Instituto
Politécnico
Nacional

Temas de Tesis

Raúl Acosta Bermejo

2024-A

Centro de
Investigación en
Computación

Instituto
Politécnico
Nacional

Centro de
Investigación en
Computación

IA en Ciber

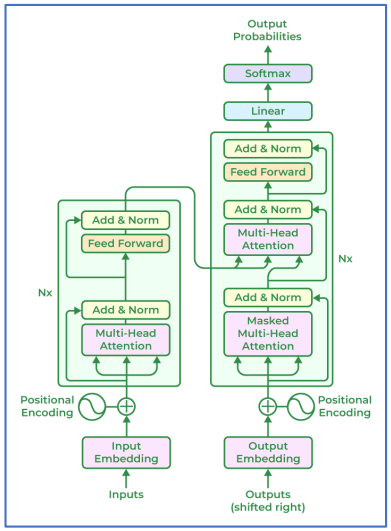
- Introduccion a LLM
  - Large Language Model
- Propuestas de Tesis
  - Generación
  - Ataque / Defensa

LLM (Large Language Model, Enormes Modelos del Lenguaje)

Descripción

Procesamiento de Lenguaje Natural

- Son modelos que:
  - Se crearon para mejorar los algoritmos de PLN
    - Usan Redes Neuronales (Recurrentes)
    - Aprendizaje supervisado para Clasificar
    - Datasets específicos etiquetados por Expertos => Corpus
  - Reúsan conceptos de:
    - Procesamiento de imágenes
    - Modelos pre-entrenados (fine-tuning)
    - Transformers creados en el 2017 para procesar secuencias de textos.
  - Modelos del Lenguaje basados en Transformers.
    - Ahora ya no se requiere crear datasets específicos
    - Se cambio el paradigma a un Aprendizaje auto-supervisado.
    - Los datos de entrenamiento se tomaron de documentos de internet

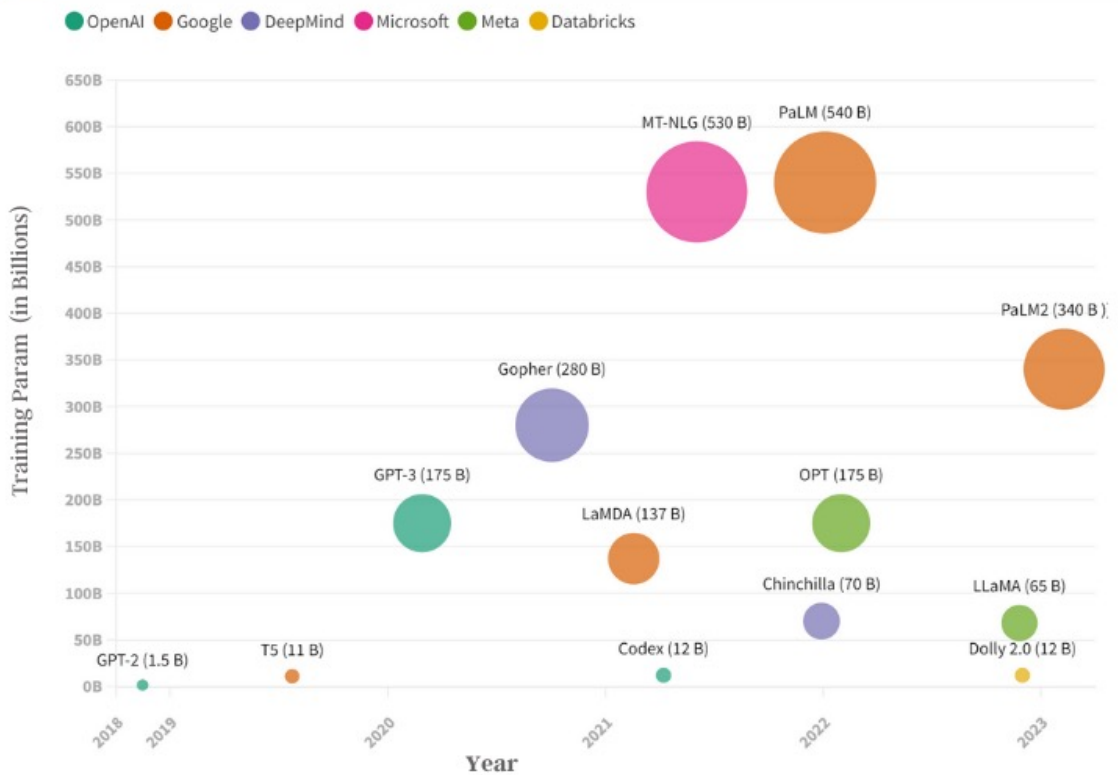


LLM

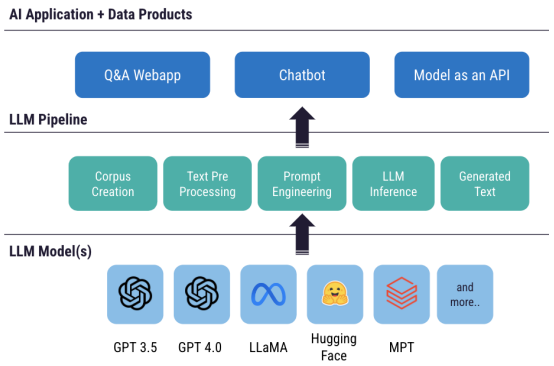
Empresas

- ❖ Ejemplos de modelos abiertos
  - OpenAI
    - Creo GPT (ver. 3.5 y 4)
    - Usados en ChatGPT.
  - Google
    - PaLM y Gemini (antes Bard).
  - Meta
    - LLaMA

La comunidad de Hugging Face
Portal con muchos de los LLMs
Python y la librería Transformers



LLM



Descripción

- Aplicaciones, IA generativa
  - Crear imágenes
  - Crear audios
  - Se puede usar un modelo existente y reentrenarlo para un uso especifico:
    - Crear programas, crear consultas a un BD, etc.

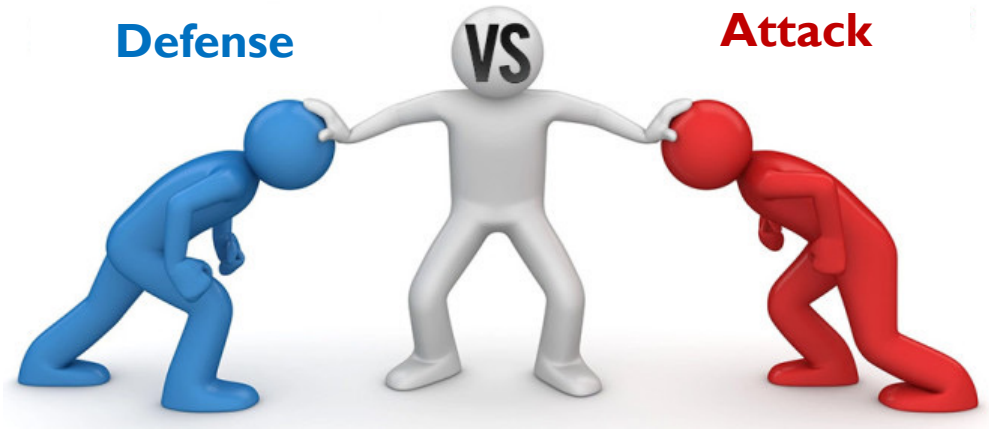
¿Y en ciberseguridad como se podrían usar?

Problema: no puede uno pedirle que haga cosas malas

¡ Por ejemplo, pedir le cómo crear una bomba molotov !

Te dice que no sabe, te recomienda una línea de ayuda (URL y teléfonos), formas de manejar la frustración.

Tesis



countermeasures

- La **escalada**, espiral de violencia es una peculiaridad de los conflictos que implica un enlazamiento de actos progresivamente más violentos de las partes en pugna.
- Se basa en un mecanismo de **acción-reacción**. a una amenaza pequeña se responde con una amenaza ligeramente superior, así indefinidamente.

Evolución
Técnicas de ataque (más sofisticadas).
Mecanismos de Defensa (más complejos).

Tesis

MAREA CISEG
Cibersecurity Laboratory

MAlware REpository for the Academy

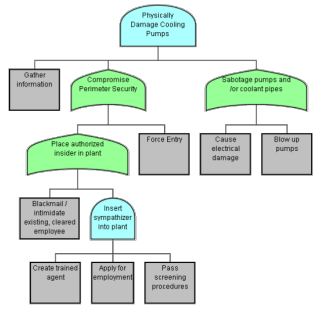
+34 millones de muestras
VirusShare
VirusSign
VxHeaven
Drebin
Wuhan
Etc.

Ataque

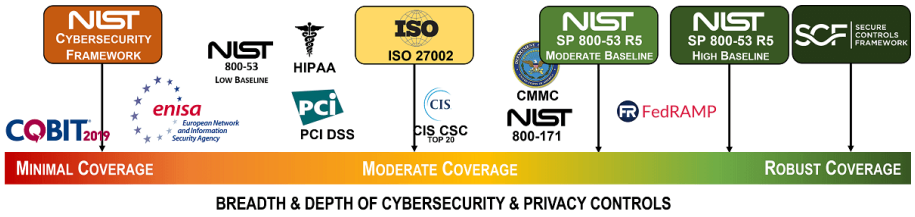
❖ Usarlos para generar:

- Malware
  - Retos: generar datos binarios
  - Archivos ejecutables, imágenes ofuscadas, etc.
- Modelos de amenazas / Árboles de Ataque
  - Tomar como base un estándar. TTP (12 tac, +200tec, Proc)
- Ejemplo de un prompt
  - Crear un virus metamórfico para Windows, y
  - Que use una técnica de ofuscación basada en UPX, y
  - Que use como contramedida un UPX no vanilla.

MITRE ATT&CK



Tesis



Defensa

❖ Usarlos para generar:

- Controles de seguridad
  - Reglas de un firewall, un IDS (Net, Snort), un IPS.
  - Las políticas de un EDR (*Endpoint Detection and Response*).
- Árboles de defensa
  - Tomar como base el CIS Controls
- Ejemplo de un prompt
  - Crear un control para proteger los datos,
  - ¿Qué nivel de madurez se requiere?
  - Usar el grupo 1, cifrar datos en el File System.
  - ¿Qué SO y se puede usar software comercial?

Requerimientos de Seguridad

Preguntas o Comentarios

Personal Web Page

https://www.cic.ipn.mx/~racostab

Email

racostab@ipn.mx

racostab@cic.ipn.mx

Tel.

55-57-59-60-00

Ext. 56652

Ver la sección de
Tesis / Nuevos temas