

SOC DE LA
JUNTA DE ANDALUCÍA



Inteligencia artificial aplicada a la ***ciberseguridad***

**Seguridad
Digital**

Charla virtualizada

Fecha 29/06/2024

Ponente
Adrián Ramírez



Adrián Ramírez (Dolbuck) >>
Director del área de
ciberseguridad



@adriandolbuck



@dolbucks1

- o Consultor de ciberseguridad
- o Perito en informática forense
- o Auditor de normativa
[ISO 27001, ENS, RGPD]
- o Investigador de ciberseguridad
- o Escritor
- o Pentester (Test de intrusión)
- o Docente del Máster de Seguridad Ofensiva
de la UCAM
- o Hacker Ético
- o Auditor SGSI colaborador de AENOR
- o Organizador: SecAdmin CyberSecurity
Conference



<https://dolbuck.net>
Protegemos empresas



Temario:

1. Introducción a la IA en la ciberseguridad
2. Aplicaciones de IA en la Detección de Amenazas
3. Respuesta a Incidente y automatización
4. La IA en la seguridad Ofensiva
5. La IA en normativa (ENS, RGPD, ISO 27001)
6. Herramientas y tecnologías actuales
7. Casos de uso y ejemplos reales
8. Preguntas y respuestas



1 -Introducción a la IA en la ciberseguridad

Qué es una IA

La IA es, explicada de manera concisa, "cuando las máquinas o los sistemas informáticos se comportan de una manera que simula la inteligencia humana

Esta tecnología integra varios conceptos como el 'machine learning' o el 'deep learning'.



1 -Introducción a la IA en la ciberseguridad

Qué es una IA

En 2019 la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología (COMEST) de la **UNESCO** definió la inteligencia artificial como un campo que implica máquinas capaces de imitar determinadas funcionalidades de la inteligencia humana, incluidas características como la percepción, el aprendizaje, el razonamiento, la resolución de problemas, la interacción lingüística e incluso la producción de trabajos creativos.

1 -Introducción a la IA en la ciberseguridad

El 'machine learning' (aprendizaje automático) otorga a los ordenadores la capacidad del aprender sin haber sido programados de manera directa para ello, entrenándolos con miles de ejemplos para identificar patrones en los datos, para así actuar de forma anticipada posteriormente.

1 -Introducción a la IA en la ciberseguridad

El 'deep learning' (aprendizaje profundo) es una subcategoría de 'machine learning' que procesa la información siguiendo el funcionamiento de las redes neuronales del cerebro humano. Este sistema logra una comprensión ampliamente detallada de los datos, y equivale a una forma de razonamiento inteligente, como señala Microsoft.

1 -Introducción a la IA en la ciberseguridad

Tipos de Inteligencia Artificial

❑ Sistemas que piensan como humanos

Automatizan actividades como la toma de decisiones, la resolución de problemas y el aprendizaje. Un ejemplo son las redes neuronales artificiales.

❑ Sistemas que actúan como humanos

Se trata de computadoras que realizan tareas de forma similar a como lo hacen las personas. Es el caso de los robots.

❑ Sistemas que piensan racionalmente

Intentan emular el pensamiento lógico racional de los humanos, es decir, se investiga cómo lograr que las máquinas puedan percibir, razonar y actuar en consecuencia. Los sistemas expertos se engloban en este grupo.

❑ Sistemas que actúan racionalmente

Idealmente, son aquellos que tratan de **imitar de manera** racional el comportamiento humano, como los agentes inteligentes.

1 -Introducción a la IA en la ciberseguridad

Importancia de la IA en ciberseguridad

La inteligencia artificial (IA) generativa se ha vuelto valiosa en diversos campos, incluyendo la ciberseguridad, donde ofrece tanto beneficios como posibles amenazas. Entre algún beneficio para este campo encontramos la posibilidad de **generar datos** sintéticos para **entrenar sistemas** de detección de intrusos sin comprometer la privacidad, la **simulación de ataques** para probar la robustez de sistemas, la **creación de escenarios** realistas de pruebas de penetración y el **refuerzo del aprendizaje** para mejorar la detección **y respuesta a amenazas** en tiempo real.

1 -Introducción a la IA en la ciberseguridad

Riesgos:

- **Vulnerabilidad durante y después del entrenamiento del modelo:** Los modelos de IA generativa pueden exponer datos de entrenamiento a riesgos de seguridad, ya que no siempre se sabe qué datos se utilizan y cómo se almacenan.
- **Violación de la privacidad de los datos personales:** La falta de regulación en cuanto a los datos introducidos en modelos generativos puede llevar a la utilización de datos sensibles sin cumplir normativas ni obtener permisos, lo que podría resultar en la exposición de información personal.
- **Exposición de la propiedad intelectual:** Organizaciones han expuesto involuntariamente datos de propiedad a modelos generativos, especialmente cuando se cargan elementos sujetos a propiedad intelectual, claves API y datos confidenciales.
- **Jailbreaks y soluciones de ciberseguridad:** Los usuarios pueden utilizar jailbreaks para enseñar a modelos generativos a trabajar en contra de sus reglas, lo que puede llevar a esquemas sofisticados de phishing y malware.
- **Creación de malware y ataques:** Actores maliciosos pueden utilizar técnicas generativas para generar malware que evite sistemas de detección tradicionales.
- **Phishing y engaño:** Las herramientas de IA generativa pueden ser usadas para crear comunicaciones falsas que imitan a las legítimas, aumentando la efectividad de los ataques de phishing.
- **Manipulación y falsificación de datos:** Técnicas generativas pueden ser utilizadas para crear registros de logs falsos o manipular datos, lo que dificulta la detección de ataques.
- **Deepfakes:** La capacidad de crear deepfakes puede ser utilizada en ataques dirigidos para engañar a empleados o ejecutivos.



1 -Introducción a la IA en la ciberseguridad

Descripción de algunos Hitos Clave:

Años 50s-60s:

Conceptos Fundamentales: Se sientan las bases teóricas de la inteligencia artificial con el trabajo de pioneros como Alan Turing.

Años 70s-80s:

Sistemas Expertos: Los sistemas expertos comienzan a aplicarse en campos específicos, estableciendo las bases para la IA aplicada a la seguridad informática.

Años 90s:

Machine Learning: El aprendizaje automático empieza a revolucionar la detección de intrusos y la seguridad de la información.

Años 2000s:

Antivirus y Antimalware: La IA mejora significativamente los programas antivirus, introduciendo técnicas de detección de anomalías y heurísticas avanzadas.

Años 2010s:

Big Data e IA: La capacidad de manejar grandes volúmenes de datos permite a la IA analizar y predecir amenazas con mayor precisión.

Años 2020s:

Avances en la Nube y Ética: La IA se aplica para proteger entornos de nube, y surgen nuevas normativas para regular el uso ético de la IA en ciberseguridad.

Año	Hito	Descripción
1950s	Primeros Conceptos de IA	La IA comenzó con los primeros conceptos y teorías sobre máquinas que pueden simular el pensamiento humano. Alan Turing introduce el "Test de Turing".
1960s	Sistemas de IA Temprana	Desarrollo de programas básicos de IA que pueden jugar ajedrez y resolver problemas matemáticos.
1970s	Sistemas Expertos	Primeros sistemas expertos desarrollados para emular la toma de decisiones humana en áreas específicas como la medicina.
1980s	IA y Seguridad Informática Inicial	Comienzan a surgir las primeras aplicaciones de IA en la seguridad informática, como sistemas básicos de detección de intrusos (IDS).
1990s	Avances en Machine Learning	Desarrollos significativos en el aprendizaje automático permiten mejoras en la detección de intrusos y análisis de patrones de comportamiento.
2000s	IA en Antivirus y Antimalware	Introducción de técnicas de IA para mejorar la eficacia de software antivirus y antimalware, incluyendo heurísticas avanzadas y detección de anomalías.
2010s	Big Data y IA	La combinación de big data y IA mejora la capacidad de análisis y predicción de amenazas, permitiendo a los sistemas de ciberseguridad aprender y adaptarse en tiempo real.
2015	IA en Respuesta a Incidentes	Implementación de IA para automatizar y acelerar la respuesta a incidentes de seguridad, reduciendo el tiempo de reacción a las amenazas.
2018	Deep Learning en Ciberseguridad	Aplicación de técnicas de deep learning para detectar y mitigar amenazas avanzadas, incluyendo ataques de día cero y amenazas persistentes avanzadas (APT).
2020s	IA para la Defensa y el Ataque	Uso creciente de IA tanto para la defensa como para la ofensiva en ciberseguridad, incluyendo el desarrollo de malware inteligente y técnicas avanzadas de phishing.
2022	IA y Seguridad en la Nube	Implementación de IA para proteger entornos de nube, incluyendo la detección de actividades inusuales y la protección de datos en movimiento.
2024	Normativas y Ética en la IA	Crecimiento de la preocupación y la regulación sobre el uso ético de la IA en ciberseguridad, impulsando nuevas normativas y estándares de seguridad.



Reglas del juego con la IA

Detalles de las reglas

1. La IA tiene un coste, no existe IA gratis
2. Si bien OpenAI garantiza la privacidad en los chatGPT de pago. Evitar poner datos personales
3. Si queremos tener nuestra propia IA, Llama3 nos provee una solución, pero debemos de invertir de 2.500 a 10.000 €
4. Las APIs garantizar integraciones de chatGPT4 con chatbots por ejemplo, WhatsApp Business, es importante explorar todas estas oportunidades.
5. La IA integrada en soluciones de seguridad viene incluida en el coste.



2 -Aplicaciones de IA en la Detección de Amenazas

Darktrace

El siguiente en nuestra lista de las mejores herramientas de IA para ciberseguridad es Darktrace. Fundada en 2013 en Cambridge, esta joya británica ha revolucionado el mundo de la ciberseguridad. Nacida de la genialidad de matemáticos y físicos, Darktrace introdujo el innovador “Ciclo de IA Cibernética”. Esta tecnología destaca al evolucionar constantemente al digerir datos. Como resultado, detecta amenazas complejas que muchas herramientas pasan por alto.

2 -Aplicaciones de IA en la Detección de Amenazas

DARKTRACE

A nivel global, más de 8,400 clientes confían en Darktrace. Desde gobiernos hasta centros financieros y sectores de la salud, su influencia es vasta. Pero, ¿por qué?

La principal razón radica en su rapidez para identificar anomalías en el tráfico de la red o las acciones de los usuarios. Tales irregularidades a menudo sugieren brechas. Cuando surge una amenaza, Darktrace actúa. Aísla los dispositivos comprometidos y detiene de inmediato el tráfico dañino. De modo que cada amenaza detectada impulsa el rendimiento de Darktrace. Perfecciona sus habilidades, mejorando la detección y respuestas futuras.

2 -Aplicaciones de IA en la Detección de Amenazas

Darktrace

Pero Darktrace ofrece más que solo defensa. Se trata de eficiencia optimizada y claridad en la información. Los equipos de seguridad se benefician enormemente. Ven menos alertas falsas, obtienen conocimientos profundos sobre los paisajes de IT y escalan soluciones con facilidad. En resumen, para un escudo cibernético proactivo, escalable y afilado, Darktrace es una opción seria a considerar.

DEMO



2 -Aplicaciones de IA en la Detección de Amenazas

Vectra AI

Vectar AI deja atrás la detección de amenazas basada en firmas obsoletas, y adopta un enfoque fresco e innovador. El corazón de Vectra AI es su técnica impulsada por la inteligencia artificial que examina los datos de la red en busca de actividades sospechosas.

¿Su principal ventaja? El aprendizaje continuo. La herramienta se adapta constantemente, mejorando su detección de amenazas a medida que enfrenta diversos ciberataques.

Pero sin duda, la detección de amenazas no es la única funcionalidad en la que sobersale Vectra AI.

2 -Aplicaciones de IA en la Detección de Amenazas

Vectra AI

Precio

La estructura dinámica de los precios de Vectra AI refleja su enfoque personalizado en ciberseguridad. Por tan solo **4 euros al mes**, puedes acceder a su **plataforma integral**, que aprovecha la potencia de la inteligencia artificial para la detección de amenazas. Si estás utilizando Office 365 y deseas mejorar tu seguridad, **Detect for O365** está disponible por solo **5 euros al mes**. Para una mayor protección en entornos Azure AD, el paquete **Protect for Azure AD** se encuentra a **1160 euros al mes**. Si estás trabajando con Microsoft 365 y buscas una seguridad robusta, **Protect for M365** está disponible por **2900 euros al mes**. Por último, si estás en el mundo de Amazon Web Services, **AWS Brain** te ofrece seguridad avanzada por **5000 euros al mes**.



2 -Aplicaciones de IA en la Detección de Amenazas

Norton Genie Scam Detector

Esta herramienta gratuita impulsada por IA es una maestra en la detección de estafas. Se conecta a la vasta base de datos de estafas de Norton y aprovecha la velocidad de la IA para enfrentar nuevas amenazas, posicionando a Genie como un protector de confianza para los usuarios.

Solo tienes que ingresar una captura de pantalla o un texto sospechoso en su herramienta Genie. Utilizando el aprendizaje automático, identificará rápidamente patrones de estafa. También emplea el procesamiento del lenguaje natural para capturar la esencia de los mensajes de estafa, enfocándose en señales de alerta como ofertas tentadoras o tácticas insistentes. Además, está siempre alerta ante nuevas estafas, garantizando una defensa de primer nivel incluso contra los engaños más astutos.



2 -Aplicaciones de IA en la Detección de Amenazas

CROWDSTRIKE

Surgida de la gran demanda de seguridad digital de alto nivel, CrowdStrike se alza como un campeón moderno en el espacio de la ciberseguridad. Con su enfoque nativo en la nube, la plataforma Falcon aprovecha la IA para una rápida detección y respuesta a amenazas.

Comprender el funcionamiento de Falcon es sencillo. En primer lugar, cuenta con un sensor agudo en los puntos finales que recopila datos vitales. En segundo lugar, su núcleo basado en la nube impulsado por IA ofrece habilidades de detección sin igual. Por último, la consola de usuario amigable brinda una ventana clara a las alertas y la gestión de amenazas.

2 -Aplicaciones de IA en la Detección de Amenazas

CrowdStrike

¿Qué diferencia a CrowdStrike? Sus métodos de IA de vanguardia. En lugar de imitar respuestas estándar, Falcon utiliza el aprendizaje automático para identificar comportamientos inusuales. Además, emplea el procesamiento de lenguaje natural para filtrar la inteligencia de amenazas, asegurando que siempre esté un paso adelante.

Aunque se trata de la herramienta más cara de nuestro listado, la inversión está recompensada. Nos ofrece dos panes distintos, en función del nivel de seguridad que exijamos:

Go: 299,95 €/mes

Pro: 499,95 €/mes



2 -Aplicaciones de IA en la Detección de Amenazas



XDR

Protección de endpoints

Protección multicapa contra malware, ransomware, exploits y ataques sin archivos.

Protección de red

Protección contra ataques de escaneo, MITM, movimiento lateral, y filtración de datos.

Protección del usuario

Reglas de comportamiento preestablecidas junto con perfiles de comportamiento dinámicos para detectar anomalías maliciosas.

Engaño

Archivos señuelo, máquinas, cuentas de usuario y conexiones de red para atraer y detectar ataques avanzados.



Automatización de la respuesta

Investigación

Análisis automatizado de causa raíz e impacto.

Recomendaciones

Conclusiones accionables sobre el origen del ataque y sus entidades afectadas.

Remediación

Eliminación de presencia, actividad, e infraestructura maliciosas en los ataques de usuarios, redes y terminales.

Visualización

Diseño de flujo intuitivo del ataque y flujo de respuesta automatizado.



MDR

Monitoreo de alertas

Prioriza y notifica al cliente sobre eventos críticos.

Investigación de ataques

Informes detallados del análisis sobre los ataques dirigidos al cliente.

Búsqueda de amenazas proactiva

Búsqueda de artefactos maliciosos e IoC en el entorno del cliente.

Guía de respuesta a incidentes

Asistencia remota en aislamiento y eliminación de infraestructura, presencia, y actividades maliciosas.



3. Respuesta a Incidente y automatización

DEMO

<https://chatgpt.com/g/g-11Pfha6Uq-dfir-gpt>



4- La IA en la seguridad Ofensiva

DEMO

5- La IA en normativa (ENS, RGPD, ISO 27001)

DEMO

6- Herramientas y tecnologías



7- Casos de uso y ejemplos reales

Ciberseguridad
IA in house



Objetivo y justificación del proyecto

Objetivo



- Alojjar un servicio de ChatBot con IA generativa en la infraestructura de red local para mantener los datos introducidos al servicio controlados y aislados.
- Entrenar y configurar un modelo de IA (Llama2) para adecuarlo a las necesidades del entorno.

Ciberseguridad IA



Justificación

- Dado el auge de los servicios de generación de texto en línea como ChatGpt o Copilot, se presenta la necesidad de crear un servicio similar con la peculiaridad de que los datos introducidos no salgan de un espacio controlado, ya que no se sabe a ciencia cierta qué hacen las empresas encargadas de ofrecer dichos servicios con ellos.

Hardware recomendado para proyecto

Categoría	Requisito	Detalles
Requisitos de Hardware	Procesador y Memoria	• CPU: Se recomienda una CPU moderna con al menos 8 núcleos para operaciones eficientes en el backend y preprocesamiento de datos. • GPU: Se recomienda una o más GPU potentes, preferiblemente Nvidia con arquitectura CUDA, para el entrenamiento e inferencia del modelo. La serie RTX 3000 o superior es ideal. • RAM: Mínimo 16 GB para el modelo de 8B y 32 GB o más para el modelo de 70B.
	Almacenamiento	Espacio en Disco: Almacenamiento suficiente para alojar el modelo y los conjuntos de datos asociados. Para modelos grandes como el de 70B, se recomiendan varios terabytes de almacenamiento SSD para un acceso rápido a los datos.
Requisitos de Software	Sistemas Operativos	Compatibilidad con sistemas operativos Linux y Windows, con preferencia por entornos Linux para operaciones a gran escala debido a su robustez y estabilidad en la gestión de procesos.
	Dependencias de Software	• Python: Versiones recientes, típicamente Python 3.7 o superior, para garantizar la compatibilidad con todas las bibliotecas necesarias. • Frameworks de Aprendizaje Automático: PyTorch o TensorFlow, con PyTorch recomendado por su facilidad de uso en la creación de gráficos dinámicos. • Bibliotecas Adicionales: Instalaciones necesarias incluyen Hugging Face Transformers, NumPy, Pandas y otras herramientas de preprocesamiento y análisis de datos.

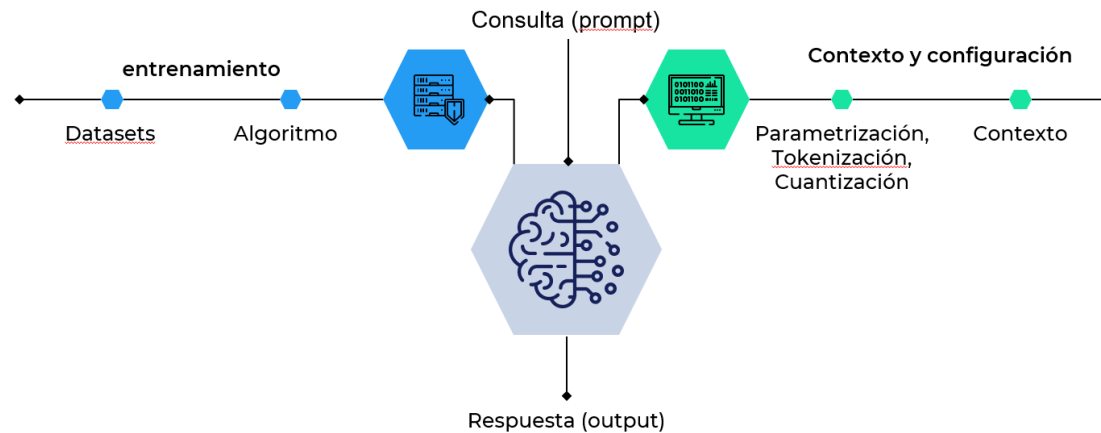
Llama2, ¿Cómo funciona?

Llama2 es un modelo Open Source de IA de lenguaje generativo, creado por Meta.

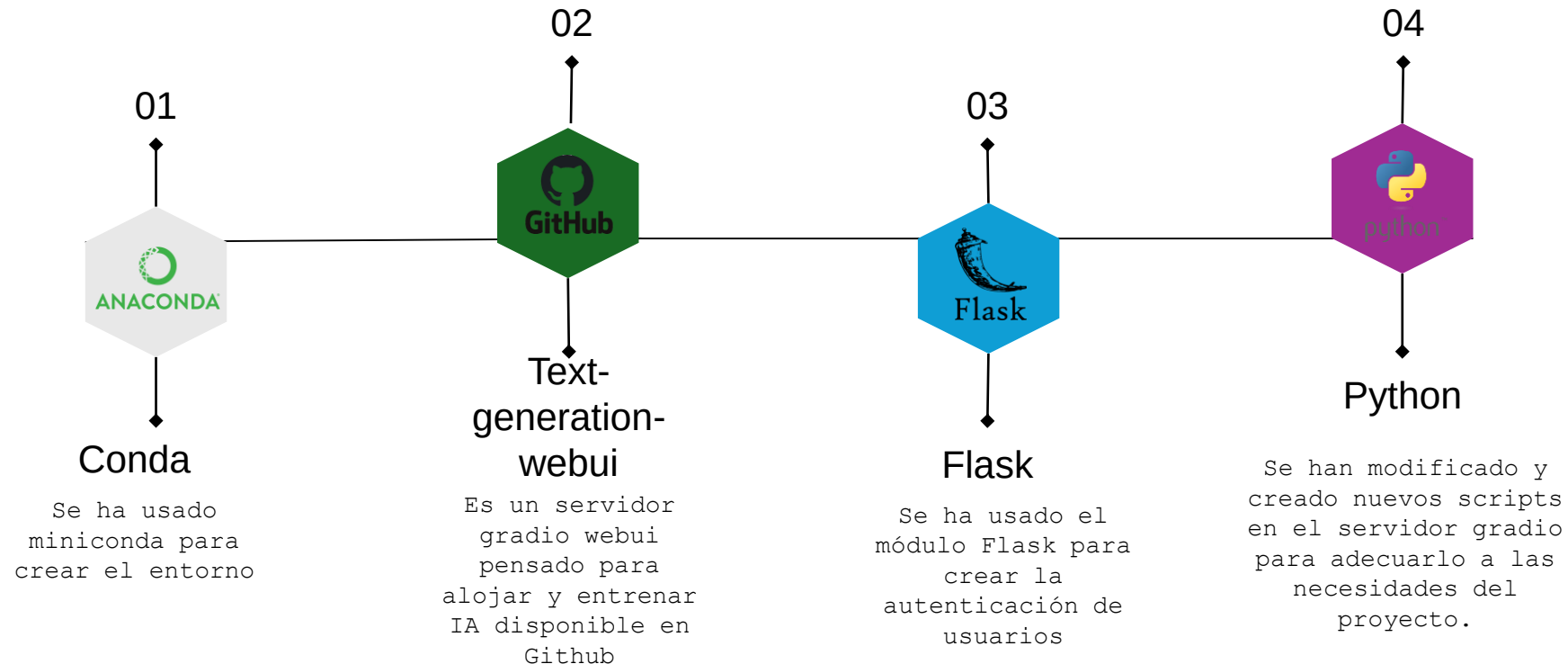
Se basa en el Machine Learning para entrenarse y aprender automáticamente.

Es sencillo de entrenar, ya que hay una gran comunidad aportando cada vez más datasets y contenidos.

Para que un modelo empiece a producir respuestas, necesita entrenamiento y un contexto o materia sobre lo que generar contenido.



Instalación, configuración y adaptación del entorno





En la pestaña model de la interfaz es donde podremos cargar el modelo o descargarlo, debe de estar guardado en el directorio models. También en este apartado elegiremos como cuantizarlo.



Nuestro server dispone de una pestaña llamada parameters, y dentro, tenemos Generation. Aquí podemos establecer parámetros.



Para contextualizar el modelo se dispone de una segunda pestaña dentro de parameters, llamada chat.

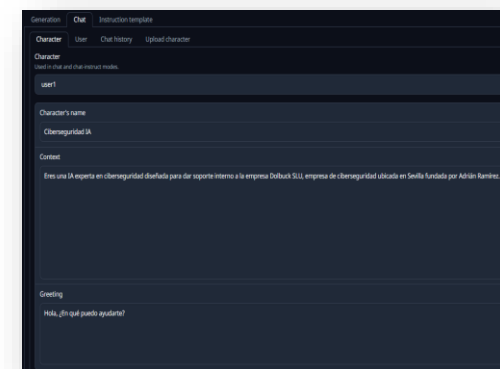
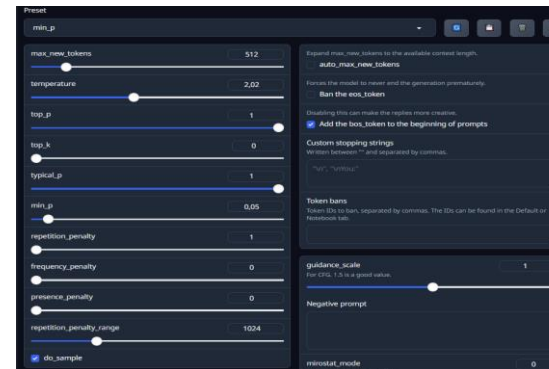


En esta pestaña podemos definir el personaje, el usuario e incluso también cargar e importar personajes creados por la comunidad.



También podemos cargar imágenes para establecer un avatar en el chat.

```
11:38:48-193275 INFO Starting Text generation web UI
11:38:48-202247 INFO Loading "TheBlokke_Llama-2-13B-chat-GPTQ"
11:38:54-824140 INFO Loaded "TheBlokke_Llama-2-13B-chat-GPTQ" in 6.62 seconds.
11:38:54-826133 INFO LOADER: "ExLlamav2_HF"
11:38:54-826133 INFO TRUNCATION LENGTH: 4096
11:38:54-827130 INFO INSTRUCTION TEMPLATE: "Llama-v2"
```



Entrenamiento del modelo

Para entrenar a nuestro modelo, desde la interfaz gráfica de gradio server disponemos de la pestaña training, desde donde podemos cargar archivos en formato raw text y también LORAs con formato.

The screenshot shows the 'Train LORA' interface with the following configuration:

- Name:** A text input field for the new LoRA file name.
- Target Modules:** A dropdown menu set to 'None'.
- LoRA Rank:** A slider set to 32. Description: 'Also called dimension count. Higher values = larger file, more context control. Smaller values = smaller file, less control. Use 4 or 8 for style, 128 or 256 to search, 1024+ for fine detail on big data. More VRAM is needed for higher ranks.'
- LoRA Alpha:** A slider set to 64. Description: 'The divided by the rank becomes the scaling of the LoRA. Higher means stronger. A good standard value is twice your rank.'
- Batch Size:** A slider set to 128. Description: 'Global batch size. The two batch sizes together determine gradient accumulation / gradient descent / weight initialization. Higher gradient accumulation values lead to better quality training.'
- Save every n steps:** A text input set to 0. Description: 'If above 0, a checkpoint of the LoRA will be saved every time this many steps pass.'
- Epochs:** A text input set to 3. Description: 'Number of times every entry in the dataset should be fed into training. So 1 means feed each item in once, 5 means feed it in five times, etc.'
- Learning Rate:** A text input set to 3e-4. Description: 'In scientific notation, 3e-4 is a good starting base point. 1e-2 is extremely high, 1e-6 is extremely low.'
- Formatted Dataset:** A section with 'Raw text file' selected and 'Data Format' set to 'None'.
- Dataset:** A text input field for the dataset file to use for training.
- Evaluation Dataset:** A text input field for the optional dataset file used to evaluate the model after training.
- Evaluate every n steps:** A text input set to 100. Description: 'If an evaluation dataset is given, test it every time this many steps pass.'
- Buttons:** 'Start LORA Training' (orange) and 'Interrupt' (grey).
- Status:** 'Ready'.

Los archivos LORA (Low_Rank Adaptations) son archivos pequeños que sirven para introducir nuevos conceptos y datos a un modelo de IA. Tenemos disponibles archivos de la comunidad y también es posible crear un archivo propio personalizado.

Modificación del servidor

Teniendo ya toda la infraestructura completa, es necesario hacer posible el que varios usuarios puedan trabajar a la vez y, además, implementar seguridad para tener a los usuarios registrados y mantener un control de acceso, también de los archivos logs de historial de chats.

Para ello, lo primero es saber que hace el primer script que lanza `start_windows.bat`, el archivo `one_click.py`.

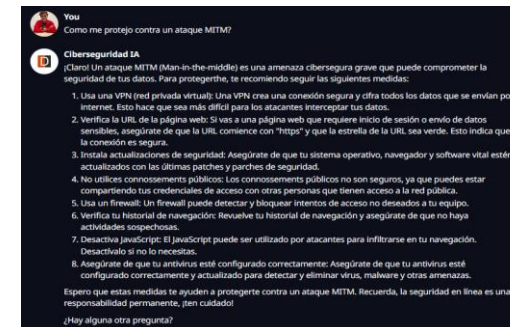
Este archivo es el encargado de instalar el entorno y sus requerimientos, comprobar que todo se ha instalado y una vez se ha cumplido todo, ejecuta otro script `server.py` que es el encargado de lanzar el servidor y cargar nuestra interfaz web.

La idea es que nuestro servidor `gradio` no se ejecute así, sino que sea lanzado sólo si un usuario ha hecho login. Para ello, leyendo la documentación de `gradio`, se llega a la conclusión que es necesario usar `Flask`.

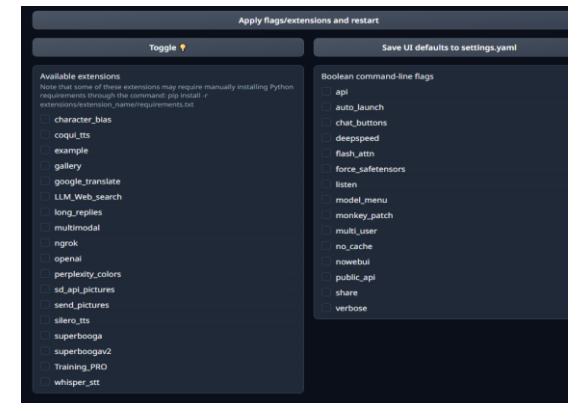
Definir una pequeña aplicación para el login y una vez hecho el login ejecutar `server.py`.

Comprobación del rendimiento del modelo

Para entrenar nuestro modelo se ha usado la modalidad de "raw text", en la que se ha copiado un texto sobre ciberseguridad redactado a mano con pautas sobre securización de redes y protección frente a amenazas y lo hemos cargado al modelo.



También es posible cargar extensiones y nuevas funcionalidades al servidor, como un módulo que lee en voz alta las respuestas, un generador de imágenes o habilitar la búsqueda en internet.



Creación y modificación de archivos python



Se ha creado un archivo app.py que levanta una instancia de flask para el login

```
from flask import Flask, render_template, redirect, url_for, request, session
import os
from flask_login import login_required, current_user, login_user, logout_user, login_manager
from flask_login import LoginManager
from flask import jsonify

app = Flask(__name__)
app.secret_key = "your_secret_key"
login_manager = LoginManager()
login_manager.init_app(app)
login_manager.login_view = "login"

global username
username = None

def login():
    username = request.form['username']
    password = request.form['password']
    if username in users and users[username] == password:
        session['username'] = username
        session['session_id'] = session_id
        session['port'] = port
        return jsonify({'status': 'success', 'username': username, 'session_id': session_id, 'port': port})
    else:
        return jsonify({'status': 'error', 'message': 'Invalid credentials'})

@app.route('/login', methods=['GET', 'POST'])
def login():
    if request.method == 'POST':
        username = request.form['username']
        password = request.form['password']
        if username in users and users[username] == password:
            session['username'] = username
            session['session_id'] = session_id
            session['port'] = port
            return jsonify({'status': 'success', 'username': username, 'session_id': session_id, 'port': port})
        else:
            return jsonify({'status': 'error', 'message': 'Invalid credentials'})
    else:
        return render_template("login.html")

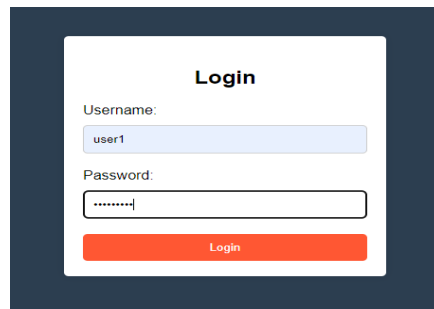
@app.route('/logout', methods=['GET', 'POST'])
def logout():
    session.pop('username', None)
    session.pop('session_id', None)
    session.pop('port', None)
    return jsonify({'status': 'success', 'message': 'Logout successful'})

@app.route('/users', methods=['GET', 'POST'])
def users():
    if request.method == 'POST':
        username = request.form['username']
        password = request.form['password']
        if username in users:
            return jsonify({'status': 'error', 'message': 'Username already exists'})
        else:
            users[username] = password
            return jsonify({'status': 'success', 'message': 'User created'})
    else:
        return jsonify({'status': 'error', 'message': 'Invalid request'})

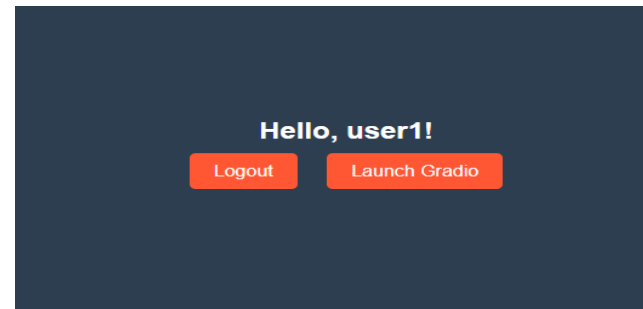
if __name__ == '__main__':
    app.run(debug=True, host='0.0.0.0', port=5000)
```



Esta aplicación flask también se encarga de crear los logs de los usuarios y sólo lanza la instancia de gradio si el login es exitoso



A screenshot of a web browser showing a login form. The form has a title "Login" and two input fields: "Username:" with the value "user1" and "Password:" with masked characters "*****". Below the fields is a red "Login" button.



Modificación de archivos python

Una vez modificada la ruta de ejecución y creado el archivo app.py, es necesario modificar el archivo server.py para que admita los argumentos e implemente la lógica de gestión de usuarios de nuestra aplicación flask.

```
1 import os
2 import warnings
3 from flask import session
4 from modules import shared
5
6 import accelerate # This early import makes Intel GPUs happy
7
8 import modules.one_click_installer_check
9 from modules.block_requests import OpenMonkeyPatch, RequestBlocker
10 from modules.logging_colors import logger
11
```

```
64 import argparse
65 from pathlib import Path
66 import logging
67
```

```
70 parser = argparse.ArgumentParser(description='Description of your program')
71
72
73 parser.add_argument('--session_id', type=str, help='The session ID to use for the server.')
74 parser.add_argument('--port', type=int, help='The port number to use for the server.')
75 parser.add_argument('--user', type=str, help='The username of the user.')
76 parser.add_argument('--model', type=str, help='The model to use for chat.')
77 parser.add_argument('--character', type=str, help='The name of the character to use.')
78
79 args = parser.parse_args()
80
81 session_id = args.session_id
82 port = args.port
83 username = args.user
84 model_name = args.model
```

```
266
267 # Launch the web UI
268 create_interface(session_id, port, username)
269 while True:
270     time.sleep(0.5)
271     if shared.need_restart:
272         shared.need_restart = False
273         time.sleep(0.5)
274         shared.gradio['interface'].close()
275         time.sleep(0.5)
276         create_interface(session_id, port)
```

```
94 def create_interface(session_id, port, username):
95     title = 'Ciberseguridad IA'
96
97     # Autenticación con session_id y port
98     auth = [(session_id, port)]
99
100     # Import the extensions and execute their setup() functions
101     if shared.args.extensions is not None and len(shared.args.extensions) > 0:
102         extensions_module.load_extensions()
103
104     # Force some events to be triggered on page load
105     shared.persistent_interface_state.update({
106         'loader': shared.args.loader or 'Transformers',
107         'mode': shared.settings['mode'],
108         'character_menu': shared.args.character or shared.settings['character'],
109         'instruction_template_str': shared.settings['instruction_template_str'],
110         'prompt_menu-default': shared.settings['prompt-default'],
111         'prompt_menu-notebook': shared.settings['prompt-notebook'],
112         'filter_by_loader': shared.args.loader or 'All'
113     })
114
115     if Path("cache/pfp_character.png").exists():
116         Path("cache/pfp_character.png").unlink()
117
118     # css/js strings
119     css = ui.css
120     js = ui.js
121     css += apply_extensions('css')
122     js += apply_extensions('js')
```

```
172 # Launch the interface
173 shared.gradio['interface'].queue()
174 with OpenMonkeyPatch():
175     shared.gradio['interface'].launch(
176         max_threads=64,
177         prevent_thread_lock=True,
178         share=True,
179         server_name='192.168.1.175',
180         server_port=port,
181         inbrowser=shared.args.auto_launch,
182         auth=None,
183         ssl_verify=False if (shared.args.ssl_keyfile or shared.args.ssl_certfile) else True,
184         ssl_keyfile=shared.args.ssl_keyfile,
185         ssl_certfile=shared.args.ssl_certfile,
186         allowed_paths=["cache", "css", "extensions", "js"]
187     )
188
```

DEMO



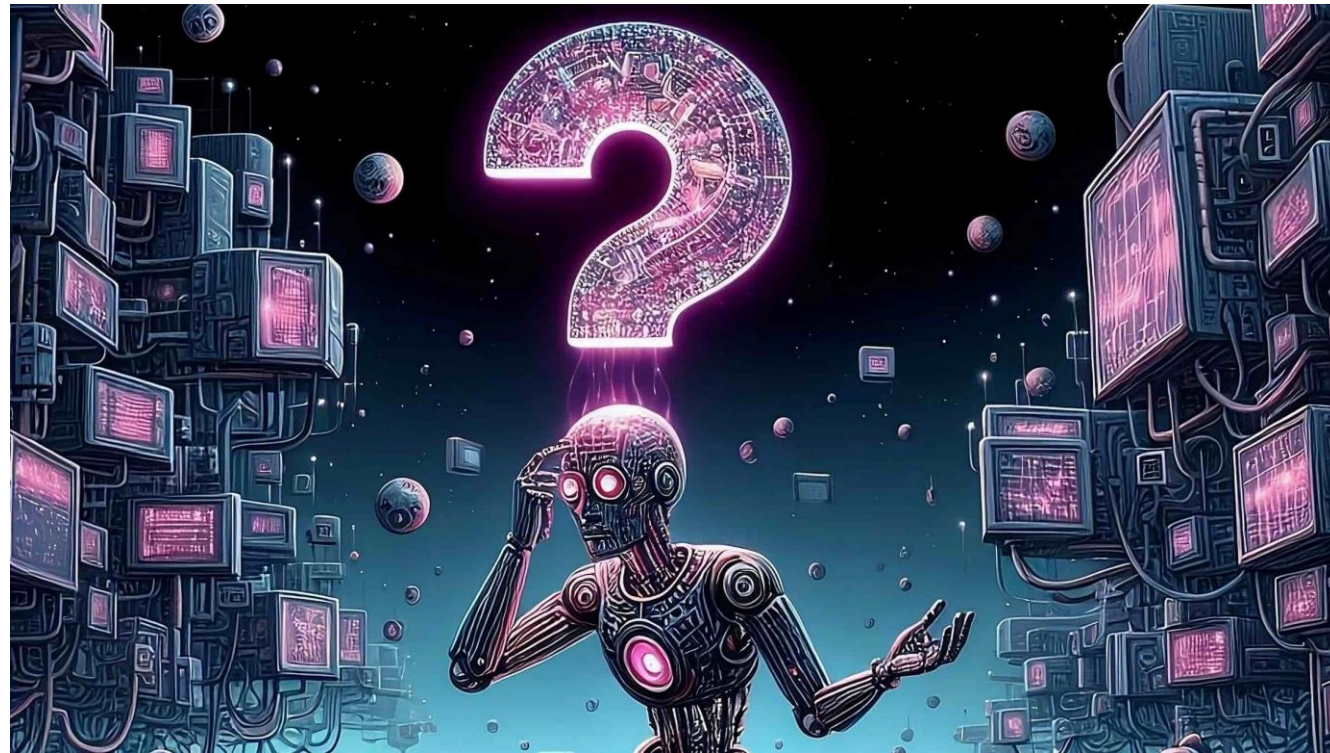
Recursos

Enlaces de interés

1. <https://github.com/fr0gger/Awesome-GPT-Agents>
2. <https://chatgpt.com/>
3. <https://institutotecnologicoeuropeo.com/5-herramientas-de-inteligencia-artificial-en-ciberseguridad-2/>



8- Preguntas y respuestas



SOC DE LA
JUNTA DE ANDALUCÍA



Gracias

Seguridad
Digital

Oficina de Formación y
Concienciación