UNIVERSIDAD CUAUHTÉMOC
VANGUARDIA Y EXCELENCIA

EDUCACIÓN A DISTANCIA
UNIVERSIDAD CUAUHTÉMOC

ACUERDO NO. 1857 CON FECHA DEL 07 DE JULIO DE 2015 DEL INSTITUTO DE EDUCACIÓN DEL ESTADO DE AGUASCALIENTES

"SISTEMA EXPERTO PARA LA DETECCIÓN DE VOCABULARIO DE CIBERGUERRA EN SITIOS WEB, UTILIZANDO TÉCNICAS DE MACHINE LEARNING Y DEEP LEARNING CON PYTHON"

TESIS PARA: **MAESTRÍA EN CIENCIAS DE LOS DATOS Y PROCESAMIENTO DE DATOS MASIVOS (BIG-DATA)**

PRESENTA(N): **ELIANA PAOLA PÉREZ ARRIETA**

DIRECTOR(A) DE TESIS: **DR. IVÁN CASTILLO ZÚÑIGA**

16 de Julio de 2020. Aguascalientes, México

ASUNTO: Carta de autorización.

Aguascalientes, Ags., 16 de julio de 2020.

LIC. ROGELIO MARTÍNEZ BRIONES
UNIVERSIDAD CUAUHTÉMOC PLANTEL AGUASCALIENTES
RECTOR GENERAL

P R E S E N T E

Por medio de la presente, me permito informar a Usted que he asesorado y revisado el trabajo de tesis titulado:

“ SISTEMA EXPERTO PARA LA DETECCIÓN DE VOCABULARIO DE CIBERGUERRA EN SITIOS WEB, UTILIZANDO TÉCNICAS DE MACHINE LEARNING Y DEEP LEARNING CON PYTHON ”

Elaborado por la Licenciada en Mátemáticas **ELIANA PAOLA PÉREZ ARRIETA**, considerando que cubre los requisitos para poder ser presentado como trabajo recepcional para obtener el grado de Maestra en Ciencia de los Datos y Procesamiento de Datos Masivos (BIG-DATA).

Agradeciendo de antemano la atención que se sirva a dar la presente, quedo a sus apreciables órdenes.

ATENTAMENTE

Dr. Iván Castillo Zúñiga
Director de tesis

EDUCACION A DISTANCIA

UNIVERSIDAD CUAUHTÉMOC
VANGUARDIA Y EXCELENCIA

Plantel Aguascalientes

Acuerdo No. 1857 del 8 de abril del 2015 del Instituto de Educación del Estado de Aguascalientes

Tesis.

Para Obtener el Grado de Maestría

en Ciencia de los Datos

y Procesamiento de Datos Masivos (BIG-DATA)

Título de la Tesis.

Sistema Experto para la Detección de Vocabulario de Ciberguerra en Sitios Web, utilizando Técnicas de Machine Learning y Deep Learning con Python.

Presenta:

Eliana Paola Pérez Arrieta.

Director:

Dr. Iván Castillo Zúñiga.

México-Colombia, agosto de 2020.

ASUNTO: Carta de autorización.

Aguascalientes, Ags., 16 de julio de 2020.

LIC. ROGELIO MARTÍNEZ BRIONES
UNIVERSIDAD CUAUHTÉMOC PLANTEL AGUASCALIENTES
RECTOR GENERAL

P R E S E N T E

Por medio de la presente, me permito informar a Usted que he asesorado y revisado el trabajo de tesis titulado:

“ SISTEMA EXPERTO PARA LA DETECCIÓN DE VOCABULARIO DE CIBERGUERRA EN SITIOS WEB, UTILIZANDO TÉCNICAS DE MACHINE LEARNING Y DEEP LEARNING CON PYTHON ”

Elaborado por la Licenciada en Mátemáticas **ELIANA PAOLA PÉREZ ARRIETA**, considerando que cubre los requisitos para poder ser presentado como trabajo recepcional para obtener el grado de Maestra en Ciencia de los Datos y Procesamiento de Datos Masivos (BIG-DATA).

Agradeciendo de antemano la atención que se sirva a dar la presente, quedo a sus apreciables órdenes.

ATENTAMENTE

Dr. Iván Castillo Zúñiga
Director de tesis

Índice.

Resumen xv
Abstract xvi
Agradecimiento xvii
Dedicatoria xviii
Introducción xix
Capítulo I. Introducción 1
1.1. Planteamiento del Problema 2
1.1.1. Contextualización. 2
1.1.2. Definición del problema. 5
1.1.3. Preguntas de Investigación. 6
1.2. Justificación 7
1.2.1. Conveniencia. 7
1.2.2. Relevancia social en Colombia. 8
1.2.3. Implicaciones educativas. 9
1.2.4. Relevancia teórica. 9
1.2.5. Utilidad metodológica. 10
1.2.6. Viabilidad. 10
1.3. Objetivos 10
1.3.1. Objetivo general. 10
1.3.2. Objetivos específicos. 11
1.4. Hipótesis 11
1.5. Descripción de la organización de la Tesis 12
Capítulo II. Estado del Arte 14
2.1. Ideas, Procedimientos y Teorías Relacionadas con la Detección de Vocabulario de Ciberguerra 15
2.1.1. Ciberguerra. 15
2.1.2. Web Scraper. 17
2.1.3. Analítica de Big Data. 20
2.1.4. Procesamiento de Lenguaje Natural. 23

2.1.5. Web Semántica. 25
2.1.6. Metadatos. 27
2.1.7. Ontologías. 28
2.1.8. Minería Web. 29
2.1.9 Paradigma metodológico de Minería de texto. 30
2.1.10. Aprendizaje de Máquina. 32
2.2. Descripción de los trabajos relacionados. 38
2.2.1. Aplicación de técnicas de Machine Learning a la detección de ataques. 38
2.2.2. Machine Learning aplicado a la seguridad. 39
2.2.3. Detección de intrusiones en sistemas web utilizando técnicas de aprendizaje profundo. 40
2.2.4. Ciberguerra y su efecto en la seguridad nacional. 41
2.2.5. La ciberguerra como realidad posible contemplada desde la prospectiva. 41
2.2.6. Ciberguerra... ¿Dudáis? 42
2.2.7. Aplicación de técnicas de minería de datos en Ciberseguridad y Ciberdefensa - una breve revisión. 43
2.3. Análisis de trabajos relacionados 43
2.4. Variables 44
Capítulo III. Materiales y Métodos 46
3.1. Materiales (Hardware y software utilizado en la investigación). 47
3.1.1. Hardware utilizado en la investigación. 47
3.1.2. Software utilizado para las pruebas. 47
3.2. Método. 47
3.2.1. Búsqueda de datos (Etapa 1). 48
3.2.2. Preprocesamiento de datos (Etapa 2). 49
3.2.3. Machine Learning (Etapa 3). 50
3.2.4. Optimización de los algoritmos de Machine Learning (Etapa 4). 50
3.3. Conjunto de datos de pruebas, población y muestra. 50
3.4. Algoritmos de Aprendizaje Supervisado desarrollados en Python para las pruebas 51
3.4.1. Árboles de decisión. 51
3.4.2. Naïve Bayes. 54

3.4.3. El K-ésimo vecino más cercano. 56
3.4.4. Redes neuronales. 58
3.4.5. Máquinas de Soporte Vectorial. 62
3.4.6. Ada Boosting. 64
3.4.7. Regresión Logística. 67
3.4.8. Random Forest. 68
3.4.9. Aprendizaje Profundo (Deep Learning). 69
3.5. Procedimiento de Pruebas 72
3.5.1. Librerías. 74
3.5.2. Cargue del Dataset y exploración de datos. 75
3.5.3. Definición de los conjuntos de entrenamiento y prueba. 77
3.6. Descripción de las pruebas con los algoritmos seleccionados 78
3.6.1. Árboles de Decisión. 79
3.6.2. Naïve Bayes. 80
3.6.3. El K-Vecino más Cercano – Knn. 82
3.6.4. Redes Neuronales. 83
3.6.5. Máquinas de Soporte Vectorial. 85
3.6.6. Ada Boosting. 87
3.6.7. Regresión Logística. 89
3.6.8. Random Forest. 91
3.6.9. Aprendizaje profundo con Redes Neuronales. 93
3.7. Consideraciones éticas. 99
Capítulo IV. Resultados 100
4.1. Resultados 101
4.1.1. Árboles de decisión. 101
4.1.2. Naïve Bayes. 102
4.1.3. El K-vecino más cercano – Knn 103
4.1.4. Redes Neuronales. 104
4.1.5. Máquinas de Soporte Vectorial. 105
4.1.6. Ada Boosting. 106
4.1.7. Regresión Logística. 107

4.1.8. Random Forest. 108
4.1.9. Aprendizaje profundo con Redes Neuronales. 109
4.2. Explicación de resultados 110
Capítulo V. Discusión 114
5.1. Análisis y comparación entre los algoritmos utilizados en la investigación. 115
5.2. Discusión con trabajos relacionados. 117
5.3. Discusión de los objetivos. 121
5.4. Respuesta a la hipótesis 123
Capítulo VI. Conclusiones 124
6.1. Conclusiones y trabajos futuros. 125
6.1.1. Conclusiones. 125
6.1.2. Fortalezas y debilidades. 125
6.1.3. Trabajos futuros (Generación de nuevas investigaciones). 126
Referencias 127
Anexos 135
Anexo A. Páginas Web utilizadas en el análisis de Ciberguerra……..….136

Índice de figuras.

Figura 1. Web Scraping 19
Figura 2. Arquitectura de un Sistema de Procesamiento de Lenguaje Natural 24
Figura 3. Evolución y predicción de la Web. 25
Figura 4. Diagrama de división en sub-conjuntos de los datos 34
Figura 5. Ejemplo de algoritmos Clustering. 36
Figura 6. Ejemplos de algoritmos de reducción de la dimensión 36
Figura 7. Análisis de variables de la investigación. 45
Figura 8. Diagrama del método de solución. 48
Figura 9. Pseudocódigo del algoritmo kDB 56
Figura 10. Ejemplo de clasificación de vecino más cercano basado en el peso uniforme y la distancia 58
Figura 11. Representación de una neurona biológica 59
Figura 12. Representación de una neurona artificial 60
Figura 13. MLP de una capa oculta con salida escalar. 62
Figura 14. Ejemplo de clasificación de dos clases con vectores de soporte. 63
Figura 15. Ejemplo del truco del Kernel 63
Figura 16. Diagrama de flujo del algoritmo del modelo Adaboost. 66
Figura 17. Diagrama de flujo del algoritmo Random Forest. 69
Figura 18. Esquema de capas de las Redes Neuronales. 71
Figura 19. Dataset original con las páginas Web a analizar. 73
Figura 20. Dataset con valores de variables predictoras en formato CSV. 73
Figura 21. Librerías de Machine Learning en Python. 74
Figura 22. Código para cargue del Dataset y código de exploración de datos. 75
Figura 23. Vista general del Dataset. 76
Figura 24. Resumen de los datos y variables. 77
Figura 25. Definición de los conjuntos para entrenamiento y pruebas. 78
Figura 26. Importe de la librería DecisionTreeClassifier. 79
Figura 27. Modelo Arboles de Decisión. 80
Figura 28. Importe de la librería GaussianNB. 80
Figura 29. Modelo Naïve Bayes. 81

Figura 30. Importe de la librería KneighborsClassifier. 82
Figura 31. Modelo Knn - vecino más cercano 83
Figura 32.Importe de la librería MLPClassifier. 84
Figura 33. Modelo Redes Neuronales. 85
Figura 34. Importe de la librería SVC 86
Figura 35. Modelo Máquinas de Soporte Vectorial. 87
Figura 36. Importe de librerías para desarrollar Ada Boosting. 88
Figura 37. Modelo Ada Boosting. 89
Figura 38. Importe de la librería LogisticRegression 90
Figura 39. Modelo de Regresión Logística. 91
Figura 40. Importe de la librería RandomForestClassifier. 92
Figura 41. Modelo Random Forest. 93
Figura 42. Importe de librerías para Deep Learning. 94
Figura 43. Modelo Redes Neuronales con Deep Learning 95
Figura 44. Parte del algoritmo inicial que sería intervenida para mejoras en el rendimiento. 96
Figura 45. Resultados del algoritmo de Redes Neuronales con Deep Learning en su primera versión. 96
Figura 46. Evidencia de presiones binarias para iteraciones entre la 46 y la 55 de 1000 97
Figura 47. Resultados arrojados por el modelo Árboles de Decisión. 102
Figura 48. Resultados Arrojados por el modelo Naïve Bayes. 103
Figura 49. Resultados arrojados por el modelo del K-vecino más cercano. 104
Figura 50. Resultados arrojados por el modelo Redes Neuronales. 105
Figura 51. Resultados arrojados por el modelo Máquinas de Soporte Vectorial. 106
Figura 52. Resultados arrojados por el modelo Ada Boosting. 107
Figura 53. Modelo de Regresión Logística. 108
Figura 54. Resultados del modelo Random Forest. 109
Figura 55. Resultados del modelo Aprendizaje profundo con Redes Neuronales 110
Figura 56. Gráfica de precisión en la detección de vocablos de Ciberguerra en sitios Web. 113

Índice de tablas.

Tabla 1. Comparativo de características de trabajos relacionados para la detección de ciberguerra. 44

Tabla 2. Variables organizadas por grupos. 45

Tabla 3. Evidencia de la toma de muestras para validación del algoritmo de Redes Neuronales con Deep Learning 98

Tabla 4. Resultados de precisión de los Algoritmos de Aprendizaje Supervisado realizados en Python. 112

Tabla 5. Resumen de resultados de los modelos implementados. 115

Tabla 6. Comparativa de resultados de la investigación de Rodríguez (2018) y la presente investigación. 118

Tabla 7. Comparativa de resultados de la investigación de Cuenca (2019) con la presente investigación. 119

Glosario de términos.

Adaboost.	Es una técnica de Boosting conocido como algoritmo Boosting Adaptativo, que mediante un entrenamiento iterativo de los clasificadores débiles o de base, les asigna mayor importancia a los datos mal clasificados anteriormente, y de esta manera obtiene un nuevo clasificador.
Algoritmo.	Conjunto ordenado de funciones que permite realizar cálculos y encontrar la solución de un problema específico.
Aprendizaje de Máquina.	Es un área de las ciencias de la computación cuyo objetivo es desarrollar técnicas que permitan aprender a los computadores.
Aprendizaje Profundo.	Tambien llamado Deep Learning, es un subconjunto de los algoritmos usados en Machine Learning, y se caracterizan por tener una arquitectura en capas donde cada capa aprende patrones más complejos según la profundidad de esta.
Aprendizaje Supervisado.	El aprendizaje supervisado es un tipo de algoritmo de Machine Learning que emplea un conjunto de datos conocidos (el denominado conjunto de datos de entrenamiento) para realizar predicciones.
Arboles de Decisión.	Es un modelo esquemático de las alternativas disponibles y de las posibles consecuencias de cada una, conformado por múltiples nodos cuadrados que representan puntos de decisión y de los cuales surgen ramas que representan las distintas alternativas, las ramas

que salen de los nodos circulares, o causales, representan los eventos.

Big Data. Análisis masivo de datos, una cuantía de datos, tan sumamente grande, que las aplicaciones de software de procesamiento de datos que tradicionalmente se venían usando no son capaces de capturar, tratar y poner en valor en un tiempo razonable. Igualmente, el mismo término se refiere a las nuevas tecnologías que hacen posible el almacenamiento y procesamiento y uso que se hace de la información obtenida a través de dichas tecnologías.

Big Data Analytics. Es la unión de las técnicas de la analítica con las técnicas de Big Data. Es un concepto que agrupa las tecnologías y desarrollos matemáticos que se dedican a almacenar, analizar y cruzar toda esa información para intentar encontrar patrones de comportamiento.

Bosques Aleatorios. Es una técnica de aprendizaje automático, basada en un conjunto de árboles de decisión combinados con bagging.

C4.5. Es un algoritmo usado para generar un árbol de decision, el cual puede ser usado para la clasificación, y por esta razón, C4.5 está casi siempre referido como un clasificador estadístico.

Ciberdefensa. Es el conjunto de acciones y operaciones desarrolladas en el ámbito informático y telemático de la defensa a fin de asegurar el cumplimiento de las misiones o servicios para los que fueron concebidos e impide que fuerzas enemigas los utilicen para cumplir los suyos.

Ciberguerra. Conflictos que se desarrollan en torno al Internet a través de ataques informáticos y a la información en sí misma y que involucran distintos intereses, generalmente políticos.

Ciberseguridad. Se conoce como la seguridad de la tecnología de la información, puesto que engloba un gran número de técnicas y métodos para proteger nuestro sistema, así como otros dispositivos o las redes.

Conpes. Consejo Nacional de Política Económica y Social. Es para Colombia, la máxima autoridad nacional de planeación y se desempeña como organismo asesor del Gobierno en todos los aspectos relacionados con el desarrollo económico y social del país.

Corpus Lingüístico. Conjunto amplio y estructurado de ejemplos reales de uso de la lengua.

CSV. Comma-Separated Values.

Dataset. Término en inglés para un conjunto de datos o colección de datos habitualmente tabulada.

HTML. Es un lenguaje de marcado de hipertexto o "HyperText Markup Language", que básicamente se escribe en su totalidad con elementos constituidos por etiquetas, contenido y atributos.

HTTP. HyperText Transfer Protocol o Protocolo de Transferencia de Hiper Textos, es el protocolo de transmisión de información de la World Wide Web.

Infraestructuras críticas.	Instalaciones, redes, servicios y equipos físicos y de tecnología de la información cuya interrupción o destrucción tendría un impacto mayor en la salud, la seguridad o el bienestar económico de los ciudadanos o en el eficaz funcionamiento de las instituciones del Estado y de las Administraciones Públicas.
Inteligencia Artificial.	Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico, incluye tareas como la planificación, el reconocimiento de objetos y sonidos, hablar, traducir, realizar actividades creativas, etc.
K-nn.	El K-nn es un algoritmo de aprendizaje supervisado, que a partir de un juego de datos inicial su objetivo será el de clasificar correctamente todas las instancias nuevas. El juego de datos típico de este tipo de algoritmos está formado por varios atributos descriptivos y un solo atributo objetivo.
Metadatos.	Datos que describen otros datos o "datos sobre datos".
Minería de datos.	Es una tecnología emergente cuyo objeto es la búsqueda de conocimiento en grandes colecciones de documentos no estructurados.
Naïve Bayes.	Son una clase especial de algoritmos de clasificación de Aprendizaje Automático, que se basan en una técnica de clasificación estadística llamada “teorema de Bayes”.

Ontología.	Especificación formal de conceptos y relaciones entre los mismos de forma jerarquizada basada en el conocimiento, restringidos por axiomas.
OTAN.	Son las siglas de la Organización del Tratado del Atlántico Norte. Es una alianza político-militar creada durante la Guerra Fría en 1949 y se fundamenta en el mutuo apoyo militar a los países miembros en caso de agresión por parte de terceros.
Procesamiento de lenguaje natural.	Definido como la manipulación automática del lenguaje natural, como el habla y el texto, por el software.
Python.	Es un lenguaje de scripting independiente de plataforma y orientado a objetos, preparado para realizar cualquier tipo de programa, desde aplicaciones Windows a servidores de red o incluso, páginas Web.
Regresión logística.	Son modelos de regresión que permiten estudiar si una variable binomial depende, o no, de otra u otras variables.
MSV.	Maquinas de Soporte Vectorial. Es un algoritmo de aprendizaje supervisado que se puede emplear para clasificación binaria o regresión, usadas principalmente en aplicaciones como el procesamiento del lenguaje natural, el habla, el reconocimiento de imágenes y la visión artificial.
TIC.	Tecnologías de Información y Comunicaciones.

Web Crawler. Es un algoritmo utilizado para analizar el código de páginas Web, generando insigths importantes para las estrategias digitales.

Web Scraper. Es una técnica que sirve para extraer información de páginas Web de forma automatizada.

Web semántica. Conjunto de actividades desarrolladas en el seno del entorno World Wide Web (WWW), para crear tecnologías de publicación de datos que sean más fácilmente legibles para las aplicaciones informáticas.

Weka. Es una plataforma de software para el aprendizaje automático y la minería de datos escrito en Java.

Resumen.

El crecimiento a gran escala del Internet con su multiplicidad de servicios y beneficios, ha traído consigo grandes retos por afrontar, entre ellos se encuentran algunos relacionados con la seguridad. Problemática en la que se enfoca este trabajo cuyo objetivo primordial es el desarrollo de un sistema experto que permita clasificar vocabulario relacionado con Ciberguerra en páginas Web, utilizando técnicas de aprendizaje supervisado y aprendizaje profundo, identificando los algoritmos adecuados para generar conocimiento y valor agregado. Para las pruebas, se construyó un Dataset que representa los distintos sitios Web en español tocantes al tema de Ciberguerra basado en un corpus lingüístico. Posteriormente, se implementaron distintos modelos predictivos en Python para su evaluación, con los algoritmos Knn, Naïve Bayes, Arboles de Decisión, Bosques Aleatorios, SVM, Regresión Logística, Redes Neuronales, Redes Neuronales con Deep Learning y Adaboost.

Entre los productos conseguidos, cabe resaltar que todos los algoritmos mencionados anteriormente, tuvieron una precisión mayor al 97.7%, considerando a nivel general un resultado de detección de vocabulario de Ciberguerra alto, sustentado con la combinación de técnicas de Big Data Analytics, Procesamiento de Lenguaje Natural, Web Semántica, Aprendizaje de Máquina y Aprendizaje Profundo, logrando una correcta clasificación de los sitios Web de Ciberguerra y proporcionando alternativas de solución a esta problemática. Destacando la efectividad de Naïve Bayes y Ada Boosting con 99.5%.

Palabras Clave: Ciberguerra, Aprendizaje de Máquina, Aprendizaje Supervisado, Aprendizaje Profundo, Grandes Volúmenes de Datos, Minería Web, Corpus Lingüístico, Rastreador Web.

Abstract.

The large-scale growth of the Internet with its multiplicity of services and benefits has brought with it great challenges to be faced, among them some related to security. Problems on which this work is focused, whose main objective is based on the development of an expert system that allows classifying vocabulary related to Cyberwar on Web pages, using supervised learning and deep learning techniques, identifying the appropriate algorithms to generate knowledge and added value. For testing, a Dataset portraying different websites in Spanish that deal with Cyberwar in a linguistic corpus was built. Later on, several Python-based predictive models were implemented for further assessment through Knn, Naïve Bayes, Decision Trees, Random Forests, SVM, Logistic Regression, Neural Networks, Neural Networks with Deep Learning and AdaBoost.

Among the obtained products, it is worth highlighting that all the aforementioned algorithms had an accuracy rate higher than 97.7% which can be regarded, in overall, as a high rate detection of Cyberwar vocabulary. Such result is supported on the combination of Big Data Analytics, Natural Language Processing, Semantic Web, Machine Learning and Deep Learning techniques which resulted on the correct classification of Cyberwar websites as well as on providing choices to solve this issue. highlighting Naïve Bayes and Ada Boosting with 99.5% effectiveness.

Key Words: Cyberwarfare, Machine Learning, Supervised Learning, Deep Learning, Big Data, Web Mining, Linguistic Corpus, Web Crawler.

Agradecimiento.

Agradezco a la Universidad Cuauhtémoc de México por la organización del programa de Maestría y haberme permitido ser parte de él.

Agradezco a todos los profesores que hicieron parte de mi proceso de aprendizaje, por acompañarme en todas las etapas de formación de esta parte fundamental de mi proyecto de vida. De manera muy especial expreso mi agradecimiento a mi director de tesis, el Dr. Iván Castillo Zúñiga, por permitirme acceder a su valioso conocimiento científico y que con toda su paciencia y dedicación me ayudó a superar las dificultades a lo largo del trabajo.

Expreso mi más infinito agradecimiento a mis padres, Tulia y Manuel, por ser los principales forjadores de mi espíritu de entrega por los sueños y al resto de mi familia, Tibizay, José D, Roberto, Rodrigo, Gustavo y James por creer en mí, por leerme, opinar y por su apoyo incondicional en distintas formas durante todo el proceso, y en especial a la abuela Luz por llenarnos de alegría día tras día.

Gratitud de manera especial para mis compañeros de maestría, en especial a Jonathan, Felipe, Víctor y Daniel con quienes pude fortalecer lazos de compañerismo, amistad y apoyo mutuo y especialmente motivación para continuar dando lo mejor de mi durante todo el transcurso de la maestría.

Finalmente agradezco a mis a amigos Ana, Margarita, Carolina, Eny, Manira, Arturo P, Arturo C y Daimer por acompañarme en cualquier manera en todos mis caminos.

Dedicatoria.

A la vida por darme la oportunidad
de adquirir nuevos conocimientos
para aportar a la sociedad.

A mi familia y amigos por motivarme a
seguir adelante.

A los cuatro abuelos por darme
la oportunidad de estar aquí.

Introducción.

Las Tecnologías de Información y Comunicaciones (TICs), han desarrollado con gran potencialidad a través de los años, la capacidad de conectar millones de redes y hacer funcionar estructuras y servicios esenciales, convirtiéndose en un gran aporte para la sociedad y a la vez trayendo múltiples retos que afrontar. Algunos de los cuales están ligados con la seguridad, ya que esta creciente dependencia de la tecnología, de los estados, empresas y personas, ha generado del ciberespacio un lugar propicio para la guerra, facilitando el uso de artimañas informáticas para robar información sensible, bloquear o dañar sus sistemas informáticos o atacar las infraestructuras de otros países. En la actualidad, ya se han presentado ataques aislados en distintos países, que han alertado a los organismos relacionados con la seguridad, y esta situación ha conllevado a los gobiernos a preparar sus herramientas informáticas, para defenderse y atacar cuando fuese necesario.

El problema que se plantea en la presente investigación está relacionado con la detección de vocabulario de Ciberguerra a partir de datos obtenidos en sitios Web, a través del análisis, clasificación y predicción con herramientas de Aprendizaje Supervisado, y aunque ya existen una gran cantidad de investigaciones alrededor de la detección de vocabulario, la relación con el tema de Ciberguerra le da gran relevancia, teniendo en cuenta que es un tema de importancia mundial.

Generar conocimiento a partir de la información ha sido una tarea que siempre ha desarrollado la humanidad y que se ha venido perfeccionando con la evolución de la Web y los avances de la informática. Estos se han usado en beneficio de gobiernos, empresas y otras entidades que analizan las tendencias de gustos, comportamientos, opiniones de millones de datos en la Web de manera rápida y efectiva. De lo anterior nace la idea de esta investigación,

la cual pretende generar conocimiento en torno a los datos que circulan en Internet alrededor de lo que se denomina Ciberguerra, convirtiéndose en una base importante para la investigación alrededor de esta temática.

En este sentido se propone entonces, desarrollar un sistema capaz de detectar léxico de Ciberguerra dentro del contenido de Internet, usando técnicas de Procesamiento de Lenguaje Natural, Web Semántica, Aprendizaje Supervisado, Machine Learning y Deep Learning, y de esta manera poder identificar entre los algoritmos implementados, cuál o cuáles son más efectivos para generar valor agregado sobre las páginas que circulan en Internet con relación a la problemática mencionada.

El trabajo investigativo se estructuró en etapas organizadas de manera secuenciada que permitieran cumplir los objetivos propuestos. Dando lugar primero, a la búsqueda de información en la Web para fundamentar con bases y recursos solidos el estudio y las investigaciones afines en algún aspecto. Posteriormente, a través de un Crawler se realizó la extracción de las páginas Web que servirían de base para la construcción de un Dataset. Seguido del preprocesamiento y tratamiento de los datos con herramientas de minería de texto, utilizando los elementos que se proponen en la herramienta ADVI. Una vez estructurados los datos del Dataset en formato CSV, se realizó la implementación con las herramientas seleccionadas y con cada uno de los modelos de aprendizaje supervisado y aprendizaje profundo escogidos para la presente investigación.

Con la base que aportaron los fundamentos de trabajos relacionados y demás trabajos incluidos en el estado del arte, se seleccionaron los modelos, Knn – Vecino más cercano, Naïve Bayes, Arboles de Decisión, Bosques Aleatorios, SVM, Regresión Logística, Redes

Neuronales, Redes Neuronales con Deep Learning y Adaboost, siendo considerados posiblemente los más apropiados para detección de vocabulario de Ciberguerra.

Finalmente se muestran los resultados de todos los modelos implementados, y a través de una discusión se comparan los resultados arrojados por cada uno de los modelos. De la misma manera, se comparan con los trabajos relacionados y estado del arte, y se establecen las conclusiones a partir de los hallazgos encontrados y se exponen las posibles líneas de investigación para futuras investigaciones.

Capítulo I. Introducción.

1.1. Planteamiento del problema.

1.1.1. Contextualización.

Las Tecnologías de Información y Comunicaciones (TICs), han venido a través de los años tomando una posición muy importante en distintos aspectos sociales, trayendo consigo un sin número de cambios y retos permanentes para nuestra sociedad. En la década de los 90, el mundo entero disfrutaba de una comunicación instantánea a bajo costo, disponibilidad de información, mayor visualización y otras ventajas que fueron creciendo rápidamente, generando así un crecimiento a gran escala del uso del mismo, como lo plantea Evans (2011), en su informe, quien pronóstica que el número de aparatos conectados a Internet por persona pasaría de 0.08 en 2013 a 6.58 en 2020, es decir, el Internet ha estado pasando de ser usado por personas en aparatos tecnológicos a controlar redes eléctricas, hospitales, cámaras de seguridad, transacciones y un gran número de actividades fundamentales para la organización y estabilidad de aspectos vitales de la economía, la educación, el arte, la política, la salud, la seguridad, la defensa, entre otros. Este uso creciente del Internet, así como ha generado grandes beneficios sociales, ha traído consigo también actividades ilícitas llevadas a cabo por ciberdelincuentes que con distintos propósitos aprovechan las vulnerabilidades del sistema. Cuando estas actividades delincuenciales tienen como finalidad atacar un país o un gobierno, robando información sensible, alterando o degradando los sistemas informáticos, entre otros ataques a través del ciberespacio pasa a denominarse Ciberguerra.

La Ciberguerra, es un tema de gran relevancia en la actualidad teniendo en cuenta, que está relacionada directamente la seguridad y afecta aspectos económicos, políticos, entre

otros, que podrían conllevar a una inestabilidad social. La Ciberguerra "se puede definir como el conjunto de acciones que se realizan para producir alteraciones en la información y en los sistemas del enemigo, a la vez que se protege la información y los sistemas del atacante" (Ferrero, 2013, p.86). Sin embargo, a pesar de que su naturaleza virtual y sus efectos podrían hacerse sentir en forma drástica en la sociedad. Torres (2011) afirma: "los actos de Ciberguerra pueden provocar unos daños que exceden las posibilidades de la mayoría de las armas convencionales" (p. 15).

Aunque hasta la fecha no se ha producido un ataque de Ciberguerra de gran impacto, desde hace aproximadamente 30 años vienen presentándose aisladamente variedad de ataques en diferentes países, difícilmente atribuibles, que inducen a este concepto. Entre los ataques más representativos cabe mencionar que Estonia en 2007, sufrió un ataque a sus sistemas de información y comunicaciones que es considerado uno de los mayores ataques de la historia hasta la fecha, ya que generó una crisis que necesitó intervención internacional Conpes (2011). Por otro lado, Lejarza (2014), afirma que en Norte América "durante dos años, hackers saquearon datos procedentes de los ordenadores del Pentágono, la NASA, Universidades y Centros de Investigación Norteamericanos, así como del Departamento de Energía, accediendo a miles de documentos de relevancia nacional." (p.8)., las investigaciones alrededor de este hecho apuntaron a que los ataques eran dirigidos por Rusia, aunque el Gobierno Ruso negó su participación en esta actividad. "Los ciberataques continuaron evolucionando rápidamente en frecuencia y complejidad, como demostraron los casos de Lituania y Georgia en julio de 2008 y el ciberataque a Kirguistán en enero de 2009". (Bejarano, 2011, p.2)

Otro ataque sucedió en 2010 en Irán, el cual, a través de un virus informático conocido actualmente como Stuxnet, dañó alrededor de 1000 centrifugadoras usadas para enriquecer uranio, cuyo presunto objetivo "era destruir las centrifugadoras nucleares iraníes de Natanz, evitando así la fabricación de armamento atómico por parte de este país" (Lejarza, 2014, p.9).

Frente a este panorama en el que el uso creciente y desmedido de las tecnologías digitales inserta en una situación de riesgo a los Gobiernos, se han generado alarmas a nivel internacional. En junio de 2011 la OTAN aprobó una Revisión de la Política de Ciberdefensa y un Plan de Acción de Ciberdefensa, propuesto a partir del Concepto Estratégico de la Alianza desde la Cumbre de Lisboa en 2010, luego de que en 2008 esta misma organización hiciera pública su Política de Ciberdefensa con el propósito de proteger los sistemas de información y comunicaciones.

Colombia, no está fuera de este escenario, teniendo en cuenta que también ha incrementado considerablemente el uso de las tecnologías de información y comunicaciones en diversidad de ámbitos en los últimos años. Conpes (2011) señala que, "el número de usuarios de Internet aumentó en 354% entre el 2005 y el 2009. El número de suscriptores a Internet se incrementó en 101% entre el 2008 y el 2010 alcanzando un total de 4384181 suscriptores de Internet fijo y móvil" (p.7). Este incremento en el uso de las nuevas tecnologías en el país, implica también incrementó en la dependencia de los mismos y por tanto el aumento de vulnerabilidad ante ataques informáticos, de los cuales ya ha venido siendo víctima, un ataque a resaltar, se dio en el primer semestre de 2011 cuando un grupo en protesta ante el Proyecto de Ley, "por el cual se regula la responsabilidad por las infracciones al derecho de autor y los derechos conexos en Internet" atacó a los portales de

la Presidencia de la República, el Senado de la República, Gobierno en línea y algunos Ministerios, dejándolos fuera de servicio por varias horas (Conpes, 2011).

1.1.2. Definición del problema.

La actualidad está marcada por la creciente demanda del uso de información que se ha venido dando a raíz del manejo de las nuevas tecnologías, este hecho ha favorecido el desarrollo del Aprendizaje de Máquina (del inglés Machine Learning), con gran potencialidad para variedad de campos. Una de las aplicaciones de gran relevancia es el análisis de datos en Internet gracias al aporte que ha venido generando en la toma de decisiones en distintos ámbitos, como el análisis de sentimientos en las redes sociales, detección y clasificación de vocabulario en los distintos Sitios Web. Además, la literatura relacionada con la aplicación de técnicas de Aprendizaje Supervisado, para la detección de los distintos vocablos generalmente están relacionados con temáticas de gran impacto social, como en este caso lo es la Ciberguerra.

En esta investigación se aborda la problemática relacionada con la necesidad de dar un buen uso al creciente volumen de información en Internet, haciendo uso de técnicas de la analítica de Big Data, para lo cual se propone analizar, clasificar y detectar el vocabulario de Ciberguerra a partir de datos extraídos de páginas de Internet.

Problemática de manera específica:

1. Dado que la información obtenida de páginas Web está mezclada con código HTML y PHP, y gran parte de estas hacen uso de fuentes de datos asíncronas, se presentan dificultades para transformarla a datos estructurados.

2. Teniendo en cuenta que la información extraída de millones de páginas Web no está estructurada y vienen en distintos formatos, se presentan inconvenientes para realizar un adecuado tratamiento de datos que permita optimizar el proceso.
3. Se encuentran obstáculos en la toma de decisiones, con base en la información recolectada, debido a la complejidad de las técnicas de aprendizaje supervisado que son implementadas.
4. Problemas al clasificar las páginas Web, dada la gran cantidad de variables que intervienen en dicha clasificación y la compleja implementación de algoritmos de Aprendizaje de Máquina (Machine Learning).
5. Dificultad para detectar el vocabulario de Ciberguerra.

1.1.3. Preguntas de investigación.

1. ¿Cuáles son las principales dificultades y qué retos deben asumirse en torno a la literatura en la construcción de vocabularios u ontologías para la clasificación de páginas Web?
2. ¿Cuál de las técnicas de Machine Learning usadas en la investigación, es la más apropiada para detectar vocabulario relacionado a la temática de Ciberguerra?
3. ¿Cuál es la efectividad de las técnicas de Big Data Analytics en la predicción de vocabulario relacionado con Ciberguerra?
4. ¿Cuáles son las principales dificultades de las técnicas de Aprendizaje de Máquina, para detectar vocabulario relacionado con Ciberguerra?
5. ¿De qué manera contribuye el análisis, la clasificación y la predicción de datos de páginas de Internet sobre Ciberguerra en la investigación alrededor de esta problemática?

1.2. Justificación.

El ritmo de vida de la actual sociedad no sólo genera un volumen de datos inmenso a gran velocidad, sino que además lo produce en una variedad tan grande, que el análisis de datos tradicional queda sin herramientas suficientes para hacer análisis que permitan sacarle provecho en la toma de decisiones fundamentadas en distintos campos. Lo anterior ha favorecido el desarrollo de herramientas de Machine Learning, con las cuales se aborda la necesidad de poder darle significado a esos datos.

La Ciberguerra por su parte, es una problemática de interés mundial, y teniendo en cuenta que esta problemática tiene un gran impacto social, la posibilidad de explotar estos grandes volúmenes de datos en pro de favorecer los estudios alrededor de la Ciberguerra, genera una gran motivación para el desarrollo del mismo.

Se pretende parametrizar a través de técnicas de extracción de datos, miles de sitios Web y utilizar distintas técnicas de aprendizaje supervisado para detectar léxicos que determinen y permitan hacer una clasificación de múltiples variables de todo lo relacionado directamente con el tema, con el fin de poder descubrir esta información y facilitar a los interesados en investigar sobre el tema.

1.2.1. Conveniencia.

La Ciberguerra, es un concepto que está creciendo como una amenaza a la seguridad de los estados, de los ciudadanos, y de todo el sistema en general. Este es uno de los principales motivos por los que los gobiernos actualmente están invirtiendo en mayor medida en la defensa de ataques y detección de vulnerabilidades en sus sistemas informáticos. La mayor conveniencia de esta propuesta radica en que el análisis, clasificación y predicción de vocabulario en torno a Ciberguerra, a partir de la utilización de herramientas de Machine

Learning, puede convertirse en una herramienta que facilite la comprensión y manejo de esta problemática.

1.2.2. Relevancia social en Colombia.

Colombia ante la defensa de la vida y la seguridad ha venido haciendo frente a esta problemática, estableciendo políticas públicas con miras a fortalecer su ciberdefensa. (Cortés, 2015), plantea que en el Decreto 1512 de 2000, art. 5, se le invita al Ministerio de Defensa Nacional a "concretar los lineamientos que materialicen la ciberseguridad y la ciberdefensa en el estado Colombiano, en conjunto con las fuerzas militares y de policía, así como los organismos de seguridad e inteligencia".

En 2011 con el documento Conpes 3701 se establecieron lineamientos para tomar medidas en cuanto a la seguridad digital, cuyo objetivo general es "el fortalecimiento de la capacidad del Estado para enfrentar las amenazas que atentan contra su seguridad y defensa en el ámbito cibernético" (Conpes, 2011, p.5) y en 2016 con el documento Conpes 3854 se complementa el objetivo con la gestión de riesgos.

Además, se resalta que Colombia tuvo su representación en octubre de 2018 en una reunión con el Secretario General de la OTAN en Bruselas, en la que la Presidencia (2018) refiere que se tuvo como objetivo principal fortalecer la capacidad del estado para minimizar el nivel de riesgo de sus ciudadanos y las entidades públicas y privadas, ante amenazas o incidentes relacionados con el ciberespacio a través de una estrategia conjunta.

Todo lo anterior permite poner en evidencia la importancia creciente ante los temas de seguridad digital, y a su vez resaltar la importancia social que conlleva poder contar con

distintas herramientas que permitan la defensa ante ciberataques, como lo son las herramientas de análisis de vocabulario a partir de técnicas de Big Data.

1.2.3. Implicaciones educativas.

Con la realización de esta investigación se dejan evidencias de aplicación de las técnicas Aprendizaje Supervisado utilizadas para realizar análisis, clasificación y predicciones de vocabulario de Ciberguerra a partir de miles de páginas de Internet, lo cual puede aportar algunas pautas para futuros estudios y profundización en el área de análisis, clasificación y predicción de vocabularios.

1.2.4. Relevancia teórica.

Al ahondar en el estudio de temáticas relacionadas con Ciberguerra, se desprenden multiplicidad de aspectos sociales que hacen relevante la investigación. Por otra parte, los estudios con base técnicas de Machine Learning para predecir vocabulario, con el incremento del uso de las nuevas tecnologías, ha tomado gran relevancia para distintos ámbitos como la robótica, el marketing, el análisis de sentimientos, entre otros. La importancia teórica de esta investigación radica en la articulación de los temas de Ciberguerra, Big Data Analytics, vocabulario en Dataset y técnicas de Machine Learning.

De los principales aportes teóricos de esta investigación, se resaltan:

1. Dar a conocer y aplicar distintas técnicas de Machine Learning y Deep Learning en la detección de léxico.
2. Proponer nuevas herramientas para la detección de vocabulario relacionado con Ciberguerra, a partir de lo cual se generará conocimiento para distintas instituciones de tipo educativo, social y gubernamental.

1.2.5. Utilidad metodológica.

La realización de esta investigación coadyuva al diseño e implementación de otros instrumentos que permitan recolectar, analizar y clasificar vocabulario o léxico referente a Ciberguerra, u otras temáticas a partir de la combinación de distintas técnicas del Big Data Analytics, Web semántica y Procesamiento de Lenguaje Natural.

1.2.6. Viabilidad.

La viabilidad de este proyecto está dada por los bajos costos y la asesoría personalizada para realizar una investigación alrededor de un tema de investigación tan amplio y de tanta trascendencia social como lo es el de Ciberguerra.

Además, el dinamismo en las soluciones presentadas marca otro punto de viabilidad del proyecto, teniendo en cuenta que los datos en Internet varían de manera extremadamente rápida. La propuesta permite que sea aplicada a un grupo de datos diferente y se puedan conseguir distintas aplicaciones, estudios y resultados totalmente diferenciados en poco tiempo, o sobre el mismo grupo de datos hacer variaciones en la aplicación, métodos para obtener resultados con mayor favorabilidad para futuras investigaciones.

1.3. Objetivos.

1.3.1. Objetivo general.

Desarrollar un Sistema Experto para clasificar vocabulario relacionado con Ciberguerra dentro del contenido de páginas Web, a través de técnicas de Machine Learning y Deep Learning, identificando los algoritmos adecuados para generar conocimiento y valor agregado.

1.3.2. Objetivos específicos.

1. Construir un modelo predictivo para identificar si el vocabulario dentro de una página Web, está relacionado con la temática de Ciberguerra.
2. Obtener una base de datos de páginas Web relacionadas con la información de Ciberguerra que circula en la Web.
3. Establecer los elementos de la base de datos que serán utilizados para las pruebas de detección de Ciberguerra, mediante la transformación de datos semiestructurados a datos estructurados en formato .csv.
4. Seleccionar distintos algoritmos de aprendizaje supervisado que van a utilizarse para detección de vocabulario de Ciberguerra.
5. Implementar los algoritmos seleccionados para detectar vocabulario de Ciberguerra.
6. Comparar los resultados arrojados por los distintos algoritmos, para establecer cuál es el más adecuado para la detección de Ciberguerra.

1.4. Hipótesis.

El desarrollo de un sistema capaz de detectar vocabulario de Ciberguerra dentro del contenido de páginas Web, usando técnicas de Procesamiento de Lenguaje Natural, Web Semántica, Aprendizaje Supervisado, Machine Learning y Deep Learning, permitirá identificar cual o cuales son los algoritmos más efectivos para generar valor agregado sobre las páginas que circulan en Internet en relación con dicha problemática.

1.5. Descripción de la organización de la Tesis.

La tesis está organizada en seis capítulos y adicionalmente las referencias bibliográficas y anexos.

El capítulo 1, Introducción, tiene como objetivo proveer al lector del contexto de la investigación, resumiendo el problema de estudio atacado, las aportaciones previas en el tema investigado, definición de conceptos clave y todo aquello que contribuya al entorno mencionado al inicio del párrafo.

El capítulo 2, Estado del arte, se describen los antecedentes de la investigación, los trabajos relacionados con la problemática que aborda el proyecto actual. Se establecen las bases teóricas en las que está sustentado el trabajo de tesis, se enfoca en conceptos relacionados con Big Data, Web Semántica, Procesamiento de Lenguaje Natural, Pre-Procesamiento de Datos y Aprendizaje Automático. Por último, se aborda de manera puntual el tema de Ciberguerra, el cual es el caso de estudio en la investigación.

El capítulo 3, Materiales y métodos, presenta una descripción detallada de la propuesta de solución a la problemática planteada y a los objetivos establecidos al inicio de la investigación.

El capítulo 4, Resultados, se muestran y explican los productos obtenidos tras la evaluación/validación de la propuesta de solución desarrollada.

El capítulo 5, Discusión, se describen los resultados obtenidos y se hace un análisis sobre la efectividad de los métodos implementados, estableciendo cuales son los más adecuados para el reconocimiento de vocabulario de Ciberguerra. En el mismo sentido, se realiza un contraste de los resultados con los objetivos, trabajos relacionados e hipótesis.

El capítulo 6, Conclusiones, se puntualizan las ventajas y desventajas de la investigación, considerando la problemática y los objetivos establecidos al inicio del estudio. Además, se presentan las posibilidades de trabajo futuro.

Las referencias bibliográficas se muestran en formato APA en orden alfabético y para finalizar se incluyen los anexos.

Capítulo II. Estado del Arte.

2.1. Ideas, procedimientos y teorías relacionadas con la detección de vocabulario de Ciberguerra.

2.1.1. Ciberguerra.

El ciberespacio pasa a ser una nueva dimensión de guerra junto a las tradicionales; terrestre, marítima, aérea y espacial, y está siendo objeto de estudio por parte de organizaciones internacionales, empresas públicas y privadas, instituciones de educación superior, entre otras entidades que visualizan la importancia de prepararse para las implicaciones que tendría, que esos conflictos que se han venido produciendo a través de la informática en el Ciberespacio en una pequeña proporción aumenten su capacidad y se dé cabida al desarrollo de una Ciberguerra entendida como "una agresión promovida por un Estado y dirigida a dañar gravemente las capacidades de otro para tratar de imponerle la aceptación de un objetivo propio o, simplemente, para sustraerle información, cortar o destruir sus sistemas de comunicación" (Sánchez, 2018, p.225).

En este sentido, la Ciberguerra comparte las características principales de una guerra convencional en cuanto a los fines, pero el escenario en el que se produciría sería el ciberespacio, el cual posee unas características diferentes del resto de los demás espacios. En torno a lo cual Corredera (2012) de manera resumida establece las siguientes:

1. El ciberespacio es un entorno único, en el que el atacante puede estar en cualquier parte del globo.
2. En la defensa intervienen muchos factores, y no sólo elementos estatales sino también privados. Se exige pues una estrecha coordinación entre todos ellos.

3. La confrontación en el ciberespacio presenta frecuentemente las características de un conflicto asimétrico; y es frecuentemente anónimo y clandestino.
4. Permite obtener información sobre objetivos sin necesidad de destruir ni neutralizar ningún sistema, y a menudo, sin delatarse.
5. Permite también ejercer el chantaje; pero al mismo tiempo la defensa puede utilizarlo para la disuasión.
6. Evoluciona rápidamente siguiendo la evolución tecnológica de las TIC. (p.15)

En cuanto a las características de la Ciberguerra, Sánchez (2018), destaca la asimetría como una de las principales características de la misma, teniendo en cuenta que proporciona a los más pequeños herramientas necesarias para que puedan enfrentarse, vencer y mostrarse superiores a los más grandes. A lo cual Rosas (2014), refiere que la Ciberguerra se hace atractiva gracias a que no requiere de la existencia de un enorme ejército ni de grandes sistemas de armamento, en cambio, los recursos humanos debidamente calificados y una cierta infraestructura de TIC, con ciberguerreros muy versados sería suficiente para perpetrar ataques contra determinados objetivos, lo cual podría considerarse relativamente económico en comparación con los costos generados por portaaviones, buques, armamento pesado y fuerzas armadas, inevitables en una guerra convencional.

Desde otra perspectiva, la Ciberguerra se muestra como un fenómeno caracterizado como, “el ejercicio de emplear los computadores, la Internet y la dimensión ciberespacial por parte de un Estado con el objetivo de causar daños suntuarios sobre otro, mediante el desarrollo de ataques cibernéticos que van dirigidos hacia su infraestructura crítica” (Gaitán, 2013, p.5). Entendiendo las Infraestructuras Críticas como, “el conjunto de recursos, servicios, tecnologías de la información y redes, que en caso de sufrir un ataque causarían

gran impacto en la seguridad, tanto física como económica de los ciudadanos, o en el normal funcionamiento del gobierno" (Ferrero, 2013, p.84). Este concepto es importante dado que estás serían la entrada principal para la Ciberguerra a causar daños físicos por fuera del campo virtual, trascendiendo del daño a sistemas informáticos, espionaje o robo de información sensible hacia la afectación de sistemas ligados con la vida misma o aspectos muy cercanos a la calidad de esta.

No obstante, otros autores se muestran escépticos ante esta perspectiva, como Rid (como se citó en Gómez, 2017), quien alrededor de las interacciones que se dan en el ciberespacio, señala que: "no son ni serán guerras, ya que las ciberoperaciones simplemente no son lo suficientemente letales para generar bajas significativas, por lo que nunca reemplazarán la violencia y, por lo mismo, están lejos de ser un acto de guerra". Ante lo cual Gómez (2017), concluye que el ciberespacio es un lugar en que las manifestaciones de violencia pueden darse con el potencial de causar daños físicos, heridos y muertos, lo que se prevé irá en aumento.

2.1.2. Web Scraper.

Poder utilizar la inmensidad de información que circula en Internet de manera veloz y en gran variedad, ofrece gran cantidad de ventajas a empresas, gobiernos, instituciones, entre otros, para obtener beneficios propios, sociales o de otra índole. Web Scraper o Web scraping, también conocido como Web Data Extraction o Web Harvesting, es una de las oportunidades para explotar el potencial de la información en la Web, "es una técnica para extraer datos de la World Wide Web (WWW) y guardarlos en un archivo de sistema o base de datos para su posterior recuperación o análisis" (Zhao, 2017, p.1).

Castillo (2015), señala que el uso de Web Scraper, facilita la tarea de obtener información de las páginas Web utilizando herramientas y robots, ahorrando a las personas ese trabajo aburrido, repetitivo y propenso a errores de la recuperación manual de datos, ofreciendo también, datos en formatos amigables para su posterior procesamiento e integración como lo son JSON, XML, CSV, XLS o RSS.

Butler (como se citó en Zhao, 2017), plantea que las herramientas de Web scraper de última generación son capaces no sólo de analizar lenguajes de marcado o Archivos JSON, sino que también se integran con el análisis visual. Además, Yi et al (como se citó en Zhao, 2017), señala que el procesamiento del lenguaje natural es usado para simular cómo los usuarios humanos navegan por el contenido Web.

Krotov y Silva (2018), establecen que el Web Scraping consta de tres fases entrelazadas (Figura 1). La fase 1, es el análisis del sitio Web, para la cual se requiere examinar la estructura subyacente de un sitio Web o un repositorio Web con el fin de comprender cómo se almacenan los datos necesarios, requiriéndose para esto una comprensión básica de la arquitectura de la World Wide Web y varias bases de datos Web. La fase 2, es el rastreo del sitio Web, que implica el desarrollo y la ejecución de un script que navega automáticamente por el sitio Web y recupera los datos necesarios, desarrollados frecuentemente con R y Python. La fase 3, es la de la organización de los datos, para facilitar un análisis posterior de estos, para lo cual muchos lenguajes de programación, como los mencionados anteriormente, contienen bibliotecas de procesamiento de lenguaje natural y funciones de manipulación de datos que son útiles para limpiar y organizarlos. Estas tres fases de Web Scraping no pueden automatizarse completamente, por lo que frecuentemente requieren al menos algún grado de participación y supervisión humana.

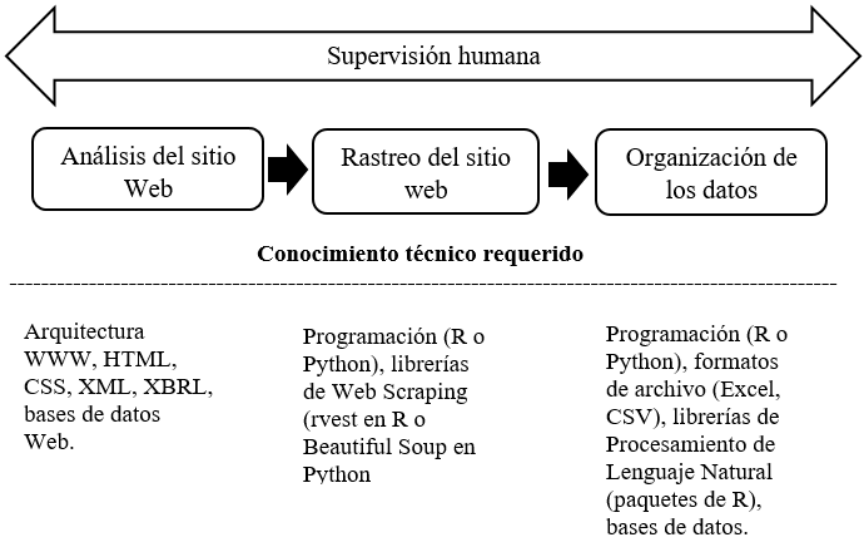


Figura 1. Web Scraping. Adaptado de Krotop y Silva (2018).

Pese a todos los beneficios que aporta este tipo de técnicas a los investigadores, no existen hasta el momento leyes sobre el uso Web Scraping directamente. Sin embargo, Dreyer y Stockton y Snell y Menaldo (como se citó en Krotop y Silva, 2018), refieren que estos procesos están guiados por un conjunto de teorías y leyes legales fundamentales relacionadas, como las infracciones de derechos de autor, incumplimiento de contrato, la Ley de fraude y abuso informático y traspaso a bienes muebles. A continuación, se proporcionan algunos detalles específicos sobre cómo se aplican estas teorías legales fundamentales al Web Scraping.

> El principal inconveniente que se da en el proceso de Web scraping es la falta de formato de la información que se desea obtener. Aunque pueda que existan APIs que entreguen los datos en un formato estandarizado o documentado, muchas veces son poco permisivos y limitan la cantidad de datos o la frecuencia con la que se solicitan. (López, 2017, p.23).

2.1.3. Analítica de Big Data.

Recolectar, organizar y analizar información sobre clientes, pacientes, estudiantes, docentes, entre otros, con el fin de generar información de importancia para mejorar los servicios e ingresos, o disminuir los gastos, ha sido un trabajo del que la sociedad desde sus distintos ámbitos se ha ocupado desde hace mucho tiempo. La información sobre gustos, edades, preferencias, enfermedades, entre otros, siempre ha estado entre nosotros dispuesta a ser recolectada en encuestas, censos u otras metodologías, pero es con el avance de las tecnologías que se han abierto posibilidades de uso de información capturada directamente de la interacción de los usuarios de Internet, aplicando la analítica de Big Data para obtener valor agregado de esta, que sea un insumo valioso para la toma de decisiones más certeras y acordes con la realidad de manera más eficiente.

Aproximadamente desde el año 2011, el concepto ha tomado fuerza, debido a ser reconocido como uno de los grandes avances tecnológicos de la actualidad y con diversas utilidades para las industrias, la economía, la política, la educación y otros sectores, por esto se ha convertido en tendencia y en fuerte motivo de estudio para la ciencia (Jiménez, 2014).

2.1.3.1. Big Data.

Existe una gran cantidad de definiciones o conceptos que se encuentran relacionados con este término, debido a sus alcances y la importancia que alberga para el desarrollo actual de diversidad de sectores. Joyanes (2013), señala que no existe unanimidad en la definición del término, pero destaca la existencia de un consenso en relación con la fuerza disruptiva que suponen los grandes volúmenes de datos y la necesidad de su captura, almacenamiento y análisis.

Camargo, Camargo y Joyanes (2014), hacen una recopilación de definiciones de este término según diversos autores, que van desde las más simples, como las que lo describen desde la imposibilidad de manejar gran cantidad de datos, por medios convencionales así: "cantidades masivas de datos que se acumulan con el tiempo que son difíciles de analizar y manejar utilizando herramientas comunes de gestión de bases de datos", o "tratamiento y análisis de enormes repositorios de datos, tan desproporcionadamente grandes que resulta imposible tratarlos con las herramientas de bases de datos y analíticas convencionales". Por otra parte, desde la perspectiva de Forrester (como se citó en Camargo, Camargo & Joyanes, 2014, p.65), Big Data hace referencia a: "las técnicas y tecnologías que hacen que sea económico hacer frente a los datos a una escala extrema; Big Data trata de tres cosas: 1) Las técnicas y la tecnología, lo que significa que la empresa tenga personal, el cual tenga gran representación y análisis de datos para tener un valor agregado con información que no ha sido manejada. 2) Escala extrema de datos que supera a la tecnología actual debido a su volumen, velocidad y variedad. 3) El valor económico, haciendo que las soluciones sean asequibles y ayuden a la inversión de los negocios"

Por su parte Benítez (2016), añade una definición que trata de abarcar lo esencial y complejo que resulta ser el uso de la Big Data afirmando que es, "cualquier característica sobre los datos que represente un reto para las capacidades de un sistema".

De acuerdo con Alla (citado por Barrón, 2019), Big Data como término tuvo sus orígenes en la época de los 90´s y año tras año se fueron presentando ciertas características para cubrir las necesidades del Big Data, comenzando en 2001 con las apreciaciones realizadas por Doug Laney, quien sostenía la necesidad de las 3 Vs del Big data, variedad,

velocidad y volumen, sin embargo, estas fueron siendo insuficientes, agregándose cada vez más, de la siguiente manera:

Variedad: Los datos se pueden venir de una gran variedad de fuentes, como sensores, posts en redes sociales, transacciones bancarias, ventas y marketing.

Velocidad: Se refiere tanto a la velocidad con que se crean los datos como a la velocidad con los que se pueden procesar, almacenar y analizar.

Volumen: El volumen se refiere a la escala de los datos (cientos de TB para muchas empresas), que es parte de lo que hace que los grandes datos sean masivos.

Veracidad: Se trata de obtener los datos que realmente contengan información procesable, lo cual es difícil de garantizar debido a la gran variedad de fuentes.

Variabilidad: Se refiere a los datos cuyo significado cambia constantemente. Es necesario entender su contexto para poder ser interpretados.

Visualización: Representar los datos con imágenes (gráficas generalmente), lo que permite interpretar información de una manera sencilla.

Valor: El gran volumen puede ocasionar que los datos sean confusos, de ahí que es importante que los datos realmente puedan ser procesados y entendidos (Alla, 2018, citado por Barrón, 2019, p 10).

2.1.3.2. Big Data Analytics.

En concordancia con lo que señala Russom (2011), el término Big Data Analytics hace referencia al uso de las técnicas de la analítica avanzada operando en conjunto con las técnicas Big Data. Es decir, Big Data Analytics se trata realmente de cómo se han unido Big

Data y la analítica para crear una de las tendencias más profundas en inteligencia empresarial (BI). Destacando, además, que algunos de los principales beneficios que aporta esta unión están relacionados, por una parte, con la proporción de muestras gigantescas por parte de Big Data, que mejoran los resultados de las herramientas analíticas, lo cual, revela y aprovecha el cambio comercial. Además, las herramientas analíticas y las bases de datos ahora pueden manejar Big Data y los procesos de la analítica se hacen más económicos.

2.1.4. Procesamiento de lenguaje natural.

El procesamiento del lenguaje natural consiste en la utilización de un lenguaje natural para comunicarnos con la computadora, debiendo ésta entender las oraciones que le sean proporcionadas, el uso de estos lenguajes naturales facilita el desarrollo de programas que realicen tareas relacionadas con el lenguaje o bien, desarrollar modelos que ayuden a comprender los mecanismos humanos relacionados con el lenguaje (Cortez, Huerta, & Pariona, 2013, p.48).

En este sentido, puede decirse que la razón de ser del procesamiento del lenguaje natural está basada en construir sistemas y mecanismos para dar cabida a la comunicación entre personas y máquinas haciendo uso del lenguaje conocido y dominado por los humanos y sistematizado en las máquinas.

Hernández y Gómez (2013), señalan que las áreas de investigación del procesamiento del lenguaje natural giran en torno a la recuperación y extracción de información, minería de datos, traducción automática, sistemas de búsquedas de respuestas, generación de resúmenes automáticos, análisis de sentimientos, entre otras.

Cortez, Huerta, y Pariona (2013), plantean la arquitectura de un sistema de procesamiento de lenguaje natural sustentada en definiciones del lenguaje natural por niveles, de la siguiente manera (sintetizado en la Figura 2):

a. Nivel Fonológico: trata de cómo las palabras se relacionan con los sonidos que representan.
b. Nivel Morfológico: trata de cómo las palabras se construyen a partir de unas unidades de significado más pequeñas llamadas morfemas.
c. Nivel Sintáctico: trata de cómo las palabras pueden unirse para formar oraciones, fijando el papel estructural que cada palabra juega en la oración y que sintagmas son parte de otros sintagmas.
d. Nivel Semántico: trata del significado de las palabras y de cómo los significados se unen para dar significado a una oración, también se refiere al significado independiente del contexto, es decir de la oración aislada.
e. Nivel Pragmático: trata de cómo las oraciones se usan en distintas situaciones y de cómo el uso afecta al significado de las oraciones. Se reconoce un subnivel recursivo: discursivo, que trata de cómo el significado de una oración se ve afectado por las oraciones inmediatamente anteriores (p.48).

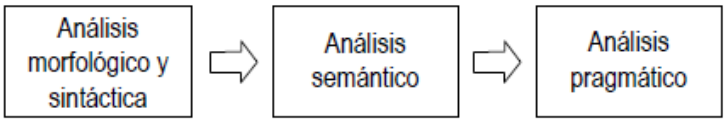


Figura 2. Arquitectura de un sistema de procesamiento de lenguaje natural. Recuperado de: Cortez, Huerta, y Pariona (2013). Aplicaciones de procesamiento de lenguaje natural (p.48)

Por otra parte, Cortez, Huerta, & Pariona (2013), establecen como una ventaja del procesamiento del lenguaje natural, el poco esfuerzo que requiere el interlocutor para aprender el medio de comunicación y como una desventaja del mismo las limitaciones que tiene la máquina en la comprensión del lenguaje natural. Convirtiéndose la ventaja de la comunicación humana en un problema en la interacción con la computadora, teniendo en cuenta que implican conocimiento y procesos de razonamiento que aún no se ha determinado cómo caracterizarlos ni cómo formalizarlos.

2.1.5. Web semántica.

Desde su nacimiento hasta la actualidad la Web ha venido evolucionando de acuerdo con las necesidades que en cada momento se han ido presentando. En la Figura 3, se puede ver gráficamente como ha venido evolucionando la Web y lo que se espera de ella en los próximos años.

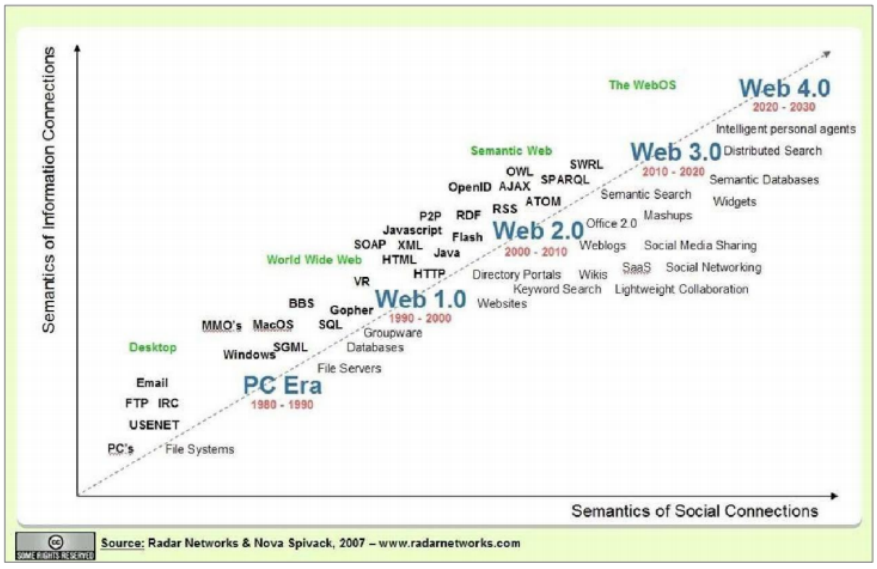


Figura 3. Evolución y predicción de la web. Recuperado de: Radar Network & Nova Spivack 2007 (como se citó en Hidalgo y Rodríguez, 2013).

Sobre algunos problemas que presenta la Web en relación con la exactitud de los resultados de búsqueda, Fensel et al. (como se citó en Villada & Jiménez, 2017), refiere que el problema radica en que el volumen de información en la Web crece con gran rapidez, pero no ocurre lo mismo con las herramientas para consultarla, creándose entonces una brecha entre lo disponible y lo asequible. Resalta también que, para la mayoría de usuarios esto no representa un verdadero problema teniendo en cuenta que sus necesidades de búsqueda se satisfacen fácilmente. Sin embargo, para científicos y gobiernos si se generan dificultades al realizar la mayoría de sus consultas de gestión del conocimiento.

Hidalgo y Rodríguez (2013), plantean que el concepto de Web semántica ha tomado relevancia en los últimos años, y resaltan que no constituye una Web totalmente nueva, sino una evolución natural de la Web tradicional, en la que los datos poseen un significado comprensible por sistemas informáticos, soportada sobre la plataforma tecnológica de la Web actual manteniéndose el uso del protocolo HTTP para el intercambio de datos.

"El término Web Semántica fue propuesto por Tim Berners-Lee, el propio inventor de la Web y director del W3C (World Wide Web Consortium). Él define la Web Semántica como una Web de datos que puede ser procesada directa e indirectamente por máquinas" (Roig, 2011, p.10).

Además, Roig (2011) destaca, que se trata de estructurar la información que circula en la Web, de tal forma que las máquinas puedan solucionar problemas previamente definidos, utilizando operaciones bien definidas, sobre unos datos bien definidos.

Pastor (2011), describe la arquitectura de la Web semántica y plantea que, para comprenderla, es necesario recordar que los tres pilares fundamentales de la Web original

son la localización de objetos y recursos mediante URL, el uso de protocolos HTTP para establecer la información entre los clientes y los servidores y el marcado de documentos con lenguaje HTML. Teniendo en cuenta que la Web semántica amplía la localización de recursos a otro concepto más general, la identificación mediante el uso de URI, XML sustituye a HTML como lenguaje de marcas utilizados para almacenar datos y el protocolo HTTP sigue siendo el encargado de la trasmisión de los objetos entre clientes y servidores. Resaltando, que con esto se busca principalmente la legibilidad de estos datos empleando especificaciones para la representación de la información.

Por último, Lozano (2019), plantea que lo que se pretende principalmente con la Web semántica es que los datos puedan ser utilizados y comprendidos por ordenadores sin supervisión humana, convirtiendo la información en conocimiento, referenciando datos dentro de las páginas Web a metadatos con un esquema común consensuado.

2.1.6. Metadatos.

Los metadatos forman parte de los conceptos fundamentales para una mejor comprensión de la Web semántica, “Desde un punto de vista general… los metadatos son elementos que describen un determinado objeto siguiendo algún modelo o conjunto de reglas” (Pastor, 2011, p. 21).

Definiendo los metadatos a partir de lo que se plantea desde la Web semántica podría decirse que: “Son descripciones estructuradas y codificadas que describen características y propiedades de objetos y recursos para facilitar su localización, recuperación, valoración administración, persistencia e interoperabilidad. No sólo son capaces de realizar una descripción de dichos recursos, sino que también pueden hacer lo propio con los procesos en

los que intervienen, sus componentes, las restricciones y las relaciones que se establecen entre ellos. La dispersión de los recursos, así como el gran volumen de información contenida en ellos y la diversidad en cuanto a su actualización y estructura son algunos de los factores que han hecho los modelos tradicionales de descripción bibliográfica que se hayan demostrado ineficaces al aplicarse a los recursos disponibles en la Web" (Pastor, 2011, p.23).

Según Audit Commission Publishing Team (como se citó en Castro, González & Ballesteros, 2015), los metadatos deben ser precisos, validos, fiables, oportunos e integrales, es decir que, al ser tan importantes para el manejo de la gestión de datos, estos deben ser lo suficientemente detallados, consistentes con lo que pretenden representar, estables, creíbles, actualizados y contener la información que dicen tener.

2.1.7. Ontologías.

La palabra ontología, de raíces griegas, desde tiempos antiguos está relacionada con el estudio del ser, en la actualidad pasa a jugar un papel fundamental en la gestión de información. Según Pedraza, Codina y Rovira (2007), en este nuevo contexto, las ontologías pueden considerarse lenguajes documentales con distintos niveles de estructura, elaboradas con una sintaxis comprensible para los ordenadores. Además, contemplan un conjunto amplio de relaciones, ya que a menudo dependen de las relaciones reales que se den entre las clases y los individuos del dominio modelado por la ontología, es por esto que una ontología permite gran riqueza en la definición de sus conceptos y sus relaciones. Es importante resaltar que están expresadas mediante una rigurosa lógica formal y, por tanto, no solamente pueden ser procesadas por aplicaciones informáticas, sino que, pueden soportar bajo ciertas restricciones, procesos de inferencia automáticos.

Las definen como los mecanismos que proporcionan la fiabilidad en cuanto a la semántica de lo expresado, sobre lo que se pretende comunicar

En este mismo sentido, “las ontologías proporcionan un vocabulario común en un área de conocimiento, definiendo, en diferentes niveles de formalización, el significado de los términos y las relaciones entre ellos” (Pastor, 2011, p.29).

Pastor (2011), hace una comparación con términos ya conocidos con el fin de facilitar el entendimiento de estas, afirmando que una de las nociones más simples de una posible ontología puede relacionarse con un vocabulario controlado, es decir, una lista finita de términos o con glosario, como una lista de términos y significados especificados como declaraciones en lenguaje natural.

2.1.8. Minería Web.

Etzioni (como se citó en Fuentes y Ruíz, 2007), define la minería Web como el empleo de las técnicas de la minería de datos, para descubrir y extraer información útil automáticamente desde la World Wide Web.

Mendoza (2011), en relación con la minería de datos, señala que existen diversos desafíos relacionados principalmente con la naturaleza volátil de su contenido y a la heterogeneidad de formatos y estilos que se entregan a los usuarios y con el enorme tamaño de este repositorio. Los cuales han generado la necesidad de crear nuevos métodos, estrategias y algoritmos especialmente desarrollados para la minería de datos en la Web, mucho más complejos que la simple aplicación de técnicas y métodos estándar de minería de datos.

Fuentes y Ruíz (2007), destacan la importancia de la aplicación de la minería Web en Internet, en los relacionado con el acceso, la recuperación y la organización de información. Ayudando a descubrir conocimientos potencialmente útiles para las organizaciones.

"La minería de datos en la Web ha sido categorizada en tres áreas de interés distintas dependiendo del tipo de fuentes de información que han sido exploradas. Estas áreas son: Minería de contenido, minería de estructura, y minería de uso" (Mendoza, 2011, p.4).

2.1.9 Paradigma metodológico de minería de texto.

La segunda fase permitió dar cabida a la minería de texto, que como señala Rochina (2017), es un proceso que permite obtener información nueva a partir de grandes cantidades de texto sobre información no estructurada.

"Por tanto, en la minería de textos los datos a tratar serán los documentos y textos de las organizaciones, en lugar de los datos de las bases de datos, llevando a cabo un análisis de los datos compartidos por todos los textos de la colección" (Rochina, 2017).

Esta investigación se basó principalmente en la minería de textos basada en lingüística, a través del Procesamiento de Lenguaje Natural (PLN) con lo cual facilita "encontrar significado en el texto del modo en que lo hacen las personas, reconociendo una variedad de formas de palabra como similares en su significado y analizando la estructura de la oración para proporcionar una infraestructura para entender el texto" (IBM, s.f.).

IBM (s.f) señala, además, que la minería de texto basada en lingüística "ofrece la velocidad y la rentabilidad propia de los sistemas basados en estadísticas, pero proporciona un grado de precisión mucho más alto y menos intervención humana".

Por otra parte, se destaca que, la minería de textos comprende tres actividades fundamentales (Rochina, 2017):

1. Recuperación de la información: Consiste en seleccionar los textos pertinentes.
2. Extracción de la información incluida en esos textos mediante el procesamiento del lenguaje natural: Hechos, acontecimientos, datos clave, relaciones entre ellos, etc.
3. Minería de datos para encontrar asociaciones entre los datos clave previamente extraídos de entre los textos.

Rochina (2017), divide las actividades anteriores en tres fases, las cuales se han tomado como referencia para plantar las cuatro etapas de la presente investigación, en primer lugar, se menciona la etapa de pre-procesamiento, en donde se define el conjunto de documentos (corpus), el cual debe ser representativo y seleccionarse aleatoriamente o mediante algún método de muestreo probabilístico, facilitando el análisis que se realizará posteriormente. Seguidamente estos serán convertidos a un formato analizable, para poder crear una representación estructurada o semi-estructurada de los mismos. Luego de tener seleccionado y estructurado el corpus lingüístico, es necesario reconocer las unidades gramaticales más pequeñas, lo que implica representar el texto como una lista de palabras mediante una representación vectorial. En esta última etapa se destaca que las representaciones internas se analizan con el objetivo de descubrir en ellas algunos patrones interesantes o nueva información, y es la etapa en la que los usuarios pueden observar y explorar los resultados.

2.1.10. Aprendizaje de máquina.

El Aprendizaje de Máquina (del inglés Machine Learning), también llamado Aprendizaje Automático, es generalmente considerado un área de la minería de datos o de la inteligencia artificial. Sin embargo, de acuerdo con lo que plantea Mitchell (como se citó Godoy, 2015), el aprendizaje de máquina es un área que, con la ayuda de diversas disciplinas como la inteligencia artificial, estadística y probabilidad, teoría de la información, psicología, neurobiología y otras, estudia cómo construir programas informáticos que mejoren su desempeño en alguna tarea gracias a la experiencia.

Por otra parte, Sandoval (2018), señala que esta es una técnica con la que se pueden detectar patrones "a bajo nivel" en miles de datos individuales. Resalta, además, que una de las principales potencias a destacar de esta, es el desarrollo de aplicaciones predictivas, porque da cabida a que se dé la automatización de procesos, la toma de decisiones y el continuo aprendizaje basado en datos.

Machine Learning, suele tener estrecha relación con otras tecnologías, por ejemplo, con la minería de datos existe gran similitud. Cuenca (2019), aclara que mientras en la minería de datos se extraen datos para la comprensión humana, Machine Learning se usa para detectar patrones y ajustar las acciones propias del programa. De modo similar pasa con el término Business Intelligence, ante lo cual Sandoval (2018), plantea que la principal diferencia entre los dos, radica en la detección de patrones entre millones de datos, por parte de Machine Learning, y establece tres aspectos a tener en cuenta para clarificar las diferencias. Machine Learning usa miles de variables al momento de detectar patrones, gracias a que, al usar datos agregados, utiliza datos individuales con características definitorias de cada una de las instancias, también ofrece una analítica predictiva en lugar de

basarse en una analítica descriptiva y uno de los mayores potenciales de Machine Learning, es que los algoritmos predictivos aprenden automáticamente de los datos y sus modelos se pueden integrar en aplicaciones para dotarlas de capacidades predictivas.

Cabezón (2018), menciona que Machine Learning se destaca por dar solución a problemáticas alrededor de gran variedad de ámbitos, por ejemplo, a problemas en los que existe solución, pero es muy difícil su mantenimiento o problemas con muchas reglas que dificultan su programación. También da soluciones a problemas complejos para los que no existe buena solución. El sistema de Machine Learning se adapta a los nuevos datos, dando soluciones a problemas cuyo entorno va cambiando. Además, puede recoger información sobre problemas muy complejos y con gran cantidad de datos.

Una de las clasificaciones de Machine Learning, es el Aprendizaje Supervisado, Aprendizaje No Supervisado y el Aprendizaje Profundo. Los datos que se usan para entrenar la máquina pueden presentarse en variedad de formatos como, videos, fotos, audios, Dataset y como metodología básica el total de los datos se divide en dos o tres sub-conjuntos, llamados entrenamiento, validación y test como se puede observar en la Figura 4, con el objetivo de garantizar que el modelo se generalice para trabajar con otros conjuntos de datos diferentes a los que se usaron para crearlo.

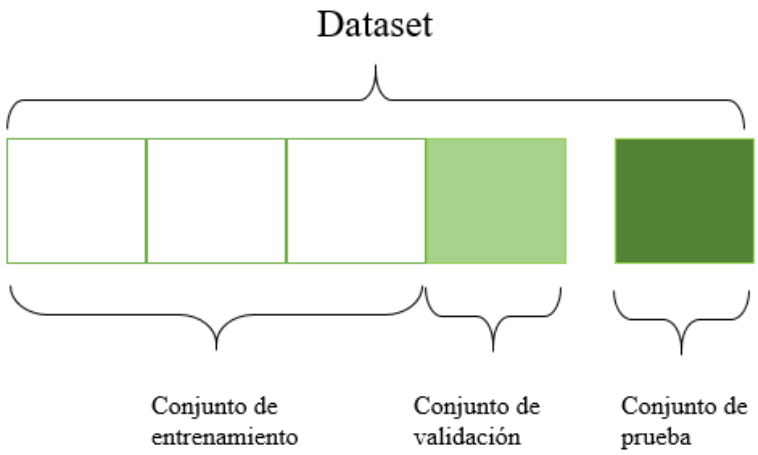


Figura 4. Diagrama de división en sub-conjuntos de los datos. Adaptado de Cuenca (2019).

2.1.10.1 Aprendizaje supervisado.

Los sistemas de aprendizaje supervisados según Sancho (2018), son aquellos en los cuales el aprendizaje de la máquina se da a partir de un conjunto de ejemplos de los cuales se conoce su valor objetivo, al cual se le llama conjunto de entrenamiento, y se intenta encontrar una función que permita asignar un valor objetivo a ejemplos que el sistema no ha visto anteriormente, ya sea como clasificación o como regresión. En cuanto a estos tipos de aprendizaje supervisado, Rodríguez (2018), establece que los algoritmos de clasificación, son necesarios cuando los datos se usan para predecir una categoría multiclass y se clasifican varias clases y los de regresión, cuando se predice un valor y se detectan anomalías, los cuales buscan identificar datos que no son habituales, es decir, que difieren unos patrones identificados previamente como actividad normal.

Los algoritmos de aprendizaje supervisado realizan sus predicciones basándose en un conjunto de ejemplos etiquetados. El algoritmo busca a los patrones en esas etiquetas y una

vez que encuentre el mejor patrón posible, puede ser usado para hacer predicciones de datos de prueba sin etiquetar (Rodríguez, 2018, p.25).

Algunos ejemplos de técnicas de aprendizaje de máquina supervisado son Árboles de Decisión, Máxima Entropía, Naive Bayes, SVM, K-nn, Regresión Lineal, Regresión Logística, Random Forest, entro otros (Godoy, 2015; Cuenca, 2019).

2.1.10.2 Aprendizaje no supervisado.

Sancho (2018), señala que los sistemas de aprendizaje no supervisados son aquellos en los que no se dispone de una salida esperada que asociar a los ejemplos con los que se trabaja, sino que únicamente a partir de las propiedades de los ejemplos intentamos dar una agrupación o caracterización de los ejemplos según la similitud entre sus propiedades.

Peláez (2012), en cuanto a la aplicación de Redes Neuronales con aprendizaje no supervisado, explica que consiste en que la red pueda descubrir por sí misma características, regularidades, correlaciones o categorías en los datos de entrada y a partir de ello obtengan de forma codificada los de salida. En algunos casos, la salida representa el grado de similitud entre la información que se le está presentado en la entrada y la que se le ha mostrado en el pasado. En otros casos, se podría realizar un establecimiento de categorías, indicando con la salida de la red a qué categoría pertenece la información presentada como entrada, siendo la propia red quien deba encontrar las categorías apropiadas a partir de correlaciones en las entradas presentadas.

Los algoritmos de aprendizaje no supervisado, generalmente se dividen en dos grupos. Los clustering y los de visualización y reducción de la dimensión. Los de clustering, son técnicas exploratorias de análisis de datos que se usan para organizar la información en

grupos similares sin tener conocimiento previo de sus estructuras (Figura 5). Algunos de los modelos más usados en esta clasificación son: K- Means, Hierarchical, Cluster Analysis y Expectation Maximization (Cuenca 2019).

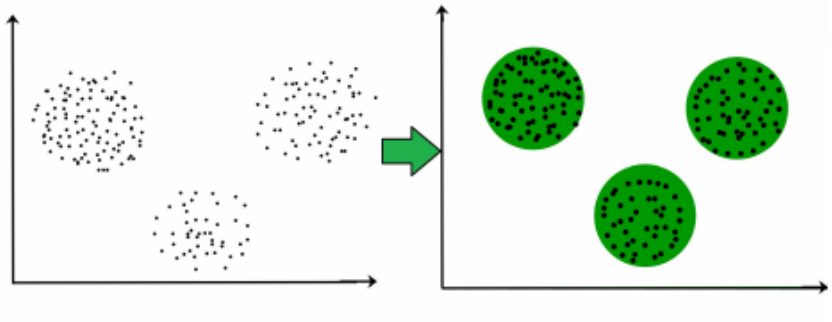


Figura 5. Ejemplo de algoritmos Clustering.
Fuente: Cuenca (2019).

Los de visualización y reducción de la dimensión son técnicas con las que se reduce la dimensión de los datos cuando estos presentan un alto número de características que dificultan la capacidad de procesamiento (Figura 6). Los más utilizados son: PCA, LLE, t-SNE.

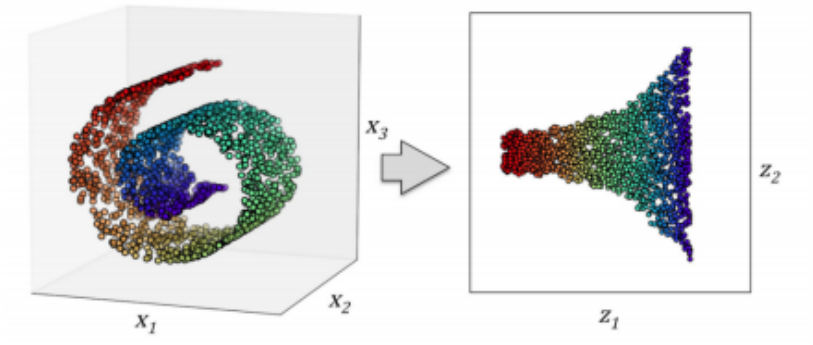


Figura 6. Ejemplos de algoritmos de reducción de la dimensión. Fuente: Cuenca (2019).

2.1.10.3 Aprendizaje profundo.

El moderno termino Deep Learning, va más allá de los estrechos propósitos científicos en el cerebro actual de los modelos de Machine Learning. Apela a un principio más general de aprendizaje de múltiples niveles de composición, los cuales pueden ser aplicados en Machine Learning marcos que no están necesariamente inspirados neuralmente (Goodfellow, Bengio, & Courville, 2016, p.14).

Hinton (como se citó en Rodríguez, 2018), acuña el término Deep Learning o aprendizaje profundo en 2006, para explicar nuevas arquitecturas de redes neuronales profundas.

Rodríguez (2018), plantea que Deep Learning, "es una de las áreas de mayor crecimiento, utiliza algoritmos multinivel para crear diferentes niveles de abstracción de la información. Algunas aplicaciones son reconocimiento de imágenes, voz, predicción de tendencias digitales y perfeccionamiento de procesos (p.12).

De acuerdo con lo que plantea Cabezón (2018), los campos de Machine Learning y Deep Learning a menudo suelen parecer indistinguibles, debido a que las dos tecnologías hacen referencia a sistemas capaces de aprender por sí solos. Sin embargo, señala que la diferencia entre ambos está en el método de aprendizaje, siendo el de Deep Learning el más complejo, sofisticado y con mayor autonomía, teniendo en cuenta que, una vez programado el sistema, la intervención del ser humano es mínima.

2.2. Descripción de los trabajos relacionados.

2.2.1. Aplicación de técnicas de Machine Learning a la detección de ataques.

Este proyecto realizado por Rodríguez (2018), pretende demostrar la utilidad de Machine Learning y su aplicación a la seguridad informática, concretamente a las conexiones maliciosas. Para lo cual, utilizó un Dataset con un aproximado de 4900000 vectores conexiones con 41 atributos, etiquetados como normal o como un ataque determinado, capturados en la red por 7 semanas. Con base en la naturaleza de este Dataset, usó algoritmos de Aprendizaje Supervisado binarios y multiclase y a su vez usó Weka para seleccionar los mejores algoritmos de precisión y rendimiento, seleccionando los Árboles de Decisión con el objetivo de mejorar los valores de Weka.

Organizó su trabajo en tres grandes etapas; formación e investigación, desarrollo del sistema y documentación, la cuales le permitieron: exponer la utilidad del Aprendizaje Automático y particularmente su aplicación a la seguridad. Estudiar diversas opciones de Weka y Scikit-Learn de análisis y visualización. Realizar distintas modificaciones a los datos basándose en una versión inicial del código y se realizaron diversos ajustes a las mismas para obtener mejores resultados en la solución final. Optimizar los parámetros de entrenamiento, obteniendo un 99.9916% de precisión en un tiempo de 45 segundos. Obtuvo un script en el que utiliza un Árbol de Decisión para detectar los ataques con unos resultados finales favorables. Incluso presenta un resultado como mayor precisión al hacer clasificación binaria la precisión es de 99.9955%.

2.2.2. Machine Learning aplicado a la seguridad.

El trabajo realizado por Cuenca (2019), en atención a algunos de los problemas de seguridad que se han venido presentando con el incremento del uso de Internet, se centra en hacer un análisis que permita generar un modelo de Machine Learning que sea capaz de detectar contenido falso o bots en la red social Twitter. Para ello utiliza el proceso de minería de datos CRISP-DM, con información de los usuarios distribuida en dos Dataset; Detecting twitter bot y PHEME, para detectar boot y noticias falsas en las redes sociales respectivamente.

A partir de la información de los Dataset mencionados, se entrenaron varios modelos de Aprendizaje Supervisado y No Supervisado, para finalmente seleccionar el modelo que con mayor eficiencia permitiera cumplir el objetivo propuesto en el proyecto.

Para iniciar el proceso de entrenamiento dividió el Dataset en dos partes; train con un 70% y test con 30%. El proceso de entrenamiento se da con los modelos Decision Tree, Random Forest, SVM y K-Means y para definir el modelo más adecuado, para cada Dataset utilizó el área bajo la curva (ROC), como mecanismo para seleccionarlo de forma objetiva. Arrojando esto, que el mejor modelo para ambos Dataset resultó ser Random Forest, aunque Decisión Tree ofrece valores similares y el modelo que proporciona valores más bajos es K-Means en el proceso de entrenamiento. Seleccionó Random Forest y tras seleccionarlo evaluó su capacidad para generalizar usando datos que con los que no había trabajado antes. Para ellos empleó conjuntos de datos previamente seleccionados para Test, obteniendo de esta práctica valores aceptables para la generalización, 0.81 y 0.96 para Bot y Fake News respectivamente.

Dentro de las sugerencias para trabajos futuros, se destaca la recomendación del uso de otro tipo de algoritmos usados con menor frecuencia, dentro de los cuales destaca, redes neuronales o combinaciones de modelos entre sí.

2.2.3. Detección de intrusiones en sistemas Web utilizando técnicas de aprendizaje profundo.

Este trabajo de investigación realizado por Hidalgo (2019), tiene como objetivo principal desarrollar un diseño particular e implementar un detector de intrusión, usando una arquitectura de aprendizaje profundo basada en Redes Neuronales Artificiales, luego de entrenarla usando un Dataset creado previamente en el que se han registrado y etiquetado las operaciones normales y los diferentes tipos de ataques de piratería por parte de intrusos hostiles. Para ello usó un Dataset desarrollado en 2017 por el Instituto Canadiense de Ciberseguridad (CIC) de la Universidad de New Brunswick, llamado `Intrusion Detection Evaluation Dataset (CICIDS2017)'. A través de Matlab, realizó un conjunto de pruebas con diferentes profundidades de la arquitectura, con el fin de ir fortaleciendo el modelo, para finalmente, presentar los resultados de un entrenamiento completamente capacitado en el campo de las Ciencias de la Computación, específicamente en ciberseguridad llevando a cabo, de manera eficiente, un sistema de detección de intrusos, basado en arquitecturas de aprendizaje profundo. Entre sus resultados se destaca que la introducción de un autoencoder como procesador intermedio tuvo un buen funcionamiento, dando estabilidad a las Redes Neuronales Artificiales y elevando el rendimiento general y la precisión del detector de intrusos a un 99.4%, de un rendimiento inicial inestable del 95% sin usar el autoencoder, concluyendo que las arquitecturas profundas, con entrenamiento separado con datos sin etiquetar y etiquetados, producen procesadores neurales mejorados.

2.2.4. Ciberguerra y su efecto en la seguridad nacional.

En este documento escrito por Blanca Cuji (2018) y publicado por la Revista de Ciencias de Seguridad y Defensa, principalmente se hace un análisis de artículos de revisión sobre la Ciberguerra, con el objetivo de determinar las tecnologías utilizadas y las implicaciones económicas de la Ciberguerra en la seguridad nacional. Para esto, ella hizo una recolección de ochenta artículos científicos y seleccionó once relacionados con la temática mencionada, publicados entre 2013 al 2018 en revistas indexadas o congresos reconocidos internacionalmente, y con base en ellos dio respuesta a las preguntas de investigación propuestas en su trabajo: ¿Qué tipo de tecnologías y herramientas informáticas son utilizadas en las Ciberguerras? y ¿Cuáles son las implicaciones económicas de las Ciberguerras en la economía de un país?

Concluyendo entre otras cosas que gracias a este fenómeno el ámbito militar, requieren cambios en las estrategias militares, en muchos países no se han elaborado normas básicas para enfrentar dicha problemática, los países desarrollados tienden, a diseñar sistemas tecnológicos que les permiten alejar posibles agresiones en el ámbito del ciberespacio. Internet ha pasado a ser un escenario ideal para los ciberataques de organizaciones terroristas y delictivas, debido a que tienen fácil acceso, escaso o nulo control gubernamental, anonimato, flujo de información y poco riesgo.

2.2.5. La Ciberguerra como realidad posible contemplada desde la prospectiva.

Este artículo escrito por Domínguez (2016), publicado por la revista de Pensamiento Estratégico y Seguridad CISDE, contiene un análisis y comparación de varios conceptos

relacionados o derivados de la Ciberguerra, para lo cual desde un contexto general alrededor de lo que es la guerra y sus implicaciones, cita artículos científicos u otros documentos publicados entre 2001 y 2015, que desde diversidad de posturas plantean los principales aspectos que se resaltan sobre la Ciberguerra, en relación con la posibilidad o imposibilidad de que bajo las condiciones económicas, políticas, sociales y de seguridad actuales, pueda desencadenarse en el mundo la Ciberguerra. De sus conclusiones se destaca el considerar la Ciberguerra como un tema de importancia en la investigación científica en constante evolución, y señala, además, que los países desarrollados tienden hoy a construir todo un sistema de medios que les permitan repeler posibles agresiones en el ámbito del ciberespacio, y a fortalecerse en cuanto a su capacidad de ataque.

2.2.6. Ciberguerra... ¿Dudáis?

Este artículo escrito por Gómez (2017) y publicado por la Revista de Marina Nº 959, fundamenta por qué las ciberoperaciones pueden ser un instrumento útil para la coerción, para lo cual inicia citando autores escépticos frente a la temática que desvirtúan con distintos argumentos que el ciberespacio pueda ser un espacio destinado para guerra. Seguidamente cita autores que desde otros enfoques analizan la naturaleza del ciberespacio como un posible campo de guerra entre dos estados, basados en la evolución del concepto en la misma manera que ha evolucionado la sociedad, dándole más soporte al concepto de Ciberguerra, desde una aproximación basada en los efectos finales que pueden alcanzarse al aplicar estas técnicas, partiendo de que aunque su potencial destructivo en la actualidad es aún incipiente y muy limitado, plantea a la Ciberguerra como una realidad y prevé que tanto su empleo como sus efectos físicos se irán incrementando con el tiempo.

2.2.7. Aplicación de técnicas de minería de datos en ciberseguridad y ciberdefensa - una breve revisión.

En este trabajo publicado por Monteiro (2017), se hace una breve revisión de diferentes técnicas de minería de datos, desde la visión de Data Mining como la integración de múltiples tecnologías y la confluencia de diferentes disciplinas, como Machine Learning, Estadística, Visualización, Inteligencia Artificial, Data Warehouse, para luego evaluar su aplicabilidad en la ciberdefensa y en el desarrollo de plataformas y herramientas para ciberseguridad. En sus conclusiones destaca la importancia de algunos aspectos relacionados con las técnicas de Data Mining, clasificadas según asociación y correlación, clasificación, agrupación y detección de anomalías, destacando las principales necesidades y formas de aplicación de las técnicas de minería de datos en seguridad cibernética y defensa cibernética, con las cuales se pueden realizar un conjunto de tareas sobre los datos de las redes, computadoras y servidores, bases de datos, sistemas operativos, etc., con el fin de obtener información útil para garantizar la protección de los sistemas.

2.3. Análisis de trabajos relacionados.

La Tabla 1, muestra un análisis de las características sobre la descripción de trabajos relacionados, usados como punto de comparación con la investigación actual.

Tabla 1. *Comparativo* de *características de trabajos relacionados para la detección de Ciberguerra.*

Trabajos relacionados	Tipo de investigación		Orientación del estudio	Técnicas para la detección vocabulario en temas afines a la Ciberguerra																														Herramientas de minería de datos o descarga de páginas web
				Fuentes de datos (Corpus)	Fuentes de datos (dataset)	Técnicas de procesamiento de datos									Aprendizaje supervisado										Herramientas tecnológicas para procesamiento											
	Cuantitativa	Cualitativa				TF y TF - IDF	Lematizar	Stop word	Tokenizar	Sinónimos	Crawler	PLN	Análisis de sentimientos	Vecino más cercano	Árboles de decisión	J48	Naive Bayes	SVM	Regresión logística	Ada Bosting	Random Forest	Redes neuronales	Redes Neuronales Deep Learning	Python	Java	Matlab	TensorFlow	Keras	NumPy	Pandas	Scipy	Matplotlib	Seanborn	Scikit - Learn	
Rodríguez (2018)	✓		Vocabulario detección de ataques	IDS DAPRA' 98	KDD99 Cup			✓							✓	✓	✓	✓	✓	✓		✓			✓	✓				✓	✓	✓	✓		✓	Weka
Cuenca (2019)	✓		Detectar contenido falso o bots	Twitter	BOTS y PHEME	✓	✓		✓	✓			✓	✓	✓			✓			✓			✓						✓	✓	✓	✓	✓	✓	CRISP- DM
Hidalgo (2019)	✓		Detectar intrusos	Tráfico en la red	CICIDS2017																		✓			✓										
Presente investigación.	✓	✓	Vocabulario de ciberguerra	Páginas Web	Páginas Web	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓		✓	ADVI

2.4. Variables.

La Tabla 2, muestra un ***análisis de las variables*** seleccionadas para el análisis de vocablos del tema de Ciberguerra, las cuales tienen sustento en los libros: "Ciberguerra" de Yolanda Quintana (2016) y "El quinto elemento: Espionaje, Ciberguerra y Terrorismo. Una amenaza real e inminente" de Alejandro Suarez Sánchez-Ocaña (2015).

Con el fin de facilitar la construcción del conjuntos de datos "dataset" que será utilizado en la analítica de datos con los algoritmos de Machine Learning, las variables son administradas por medio de ontologías semánticas, la cuales están formadas por metadatos, conformados a su vez por clases, subclases y atributos, de la siguiente forma: la clase representa el tema principal, la subclase a los subtemas y los atributos son cada palabra que permite identificar cada característica, como se aprecia en la Figura 7.

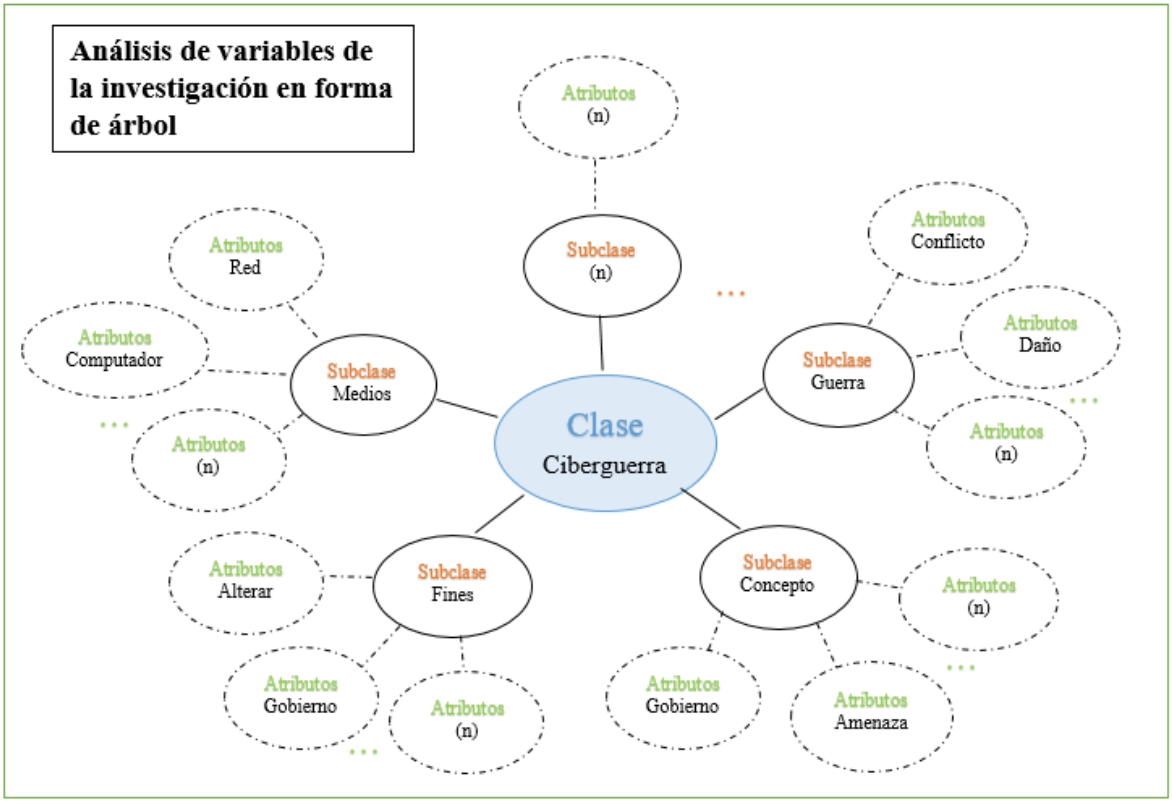


Figura 7. Análisis de variables de la investigación (Clases – subclases – atributos). Fuente: desarrollo propio.

La tabla 2, describe el análisis de las variables de Ciberguerra en español, organizado por grupos de variables.

Tabla 2. Variables organizadas por grupos (*Corpus lingüístico de Ciberguerra).*

Categorización del corpus lingüístico de Ciberguerra											
No.	**Grupo/término**	12.	Delincuencia.	23.	Computador.	36.	Robot.	47.	Electrónico.	59.	Energía.
	1. De Ciberguerra.	13.	Destrucción.	24.	Comunicación.	37.	Servidor.	48.	Envenenamiento.	60.	Expertos.
1.	Alterar.	14.	Ejercito.	25.	Cosas.	38.	Sistemas.	49.	Gusanos.	61.	Financiero.
2.	Amenaza.	15.	Enemigo.	26.	Digital.	39.	Social.	50.	MasterCard.	62.	Gobierno.
3.	Anonymous.	16.	Guerra.	27.	Espacio.	40.	Tecnología.	51.	PayPal.	63.	Grupo.
4.	Atacar.	17.	Masivo.	28.	Facebook.	41.	Twitter.	52.	Satélites.	64.	Hackers.
5.	Ataque.	18.	Ofensiva.	29.	Información.	42.	Web.	53.	Virus.	65.	Militares.
6.	Ciberataque.	19.	Penetrar.	30.	Informática.		**3. formas de**	54.	Visa.	66.	Nuclear.
7.	Cibercomando.	20.	Terrorismo.	31.	Infraestructura.		**Atacar.**		**4. Dirigido**	67.	País.
8.	Ciberdefensa.		**2. Medios**	32.	Internet.	43.	Antivirus.	55.	Agua.	68.	Petróleo.
9.	Ciberguerra.		**Informáticos.**	33.	Mensajes.	44.	Correo.	56.	Armado.	69.	Población.
10.	Conflicto.	21.	Ciberespacio.	34.	Ordenador.	45.	DNS.	57.	Armas.	70.	Represas.
11.	Daño.	22.	Cibernético.	35.	Red.	46.	Dominio.	58.	Cosechas.	71.	Seguridad.

Capítulo III. Materiales y Métodos.

3.1. Materiales (hardware y software utilizado en la investigación).

3.1.1. Hardware utilizado en la investigación.

Para la realización de este proyecto se utilizó un computador portátil HP con las siguientes características:

- Procesador Intel (R) Core (TM) i5-8250U CPU @ 1.60GHz 1.80GHz.
- 6 MB de memoria caché.
- GPU con 960 CUDA Cores (Nvidia GeForce Experience versión 3.20.2.34)
- Memoria RAM de 8 GB DDR4 de 2400MHZ.

3.1.2. Software utilizado para las pruebas.

El sistema operativo usado en la investigación es Windows 10 Home Single Language de 64 bits (10.0, compilación 18362). Para las pruebas de detección de Ciberguerra se utilizó el lenguaje de programación Python versión 3.7.4, en donde se construyó el código para los algoritmos de Arboles de Decisión, Naive Bayes, Knn, Red Neuronal, MSV, Ada Boosting, Regresión logística, Random Forest y Deep Learning. Las librerías de apoyo son: Scikit-Learn, Pandas, Nampy y matplotlib, Keras y TensorFlow).

3.2. Método.

En este apartado se exponen los métodos utilizados para analizar las variables relacionadas con el problema de investigación; de igual manera los procedimientos adoptados para dar respuesta a las preguntas de investigación y comprobar la hipótesis. La Figura 8, describe el proceso del método propuesto para la detección de vocabulario de Ciberguerra con sus etapas y respectivas fases.

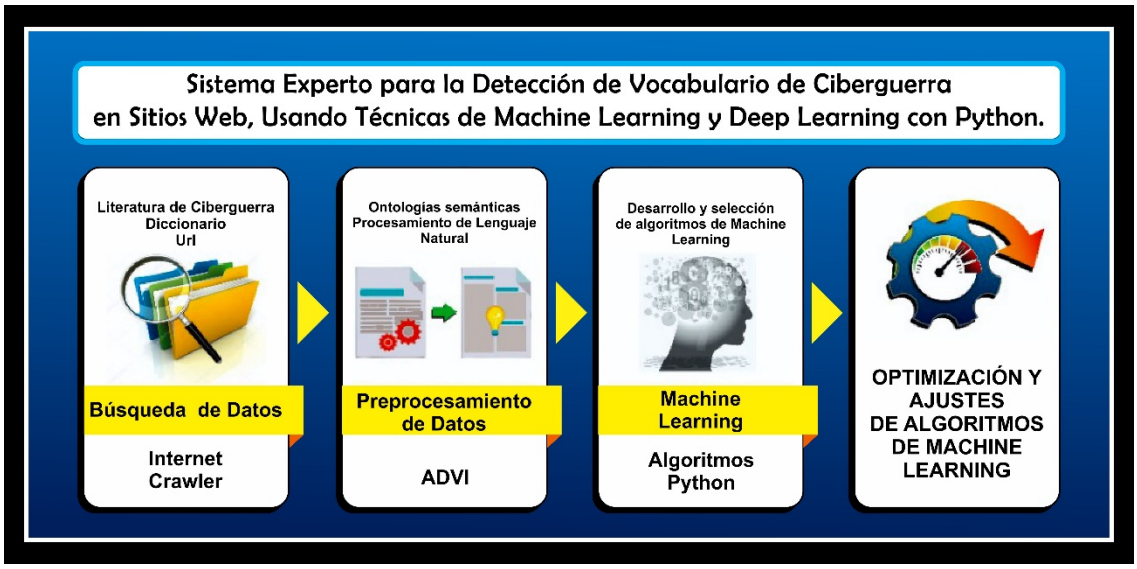


Figura 8. Diagrama del método de solución. Fuente: Desarrollo propio.

3.2.1. Búsqueda de datos (Etapa 1).

Esta etapa se basó principalmente en la búsqueda de información en Internet relacionada con la problemática seleccionada "Ciberguerra". Para identificar el contexto, las palabras relacionadas, el propósito, los medios a través de los cuales se genera y las formas en que es posible hacer ataques a través de esta. Seleccionando los sitios Web para generar el corpus lingüístico y a su vez el diccionario de Ciberguerra base para la investigación. Para ello, se localizaron y descargaron más de 1000 páginas Web haciendo uso de un Crawler, los cuales según Chatterjee y Nath (2017), están diseñados para recuperar páginas Web e insertarlas en el repositorio local. Con el propósito de crear una réplica de todas las páginas visitadas que luego son procesadas por un motor de búsqueda que indexará las páginas descargadas que ayudan en búsquedas rápidas.

3.2.2. Preprocesamiento de datos (Etapa 2).

Esta fase permite realizar el preprocesamiento de los datos obtenidos con la arquitectura ADVI desarrollada en **Java**, basada en la propuesta de Castillo, Luna, Muñoz y López (2016), con la intención de garantizar la limpieza de los datos. En donde, implementaron una estrategia genética integrando técnicas de Web semántica y procesamiento de lenguaje natural para detectar vocabulario de Ciberbullyng, a partir de:

- Tokenización: parte del proceso con el cual se segmentan las palabras que provienen de las páginas Web descargadas.
- Stop Word: permite eliminar las palabras vacías, es decir, las que no son esenciales para la creación de significados como preposiciones, pronombres, artículos, adverbios, conjunciones y algunos verbos. Esto se hace con el fin de optimizar las consultas del vocabulario.
- Frecuencia de Término (TF): permite almacenarlas en una base de datos para calcular el número de palabras y los tiempos que se repiten.
- Frecuencia de Término con Frecuencia Inversa del Documento (TF-IDF): permite asignar un peso, además de indicar la importancia de cada uno en los archivos que se están analizando.
- Métodos de lematización y sinónimos: Para obtener una mayor precisión en las búsquedas, la raíz de cada término se obtuvo utilizando la técnica de lematización con el algoritmo de Porter.

3.2.3. Machine Learning (Etapa 3).

Esta fase del proyecto dará cabida a la implementación de máquinas de aprendizaje supervisado con el fin de determinar cuál técnica de Machine Learning es la más pertinente para la detección de vocabulario referente a Ciberguerra. Para esto se desarrollaron códigos en Python, de los algoritmos: Árboles de Decisión, Naive Bayes, K-ésimo vecino más cercano, Redes Neuronales, Máquinas de Soporte Vectorial, Random Forest y Ada boosting. Por último, con el objetivo de profundizar en la investigación, se construirá un algoritmo con aprendizaje profundo (del inglés Deep Learning) basado en Redes Neuronales. Cabe señalar, que los algoritmos son seleccionados a partir de los trabajos relacionados y sus resultados.

3.2.4. Optimización de los algoritmos de Machine Learning (Etapa 4).

A partir de los resultados obtenidos, y luego de la implementación en Jupyter Notebook de las técnicas de Machine Learning mencionadas en la fase anterior, se exponen las soluciones generadas y se comparan con otras investigaciones similares de manera que se pueda determinar cuáles son las más adecuadas para la presente investigación, y enfocando el análisis en las técnicas que presenten un porcentaje mayor de aceptabilidad.

Por otra parte, se presentarán propuestas a tener en cuenta para futuras investigaciones partiendo de los resultados de la investigación actual y los respectivos análisis realizados con trabajos de investigación relacionados.

3.3. Conjunto de datos de pruebas, población y muestra.

La población sobre la que se trabajó en la presente investigación, constituye 1108 páginas Web de Ciberguerra (**el listado de las Url se encuentra en el anexo A**), representando a todas las páginas del tema en español, debido a que el Crawler en sus

búsquedas después de las 1000 páginas localizadas, la mayoría es muy repetitiva. La muestra está conformada por un total de 1108 sitios Web (localizados y descargados) y 71 variables predictoras (definidas en el objetivo 2.4, pág. 43)), así también una variable categórica, utilizada como apoyo para la predicción, con los valores "Si" y "No". Para la clase que corresponde a las páginas Web que "Si" son de Ciberguerra, 842; y para la clase que corresponde a que las páginas Web, que "No" son de Ciberguerra, 266.

3.4. Algoritmos de aprendizaje supervisado desarrollados en Python para las pruebas.

Análisis y descripción de las máquinas con aprendizaje supervisado desarrolladas en Python para las pruebas de esta investigación.

- Arboles de Decisión.
- Naïve Bayes.
- El K-vecino más Cercano – Knn.
- Redes Neuronales.
- Redes Neuronales con aprendizaje profundo.
- Máquinas de Soporte Vectorial.
- Ada Boosting.
- Regresión Logística.
- Random Forest.

3.4.1. Árboles de Decisión.

Dentro de los métodos de aprendizaje supervisado se destacan los Árboles de Decisión del inglés Decision Trees (DTs) en torno a los cuales Maya (2018), manifiesta que en lo referente a la tecnología de inteligencia Artificial (IA), el método de árbol de decisiones

“es usado para la parametrización y desarrollo de aplicaciones con el fin de indicarle a un programa cómo comportarse frente a una situación que se le presente” (p 17). En este mismo sentido, según lo planteado por Pedregosa et al., (2011), es un método de aprendizaje supervisado no paramétrico que se usa para la clasificación y la regresión, y cuyo objetivo es crear un modelo que prediga el valor de una variable objetivo luego de haber aprendido de reglas de decisión simples inferidas de las características que tienen los datos.

Pedregosa et al., (2011), destaca que dentro de las principales ventajas de los árboles de decisión están la facilidad de entender e interpretar, gracias a que pueden ser visualizados, además requiere poca preparación de datos aunque no admite valores faltantes, tiene la capacidad de manejar datos numéricos y categóricos, es capaz de manejar problemas de salida múltiple y funciona bien incluso si sus supuestos son violados de alguna manera por el modelo verdadero a partir del cual se generaron los datos.

Por otra parte, Bouza & Santiago (2012), señalan que esta técnica permite:

- Segmentación: establecer que grupos son importantes para clasificar un cierto ítem.
- Clasificación: asignar ítems a uno de los grupos en que está particionada una población.
- Predicción: establecer reglas para hacer predicciones de ciertos eventos.
- Reducción de la dimensión de los datos: Identificar que datos son los importantes para hacer modelos de un fenómeno.
- Identificación-interrelación: identificar qué variables y relaciones son importantes para ciertos grupos identificados a partir de analizar los datos.

- Recodificación: discretizar variables o establecer criterios cualitativos perdiendo la menor cantidad posible de información relevante. (p. 66)

Hssina, Merbouha, Ezzikouri, & Erritali (s,f), señalan que ID3 y C4.5 introducidos por J.R Quinlan, como dos importantes algoritmos para producir árboles de decisión razonables. Resaltando sus ventajas, desventajas y haciendo comparaciones entre ellos. Entre las cuales se destacan las siguientes:

- ID3: basado en el algoritmo Sistema de Aprendizaje Conceptual (CLS), el cual construye el árbol basado en la información obtenida de las instancias de entrenamiento y luego usa lo mismo para clasificar los datos de prueba. El algoritmo ID3 generalmente usa atributos nominales para clasificación sin valores faltantes.
- C4.5: fue hecho para superar las limitaciones del algoritmo ID3, para lo cual C4.5 utiliza "Ganancia de información" para evaluar el atributo de división. Los arboles de decisión se construyen en C4.5 usando un conjunto de datos de entrenamiento o conjuntos de datos como en ID3, y en cada nodo del árbol, C4.5 elige un atributo de los datos que más efectivamente divide su conjunto de muestras en subconjuntos enriquecidos en una clase u otra. Su criterio es la ganancia de información normalizada que resulta de elegir un atributo para dividir los datos. El atributo con la mayor ganancia de información normalizada se elige para tomar la decisión.

Pedregosa et al., (2011), destaca como algoritmos para arboles de decisión el C5.0 y CART, de los cuales mencionan lo siguiente:

- C5.0: es el último lanzamiento de la versión de Quinlan bajo una licencia patentada. Utiliza menos memoria y crea conjuntos de reglas más pequeños que C4.5 a la vez que es más preciso.
- CART: es muy similar a C4.5, pero difiere en que admite variables objetivo numéricas y no calcula conjuntos de reglas. CART construye árboles binarios utilizando la característica y el umbral que producen la mayor ganancia de información en cada nodo.

Además, Pedregosa et al., (2011), manifiesta que esta librería actualmente utiliza una versión optimizada del algoritmo CART; sin embargo, la implementación de scikit-learn no admite variables categóricas por ahora.

3.4.2. Naïve Bayes.

Šlibar (2019), señala a Naïve Bayes como un algoritmo de clasificación que se basa en la aplicación del teorema bayesiano, en el que cada atributo es una clase condicional independiente. En concordancia con lo anterior, Sammut y Webb (2017), establecen que Naïve Bayes es un algoritmo de aprendizaje simple que utiliza el teorema de Bayes junto con una fuerte suposición de que los atributos son condicionales independientes dada la clase. Mientras esta supuesta independencia es frecuentemente violada en la práctica. Naive Bayes sin embargo a menudo ofrece precisión de clasificación competitiva.

"Los modelos ingenuos de Bayes se pueden usar para abordar problemas de clasificación a gran escala para los cuales el conjunto de entrenamiento completo podría no encajar en la memoria. Para manejar este caso, MultinomialNB, BernoulliNB y GaussianNB, exponen un método partial_fit que puede usarse de forma incremental como se hace con otros

clasificadores como se demuestra en la clasificación de documentos de texto fuera del núcleo. Todos los clasificadores ingenuos de Bayes admiten ponderación de muestra" (Pedregosa et al., 2011).

En estadística y teoría de la probabilidad en torno al teorema de Bayes, Keshari (2018), plantea que este "describe la probabilidad de un evento, basado en el conocimiento previo de las condiciones que podrían estar relacionadas con el evento. Y sirve como una forma de averiguar la probabilidad condicional."

Keshari (2018), refiere que dada una Hipótesis (H) y evidencia (E), el teorema de Bayes relaciona la probabilidad de la hipótesis antes de obtener la evidencia, P (H), y la probabilidad de la hipótesis después de obtener la evidencia, P (H | E), de la siguiente manera:

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

Por otra parte, Sahami (como se citó en Larrañaga, Inza y Moujahid, s.f), presenta un algoritmo denominado k Dependence Bayesian classifier (kDB), este se fundamenta en el concepto de clasificador Bayesiano k-dependiente, el cual contiene la estructura del clasificador Naïve Bayes y permite a cada variable predictora tener un máximo de k variables padres sin contar a la variable clase. El pseudocódigo del algoritmo kDB puede consultarse en la Figura 9.

Paso 1. Para cada variable predictora X_i, $i = 1, \ldots, n$, calcular la cantidad de información mútua con respecto a la clase C, $I(X_i, C)$
Paso 2. Para cada par de variables predictoras calcular la cantidad de información mútua condicionada a la clase, $I(X_i, X_j|C)$, con $i \neq j$, $i, j = 1, \ldots, n$
Paso 3. Inicializar a vacío la lista de variables usada $\aleph$
Paso 4. Inicializar la red Bayesiana a construir, BN, con un único nodo, el correspondiente a la variable C
Paso 5. Repetir hasta que $\aleph$ incluya a todas las variables del dominio:
  Paso 5.1. Seleccionar de entre las variables que no están en $\aleph$, aquella X_{max} con mayor cantidad de información mútua respecto a C, $I(X_{max}, C) = \text{máx}_{X \notin \aleph} I(X, C)$
  Paso 5.2. Añadir un nodo a BN, $X \notin \aleph$ representando X_{max}
  Paso 5.3. Añadir un arco de C a X_{max} en BN
  Paso 5.4. Añadir $m = \text{mín}(|\aleph|, k)$ arcos de las m variables distintas X_j en $\aleph$ que tengan los mayores valores $I(X_{max}, X_j|C)$
  Paso 5.5. Añadir X_{max} a $\aleph$
Paso 6. Computar las probabilidades condicionadas necesarias para especificar la red Bayesiana BN

Figura 9. Pseudocódigo del algoritmo kDB (Sahami, 1996).

3.4.3. El K-ésimo vecino más cercano.

El vecino más cercano (K-nn), es un método que utiliza ciertos principios con el fin de clasificar un objeto en una clase distinta cuando hay dos o más grupos de objetos de clase conocida (Vidueira, Souza, Miranda, & Figueiredo, 2015). Pedregosa et al., (2011), aclara que para esto se utiliza un principio que consiste en "encontrar un número predefinido de muestras de entrenamiento más cercanas en distancia al nuevo punto y predecir la etiqueta a partir de ellas". Además, manifiestan que dicho número de muestras puede "ser una constante definida por el usuario (k-aprendizaje vecino más cercano), o puede variar en función de la densidad local de puntos (aprendizaje vecino basado en el radio)", asimismo, "la distancia puede ser, en general, cualquier medida métrica: la distancia euclidiana estándar es la opción más común".

De acuerdo con esto Miller y Miller, 2005 citados por Vidueira et.al (2015), manifiestan que este método "se basa en el concepto de proximidad y no presupone la

distribución en las clases” y que “esta técnica de similitud supone que cuanto más cerca se encuentran los objetos en el espacio de medición, más probable es que pertenezcan a la misma categoría o sean similares con respecto a las variables en estudio”. Pedregosa et al., (2011), declaran que los “métodos basados en los vecinos se conocen como métodos de aprendizaje automático no generalizados, ya que simplemente "recuerdan" todos sus datos de entrenamiento (posiblemente transformados en una estructura de indexación rápida como un árbol de bolas o un árbol KD)”.

El tipo de aprendizaje basado en instancias o aprendizaje no generalizador según lo explica Pedregosa et al., (2011), almacena las solicitudes de los datos de capacitación, agregando que la “clasificación se calcula a partir de un voto de mayoría simple de los vecinos más cercanos de cada punto: a un punto de consulta se le asigna la clase de datos que tiene la mayor cantidad de representantes dentro de los vecinos más cercanos del punto”. Dejando claro que:

> La clasificación básica de vecinos más cercanos utiliza pesos uniformes: es decir, el valor asignado a un punto de consulta se calcula a partir de un voto mayoritario simple de los vecinos más cercanos. En algunas circunstancias, es mejor pesar a los vecinos de manera que los vecinos más cercanos contribuyan más al ajuste. Esto se puede lograr a través de la palabra clave weights. El valor predeterminado, asigna pesos uniformes a cada vecino, a su vez, asigna pesos proporcionales al inverso de la distancia desde el punto de consulta. Alternativamente, se puede suministrar una función de distancia definida por el usuario para calcular los pesos.weights = 'uniform'weights = 'distance' (Pedregosa et al., 2011).

La Figura 10, muestran la clasificación basada en los vecinos más cercanos, haciendo diferencia entre clasificación con peso uniforme y distancia.

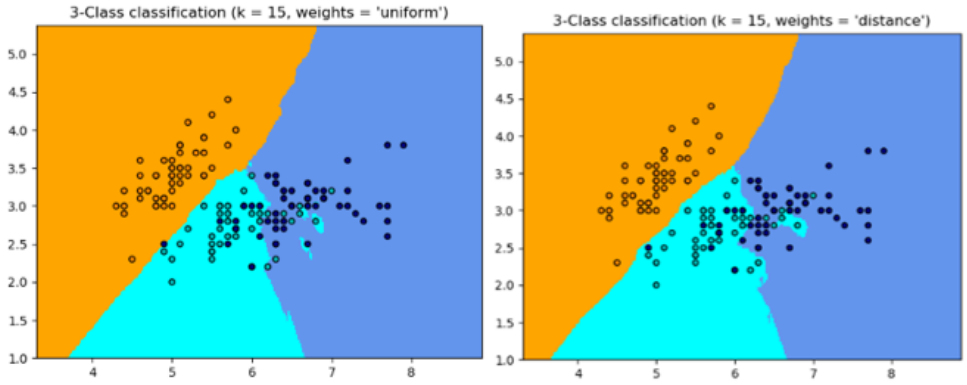


Figura 10. Ejemplo de clasificación de vecino más cercano basado en el peso uniforme y la distancia. Recuperado de: Pedregosa et al. (2011).

3.4.4. Redes Neuronales.

Como el método de aprendizaje del cerebro humano, es el que inspira la estructura de las Redes Neuronales, es pertinente y necesario que se presente primero algunos conceptos básicos en torno a las neuronas, para luego dar cabida a la comparación con la forma en la que funcionan los algoritmos basados en las redes neuronales.

Adserá (2010), refiere que las neuronas son las células del sistema nervioso que transmiten los impulsos nerviosos, están conectadas entre ellas formando circuitos neuronales. Cada neurona recibe impulsos nerviosos por sus dendritas, que son las prolongaciones ramificadas que salen de su cuerpo. El axón de la neurona es la prolongación larga que transmite la excitación neuronal y que permite enviar la información nerviosa a otras neuronas. El cuerpo y el axón de la neurona está recubierto de unas células de soporte llamadas células gliares, que están rellenas de mielina, una substancia grasa que hace de

aislante y permite que los impulsos nerviosos circulen mejor por el axón. Como se observa en la Figura 11.

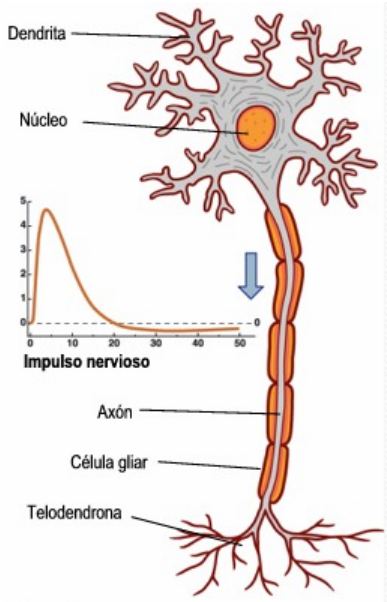


Figura 11. Representación de una neurona biológica. Recuperado de: Adserá (2010).

> Cuando un impulso nervioso llega hasta el final del axón (el telodendrona), provoca que se segreguen neurotransmisores. Los neurotransmisores son las substancias químicas que permiten transmitir la excitación nerviosa de una neurona a otra, a través de las sinapsis. Una sinapsis es la pequeñísima separación que hay en la zona de contacto entre la telodendrona de la neurona emisora y la dendrita de la neurona receptora, espacio en el que se segregan y trasladan los neurotransmisores. (Adserá, 2010).

Existe una estrecha relación entre la estructura de una neurona biológica y una neurona artificial. Cabezón (2018), señala que las neuronas artificiales que se interconectan en diversas redes son las encargadas del aprendizaje, recibiendo y combinando señales con otras neuronas. A través de las dendritas es posible la transmisión de información entre estas

neuronas, mediante la sinapsis y si la señal combinada es lo suficientemente fuerte, el nervio libera neurotransmisores. Según el tipo de neurotransmisor, las neuronas pueden excitarse si reciben el estímulo o inhibirse si no llega, generando una respuesta u otra según cada caso. En la Figura 12, se observa la representación de una neurona artificial.

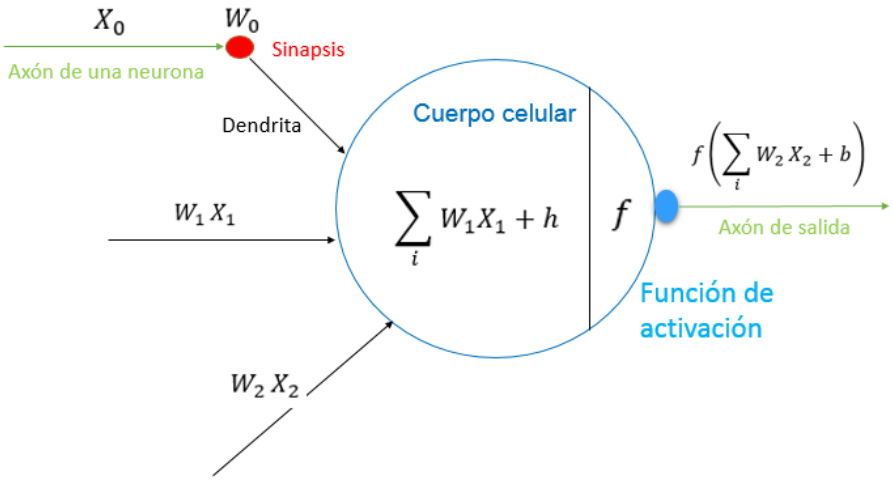


Figura 12. Representación de una neurona artificial. Adaptado de Cabezón (2018)

Cabezón (2018), expresa que existen distintas clasificaciones de Redes Neuronales, sin embargo, hace énfasis en la clasificación según su topología y la clasificación según su algoritmo de aprendizaje.

Para la clasificación según su topología, se tiene en cuenta que las neuronas se agrupan para formar capas y que estas se unen formando Redes Neuronales, y se clasifican en:

- Redes Monocapa: Son aquellas con una sola capa. Para unirse se crean conexiones entre las neuronas de la capa. Entre las redes monocapa existen algunas en las que las neuronas pueden estar conectadas consigo mismas y se denominan autorecurrentes.

- Redes Multicapa: Las redes multicapa están formadas por varias capas de neuronas. Estas redes se pueden a su vez clasificar atendiendo a la manera en que se conectan sus capas.

La clasificación según su algoritmo de aprendizaje es muy parecida a la de Machine Learning, en las Redes Neuronales también se pueden distinguir el aprendizaje supervisado, aprendizaje no supervisado y aprendizaje reforzado.

Por otro lado, Pedregosa et al., (2011), destaca del modelo de Redes Neuronales el algoritmo Multi-layer Perceptron, que es un algoritmo de aprendizaje supervisado que aprende una función $f: R^m \rightarrow R^o$ para una entrenamiento en un conjunto de datos, donde m es el número de dimensiones de entrada, y/o es el número de dimensiones para la salida. Para un conjunto de características dadas $X = x_1, x_2, x_3, x_4, \ldots x_m$ y un objetivo y, puede aprender un aproximador de función no lineal para clasificación o regresión, entre la capa de entrada y la de salida, puede haber una o más capas no lineales, llamadas capas ocultas, lo cual se puede observar en la Figura 13, en la cual se esquematiza un MLP de una capa oculta con salida escalar.

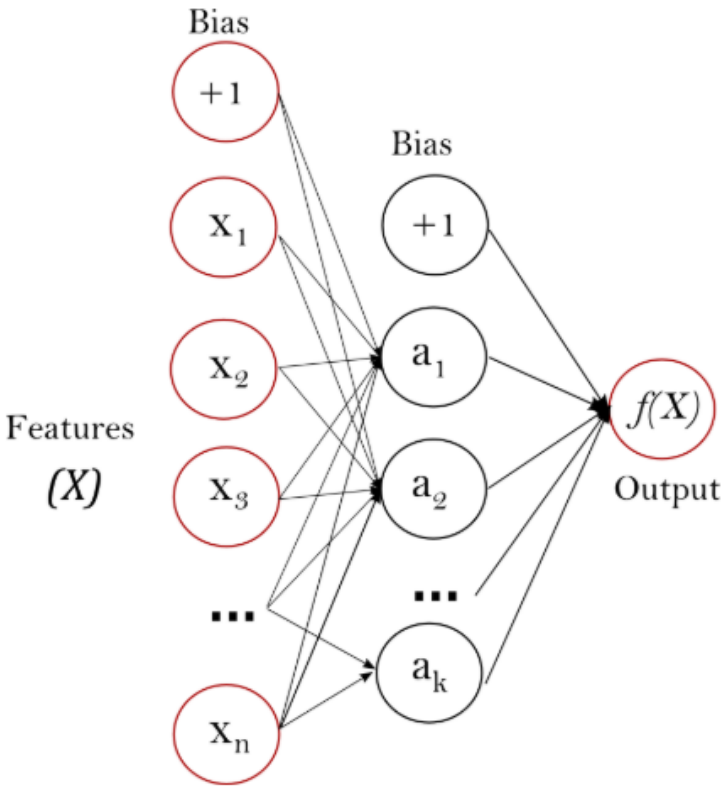


Figura 13. MLP de una capa oculta con salida escalar.
Recuperada de Pedregosa et al., (2011).

3.4.5. Máquinas de Soporte Vectorial.

Según lo que plantea Pedregosa et al., (2011), las Máquinas de Soporte Vectorial (SVM), son un conjunto de métodos de aprendizaje supervisado utilizados para la clasificación, regresión y detección de valores atípicos. Martínez (s.f), comenta que son una opción bastante usada para la clasificación, debido a que permiten encontrar la forma óptima de clasificar entre varias clases, maximizando el margen de separación entre las mismas, a través de los vectores de soporte que son los que definen el borde de esta separación del hiperplano que separa las clases. Los puntos multi-dimensionales se representan con vector de n dimensiones. Como puede apreciarse en la Figura 14.

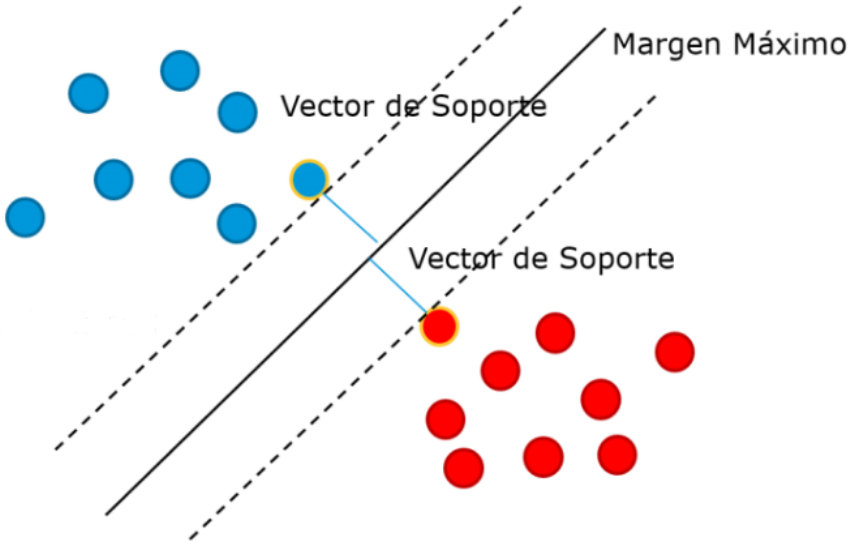


Figura 14. Ejemplo de clasificación de dos clases con vectores de soporte. Recuperado de Martínez (s.f).

Martínez (s.f) señala, además, que en el caso en que las clases no sean linealmente separables, es común hacer uso de el truco del Kernel. Cuando se hace referencia al Kernel conforme con lo que plantea Ferreira (2016), se habla de una función con la que se mapean los datos de entrada a otro espacio dimensional, es decir, que un set de entrenamiento x es mapeado a un nuevo set a partir de una operación, estos últimos son pasados al algoritmo de aprendizaje. Como se muestra en el ejemplo de la Figura 15.

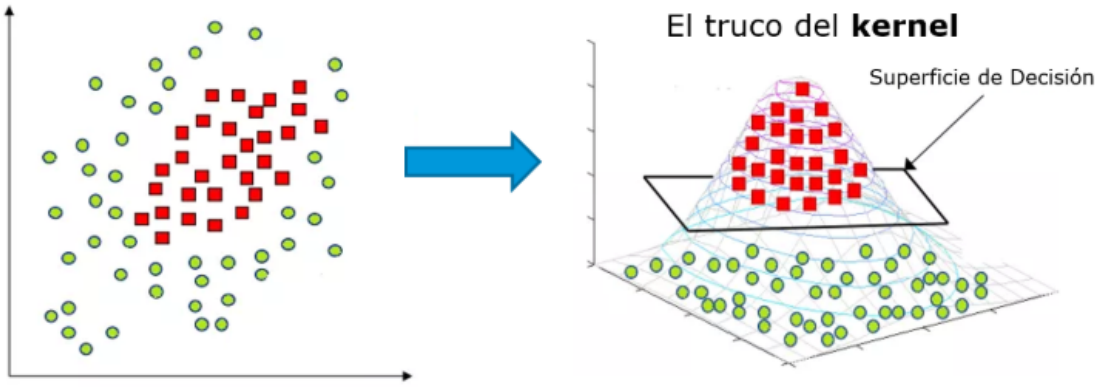


Figura 15. Ejemplo del truco del Kernel. Recuperado de Martinez (s.f)

Pedregosa et al., (2011), destaca algunas ventajas de las Máquinas de Soporte Vectorial, como su efectividad en espacios de altas dimensiones y en los casos donde el número de dimensiones es mayor que el número de muestras, eficiencia en la memoria debido a que utiliza un subconjunto de puntos de entrenamiento en la función de decisión y la versatilidad en cuanto a que se pueden especificar diferentes funciones de Kernel para la función de decisión.

Algunos algoritmos de Máquinas de Soporte Vectorial son usados con gran éxito para la clasificación o regresión de grandes datos. Gala (2013), enfatiza en Dual Coordinate Descent method (DCD) y Stochastic Sub-gradient Descent for Primal Problem (Pegasos). De los cuales destaca que DCD, es un método batch que trata mediante un método iterativo de descenso coordinado de encontrar la solución en el problema dual, mostrando con claridad las condiciones de convergencia. Pegasos por su parte, es un método online que resuelve el problema primal y en el que se alternan dos pasos, un método de descenso por subgradiente aproximado y una proyección sobre un subconjunto, afirmándose por consiguiente que el aprendizaje online es más rápido que el batch.

3.4.6. Ada Boosting.

Según lo que plantea Pedregosa et al., (2011), dentro de la clasificacipon de metodos de ensamble, se encuentra el algoritmo de refuerzo AdaBoost, introducido en 1995 por Freund y Schapire. El cual, es comúnmente usado tanto para la regresión como para la clasificación y se destaca, además, que su principio básico es ajustar una secuencia de modelos que son solo un poco mejores que las suposiciones aleatorias, como pequeños

árboles de decisión, en versiones de datos modificadas repetidamente. Combinando todas las predicciones mediante un voto mayoritario ponderado para producir la predicción final.

> AdaBoost, es un metaestimulador que comienza ajustando un clasificador en el conjunto de datos original y luego ajusta copias adicionales del clasificador en el mismo conjunto de datos, pero donde los pesos de las instancias clasificadas incorrectamente se ajustan de modo que los clasificadores posteriores se centren más en casos difíciles. Esta clase implementa el algoritmo conocido como AdaBoost-SAMME (Pedregosa et al., 2011).

En cuanto a ventajas y desventajas de este algoritmo, Emer (s.f), propone como ventajas que es rápido, simple y fácil de programar, no hay parámetros para sintonizar (excepto T), no se necesitan conocimientos previos sobre el alumno débil, probablemente eficaz dada la suposición de aprendizaje débil y versátil y en relación con sus oportunidades de mejora destaca que los clasificadores débiles, si son demasiado débiles puede conducir a bajos márgenes o sobreajustes y AdaBoost es particularmente vulnerable al ruido uniforme.

El algoritmo del modelo Adaboost se resume en el esquema que se observa en la Figura 16.

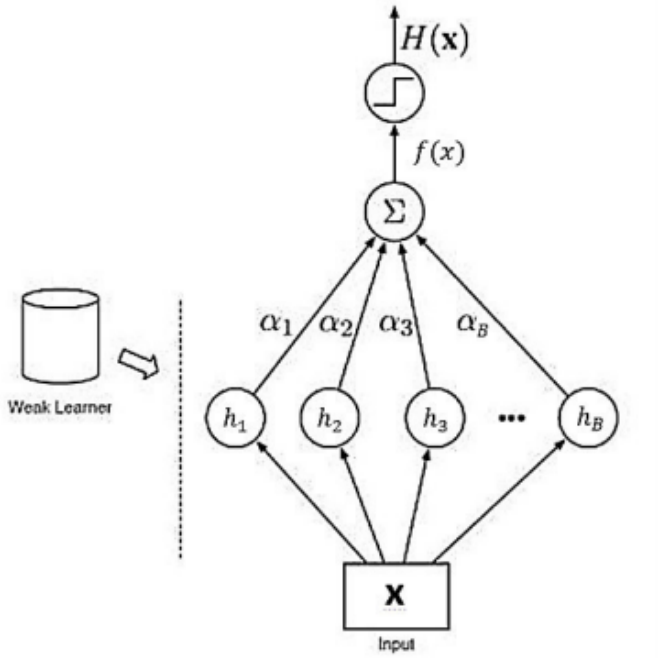


Figura 16. Diagrama de flujo del algoritmo del modelo adaboost.
Adaptado de la Figura 1 de Bailly y Milgram (2009).
Recuperado de Alarcón (2017).

Es importante destacar que existen algunas herramientas que permiten el mejor desarrollo de este modelo, dentro de los cuales se destacan, GradientBoostingClassifier y AdaBoostClassifier. Pedregrosa et al. (2011), señala que GradienBoostingClassiier construye un modelo aditivo de manera progresiva y los árboles de regresión se van ajustando en el gradiente negativo de la función de pérdida de desviación binomial o multinomial. Por otra parte, define al clasificador AdaBoost como "un metaestimulador que comienza ajustando un clasificador en el conjunto de datos original y luego ajusta copias adicionales del clasificador en el mismo conjunto de datos, pero donde los pesos de las instancias clasificadas incorrectamente se ajustan de modo que los clasificadores posteriores se centren más en casos difíciles". Esta clase implementa el algoritmo conocido como AdaBoost-SAMME.

3.4.7. Regresión Logística.

Según Fernandez (como se citó en Menes, Arcos, Moreno, & Gallegos, 2015), "El algoritmo de regresión logística, es un tipo de análisis estadístico orientado a la predicción de una variable categórica en función de otras variables consideradas como parámetros predictores". Por su parte, Silva (1994), plantea que la regresión logística es uno de los instrumentos estadísticos más expresivos y versátiles, y señala que tiene sus orígenes en 1961 con el trabajo de Cornfield, Gordon y Smith con una investigación sobre el riesgo de padecer una enfermedad coronaria y con el pasar de los años ha venido perfeccionándose con los aportes de otros investigadores como Walker y Duncan en 1967, abordando el tema de estimar la probabilidad de ocurrencia de un acontecimiento específico en función de varias variables y finalmente con los avances tecnológicos a partir de los 80 se expande gracias a los avances tecnológicos.

> Entre los usos más frecuentes de la regresión en los estudios biomédicos, está el de intentar describir cómo los valores de la variable dependiente están relacionados con el de la variable predictora o predictoras, el de intentar explicar cuáles son las predictoras de mayor interés y finalmente el intentar predecir, cuantificando su riesgo de aparición, los casos futuros (Ortega y Cayuela, s.f).

Un algoritmo de la regresión logística que se destaca para la clasificación binaria es el algoritmo de regresión logística bayesiana (BBR), que según plantean (Ortiz, Martín, Ureña, & García, 2005), su clave es la utilización de una distribución de probabilidad previa y algoritmos de optimización sucesiva de los ejemplos de entrenamiento suministrados. Este algoritmo inicialmente realiza una regresión logística de los datos de entrenamiento a partir

de la distribución de probabilidad elegida (Gausiana o Laplace), luego se va optimizando sucesivamente a través de la aplicación de un algoritmo de regresión logística en cadena. Se comienza poniendo todas las variables a algún valor inicial, y se busca qué valor de la primera variable minimiza la función objetivo, asumiendo que todas las otras variables mantienen constantes sus valores iniciales. El mismo método se lleva a cabo con la segunda variable, y así sucesivamente hasta que se han cruzado todas las variables. Este proceso se repite varias pasadas hasta encontrar un criterio de convergencia.

3.4.8. Random Forest.

Random forest, es un modelo de minería de datos que "surge como combinación de las técnicas de Classification And Regression Tree (CART) y Booststrap Aggregating (Bagging) para realizar la combinación de árboles predictores en la que cada árbol depende de los valores de un vector aleatorio probado independientemente y con la misma distribución para cada uno de estos (Alarcón, 2017, p.29).

Alarcón (2017), destaca que el modelo Random Forest aumenta la precisión en la clasificación a través de aleatoriedad en la construcción de cada clasificador individual, esta aleatorización puede introducirse en la partición del espacio, tanto en las etapas de entrenamiento, convalidación y de prueba.

Montillo y Ling (s.f), señalan como una característica importante del bosque aleatorio "que mientras crece cada árbol, se puede calcular un error de prueba construido a partir de las muestras de entrenamiento", Planteando que:

> Por cada árbol cultivado, aproximadamente el 36% de las muestras en el conjunto de entrenamiento no son seleccionado en bootstrap y se conoce como fuera de bootstrap

(OOB) muestras. Usando estos como entrada para el árbol correspondiente permite que el bosque aleatorio obtenga respuestas estimadas para ellos, como si no estuvieran entrenados muestras de prueba (Montillo & Ling, s.f, p.2).

El algoritmo del modelo Random Forest se resume en el esquema que se observa en la Figura 17.

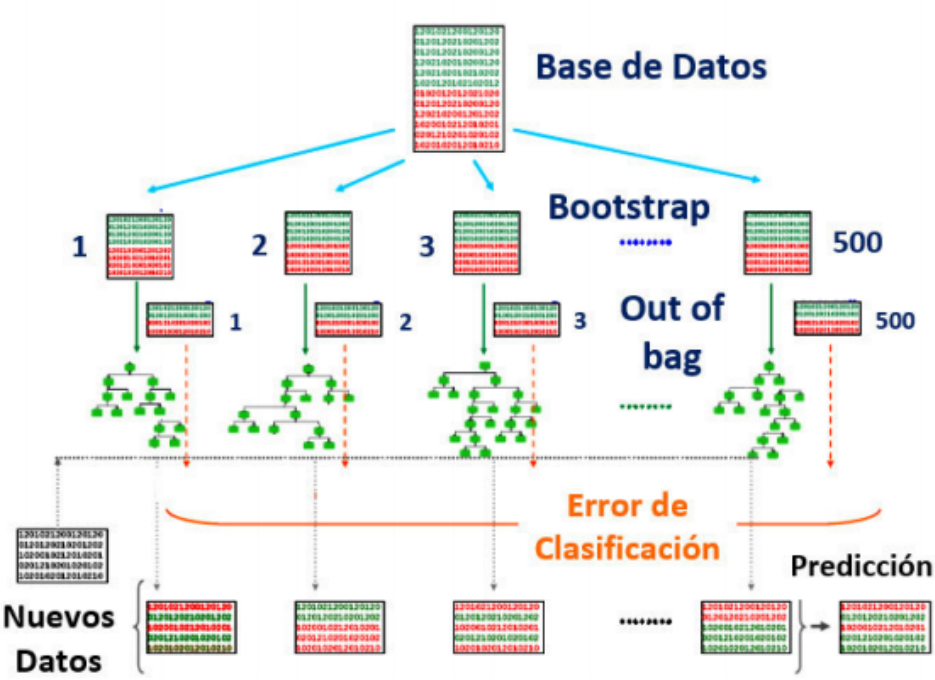


Figura 17. Diagrama de flujo del algoritmo Random Forest.
Adaptado de la Figura 1.2 de K. Hultstrom (2013).
Recuperado de Alarcón (2017).

3.4.9. Aprendizaje Profundo (Deep Learning).

Montañés, Aznar y Del Hoyo (2018), afirman que las Redes Neuronales con aprendizaje profundo han venido presentando una amplia variedad de soluciones con gran potencialidad en distintas áreas. Se han desarrollo diversos tipos de Redes Neuronales que presentan unas características óptimas para el aprendizaje de ciertos tipos de información, de los cuales resaltan las redes convolucionales (CNNs) y las Long Short-term Memory RNN

(LSTM). Las CNNs han demostrado su excepcional capacidad de aprendizaje en el área de visión por computador, para el reconocimiento y clasificación de objetos y en PLN, son capaces de extraer automáticamente los vectores de características sobre n-gramas mediante filtrado convolucional y las LSTM, que incorporan un novedoso mecanismo de memoria para el modelado de dependencias a largo plazo, lo que resuelve algunos de los problemas derivados del cálculo del gradiente en redes recurrentes más simples, convirtiéndose en una de las mejores opciones para el desarrollo de modelos de minería.

En los relacionado con la clasificación de imágenes las Redes Neuronales con aprendizaje profundo Arista, Calderón, Fierro y Nakano (2017), señalan que el "aprendizaje profundo consiste en realizar operaciones de convolución en una imagen, y resumir las características más relevantes, acentuando aquellas que tienen mayor preponderancia" (p. 214). Sostienen, además, que en una Red Neuronal convolucional, las neuronas se encuentran implementadas en una estructura bidimensional a manera de matriz, teniendo en cuenta que la imagen de entrada también se considera una retícula de neuronas, donde cada neurona corresponde a un pixel de la imagen. Cada neurona de una capa posterior cuenta con un rango de visión de la capa anterior, de tal manera que las neuronas de la capa anterior que presenten una activación adecuada contribuirán a la entrada de la capa posterior. El objetivo de dichas neuronas en la capa posterior es detectar una cierta característica en un área de la imagen.

Por otra parte, Arista et al. (2017), destacan que las Redes Neuronales convolucionales de aprendizaje profundo basan su eficacia en la cantidad de capas dentro de su topología. Entre más cantidad de capas posea una red convolucional, mejor será su desempeño en el reconocimiento, y su capacidad de clasificación.

En este mismo sentido, Santiago y Sánchez (2016), realizaron el diseño de una solución de seguridad informática integrando dos tecnologías: los sistemas multiagentes reactivos y BDI con las Redes Neuronales de aprendizaje profundo para la detección y contención de ataques informáticos tradicionales y avanzados. Para lo cual, en la capa de análisis, usaron el aprendizaje automático supervisado compuesto por dos Redes Neuronales multicapa de aprendizaje profundo – DNN discretas, compuestas por perceptrones con cinco capas de neuronas: una de entrada, tres ocultas y una de salida, como se puede visualizar en la Figura 18. La distribución de la cantidad de neuronas es el resultado del número de elementos Xi de los vectores de entrada y del cálculo del número de neuronas efectivas requeridas por capa considerando los umbrales de sobre entrenamiento e inefectividad del proceso de clasificación. Cada perceptron con una función de activación diferente le da un mayor grado de confiabilidad a la clasificación. Teniendo como objetivo clasificar el tráfico monitorizado como normal, ilegal o sospechoso de acuerdo a los datos recibidos en el Dataset.

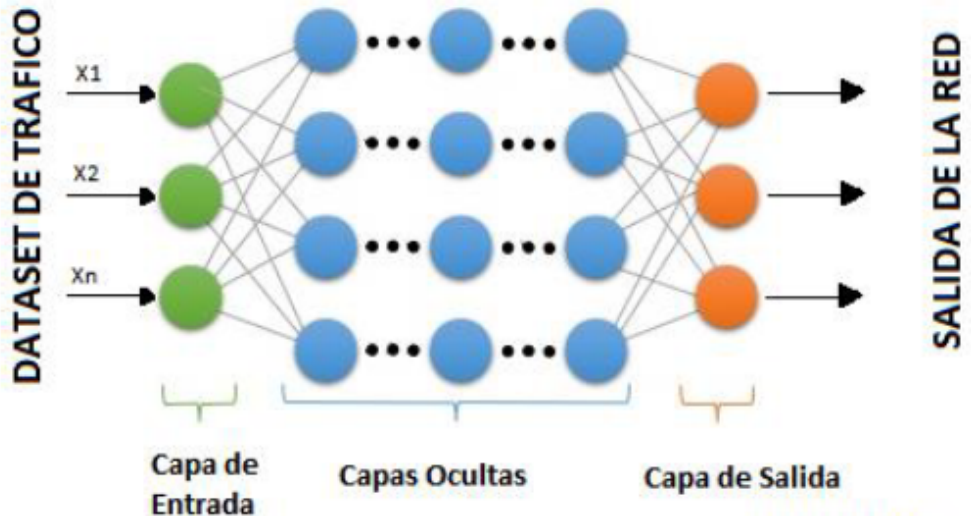


Figura 18. Esquema de capas de las Redes Neuronales.
Recuperado de Santiago y Sánchez (2016)

3.5. Procedimiento de pruebas.

El desarrollo e implementación de lo que se propone en esta investigación se llevó a cabo con el uso de un conjunto de herramientas que permitieron convertir la información contenida en más de 1000 páginas de Internet, en una base de datos que pudiera ser analizada para obtener valor agregado de las mismas, detectando vocabulario relacionado con Ciberguerra. Los fundamentos fueron basados en la implementación de la herramienta ADVI propuesta por Castillo et al. (2016).

En primer lugar, se estableció el corpus lingüístico de Ciberguerra, sustentado principalmente en los libros: "Ciberguerra" de Yolanda Quintana (2016) y "El quinto elemento: Espionaje, Ciberguerra y Terrorismo. Una amenaza real e inminente" de Alejandro Suarez Sánchez-Ocaña (2015). Lo cual permitió definir las variables predictoras que con el resto del desarrollo de la investigación serían fundamentales para cumplir los propósitos de la misma.

Para la limpieza y preprocesamiento de los datos se trabajó con las técnicas de Tokenización, Stop Word, TF, TF-IDF, métodos de lematización y sinónimos. Con los cuales, se da la segmentación de palabras provenientes de las páginas Web, se eliminan las palabras vacías, se calculan número de palabras y los tiempos que se repiten, se les asigna un peso y se obtiene la raíz de cada término. Asegurando que la información del Dataset se encuentre integra, y que no haya presencia de datos sucios, como filas vacías, datos alfanuméricos en variables numéricas, valores adicionales a lo esperado en la variable a predecir, entre otros. Todo lo anterior basado en la herramienta ADVI. Parte del Dataset se puede observar en las Figuras 19 y 20.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Tema	Subtema	Concepto	URL	URL_tit	guerra	ciberguer	amenaza	ciberataq	ataque	ciberdefe	atacar	terrorismo	daño	conflicto
2	ciberguer	ciberguer	ciberguer	3887	https://www.marxists.org/espanol/marigh/obras/mini.htm	406	0	1	0	33	0	24	14	21	2
3	ciberguer	ciberguer	ciberguer	3874	http://chiapas.laneta.org/desmilitarizacion/encuentro/ponencias/sandoval.htm	129	0	7	0	10	0	1	12	0	78
4	ciberguer	ciberguer	ciberguer	1529	https://www.monografias.com/trabajos14/histcomput/histcomput2.shtml	4	0	0	0	1	0	0	0	0	0
5	ciberguer	ciberguer	ciberguer	394	https://es.gizmodo.com/el-ejercito-de-estados-unidos-se-esta-preparando-para	3	0	1	1	11	0	0	0	0	1
6	ciberguer	ciberguer	ciberguer	1047	http://www.belt.es/noticias/especial/Ciberguerra/index.asp	85	46	19	38	57	18	4	50	0	3
7	ciberguer	ciberguer	ciberguer	3670	https://es.gizmodo.com/el-ataque-ddos-que-tumbo-medio-internet-es-un-adelan	9	8	0	0	29	0	7	0	0	0
8	ciberguer	ciberguer	ciberguer	3855	http://www.rs-labs.com/papers/tacticas/	6	0	1	0	72	0	46	0	0	0
9	ciberguer	ciberguer	ciberguer	853	https://es.gizmodo.com/descubren-un-nuevo-vehiculo-oculto-en-el-codigo-de-g	0	0	0	0	0	0	2	0	0	0
10	ciberguer	ciberguer	ciberguer	3428	http://www.razonypalabra.org.mx/anteriores/n36/jesteinou.html	2	0	1	0	3	0	0	1	0	1
11	ciberguer	ciberguer	ciberguer	832	http://clio.rediris.es/n40/articulos/alumnos_soldados.html	187	0	1	0	22	0	2	2	8	15
12	ciberguer	ciberguer	ciberguer	2306	https://www.um.es/docencia/barzana/II/II04.html	7	0	0	0	0	0	0	0	0	0
13	ciberguer	ciberguer	ciberguer	815	https://journals.openedition.org/polis/6306	88	0	4	0	0	0	0	5	0	16
14	ciberguer	ciberguer	ciberguer	3705	https://sistemas.uniandes.edu.co/es/foros-isis/temas-foros-isis/bpm/foro-2/90-	5	2	1	0	2	7	5	0	0	0
15	ciberguer	ciberguer	ciberguer	885	https://es.wikipedia.org/wiki/Guerra	241	0	0	0	1	0	1	2	0	62
16	ciberguer	ciberguer	ciberguer	1539	https://www.brookings.edu/es/articles/la-revolucion-de-la-robotica-y-el-conflic	72	0	0	0	10	0	4	3	0	6
17	ciberguer	ciberguer	ciberguer	18	https://www.muyinteresante.es/tecnologia/articulo/iesto-es-la-ciberguerra	5	3	5	0	11	0	3	0	0	0
18	ciberguer	ciberguer	ciberguer	976	https://www.gbm.net/es/de-cero-a-uno-para-convertirse-en-el-proximo-sun-tz	2	0	0	0	0	0	0	0	0	0
19	ciberguer	ciberguer	ciberguer	1869	https://es.slideshare.net/joelgomezmx/hacktivismo-ciberterrorismo-y-ciberguerr	17	6	3	2	36	0	4	19	0	2
20	ciberguer	ciberguer	ciberguer	1655	https://observatorio.cisde.es/archivo/la-amenaza-cibernetica-ciberguerra-y-cib	12	6	25	10	23	10	12	6	0	3
21	ciberguer	ciberguer	ciberguer	66	http://www.revistadon.com/16670/cinco-claves-sobre-la-ciberguerra-yolanda-	41	21	12	0	28	0	7	6	3	6
22	ciberguer	ciberguer	ciberguer	3858	https://www.redalyc.org/html/815/81501508/	145	0	3	0	7	0	1	2	0	9
23	ciberguer	ciberguer	ciberguer	421	http://www.capesic.cat/es/2017/05/31/como-italia-se-prepara-para-afrontar-el	3	0	12	0	20	0	1	58	0	2
24	ciberguer	ciberguer	ciberguer	63	http://www.iniseg.es/ciberseguridad/masteres-oficiales/master-ciberterrorismo	33	27	0	4	0	4	0	34	0	2
25	ciberguer	ciberguer	ciberguer	78	https://www.nytimes.com/es/2017/01/12/ciberguerra-a-la-venta/	0	3	2	0	3	0	5	7	0	2
26	ciberguer	ciberguer	ciberguer	295	https://laklave.wordpress.com/tag/guerra-cibernetica/	34	3	1	6	29	0	3	4	0	8
27	ciberguer	ciberguer	ciberguer	1581	http://www.fao.org/docrep/003/x7352s/X7352s02.htm	4	0	0	0	0	0	0	0	0	0
28	ciberguer	ciberguer	ciberguer	472	http://www.academia.edu/Documents/in/Ciberguerra	59	31	6	8	18	1	1	0	0	11
29	ciberguer	ciberguer	ciberguer	2193	https://moscovita.org/moscownews/venezuela-paisaje-politico-para-entender-el-	22	1	2	0	2	0	0	10	0	8

Figura 19. Dataset original con las páginas Web a analizar.

	A	B	C	D	E	F	G	H	I	J	K	L
1	guerra,ciberguerra,amenaza,ciberataque,ataque,ciberdefensa,atacar,terrorismo,dano,conflicto,anonymous,penetrar,ejercito,seguridad,ofensiva,delincuer											
2	406,0,1,0,33,0,24,14,21,2,0,9,12,29,3,6,104,4,0,1,0,24,0,0,0,1,1,0,4,37,8,0,2,0,5,0,1,0,0,2,0,33,0,1,0,5,17,1,13,33,16,1,87,29,1,29,27,0,1,0,7,2,0,0,0,0,0,0,13,0,Si											
3	129,0,7,0,10,0,1,12,0,78,0,0,47,68,15,2,6,1,0,1,7,103,0,3,15,9,0,0,48,0,0,5,2,0,13,0,5,0,0,1,0,4,0,0,3,34,3,7,16,75,49,10,10,107,0,32,27,0,5,0,5,4,1,2,0,3,0,0,0,0,Si											
4	4,0,0,0,1,0,0,0,0,0,0,1,2,13,0,0,0,0,0,2,0,114,0,1,12,25,13,5,1,0,0,262,2,30,27,0,11,0,0,0,0,224,2,0,25,0,0,0,2,6,4,1,0,1,0,16,1,0,2,0,18,34,0,1,0,2,0,0,0,0,Si											
5	3,0,1,1,11,0,0,0,0,1,0,0,5,0,0,0,1,1,0,1,0,479,0,1,0,1,0,137,0,0,0,0,0,0,0,0,15,4,3,0,0,2,0,0,0,0,0,0,0,0,0,0,0,1,2,0,0,3,0,0,0,0,1,0,3,0,0,0,0,0,0,Si											
6	85,46,19,38,57,18,4,50,0,3,1,1,0,153,1,0,1,0,0,3,0,14,21,0,0,23,12,8,4,0,0,4,0,1,7,0,0,1,1,10,0,9,0,0,6,0,0,0,0,6,0,0,4,18,0,3,0,9,15,2,1,0,8,0,2,1,0,0,5,0,Si											
7	9,8,0,0,29,0,7,0,0,0,0,0,1,9,0,0,0,0,0,0,0,0,478,0,0,1,13,0,0,0,0,0,0,0,0,8,1,0,2,3,0,2,0,4,0,0,0,0,0,0,0,1,5,0,0,0,0,0,0,0,0,4,0,0,0,1,0,0,0,1,0,0,0,Si											
8	6,0,1,0,72,0,46,0,0,0,0,0,0,15,1,0,0,0,0,0,9,89,0,0,0,10,4,8,0,0,0,0,26,153,24,0,2,0,0,0,0,20,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,0,2,0,0,34,0,4,0,0,6,9,0,0,0,Si											
9	0,0,0,0,0,0,2,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,486,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,3,0,0,0,3,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,Si											
10	2,0,1,0,3,0,0,1,0,1,0,1,1,5,0,0,0,0,0,5,7,34,18,6,10,80,6,1,74,0,0,10,1,1,7,0,51,0,0,5,0,20,2,0,2,1,10,4,11,4,6,25,0,0,0,14,1,0,0,0,5,21,1,5,0,0,0,0,0,0,Si											
11	187,0,1,0,22,0,2,2,8,15,0,0,20,8,9,0,30,3,0,0,0,12,0,0,1,0,0,6,11,0,0,0,0,0,4,0,2,1,1,1,0,6,0,0,0,17,0,0,9,4,22,0,22,12,0,4,6,0,1,0,12,1,0,0,0,0,0,0,0,0,Si											
12	7,0,0,0,0,0,0,0,0,0,0,0,0,3,11,0,0,0,0,0,4,0,70,0,4,33,56,6,19,1,0,0,21,8,3,92,0,4,0,0,0,0,28,9,0,3,0,0,0,0,1,7,0,0,1,0,8,1,0,4,0,4,12,1,2,0,1,0,0,0,0,Si											
13	88,0,4,0,0,0,0,5,0,16,0,0,0,31,0,0,8,3,0,2,7,16,0,0,0,1,0,0,21,0,0,0,3,0,6,0,3,1,1,0,0,10,0,0,4,7,0,50,4,9,53,4,4,14,0,4,9,0,2,0,3,4,0,0,0,5,0,0,0,0,Si											
14	5,2,1,0,2,7,5,0,0,0,0,0,0,84,0,0,0,0,0,0,0,0,20,0,13,13,13,35,5,7,0,0,90,0,0,2,0,2,0,2,9,0,33,2,0,2,0,0,0,0,7,8,2,2,0,0,5,0,0,9,2,0,6,0,1,0,0,0,0,0,0,Si											
15	241,0,0,0,1,0,1,2,0,62,0,0,3,0,0,0,0,5,0,0,1,0,0,0,0,0,0,2,0,1,5,0,0,0,0,0,0,2,0,1,0,0,0,0,5,0,0,3,0,0,1,3,1,12,0,4,2,0,11,14,0,0,0,0,1,0,0,0,1,0,0,0,0,Si											
16	72,0,0,0,10,0,4,3,0,6,0,0,0,10,5,1,0,10,2,0,1,1,45,0,0,0,1,2,5,0,0,0,11,0,0,4,0,3,7,7,0,0,29,74,0,0,1,0,0,0,1,3,0,10,16,0,6,6,0,10,0,2,2,0,2,0,0,0,0,0,0,Si											
17	5,3,5,0,11,0,3,0,0,0,0,0,1,10,6,0,0,0,0,0,0,28,4,1,5,13,6,126,0,0,0,0,0,10,6,0,5,8,18,0,0,3,16,0,0,1,0,0,1,4,7,0,0,0,0,0,0,2,1,0,4,1,0,0,1,3,1,0,39,0,Si											
18	2,0,0,0,0,0,0,0,0,0,0,0,0,0,5,0,0,0,0,0,1,0,13,0,0,0,1,57,115,39,0,0,5,2,2,0,0,0,13,14,3,0,2,14,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,7,2,0,0,0,0,0,0,59,1,Si											
19	17,6,3,2,36,0,4,19,0,2,1,5,0,10,0,5,1,0,0,3,4,11,2,3,42,12,2,16,5,0,0,15,1,4,0,0,0,2,2,3,0,28,0,0,3,1,0,1,1,8,8,3,0,2,0,9,0,4,1,3,6,10,8,0,0,1,0,0,0,0,Si											
20	12,6,25,10,23,10,12,6,0,3,0,2,0,39,2,12,0,0,0,4,0,10,12,13,13,6,2,5,4,0,0,5,2,2,4,0,15,0,0,2,0,27,0,0,0,0,0,0,1,2,5,0,0,1,0,3,2,0,1,0,8,10,5,0,0,0,0,0,0,0,Si											
21	41,21,12,0,28,0,7,6,3,6,1,0,5,7,3,7,2,3,0,1,4,25,8,0,9,15,7,8,4,0,0,0,2,1,4,0,0,3,0,3,0,10,8,0,1,5,0,0,1,0,4,0,2,4,0,11,3,2,2,0,0,0,5,0,2,1,0,0,0,0,Si											
22	145,0,3,0,7,0,1,2,0,9,0,1,35,4,5,0,11,2,0,2,0,15,0,0,0,0,0,0,0,4,0,0,0,0,0,0,6,0,3,0,0,0,0,4,0,0,0,0,0,0,0,1,3,7,0,4,15,1,2,6,0,2,0,0,0,0,0,0,4,0,0,0,0,Si											

Figura 20. Dataset con valores de variables predictoras en formato CSV.

Una vez que, el Dataset se encuentra en condiciones óptimas, es decir, que los datos se encuentran limpios, a través del preprocesamiento mencionado en el párrafo anterior se da paso a la realización de un análisis que permite explorar la información motivo de investigación, generando a partir de esta una descripción que permita obtener datos estadísticos básicos relacionados con el Dataset y a la implementación de los distintos modelos seleccionados para esta investigación.

3.5.1. Librerías.

El código se escribió usando Jupyterlab, una aplicación que permite ordenar y documentar el código y a la vez realizar ejecuciones por partes de las librerías. Para la implementación de cada modelo, se utilizó el lenguaje de scripting Python y la librería de Scikit – Learn de Machine Learning y otras librerías para tratamiento y computación de los datos como: Pandas, Numpy y la herramienta de gráficos Matplotlib. Las cuales serían de utilidad para el conteo del tiempo, separación de los conjuntos de entrenamiento y prueba, visualización de la matriz de confusión. En la Figura 21, se muestra el código de importación de las librerías necesarias para la implementación de los modelos a utilizar en la presente investigación. Por otro lado, las que son requeridas en particular para cada modelo serán especificadas en cada uno de ellos.

```
# Librerias
# *********************************************************

# Libreria para el tiempo
import time

# Libreria para algoritmos de machine learning
import sklearn

# Importamos el train_test_split para dividir los datos de entrenamiento y prueba
from sklearn.model_selection import train_test_split
# Libreria para exportar el gráfico generado del arbol de decisión
from sklearn.tree import export_graphviz

# Libreria para construir la matriz de confusión
from sklearn.metrics import confusion_matrix

# Libreria para ignorar los warnings
import warnings
warnings.filterwarnings("ignore", category=FutureWarning)

# Libreria para cargar el dataset o dataframe
import pandas as pd
import seaborn as sns

# Libreria para usar graficación de la matríz de confusión
import matplotlib.pyplot as plt
```

Figura 21. Librerías de Machine Learning en Python.

3.5.2. Cargue del Dataset y exploración de datos.

La Figura 22, muestra el código en lenguaje de programación Python que ejecuta el cargue del conjunto de datos y otros comandos que permiten visualizar la información con la cual se está trabajando, como: encabezados de los datos, dimensiones, variables y características utilizadas en el análisis.

```
# Inicia del contador de tiempo
inicio_tiempo=time.time()

# Carga el dataset
dataset = 'dataset_ciberguerra.csv'
df = pd.read_csv(dataset)

# Vistazo general del dataset
print("información\n", df.head())
print(df.head())

#Resumen de los datos
summary = df.describe()
arregloCorr= df[df.columns[-1]] # Especifica tomar variable a predecir (última posición)
print("Valores en Correcto\n",arregloCorr.value_counts())
summary = summary.transpose()
print(summary.head())

#Dimensiones
print("\n")
print ("Dimensión del Dataset: ")
print ("Registros: ", df.shape[0])
print ("Columnas:  ", df.shape[1])
```

```
#Variables
print("\n")
print ("Variables predictoras (Valores)")
arreglox = df[df.columns[:-1]].as_matrix() # Específica tomar variables predictoras
print (arreglox)
print("\n")
print ("Variable a predecir (Valores)")
arregloy= df[df.columns[-1]].as_matrix() # Específica tomar variable a predecir (última posición)
print (arregloy)

#Muestra las caracteristicas utilizadas en el análisis
print("\n")
print("Características utilizadas en el análisis")
print (df.keys())
```

Figura 22. Código para cargue del Dataset y código de exploración de datos.

Los resultados de los códigos anteriores ofrecen la posibilidad de reconocer y visualizar la información con la cual se está trabajando. Por una parte, los encabezados dan una mejor idea de los datos del Dataset, permitiendo mejorar la experiencia visual en torno

al contenido de este. La Figura 23 permite la visualización de los encabezados, mostrando información del Dataset en las 5 primeras filas.

El conjunto de datos con los cuales se realiza la presente investigación consta de un total de 1108 sitios Web localizados y descargados por el Web Crawler de la herramienta ADVI, con 71 variables predictoras (relacionadas en el corpus lingüístico), y en la variable categórica a predecir se incluyen los valores “Si” y “No”. Para la clase “Si” hay 842 y para la clase “No”, 266. En la Figura 23 y 24, es posible apreciar sus dimensiones, la distribución de la variable categórica, así como también un resumen estadístico que permite entender la escala de los valores y verificar posibles irregularidades que presenten los datos. Cabe señalar que la información que se muestra dichas figuras, presentan un fragmento del conjunto de datos.

DETECCIÓN DE VOCABULARIO DE CIBERGUERRA

información

	guerra	ciberguerra	amenaza	ciberataque	ataque	ciberdefensa	atacar	\
0	406	0	1	0	33	0	24	
1	129	0	7	0	10	0	1	
2	4	0	0	0	1	0	0	
3	3	0	1	1	11	0	0	
4	85	46	19	38	57	18	4	

	terrorismo	dano	conflicto	...	electronico	virus	satelites	antivirus	\
0	14	21	2	...	2	0	0	0	
1	12	0	78	...	4	1	2	0	
2	0	0	0	...	34	0	1	0	
3	0	0	1	...	1	0	3	0	
4	50	0	3	...	0	8	0	2	

	dominio	dns	paypal	visa	mastercard	correcta
0	0	0	0	13	0	Si
1	3	0	0	0	0	Si
2	2	0	0	0	0	Si
3	0	0	0	0	0	Si
4	1	0	0	5	0	Si

[5 rows x 71 columns]

Figura 23. Vista general del Dataset.

```
Valores en Correcto
 Si    842
No    266
Name: correcta, dtype: int64
             count      mean        std  min  25%  50%  75%    max
guerra      1108.0  6.359206  18.413474  0.0  0.0  3.0  7.0  406.0
ciberguerra 1108.0  1.415162   3.334827  0.0  0.0  0.0  2.0   46.0
amenaza     1108.0  0.846570   2.148068  0.0  0.0  0.0  1.0   25.0
ciberataque 1108.0  0.639892   2.131053  0.0  0.0  0.0  0.0   38.0
ataque      1108.0  2.914260   5.888520  0.0  0.0  0.0  3.0   72.0


Dimensión del Dataset:
Registros:  1108
Columnas:   71


Variables predictoras (Valores)
[[406   0   1 ...   0  13   0]
 [129   0   7 ...   0   0   0]
 [  4   0   0 ...   0   0   0]
 ...
 [  0   0   0 ...   0   0   0]
 [  0   0   0 ...   0   0   0]
 [  0   0   0 ...   0   0   0]]


Variable a predecir (Valores)
['Si' 'Si' 'Si' ... 'No' 'No' 'No']
```

Figura 24. Resumen de los datos y variables.

3.5.3. Definición de los conjuntos de entrenamiento y prueba.

El conjunto de datos fue dividido en un conjunto de entrenamiento y un conjunto de prueba. Según lo que plantea Castillo (s.f), la muestra tomada para entrenamiento permite determinar los parámetros del clasificador y el conjunto de prueba se utiliza para estimar el error de generalización. Por lo que para este estudio se tomó para el conjunto de entrenamiento un 80% de los datos y para el conjunto de prueba un 20%. La Figura 25, muestra la preparación de los datos, en cuanto a la separación de las muestras que serán usadas para entrenamiento y las que serán usadas para pruebas, esto será una condición importante que permitirá posteriormente el desarrollo cada una de las distintas máquinas de aprendizaje supervisado.

```
# Especificación del modelo de datos para entrenamiento
X_train, X_test, y_train, y_test = train_test_split(arreglox , arregloy,test_size = 0.2)
# se dividen los datos, 80% para el entrenamient (X_train) y 20% para pruebas (x_test)
# X_train contiene las variables predictoras, y_train las variables a predecir
# X_test contiene las variables predictoras, y_test las variables a predecir

print("\n")
print ("Entrenamiento: ",len(X_train), "registros (80%)")
print ("Pruebas:       ",len(X_test), "registros (20%)")
```

Figura 25. Definición de los conjuntos para entrenamiento y pruebas.

El conjunto de prueba (test o conjunto de generalización), se utiliza para estimar el error de generalización de la muestra ya que el objetivo final, es que el clasificador consiga un error de generalización pequeño evitando el sobreajuste (o sobre - entrenamiento), que consiste en una sobrevaloración de la capacidad predictiva de los modelos obtenidos: en esencia, no tiene sentido evaluar la calidad del modelo sobre los datos que han servido para construirlo ya que esta práctica nos lleva a ser demasiado optimistas acerca de su calidad, Castillo (s.f).

Se desataca, además, que todas las máquinas de aprendizaje fueron entrenadas con el mismo conjunto de datos de entrenamiento y probadas con el mismo conjunto de pruebas establecido previamente.

3.6. Descripción de las pruebas con los algoritmos seleccionados.

Seguidamente se desarrollaron las pruebas con los distintos modelos predictivos de Machine Learning y Deep Learning, propuestos en la investigación, los cuales incluyen Arboles de Decisión, Naive Bayes, k-ésimo vecino más cercano, Redes Neuronales, Redes Neuronales con Aprendizaje Profundo, Maquinas de Soporte Vectorial, Adaboost, Regresión Logística y Random Forest. Lo cual brinda una mayor posibilidad de encontrar un algoritmo que favorezca con un gran porcentaje de acierto la predicción de vocabulario de Ciberguerra,

ya que se parte de probar con una amplia gama de algoritmos. A continuación, se muestra una descripción de cada uno de los modelos implementados en las pruebas de detección de vocabulario de Ciberguerra en los sitios de Web, describiendo el algoritmo y a su vez exponiendo los comandos y estructura.

3.6.1. Árboles de Decisión.

Para el desarrollo del modelo de Árboles de Decisión, se cargan las librerías DecisionTreeClassifier, como se aprecia en la Figura 26.

```
# Libreria para utilizar el árbol de decisión
from sklearn.tree import DecisionTreeClassifier
```

Figura 26. Importe de la librería DecisionTreeClassifier.

Posteriormente, se inicia el contador de tiempo y se carga el Dataset, se muestran características básicas de los datos y se específica la información para entrenamiento y prueba, como se muestra en la Figura 27. Se establece la variable 'Model', que permitirá el cargue del modelo, así mismo, se establece el modelo para el conjunto de entrenamiento, se obtiene la predicción, la matriz de confusión y se termina el conteo del tiempo. Por último, se da paso a la visualización de la construcción y entrenamiento del modelo con el algoritmo mencionado y se valida su precisión con el conjunto de datos de prueba, finalizando el proceso con la construcción de la matriz de confusión mostrando aciertos y errores.

```
# Establece en el modelo los registros para entrenamiento
Model =arbol.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(arbol.score(X_train, y_train)),"%")
print("\n")

# Obtiene la predicción
y_pred = Model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("--------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                  {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("--------------------------------------------------------------------------------------------")
```

Figura 27. Modelo Arboles de decisión.

3.6.2. Naïve Bayes.

Para el desarrollo de este modelo además de cargar las librerías mencionadas al inicio del capítulo, se carga la librería GaussianNB, la cual específica que permitirá correr el modelo de Naïve Bayes, como se observa en la Figura 28.

```
# Libreria para utilizar naive bayes
from sklearn.naive_bayes import GaussianNB
```

Figura 28. Importe de la librería GaussianNB.

Luego de esto, con el algoritmo se inicia el conteo del tiempo, la ubicación del Dataset en el equipo, datos generales del Dataset y especificaciones de los conjuntos usados para entrenamiento y prueba, como se observa en la Figura 29. Se define la variable que cargará el modelo, que en este caso es la variable 'bayes'. Luego de esto se da paso a la definición de las variables para calcular el porcentaje de aprendizaje del algoritmo, calcular la predicción y construcción de la matriz de confusión con los aciertos y fallos del modelo en la predicción de vocabulario de Ciberguerra y finalmente, se presentan los comandos que permitirán la visualización de los resultados.

```
# Variable bayes carga el MODELO del algoritmo Naive Bayes
bayes = GaussianNB()

# Establece en el modelo los registros para entrenamiento
Model= bayes.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(bayes.score(X_train, y_train)),"%")

# predicción
y_pred = Model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("-------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                  {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("-------------------------------------------------------------------------------------------")
```

Figura 29. Modelo Naïve Bayes.

3.6.3. El K-Vecino más Cercano – Knn.

Para la implementación del modelo Knn – vecino más cercano, se cargan todas las librerías que se han mencionado para los modelos anteriores y en específico la librería KneighborsClassifier, que será la que permitirá el cargue del mismo, como se observa en la Figura 30.

```
# Libreria para utilizar K-nn vecino más cercano
from sklearn.neighbors import KNeighborsClassifier
```

Figura 30. Importe de la librería KneighborsClassifier.

Luego de iniciar el contador de tiempo, se carga el Dataset, se muestran algunas generalidades de los datos y de las variables predictoras, y se especifican los datos para el entrenamiento, como se observa en la Figura 31. Se da paso a la definición de la variable 'knn' que carga el modelo, haciéndole una relación con nueve vecinos y se establece 'Model', como el modelo de registros para entrenamiento, de acuerdo con los conjuntos de entrenamiento y prueba definidos previamente. A partir de aquí se da paso la verificación del porcentaje de aprendizaje del algoritmo, y a definir las variables que generarán la predicción y la matriz de confusión y luego de que se finalice el conteo del tiempo. Los comandos apuntarán a mostrar los resultados.

```
# Variable knn carga el MODELO del algoritmo K-nn Vecino más cercano
knn= KNeighborsClassifier(n_neighbors=9) #Indica relación con 9 vecinos

# Establece en el modelo los registros para entrenamiento
Model= knn.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(knn.score(X_train, y_train)),"%")

# predicción
y_pred = Model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("--------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                    {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("--------------------------------------------------------------------------------------------")
```

Figura 31. Modelo Knn - vecino más cercano

3.6.4. Redes Neuronales.

Existen distintos tipos de Redes Neuronales, en este caso se trabaja con las Redes Neuronales Multicapa, por lo cual, al momento de cargar las librerías correspondientes para el conteo del tiempo, visualización, carga del Dataset, entre otras. MLPClassifier específica que se va a usar la Red Neuronal Perceptron Muticapa, lo cual se puede visualizar en la Figura 32.

```
# Especificaciones para usar la Red Neuronal Perceptron Muticapa
from sklearn.neural_network import MLPClassifier
```

*Figura 32.*Importe de la librería MLPClassifier.

Luego de esto se determinan las variables que permitirán el inicio del conteo del tiempo, el cargue del Dataset, la visualización de datos generales y descriptivos del Dataset y de las variables predictoras, así mismo se determinan las condiciones del modelo de datos para entrenamiento y pruebas. A partir de esta parte del algoritmo, como se muestra en la Figura 33, se define la variable 'red' la cual determinará el cargue del modelo de redes neuronales, especificando, además, un máximo de 1000 iteraciones y 4 capas ocultas. Posteriormente, se específica el modelo de datos con el cual se clasificarían las pruebas y se muestran el porcentaje de aprendizaje del modelo, así como también el ajuste de los datos de entrenamiento. Para luego definir las variables que permitirán medir la precisión del modelo y construir la matriz de confusión del entrenamiento y de las pruebas y se finaliza el conteo del tiempo, para dar cabida a mostrar los resultados.

```
# Variable red carga el modelo de Red Neuronal, parametros (iteraciones y capas ocultas en la red)
# El modelo se ajusta hasta que se llegue a una solucion óptima
red=MLPClassifier(max_iter=1000, hidden_layer_sizes=(4))

# Establece en el modelo número de registros para entrenamiento y para pruebas
modelo = red.fit(X_train,y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format (red.score(X_test,y_test)),"%")
print("\n")

# Ajusta el modelo a los datos de entrenamiento
print ("Ajusta el modelo de datos al proceso de entrenamiento")
#print (model.fit(X_train, y_train))

#Matríz de confusión de pruebas
y_pred = modelo.predict(X_test)
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("---------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Porcentaje de precisión:   {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:  {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de: {0:.2f}"
      .format(tiempo),"seg")
print ("---------------------------------------------------------------------")
```

Figura 33. Modelo Redes Neuronales.

3.6.5. Máquinas de Soporte Vectorial.

Para el desarrollo del modelo Maquinas de Soporte Vectorial, se cargan las librerías mencionadas al inicio del presente capítulo y en específico para este modelo la librería SVC, como se observa en la Figura 34.

```
# Libreria para trabajar con Máquinas de Soporte Vectorial
from sklearn.svm import SVC
```

Figura 34. Importe de la librería SVC

Luego de esto, se inicia el contador de tiempo y se carga el Dataset, se muestran algunas características básicas de los datos y se especifican los datos para entrenamiento y prueba. Se da pasó a establecer la variable 'model' que permitirá el cargue del modelo, así mismo, se establece el número de muestras para el conjunto de entrenamiento y se verifica el porcentaje de aprendizaje del algoritmo. Para dar paso a la definición de las variables con las cuales se obtiene la predicción, la matriz de confusión, y se termina el conteo del tiempo. Finalmente se da paso a la visualización de la construcción y entrenamiento del modelo con el algoritmo de Maquinas de Soporte Vectorial, se valida su precisión con el conjunto de datos de prueba, para finiquitar el proceso con la construcción de la matriz de confusión con aciertos y errores. En la Figura 35, puede observarse con detalles todo el proceso descrito en este párrafo.

Para la optimización de las Máquinas de Soporte Vectorial, se utiliza el parametro "C", el cual le indica cuanto desea clasificar mal un dato. El proceso consiste, en si el valor de C es más pequeño que (10), elige un hiperplano de separación de mayor margen haciendo un mejor trabajo de clasificación. Por otro lado, cuando C, es más grande elige un hiperplano de separación de menor margen, generando un mayor número de errores de clasificación.

```
# Optimización (Parametro de C, le indica cuanto desea evitar clasificar mal un dato)
# Cuando el valor de C es más pequeño, elige un hiperplano de separación de mayor margen
# haciendo un mejor trabajo de clasificación
maquina = SVC(C=1)
# Variable model carga el modelo entrenado de la Máquina de Soporte Vectorial
model = maquina.fit(X_train, y_train)
# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(model.score(X_test,y_test)),"%")
print("\n")
# Matriz de confusión de entrenamiento
y_pred = model.predict(X_test)
# Matriz de confusión de pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)
#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("--------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                    {0:.1f}"
      .format(porcentaje * 100),"%")
print("Precisión optimizada (C=10):  {0:.1f}"
      .format(maquina.score(X_test, y_test) * 100),"%   **Optimiza la clasificación generando ")
print ("                                        un hiperplano de separación de mayor margen")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("--------------------------------------------------------------------------------------------")
```

Figura 35. Modelo Máquinas de Soporte Vectorial.

3.6.6. Ada Boosting.

Para la implementación del modelo Ada Boosting, se cargan las librerías necesarias mencionadas al inicio del presente capítulo, las cuales permitirán el conteo del tiempo, la visualización de resultados, entre otros y en específico, como se muestra en la Figura 36, se cargan GradientBoostingClassifier, AdaBoostClassifier, accuracy_score y xgboost, que en conjunto permiten gran precisión para el modelo.

```
# Librerias para utilizar AdaBoost (Boosting)
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn.metrics import accuracy_score
import xgboost as xgb   # Se debe instalar pip install xgboost
```

Figura 36. Importe de librerías para desarrollar Ada Boosting.

Luego de esto el algoritmo inicia el conteo del tiempo, se carga el Dataset, se permite la visualización de algunas generalidades del Dataset y así mismo se especifican las condiciones para la selección de los conjuntos de entrenamiento y prueba. Posteriormente, como se puede apreciar en la Figura 37, se define la variable 'ada' que será con la cual se cargue el modelo de Ada Boosting, y se estipula 500 como número máximo de estimadores en los que finaliza el refuerzo, una taza de ganancia de 1.5 y un generador de números aleatorios de 1. Seguidamente, se establece el modelo de los registros para entrenamiento, se verifica el porcentaje de aprendizaje para el algoritmo y se definen las variables que darán paso a la predicción y a la construcción de la matriz de confusión, se finaliza el conteo del tiempo y finalmente se da visualización de los resultados.

```
# Variable ada carga el MODELO del algoritmo adabooost
ada = AdaBoostClassifier( n_estimators=500,learning_rate=1.5, random_state=1)

# Establece en el modelo los registros para entrenamiento
model= ada.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(ada.score(X_train, y_train)),"%")
print("\n")

# predicción
y_pred = model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("--------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                    {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("--------------------------------------------------------------------------------------------")
```

Figura 37. Modelo Ada Boosting.

3.6.7. Regresión Logística.

Para poner en funcionamiento el modelo de regresión logística se requiere importar la librería LogisticRegression, como se observa en la Figura 38, además del importe del resto de las librerías necesarias para el conteo del tiempo, construcción y visualización de la matriz de confusión, entre otras.

```
# Libreria para utilizar regresión logistica
from sklearn.linear_model import LogisticRegression
```

Figura 38. Importe de la librería LogisticRegression

Luego de realizar el importe de las librerías mencionadas se inicia el proceso de definición de variables y condiciones para implementar el modelo de Regresión Logística. Para lo cual una vez se inicia el conteo del tiempo, se carga el Dataset, se muestran características generales de los datos, las variables con las que se está trabajando y se especifican las condiciones para los conjuntos de entrenamiento y prueba. Seguidamente como se observa en la Figura 39, se define la variable 'logisticr' con la cual se carga el modelo, se establece el modelo para el conjunto de entrenamiento, y se definen las variables para la predicción y la matriz de confusión. Para finalmente terminar el conteo del tiempo y mostrar los resultados.

```
# Variable bayes carga el MODELO del algoritmo Naive Bayes
logisticr = LogisticRegression()

# Establece en el modelo los registros para entrenamiento
model= logisticr.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(logisticr.score(X_train, y_train)),"%")

# predicción
y_pred = model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("-------------------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                    {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("-------------------------------------------------------------------------------------------------")
```

Figura 39. Modelo de Regresión Logística.

3.6.8. Random Forest.

Para la implementación del modelo Random Forest se instalan las librerías necesarias para contabilizar el tiempo, visualización de datos, cargue del Dataset y otras generales para todos los modelos, y como se observa en la Figura 40, RandomForestClassifier en específico para este.

```
# Libreria para utilizar bosques aleatorios
from sklearn.ensemble import RandomForestClassifier
```

Figura 40. Importe de la librería RandomForestClassifier.

Se inicia el conteo del tiempo, se cargan el Dataset y se muestran algunas características generales del Dataset, así como también se definen la cantidad de datos tomados para entrenamiento y pruebas. Como se observa en la Figura 41, se define la variable 'bosque' que carga el modelo, con número máximo de estimadores de 10, se establecen los registros para entrenamiento, se definen las variables para la predicción y la construcción de la matriz de confusión, y se finaliza el conteo del tiempo para finiquitar el proceso con la muestra de los resultados.

```
# Variable arbol carga el MODELO del Árbol de Desición
bosque = RandomForestClassifier(n_estimators=10)

# Establece en el modelo los registros para entrenamiento
model= bosque.fit(X_train, y_train)

# Verifica el porcentaje de aprendizaje del algoritmo
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}"
       .format(bosque.score(X_train, y_train)),"%")
print("\n")

# predicción
y_pred = model.predict(X_test)

# Matriz de confusión en pruebas
cnf_matrix = confusion_matrix(y_test, y_pred)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)

# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("-----------------------------------------------------------------------------------------")
porcentaje=(1-(y_test != y_pred).sum()/y_test.shape[0])
print("Precisión:                  {0:.1f}"
      .format(porcentaje * 100),"%")
print("Errores de clasificación:     {1} errores, sobre un total de {0} casos"
      .format(y_test.shape[0],(y_test != y_pred).sum()))
print("Tiempo de ejecución es de:    {0:.2f}"
      .format(tiempo),"seg")
print ("-----------------------------------------------------------------------------------------")
```

Figura 41. Modelo Random Forest.

3.6.9. Aprendizaje profundo con Redes Neuronales.

Para la implementación del modelo de aprendizaje profundo, se utilizó como base la estructura propuesta de Redes Neuronales, por lo cual se hace el importe de las mismas librerías en cuanto al conteo del tiempo, visualización, carga del Dataset, entre otras. En el mismo sentido, para la utilización de las técnicas de Deep Learning, se importan las librerías Tensorflow y Keras, este proceso se puede observar en la Figura 42.

```
# Se importan las librerias Tensorflow y Keras para el deep learning
import tensorflow as tf
from keras.models import Sequential
from keras.layers.core import Dense
```

Figura 42. Importe de librerías para Deep Learning.

El presente algoritmo se estableció como uno de los aportes más importante a esta investigación, tras un proceso de optimización y validación con las pruebas del Dataset de Ciberguerra. Por último, se estableció el algoritmo que se muestra en la Figura 43, el cual contiene elementos como la definición del modelo, definición del número de neuronas de entrada y de salida para la función de activación en las capas ocultas “relu” y para la capa de salida “sigmoid”. En el algoritmo se puede apreciar el modelo de compilación seleccionado, el cual se considera el más propicio para el tipo de datos a tratar. Se define, además, 50 como el número de iteraciones, y luego como en el resto de los algoritmos, se verifica el porcentaje de aprendizaje del algoritmo, se genera la matriz de confusión, se finaliza el conteo del tiempo y se imprimen los resultados.

```
model = Sequential()
model.add(Dense(16, input_dim=cantidad_de_variables_predictoras, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='mean_squared_error', optimizer='adam', metrics=['binary_accuracy'])

#Definición del número de iteraciones
model.fit(predictores_entrenamiento, valores_a_predecir_entrenamiento, epochs=50, verbose=0 )
print("\n")

# Verifica el porcentaje de aprendizaje del algoritmo
scores  = model.evaluate(predictores_pruebas , valores_a_predecir_pruebas )
print("\n\n")
print ("Aprendizaje del algoritmo en el entrenamiento: {0:.2f}" .format (scores[1]*100),"%")

#Matríz de confusión de entrenamiento
valores_predichos = model.predict(predictores_pruebas ).round()
cnf_matrix = confusion_matrix(valores_a_predecir_pruebas , valores_predichos)

#Finaliza el tiempo del proceso
tiempo=(time.time()-inicio_tiempo)
```

```
# Matriz de confusión de pruebas
print("\n")
print ("MATRIZ DE CONFUSIÓN")
sns.heatmap(cnf_matrix.T, square=True, annot=True, fmt='d', cbar=True)
plt.xlabel('Clase verdadera')
plt.ylabel('Clase predecida')
plt.title('Matriz de Confusión')
plt.show()
print ("Descripción")
print ("Verdaderos Negativos: ",cnf_matrix[0,0])
print ("Verdaderos Positivos: ",cnf_matrix[1,1])
print ("Falsos Positivos:     ",cnf_matrix[1,0])
print ("Falsos Negativos:     ",cnf_matrix[0,1])

# Visualizando los resultados
print("\n\n")
print ("RESULTADOS DE CLASIFICACIÓN")
print ("---------------------------------------------------------------------")
valores_predichos =   [x[0] for x in valores_predichos]
errores = (valores_a_predecir_pruebas  != valores_predichos).sum()
total_de_casos = valores_a_predecir_pruebas .shape[0]
porcentaje = (100 / total_de_casos) * (total_de_casos - errores)
print("Porcentaje de precisión:   {0:.1f}".format(porcentaje),"%")
print("Errores de clasificación:  {1} errores, sobre un total de {0} casos".format(total_de_casos,errores))
print("Tiempo de ejecución es de: {0:.2f}".format(tiempo),"seg")
print ("---------------------------------------------------------------------")
```

Figura 43. Modelo Redes Neuronales con Deep Learning

3.6.9.1. Optimización del modelo de Deep Learning.

El modelo de aprendizaje profundo, inicialmente se basó en la misma estructura usada para Redes Neuronales y se tomó como base la propuesta de Bagnato (2018), adaptando el algoritmo propuesto en su ejercicio a la estructura y las necesidades del presente ejercicio. Inicialmente el algoritmo contenía los elementos que se muestran en la Figura 44.

```
model = Sequential()
model.add(Dense(64, input_dim=cantidad_de_variables_predictoras, activation='relu'))
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='mean_squared_error', optimizer='adam', metrics=['binary_accuracy'])

#Definición del número de iteraciones
model.fit(predictores_entrenamiento, valores_a_predecir_entrenamiento, epochs=1000)
print("\n")
```

Figura 44. Parte del algoritmo inicial que sería intervenida para mejoras en el rendimiento.

Implementado el algoritmo con estos valores los resultados de eficiencia y rendimiento son regulares en el caso de la precisión. Sin embargo, en cuanto al tiempo de ejecución que requiere están muy alejados de lo que se pretende obtener. Los valores pueden apreciarse en la Figura 45.

```
RESULTADOS DE CLASIFICACIÓN
----------------------------------------------------------------
Porcentaje de precisión:   98.6 %
Errores de clasificación:  3 errores, sobre un total de 222 casos
Tiempo de ejecución es de: 49.41 seg
----------------------------------------------------------------
```

Figura 45. Resultados del algoritmo de Redes Neuronales con Deep Learning en su primera versión.

Partiendo de estos resultados, se hizo un análisis sobre los valores o comandos que sería necesario cambiar para hacer el algoritmo más eficiente con el tiempo ya que al estar muy elevado en segundos, con una gran cantidad de valores, esto genera sobre costos

importantes. Dicho análisis comenzó por disminuir el número de iteraciones teniendo en cuenta que 1000 iteraciones sería un valor muy exagerado conociendo que desde la iteración 50 se comienzan a obtener precisiones binarias con 1 o muy cercanas a 1 (ver Figura 46), por lo cual se seleccionó el valor 50 para el número de iteraciones.

```
Epoch 46/1000
886/886 [==============================] - 0s 46us/step - loss: 0.0023 - binary_accuracy: 0.9989
Epoch 47/1000
886/886 [==============================] - 0s 46us/step - loss: 0.0018 - binary_accuracy: 1.0000
Epoch 48/1000
886/886 [==============================] - 0s 45us/step - loss: 0.0016 - binary_accuracy: 1.0000
Epoch 49/1000
886/886 [==============================] - 0s 45us/step - loss: 0.0015 - binary_accuracy: 1.0000
Epoch 50/1000
886/886 [==============================] - 0s 45us/step - loss: 0.0014 - binary_accuracy: 1.0000
Epoch 51/1000
886/886 [==============================] - 0s 48us/step - loss: 0.0014 - binary_accuracy: 1.0000
Epoch 52/1000
886/886 [==============================] - 0s 47us/step - loss: 0.0013 - binary_accuracy: 1.0000
Epoch 53/1000
886/886 [==============================] - 0s 46us/step - loss: 0.0012 - binary_accuracy: 1.0000
Epoch 54/1000
886/886 [==============================] - 0s 48us/step - loss: 0.0012 - binary_accuracy: 1.0000
Epoch 55/1000
886/886 [==============================] - 0s 48us/step - loss: 0.0011 - binary_accuracy: 1.0000
```

Figura 46. Evidencia de presiones binarias para iteraciones entre la 46 y la 55 de 1000.

Al realizar el cambio del número de iteraciones mencionado, se disminuyó en gran medida el tiempo, con lo cual se logró el propósito de este cambio. El tiempo de ejecución de disminuyó hasta alrededor de 5 segundos, como se puede observar en la Figura 46. Sin embargo, todavía se considera que es necesario seguir disminuyendo el tiempo, puesto que para el número de datos manejados sigue siendo un tiempo muy extenso. Para esto se propone hacer una disminución del número de neuronas de entrada y pasarle un parámetro (verbose = 0) a la función de entrenamiento, en el cual elimine el proceso de visualización de las iteraciones, ya que estas al contener una barra de progreso en tiempo real afectaban el desempeño del procesador y por lo tanto opacaban el rendimiento del algoritmo, esto teniendo en cuenta que el algoritmo estaba siendo intervenido cada micro segundos para revisar el porcentaje de avance para alimentar a barra de progreso. Una vez realizados estos

cambios, el tiempo de ejecución del algoritmo se reduce a tiempos entre 1 y 2 segundos, considerando este resultado un tiempo aceptable en relación con los demás algoritmos y la precisión del mismo.

3.6.9.2.Validación del algoritmo.

Para la validación del algoritmo, se implementó 32 veces y se obtuvo la media de las precisiones en esa muestra tomada, lo cual según el teorema del límite central permitirá establecer una gran proximidad con la media de la población, como se evidencia en la tabla 3.

Tabla 3. *Evidencia de la toma de muestras para validación del algoritmo de Redes Neuronales con Deep Learning*

Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 2.00 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.87 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 1.83 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.82 seg
Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.92 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.82 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 1.92 seg	Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.94 seg
Porcentaje de precisión: 98.2 % Errores de clasificación: 4 errores Tiempo de ejecución es de: 1.92 seg	Porcentaje de precisión: 98.2 % Errores de clasificación: 4 errores Tiempo de ejecución es de: 1.93 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.74 seg	Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.87 seg
Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.79 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 1.77 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.88 seg	Porcentaje de precisión: 98.2 % Errores de clasificación: 4 errores Tiempo de ejecución es de: 1.69 seg
Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.88 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.80 seg	Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.97 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.79 seg
Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.90 seg	Porcentaje de precisión: 98.2 % Errores de clasificación: 4 errores Tiempo de ejecución es de: 1.80 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 2.07 seg	Porcentaje de precisión: 100.0 % Errores de clasificación: 0 errores Tiempo de ejecución es de: 1.88 seg
Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.79 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 1.81 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.96 seg	Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 1.78 seg
Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.84 seg	Porcentaje de precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.82 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.89 seg	Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.70 seg

Una vez tomada la muestra, se da paso a obtener la media de los valores de la precisión, obteniendo el valor de $\bar{X} = 99.009375 \approx 99.1$. Con este proceso podemos entonces inferir que la media poblacional se acerca a este valor, validando el 99.1% como la precisión del algoritmo propuesto en este trabajo de investigación para detectar vocabulario de Ciberguerra con Redes Neuronales Multicapa con Deep Learning, el cual podrá constatarse en los resultados.

3.7. Consideraciones éticas.

En ocasiones, la información utilizada en las investigaciones de carácter científico, pueden incurrir en errores relacionados con el tratamiento de los datos, teniendo en cuenta que pueden estar basados en datos sensibles o que simplemente de una u otra forma violentan la privacidad de personas o entidades. La presente tesis, recolecta información libre de la red, la cual ha sido publicada con pleno consentimiento del autor y no incurre en recolección de datos personales o sensibles, ya que toma como referencia sólo el vocabulario usado en las páginas Web seleccionadas para dar fundamento a la creación de ontologías semánticas sobre el tema de Ciberguerra. Dentro del mismo marco, el uso que se le da a las mismas, es análisis de conexiones y relaciones entre ellas para generar conocimiento o valor agregado sobre la temática en particular. Por último, es importante mencionar, que los resultados son mostrados en forma general y son públicos.

En el contexto colombiano se reglamenta el marco general de la protección de los datos personales en la nación a través de la Ley 1581 de 2012 y en virtud de lo que la misma reglamenta y lo expuesto en el párrafo precedente, puede afirmarse que no se incurre en tratamiento inadecuado de datos.

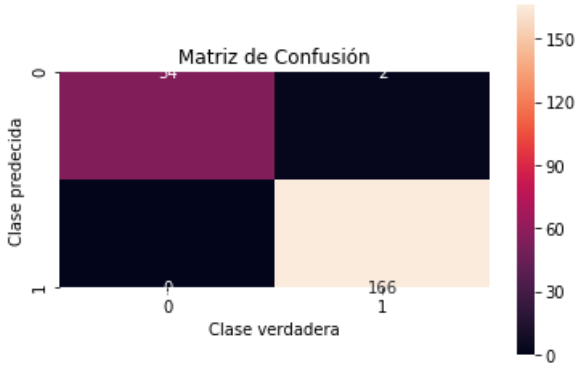
Capítulo IV. Resultados.

4.1. Resultados.

Luego de realizar las pruebas con los algoritmos correspondientes, expuestos en el objetivo anterior, bajo las condiciones establecidas en el mismo. Se da cabida a exponer y explicar las salidas que generan en el desarrollo de cada uno de ellos. Para medir la eficiencia de cada modelo de aprendizaje, se obtienen los resultados de los porcentajes de precisión de cada modelo empleado, los tiempos que tardó en ejecutarse y también la matriz de confusión, de cada uno, la cual permite establecer el número de aciertos y errores para las distintas clases.

4.1.1. Árboles de Decisión.

En los resultados que ofrece este modelo se observa que de los 222 registros que fueron utilizados para las pruebas, correspondiente al 20% de todos los datos del Dataset, el modelo pudo predecir un total de 220 registros, alcanzando un 99.1% de precisión. Realizando este proceso en un tiempo de 0.04 segundos. Puede decirse, además, que todas las páginas que el modelo consideró que no eran de Ciberguerra, efectivamente no lo eran y solo en dos casos de los que consideró que si tenían un vocabulario correspondiente al de Ciberguerra falló. En la Figura 47, pueden observarse la matriz de confusión y los demás resultados arrojados por este modelo.

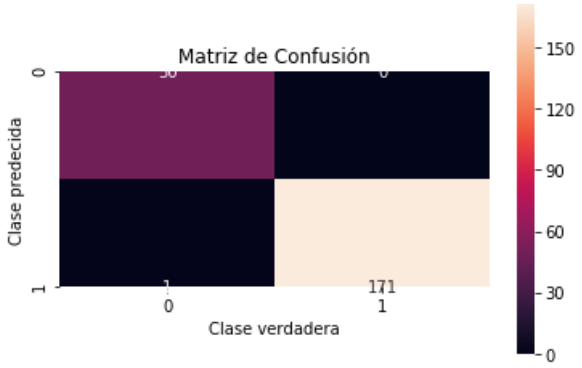



```
Descripción
Verdaderos Negativos:  54
Verdaderos Positivos:  166
Falsos Positivos:      2
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    99.1 %
Errores de clasificación:     2 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.04 seg
--------------------------------------------------------------------------------
```

Figura 47. Resultados arrojados por el modelo Árboles de Decisión.

4.1.2. Naïve Bayes.

Sobre el total de los datos tomados para las pruebas que corresponde a 222, el algoritmo Naïve Bayes presenta una precisión de 99.5% logrando predecir 221 de los casos estudiados. En este caso el modelo pudo predecir todo el conjunto de positivos, sin embargo, falló al considerar negativo un sitio Web que estaban clasificados como positivo. En la Figura 48, se muestra la matriz de confusión y demás resultados arrojados por el modelo.

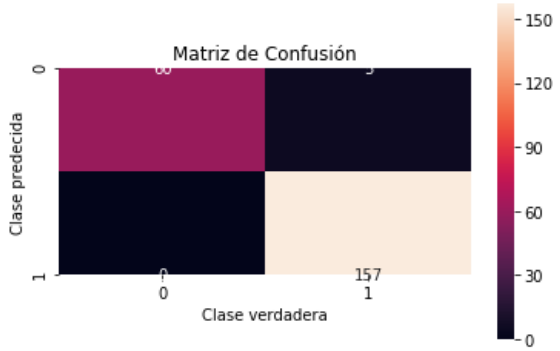



```
Descripción
Verdaderos Negativos:  50
Verdaderos Positivos:  171
Falsos Positivos:      0
Falsos Negativos:      1


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    99.5 %
Errores de clasificación:     1 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.05 seg
--------------------------------------------------------------------------------
```

Figura 48. Resultados Arrojados por el modelo Naïve Bayes.

4.1.3. El K-vecino más cercano – Knn.

La Figura 49, muestra los resultados del modelo Knn – vecino más cercano, en los cuales puede observarse que el modelo tuvo una precisión del 97.7%, resaltando que el error estuvo en que de los 162 positivos que detectó el modelo, 5 de estos estaban clasificadas dentro del grupo de los negativos y su implementación tuvo una duración de 0.26 segundos.

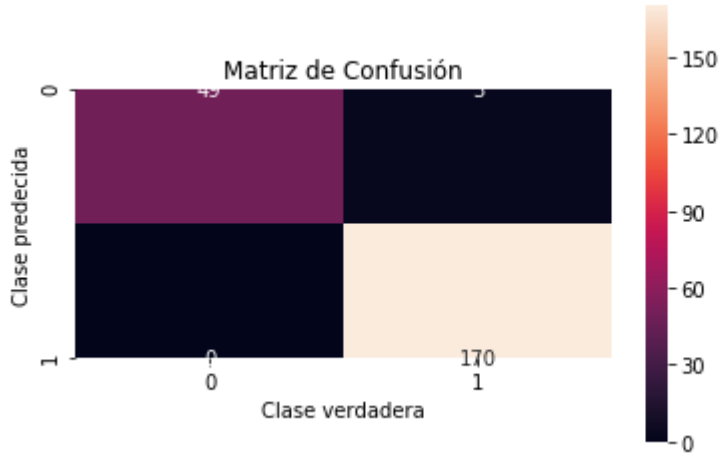



```
Descripción
Verdaderos Negativos:  60
Verdaderos Positivos:  157
Falsos Positivos:      5
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    97.7 %
Errores de clasificación:     5 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.26 seg
--------------------------------------------------------------------------------
```

Figura 49. Resultados arrojados por el modelo del Knn-vecino más cercano.

4.1.4. Redes Neuronales.

La Figura 50, muestra los resultados arrojados por el modelo de Redes Neuronales, con el algoritmo Perceptron Multicapa, en la cual se puede observar que el modelo tuvo un acierto de 219 casos de los 222 tomados para probar la efectividad del modelo, cometiendo el error de etiquetar como positivos 3 casos de páginas Web que no pertenecían al grupo de páginas que contienen el vocabulario de Ciberguerra establecido en el corpus lingüístico aprendido por el modelo. De modo que se evidencia que el modelo tuvo una precisión del 98.6% y se ejecutó en un tiempo de 0.68 segundos.



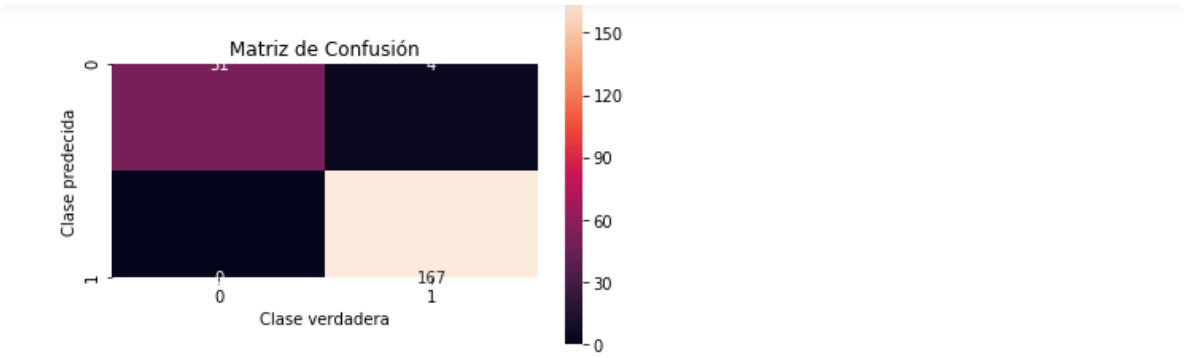

```
Descripción
Verdaderos Negativos:  49
Verdaderos Positivos:  170
Falsos Positivos:      3
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
---------------------------------------------------------------
Porcentaje de precisión:   98.6 %
Errores de clasificación:  3 errores, sobre un total de 222 casos
Tiempo de ejecución es de: 0.68 seg
---------------------------------------------------------------
```

Figura 50. Resultados arrojados por el modelo Redes Neuronales.

4.1.5. Máquinas de Soporte Vectorial.

En la Figura 51, puede observarse que el modelo de Maquinas de Soporte Vectorial tuvo una efectividad del 98.2%, teniendo en cuenta que cometió 4 errores identificando como positivas 4 páginas categorizadas como negativas. Utilizando para esto 0.04 segundos.

En relación con el proceso de optimización con el parametro “C”, no se dio mejora obteniendo también un 98.2% de efectividad.

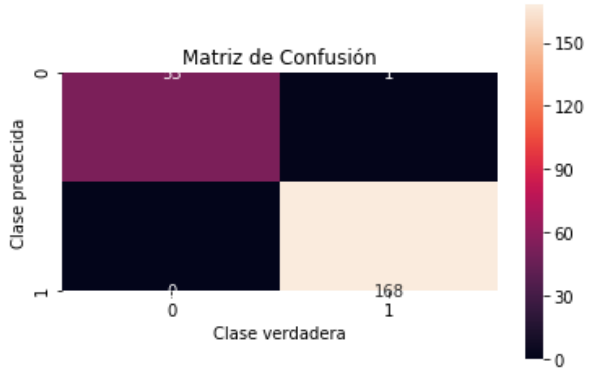



```
Descripción
Verdaderos Negativos:  51
Verdaderos Positivos:  167
Falsos Positivos:      4
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------------
Precisión:                    98.2 %
Precisión optimizada (C=10):  98.2 %   **Optimiza la clasificación generando
                                         un hiperplano de separación de mayor margen
Errores de clasificación:     4 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.04 seg
--------------------------------------------------------------------------------------
```

Figura 51. Resultados arrojados por el modelo Maquinas de Soporte Vectorial.

4.1.6. Ada Boosting.

En la Figura 52, se aprecian los resultados arrojados por el modelo Ada Boosting, el cual obtuvo una precisión del 99.5%, pudiendo predecir 53 negativos y 168 positivos para un total de 221 datos en total tomados para la realización de las pruebas. El tiempo de ejecución del modelo fue de 1.15 segundos.

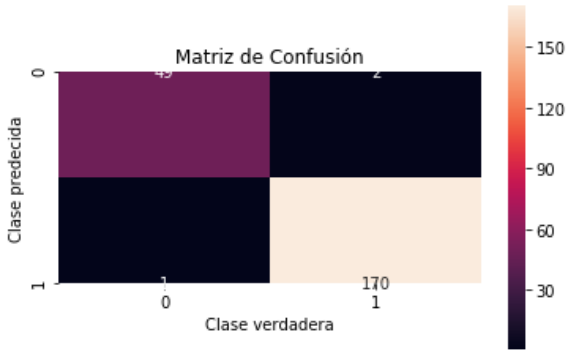



```
Descripción
Verdaderos Negativos:  53
Verdaderos Positivos:  168
Falsos Positivos:      1
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    99.5 %
Errores de clasificación:     1 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    1.15 seg
--------------------------------------------------------------------------------
```

Figura 52. Resultados arrojados por el modelo Ada Boosting.

4.1.7. Regresión Logística.

Los resultados arrojados por el modelo de Regresión Logística se observan en la Figura 53, entre los cuales se destaca que en un tiempo de 0.04 segundos, obtuvo una efectividad del 98.6%, obteniendo errores, clasificando como positivos 2 páginas etiquetadas como negativas y clasificando como negativa una página Web etiquetada como de Ciberguerra en el conjunto de pruebas.

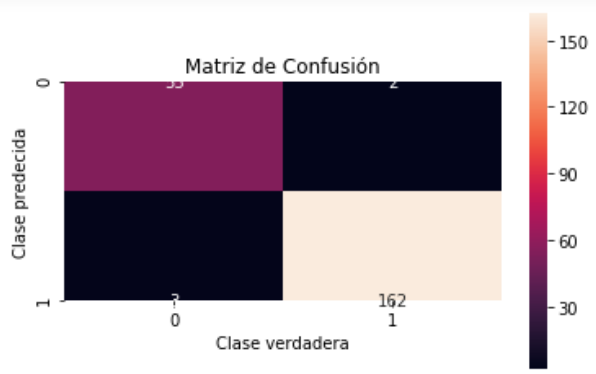



```
Descripción
Verdaderos Negativos:  49
Verdaderos Positivos:  170
Falsos Positivos:      2
Falsos Negativos:      1


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    98.6 %
Errores de clasificación:     3 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.04 seg
--------------------------------------------------------------------------------
```

Figura 53. Modelo de Regresión Logística.

4.1.8. Random Forest.

La Figura 54, muestra los resultados arrojados por el modelo Random Forest, con el cual se obtuvo una precisión del 97.8% en la detección de vocabulario de Ciberguerra en un conjunto de 222 páginas Web de las cuales, el modelo pudo predecir y clasificar 55 negativas y 162 positivas y falló detectando como positivas 2 páginas clasificadas dentro de las negativas y 3 negativas que estaban etiquetadas como positivas, utilizando para esto 0.03 segundos.

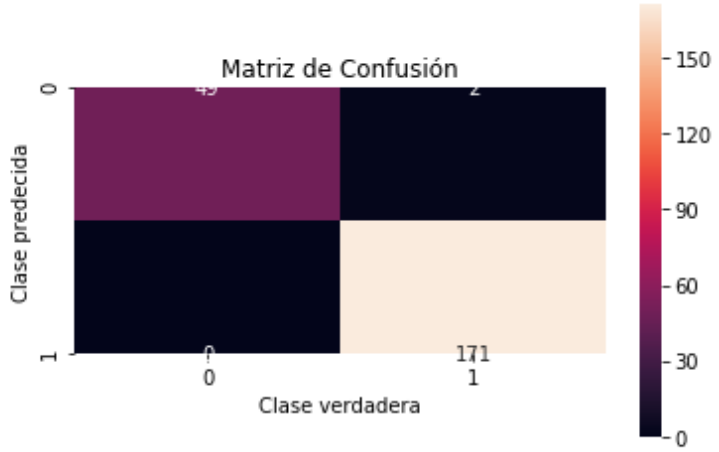



```
Descripción
Verdaderos Negativos:  55
Verdaderos Positivos:  162
Falsos Positivos:      2
Falsos Negativos:      3


RESULTADOS DE CLASIFICACIÓN
--------------------------------------------------------------------------------
Precisión:                    97.7 %
Errores de clasificación:     5 errores, sobre un total de 222 casos
Tiempo de ejecución es de:    0.03 seg
--------------------------------------------------------------------------------
```

Figura 54. Resultados del modelo Random Forest.

4.1.9. Aprendizaje profundo con Redes Neuronales.

En la Figura 55, puede apreciarse que los resultados arrojados por el modelo de Redes Neuronales con Deep Learning, que de un total de 222 páginas Web pudo predecir 220, los dos errores que resultaron se dieron al identificar como positivas dos páginas que estaban etiquetadas como negativas. Manteniendo una efectividad del 99.1% y el tiempo de ejecución de dicha predicción fue de 1.70 segundos.




```
Descripción
Verdaderos Negativos:  49
Verdaderos Positivos:  171
Falsos Positivos:      2
Falsos Negativos:      0


RESULTADOS DE CLASIFICACIÓN
---------------------------------------------------------------
Porcentaje de precisión:   99.1 %
Errores de clasificación:  2 errores, sobre un total de 222 casos
Tiempo de ejecución es de: 1.70 seg
---------------------------------------------------------------
```

Figura 55. Resultados del modelo Aprendizaje profundo con Redes *Neuronales*

4.2. Explicación de resultados.

En esta sección se hará una breve descripción de lo que se presenta en los resultados de los algoritmos que se seleccionaron para la implementación, que permitirán a su vez una mejor comprensión de los mismos. Para cada modelo se presenta:

- Una matriz de confusión que nos aportará información sobre el número de verdaderos negativos, verdaderos positivos, falsos positivos y falsos negativos, que cada modelo arrojó. Tomando en cuenta que los verdaderos negativos representan el número de páginas Web que el modelo identificó como negativos y según las etiquetas previas

que ya contenían los registros del dataset son negativas, los verdaderos positivos son los registros que el modelo identificó como páginas con vocabulario de Ciberguerra y su predicción fue efectiva. Los falsos positivos representan el número de registros que fueron identificados como positivos no siéndolo y los falsos negativos el número de registros identificados como negativos que realmente estaban etiquetados como positivos y el modelo no logró identificarlos como tal.

- Cada modelo, además, muestra el porcentaje de precisión, que representa la proporción de verdaderos positivos y verdaderos negativos, frente al total de registros. Este es el dato que permite hacer una comparación más directa entre los modelos, ya que presenta un porcentaje del acierto que cada uno tuvo al enfrentarse al 20% del dataset que se tomó para evaluarlo luego de ser entrenado con el 80% del mismo.
- Se presenta, además, los errores de clasificación, que representa el desacierto del modelo, siendo esto la suma de los falsos positivos y falsos negativos.
- Por último, se evidencia el tiempo de ejecución, con el cual se muestra en segundos el tiempo que demoró el modelo en hacer el proceso de predicción.

La tabla 4 muestra un consolidado de los resultados arrojados por los nueve modelos motivo de estudio en esta investigación, en la cual se ponen en evidencia los aspectos explicados en el párrafo precedente y se facilita la comparación entre los mismos.

Tabla 4. *Resultados de precisión de los Algoritmos de aprendizaje supervisado realizados en Python.*

Árboles de Decisión	Naïve Bayes	KNN
Descripción Verdaderos Negativos: 54 Verdaderos Positivos: 166 Falsos Positivos: 2 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 0.04 seg	Descripción Verdaderos Negativos: 50 Verdaderos Positivos: 171 Falsos Positivos: 0 Falsos Negativos: 1 RESULTADOS DE CLASIFICACIÓN Precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 0.05 seg	Descripción Verdaderos Negativos: 60 Verdaderos Positivos: 157 Falsos Positivos: 5 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Precisión: 97.7 % Errores de clasificación: 5 errores Tiempo de ejecución es de: 0.26 seg
Redes Neuronales	**SVM**	**Ada Boosting**
Descripción Verdaderos Negativos: 49 Verdaderos Positivos: 170 Falsos Positivos: 3 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Porcentaje de precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 0.68 seg	Descripción Verdaderos Negativos: 51 Verdaderos Positivos: 167 Falsos Positivos: 4 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Precisión: 98.2 % Precisión optimizada (C=10): 98.2 % Errores de clasificación: 4 errores Tiempo de ejecución es de: 0.04 seg	Descripción Verdaderos Negativos: 53 Verdaderos Positivos: 168 Falsos Positivos: 1 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Precisión: 99.5 % Errores de clasificación: 1 errores Tiempo de ejecución es de: 1.15 seg
Regresión Logística	**Random Forest**	**Aprendizaje profundo**
Descripción Verdaderos Negativos: 49 Verdaderos Positivos: 170 Falsos Positivos: 2 Falsos Negativos: 1 RESULTADOS DE CLASIFICACIÓN Precisión: 98.6 % Errores de clasificación: 3 errores Tiempo de ejecución es de: 0.04 seg	Descripción Verdaderos Negativos: 55 Verdaderos Positivos: 162 Falsos Positivos: 2 Falsos Negativos: 3 RESULTADOS DE CLASIFICACIÓN Precisión: 97.7 % Errores de clasificación: 5 errores Tiempo de ejecución es de: 0.03 seg	Descripción Verdaderos Negativos: 49 Verdaderos Positivos: 171 Falsos Positivos: 2 Falsos Negativos: 0 RESULTADOS DE CLASIFICACIÓN Porcentaje de precisión: 99.1 % Errores de clasificación: 2 errores Tiempo de ejecución es de: 1.70 seg

La Figura 56, muestra un resumen de los resultados en relación con las precisiones de los nueve algoritmos trabajados en la presente investigación, para una mejor visualización de los mismos.

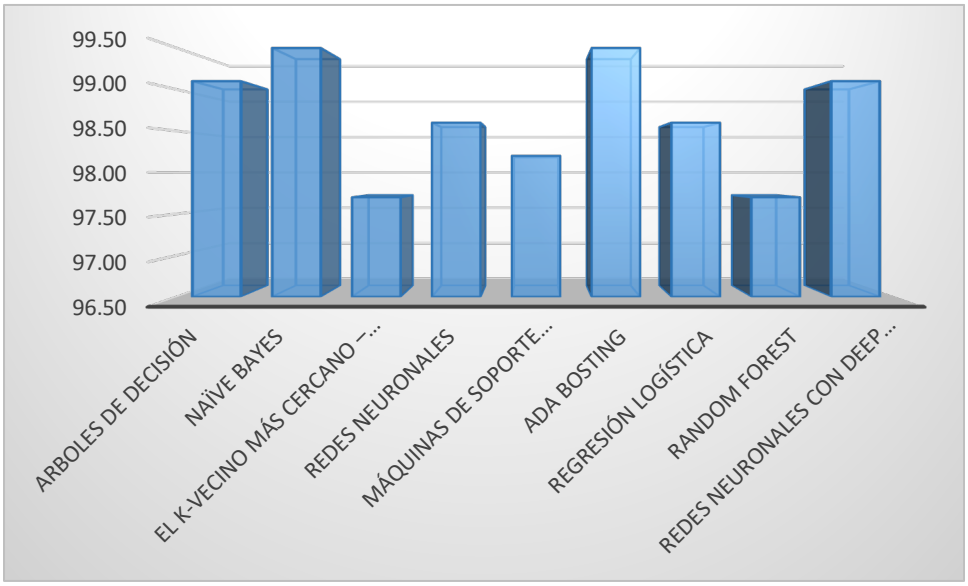


Figura 56. Gráfica de precisión en la detección de vocablos de Ciberguerra en sitios Web.

Capítulo V. Discusión.

5.1. Análisis y comparación entre los algoritmos utilizados en la investigación.

La Tabla 5, presenta un resumen de los resultados obtenidos con los algoritmos de aprendizaje supervisado utilizados para las pruebas en esta tesis, resaltando los parámetros: modelo de datos, números correspondientes a la matriz de confusión, porcentaje de precisión y el tiempo que tardó en realizar el proceso. Lo que permite identificar cual o cuales técnicas ofrecen mayores ventajas en torno a la detección de vocabulario de Ciberguerra. En donde el porcentaje mayor, obtuvo un 99.5% de acierto con los algoritmos Ada Boosting y Naïve Bayes, con tiempos de ejecución de 1.15 y 0.05 segundos respectivamente; y el porcentaje menor, se dio con los algoritmos de Knn – vecino más cercano y Random Forest con un 97.7% de precisión, con tiempos de ejecución 0.26 y 0.03 segundos respectivamente.

El algoritmo con el mejor tiempo fue Random Forest con 0.03 segundos. Sin embargo, como se menciona en el párrafo anterior, su precisión fue la más baja.

Tabla 5. *Resumen de resultados de los modelos implementados.*

Modelo	Verdaderos negativos	Verdaderos positivos	Falsos positivos	Falsos negativos	Precisión (%)	Tiempo (Seg.)
Arboles de decisión	54	166	2	0	99.1	0.04
Naïve Bayes	50	171	0	1	**99.5**	0.05
El K-vecino más cercano – KNN	60	157	5	0	97.7	0.26
Redes neuronales	49	170	3	0	98.6	0.68
Máquinas de soporte vectorial	51	167	4	0	98.2	0.04
Ada Boosting	53	168	1	0	**99.5**	1.15
Regresión Logística	49	170	2	1	98.6	0.04
Random Forest	55	162	2	3	97.7	**0.03**

Redes neuronales con Deep Learning	49	171	2	0	99.1	1.70

Con base en el análisis anterior, cabe destacar que el algoritmo que obtiene mejores resultados, es Naïve Bayes con un 95.5 de precisión y 0.05 segundos de ejecución en el proceso. En segundo lugar, el algoritmo de Arboles de Decisión con una precisión de 99.1%, muy cercana al anterior, pero con un mejor tiempo de ejecución de 0.04 segundos. Los algoritmos Ada Boosting y Redes Neuronales con Deep Learning, se destacan al igual que Naïve Bayes y Arboles de Decisión, en cuanto a la precisión, sin embargo, quedan un poco atrás en el tiempo de ejecución, aunque sus resultados son muy competitivos.

Una comparativa interesante se da entre los modelos Redes Neuronales y Redes Neuronales con Deep Learning, en donde en un inicio fue superior el primero que el segundo. *Es importante resaltar el proceso de optimización que se realizó al algoritmo de Redes Neuronales con Deep Learning, cuando en un inicio de su programación el tiempo de ejecución obtenido fue mayor a 150 segundos. Proceso, que tuvo que ser mejorado, realizando distintas combinaciones en el número de iteraciones y número de capas de entrada, para finalmente lograr reducirlo hasta 1.60 segundos, incluso mejorando el promedio de precisión obtenida con el mismo.*

Cabe resaltar, de manera general que, todos los algoritmos implementados, tuvieron porcentajes altos de precisión y se ejecutaron en tiempos razonables en relación con la calidad y cantidad de datos.

5.2. Discusión con trabajos relacionados.

En este espacio se dará revisión a las conexiones entre los resultados de la presente investigación y los trabajos relacionados que se encuentran referenciados en el estado del arte.

En relación con el trabajo de Rodríguez (2018), ambas investigaciones se relacionan con la seguridad informática, con la diferencia que, este último se enfoca en las conexiones maliciosas. La naturaleza de los Dataset es similar basada en aprendizaje supervisado y ontologías semánticas, donde cada conexión o página Web está etiquetada. Así también, usaron el lenguaje Python en la programación de sus algoritmos, librerías similares para el análisis y visualización de los datos, y para el aprendizaje profundo sólo fueron utilizadas en la presente investigación, Keras y TensoFlow.

La Tabla 6, muestra un comparativo entre los resultados arrojados por el trabajo de Rodríguez (2018) y la presente investigación, destacando en negrita los mejores resultados en tiempo y precisión. Cabe señalar, que el volumen y estructura de los datos es distinta, por lo cual la comparación se centrará en identificar las potencialidades de los algoritmos que en común tuvieron excelentes resultados.

Se puede apreciar en esta comparativa, que en cuanto a precisión se evidencia total oposición, teniendo en cuenta que en la presente investigación entre sus mejores resultados presenta a Naïve Bayes con un 99.5% de precisión, mientras que para la investigación de Rodríguez (2018), se encuentra como el resultado más bajo con un 93.01% de precisión. En ese mismo sentido, el mejor resultado para la investigación de Rodríguez (2018), es Random Forest con un 99.96% de precisión y en la presente investigación resultó uno de los más bajos

con un del 97.7% de precisión. En cuanto al tiempo de ejecución del proceso, teniendo en cuenta las proporciones, se puede decir que en la presente investigación pudo mejorar el tiempo de los algoritmos llevándolos a tiempos más homogéneos en relación con la cantidad y estructura de sus datos. Haciendo notar, que los algoritmos con resultados de tiempo muy elevados en la investigación de Rodríguez (2018), no son homogéneos en comparación con sus propios resultados; contrario a la presente investigación, donde se muestran homogenizados con los demás algoritmos. Por último, los algoritmos que se muestran con resultados más elevados en relación con el tiempo en la presente investigación, no fueron objeto de estudio de la investigación de Rodríguez (2018).

Tabla 6. Comparativa de resultados de la investigación de Rodríguez (2018) y la presente investigación.

Modelo	Rodríguez (2018)		Presente investigación	
	Precisión (%)	Tiempo (Seg.)	Precisión (%)	Tiempo (Seg.)
Arboles de decisión	99.92	**7.71**	99.1	**0.04**
Naïve Bayes	93.01	**5.1**	**99.5**	0.05
El K-vecino más cercano – KNN	-----	------	97.7	0.26
Redes neuronales	-----	------	98.6	0.68
Máquinas de soporte vectorial	99.90	486.68	98.2	**0.04**
Ada Boosting	------	------	**99.5**	1.15
Regresión Logística	99.89	6016.51	98.6	**0.04**
Random Forest	**99.96**	387.39	97.7	**0.03**
J48	**99.95**	64.68	-----	-----
Redes Neuronales con Deep Learning	-----	-----	99.1	1.70

Otro algoritmo a destacar en ambas investigaciones, es “Arboles de Decisión”, el cual presento buen rendimiento y alto porcentaje de precisión, lo que genera alta confiabilidad en este modelo, apoyado con información de trabajos relacionados según Pedregosa et al., (2011), en donde destaca como una excelente técnica para clasificación de datos con las ventajas de: facilidad de entender e interpretar la poca preparación de datos que requiere, la capacidad de manejar datos numéricos y categóricos y la capacidad de manejar problemas de salida múltiple; razones, por las cuales fue seleccionado para realizar un proceso de optimización, que finalmente tuvo gran relevancia en su investigación.

El trabajo realizado por Cuenca (2019), tiene como objetivo primordial el uso de diversas técnicas de Machine Learning para detectar una posible amenaza en redes sociales o acciones realizadas por Bots, a través de un modelo matemático, utilizando Python como lenguaje de programación y Anaconda como entorno de trabajo, compartiendo además la mayor parte de las librerías para tratamiento de datos y visualización de resultados.

Para las pruebas utiliza dos Dataset, PHEME (PH) y Detecting Twitter Bot (DTB), con los algoritmos, Arboles de Decisión, Bosque Aleatorios, Máquinas de Soporte Vectorial y Knn – Vecino más Cercano, seleccionando a Random Forest como la mejor opción con una efectividad del 81% y 96% para la identificación de noticias falsas o la presencia de Bots respectivamente. La Tabla 7, muestra un comparativo con la presente investigación, resaltando en negrita los resultados más altos.

Tabla 7. Comparativa de resultados de la investigación de Cuenca (2019) con la presente investigación.

Modelo	Cuenca (2019)		Presente investigación
	PH	DTB	

Arboles de Decisión	91	**98**	99.1
Naïve Bayes	**---**	**---**	**99.5**
El K-vecino más cercano – Knn	49	48	97.7
Redes Neuronales	**---**	**---**	98.6
Máquinas de Soporte Vectorial	**100**	91	98.2
Ada Boosting	**---**	**---**	**99.5**
Regresión Logística	**---**	**---**	98.6
Random Forest	**96**	**99**	97.7
Redes Neuronales con Deep Learning	**---**	**---**	99.1

Se puede observar que el modelo Knn - vecino más cercano, presenta poca efectividad en ambas investigaciones, teniendo en cuenta que en el trabajo de Cuenca (2019), obtuvo una efectividad por debajo del 50% tanto en la detección de noticias falsas como en la de presencia de Bots, y en la investigación actual, es uno de los algoritmos que presenta menor efectividad al momento de identificar páginas Web relacionadas con el vocabulario de Ciberguerra.

Por otra parte, cabe mencionar que el algoritmo seleccionado como el más apropiado para la detección de los dos tipos de datos manejados en el trabajo de Cuenca (2019), no resalta como una de las mejores opciones de efectividad en la presente investigación.

Otro aspecto importante, con relación al trabajo mencionado, es que dio cabida a la selección del modelo de Redes Neuronales y Redes Neuronales con Deep Learning, de acuerdo a sus recomendaciones de investigaciones futuras, en las cuales plantea la importancia de emplear otro tipo de algoritmos o la combinación de algunos de ellos para hacer más rica la investigación.

Por último, Hidalgo (2019), presenta algunas similitudes en los objetivos con la presente investigación, y aporto ideas para el tratamiento de los datos con Redes Neuronales Artificiales, aunque no se trabajó con los mismos lenguajes de programación, la técnica utilizada permitió elevar el rendimiento general y la precisión del detector de intrusos desde un 95% hasta un 99.4%, hecho que propicio establecer en la presente tesis el filtro Relu que fue fundamental en la programación y optimización del algoritmo de Redes Neuronales con Deep Learning.

5.3. Discusión de los objetivos.

1. *Construir un modelo predictivo que sea capaz de identificar si el vocabulario dentro de una página Web, está relacionado con la temática de Ciberguerra.*

En la presente investigación se creó un modelo predictivo capaz de categorizar si una entrada es tocante a la temática de Ciberguerra dentro del contenido de páginas Web. El procedimiento está formado por varios algoritmos con altos porcentajes de precisión. Así también, se consideró la terminología asociada a esta problemática, sustentado con trabajos relacionados tanto para la identificación del vocabulario como para las diferentes técnicas a utilizar, como: la limpieza, procesamiento de la información e implementación de los algoritmos de Machine Learning.

2. *Obtener una base de datos de páginas Web relacionadas con la información de Ciberguerra que circula en la Web.*

En este trabajo fue fundamental la búsqueda de información en Internet relacionada con la problemática de “Ciberguerra”. Lo que permitió, identificar el contexto, las palabras relacionadas, el propósito, los medios a través de los cuales se genera y las formas en que es

posible realizar ataques. Para formar la muestra de datos para las pruebas, fueron seleccionadas 1108 páginas Web, de las cuales 842 están relacionadas con el tema de Ciberguerra y 266 no lo están.

3. *Establecer cuales elementos de la base de datos será utilizado para las pruebas de detección de Ciberguerra, mediante la transformación de datos semiestructurados a datos estructurados en formato .csv.*

Basado en el pre-procesamiento de datos sobre las páginas Web localizadas y descargadas por el Crawler, se estableció el corpus lingüístico que permitió a los distintos modelos identificar si el contenido de la página Web estaba o no relacionado con la temática de Ciberguerra, posteriormente se generó el Dataset en formato estándar con extensión .csv.

4. *Seleccionar distintos algoritmos de aprendizaje supervisado que van a utilizarse para detección de vocabulario de Ciberguerra.*

Con base en los trabajos relacionados se seleccionaron los algoritmos: Arboles de Decisión, Naïve Bayes, El Knn - vecino más cercano, Redes Neuronales, Redes Neuronales con Aprendizaje Profundo, Máquinas de Soporte Vectorial, Ada Boosting, Regresión Logística y Random Forest, los cuales presentan similitud con la presente investigación en el tratamiento de datos y su porcentaje de predicción es competitivo.

5. *Implementar los algoritmos seleccionados para detectar vocabulario de Ciberguerra.*

En las pruebas para detectar vocabulario de Ciberguerra en sitios Web (con los algoritmos mencionados en el punto anterior), se obtuvieron resultados de precisión entre el 97.7 y 99.5%, y tiempos de ejecución entre 0.03 y 1.70, resultados que se consideran

excelentes y permiten hacer un aporte significativo a los objetivos propuestos en esta investigación.

6. *Comparar los resultados arrojados por los distintos algoritmos, para establecer cuál es el más adecuado para la detección de Ciberguerra.*

Para establecer cual algoritmo es el más adecuado, se tomaron como base los parametros: efectividad y rendimiento. Destacando que todos los algoritmos implementados tuvieron porcentajes altos de precisión y se ejecutaron en tiempo razonables en relación con la calidad y cantidad de datos al detectar vocabulario de Ciberguerra en páginas Web. El algoritmo Naïve Bayes, obtuvo el porcentaje de precisión más alto (99.5%), sumado a un excelente rendimiento (0.04 segundos).

5.4. Respuesta a la hipótesis.

Precisando la hipótesis que a la letra dice "*El desarrollo de un sistema capaz de detectar vocabulario de Ciberguerra dentro del contenido de páginas Web, usando técnicas de Procesamiento de Lenguaje Natural, Web Semántica, Aprendizaje Supervisado, Machine Learning y Deep Learning, permitirá identificar cuál o cuáles son los algoritmos más efectivos para generar valor agregado sobre las páginas que circulan en Internet en relación con dicha problemática*". La presente investigación encuentra que los algoritmos: Naïve Bayes y Ada Boosting, son los algoritmos con mayor precisión para tal efecto, con 99.5% de efectividad, apoyados en el proceso con técnicas de Procesamiento de Lenguaje Natural, Web Semántica, Aprendizaje Supervisado, Machine Learning y Deep Learning. Con lo cual puede afirmarse que la hipótesis planteada inicialmente corresponde con los resultados obtenidos.

Capítulo VI. Conclusiones.

6.1. Conclusiones y trabajos futuros.

6.1.1. Conclusiones.

De acuerdo a los resultados obtenidos, se puede concluir que el sistema experto para la detección de vocabulario de Ciberguerra en Sitios Web, basado en Técnicas de Machine Learning y Deep Learning con Python, es competitivo, debido a que sus resultados oscilan entre el 97.7% y el 99.5%, de precisión. El estudio se encuentra fundamentado con una metodología organizada en etapas, que permitió la construcción del Dataset combinando distintas técnicas, como: Big Data Analytics, Procesamiento de Lenguaje Natural, Web Semántica, Minería de Texto, así como, la implementación de la herramienta ADVI para la creación de un corpus lingüístico, permitiendo un proceso exitoso en la clasificación de los sitios Web de Ciberguerra.

6.1.2. Fortalezas y debilidades.

Una de las principales ventajas que tiene esta propuesta, radica en la escogencia del tema "Ciberguerra" de importancia para la ciencia, las naciones, gobiernos, fuerzas armadas, y demás entes relacionados con la seguridad a nivel internacional, tal como está sustentado en los aportes científicos que han sido citados a través de la misma.

Otra de las fortalezas que se tienen, son los altos porcentajes de precisión en la detección de vocablos de Ciberguerra en los sitios Web, a través de todos los algoritmos implementados por arriba del 97.7%, por lo que se tiene gran confiabilidad en el modelo de datos generado.

Otra de las fortalezas, es la profundidad de las técnicas aplicadas en la investigación, destacando el uso de algoritmos de Redes Neuronales con Deep Learning, una de las técnicas

más avanzadas en la analítica de datos en comparación con los algoritmos de Machine Learning, en donde se obtuvo un algoritmo inicial que fue perfeccionado hasta obtener una versión tan competitiva como la mejor de Machine Learning.

Por otro lado, uno de los inconvenientes que se presentaron, fue la falta de experticia y práctica en relación con las herramientas de programación utilizadas, ya que causó retrasos importantes en muchos momentos cruciales de la investigación.

6.1.3. Trabajos futuros (Generación de nuevas investigaciones).

Existen posibles líneas de trabajo futuro, identificadas en el desarrollo de la presente investigación. Para lo cual se propone:

- Establecer mecanismos para desarrollar herramientas y procedimientos que permitan la construcción de conjuntos de datos (dataset) basados en la Web profunda y puedan ser analizados por el modelo de analítica de datos generado en esta investigación.
- Construir nuevos modelos de análisis con aprendizaje profundo (Deep Learning), combinando con tecnologías como Spark mediante el uso de GPUs y procesamiento distribuido, buscando mejorar el comportamiento en tiempo y predicción.

Referencias.

Adserá, A. (21 de Abril de 2010). Cerebro y Sistema Nervioso. Obtenido de https://www.enciclopediasalud.com/categorias/cerebro-y-sistema-nervioso/articulos/la-neurona

Alarcón, J. (2017). Modelos de minería de datos: random forest y adaboost, para identificar los factores asociados al uso de las TIC (internet, telefonía Fija y televisión de paga) en los hogares del Perú. 2014. Obtenido de https://cybertesis.unmsm.edu.pe/bitstream/handle/cybertesis/7404/Alarcon_fj.pdf?sequence=3&isAllowed=y

Arista, A., Calderón, G., Fierro, A., & Nakano, M. (6 de Febrero de 2017). Clasificación de Imágenes Urbanas Aéreas: Comparación entre Descriptores de Bajo Nivel y Aprendizaje Profundo. Obtenido de https://scielo.conicyt.cl/pdf/infotec/v28n3/art21.pdf

Bagnato, J. (29 de Mayo de 2018). Aprende Machine Learing. Obtenido de Crear una sencilla red nueronal: https://www.aprendemachinelearning.com/una-sencilla-red-neuronal-en-python-con-keras-y-tensorflow/

Barrón, I. (2019). Detección y Análisis de Vocabulario de Ciberterrorismo en la Detección y Análisis de Vocabulario de Ciberterrorismo en la Web, a través del uso de Modelos Predictivos de Machine Learning. México.

Bejarano, M. C. (Septiembre de 2011). Instituto Español de Estudios Estrategicos. Obtenido de http://www.ieee.es/Galerias/fichero/docs_informativos/2011/DIEEEI09-2011ConceptoCiberdefensaOTAN.pdf

Benítez, J. (Julio de 2014). Big Data: Algoritmos, tecnología y aplicaciones. Obtenido de Depto. Ciencias de la Computación e I.A. Universidad de Granada: http://madm.uib.es/wp-content/uploads/2016/06/Jose-Manuel-Benitez-Sanchez-Big-Data-Algoritmos_tecnologia_y_aplicaciones.pdf

Bouza, C., & Santiago, A. (2012). La Minería De Datos: Arboles De Decisión y su Aplicación en Estudios Médicos. Obtenido de Universidad de la Habana: https://s3.amazonaws.com/academia.edu.documents/43713947/MINERIA_DE_DATOS_MEDICOS.pdf?response-content-disposition=inline%3B%20filename%3DLA_MINERIA_DE_DATOS_ARBOLES_DE_DECISION.pdf&X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAIWOWYYGZ2Y53UL3A%2F

Cabezón, M. (1 de Septiembre de 2018). Predicción demanda eléctrica española. Implementación de redes neuronales recurrentes en Python. Obtenido de Trabajo de fin de máster en Ingeniería Matemática [Universidad Complutense de Madrid]: https://eprints.ucm.es/49444/1/2018-MIGUEL%20CABEZON%20Memoria.pdf

Camargo, J., Camargo, J., & Joyanes, L. (1 de Diciembre de 2014). Conociendo Big Data. Obtenido de Revista Scielo: http://www.scielo.org.co/pdf/rfing/v24n38/v24n38a06.pdf

Castañeda, J. (2019). Análisis, Clasificación y Predicción del Vocabulario de Cibercrimen en Internet Usando Modelos Predictivos de Machine Learning. CDMX, México.

Castillo, I. (s.f.). Minería de datos con R. Aguascalientes: Instituto Tecnologico Aguascalientes.

Castillo, I., Luna, F., Muñoz, J., & López, J. (2 de Septiembre de 2016). Arquitectura (advi) para la Detección de Vocabulario de Ciberbullying en Internet Combinando Técnicas de Big Data Analytics y Web Semántica. Obtenido de https://www.dyna-newtech.com/search-content-NT/architecture-advi-for-the-detection-of-cyberbullying-vocabulary-in-internet-combining-techniques-of

Castrillo, O. (10 de Diciembre de 2015). Web Scraping: Applications and Tools. Obtenido de EPSi Platform: https://www.europeandataportal.eu/sites/default/files/2015_web_scraping_applications_and_tools.pdf

Castro, A., González, J., & Ballesteros, J. (26 de Junio de 2015). Technologies for metadata management in scientific articles. Obtenido de http://web.a.ebscohost.com.dibpxy.uaa.mx/ehost/pdfviewer/pdfviewer?vid=6&sid=ea667199-77e8-4dfb-b5ee-bfbf20f537a3%40sdc-v-sessmgr01

Chatterjee, S., & Nath, A. (4 de Abril de 2017). Crawler, Auto-Explore the Web – Web . Obtenido de https://www.researchgate.net/publication/316601171_Auto-Explore_the_Web_-_Web_Crawler

Conpes. (14 de julio de 2011). MINTIC. Obtenido de Lineamientos de Política para Ciberseguridad y Ciberdefensa: https://www.mintic.gov.co/portal/inicio/3510:Conpes-3701-de-2011

Corredera, J. C. (2012). Introducción. En Ceseden, El ciberespacio. Nuevo esenario de confrontación. (págs. 9-34). Madrid.

Cortés, R. (2015). Estado actual de la política pública de ciberseguridad y ciberdefensa en Colombia. Revista de Derecho Comunicaciones y Nuevas tecnologías.

Cortez, A., Huerta, H. V., & Pariona, J. (2013). Procesamiento de lenguaje natural. Revista de Ingeniería de Sistemas e Informática vol. 6, N.º 2, 45-54. Obtenido de Sistema de Bibliotecas y Biblioteca Central.

Cuenca, W. (Junio de 2019). Machine Learning aplicado a la seguridad. Obtenido de Universidad oberta de Cataluña: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/97009/6/wcuencaTFM0619memoria.pdf

Cuji, B. (11 de 03 de 2018). Ciberguerra y su Efecto en la Seguridad Nacional. Obtenido de Revista de Ciencias de Seguridad y Defensa (Vol. III, No. 2, 2018): http://geo1.espe.edu.ec/wp-content/uploads//2018/04/109-117.pdf

Domínguez, J. (23 de Mayo de 2016). La ciberguerra como realidad posible contemplada desde la prospectiva. Obtenido de Revista de Pensamiento Estratégico y Seguridad CISDE: http://uajournals.com/ojs/index.php/cisdejournal/article/view/146

Emer, E. (s.f.). BOOSTING (ADABOOST ALGORITHM) . Obtenido de http://math.mit.edu/~rothvoss/18.304.3PM/Presentations/1-Eric-Boosting304FinalRpdf.pdf

Evans, D. (2011). Internet de las cosas Cómo la próxima evolución. Cisco Internet Business Solutions Group (IBSG).

Ferreira, A. (2016). Computacion paralela aplicada a la detección de anormalidades en la fermentacion vínica mediante Maquinas de Soporte Vectorial. Obtenido de http://noesis.uis.edu.co/bitstream/123456789/20097/1/165286.pdf

Ferrero, A. (2013). La Ciberguerra. Génesis y evolución,. Revista General de Marina, 81-97.

Fuentes, S., & Ruíz, M. (Octubre de 2007). Minería Web: un recurso insoslayable para el profesional de la información. Obtenido de scielo: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352007001000011

Gaitán, A. (2013). La ciberguerra y sus generaciones: un enfoque para comprender la incidencia de las tic en la guerra regular. Revista Estudios en Seguridad y Defensa, 5-18. Obtenido de La ciberguerra y sus generaciones: un enfoque para comprender la incidencia de las tic en la guerra regular: https://esdeguerevistacientifica.edu.co/index.php/estudios/article/view/194/279

Gala, Y. (25 de Septiembre de 2013). Algoritmos SVM para problemas sobre big data. Obtenido de https://repositorio.uam.es/bitstream/handle/10486/14108/66152_Yvonne_Gala_Garcia.pdf?sequence=1

García, Á. (Junio de 2016). Machine Learning en Bases de Datos de Lenguaje Natural. Obtenido de UNIVERSIDAD AUTÓNOMA DE MADRID: https://repositorio.uam.es/bitstream/handle/10486/676778/Garcia_Gutierrez_Alvaro_tfg.pdf?sequence=1&isAllowed=y

García, O. (s.f.). Big Data. Obtenido de Universidad Nacional de Lujan: http://www.basicas.unlu.edu.ar/sites/www.basicas.unlu.edu.ar/files/site/Big-Data-Machine_Learning.pdf

Godoy, Á. (5 de Noviembre de 2015). Técnicas de aprendizaje de maquina utilizadas para la minería de texto. Obtenido de Scielo: http://www.scielo.org.mx/pdf/ib/v31n71/2448-8321-ib-31-71-00103.pdf

Gómez, H. (2017). Ciberguerra...¿Dudáis? Revista de Marina Nº 959, 34-39.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. Cambridge Massachusets: Massachusets Institute of tecnology.

Hernandez, M., & Gomez, J. (2013). Aplicaciones de Procesamiento de Lenguaje Natural. Revista Politécnica Vol. 32, No. 1, 87-96.

Hidalgo, Y., & Rodríguez, R. (2013). La web semántica: una breve revisión. Revista Cubana de Ciencias Informáticas , 76-85.

Hssina, B., Merbouha, A., Ezzikouri, H., & Erritali, M. (s.f.). A comparative study of decision tree ID3 and C4.5. Obtenido de https://saiconference.com/Downloads/SpecialIssueNo10/Paper_3-A_comparative_study_of_decision_tree_ID3_and_C4.5.pdf

IBM. (s.f.). Acerca de la minería de textos. Obtenido de https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/ta_guide_ddita/textmining/shared_entities/tm_intro_tm_defined.html

Jímenez, C. (Diciembre de 2014). Big data. Un nuevo paradigma de análisis de datos. Obtenido de https://www.iit.comillas.edu/docs/IIT-14-153A.pdf

Joyanes, L. (2013). Big Data Analisisde grandes volumenes de datos en organizaciones. Mexico D F: Alfaomega Grupo editor.

Krotov, V., & Silva, L. (2018 de Septiembre de 2018). Legality and Ethics of Web Scraping. Obtenido de Researchgate: https://www.researchgate.net/publication/324907302_Legality_and_Ethics_of_Web_Scraping

Larrañaga, P., Inza, I., & Moujahid, A. (s.f.). Clasificadores Bayesianos. Obtenido de http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t6bayesianos.pdf

Lejarza, E. (21 de Febrero de 2014). Instituto Español de Estudios Estratégicos. Obtenido de http://www.ieee.es/Galerias/fichero/docs_opinion/2014/DIEEEO18-2014_Ciberguerra_EscenariosConfrontacion_EguskineLejarza.pdf

López, Y. (2017). Análisis y Diseño de un Sistema para la Extracción, Análisis y Comparación de Precios de Tiendas en la Ciudad de Bogotá D.C. Obtenido de Universidad Industrial: http://repository.udistrital.edu.co/bitstream/11349/6853/2/LopezPovedaYair2017.pdf

Lozano, A. (s.f.). Ontologías en la Web Semántica. Obtenido de Universidad de Zaragoza: http://eolo.cps.unizar.es/docencia/MasterUPV/Articulos/Ontologias%20en%20la%20Web%20Semantica.pdf

Martinez, J. (s.f). Máquinas de Vectores de Soporte (SVM). Obtenido de https://iartificial.net/maquinas-de-vectores-de-soporte-svm/

Maya, E. (2018). Los Árboles de Decisión como Herramienta para el Analisis de Riesgos de los Proyectos. Obtenido de https://repository.eafit.edu.co/: https://repository.eafit.edu.co/bitstream/handle/10784/12980/Elena_MayaLopera_2018.pdf?sequence=2&isAllowed=y

Mendoza, M. (1 de Octubre de 2011). Minería de datos en la Web. Obtenido de researchgate: https://www.researchgate.net/publication/229068670_Mineria_de_datos_en_la_Web

Menes, I., Arcos, G., Moreno, P., & Gallegos, K. (14 de Septiembre de 2015). Desempeño de algoritmos de minería en indicadores académicos: Árbol de Decisión y Regresión Logística . Obtenido de http://scielo.sld.cu/pdf/rcci/v9n4/rcci08415.pdf

Montañés, R., Aznar, R., & Hoyo, R. d. (Septiembre de 2018). Aplicaci´on de un modelo h´ıbrido de aprendizaje profundo para el An´alisis de Sentimiento en Twitter. Obtenido de http://ceur-ws.org/Vol-2172/p4_ittainnova_tass2018.pdf

Monteiro, P. (Septiembre de 2017). Aplicação de técnicas de Data Mining na Cibersegurança e Ciberdefesa – uma breve revisão. Obtenido de Pós-Graduação em Cibersegurança e Ciberdefesa: https://www.academia.edu/37596707/Aplica%C3%A7%C3%A3o_de_t%C3%A9cnicas_de_Data_Mining_na_Ciberseguran%C3%A7a_e_Ciberdefesa_uma_breve_revis%C3%A3o

Montillo, A., & Ling, H. (s.f.). Age Regression From Faces Using Random Forests . Obtenido de http://www.dabi.temple.edu/~hbling/publication/MontilloL09icip.pdf

Ortega, M., & Cayuela, A. (s.f.). Regresión logística no condicionada y tamaño de muestra: una revisión bibliográfica. Obtenido de https://www.scielosp.org/article/resp/2002.v76n2/85-93/

Ortiz, A., Martín, M., Ureña, A., & García, M. (1 de Junio de 2005). Detección automática de Spam utilizando Regresión Logística Bayesiana. Obtenido de https://www.redalyc.org/pdf/5157/515751735016.pdf

Pastor, J. (Noviembre de 2011). Tecnologías de la web semántica. Obtenido de https://ebookcentral.proquest.com/lib/univeraguascalientessp/reader.action?docID=3201739

Pedraza, R., Codina, L., & Rovira, C. (4 de Noviembre de 2007). Web semántica y ontologías en el procesamiento de la información documental. Obtenido de http://eprints.rclis.org/14298/1/webSemanticaOntologias2007.pdf

Pedregosa. (2011). scikit-learn. Obtenido de https://scikit-learn.org/stable/modules/tree.html#tree-algorithms-id3-c4-5-c5-0-and-cart

Pelaez, N. (2012). Aprendizaje no Supervisado y el Algoritmo Wake-Sleep en Redes Neuronales. obtenido de universidad tecnologica de la mixteca: http://jupiter.utm.mx/~tesis_dig/11612.pdf

Presidencia. (23 de Octubre de 2018). Presidencia de la República . Obtenido de https://id.presidencia.gov.co/Paginas/prensa/2018/181023-Colombia-y-OTAN-emprenderan-estrategia-conjunta-para-prevenir-ciberataques-y-fortalecer-la-ciberseguridad-en-el-pais.aspx

Quintana, Y. (2016). Ciberguerra. Los Libros de las Cataratas.

Rochina, P. (25 de Abril de 2017). ¿Qué es y cuáles son las aplicaciones del Text Mining? Obtenido de https://revistadigital.inesem.es/informatica-y-tics/text-mining/

Rodriguez, J. (4 de Junio de 2018). Aplicación de tencnicas de Machine Learning a la detección de ataques. Obtenido de openaccess: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/81126/11/jmrodriguez85TFM0618memoria.pdf

Roig, C. (Junio de 2011). TFC – XML y Web Semántica . Obtenido de openaccess: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/8086/1/croiggTFC0611memoria.pdf

Rosas, M. C. (1 de 10 de 2014). BiblioMedia. Obtenido de https://www-bibliomedia-com-mx.dibpxy.uaa.mx/

Russom, P. (2011). BIG DATA ANALYTICS. Obtenido de https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf

Sammut, C., & Webb, G. (2017). Enciclopedy of Machine Learning and Data Mining. Obtenido de https://books.google.com.co/books?id=fjdhDwAAQBAJ&pg=PA894&lpg=PA894&dq=sammut+y+webb+naive+bayes&source=bl&ots=5U8piWhb0i&sig=ACfU3U0SBUdhn79xZkUcrr5g6GldKKxoPg&hl=es&sa=X&ved=2ahUKEwi4yrvXqdTmAhXFjVkKHcp6BoIQ6AEwAnoECAkQAQ#v=onepage&q&f=false

Sánchez, G. (13 de Febrero de 2018). Internet: Una Herramienta para las Guerras en el Siglo XXI. Obtenido de Revista Política y Estrategia N˚ 114, 224-242.: https://www.researchgate.net/publication/323169876_INTERNET_UNA_HERRAMIENTA_PARA_LAS_GUERRAS_EN_EL_SIGLO_XXI/citation/download

Sancho, F. (26 de Diciembre de 2018). Clasificación Supervisada y No Supervisada. Obtenido de Dpto. Ciencias de la computación e inteligencia artificial, Universidad de Sevilla: http://www.cs.us.es/~fsancho/?e=77

Sandoval., L. (2018). Algoritmos de Aprendizaje Automático para Análisis y Predicción de Datos. Revista tecnológica, 36-40.

Santiago, E., & Sánchez, J. (10 de Abril de 2016). Diseño de un Sistema Multiagentes Híbrido Basado en Aprendizaje Profundo para la Detección y Contención de Ciberataques. Obtenido de http://revistas.unipamplona.edu.co/ojs_viceinves/index.php/RCTA/article/view/2495

Silva, L. (1994). Excursión a la regresión logística en ciencias de la salud. Ediciones Díaz de SantosEdiciones Díaz de Santos.

Šlibar, B. (Noviembre de 2019). Predicting the Number of Downloads of Open Datasets by Naïve Bayes Classifier. Obtenido de University of Zagreb, Faculty of Organization and Informatics Varaždin, Pavlinska 2, 42 000 Varaždin, Croatia: http://web.a.ebscohost.com.dibpxy.uaa.mx/ehost/pdfviewer/pdfviewer?vid=3&sid=431913db-12f8-45af-ba6f-2bc47ce01988%40sdc-v-sessmgr01

Suárez, A. (2015). El quinto elemento : espionaje, ciberguerra y terrorismo : una amenaza real e inminente. Deusto.

Torres, M. (2011). ASPFOR XXVI, Un paso adelante. Ejercito, 14-19.

Vidueira, J., Souza, C., Miranda, R., & Figueiredo, A. (10 de Junio de 2015). The use of the k nearest neighbor method to classify the representative elements. Obtenido de http://www.scielo.org.mx/pdf/eq/v26n3/0187-893X-eq-26-03-00195.pdf

Villada, D., & Jiménez, A. (12 de Mayo de 2017). La Web Semántica y la Web Profunda como Sistemas de Información: Análisis a una realidad. Obtenido de Revista Antioqueña de las Ciencias Computacionales y la Ingeniería de Sofware: http://web.b.ebscohost.com.dibpxy.uaa.mx/ehost/pdfviewer/pdfviewer?vid=3&sid=a07c24f7-60b6-4831-b681-743c35429348%40pdc-v-sessmgr01

Zhao, B. (Mayo de 2017). Researchgate. Obtenido de Web Scraping: https://www.researchgate.net/publication/317177787_Web_Scraping

Anexos.

Anexo A.

No.	Páginas Web utilizadas en el análisis de Ciberguerra
1	https://www.marxists.org/espanol/marigh/obras/mini.htm
2	http://chiapas.laneta.org/desmilitarizacion/encuentro/ponencias/sandoval.htm
3	https://www.monografias.com/trabajos14/histcomput/histcomput2.shtml
4	https://es.gizmodo.com/el-ejercito-de-estados-unidos-se-esta-preparando-para-d-1795927413
5	http://www.belt.es/noticias/especial/Ciberguerra/index.asp
6	https://es.gizmodo.com/el-ataque-ddos-que-tumbo-medio-internet-es-un-adelanto-1788144314
7	http://www.rs-labs.com/papers/tacticas/
8	https://es.gizmodo.com/descubren-un-nuevo-vehiculo-oculto-en-el-codigo-de-gta-1828162566
9	http://www.razonypalabra.org.mx/anteriores/n36/jesteinou.html
10	http://clio.rediris.es/n40/articulos/alumnos_soldados.html
11	https://www.um.es/docencia/barzana/II/Ii04.html
12	https://journals.openedition.org/polis/6306
13	https://sistemas.uniandes.edu.co/es/foros-isis/temas-foros-isis/bpm/foro-2/90-foros-isis/revista
14	https://es.wikipedia.org/wiki/Guerra
15	https://www.brookings.edu/es/articles/la-revolucion-de-la-robotica-y-el-conflicto-en-el-siglo-xxi/
16	https://www.muyinteresante.es/tecnologia/articulo/iesto-es-la-ciberguerra
17	https://www.gbm.net/es/de-cero-a-uno-para-convertirse-en-el-proximo-sun-tzu-digital
18	https://es.slideshare.net/joelgomezmx/hacktivismo-ciberterrorismo-y-ciberguerra
19	https://observatorio.cisde.es/archivo/la-amenaza-cibernetica-ciberguerra-y-ciberdefensa/
20	http://www.revistadon.com/16670/cinco-claves-sobre-la-ciberguerra-yolanda-quintana
21	https://www.redalyc.org/html/815/81501508/
22	http://www.capesic.cat/es/2017/05/31/como-italia-se-prepara-para-afrontar-el-terrorismo-cibernetico/

23	http://www.iniseg.es/ciberseguridad/masteres-oficiales/master-ciberterrorismo.html
24	https://www.nytimes.com/es/2017/01/12/ciberguerra-a-la-venta/
25	https://laklave.wordpress.com/tag/guerra-cibernetica/
26	http://www.fao.org/docrep/003/x7352s/X7352s02.htm
27	http://www.academia.edu/Documents/in/Ciberguerra
28	https://moscovita.org/mosconews/venezuela-paisaje-politico-para-entender-el-23-de-febrero/
29	https://www.eldiario.es/turing/escenarios-ciberguerra-nuevo-orden-mundial_0_129837338.html
30	https://elpais.com/elpais/2018/01/22/eps/1516637253_754345.html
31	https://es.slideshare.net/luisfernandolandaeta/nuevoenfoqueguerrasmarykaldor
32	http://www.revistareflexiones.mx/index.php/guerra-cibernetica-situacion-actual-y-perspectiva-futura
33	https://journals.openedition.org/alhim/533
34	http://www.tiempodehoy.com/mundo/la-primera-ciberguerra-mundial
35	http://www.cubadebate.cu/opinion/2017/12/22/guerras-ciberneticas-nuevas-formas-de-guerra/
36	https://www.almendron.com/tribuna/etiqueta/ciberguerra/
37	https://www.dw.com/es/ciberguerra-peligro-de-muerte/a-38850521
38	http://www.planv.com.ec/investigacion/investigacion/la-ciberguerra-correa
39	http://www.thiber.org/1197-2/
40	https://www.elespanol.com/reportajes/grandes-historias/20161216/178733050_0.html
41	https://www.researchgate.net/publication/281714598_Conflictos_contemporaneos_y_ciberespacio
42	https://moscovita.org/mosconews/mi-homenaje-a-antonio-machado/
43	https://civismundirevista.wordpress.com/2015/09/26/conflictos-contemporaneos-y-ciberespacio/
44	http://agendaglobal.redtercermundo.org.uy/2014/05/28/la-ciberguerra-se-agudiza/
45	https://www.nato.int/docu/review/2011/11-september/cyber-threads/es/index.htm
46	https://www.elmundo.es/papel/historias/2017/03/12/58c151f522601dab398b45dc.html
47	http://katehon.com/es/article/eeuu-y-rusia-de-la-guerra-fria-la-guerra-cibernetica

48	https://www.redpapaz.org/el-tiempo-los-papas-declaran-la-guerra-a-la-comida-chatarra/
49	http://www.fundaceic.org/2014/08/11/hostilidades-en-el-ciberespacio-entre-estados-naciones/
50	https://es.wikipedia.org/wiki/Internet_profunda
51	https://www.muycomputerpro.com/2018/07/03/gran-guerra-ciberespacio
52	https://www.eldiario.es/turing/Iran-Washington-ciberguerra-Tel-Aviv_0_396861028.html
53	https://www.universidadviu.com/ciberespacio-inseguridad-carrera-armamentos-disuasion-siglo-xxi/
54	https://www.technologyreview.es/s/10883/china-contra-el-mundo-asi-funciona-su-ciberguerra
55	https://www.proceso.com.mx/92367/guerra-en-el-ciberespacio
56	http://www.iicybersecurity.com/empresa-de-seguridad-informatica.html
57	https://marketing4ecommerce.net/historia-de-internet/
58	https://elusuariofinal.wordpress.com/category/ciberguerra/
59	https://elpais.com/elpais/2012/05/28/opinion/1338198139_520582.html
60	http://www.rtve.es/noticias/20110923/ciberguerra-asi-se-libraran-nuevas-batallas/463576.shtml
61	http://www.juventudrebelde.cu/cuba/2011-03-21/ciberguerra-mercenarismo-en-la-red
62	http://www.ejercito.mil.pe/
63	https://elpais.com/diario/2001/09/18/opinion/1000764008_850215.html
64	https://www.esglobal.org/juegos-de-ciberguerra/
65	https://tdhpe.techdata.eu/es/Tech-Data-Blog/Conoce-el-arte-de-la-guerra-cibernetica/
66	https://lasverdadesdemiguel.tv/la-ciberguerra-tu-hogar-sera-el-campo-de-batalla/
67	http://blogs.lavanguardia.com/washington/apuntes-sobre-la-ciberguerra-39643
68	https://xombit.com/2013/02/analisis-2012-seguridad-informatica
69	https://eugene.kaspersky.com.mx/2016/12/06/una-breve-historia-de-los-ataques-ddos/
70	https://www.eldiario.es/cultura/tecnologia/Primera-Ciberguerra-Mundial_0_598790464.html
71	http://www.accionecologica.org/petroleo/politicas-petroleras/191-la-guerra-petrolera-en-irak

72	http://www.lostiempos.com/oh/actualidad/20180226/guerra-digital-poder-traves-redes-sociales
73	https://www.bbc.com/mundo/noticias-38926665
74	http://www.bancomundial.org/es/results/2013/04/13/ict-results-profile
75	http://www.ahorasemanal.es/la-guerra-tecnologica-de-daesh
76	https://www.monografias.com/trabajos109/guerra-secreta-red/guerra-secreta-red.shtml
77	https://www.infolibre.es/noticias/politica/2016/06/21/ciberguerra_51556_1012.html
78	https://dragonball.fandom.com/es/wiki/Guerra_del_Red_Ribbon
79	https://www.alground.com/site/cose-guerra-cibernetica/45247/
80	https://www.bbc.com/mundo/noticias/2011/07/110722_eeuu_pentagono_ciberespacio_estrategia_wbm
81	https://www.elconfidencial.com/mundo/2016-11-30/alemania-se-prepara-para-la-ciberguerra_1296637/
82	https://searchdatacenter.techtarget.com/es/reporte/Por-que-necesitamos-jugar-a-la-guerra-cibernetica
83	https://www.abc.es/tecnologia/redes/20150615/abci-ciberguerra-ciberataque-china-201506131801.html
84	https://elpais.com/tecnologia/2012/06/13/actualidad/1339595642_413641.html
85	http://www.rtve.es/alacarta/videos/en-portada/portada-64/3892038/
86	https://www.curiosfera.com/historia-de-la-computadora-y-su-inventor/
87	http://www.iniseg.es/blog/ciberseguridad/en-plena-ciberguerra-mundial/
88	http://www.expansion.com/actualidadeconomica/analisis/2017/09/07/59b0245646163fdf1b8b457a.html
89	https://elpais.com/tecnologia/2017/05/25/actualidad/1495714364_501019.html
90	http://computerworldmexico.com.mx/notpeyta-acto-ciberguerra/
91	https://www.lanacion.com.ar/1953073-la-primera-ciberguerra-mundial
92	https://mundo.sputniknews.com/blogs/201704121068314836-ffaa-defensa-rusia-seguridad/
93	https://www.swissinfo.ch/spa/mediciones_-bien-preparada-suiza-en-caso-de-desastre-nuclear--/43688990
94	https://www.genbeta.com/genbeta/el-legado-de-la-guerra-fria-en-la-informatica
95	http://razonesdecuba.cubadebate.cu/articulos/la-primera-ciberguerra-mundial/
96	https://www.20minutos.es/noticia/897438/0/guerra/informacion/hackers/

97	https://elpais.com/tag/internet_profunda/a
98	https://elcomercio.pe/tecnologia/actualidad/trae-guerra-cibernetica-anonymous-islamico-389141
99	https://www.cic.es/resumen-del-systems-security-day-ssday16/
100	http://www.iniseg.es/blog/ciberseguridad/ciberterrorismo-y-ciberguerra/
101	https://omicrono.elespanol.com/2017/03/ciberguerra-ejercito-hackers/
102	https://www2.latercera.com/noticia/laboratorio-pruebas-la-ciberguerra/
103	http://razonesdecuba.cubadebate.cu/articulos/escarbando-en-la-genisis-de-la-ciberguerra-parte-i/
104	https://www.dw.com/es/rusia-hackea-y-hostiga-a-soldados-de-la-otan-seg%25C3%25BAn-informe/a-40852655
105	http://disparamag.com/politica/extramuros/ciberespacio-el-nuevo-tablero-la-guerra/index.html
106	http://blogs.eltiempo.com/wikimundo/2018/08/23/3-razones-ocultas-la-guerra-espacial-trump/
107	https://www.linguee.es/espanol-ingles/traduccion/ciberguerra.html
108	https://www.agendadigitale.eu/sicurezza/cyber-guerra-alla-ricerca-governance-globale-sta-succedendo/
109	https://actualidad.rt.com/actualidad/view/92257-guerras-futuro-contacto-informacion
110	https://dialogo-americas.com/es/articles/brasil-se-prepara-para-la-guerra-cibernetica
111	https://www.eldiario.es/turing/Nuevas-tecnologias-tiempos-guerra_0_172183356.html
112	https://llamaloy.wordpress.com/2016/06/12/dia-o-comienza-la-ciberguerra/
113	https://www.eldiariomontanes.es/cantabria/ciberespacio-20170926151902-nt.html
114	http://andinalink.com/existe-el-riesgo-de-una-ciberguerra-que-acabe-con-internet/
115	https://www.daypo.com/introduccion-ciberseguridad.html
116	https://www.latercera.com/noticia/el-ciberespacio-es-el-nuevo-ambito-de-la-guerra-para-el-pentagono/
117	https://www.vice.com/es_mx/article/qv5xwq/activismo-ciberguerra-ucrania
118	https://elpais.com/tag/guerra_digital/a
119	https://www.contralinea.com.mx/archivo-revista/2011/08/07/guerra-cibernetica-la-nueva-amenaza/

120	https://jitorreblanca.wordpress.com/tag/guerra-digital/
121	http://www.derechos.org/nizkor/espana/doc/usa10.html
122	https://www.eldiario.es/tecnologia/informatica-trabajo-mujeres-dominada-hombres_0_679282351.html
123	https://www.genbeta.com/seguridad/stuxnet-historia-del-primer-arma-de-la-ciberguerra
124	https://www.elespanol.com/ciencia/tecnologia/20160629/136236887_0.html
125	https://www.uv.es/~pla/solidaritat/marset.html
126	https://www.bbc.com/mundo/noticias-internacional-43792225
127	https://www.bbc.com/mundo/noticias/2012/04/120430_tecnologica_ciber_armas_aa
128	https://www.nextu.com/blog/generaciones-de-las-computadoras/
129	https://mundo.sputniknews.com/seguridad/201801051075219335-ciberseguridad-washington-ciberguerra/
130	http://www.gees.org/articulos/contrarrestando-el-arte-de-la-guerra-informatica
131	https://www.diariosur.es/tecnologia/201612/18/guerra-fria-libra-20161216165201-rc.html
132	https://rpp.pe/tecnologia/mas-tecnologia/la-ciberguerra-ya-es-una-realidad-noticia-224697
133	https://cso.computerworld.es/cibercrimen/la-ciberguerra-a-examen
134	https://actualidad.rt.com/themes/view/45214-ciberguerra
135	https://www.bbc.com/mundo/noticias/2011/11/111101_mexico_anonymous_zetas_opcartel_irm
136	https://www.larazon.es/historico/3986-la-primera-ciberguerra-mundial-NLLA_RAZON_347513
137	http://www.tiempodehoy.com/espana/la-ciberguerra-aterriza-en-cataluna
138	https://www.larazon.es/economia/asi-sera-la-i-guerra-cibernetica-GB15140116
139	http://www.juventudrebelde.cu/cuba/2011-03-24/nuevo-ejemplo-de-guerra-cibernetica
140	http://www.egov.ufsc.br/portal/conteudo/ciberguerra
141	https://elminuto.cl/ciberguerra-el-campo-de-batalla-digital-nuevo-desafio-para-la-region/
142	https://www.silicon.es/2268086-2268086
143	https://www.eldiario.es/temas/ciberguerra/
144	https://elpais.com/internacional/2016/04/28/actualidad/1461860956_005713.html

145	https://www.diariosur.es/sociedad/201501/05/ciberguerra-20150105173302.html
146	http://www.cubadebate.cu/etiqueta/ciberguerra/
147	https://www.cronista.com/3dias/La-amenaza-global-las-nuevas-guerras-del-siglo-XXI-20150220-0011.html
148	http://www.cioal.com/2015/03/24/china-posee-una-unidad-de-guerra-cibernetica/
149	https://www.elconfidencial.com/mundo/2017-10-02/batalla-estatua-estonia-ciberguerra-rusia_1451408/
150	http://www.razonypalabra.org.mx/comunicarte/2010/carte10_agosto_ciberguerra.html
151	https://www.bbc.com/mundo/noticias/2011/10/111020_tecnologia_duqu_sotware_malicioso_mr
152	https://www.eldiario.es/zonacritica/Neutralidad-Red-Internet-Libertad-ITU-WCIT-12_6_86851317.html
153	https://www.advantio.com/es/blog/securing-your-payments-in-the-age-of-cyber-warfare
154	https://elpais.com/tecnologia/2015/11/19/actualidad/1447949979_838506.html
155	http://paginaspersonales.deusto.es/airibar/Ed_digital/INF/Intro/Historia.html
156	http://blogs-lectores.lavanguardia.com/colaboraciones/la-guerra-de-redes
157	http://blog.trendmicro.es/%3Ftag%3Dciberguerra
158	http://www.ult.edu.cu/ciberguerra-de-las-novelas-de-c-f-a-la-realidad/
159	https://www.urgente24.com/257826-usa-vs-rusia-como-en-la-guerra-fria-pero-cibernetico
160	https://www.lanacion.com.ar/1626859-inglaterra-crea-una-division-militar-para-la-guerra-cibernetica
161	https://elpais.com/tecnologia/2003/02/07/actualidad/1044610080_850215.html
162	http://esmateria.com/2013/06/04/la-ciberguerra-es-inevitable/
163	https://www.dw.com/es/guerras-cibern%25C3%25A9ticas-nerds-al-ataque/a-17005878
164	http://www.emprendeseguridad.com/magazine/cuidado-con-la-guerra-cibernetica-2018/
165	https://mundo.sputniknews.com/europa/201512231055161917-anonumous-ciberguerra-turquia-daesh/
166	https://www.ccn-cert.cni.es/seguridad-al-dia/noticias-seguridad/1179-ciberguerra-en-espana.html
167	https://revistasic.es/revista-sic/sic-128/bibliografia/
168	http://www.cubadebate.cu/

169	https://www.insumisos.com/diplo/NODE/1989.HTM
170	https://www.nytimes.com/es/2018/05/28/bancos-delitos-ciberneticos/
171	https://www.dw.com/es/europa-y-la-guerra-por-el-ciberespacio/a-42621169
172	http://www.cubadebate.cu/etiqueta/ciberguerra/page/11/
173	https://ropa-militar.com/es/blog/tacticas-militares-que-te-ayudaran-en-combate-n216
174	https://www.silicon.es/anonymous-ciberguerra-isis-2297112
175	https://diarioti.com/notpeyta-un-acto-de-ciberguerra/104869
176	https://www.merca2.es/nueva-guerra-cibernetica-redes-sociales/
177	https://www.telesurtv.net/pages/Especiales/GuerrasDeAfrica/index.jsp
178	https://www.eoi.es/blogs/ciberseguridad/2016/04/17/turquia-una-ciberguerra-abierta/
179	https://mundo.sputniknews.com/firmas/201803151077030086-politica-washington-asia-pekin/
180	https://noticias.canalrcn.com/tags/guerra-cibernetica
181	http://espacioestrategico.blogspot.com/2013/01/brasil-y-su-politica-cibernetica-de.html
182	https://www.nytimes.com/es/2017/08/04/solucion-contra-los-hackers-hackeo-defcon/
183	https://www.vice.com/es_mx/article/9kje4e/ciberterrorismo-propaganda-de-estado-islamico
184	http://cuatrof.net/2016/12/14/guerra-en-la-web-bombardeo-cibernetico-contra-venezuela/
185	https://4grandesverdades.wordpress.com/2012/04/13/la-guerra-y-la-paz-en-el-ciberespacio/
186	http://katehon.com/es/article/informe-de-la-guerra-de-yemen
187	https://www.contralinea.com.mx/archivo-revista/2011/02/20/la-nueva-guerra-digital/
188	https://insurgente.org/eeuu-china-la-guerra-comercial-entre-ambos-paises-no-es-lo-que-te-dijeron/
189	https://www.definicionabc.com/historia/ciberguerra.php
190	https://es.weforum.org/agenda/2017/03/un-enfoque-normativo-para-la-prevencion-de-la-ciberguerra/
191	http://www.aucal.edu/grados/master-ciberterrorismo.html
192	https://actualidad.rt.com/actualidad/view/121546-ciberguerra-conflicto-ucrania-hackers

193	http://www.cubadebate.cu/noticias/2014/08/05/anonymous-le-declara-la-ciberguerra-a-israel/
194	https://psicologiaymente.com/miscelanea/consecuencias-de-segunda-guerra-mundial
195	http://www.infosecurityvip.com/newsletter/efraude_jul17.html
196	http://toyoutome.es/blog/como-ser-testigo-de-la-ciberguerra-a-traves-de-9-mapas-interactivos/34451
197	https://as.com/meristation/2018/10/26/noticias/1540549600_668706.html
198	https://marketing4ecommerce.net/facebook-y-adblock-plus-siguen-en-guerra/
199	https://www.lanacion.com.ar/1201067-el-mundo-a-las-puertas-de-una-guerra-cibernetica
200	https://blogs.20minutos.es/enguerra/2009/03/27/guerra-drones-entre-eeuu-iraan-y-pakistaan/
201	https://www.ivoox.com/ciberataques-malwares-ransomwares-mente-digital-audios-mp3_rf_18762847_1.html
202	https://es.rbth.com/articles/2012/07/03/los_peligros_de_la_ciberguerra_17711
203	https://www.estrategiaynegocios.net/opinion/477033-345/ciberguerra-riesgo-para-la-economia-mundial
204	https://www.securityartwork.es/2018/06/11/cyber-guerra-fria-iv-iran/
205	https://www.technologyreview.es/s/10871/como-las-armas-cuanticas-cambiaran-las-guerras-del-futuro
206	http://www.techweek.es/seguridad/informes/1006261004801/mcafee-advierte-ciberguerra.1.html
207	https://www.univision.com/noticias/noticias-de-latinoamerica/ciberguerra-de-blogueros-en-cuba
208	https://carlosguerraterol.com/como-hacer-un-plan-de-marketing-digital/
209	http://nuso.org/articulo/ciberguerra-mas-y-menos-de-lo-que-parece/
210	https://elpais.com/elpais/2017/08/14/opinion/1502725754_912210.html
211	https://www.lavoz.com.ar/mundo/asi-se-prepara-rusia-para-combatir-en-ciberguerra
212	https://securelist.lat/one-night-to-hack-in-paris/76589/
213	https://www.bbc.com/mundo/noticias/2010/10/101018_1451_guerra_cibernetica_stuxnet_virus_dc
214	http://www.cioal.com/2012/09/25/el-mundo-esta-en-ciberguerra/
215	https://es.weforum.org/agenda/2015/06/la-banca-esta-en-riesgo-en-la-era-de-la-ciberguerra/

216	https://mundo.sputniknews.com/politica/201810141082708828-europa-amsterdam-moscu-acusaciones/
217	https://www.elmundo.es/cataluna/2014/09/08/540cebeb268e3eaa5f8b457d.html
218	http://institutointerglobal.org/2010/05/25/la-guerra-informatica-cyber-warfare/
219	https://actualidad.rt.com/tag/Ciberguerra
220	https://elpais.com/diario/2010/01/30/opinion/1264806004_850215.html
221	https://www.proceso.com.mx/215917/africa-la-guerra-por-el-coltan
222	http://www.rcci.net/globalizacion/2003/fg336.htm
223	https://www.genwords.com/blog/marketing-guerrilla
224	http://www.datagora.es/ataques-ocultos-y-sabotajes-la-guerra-cibernetica-que-no-cesa/
225	http://diariouno.pe/columna/la-guerra-cibernetica-la-nueva-blitzkrieg-del-siglo-xxi/
226	http://www.ordenadores-y-portatiles.com/ciberguerra.html
227	http://www.elojodigital.com/categoria/tags/ciberguerra
228	https://smart-lighting.es/telefonica-ciberataque-guerra-informatica/
229	http://www.ipsnoticias.net/1997/03/comunicacion-el-general-sun-tzu-y-la-guerra-del-ciberespacio/
230	https://elpais.com/diario/2000/10/30/internacional/972860407_850215.html
231	https://www.bbc.com/mundo/noticias/2012/01/120120_tecnologia_guerra_por_internet_aa
232	https://markaabyayala.wordpress.com/2018/10/18/tecnicas-de-la-guerra-digital-en-brasil/
233	http://www.iniseg.es/blog/seguridad/ciberguerra-las-armas-del-futuro/
234	https://www.xataka.com/tag/ciberguerra
235	https://actualidad.rt.com/actualidad/view/131931-mapa-ciberguerra-mundial
236	https://www.monografias.com/trabajos20/guerra-irak/guerra-irak.shtml
237	https://www.bez.es/678570999/Guerra-digital-contra-Daesh-necesitamos-creatividad.html
238	http://news.bbc.co.uk/hi/spanish/science/newsid_7560000/7560462.stm
239	https://www.bbc.com/mundo/noticias-47199738
240	https://mundo.sputniknews.com/politica/201803111076918956-eeuu-rusia-obama-ciberataques-injerencia/
241	https://www.vernegroup.com/actualidad/red-team-vs-blue-team-simulacros-de-ciberguerra-por-kinomakino

242	https://www.elespanol.com/mundo/europa/20170109/184732377_0.html
243	https://mundo.sputniknews.com/politica/201508071040099682/
244	http://www.redseguridad.com/actualidad/editoriales/tierra-mar-aire-espacio-y-ciberespacio
245	https://www.forte.jor.br/
246	https://www.incibe-cert.es/blog/ctf-entrenamiento-seguridad-informatica
247	http://fernandosantamaria.com/blog/ideas-de-la-guerra-red-tras-la-muerte-de-bin-laden/
248	https://actualidad.rt.com/actualidad/163692-snowden-nsa-armas-ciberneticas-guerra
249	https://www.elcomercio.com/guaifai/kaspersky-seguridad-internet-guerra-hackers.html
250	https://www.technologyreview.es/s/2431/deberiamos-ser-los-primeros-en-disparar-en-una-ciberguerra
251	https://lapupilainsomne.wordpress.com/2019/02/23/abismos-entre-el-si-y-el-no-por-luis-toledo-sande/
252	http://www.redcross.int/ES/mag/magazine2003_3/4-9.html
253	http://atalayar.com/content/la-guerra-contra-el-yihadismo-hay-que-ganarla-en-la-red
254	http://www.thiber.org/2014/08/29/la-guerra-digital-en-el-conflicto-arabe-israeli/
255	https://www.technologyreview.es/s/4022/si-estamos-en-ciberguerra-donde-estan-las-ciberarmas
256	http://telocuentonews.com/site/nuevos-campos-de-batalla-la-guerra-informatica-digital/
257	http://www.infoguerra.es/
258	https://twitter.com/i/web/status/1099357806753398784
259	http://www.pensamientocritico.org/enrfoj0915.htm
260	http://guillermomarcet.com/la-guerra-por-la-neutralidad-de-la-red/
261	https://observatorio.cisde.es/actualidad/la-ciberguerra-la-amenaza-de-la-quinta-dimension/
262	http://www.planv.com.ec/ideas/ideas/la-guerra-cibernetica-ya-esta-aqui
263	https://elpais.com/internacional/2018/02/19/actualidad/1519058033_483850.html
264	https://www.spindox.it/it/blog/guerra-informatica-stati-uniti-russia/
265	https://mundo.sputniknews.com/tags/keyword_guerra_cibernetica/
266	https://www.adslzone.net/2018/03/16/rusia-ataca-eeuu-confirmado/

267	https://www.larazon.es/internacional/del-califato-en-siria-e-irak-a-la-ciberguerra-LF16648047
268	https://cepymenews.es/estamos-preparados-ciberguerra/
269	http://www.intact01.net/blog/2012/01/el-laberinto-cibernetico-una-introduccion-al-neo-nomadismo/
270	https://www.elcolombiano.com/opinion/columnistas/como-prevenir-una-ciberguerra-MB7102068
271	http://www.innovaes.com/ciberguerra-un-cambio-en-la-estructura-del-conflicto/
272	https://blogs.deusto.es/master-informatica/comienza-la-ciber-guerra/
273	http://www.juegos.com/juegos/juegos-de-guerra
274	https://www.elsaltodiario.com/economia/guerra-contra-dinero-efectivo
275	https://mundo.sputniknews.com/economia/201809281082330071-resistencia-de-pekin-presion-eeuu/
276	https://diarioti.com/la-ciberguerra-paso-de-la-ficcion-a-la-realidad/25060
277	https://www.elcorreo.com/bizkaia/sociedad/201705/13/mentiras-20170512181509.html
278	https://info.nodo50.org/
279	http://www.eveliux.com/mx/El-origen-de-la-Computadoras.html
280	https://www.laondadigital.uy/archivos/10357
281	https://www.elespectador.com/opinion/la-guerra-cibernetica-columna-31699
282	https://www.cybertechprojects.com/news/ciberataques-el-arma-perfecta/
283	http://zetatijuana.com/2018/10/cientificos-combaten-guerra-cibernetica/
284	https://www.elnuevodiario.com.ni/nacionales/48863-eu-guerra-ciberespacio/
285	http://www.ciberderecho.com/anonymous-y-daesh-la-ciberguerra-en-forma-de-ciberguias/
286	http://fernandosantamaria.com/blog/tag/guerra-red/
287	https://lateralia.es/darknet-inmersion-web/
288	https://securelist.lat/flame-un-nuevo-malware-espa-que-esparce-rumores-de-ciberguerra/76388/
289	http://alvaroecija.com/es/ciberseguridad-el-arte-de-la-guerra-2-0/
290	https://www.univision.com/noticias/tecnologia/lista-de-paises-protegidos-en-una-ciberguerra
291	https://www.farodevigo.es/sociedad/2019/02/17/rusia-lista-guerra-cibernetica/2053331.html

292	https://www.clarin.com/ediciones-anteriores/rusia-estonia-guerra-ciberespacio_0_BJgW1W1CYx.html
293	https://piedepagina.mx/guanajuato-la-presa-oculta-del-huachicol.php
294	http://www.sinpermiso.info/textos/las-guerras-mienten
295	https://www.computerworld.es/sociedad-de-la-informacion/una-leccion-de-anonymous-al-colectivo-ti
296	https://peromatech.wordpress.com/2013/07/10/el-ciberespacio-el-quinto-dominio-de-la-guerra/
297	http://uajournals.com/ojs/index.php/cisdejournal/article/view/146
298	https://www.hoy.es/prensa/20061029/sociedad/guerra-navegar-desatado-ciberespacio_20061029_amp.html
299	http://www.iingen.unam.mx/es-mx/Publicaciones/GacetaElectronica/Mayo2016/Paginas/Ciberguerra.aspx
300	https://actualidad.rt.com/programas/zoom/219366-zoom-rt-guerra-ciberespacio
301	https://www.genbeta.com/activismo-online/juegos-de-ciber-guerra
302	https://peru21.pe/mundo/infraestructura-global-internet-ataques-masivos-461576
303	https://nopasanada.mx/2018/03/guerra-cibernetica-electoral/
304	http://iela.ufsc.br/noticia/24-cosas-sobre-isis-y-al-qaeda-que-no-quieren-que-sepas
305	http://www.chilexpress.cl/restricciones-encomiendas-internacionales-envios-extranjero
306	http://www.el-nacional.com/noticias/columnista/guerra-cibernetica-trump_74704
307	https://actualidad.rt.com/actualidad/view/87457-eeuu-perder-ciberguerra-china
308	https://cadenaser.com/programa/2017/08/08/la_ventana/1502212235_357268.html
309	https://www.diariodeibiza.es/opinion/2019/02/18/rusia-lista-guerra-cibernetica/1048200.html
310	https://www.elmundo.es/internacional/2017/01/06/586ea8e422601d34698b4669.html
311	https://elhabanerochekere.blogspot.com/2013/12/los-drones-y-la-guerra-colateral-en-el.html
312	http://lajerusalenfutura.blogspot.es/1517623660/
313	https://www.elespectador.com/tecnologia/ciberguerra-una-realidad-amenaza-al-planeta-articulo-366079
314	http://www.expansion.com/2012/12/27/empresas/digitech/1356640346.html
315	https://www.codigonuevo.com/sociedad/proxima-guerra-mundial-cibernetica

316	http://www.cubahora.cu/politica/operacion-a-gran-escala-de-guerra-en-red-contra-cuba-infografia
317	http://www.enhacke.com/2019/01/04/mapa-de-ciberguerra/
318	http://alvaroecija.com/es/ciberguerra-anonymous-estado-islamico/
319	https://globbsecurity.com/anonymous-quiere-que-te-unas-a-su-ciberguerra-contra-isis-36912/
320	https://www.20minutos.es/noticia/1642816/0/israel/defensa-fronteras/ciberguerra/
321	http://www.adelante.cu/index.php/es/opinion/14979-trump-lleva-la-guerra-al-ciberespacio
322	http://www.elsiglo.cl/2017/11/07/redes-de-guerra-la-militarizacion-de-twitter/
323	https://globbsecurity.com/analisis-deep-web-34788/
324	http://www.practia.global/Perspectiva-Digital/Paginas/Ciberguerra.aspx
325	https://omicrono.elespanol.com/2011/06/la-guerra-de-las-redes-sociales-inunda-la-red-de-botones/
326	https://www.elmundo.es/elmundo/2012/06/07/navegante/1339045745.html
327	https://www.libertaddigital.com/opinion/florentino-portero/la-guerra-en-red-56705/
328	https://www.elcultural.com/revista/letras/Guerra-en-la-red-El-lado-oscuro-de-Google/28576
329	https://cnnespanol.cnn.com/2013/01/26/anonymous-declara-la-ciber-guerra-al-gobierno-de-eeuu/
330	https://www.laopiniondemalaga.es/opinion/2018/05/06/ciberguerra-guerra-reglas/1004870.html
331	https://www.planetadelibros.com.mx/libro-guerra-en-la-red/150628
332	https://pablorossi.cienradios.com/la-ciberguerra-rusia-planea-desconectarse-internet/
333	https://www.eluniverso.com/noticias/2016/09/23/nota/5816866/yahoo-victima-ciberguerra-fria
334	https://laprensa.peru.com/tecnologia-ciencia/noticia-virus-hackers-74789
335	http://www.lostiempos.com/actualidad/opinion/20170707/columna/guerra-cibernetica
336	https://www.pantallasamigas.net/internet-podria-ser-campo-de-batalla-de-la-proxima-guerra-mundial/
337	http://icci.nativeweb.org/boletin/109/editorial.html

338	https://www.vice.com/es_mx/article/nngyp8/anonymous-desmiente-el-futuro-de-la-guerra-cibernetica
339	https://www.pandasecurity.com/spain/mediacenter/seguridad/preparados-contra-ciberguerra/
340	https://hipertextual.com/2014/03/mapa-tiempo-real-ciberguerra
341	https://actualidad.rt.com/actualidad/276545-pentagono-ofensiva-ciberguerra
342	https://computadorasyguerras.wordpress.com/2013/02/16/computadoras-y-conflictos-belicos/
343	http://www.cubahora.cu/politica/escandalo-millonario-en-ciberguerra-de-estados-unidos-contra-cuba
344	https://www.larioja.com/opinion/seguridad-ciberespacio-20170711233003-ntvo.html
345	https://peru21.pe/tecnologia/rusia-desconectara-internet-abril-prepararse-ciberguerra-nndc-459214
346	https://www.larazon.es/internacional/la-guerra-digital-en-el-conflicto-arabe-israeli-MI7219169
347	https://www.minutouno.com/notas/333077-monstermind-el-arma-secreta-eeuu-una-guerra-cibernetica
348	http://www.razonypalabra.org.mx/anteriores/n38/nbuitron.html
349	http://revistasumma.com/22073/
350	https://www.lanacion.com.ar/ciberguerra-t58775
351	https://www.merca2.es/2018-colisionara-internet-realidad/
352	https://www.elespectador.com/tecnologia/guerra-era-digital-articulo-639870
353	https://www.lne.es/sociedad/2017/03/17/guerra-digital-llega-asturias-350/2074459.html
354	https://hipertextual.com/2010/08/estados-unidos-alista-las-defensas-para-la-ciberguerra
355	http://www.saber.ula.ve/handle/123456789/36770
356	http://micrositios.mintic.gov.co/historias/las-ninas-de-la-guerra/
357	https://www.ticketmaster.es/artist/gira-imaginbank-ana-guerra-cepeda-entradas/1009047
358	https://www.lanacion.com.ar/1389991-guerra-en-la-red-google-vs-facebook
359	http://www.diariojornada.com.ar/66567/Sociedad/De_la_era_del_hielo_a_la_de_la_Ciberguerra
360	http://www.lahuelladigital.com/alticnativa/el-primer-manual-de-ciberguerra-por-encargo-de-la-otan/

361	http://www.serdigital.cl/tag/ciberguerra/
362	https://www.ultimahora.com/ciberguerra-n307229.html
363	https://www.elcomercio.com/actualidad/anonymous-declara-guerra-cibernetica-estadoislamico.html
364	http://entreparentesis.org/perfiles-falsos-redes-sociales/
365	https://www.prensa.com/mundo/OTAN-prepara-ciberguerra-refuerzos-Afganistan_0_4890260957.html
366	http://www.sobrepantalles.net/2012/12/neutralidad-de-la-red-la-nueva-guerra-fria/
367	https://www.heraldo.es/noticias/internacional/2017/11/05/cataluna-ciber-guerra-fria-1205864-306.html
368	http://baracuteycubano.blogspot.com/2009/01/la-guerra-informatica.html
369	http://www.ejercito.mde.es/unidades/Valencia/rt21/Historial/index.html
370	http://laescaleradeiakob.blogspot.com/2017/01/la-primera-guerra-informatica.html
371	http://vintegris.info/guerra-cibernetica-001-2/
372	https://www.silicon.es/internet-cosas-brecha-seguridad-2345132
373	https://www.elmundo.es/navegante/2007/11/29/tecnologia/1196327914.html
374	https://www.ivoox.com/podcast-guerra-digital_sq_f1638478_1.html
375	https://www.paginasamarillas.es/f/tortosa/guerra-informatica_202274429_000000002.html
376	https://www.carasycaretas.com.uy/meng-wanzhou-la-guerra-fria-del-siglo-xxi-es-digital/
377	https://www.eldiario.es/theguardian/ISIS-industrializado-ataques-suicida-tactica_0_618038376.html
378	https://www.exordio.com/1939-1945/militaris/espionaje/enigma.html
379	https://www.bloglenovo.es/amenazas-de-la-inteligencia-artificial/
380	https://gestion.pe/mundo/internacional/ee-uu-rusia-china-manejan-guerra-cibernetica-254887
381	https://elpais.com/tecnologia/2016/05/05/actualidad/1462440524_562055.html
382	http://informatica.blogs.uoc.edu/2014/06/26/informaticos-en-el-cine-juegos-de-guerra/
383	https://www.efefuturo.com/noticia/iot-objetivo-cibercrimen/
384	https://cadenaser.com/ser/2013/03/27/ciencia/1364354666_850215.html

385	https://www.silicon.es/osetia__georgia_se_une_a_la_guerra_cibernetica_contra_rusia-2170665
386	https://www.bbc.com/mundo/noticias-internacional-37779456
387	https://www.elespectador.com/tecnologia/guerra-digital-contra-el-estado-islamico-articulo-599804
388	http://www.info7.mx/seccion/amenazan-ciberactivistas-con-intensificar-guerra-informatica/507653
389	https://www.nytimes.com/es/2018/05/01/elecciones-2018-mexico-bots/
390	http://correodelsur.com/politica/20181115_evo-teme-perder-por-guerra-digital.html
391	https://www.dealerworld.es/seguridad/las-infraestructuras-criticas-objetivo-de-la-ciberguerra
392	https://gestion.pe/tecnologia/rusia-desconectara-internet-abril-prepararse-ciberguerra-nndc-258445
393	https://es.slideshare.net/elcondoro/la-guerra-informatica
394	https://invdes.com.mx/politica-cyt-i/internet-de-las-cosas-objetivo-prioritario-del-cibercrimen/
395	https://redhistoria.com/libros-gratis-arte-la-guerra-sun-tzu/
396	https://www.lne.es/sucesos/2018/12/13/gijoneses-detenidos-caer-red-traficaba/2395756.html
397	http://www.irizar.org/MLV-ROA.html
398	https://noticias.perfil.com/2018/04/28/la-guerra-en-el-ciberespacio/
399	https://hipertextual.com/2012/08/eeuu-armas-ciberguerra
400	http://razonesdecuba.cubadebate.cu/articulos/guerra-informatica-contra-cuba/
401	https://as.com/betech/2017/10/02/portada/1506941570_688811.html
402	https://laprensa.peru.com/actualidad/noticia-posible-regular-guerra-cibernetica-28115
403	https://www.sinembargo.mx/19-02-2018/3387810
404	https://rgnn.org/es/2015/07/18/la-ciberguerra-entre-anonymous-y-el-estado-islamico/
405	http://colabora.softwarelibre.gob.ve/anonymous-le-declara-la-guerra-digital-a-israel
406	http://laestrella.com.pa/panama/nacional/proxima-guerra-podria-ciberespacio/23752448
407	https://www.elcomercio.com/app_public.php/actualidad/farc-noticiero-internet-youtube.html

408	https://www.elmundo.es/economia/innovadores/2017/11/03/59fc1d3b46163fc6228b461d.html
409	https://www.muyseguridad.net/2013/08/29/batalla-siria-ciberespacio/
410	https://lampadia.com/analisis/politica/macron-propone-un-ejercito-europeo
411	https://impresa.prensa.com/mundo/agudiza-guerra-informatica-Israel_0_3301170012.html
412	https://www.prensa.com/mundo/Morales-prepararse-guerra-digital-Bolivia_0_4956254347.html
413	https://www.uv.mx/infosegura/general/noti_ciberguerra/
414	https://omicrono.elespanol.com/2014/12/la-ciberguerra-un-antes-y-un-despues-en-la-forma-de-luchar/
415	http://mujeresdenegromadrid.blogspot.com/2019/01/carta-la-red-internacional-de-mujeres.html
416	https://www.casadellibro.com/libro-ciberguerra/9788490971260/3025391
417	https://lasillarota.com/mundo/asi-afectara-a-mexico-la-ciberguerra-que-viene/127849
418	https://globbsecurity.com/topics/ciberguerra/
419	https://www.muyseguridad.net/2015/11/18/anonymous-ciberguerra-isis/amp/
420	http://blog.consultorartesano.com/2010/06/lucha-en-el-campo-de-batalla-empresarial.html
421	https://actualidad.rt.com/actualidad/298389-musk-explicar-responder-criticos-twitter
422	https://actualidad.rt.com/ciencias/view/11691-Apple-y-Adobe-inician-su-particular-guerra-digital
423	https://www.sisap.com/ciber-ataques/ciberguerra/
424	https://www.bbc.com/mundo/internacional/2009/10/091007_1730_primer_terrorista_jg
425	https://www.elmundo.es/economia/2017/03/28/58da300a22601d4a3e8b4649.html
426	http://cronica.uno/fanb-constituira-comando-contraataques-ciberneticos-iii/
427	http://blogs.uab.cat/gmartinez/2014/03/28/cyberwar-la-guerra-en-el-ciber-espacio/
428	https://www.elnuevodiario.com.ni/nacionales/486316-crisis-nicaragua-donald-trump-daniel-ortega/
429	https://israelnoticias.com/militar/f-35-redefine-guerra-f-22/

430	https://www.muyseguridad.net/2019/02/18/guerra-comercial-tambores-ciberguerra/
431	http://seguridadglobalyderechos.com/tag/ciberespacio/
432	https://elcomercio.pe/huellas-digitales/guerra-futbol-noticia-442226
433	https://www.dot1.com.mx/blog/estas-preparado-para-una-guerra-cibernetica
434	http://www.occhidellaguerra.it/olanda-hacker-russi/
435	https://globbsecurity.com/del-cibercrimen-a-la-ciberguerra-a-debate-26131/
436	https://www.prensa.com/tecnologia/Pentagono-prepara-comando-guerra-ciberespacio_0_2576742746.html
437	https://www.laprovincia.es/opinion/2017/11/04/facebook-google-ciberguerra-fria/993968.html
438	http://elheraldoslp.com.mx/2015/08/02/anonymous-amenaza-a-canada-con-iniciar-una-ciberguerra/
439	https://www.pandasecurity.com/spain/mediacenter/tecnologia/pentagono-proxima-gran-guerra-ciberarmas/
440	https://es.rbth.com/articles/2012/03/21/rusia_crea_una_estrategia_de_ciberguerra_16580
441	http://planetadelibrosmexico.com/la-lucha-por-la-ciberguerra/
442	https://www.telesurtv.net/bloggers/La-guerra-cibernetica-EE.UU.--Rusia-20170301-0001.html
443	https://jornadasciberdefensa.es/2018/programa/232/es
444	https://www.diariosur.es/opinion/guerra-ciber-20171126003923-ntvo.html
445	https://www.dw.com/es/ciberguerra-contra-el-narco/a-15507515
446	https://www.abc.es/cultura/20140816/abci-espias-sovieticos-guerra-mundial-201408131403.html
447	http://www.e-tlaxcala.mx/
448	http://www.razonypalabra.org.mx/comunicarte/2003/abril.html
449	https://actualidad.rt.com/video/215321-hackers-protagonizan-guerra-cibernetica-extremistas
450	https://hipertextual.com/2010/12/la-ciberguerra-contra-los-antiwikileaks-se-endurece
451	https://esfap.fap.mil.pe/index.php/programas/gepac
452	https://impresa.prensa.com/mundo/hackers-llevan-guerra-red_0_887911235.html
453	http://www.securitycollege.us/mcne.html
454	https://www.bbc.com/mundo/noticias-internacional-41907637

455	http://ned.pe/
456	https://es.reuters.com/article/topNews/idESMAE87607M20120807
457	https://www.elespectador.com/opinion/solos-en-medio-de-la-ciberguerra-columna-693785
458	https://www.revistavanityfair.es/poder/galerias/imagenes-pascua-militar/11918/image/679410
459	https://www.softzone.es/2014/07/07/el-mapa-de-la-guerra-cibernetica-mundial/
460	https://www.kaspersky.es/blog/drone-gone-in-11-ms/10413/
461	https://peru21.pe/cheka/tecnologia/llego-primera-guerra-digital-12653
462	https://seguranca.uol.com.br/antivirus/dicas/curiosidades/ciberguerra-entenda-essa-ameaca.html
463	https://www.getabstract.com/es/resumen/seguridad-cibernetica-y-guerra-cibernetica/23817
464	https://hipertextual.com/2011/01/ciberguerra-en-dos-actos
465	http://finde.latercera.com/series-y-peliculas/58-mejores-peliculas-en-netflix-febrero-2019/
466	https://www.sia.es/phone/ciberseguridad.html
467	http://www.lavaca.org/notas/la-red-en-contra-de-la-guerra/
468	http://elperiodico-digital.com/2018/11/19/comienza-la-guerra-digital-electoral/
469	https://www.elespectador.com/tecnologia/seguridad-informatica-una-guerra-de-robots-articulo-742986
470	https://www.lne.es/economia/2019/02/06/armas-ciberguerra/2422324.html
471	https://www.levante-emv.com/economia/2019/02/06/armas-ciberguerra/1831483.html
472	https://redhistoria.com/causas-de-la-primera-guerra-mundial/
473	http://estrategia.gobiernoenlinea.gov.co/623/w3-article-73372.html
474	https://www.farodevigo.es/economia/2019/02/06/armas-ciberguerra/2046779.html
475	http://pro-universitarios.com/la-guerra-digital/
476	https://www.vice.com/es_mx/article/wd3jey/el-inquietante-mapa-de-ataques-ciberneticos-en-tiempo-real
477	https://www.prensalibre.com/internacional/eeuu-lanza-ciberguerra-agresiva-contra-estado-islamico/
478	http://www.redq.cl/

479	https://www.welivesecurity.com/la-es/2013/03/22/consecuencias-falso-ciberataque/
480	https://www.clarin.com/mundo/guerra-cibernetica-eeuu-china-juicio_0_rJPxBbacv7l.html
481	https://elpais.bo/morales-no-por-culpa-de-la-guerra-digital-podemos-perder-las-elecciones/
482	https://www.bbc.com/mundo/noticias-38393572
483	https://www.elcolombiano.com/opinion/columnistas/ataques-informaticos-y-guerra-digital-EJ6579677
484	https://www.24con.com/nota/24780-cibercriminalessoldados-de-la-guerra-informatica/
485	https://cubaenresumen.wordpress.com/tag/ciberguerra/
486	http://capitalradio.es/la-guerra-del-s-xxi-se-libra-ciberespacio/
487	https://sp.depositphotos.com/vector-images/ciberespacio-st100.html
488	http://www.resumenlatinoamericano.org/2018/02/17/egipto-al-sisi-se-va-a-la-guerra/
489	https://www.marcialpons.es/libros/ciberguerra/9788490971260/
490	http://www.elmercurio.com/blogs/2018/07/02/61447/La-ciberguerra-es-cara.aspx
491	https://www.3djuegos.com/foros/tema/49135325/1/caballo-de-guerra/
492	https://hipertextual.com/2010/12/ciberguerra-antiwikileaks-evoluciona
493	https://gestion.pe/peru/politica/pentagono-aumento-presupuesto-15-guerra-informatica-112664
494	http://redfilosofia.es/blog/2018/07/14/la-filosofia-en-pie-de-guerra/
495	http://www.expansion.com/economia-digital/innovacion/2017/09/07/59b10c12268e3eba478b4584.html
496	https://www.silicon.es/actualidad/security/cyberwar/page/2
497	https://www.fib.upc.edu/retro-informatica/avui/simulacio.html
498	http://es.rfi.fr/americas/20150114-guerra-informatica-y-terrorismo
499	http://caracol.com.co/radio/2018/01/22/internacional/1516657537_049296.html
500	https://blogs.deusto.es/master-informatica/yahoo-una-cibervictima-de-la-ciberguerra-ciberfria/
501	http://nacionalypopular.com/
502	https://www.libertaddigital.com/opinion/ideas/quien-invento-la-guerra-relampago-21.html
503	https://www.20minutos.es/noticia/534741/0/guerra/mundial/ciberespacio/

504	https://www.milenio.com/internacional/nsa-tiene-nuevo-programa-de-guerra-informatica-snowden
505	http://fic.uas.edu.mx/tag/patricia-corrales-de-guerra/
506	https://es.digitaltrends.com/videojuego/battlefield-v-segunda-guerra-mundial/
507	https://www.pulzo.com/tecnologia/nuevo-proyecto-rusia-busca-desconectarse-internet-PP641132
508	https://info.nodo50.org/Los-refugiados-de-guerra.html
509	https://www.lanacion.com.ar/2208163-la-era-de-la-ciberguerra-y-cualesson-las-herramientas-de-defensa
510	http://www.forodeseguridad.com/artic/discipl/disc_4015.htm
511	http://wradio.com.mx/radio/2009/10/06/internacional/1254862320_890877.html
512	https://www.elmundo.es/navegante/99/noviembre/18/china.html
513	https://www.lavanguardia.com/tecnologia/20161025/411295084081/internet-de-las-cosas-ciberataque.html
514	https://www.tesisenred.net/handle/10803/8607
515	https://www.elheraldo.hn/opinion/616203-210/ciberguerra-la-nueva-justicia-social
516	http://www.lamalatesta.net/product_info.php/products_id/59224
517	https://www.prensalibre.com/guatemala/comunitario/jovenes-ciberneticos-asi-es-la-generacion-z/
518	http://contenido.com.mx/2014/06/como-se-veria-un-mapa-de-la-guerra-cibernetica-global/
519	https://www.elperiodico.com/es/opinion/20101217/guerra-de-guerrillas-en-el-ciberespacio-627698
520	https://www.mineducacion.gov.co/cvn/1665/w3-article-318362.html
521	https://www.lamarea.com/2017/03/18/la-ciberguerra-electoral/
522	https://actualidad.rt.com/video/263350-acusaciones-eeuu-guerra-informatica
523	https://www.elcomercio.com/actualidad/anonymous-desenmascarar-kukluxklan-ciberguerra.html
524	https://www.lavozdegalicia.es/noticia/maritima/2016/09/25/guerra-redes-red/0003_201609G25P39991.htm
525	https://www.esglobal.org/hay-una-ciberguerra-entre-iran-y-arabia-saudi/
526	https://www.libreriatrama.com/ga/libro/ciberguerra_296644
527	http://www.cad.com.mx/historia_del_internet.htm
528	https://epampliega.com/blog/index.php/tag/ciberguerra/

529	https://latam.casadellibro.com/libro-ciberguerra/9788490971260/3025391
530	https://www.eleconomista.es/firmas/noticias/8508873/07/17/Ciberguerra-sin-victimas.html
531	https://www.sisap.com/ciber-ataques/ciberguerra-2/
532	https://www.larazon.es/historico/4384-guerras-intangibles-PLLA_RAZON_347511
533	http://xtremsecure.com.mx/pentagono-aumenta-15-su-presupuesto-para-enfrentar-guerra-informatica/
534	https://hardzone.es/2018/10/04/hackeo-china-chip-espiar-empresas/
535	https://www.ivoox.com/crisis-mundial-guerra-cibernetica-audios-mp3_rf_15850630_1.html
536	https://diarioti.com/china-admite-oficialmente-tener-ciber-ejercito/86564
537	http://www.vertigopolitico.com/articulo/47565/La-era-de-la-ciberguerra
538	http://www.t13.cl/noticia/mundo/el-populismo-y-guerra-informatica-2.0
539	http://www.elmercurio.com/blogs/2016/10/18/45878/Guerra-cibernetica.aspx
540	https://politologoenred.blogspot.com/2017/11/la-ciberguerra.html
541	https://www.vidaextra.com/listas/los-mejores-juegos-de-estrategia-para-pc
542	http://pcworld.com.mx/maven-la-inteligencia-artificial-la-guerra-nuevo-entra-en-polemica/
543	https://www.diariodemallorca.es/economia/2019/02/06/armas-ciberguerra/1389490.html
544	https://cnnespanol.cnn.com/2015/06/05/pornografia-vengativa-la-ciberguerra-contra-las-mujeres/
545	https://www.abc.es/20090330/prensa-internacional-asia/guerra-ciberespacio-20090330.html
546	https://cadenaser.com/emisora/2016/10/06/ser_madrid_sur/1475739090_860053.html
547	https://listindiario.com/tecnologia/2012/01/24/219141/ley-sopa-y-la-primera-guerra-del-ciberespacio
548	https://www.xataka.com/
549	https://israelnoticias.com/iran/iran-derrocar-netanyahu-guerra/
550	https://www.vice.com/es/article/pgmbd7/norse-ataques-ciberneticos-tiempo-real-hackers
551	https://www.puromarketing.com/48/29702/cierre-interviu-tiempo-crisis-papel-guerra-digital.html

552	https://listindiario.com/las-mundiales/2018/09/21/534095/trump-guerra-cibernetica
553	http://www.cctvradio.com/cci---central-corp-inteligence/drones-uav-hackeado-por-insurgentes-iraquies
554	https://www.eitb.eus/es/noticias/internacional/detalle/2080776/rusia--guerra-informatica-rusia/
555	http://www.habanaradio.cu/opiniones/guerra-en-el-ciberespacio/
556	https://www.pandasecurity.com/spain/mediacenter/panda-security/apoya-llamamiento-paris-ciberespacio/
557	https://www.eldiario.es/theguardian/drones-peligro-real-accidente-peor_0_492051435.html
558	https://www.elheraldo.hn/opinion/624756-368/la-ciberguerra-mundial
559	https://www.elmundo.es/elmundo/2009/10/05/navegante/1254751606.html
560	http://www.ipsnoticias.net/1997/07/espana-telefonica-gana-batalla-decisiva-en-la-guerra-digital/
561	https://cnnespanol.cnn.com/video/cnnee-pkg-burke-israel-virus/
562	http://www.prensaislamica.com/nota5459.htmL
563	https://www.babellibros.com/libro/ciberguerra_38204
564	https://hipertextual.com/2011/07/la-maquina-enigma-el-sistema-de-cifrado-que-puso-en-jaque-a-europa
565	https://www.elmundo.es/america/2010/07/26/estados_unidos/1280100715.html
566	https://www.laopinion.es/economia/2019/02/06/armas-ciberguerra/951029.html
567	https://www.clarin.com/mundo/EEUU-declara-ciberguerra-ISIS_0_VkGQklwgb.html
568	https://www.laprovincia.es/economia/2019/02/06/armas-ciberguerra/1144972.html
569	https://www.iis.unam.mx/blog/informa-afadem-sobre-un-sitio-web-con-documentos-sobre-la-guerra-sucia/
570	http://laestrella.com.pa/panama/nacional/empieza-guerra-informatica/23725061
571	http://es.rfi.fr/europa/20180921-la-guerra-cibernetica-o-la-batalla-por-los-datos
572	https://www.heraldo.es/noticias/sociedad/2016/10/29/la-gran-guerra-digital-continua-1136048-310.html
573	http://www.cubahora.cu/politica/estados-unidos-tiende-un-cerco-sobre-venezuela
574	https://hipertextual.com/tag/ciberguerra
575	https://gestion.pe/tecnologia/internet-cosas-maravilla-desprotegida-son-dispositivos-smart-232718

576	http://www.madrimasd.org/blogs/culturadered/la-guerra-de-los-navegadores/
577	https://aptie.es/usa-preparandose-para-la-guerra-del-ciberespacio/
578	https://twitter.com/actualidadrt/status/507823418704932864
579	http://www.operamundi-magazine.com/2012/01/llego-la-primera-guerra-digital.html
580	http://laestrella.com.pa/internacional/america/morales-advierte-guerra-digital/24046930
581	https://www.diarioinformacion.com/opinion/2019/02/15/rusia-lista-guerra-cibernetica/2118350.html
582	https://www.tecmundo.com.br/seguranca/122616-ciberguerra-eua-atacam-coreia-norte-ddos.htm
583	http://redsalasdeteatro.cl/evento/cesar-va-la-guerra/2019-01-07/
584	https://www.oii.ox.ac.uk/videos/la-guerra-mondiale-informatica/
585	https://www.farodevigo.es/portada-o-morrazo/2019/02/23/xose-manuel-seso-duran-carga/2056951.html
586	http://www.expansion.com/actualidadeconomica/analisis/2018/02/12/5a81672e22601d243c8b456b.html
587	http://www.resumenlatinoamericano.org/2017/03/04/la-guerra-cibernetica-eeuu-rusia/
588	https://www.itespresso.es/telefonica-a-la-caza-de-startups-en-espana-183449.html
589	http://www.rtve.es/alacarta/videos/telediario/tdfs3-ciberataque/1875952/
590	https://www.elcomercio.com/opinion/guerra-digital.html
591	http://www.nerdilandia.com/cyberwar-la-guerra-en-el-ciber-espacio/
592	https://hipertextual.com/2013/12/hackear-drones-es-posible
593	https://actualidad.rt.com/actualidad/306340-iran-hackeado-drones-eeuu
594	http://pagina10.com/web/vientos-de-guerra-3/
595	https://www.ofertia.com.mx/tiendas/los-guerra/informatica-electronica/filiales-s-100496
596	http://dinaciberdanna.blogspot.com/
597	https://twitter.com/statuses/1009928833426952195
598	http://www.fhrcuba.org/es/2017/06/la-ciberguerra-de-alejandro-castro-espin/
599	https://www.lanacion.com.ar/1457318-a-las-puertas-de-una-guerra-informatica
600	https://alicia.concytec.gob.pe/vufind/Record/ESUP_d25321ad3706ff81c8e5806112985491

601	http://www.rtve.es/alacarta/videos/telediario/guerras-abiertas-mundo-2019/4915612/
602	http://www.expansion.com/2013/04/23/entorno/aula_abierta/1366707498.html
603	https://www.linguee.com/spanish-english/translation/guerra%2Bcibern%25C3%25A9tica.html
604	https://unioninformatica.org/tag/guerra-de-patentes/
605	http://tecnomundo.es/video-drones/
606	https://www.diarioinformacion.com/politica/2019/02/21/senado-cuentas/2120524.html
607	http://mobile.cubasi.cu/component/k2/itemlist/tag/Ciberguerra%3Fstart%3D30
608	http://inversorlatam.com/se-libra-una-guerra-informatica-a-traves-de-twitter/
609	https://reportarte.es/2015/02/ciberseguridad-ciberguerra-estamos-preparados-para-el-apagon/
610	https://www.prensalibre.com/internacional/iran-defensa-informatica-mar-0-838716212/
611	http://haadas.blogspot.com/2010/08/guerra-en-el-ciberespacio.html
612	https://www.diariodemallorca.es/opinion/2019/02/15/rusia-lista-guerra-cibernetica/1392096.html
613	https://www.diariosur.es/sociedad/201604/11/infierno-hitler-20160410185853.html
614	http://www.t13.cl/noticia/mundo/10-cosas-que-quizas-no-sabias-de-la-guerra-de-vietnam
615	https://www.20minutos.es/noticia/1178696/0/congo/coltan/moviles/
616	https://www.nodo50.org/pretextos/colombia-sua.htm
617	https://www.newtonpaiva.br/home
618	https://www.paginaciudadana.com/la-guerra-digital-de-las-elecciones/
619	http://puentelibre.mx/seccion/tecnologia/1
620	https://www.laopiniondemalaga.es/opinion/2009/12/06/guerra-red/307101.html
621	http://eju.tv/2018/10/morales-pide-ayuda-a-china-para-enfrentar-la-guerra-digital/
622	https://www.telesurtv.net/news/Revelan-nuevo-programa-de-guerra-informatica-20140813-0082.html
623	https://twitter.com/jallorente/status/1063800125913538560
624	https://www.clarin.com/mundo/dilma-ciberespacio-puede-usado-guerra_0_BJuD_0mjD7e.html

625	https://fuentelatina.org/la-guerra-los-drones-ciberseguridad-frente-los-hackers/
626	https://elpais.bo/tag/guerra-informatica/
627	https://www.lne.es/sociedad/2019/02/14/rusia-lista-guerra-cibernetica/2426491.html
628	https://www.clarin.com/economia/efectos-red-guerra-taxis_0_B1iWtQ2cvmg.html
629	https://www.udllibros.com/libro-ciberguerra-7040300583
630	https://reforma.vlex.com.mx/vid/hacen-guerra-ciberespacio-81274422
631	https://twitter.com/VAguileraDiaz/status/963011799971966976/photo/1
632	https://www.urgente24.com/249553-tacticas-y-herramientas-para-la-guerra-digital
633	https://www.esglobal.org/el-oscuro-arte-de-la-ciberguerra/
634	https://www.microsoft.com/es-mx/p/red-faction-armageddon-el-camino-a-la-guerra/c3tx0kmj1v7h
635	http://www.diariosigloxxi.com/texto-diario/mostrar/431738/ciberataques-guerra-informatica
636	http://revistas.upcomillas.es/index.php/internationalrelations/article/view/10698
637	https://peru21.pe/opinion/opina21-juan-claudio-lechin/rusia-perfecciona-guerra-cibernetica-386504
638	https://israelnoticias.com/franja-de-gaza/hamas-israel-espiar-lider-gaza/
639	https://blogs.protegerse.com/2014/09/24/blackenergy-ciberguerra-en-ucrania-y-en-polonia/
640	http://radiofidescochabamba.com/morales-pide-ayuda-a-china-para-enfrentar-la-guerra-digital/
641	http://www.cefadigital.edu.ar/handle/123456789/57
642	https://www.eitb.eus/es/noticias/detalle/2906628/analisis-ana-aizpiri--guerra-red/
643	http://www.ehu.eus/ojs/index.php/HC/article/view/17711
644	https://prosulting.es/tag/ciberguerra/
645	https://heraldocubano.wordpress.com/tag/guerra-informatica/
646	https://www.bigbangnews.com/mundo/La-guerra-cibernetica-en-vivo-y-en-tiempo-real-20150605-0030.html
647	http://indicepolitico.com/programas-ciberguerra-son-un-grave-riesgo-de-proliferacion/
648	https://news.sophos.com/es-es/tag/ciberguerra/
649	https://autismodiario.org/

650 https://www.elnorte.ec/editorial/articulistas/guerra-informatica-AHEN2002

651 https://riunet.upv.es/handle/10251/51471

652 http://www.telemadrid.es/programas/zoom-telemadrid/Ciberguerra-2-1757244260--20160115082454.html

653 https://www.20minutos.es/noticia/2698935/0/anonymous/ciberguerra-total/donald-trump/

654 https://noticieros.televisa.com/ultimas-noticias/principal-reto-internet-cosas-es-seguridad/

655 https://www.libertaddigital.com/opinion/ideas/matematicas-contra-la-guerra-1232.html

656 https://lahistoriadecuba.wordpress.com/2019/02/18/almeida-vive-hoy-mas-que-nunca/

657 http://www.trabajadores.cu/20140425/alerta-la-guerra-se-da-en-el-ciberespacio-tambien/

658 http://www.lajiribilla.cu/articulo/oliver-stone-snowden-y-la-ciberguerra-mas-peligrosa-del-mundo

659 https://smokingvictims.bandcamp.com/track/guerra-en-el-ciberespacio

660 http://repositorio.uasb.edu.ec/handle/10644/969

661 https://innovadores.larazon.es/es/not/no-a-la-ciberguerra

662 http://virtual.esup.edu.pe/handle/ESUP/113

663 https://www.silicon.es/las-ventas-de-smartphones-se-estan-estancando-2391395

664 http://www.iese.edu.ar/

665 https://www.elcultural.com/revista/letras/Los-piratas-de-Somalia-dentro-de-su-mundo-oculto/29903

666 https://www.lacuarta.com/

667 https://www.muyseguridad.net/2014/12/16/drones-espia/

668 https://www.lanuevacronica.com/guerra-garrido-recorre-la-escalera-entre-las-eras-analogica-y-digital

669 http://www.redambiental.com/tag/horacio-guerra-marroquin/

670 http://www.cefadigital.edu.ar/handle/123456789/249

671 https://www.diariocordoba.com/noticias/opinion/guerra-ciberespacio_604694.html

672 https://www.elperiodico.com/es/internacional/20190117/eeuu-ampliar-sofisticar-misiles-7250700

673 http://ericacg33.wixsite.com/ciberguerra

674	https://www.laprensa.com.ni/tag/guerra-cibernetica
675	https://www.elmundo.es/deportes/formula-1/2018/04/30/5ae64670268e3e4a588b45d2.html
676	http://revistas.ujat.mx/index.php/ecosoc/article/view/2250
677	https://www.saq.com/page/en/saqcom/red-vermouth/guerra-reserva-rojo/13038981
678	https://www.telesurtv.net/news/Cuba-rechaza-ante-la-ONU-la-ciberguerra-de-EE.UU.-20140927-0040.html
679	https://www.muyinteresante.com.mx/
680	https://losandes.com.ar/article/estrategia-la-guerra-de-zapa
681	http://america24.com/news/mueller-la-russia-ha-condotto-una-guerra-informatica-contro-gli-usa
682	https://agenda.deusto.es/taller-coloquio-espionaje-ciberguerra-y-terrorismo/
683	https://espacioseuropeos.com/tag/ciberguerra-de-anonymous/
684	https://www.anahuac.mx/mexico/EscuelasyFacultades/estudiosglobales/diplomados
685	https://webs.ucm.es/info/especulo/numero8/d_nora.htm
686	https://www.harpersbazaar.com/es/cultura/ocio/a323190/cine-streaming-batalla-guerra-digital/
687	https://www.elcultural.com/revista/opinion/La-Edad-Digital/39150
688	http://planoinformativo.com/
689	https://www.meneame.net/story/ciberguerra-matar-hackers-civiles-legal
690	https://as.com/betech/2016/10/28/portada/1477685427_331526.html
691	https://www.elmundo.es/elmundo/2010/01/22/navegante/1264147906.html
692	https://as.com/betech/2017/09/04/portada/1504551983_319168.html
693	https://www.tiempoar.com.ar/nota/aguad-prepara-las-fuerzas-armadas-para-la-ciberguerra
694	http://www.magarciaguerra.com/2011/09/informatica-sjb-mi-blog-de-aula/
695	http://www.telemadrid.es/deportes/Alonso-ironiza-guerra-Red-Bull-0-1449455058--20130325020050.html
696	https://www.tiempoar.com.ar/nota/la-guerra-por-los-datos-personales-el-oro-de-la-era-digital
697	https://www.elperiodico.com/es/entre-todos/participacion/independencia-guerra-digital-144360
698	https://www.lavoz.com.ar/categoria/temas-libres-9663

699	https://www.ambito.com/anonymous-le-declaro-la-guerra-informatica-trump-n3969683
700	http://www1.udel.edu/leipzig/071198/elf071298.htm
701	http://www.elfarocultural.com/2012/02/sopa.html
702	https://www.focus.it/temi/guerra-informatica
703	http://joaquin-rodrigo.com/tienda/cantos-de-amor-y-de-guerra-piano-red.html
704	https://www.paginasiete.bo/opinion/editorial/2018/2/6/guerra-digital-est-instalada-168993.html
705	http://blogs.eltiempo.com/wikimundo/tag/internet-de-las-cosas/
706	http://clio.rediris.es/udidactica/IGM/
707	https://es.euronews.com/2018/04/10/rusia-hackea-drones-estadounidenses-en-siria
708	https://twitter.com/hashtag/ciberguerra
709	https://www.uexternado.edu.co/tag/ciberguerra/
710	https://www.elnuevodiario.com.ni/tag/guerra-informatica/
711	https://www.infodefensa.com/latam/2015/02/05/noticia-uruguay-blinda-contra-ataques-ciberneticos.html
712	http://spanish.peopledaily.com.cn/n3/2019/0222/c31617-9548901.html
713	https://blog.maestrotvsnte.mx/el-uso-de-los-drones-como-funcionan-que-hacen-tipos/
714	https://www.marca.com/2009/12/19/baloncesto/nba/1261241537.html
715	https://www.nytimes.com/es/2018/05/14/alan-turing-patrones-informatica/
716	https://blog.tacticalsecurity.net/noticias/estudiante-aleman-hackear-drones/
717	https://hipertextual.com/archivo/imagen-del-dia/ciberguerra-kaspersky-mapa/
718	https://www.mediamarkt.es/es/category/_campa%25C3%25B1as-y-ofertas-701114.html
719	https://www.ccma.cat/tv3/30-minuts/hackers-1-o-a-30-minuts/noticia/2859010/
720	https://www.elperiodicodearagon.com/noticias/temadia/guerra-ciberespacio_50460.html
721	http://www.rezagos.com/news/157-guerras-tribales-por-el-coltan-el-lado-oscuro-de-la-tecnologia.html
722	https://www.cifraclub.com.br/katalexia/ciber-guerra/letra/
723	https://diocesanos.es/blogs/equipotic/2014/05/13/el-legado-de-la-guerra-fria-en-la-informatica/

724	https://www.adslzone.net/foro/miscelanea.30/guerra-ciberespacio-esta-servida.104673/
725	http://emprenderioja.es/blog/2013/07/29/la-estrategia-y-la-tactica/
726	https://www.panorama.it/news/esteri/putin-le-elezioni-russe-e-la-guerra-informatica/
727	https://www.booksonline.es/es/libros/ciberguerra_ALO0052899
728	https://www.carrefour.es/ciberguerra/9788490971260/p
729	https://www.enimbos.com/es/ciberguerra-sin-victimas-n39
730	http://www.unalineasobreelmar.net/2011/02/14/paisaje-sonoro-de-anonymous-en-los-goya/
731	https://www.elmundo.es/opinion/2018/01/18/5a5f9677468aebac798b4586.html
732	http://www.marcotradenews.com/noticias/la-guerra-por-el-cobalto-el-oro-azul-61109
733	http://www.elhorizonte.mx/opinion/editorial/la-guerra-digital-de-2018/2011252
734	https://news.un.org/es/tags/ciberespacio
735	http://andreselias.com/2016/02/09/conoce-de-que-se-trata-el-taller-el-arte-de-la-guerra-digital/
736	http://edgar-guerra.tripod.com/
737	http://carlesm.com/2009/06/guia-de-ciberguerra-para-iranelection/
738	http://www.lr21.com.uy/mundo/86166-tv-israeli-en-lengua-arabe-nueva-arma-en-la-guerra-informatica
739	https://tecnologeria.com/bit/1x07/alan-turing-padre-informatica
740	https://www.libertaddigital.com/opinion/george-will/cuando-la-guerra-era-la-respuesta-40094/
741	https://www.netflix.com/ai-es/title/80092412
742	http://www.cimamalaga.com/primera-guerra-informatica-mundial/
743	https://www.finect.com/usuario/inveo.es/articulos/la_guerra_de_las_galaxias
744	http://cms.ual.es/UAL/universidad/organosgobierno/gabcomunicacion/noticias/31OCT2013_SEMINARIOLEGION
745	https://www.teletica.com/179413_la-guerra-sucia-en-la-red
746	http://www.cronicasdelanzarote.es/
747	http://notinerd.com/9-delitos-que-se-cometieron-con-la-ayuda-de-drones/
748	https://www.diariodeleon.es/noticias/cultura/guerra-digital-cairo_822136.html
749	http://www.aipaz.org/content/agenda/item/402-na-procura-de-alternativas

750	http://www.unionpuebla.mx/articulo/2019/02/22/espectaculos/guerra-por-rating-azteca-contra-televisa
751	https://www.sport.es/es/noticias/opinion/arturo-vidal-para-guerra-7320520
752	http://dss.collections.imj.org.il/es/war
753	https://odisea.es/programas/guerra-cibernetica/
754	http://www.noticiasurbanas.com.ar/noticias/9ad326f370afa22d54823f654cd3fc49/
755	https://www.americatv.com.pe/cuarto-poder/guerra-politica-twitter-noticia-14293
756	https://elperiodico.com.gt/opinion/2018/05/11/orwell-y-la-guerra-oscura-por-el-ciberespacio/
757	http://www.acn.cu/2011/junio/23ed-informatica.htm
758	https://www.elsotano.com/libro-ciberguerra-10496770
759	http://www.cronicaviva.com.pe/tag/guerra-informatica/
760	http://www.rae.es/
761	http://redsalud.uc.cl/ucchristus/Directorio/claudio-guerra-sanchez.act
762	http://caracol.com.co/tag/guerra_digital/a/
763	https://www.nytimes.com/es/tag/ciberguerra/
764	https://hipertextual.com/2017/07/hackear-firmware-drones-nueva-frontera-control-libertad
765	http://www.patrianueva.bo/
766	http://catedras.uca.es/jean-monnet/noticias/agora-seguridad-uca
767	https://www.elmundo.es/tecnologia/2015/05/07/554b1d1a268e3ef5028b456d.html
768	https://www.lanacion.com.ar/1844519-los-senores-de-la-guerra-digital
769	https://www.elperiodicodearagon.com/noticias/temadia/guerra-ciberespacio_48271.html
770	https://www.uoc.edu/portal/es/index.html
771	https://elcomercio.pe/noticias/guerra-cibernetica
772	https://brecha.com.uy/la-guerra-comercial-estados-unidos-china-y-la-batalla-por-la-red-5g/
773	https://www.clarin.com/tema/ciberguerra.html
774	http://hackmadrid.org/objetivos.html
775	https://www.todostuslibros.com/materia/ta-ctica-militar_JWKT
776	https://elmundo.sv/guerra-feroz-en-el-liderato-y-el-descenso/

777	https://catalog.hathitrust.org/Record/004014640
778	https://www.fnac.es/a1236228/Ciberguerra
779	https://www.clarin.com/tecnologia/Hackers-prometen-escalada-informatica-WikiLeaks_0_BJb94tpP7g.html
780	https://www.ekoi.com/en/guerra/5279-guerra-red-right-arm-4150000527901.html
781	https://www.uninorte.edu.co/web/ahernand/home/-/blogs/%25C2%25BFque-es-la-ciberguerra-
782	https://vozlibre.com/tag/guerra-cibernetica/
783	http://vtv.gob.ve/iran-hackeo-drones-ee-uu-videos-aeronaves-incautadas/
784	http://delaurbe.udea.edu.co/etiqueta/guerra-digital/
785	https://www.wdl.org/es/item/4637/
786	https://revistas.ucm.es/index.php/RUNI/article/download/49643/46161
787	https://www.20minutos.es/noticia/209221/0/bloguero/declara/guerra/
788	http://lajiribilla.cu/articulo/la-guerra-se-libra-hoy-tambien-por-el-ciberespacio
789	http://wradio.com.mx/programa/2013/09/24/audios/1380073440_978294.html
790	https://warframe.fandom.com/es/wiki/Natah
791	https://www.diariovasco.com/sociedad/salud/centros-salud-guerra-20190217005456-ntvo.html
792	https://www.leadersummaries.com/ver-resumen/como-ganar-una-guerra
793	http://www.esgcffaa.edu.ar/esp/oac-presentacion.php
794	http://www.polifemo.com/libros/ciberguerra/132037/
795	https://www.zendalibros.com/tag/ciberguerra/
796	https://www.baenegocios.com/mundo/Holanda-en-guerra-cibernetica-con-Moscu-20181014-0038.html
797	https://www.mensaje.cl/edicion-impresa/mensaje-667/la-guerra-por-las-mentes-en-el-ciberespacio/
798	https://www.tesisenred.net/handle/11268/1272
799	http://www.madrid.org/bibliotecas/
800	https://m.lavoz.com.ar/temas/guerra-cibernetica
801	https://www.uniba.it/eventi-alluniversita/2017/allegati/LocandinaGuerrainformatica.jpg/view
802	http://lirelactu.fr/source/la-vanguardia/bbdce4e9-46ec-4901-b669-58b9682e6bfd
803	https://www.pandasecurity.com/spain/mediacenter/src/uploads/2011/10/Informe-PandaLabs-Q3-2011.pdf

804 https://issuu.com/edicionesdegue/docs/boletin_ciberespacio_v-1_ed-1

805 https://peru21.pe/noticias/guerra-informatica

806 https://www.diariocordoba.com/noticias/sociedad/segunda-guerra-digital_170600.html

807 http://redmentores.3ie.cl/mentor/lautaro-guerra/

808 https://noticieros.televisa.com/videos/analisis-de-la-ciberguerra-entre-paises/

809 https://listado.mercadolibre.com.mx/spiderman-arana-escarlata-ciberguerra

810 http://www.rai.tv/dl/RaiTV/programmi/media/ContentItem-05be5c86-30c6-48a8-a198-7f63df812197.html

811 https://idus.us.es/xmlui/handle/11441/29967

812 https://prezi.com/ncnm5dmkzw0q/guerra-informatica/

813 http://www.redbolivision.tv.bo/actualidad/evo-teme-a-la-guerra-digital-62790

814 https://www.kkbox.com/tw/tc/song/Viq009H.29XF15LpF15Lp0XL-index.html

815 https://www.redbullbatalladelosgallos.com/noticias/invert-lanza-grito-de-guerra-junto-a-mowlihawk

816 https://www.netflix.com/us-es/title/80092412

817 http://www3.gobiernodecanarias.org/medusa/ecoescuela/expresate/tag/ies-valle-guerra-de-tenerife/

818 https://prezi.com/aehe3j-3ee-i/que-es-guerra-cibernetica/

819 http://revistas.ucm.es/index.php/CIYC/article/view/36997

820 https://issuu.com/178anasofiacordobaardila/docs/sin_t__tulo_1

821 https://afondo.levante-emv.com/valencia/guerra-a-las-terrazas-de-ciutat-vella.html

822 http://132.248.9.195/ptd2013/agosto/0699056/0699056.pdf

823 http://www.codajic.org/node/1861

824 http://www.bioeticayderecho.ub.edu/es/la-guerra-por-los-datos-personales-el-oro-de-la-era-digital

825 https://drone-shadow-strike.uptodown.com/android

826 https://redhydrantpress.com/

827 http://www3.gobiernodecanarias.org/medusa/ecoblog/npergue/category/informatica/

828 https://revistas.ucm.es/index.php/DCIN/article/viewFile/54420/49730

829 http://www.gacetamercantil.com/notas/62627/

830 https://prezi.com/1wc8y04m5fq5/la-gran-guerra-digital/

831 https://prezi.com/dfvyt1svqats/tacticas-de-guerra-de-la-primera-guerra-mundial/

832 https://definicion.mx/dron/

833 https://www.redentradas.com/

834 https://prezi.com/ltsrhwwzzcvl/ciberguerra/

835 https://www.msn.com/es-es/dinero/business-class/drone-selfie-on-a-waterfall/vi-BBvMCBa

836 https://www.elsalvador.com/tag/guerra-informatica/

837 https://www.colombia.com/

838 https://revistas.ucm.es/index.php/UNIS/article/download/38473/37212

839 http://www.translators.cl/es-noticia.php%3Fid%3D1831

840 http://www.maps.com.mx/index.php/component/tags/tag/ciberguerra

841 https://actualidad.rt.com/actualidad/view/82107-guerra-cibernetica-mayor-amenaza-futuro/video

842 https://glpsat.blogspot.com/2018/05/tu-router-de-casa-un-arma-de.html

843 http://www.ecodias.com.ar/art/y-trump-le-declara-la-guerra-cibern%25C3%25A9tica-rusia-y-china

844 https://prezi.com/islcabatlfrr/drones/

845 http://tuertoperoveotodo.blogspot.com/2016/08/guerra-digital-o-ciber-guerra.html

846 https://turquinauta.blogspot.com/2013/01/ciberguerra-contra-cuba-twitter-cierra.html

847 https://www.infosecuritymexico.com/es/session-details.1163.2542.Sesi%25C3%25B3nPandaSecurity.html

848 https://www.pinterest.es/pin/188940146851683702/

849 https://www.dailymotion.com/video/x6gwkem

850 https://www.redalyc.org/pdf/267/26748302002.pdf

851 https://www.reforma.com/aplicaciones/articulo/default.aspx%3Fid%3D712787

852 https://www.reforma.com/aplicaciones/articulo/default.aspx%3Fid%3D1336431

853 https://dialnet.unirioja.es/descarga/libro/413843.pdf

854 http://puentelibre.mx/noticia/57763-memes_atacaron_en_la_ciberguer/2

855 http://tuertoperoveotodo.blogspot.com/2018/07/que-es-la-ciber-guerra-dedicado-al.html

856 https://www.reforma.com/aplicaciones/articulo/default.aspx%3Fid%3D1327161

857 https://www.coe.int/es/web/compass/war-and-terrorism

858	https://diarioti.com/la-otan-publica-manual-de-ciberguerra/62351
859	https://dialnet.unirioja.es/servlet/articulo%3Fcodigo%3D5998149
860	https://blogs.ua.es/pi/2012/06/28/la-guerra-de-los-navegadores/
861	http://www.elprofesionaldelainformacion.com/contenidos/2002/julio/11.pdf
862	http://www.belt.es/noticias/2003/abril/14_18/14/Ciberguerra_15.htm
863	https://pt.scribd.com/doc/316764266/Castells-Manuel-Redes-de-indignacion-y-esperanza-pdf
864	http://dernegocios.uexternado.edu.co/comercio-electronico/nos-llego-la-ciberguerra/
865	http://www.euskomedia.org/PDFAnlt/riev/56/56520572.pdf
866	https://www.leadersummaries.com/ver-resumen/el-arte-de-la-guerra-hoy
867	https://noticieros.televisa.com/historia/requisitos-ser-maestro-universidades-amlo/
868	https://www.swhosting.com/blog/ciber-guerra-de-internet-nuestras-vidas/
869	http://www.au.af.mil/au/afri/aspj/apjinternational/apj-s/2012/2012-3/2012_3_05_lee_s.pdf
870	http://www.imaginar.org/taller/ciberdefensa/D1_04_desafios_gnavarro.pdf
871	https://www.amazon.com.br/Guerra-red-nuevos-campos-batalla/dp/843446960X
872	https://es.scribd.com/document/391053579/LA-CRIPTOMONEDA-Cibercrimen-y-el-lavado-de-activos
873	https://ar.pinterest.com/pin/564005553318961385/
874	https://www.directivosyempresas.com/internet/tecnologia/etica-en-la-inteligencia-artificial/
875	https://www.icrc.org/spa/resources/documents/interview/2013/06-27-cyber-warfare-ihl.htm
876	https://www.itespresso.es/asi-es-galaxy-fold-el-smartphone-plegable-de-samsung-183461.html
877	https://actualidad.rt.com/actualidad/251534-assange-primera-guerra-mundial-internet-cataluna
878	http://www.aedhe.es/wp-content/uploads/2018/11/HackingShodan_sin-video_AEDHE.pdf
879	http://thiber.org/wp-content/uploads/2018/10/Numero02_Octubre_2018_opt.pdf
880	https://www.muyhistoria.es/contemporanea/articulo/internet-hija-de-la-guerra-fria-691487081931
881	https://www.razon.com.mx/el-cultural/la-era-del-dron-del-ciberpunk-a-la-guerra-sin-fin/

882	https://www.elconfidencial.com/tecnologia/2018-10-04/china-amazon-apple-espionajke_1625497/
883	https://www.zaragoza.es/
884	https://moscovita.org/mosconews/cuanto-cuesta-el-concepto-avelina-lesper/
885	https://www.lanacion.com.ar/742323-con-mas-censura-china-le-declara-la-guerra-a-internet
886	http://www.computerworld.es/archive/la-guardia-civil-se-prepara-contra-el-ciberterrorismo
887	https://www.icrc.org/spa/resources/documents/faq/130628-cyber-warfare-q-and-a-eng.htm
888	http://red.pucp.edu.pe/wp-content/uploads/biblioteca/081020.pdf
889	https://www.sopitas.com/geek/internet-guerra-cibernetica-ataque-ddos-definicion/
890	https://www.diariogol.com/
891	https://dialnet.unirioja.es/descarga/articulo/4276097.pdf
892	https://www.lavoz.com.ar/temas/internet-4
893	http://theappfest.com/portfolio-item/chema-alonso-hacking-ciberguerra-y-otras-palabrotas/
894	https://www.amazon.com.mx/guerra-Internet-%25C3%25BAltima-batalla-controlar-ebook/dp/B00N84NXTK
895	http://aviles.es/
896	http://cespe.espe.edu.ec/tag/ciberguerra/
897	https://revistasic.es/wp-content/uploads/2018/03/129-bibliografia.pdf
898	https://www3.uc3m.es/reina/Fichas/Idioma_1/288.12395.html
899	http://www.ieee.es/Galerias/fichero/cuadernos/CE_149_Ciberseguridad.pdf
900	https://www.icrc.org/spa/resources/documents/misc/5tecg3.htm
901	http://bdex.eb.mil.br/jspui/bitstream/1/1015/1/TESE_DORNELES.pdf
902	http://www.bne.es/es/Micrositios/Guias/Guerra_independencia/GIInternet/
903	https://www.uoc.edu/uocpapers/7/dt/esp/tubella_tabernero_dwyer.html
904	http://mittic.gobex.es/upload/web/publico/06_uid1.SeguridadparaPYMESGuerra.pdf
905	https://twitter.com/monigps/status/1075414103102054400
906	http://polux.unipiloto.edu.co:8080/00001363.pdf
907	https://www.bbc.com/mundo/noticias-46475391
908	http://www.sinembargo.mx/02-07-2017/3253471

909	http://www.elladodelmal.com/2013/01/incidentes-de-ciberguerra-y.html
910	http://www.portalcultura.mde.es/noticias/home/2016/Junio/PresentaLibroCiberguerra.html
911	https://www.amazon.com.br/El-Arte-Guerra-Internet-Spanish-ebook/dp/B07LDMNSFJ
912	https://www.redbull.com/mx-es/total-war-arena-entrevista
913	http://www.acn.cu/2011/junio/26gct-economia-ciberguerra.htm
914	https://lapupilainsomne.wordpress.com/tag/ciberguerra/
915	https://www.amazon.com/Guerra-red-nuevos-batalla-Spanish-ebook/dp/B006GFT8F6
916	http://www.foixblog.com/2013/06/23/la-guerra-fria-se-libra-en-el-ciberespacio/
917	http://www.hugoperezidiart.com.ar/teoria-aplicada-2014/Kaldor-2006.pdf
918	https://www.amazon.de/Ciberguerra-Yolanda-Quintana-Serrano/dp/8490971269
919	https://www.elheraldo.co/internacional/corea-del-norte-armas-nucleares-o-ciberguerra-183292
920	http://biblio3.url.edu.gt/Libros/provinciales/computadoras.pdf
921	https://www.levante-emv.com/espana/2019/02/21/juicio-proces-retoma-hoy-turno/1838624.html
922	https://www.megustaleer.com/libros/guerras-de-internet/MAR-009104
923	http://www.uypress.net/auc.aspx%3F17361
924	http://www.biblioteca.org.ar/libros/656228.pdf
925	https://www.bbc.com/mundo/noticias/2013/02/130212_ciberguerra_ataque_preventivo_efectividad_wbm
926	https://www.dsn.gob.es/sites/dsn/files/ESN2017_capitulo_4.pdf
927	http://partidodeltrabajo.org.mx/2011/seminarioXXII/site/docs/2313.pdf
928	https://cnnespanol.cnn.com/2017/08/21/paises-robots-lideres-robotica-mejores-mundo/
929	http://conexioninversa.blogspot.com/2013/01/ciberguerra-ciberdefensa.html
930	https://www.amazon.de/El-Arte-Guerra-en-Internet/dp/1792922965
931	http://www.ieee.es/Galerias/fichero/docs_opinion/2016/DIEEEO86-2016_Ciberamenazas_JRuizDiaz.pdf
932	https://historiayvideojuegos.com/wp-content/uploads/attachments/50.pdf
933	http://www.vitoria-gasteiz.org/docs/a25/000000000/000550000/550749.pdf
934	http://www.academia.edu/6182513/La_Ciberguerra_la_guerra_inexistente

935	http://www.brigadasinternacionales.uclm.es/wp-content/uploads/2017/09/Anexo.pdf
936	http://sedici.unlp.edu.ar/bitstream/handle/10915/40210/Documento_completo.pdf%3Fsequence%3D1
937	https://dialnet.unirioja.es/descarga/articulo/836702.pdf
938	http://www.diputados.gub.uy/informes_viajes/528-132a-asamblea-de-la-union-interparlamentaria/
939	http://panamericana.pe/24horas/tecnologia/99740-se-inicio-la-guerra-por-controlar-el-ciberespacio
940	https://www.eldiario.es/internacional/Aplicaciones-internet-vuelven-politica-India_0_754525253.html
941	http://www.ugr.es/~gesi/yihadismo-ciberamenaza.pdf
942	http://redfilosofia.es/congreso/wp-content/uploads/sites/4/2017/07/2.9.pdf
943	https://www.edicionesobelisco.com/libro/2010/de-la-guerra-digital
944	https://www.linguee.com/spanish-english/translation/los%2Bpeligros%2Bde%2Binternet.html
945	https://www.amazon.com/Guerra-Digital/dp/B0723DZQLR
946	http://www.sinembargo.mx/08-05-2013/612555
947	http://www.expansion.com/2013/06/21/empresas/tmt/1371843239.html
948	https://cadenaser.com/emisora/2018/12/14/radio_murcia/1544815342_472230.html
949	http://www.bbc.co.uk/blogs/mundo/un_mundo_feliz/2010/12/2010_el_ano_de_la_ciberguerra.html
950	https://guerradental.com/internet-specials/
951	http://ri.ujat.mx/bitstream/20.500.12107/1688/1/250-1792-A.pdf
952	https://www.ugr.es/~gesi/Guerras-hibridas.pdf
953	https://www.jurisway.org.br/v2/dhall.asp%3Fid_dh%3D12506
954	https://www.uv.mx/infosegura/general/noti_ciberataques/
955	http://www.diariodeleon.es/noticias/deportes/vettel-desata-guerra-red-bull_781043.html
956	https://pt.scribd.com/document/381867986/Ciberespacio-y-resistencias-pdf
957	https://www.stitcher.com/podcast/anchor-podcasts/la-guerra-digital/e/57192338%3Fautoplay%3Dtrue
958	https://rpp.pe/tema-ciberguerra
959	http://www.pensamientopenal.com.ar/system/files/2016/02/doctrina42952.pdf

960 http://pablocoraje.com/el-gigante-amazon/

961 http://www.scielo.org.ar/pdf/cts/v6n18/v6n18a02.pdf

962 https://www.ivoox.com/podcast-podcast-caibarien-audio_sq_f158159_1.html

963 https://www.14ymedio.com/cienciaytecnologia/guerra-Iran-internet-victima-Telegram_0_2429157062.html

964 https://dialnet.unirioja.es/descarga/articulo/3745519.pdf

965 https://id.scribd.com/doc/316764266/Castells-Manuel-Redes-de-indignacion-y-esperanza-pdf

966 https://www.hawkersco.com/

967 https://repositorio.cepal.org/bitstream/handle/11362/21681/S2013023_es.pdf

968 https://moscovita.org/mosconews/venezuela-ser-o-no-ser/

969 https://www.genbeta.com/activismo-online/el-pentagono-recibe-luz-verde-para-la-guerra-en-internet

970 http://www.revista-redi.es/wp-content/uploads/2017/08/8_estudios_moran_blanco_ciberseguridad.pdf

971 http://www.adizesca.com/site/assets/en-el_arte_de_la_guerra_sun_tzu-ld.pdf

972 http://www.portafolio.co/economia/finanzas/tv-nuevo-frente-guerra-digital-359234

973 https://www.amazon.es/guerra-Internet-ultima-batalla-controlar/dp/1502447371

974 https://posmomagazine.wordpress.com/2015/02/19/posmo-investigaciones-drones-de-guerra/

975 https://mundo.sputniknews.com/firmas/201902211085623157-recesion-de-economia-de-alemania/

976 http://www.reclutamiento.defensa.gob.es/como-ingresar/tropa-marineria/buscador-plazas/

977 https://www.clarin.com/sociedad/guerra-internet-instagram-desaparecen-twitter_0_S1nx1R3jv7e.html

978 https://www.amazon.com.mx/cyberdelincuencia-cyberespionaje-y-Guerra-informatica-ebook/dp/B00HBT9CPW

979 https://www.cruzroja.es/

980 http://www.revistarazonypalabra.org/index.php/ryp/article/view/1156/1134

981 https://www.mediamarkt.es/es/category/_televisi%25C3%25B3n-701130.html

982 https://www.rieb.kobe-u.ac.jp/academic/ra/dp/English/dp246.pdf

983 http://www.radiorebelde.cu/noticia/alerta-investigador-cubano-sobre-peligros-ciberguerra-20110218/

984	http://www.scielo.org.mx/pdf/amdi/v14/v14a27.pdf
985	http://www.diarioveloz.com/notas/193434-rusia-preparado-la-ciberguerra-se-desconectara-internet
986	http://www.gsi.gov.br/
987	https://www.lasprovincias.es/comunitat/201705/22/ciberguerra-solo-empezado-20170522004750.html
988	https://www.icrc.org/es/download/file/14501/irrc-886-casey-maslen.pdf
989	https://rpp.pe/tema-guerra-cibernetica
990	http://razonesdecuba.cubadebate.cu/articulos/el-task-force-y-la-guerra-en-internet-contra-cuba/
991	https://www.cubanet.org/foto/guerra-en-internet-omar-santana/
992	https://www.perfil.com/
993	https://gestion.pe/tecnologia/whatsapp-debera-revisar-comparten-usuarios-india-fake-news-nndc-259416
994	https://m.eldiario.es/cultura/entrevistas/Ciberguerra-malware-cibercrimen-libro_0_528797120.html
995	https://lapupilainsomne.wordpress.com/2011/11/29/de-la-ciberguerra-a-la-ciberdefensa-activa-video/
996	http://www.pcworldenespanol.com/2011/06/07/grandes-preguntas-sobre-la-ciberguerra/
997	https://embassies.gov.il/san-jose/newsAndEvents/Pages/Nuevo-Programa-Ciberguerra.aspx
998	https://www.elmundo.es/elmundo/2010/04/21/internacional/1271823833.html
999	http://alicia.concytec.gob.pe/vufind/Record/ESUP_f28e135a98e501310031645c2495725a
1000	http://www.cubadebate.cu/etiqueta/ciberguerra/page/80/
1001	http://www.cari.org.ar/pdf/ciberdefensa_riesgos_amenazas.pdf
1002	http://www.thiber.org/wp-content/uploads/2019/01/Numero_01_Enero_Comentario.pdf
1003	https://es.wikipedia.org/wiki/Guerra_en_red
1004	https://dialnet.unirioja.es/servlet/articulo%3Fcodigo%3D4899931
1005	http://www.milenio.com/opinion/luis-eugenio-todd/ciencia-politica/la-guerra-cibernetica
1006	http://www.radiorebelde.cu/comentarios/mas-ciberguerra-20120116/
1007	https://moscovita.org/mosconews/en-defensa-del-arte-contemporaneo-avelina-lesper/

1008	https://www.20minutos.es/noticia/327899/0/socialistas/contra/canon/
1009	http://www.libreriafleming.com/libreria/product_info.php%3Fproducts_id%3D108765
1010	https://www.elmundo.es/internacional/2016/02/27/56d0d89622601d7b1e8b461a.html
1011	http://www.congreso.gob.pe/Docs/OCI/files/conferencia_congreso.pdf
1012	https://www.calameo.com/books/00399236526e1f3657318
1013	https://www.amazon.com.mx/Guerra-Cibern%25C3%25A9tica-Blu-ray-Personajes-Animados/dp/B06XYGKPJ1
1014	http://www.gandhi.com.mx/ofertas/guerra-cibernetica-1%3Flanding_source%3Ddigital
1015	http://www.cubadebate.cu/especiales/2017/03/23/wikileaks-y-las-guerras-de-cuarta-generacion/
1016	https://www.larazon.es/ciberguerra-preparados-para-combatir-en-la-r-1-CF2070646
1017	http://revistas.usta.edu.co/index.php/hallazgos/article/view/1571
1018	http://www.diariovasco.com/agencias/201511/18/fundeu-bbva-ciberguerra-espacio-537457.html
1019	http://www.techweek.es/servicios-web/informes/1007877004901/ciberactivismo-ciberguerra-debate.1.html
1020	http://cadenaser.com/tag/guerra_digital/a/
1021	http://www.revista.unam.mx/vol.9/num9/art63/int63.htm
1022	https://es.slideshare.net/mireyagonzalezhernandez/investigacion-final-ciberguerra
1023	http://revista.ieee.es/index.php/ieee/article/view/309
1024	http://www.irenees.net/bdf_fiche-documentation-145_es.html
1025	http://www.revistasculturales.com/xrevistas/PDF/72/1874.pdf
1026	http://treehoze.com/de/InfobaeTecno/1075331177282600960
1027	http://www.armada.mde.es/archivo/rgm/2013/01/cap09.pdf
1028	http://www.gees.org/articulos/que-nuevo-orden-reflexiones-sobre-el-mundo-al-que-vamos
1029	https://www.elnuevodiario.com.ni/opinion/433287-guerra-redes-sociales/
1030	http://www.telemadrid.es/noticias/internacional/Ciberguerra-2-1570062991--20140501033209.html
1031	http://martianos.ning.com/profiles/blogs/la-guerra-en-internet-contra-cuba-2-articulos

1032	http://www.saber.ula.ve/bitstream/123456789/36770/1/articulo9.pdf
1033	https://cybermambi.wordpress.com/tag/peligros-de-la-ciberguerra/
1034	http://www.abc.es/20120213/tecnologia/abci-ciberguerra-empieza-201202130459.html
1035	https://revistas.upb.edu.co/index.php/derecho/article/viewFile/668/612
1036	http://www.vsantivirus.com/mm-ciberguerra.htm
1037	https://es.euronews.com/2018/04/17/rusia-declara-la-guerra-a-telegram
1038	https://www.uexternado.edu.co/wp-content/uploads/2018/10/JUANPSALAZAR.pdf
1039	https://www.scribd.com/document/225554649/Castells-2012-Redes-de-Indignacion-y-Esperanza
1040	https://galicine.es/cine.php%3Fid%3D8
1041	http://www.elladodelmal.com/
1042	http://www.radiocubana.cu/la-opinion/28-desde-la-radio/21584-la-guerra-secreta-en-internet
1043	https://www.laprensa.hn/mundo/1009466-410/eua-y-rusia-al-borde-de-una-guerra-cibern%25C3%25A9tica
1044	http://www.noticiasrcn.com/tags/guerra-cibernetica
1045	http://geo1.espe.edu.ec/wp-content/uploads/2016/07/art15.pdf
1046	https://unesdoc.unesco.org/ark:/48223/pf0000259393_spa
1047	https://pt.scribd.com/document/349719710/Guia-Para-Ver-y-Analizar-Matrix
1048	https://www.stitcher.com/s/%3Feid%3D57756267
1049	https://yu.los40.com/
1050	https://moscovita.org/mosconews/la-mentalidad-de-la-manada-o-la-cargada-curatorial-avelina-lesper/
1051	https://icono14.net/ojs/index.php/icono14/article/view/783
1052	http://www.uniendometas.org.ar/wp-content/uploads/CS-Ciberterrorismo.pdf
1053	http://www.ara.mil.ar/archivos/Docs/015(6).pdf
1054	https://elpais.com/tecnologia/2004/08/02/actualidad/1091435285_850215.html
1055	http://www.heraldo.es/multimedia/imagenes/la-gran-guerra-digital-continua/
1056	https://www.nytimes.com/es/2017/08/11/la-nueva-guerra-fria/
1057	https://listindiario.com/
1058	https://redaccion.conclusion.com.ar/

1059	https://www.amazon.fr/Guerra-Inexistente-Ciberguerra-Ciberdefensa/dp/3659067881
1060	https://laicismo.org/los-judios-ultraortodoxos-declaran-la-guerra-a-internet/
1061	https://odisea.es/programas/cyberwar-t1/episodios/
1062	https://comum.rcaap.pt/bitstream/10400.26/8059/1/Microsoft%2520Word%2520-%2520TIA.pdf
1063	https://dialnet.unirioja.es/servlet/articulo%3Fcodigo%3D3439447
1064	https://dialnet.unirioja.es/descarga/articulo/4331298.pdf
1065	https://www.amazon.com/guerra-Internet-batalla-controlar-Spanish/dp/1502447371
1066	https://www.amazon.es/CIBERGUERRA-Yolanda-Quintana-Serrano/dp/8490971269
1067	https://www.lne.es/sociedad/
1068	http://www.defensa.gob.es/Galerias/documentacion/revistas/2013/red-294-ciberguerra.pdf
1069	https://www.elmundo.es/navegante/99/agosto/19/yakarta.html
1070	http://www.eldiariodetucuman.com/web/ciberguerra-como-prepararse-para-los-hackers-del-2019/
1071	https://www.ambito.com/el-choque-huawei-oculta-una-guerra-el-control-internet-n5004823
1072	http://capitalradio.es/estalla-la-guerra-de-internet-en-eeuu/
1073	https://www.msn.com/es-mx/noticias
1074	https://www.obs-edu.com/int/blog-investigacion/sistemas/que-es-ciberseguridad-y-de-que-fases-consta
1075	https://es.wikipedia.org/wiki/Cibercomando_de_Estados_Unidos
1076	https://www.amazon.es/Inexistente-Ciberguerra-Llongueras-Vicente-Adrianna/dp/3659067881
1077	https://www.pandasecurity.com/spain/mediacenter/noticias/las-mejores-peliculas-sobre-ciberseguridad/
1078	http://www.internetprofunda.com.ar/wp-content/uploads/2016/05/Internet-Profunda.pdf
1079	http://www.ventanapolitica.cu/articulo/la-ciberguerra
1080	http://www.ub.edu/geocrit/arac-47.htm
1081	https://centronaval.org.ar/boletin/BCN817/817lopez.pdf
1082	http://www.eldesconcierto.cl/new/tag/ciberguerra/

1083	https://nae.global/la-seguridad-digital-amenazas-y-soluciones/
1084	http://www.nos-comunicamos.com.ar/content/ciber-guerra-entre-estados-unidos-y-china
1085	http://es.catholic.net/op/articulos/9715/guerra-en-la-red.html
1086	http://www.academia.edu/32164878/EL_CIBERESPACIO_Y_LA_SEGURIDAD_NACIONAL_EN_EL
1087	https://www.amazon.com/Summer-Wars-Cibernetica-Japanese-Language/dp/B06XYGKPJ1
1088	http://www.centredelas.org/es/seguridad-y-defensa/869-ciberseguridad-y-ciberguerra
1089	https://www.ivoox.com/espana-reserva-ciberguerrera-occidente-episodio-audios-mp3_rf_18703671_1.html
1090	https://as.com/betech/2018/07/05/portada/1530814852_906273.html
1091	http://www.ceeag.cl/wp-content/uploads/2018/07/LA-CIBERGUERRA-SUS-IMPACTOS-Y-DESAFIOS.pdf
1092	https://www.icrc.org/spa/resources/documents/misc/5tedfy.htm
1093	http://www.abc.es/20120213/sociedad/abcp-primera-ciberguerra-esta-marcha-20120213.html
1094	https://www.amazon.com.mx/Guerra-Cibern%25C3%25A9tica-Varios-Personajes-Animados/dp/B06XY7L3MW
1095	https://www.labrujulaverde.com/2009/01/computadoras-por-persona
1096	http://misionverdad.com/trama-global/zero-days-o-los-limites-de-la-ciberguerra
1097	http://www.unq.edu.ar/advf/documentos/593955b92ae2c.pdf
1098	https://www.lmneuquen.com/google-y-facebook-los-gigantes-internet-guerra-europa-n623454
1099	http://internetng.dit.upm.es/stuxnet-nuevo-virus-de-sabotaje-industrial-y-ciberguerra/
1100	http://esdegue.edu.co/node/224
1101	https://dialnet.unirioja.es/servlet/articulo%3Fcodigo%3D4286731
1102	https://www.librerianacional.com/pagina%3Dproducto%26libro%3D343167
1103	https://gestion.pe/tecnologia/ciberguerra-amenaza-trascender-ciencia-ficcion-78379
1104	https://www.elespanol.com/bluper/noticias/antena3-gana-guerra-televisiones-internet-usuarios-unicos
1105	http://www.redalyc.org/articulo.oa%3Fid%3D93521629014
1106	https://www.larazon.es/opinion/columnistas/la-ciberguerra-total-KG8282581

1107	https://reedes.org/seminario-sobre-guerra-y-salud/
1108	http://campoderelampagos.org/critica-y-reviews/16/6/2018