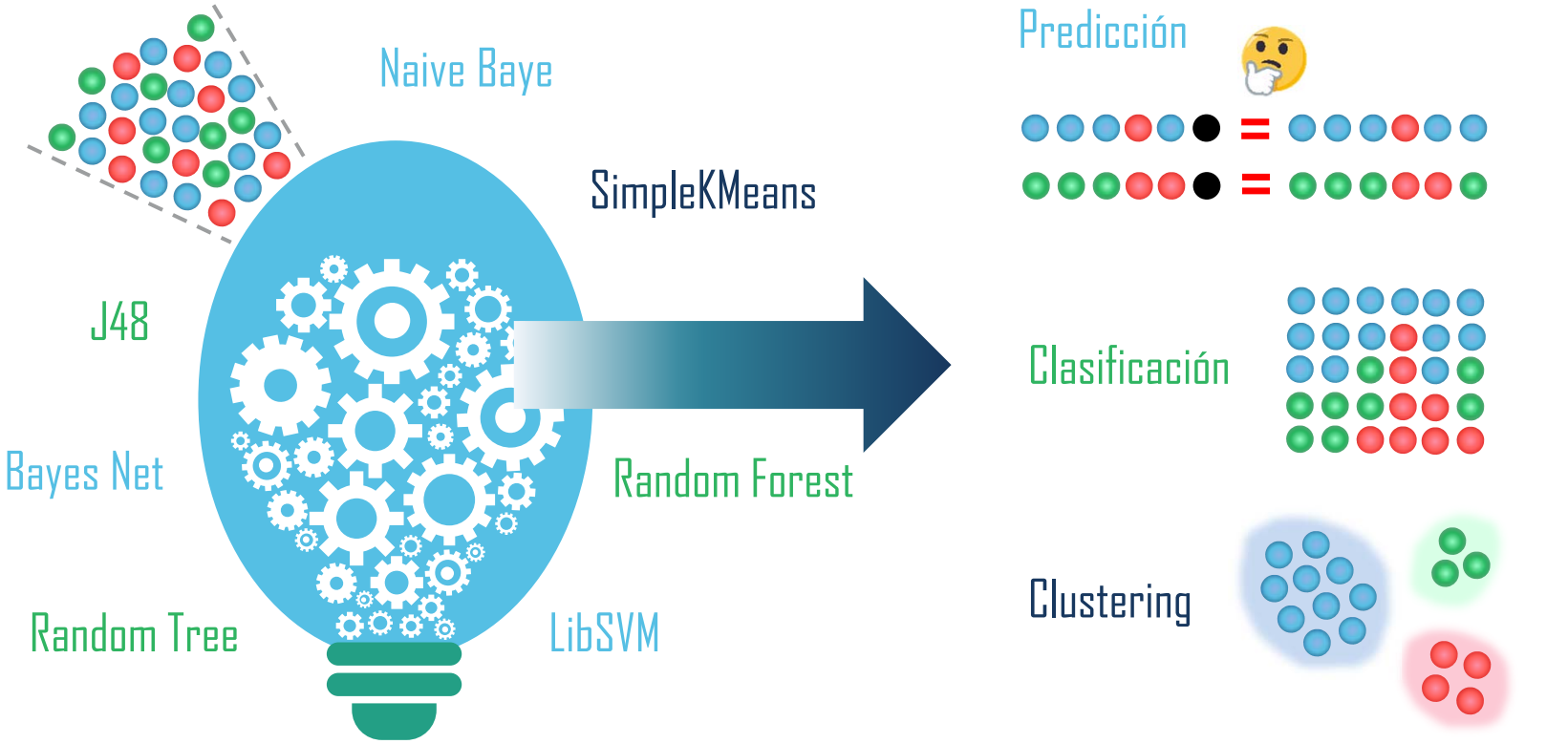
FASE 3: ANÁLISIS DE DATOS



FASE 3: ANÁLISIS DE DATOS



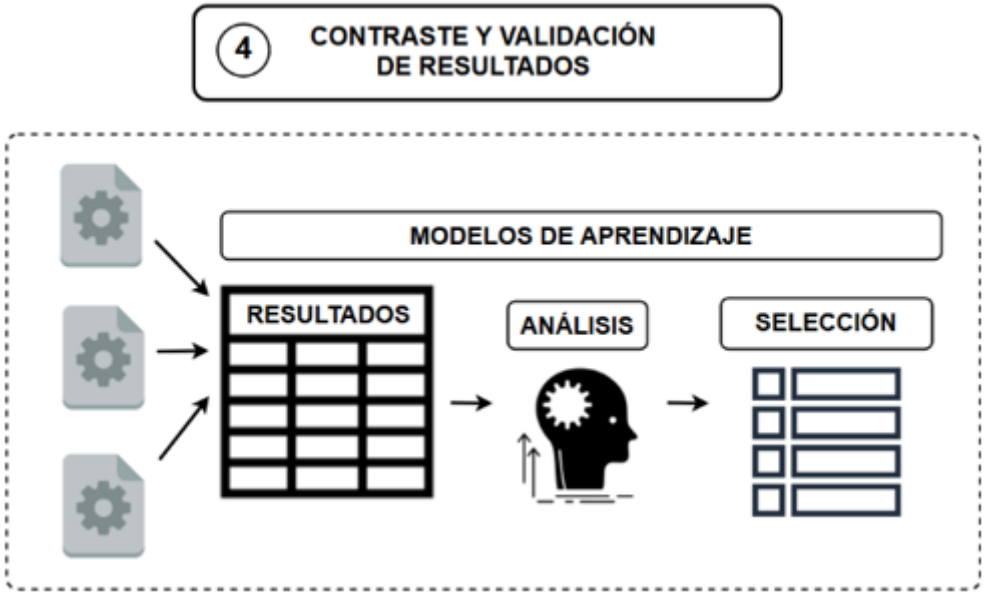
FASE 3: ANÁLISIS DE DATOS



Minería de datos

Conductas extrañas

FASE 4: VALIDACIÓN DE RESULTADOS

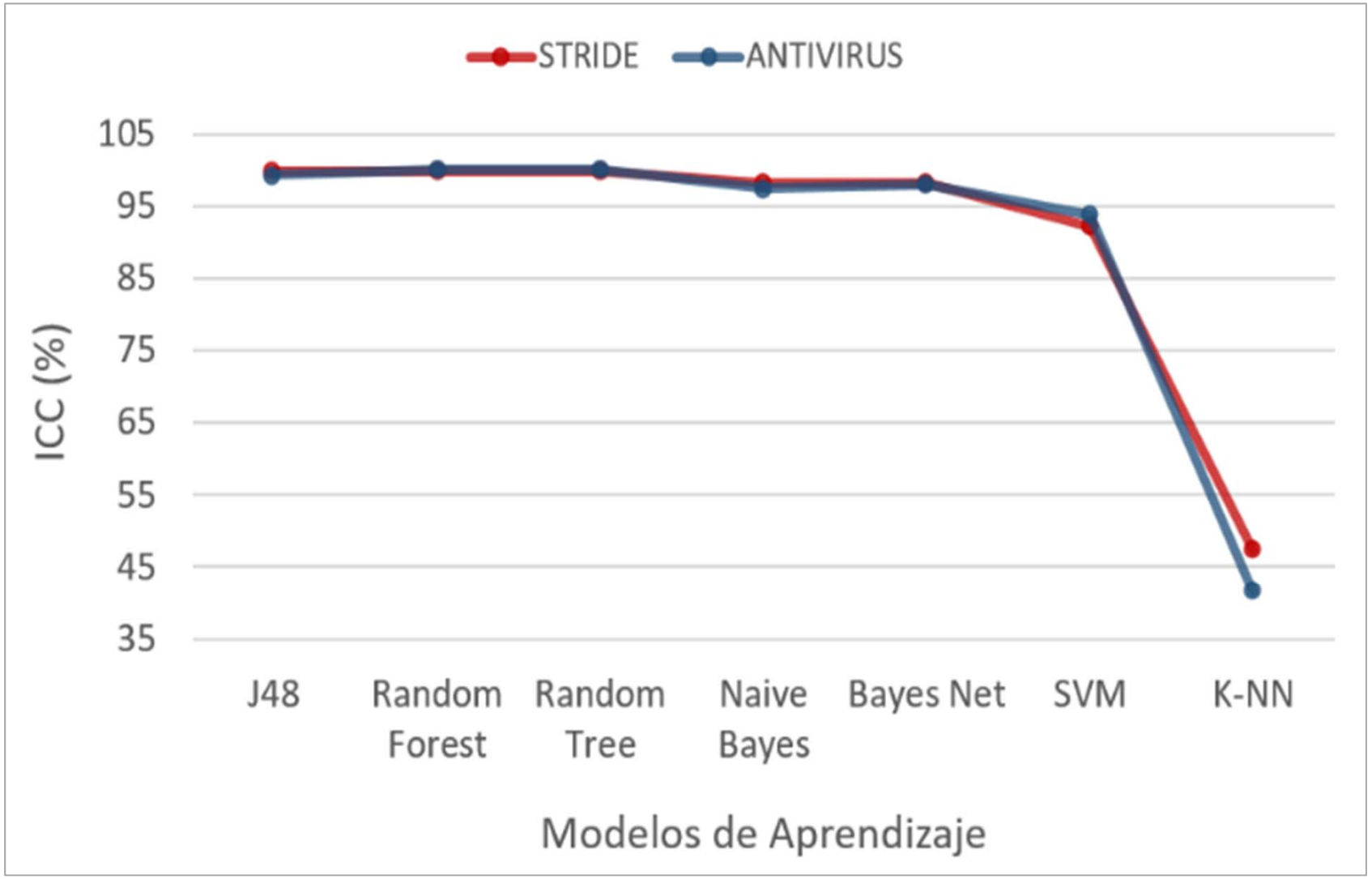


PREDECIR FUTURAS AMENAZAS

PREDECIR FALSOS POSITIVOS

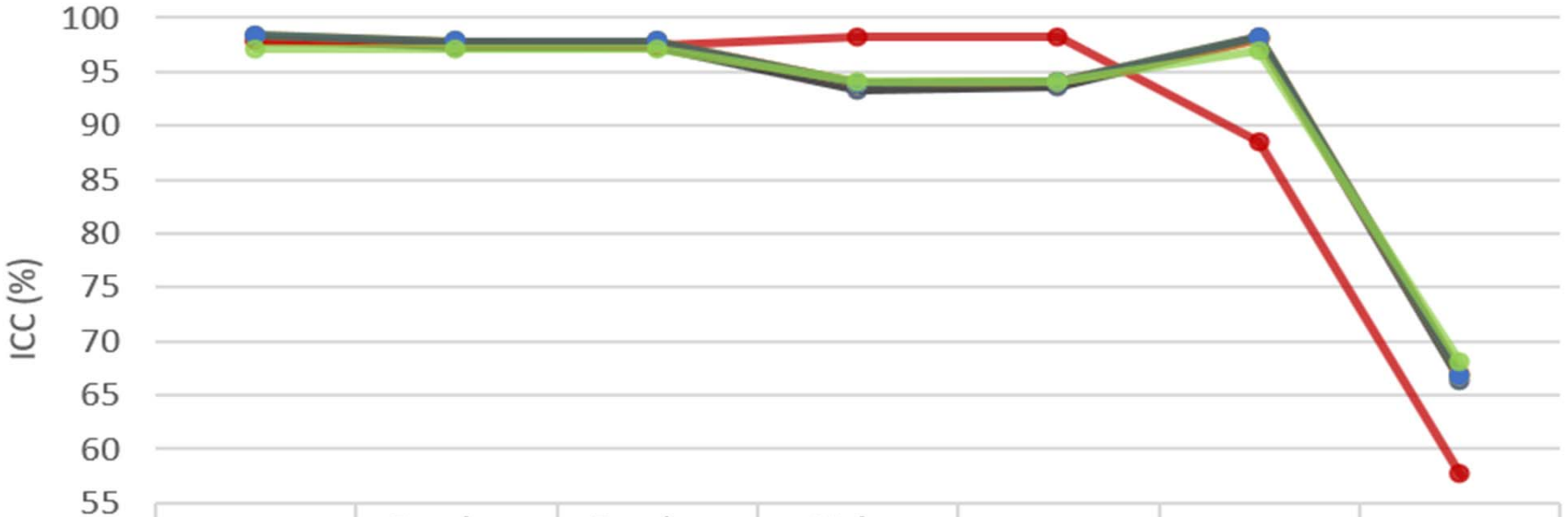
ANTICIPARSE A CIBERATAQUES

FASE 4: VALIDACIÓN DE RESULTADOS



Efectividad entre la criticidad propuesta (STRIDE) y la del proveedor de Antivirus

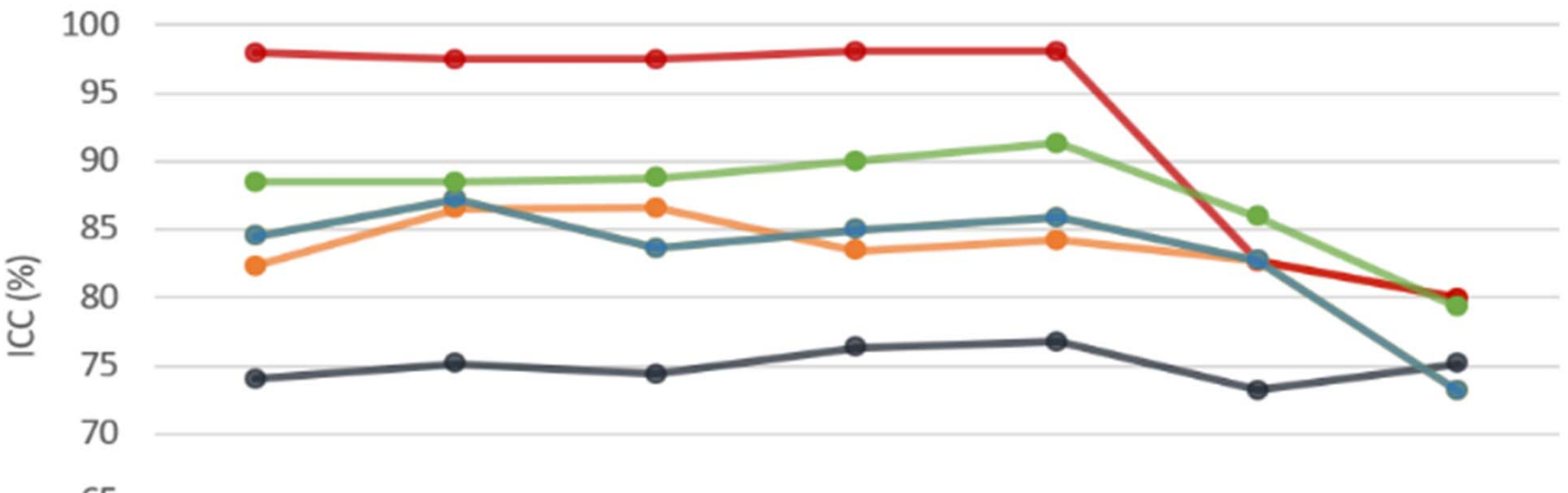
FASE 4: VALIDACIÓN DE RESULTADOS



	J48	Random Forest	Random Tree	Naive Bayes	Bayes Net	SVM	K-NN
S	98,1	97,27	97,24	93,27	93,56	98,07	66,46
T	97,92	97,72	97,71	93,92	94,13	97,99	66,97
R	97,81	97,65	97,44	98,25	98,26	88,5	57,63
I	98,47	97,75	97,75	93,9	94,08	98,29	66,83
D	98,47	97,75	97,75	93,9	94,08	98,29	66,83
E	97,06	97,15	97,08	94,04	94,04	96,96	68

Efectividad entre la categorías de STRIDE de Antivirus

FASE 4: VALIDACIÓN DE RESULTADOS



	J48	Random Forest	Random Tree	Naive Bayes	Bayes Net	SVM	K-NN
S	74,04	75,2	74,41	76,38	76,77	73,23	75,2
T	82,31	86,54	86,62	83,46	84,23	82,7	80
R	97,96	97,46	97,48	98,11	98,11	82,7	80
I	84,55	87,27	83,64	85	85,91	82,73	73,19
D	84,55	87,27	83,64	85	85,91	82,73	73,19
E	88,52	88,52	88,78	90,05	91,33	85,97	79,33

Efectividad entre la categorías de STRIDE balanceadas de Antivirus

CONCLUSIONES

Completo Sistema

- Monitorización de Eventos e Información de Seguridad

Mismas **fuentes de** datos que **SIEM Q-Radar**

- **resultados similares**

Modelo propuesto:

- **completar** y **complementar** la monitorización proporcionada por modelos comerciales
- permite la **predicción de conductas de riesgo** para anticipar respuestas ante estas situaciones

GRACIAS POR SU ATENCIÓN

@_AndresCaro_

andresc@unex.es