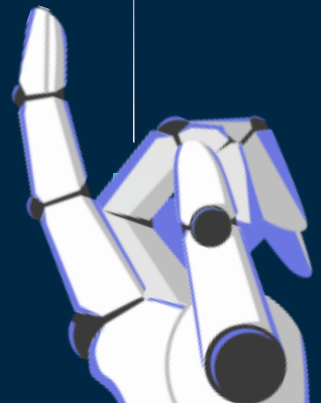
MODELO DE IA PARA AUTOMATIZAR LA RECOPILACIÓN DE INFORMACIÓN EN EL PENTESTING

Alejandro Arce Rendón
Alexander Samacá Burbano

Tutor: Christian Camilo Urcuqui López, MSc



CONTENIDO

1 MOTIVACIÓN Y ANTECEDENTES

2 DESCRIPCIÓN DEL PROBLEMA

3 OBJETIVO GENERAL Y ESPECÍFICOS

4 MARCO TEÓRICO

5 ESTADO DEL ARTE

6 METODOLOGÍA Y EXPERIMENTOS

7 ENTREGABLES POR OBJETIVOS

8 CONCLUSIONES Y TRABAJO FUTURO

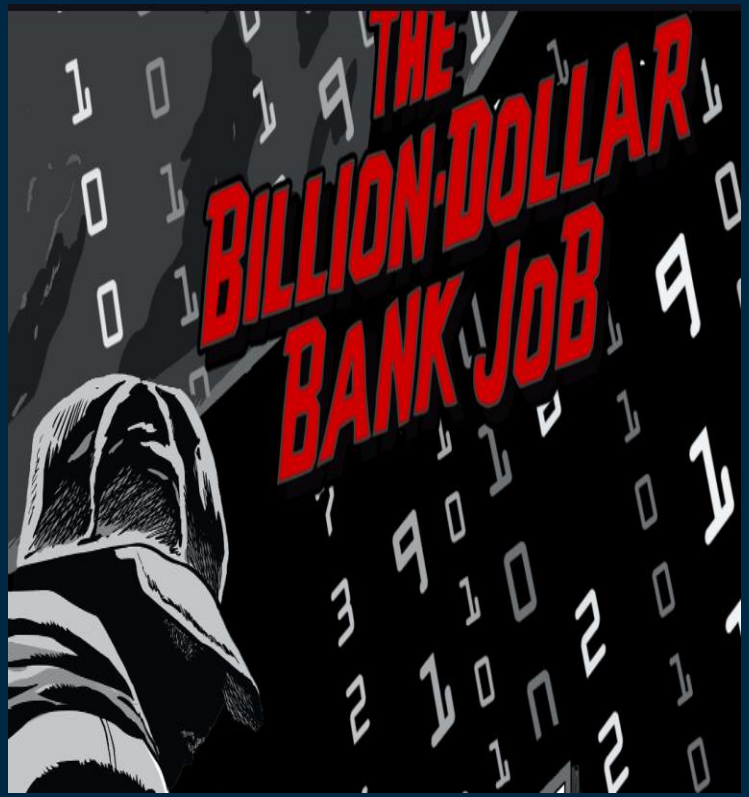
ABSTRACT

The project main idea is to improve the recon phase on pentesting using A1 models to work and interact with powerful tools that require some previous knowledge, this can obstruct the optimum time to prepare the recon phase. In the following project we developed A1 models that recognize natural language and implement them as a orchestrator to improve the use of different tools to gather information.

MOTIVACIÓN Y ANTECEDENTES

01

Caso BCB 81 millones perdidos



IN 2016, A MYSTERIOUS SYNDICATE TRIED TO STEAL $951 MILLION FROM BANGLADESH'S CENTRAL BANK – AND LAID BARE A PROFOUND WEAKNESS IN THE SYSTEM BY WHICH MONEY MOVES AROUND THE WORLD.

At 8:45 in the morning on Friday, Feb. 5, 2016, Zubair Bin Huda, a director at Bangladesh's central bank, entered the 30-story, concrete-and-glass headquarters in Dhaka. Bin Huda, slim and soft-spoken, with a thin black mustache and beard, rode an elevator to the ninth floor and eventually walked into the back office of the Accounts and Budgeting Department's "dealing room," the most restricted area of the building, accessible to only a handful of employees.

'THESE CENTRAL BANKS OFTEN CANNOT AFFORD GOOD SECURITY, GOOD SOFTWARE, OR HIRE A PROPER SPECIALIST TO CONFIGURE THEIR NETWORK.'

- https://theonebrief.com/the-bangladesh-bank-heist-lessons-in-cyber-vulnerability/
- https://www.nytimes.com/interactive/2018/05/03/magazine/money-issue-bangladesh-billion-dollar-bank-heist.html

MOTIVACIÓN Y ANTECEDENTES

02

Solo 5% de las carpetas de las organizaciones están correctamente protegidos

- https://info.varonis.com/hubfs/Varonis%202019%20Global%20Data%20Risk%20Report.pdf

MOTIVACIÓN Y ANTECEDENTES

03

"Según estimados para 2023, los cibercriminales robarán 33 billones de datos"


JUNIPER RESEARCH

Home > Press > Press releases > Cybersecurity Breaches to Result in Over 146 Billion Records Being Stolen by 2023

CYBERSECURITY BREACHES TO RESULT IN OVER 146 BILLION RECORDS BEING STOLEN BY 2023

Number of records breached each year to nearly triple over the next 5 years, while cybersecurity spend will only increase by an average of 9% per company per annum

Hampshire, UK – 8th August 2018: A new report by Juniper Research found that over 33 billion records will be stolen by cybercriminals in 2023 alone, an increase of 175% over the 12 billion records expected to be compromised in 2018, resulting in cumulative loss of over 146 billion records for the whole period.

https://www.juniperresearch.com/home

https://www.juniperresearch.com/press/cybersecurity-breaches-to-result-in-over-146-bn

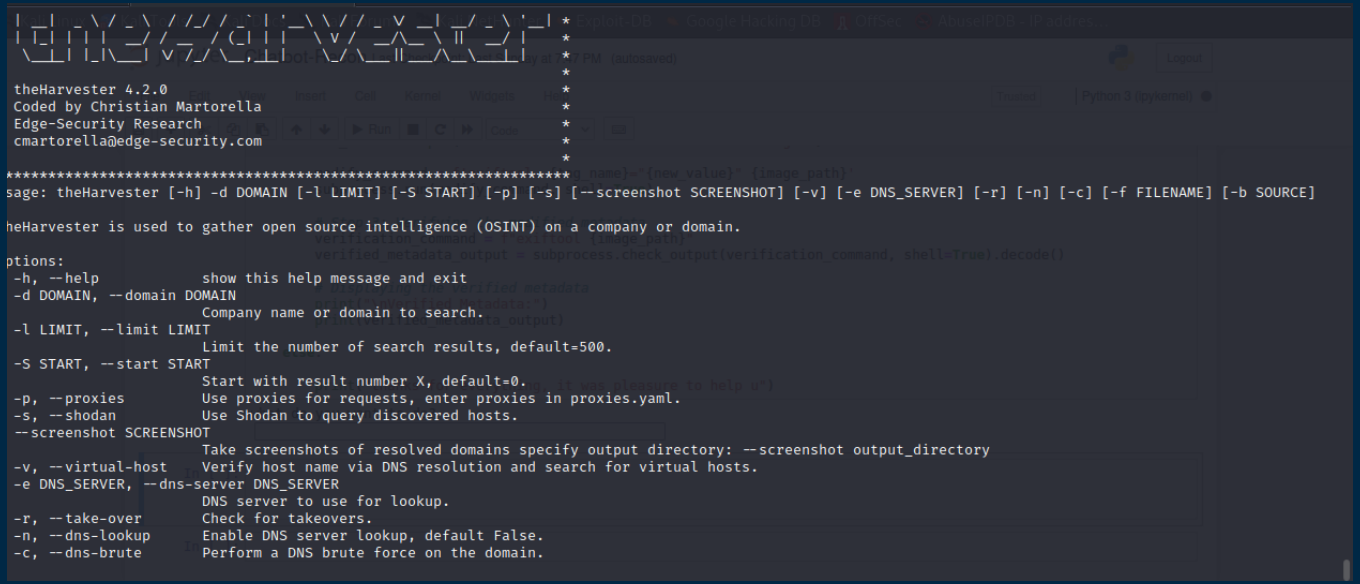
¿Qué es pentesting?

https://sp.depositphotos.com/stock-photos/pentest.html

DESCRIPCIÓN DEL PROBLEMA

Las herramientas actuales para encontrar información relevante en la fase de reconocimiento del pentesting son poco intuitivas y requieren de experticia

Ejemplo de pantalla de ayuda de theHarvester



```
 _   _                                            _
| |_| |_ ___ \ / _ \  / /_/ / _` | '__\ \ / / _ \ __| __/ _ \ '__|
| | | | | |  __/ / __  / (_| | |   \ V /  __/\__ \ ||  __/ |
 \___|_|_|\___| \/ /_/ \__,_|_|    \_/ \___||___/\__\___|_|

theHarvester 4.2.0
Coded by Christian Martorella
Edge-Security Research
cmartorella@edge-security.com

*******************************************************************
sage: theHarvester [-h] -d DOMAIN [-l LIMIT] [-S START] [-p] [-s] [--screenshot SCREENSHOT] [-v] [-e DNS_SERVER] [-r] [-n] [-c] [-f FILENAME] [-b SOURCE]

heHarvester is used to gather open source intelligence (OSINT) on a company or domain.

ptions:
 -h, --help            show this help message and exit
 -d DOMAIN, --domain DOMAIN
                       Company name or domain to search.
 -l LIMIT, --limit LIMIT
                       Limit the number of search results, default=500.
 -S START, --start START
                       Start with result number X, default=0.
 -p, --proxies         Use proxies for requests, enter proxies in proxies.yaml.
 -s, --shodan          Use Shodan to query discovered hosts.
 --screenshot SCREENSHOT
                       Take screenshots of resolved domains specify output directory: --screenshot output_directory
 -v, --virtual-host    Verify host name via DNS resolution and search for virtual hosts.
 -e DNS_SERVER, --dns-server DNS_SERVER
                       DNS server to use for lookup.
 -r, --take-over       Check for takeovers.
 -n, --dns-lookup      Enable DNS server lookup, default False.
 -c, --dns-brute       Perform a DNS brute force on the domain.
```

OBJETIVO GENERAL

Recopilar y filtrar la información en la fase de reconocimiento de los ataques de pentesting, utilizando un modelo de inteligencia artificial.

OBJETIVOS ESPECÍFICOS

1. Definir las fuentes y las variables necesarias para el desarrollo del modelo que apoyará el reconocimiento durante un ataque de pentesting.

2. Entrenar modelos de IA que permitan facilitar el proceso de reconocimiento en las fases de pentesting.

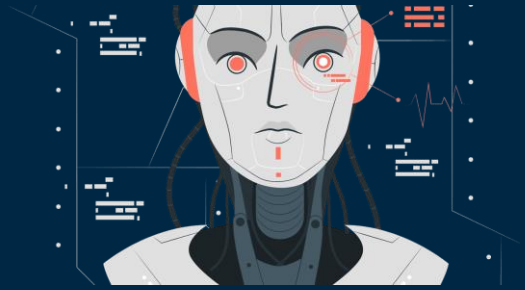
3. Validar los mejores modelos de Inteligencia Artificial usando procesamiento de lenguaje natural para la fase de reconocimiento utilizando métricas estadísticas.

4. Implementar y conectar un aplicativo entre el modelo y las herramientas para la etapa de reconocimiento

MARCO TEÓRICO

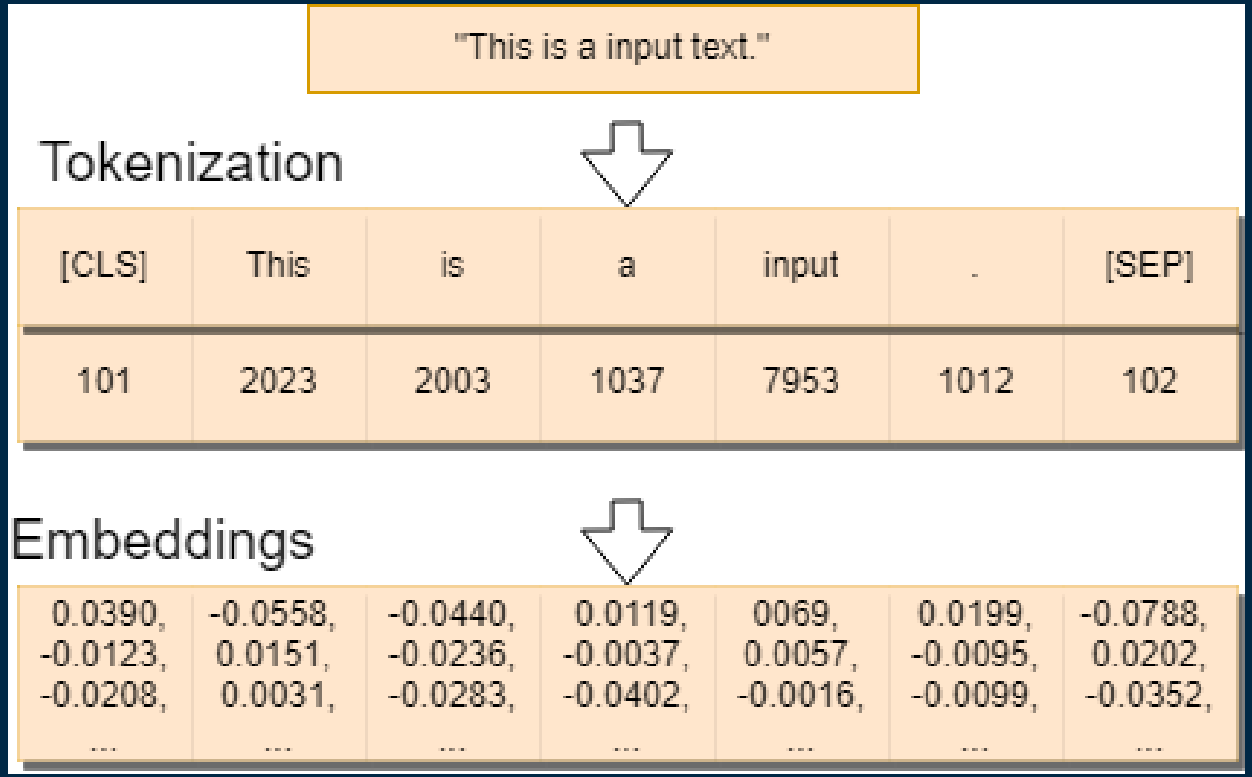
Ciberseguridad

Inteligencia Artificial



¿Qué es NLP?

Tokenización



https://vaclavkosar.com/ml/Tokenization-in-Machine-Learning-Explained

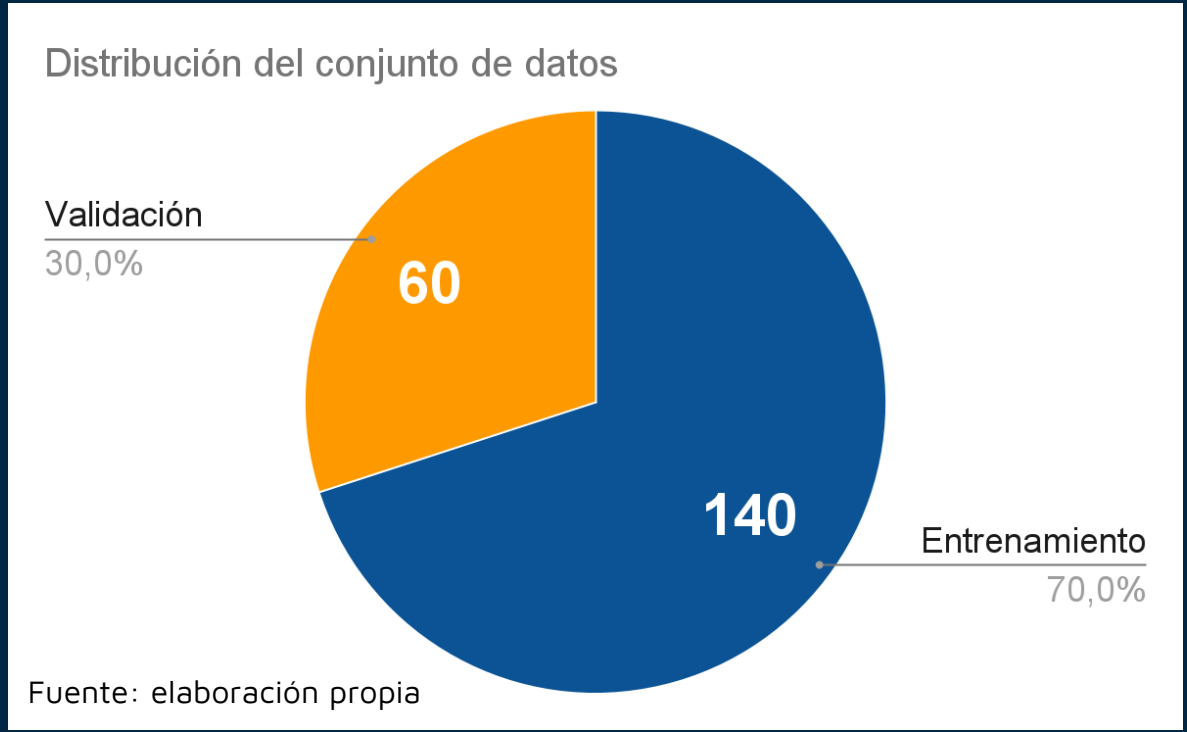
ESTADO DEL ARTE

Sistemas/ Caracteristicas	Recon Framework	Machine Learning	Enumeración	Recon pasivo
Nuestro proyecto	●	●	●	●
Autonomous Penetration Testing using RL		●	●	
Autonomous Security Analysis and Pentesting		●		
Reinforcement Learning for Efficient Network Penetration Testing		●		
Automation of Recon for Ethical Hackers	●		●	●
Automation of Cyber-Reconnaissance	●		●	●

METODOLOGÍA: CRISP-DM

ENTENDIMIENTO DEL NEGOCIO

ENTENDIMIENTO DE LOS DATOS



Fuente: elaboración propia

“I want to search for IP addresses associated with a specific domain”

PREPARACIÓN DE LOS DATOS

- Vectorizamos los textos para que los modelos pudiesen entender los textos usando un tokenizer.
- Se cambiaron los nombres de las herramientas a una representación numérica para la clasificación.

MODELOS

EXPERIMENTO 1: MODELO flan-t5

EXPERIMENTO 2: MODELO BERT-Base


```
[
    {
        "context": "Cybersecurity is a constantly evolving field, as new threats and vulnerabilities emerge with advances in technology. As such,
        "qas": [
            {
                "id": "00001",
                "is_impossible": false,
                "question": "Cibersecurity is an obsolete field?",
                "answers": [
                    {
                        "text": "Is a constantly evolving field, as new threats and vulnerabilities emerge with advances in technology.",
                        "answer_start": 14
                    },
                    {
                        "text": "Cibersecurity is a constantly evolving field, as new threats and vulnerabilities emerge with advances in technol
                        "answer_start": 0
                    },
                    {
                        "text": "it is important for individuals and organizations to stay up-to-date on the latest security trends and best prac
                        "answer_start": 106
                    }
                ]
            }
        ]
    }
]
```

EXPERIMENTO 3: MODELO secureBERT

```
Text here:      I want to get information of a target, the tool for that is <mask>
SecureBERT:
Mask  Predictions :  [':', 'nmap', '…', '', '.', ':', 'info', 'Metasploit', 'the', 'Nessus']
==========================
```

```
Text here:      I would like to get pictures from a target, the tool is <mask>
SecureBERT:
Mask  Predictions :  ['…', ':', '', '.', '…', 'not', ':', '.', '', 'a']
==========================
```

RESULTADOS

Epoch	Accuracy	Precision	F1 Score	Validation Loss
1	0.95	0.95	0.95	0.49
2	0.96	0.96	0.96	0.17
3	0.90	0.92	0.90	0.62

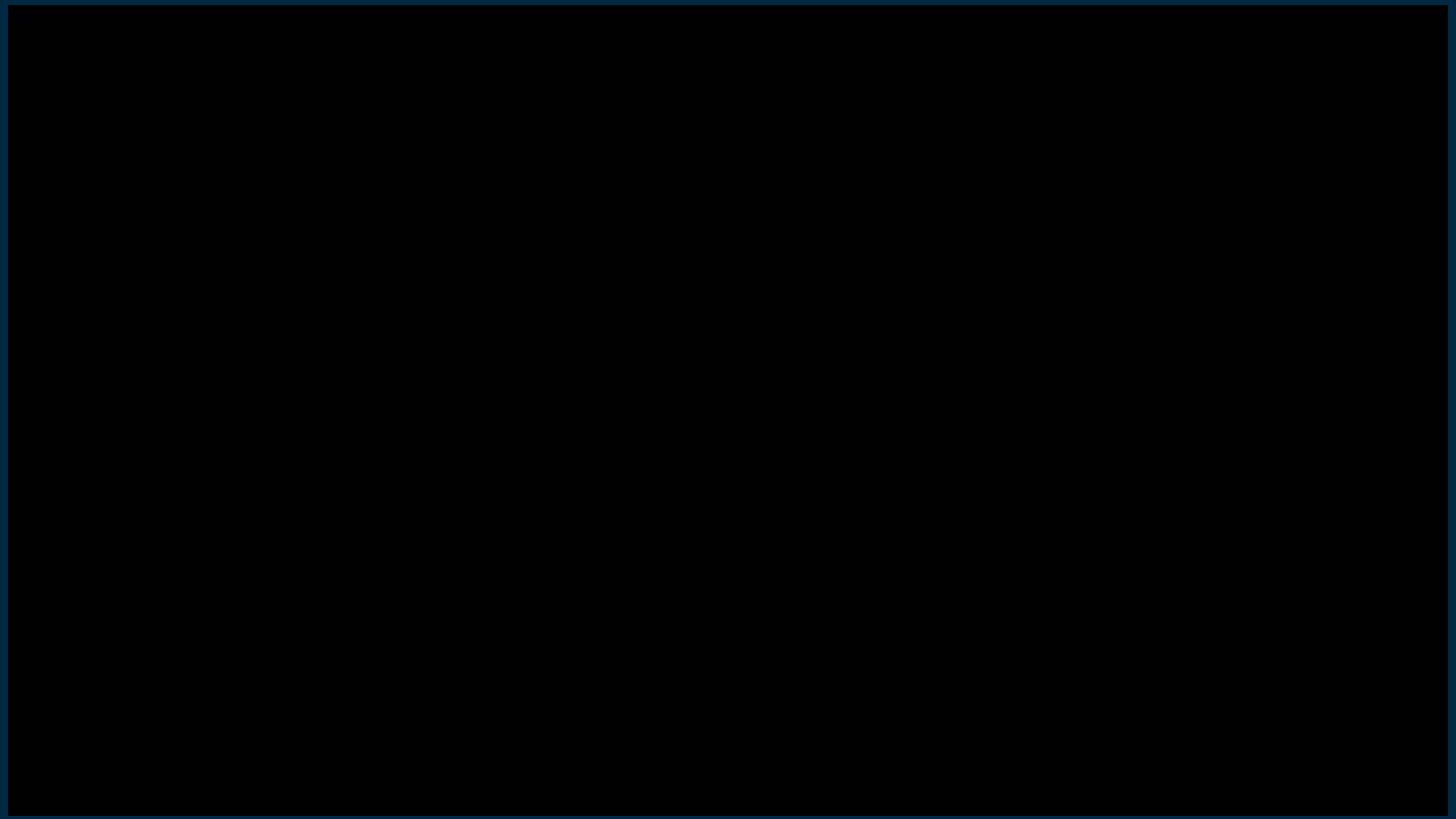
RESULTADOS: MODELO secureBERT

```
what do you want to do?
i want get info about a domain

The tools that are identified to help you are theHarvester and Shodan, we will guide u on a step by step to use them
Available parameters of search
1. Domain
2. IP address
1

Enter the target domain(www.example.com): www.google.com
```

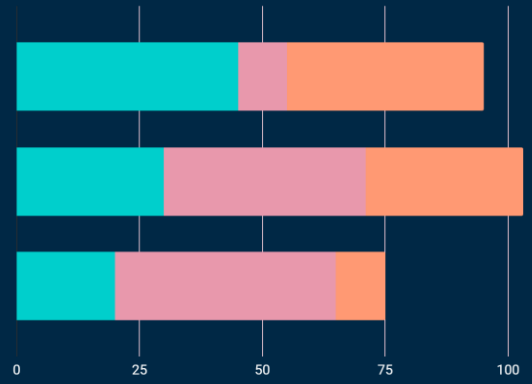
```
[*] Interesting Urls found: 1
--------------------
https://www.google.com/?gws_rd=ssl

[*] LinkedIn Links found: 0
----------------------

[*] IPs found: 1158
-------------------
1.0.0.8
1.0.0.10
1.0.0.12
1.0.0.18
1.0.0.20
1.0.0.37
1.234.65.170
3.23.157.58
```

```
[*] Emails found: 10
----------------------
ameurhosni@google.com
codyheiner@google.com
jillmckinnon@google.com
jordyzomer@google.com
kennethrosario@google.com
louischiu@google.com
markrowe@google.com
michamazur@google.com
stevenvanni@google.com
trast@google.com

[*] Hosts found: 88
---------------------
02.www.google.com
040.www.google.com
```

HERRAMIENTA RECON-NLP

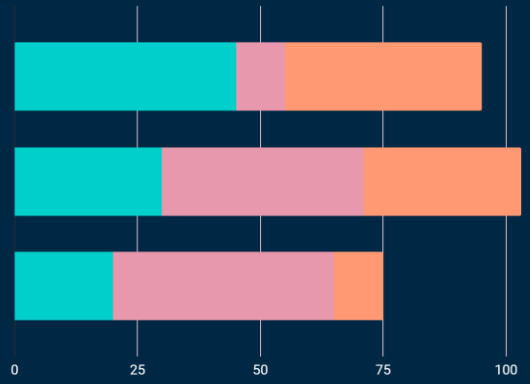


ENTREGABLES POR OBJETIVOS



Objetivos	Entregables
Definir las fuentes y las variables necesarias para el desarrollo del modelo	1. Matriz con las variables y fuentes necesarias
Entrenar modelos de IA para la fase de reconocimiento en el pentesting	1. Código de los modelos seleccionados
Validar los mejores modelos de IA	1. Código con los resultados de los modelos. 2. Resultado de las pruebas

ENTREGABLES POR OBJETIVOS



Objetivos	Entregables
4. Implementar y conectar un aplicativo entre el modelo y las herramientas para la etapa de reconocimiento	1. Aplicativo con su instructivo

CONCLUSIONES

- Gracias al uso de SecureBERT y la incorporación de oraciones poco técnicas, permite entrenar un nuevo modelo de clasificación que dé como resultado herramientas útiles de ciberseguridad.
- Se logra entrenar un modelo de NLP con un exactitud del 96%.
- Es posible utilizar herramientas de OSINT como Shodan, TheHarvester y ExploitDB en modelos de IA para la recopilación de información pública.

TRABAJO FUTURO

- Interfaz gráfica
- Orquestar otras herramientas comunes en ciberseguridad para la recopilación de datos públicos
- Incorporación de nueva información al dataset creado, con el fin de mejorar los resultados y escalar el modelo.
- Integración con ChatGPT.

¿PREGUNTAS?

REFERENCIAS BIBLIOGRÁFICAS

- Urcuqui López, C.C. y Navarro Cadavid, A. (coords.) (2022). Ciberseguridad: los datos tienen la respuesta. Cali pag: Editorial Universidad Icesi. DOI: https://doi.org/10.18046/EUI/ee.4.2022
- G., N. (2022, Agosto 5). 35 Outrageous Hacking Statistics & Predictions [2022 Update]. Review42. https://review42.com/resources/hacking-statistics/
- Gregg, M., & Santos, O. (2022, Marzo 10). Footprinting, Reconnaissance, and Scanning | “Do I Know This Already?” Quiz. Pearson IT Certification. Retrieved from https://www.pearsonitcertification.com/articles/article.aspx?p=3129461
- The Bangladesh Bank Heist: Lessons In Cyber Vulnerability. (2019, Septiembre 13). The One Brief. Retrieved from :https://theonebrief.com/the-bangladesh-bank-heist-lessons-in-cyber-vulnerability/
- IBM Education Cloud. (2021, Junio). Supervised Learning. https://www.ibm.com/cloud/learn/supervised-learning
- xiv.org/abs/1905.05965

REFERENCIAS BIBLIOGRÁFICAS

- Lanmaster53. s.f. *Open Source Intelligence Gathering Tool Aimed at Reducing the Time Spent Harvesting Information From Open Sources*. GitHub. https://github.com/lanmaster53/recon-ng
- V. R. Saraswathi, I. S. Ahmed, S. M. Reddy, S. Akshay, V. M. Reddy and S. M. Reddy. (2022). *Automation of Recon Process for Ethical Hackers*. International Conference for Advancement in Technology (ICONAT). pg. 1-6, doi: 10.1109/ICONAT53423.2022.9726077
- Roy, A., Mejia, L., Helling, P., & Olmsted, A. (2017). *Automation of cyber-reconnaissance: A Java-based open source tool for information gathering*. In 2017 12th International Conference for Internet Technology and Secured Transactions (ICITST) (pg. 424-426). IEEE. https://ieeexplore.ieee.org/abstract/document/8356437
- Schwartz, J. (2019, Mayo 15). *Autonomous Penetration Testing using Reinforcement Learning*. arXiv.org. https://ar

GRACIAS

Plantilla tomada de