Detección de phishing en etapa de detección temprana utilizando características relacionadas a la marca afectada

Daniel Alejandro Barreiro Herrera

Universidad Nacional de Colombia
Facultad de Ingenieria, Departamento de Ingeniería de Sistemas e Industrial
Maestría en ingeniería de sistemas y computación
Bogotá, Colombia
2023

Detección de phishing en etapa de detección temprana utilizando características relacionadas a la marca afectada

Daniel Alejandro Barreiro Herrera

Trabajo de grado presentado como requisito parcial para optar al título de:
Magister en Ingeniería de sistemas y computación

Director:
Doctor en Ingeniería de Sistemas y Computación, Jorge Eliecer Camargo Mendoza

Línea de Investigación:
Ciberseguridad
Grupo de Investigación:
Unsecurelab

Universidad Nacional de Colombia
Facultad de Ingenieria, Departamento de Ingeniería de Sistemas e Industrial
Maestría en ingeniería de sistemas y computación
Bogotá, Colombia
2023

Dedicatoria

Dedicado a mi familia, guía y motivación de mis mejores logros.

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

Daniel Alejandro Barreiro Herrera
Nombre

Fecha 17/07/2023

Resumen

Detección de phishing en etapa de detección temprana utilizando características relacionadas a la marca afectada

El phishing es uno de los ataques cibernéticos sufridos por los usuarios de servicios transaccionales a través de Internet, si bien existe investigación enfocada en detectar ataques de phishing y la literatura muestra resultados con alta efectividad en detección, estos estudios no permiten enfatizar en qué etapa de detección se actúa. Teniendo en cuenta la revisión sistemática de literatura realizada previamente en [Barreiro and Camargo, 2022], se presenta una descripción general actualizada de la detección de phishing, en este estudio se identificó que el 83 % de literatura consultada se centró en la fase de mitigación, donde la metodología funciona de manera reactiva utilizando características estáticas que brindan alta precisión pero fallan en el modelo con el tiempo.

Es así como en el presente documento se detallará la implementación de un modelo computacional de detección de phishing basado en la extracción de características de la marca afectada, el cual permita actuar en la etapa de prevención del ataque. Se realiza un análisis exploratorio de datasets de phishing para tres marcas, posteriormente se seleccionan las características de marca y se detallará los detalles de diseño e implementación de los modelos para las tres marcas seleccionadas, probando diferentes modelos de aprendizaje de maquina y analizando el comportamiento de sus características. Finalmente, se analizarán resultados y se presentarán conclusiones para enfatizar la importancia de usar información de marca y mezclar diferentes enfoques para mejorar la detección de etapas tempranas. La contribución de este trabajo se centra en establecer una aproximación diferente que permite construir el modelo adecuado para cada marca, incentivando futuras investigaciones y futuros trabajos relacionados para considerar sus modelos más allá de la alta precisión, y plantear cómo estos pueden proporcionar soluciones eficientes que se pueden integrar en entornos de producción reales para proteger a los usuarios.

Palabras clave: (phishing, detección, marca , etapa temprana, proactividad).

Abstract

Phishing detection in early detection stage using features related to the affected brand

Phishing is one of the cyber attacks suffered by users of transactional services over the Internet, although there is research focused on detecting phishing attacks and the literature shows highly effective results in detection, these studies do not allow emphasize at what stage of detection is acted on. Taking into account the systematic review of literature previously carried out in [Barreiro and Camargo, 2022], an updated general description of phishing detection is presented, in this study, it was identified that 83 % of the selected literature focused on the mitigation phase, where the methodology works reactively using static features that provide high accuracy but fail in the model over time.

This is how this document will detail the implementation of a phishing detection computational model based on the extraction of characteristics of the affected brand and that also allows acting in the attack prevention stage. An exploratory analysis of phishing datasets for three brands is carried out, then the brand characteristics are selected and the details of the design and implementation of the models for the three selected brands will be detailed, testing different machine learning models and analyzing the feature's performance. Finally, results will be analyzed and conclusions will be presented to emphasize the importance of using brand information and mixing different approaches to improve early-stage detection. The contribution of this work is focused on establishing another approach for building the best solution for each brand, encouraging future research and future related work to consider their models beyond high precision, and proposing how these models can provide efficient solutions that can be integrated into production environments to protect the users.

Keywords: phishing, detection, brand, early stage, proactivity

.

Este Trabajo Final de maestría fue calificado en abril de 2023 por el siguiente evaluador :

Jorge Eduardo Ortíz Triviño PhD.
Profesor Facultad de Ingeniería
Universidad Nacional de Colombia

Tabla de Contenido

Lista de Figuras

Lista de Tablas

1. Introducción

En los últimos años, nuestro entorno y la forma en que se interactúa como sociedad ha sufrido transformaciones. Los avances en la tecnología, especialmente en las comunicaciones, son pilar y motor de esta transformación, proporcionando velocidades más altas y coberturas más amplias que han permitido generar desarrollos que han facilitado la gestión de transacciones digitales, brindando una alternativa a lo que antes se hacía en un sitio físico, ahora a solo un clic de distancia. El portafolio de servicios que se ofrece en la web continua creciendo y diversificándose. La pandemia por el COVID-19 ha acelerado la transformación digital de muchas organizaciones, las cuales han sido incentivadas a ofrecer sus servicios a través de una plataforma web, por lo que la seguridad de quienes interactúan con estos sitios es un tema que se debe mantener como prioritario para asegurar la confianza del usuario a hacer uso de estos canales.

Hay dos actores principales en este escenario, de un lado están los que ofrecen servicios buscando ofrecer facilidades a sus usuarios, agilidad en sus procesos y beneficio económico al exponer nuevas formas de atraer a clientes. Y por otro lado los usuarios de estos servicios, buscando ahorro de tiempo y una experiencia de usuario agradable y segura para realizar las operaciones que necesita realizar. Esperando que del lado del prestador del servicio se le ofrezca comodidad , pero también seguridad. Según una investigación del [apwg, 2022] Anti-Phishing Working Group, en el cuarto trimestre de 2021, el phishing como ataque es uno de los ataques más sufridos por los usuarios de la web, mostrando en su reporte la cifra de 316.747 ataques de phishing reportados en diciembre de 2021, siendo el mayor número de ataques reportados en la historia de sus reportes desde 2004, año en que se empezaron a generar. El informe también menciona que la cantidad de ataques a fines de 2021 es tres veces mayor que la cantidad de ataques anteriores a principios de 2020. También analiza las industrias objetivo de estos ciberataques y encontró que el sector financiero fue uno de los más afectados con 23,2 %, seguido de SAAS/Web-mail con 19,5 % y comercio electrónico/servicios minoristas con 17,3 %.

[A, 2020] y [Patil et al., 2018] hicieron el mismo análisis en sus presentaciones hace 2 y 4 años, respectivamente, encontrando que estas estadísticas continuaron creciendo en los años siguientes. Esto sugiere que si bien la investigación sobre la detección de phishing es diversa y no es un tema reciente, se debe mejorar la aplicación de mecanismos que no solo mitiguen sino que eviten que los usuarios caigan en las primeras etapas del phishing.

En [Das et al., 2020] se describe una revisión de la investigación relacionada con el phishing y sus desafíos en detección y la Figura **1-1** muestra los aspectos que este documento desea resaltar.

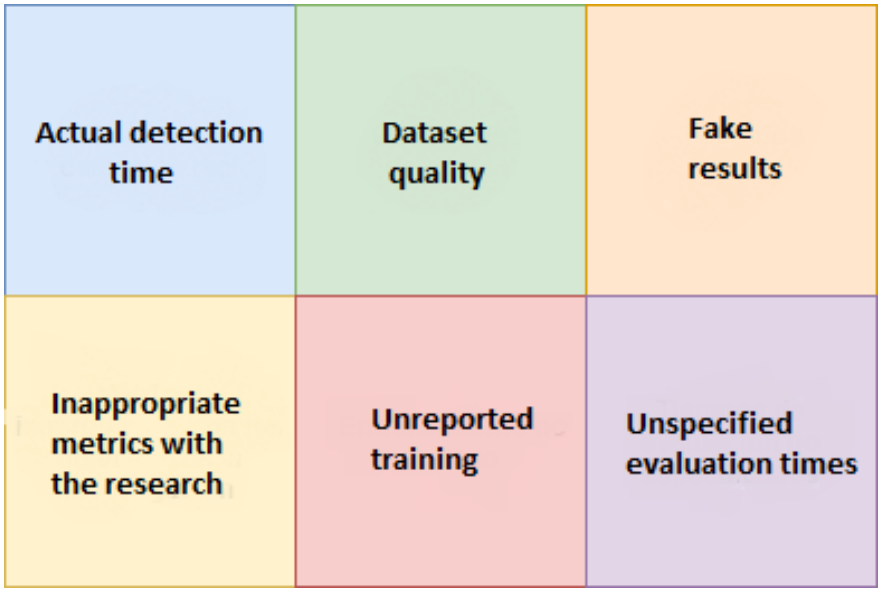


Figura 1-1.: Desafíos en detección de phishing

La investigación de detección de phishing no es particularmente nueva, y apuntar a detectar tantas URL maliciosas como sea posible parece un buen objetivo, buscando eliminar el trabajo manual de los SOC de seguridad y clasificando URL potencialmente peligrosas en lugar de falsos positivos. El propósito de este trabajo está lejos de señalar que este objetivo es malo o innecesario, por lo demás, se considera importante, y propone una metodología que ayuda a detectar phishing en etapas tempranas utilizando una combinación de modelos que pueden ser distintos según cada marca, es crucial que los componentes de estos modelos puedan actuar durante la fase de prevención en el proceso y en su conjunto.

Otro aspecto a considerar es el cambio en las características del phishing a lo largo del tiempo [Das et al., 2020]. El phishing en 2022 no es como el phishing en 2018, en solo 4 años, los atacantes han encontrado formas de robar información del usuario a través de formularios, sitios hackeados, servicios de alojamiento gratuitos y tunelización de sitios locales. Por esta razón, es necesario considerar características que no solo proporcionen la alta precisión de las detecciones actuales, sino que también habiliten mecanismos que puedan adaptarse a diferentes técnicas y que puedan identificar elementos clave el phishing para permitir su detección temprana y mantener su rendimiento en un rango de tiempo aceptable.

Contribuciones

Este trabajo explora otros enfoques distintos a los encontrados en el estado del arte en detección de phishing, identificando los puntos clave donde la investigación puede proporcionar soluciones efectivas e integrables en entornos reales. Los hallazgos permiten reflejar las características identificadas y ajustar las recomendaciones del modelo para la identificación de phishing en etapas tempranas de detección teniendo en cuenta características relacionadas a la marca.

En el transcurso de esta investigación se realizaron contribuciones como lo es un artículo de revisión de literatura en el que se expuso la problemática y la necesidad de realizar el estudio de marca en detección de phishing, para el cual se realizó una ponencia. Adicionalmente se participó en la tercera jornada de ciberseguridad de la universidad Nacional, donde el artículo fue aceptado y se realizó una ponencia con poster durante la jornada y finalmente se realizó un artículo de resultados a presentarse en el journal de inteligencia artificial de Iberamia. A continuación se exponen las contribuciones:

- **Barreiro, D. A. and Camargo, J. E. (2022)**. A systematic review on phishing detection: A perspective beyond a high accuracy in phishing detection. pages 173–188 fue publicado en Communications in Computer and Information Science book series (CCIS,volume 1643) y presentado en 5th International Conference on Applied Informatics en Arequipa , Perú.

- **Barreiro, D. A. and Camargo, J. E. (2022)**. Detección de phishing en etapas tempranas utilizando características de marca. Poster presentado en 3ra Jornada de Ciberseguridad Universidad Nacional JCUN2022.

- **Barreiro, D. A. and Camargo, J. E. (2023)**. Phishing detection in early stage based on brand features. Inteligencia Artificial - la Sociedad Iberoamericana de Inteligencia Artificial (IBERAMIA) en preparación.

2. Marco de referencia

En este capitulo se exploran algunos conceptos, así como mecanismos que permiten abordar la problemática en detección de phishing de manera adecuada. Se iniciará con definición del phishing así como diferentes formas de presentarse a los usuarios. Se continuará con explicar la detección de phishing y sus diferentes mecanismos utilizados en detección, haciendo especial énfasis en identificar etapas de detección de phishing donde se considera está la principal contribución de este capítulo.

2.1. Phishing

Se denomina phishing a un conjunto de técnicas utilizadas en el campo de la ingeniería social que como [Baig et al., 2021] comenta busca persuadir a las personas a entregar información sensible o lograr acceder a su sistema para obtener datos sensibles, cabe mencionar que aunque en la bibliografía se relaciona mucho el phishing a correos electrónicos, esta no es la única manera de dispersar URLs maliciosas, redes sociales, mensajes de texto, suplantación de identidad y anuncios son algunas de los mecanismos mayormente conocidos objeto de estudio en la literatura. Sea cual sea el mecanismo de dispersión siempre se presenta un patrón en el flujo de acción de un ataque de phishing.

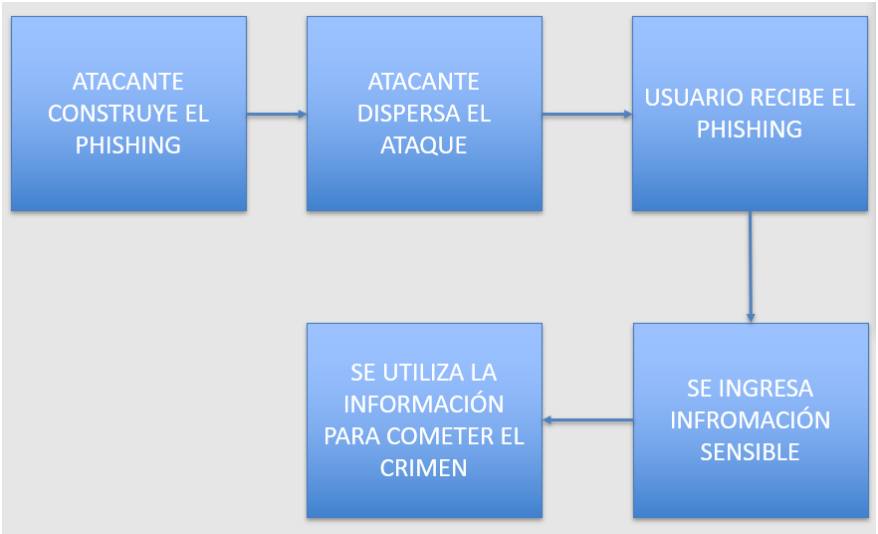


Figura 2-1.: Descripción de operación de un ataque de phishing

El phishing no es un concepto especialmente nuevo y se encuentra en constante evolución, según [JAMES, 2005] datan registros desde 1995 con ataques de phishing para AOL, también se registra el primer reporte para entidades financieras en Julio de 2003, desde entonces se han venido implementando diferentes tipos de protocolos, campañas, monitoreo y mecanismos que buscan identificar y bloquear posibles ataques de phishing. Llegando incluso a crear comunidades y entidades que luchan contra el fraude como es el caso de APWG [apwg, 2022], consorcio que concentra esfuerzos en combatir el fraude electrónico haciendo seguimiento constante del las tendencias y comportamientos del phishing. Donde cada trimestre están publicando reportes del comportamiento de los ataques de phishing donde si se compara lo que era el phishing en el 2005 con [JAMES, 2005] se observa como el crecimiento tecnológico y la expansión de servicios a plataformas web ha vuelto complejo el panorama en cuanto a analizar phishing se refiere, tomando como referencia [apwg, 2022] en su ultimo reporte de segundo trimestre de 2022 muestra como mes a mes se incrementan el numero de ataques únicos de phishing de 362,852 en abril a 381,717 en junio. Este mismo panorama se repite en cada reporte de cada año disponible para consulta al publico en su pagina web (apwg.org). En este reporte también se muestra un resumen de los sectores mayormente afectados dentro de los que se encuentran el sector financiero, SaaS y las redes sociales reportando entre estos 3 sectores el 62 % de los reportes de phishing.

2.2. Detección de phishing

2.2.1. Etapas de detección

Para comprender cuándo entra en juego la investigación sobre ataques de phishing, es importante identificar los diferentes tipos de enfoques en la literatura que aborda el problema, y analizar y presentar sus resultados en consecuencia.
Al analizar diferentes estudios que funcionan en etapas de detección completamente diferentes, estudios como [Ya et al., 2019] se enfocan en validar el phishing de las URL implementadas, un medio de ataque que ya existe. Es posible que algunas personas ya hayan recibido la misma URL, mientras que otras han recibido instrucciones para evitar caer en ella. Por otro lado, investigaciones como [Nakamura and Dobashi, 2019] se enfocan en crear algoritmos de generación de dominios que puedan funcionar en tiempo cero, lo que significa que el algoritmo puede identificar posibles ataques de phishing incluso antes de que se implemente este ataque.

Aunque todos están diseñados para detectar phishing, varían ampliamente en el método, la etapa de detección, la fuente y las técnicas utilizadas. Teniendo en cuenta que los resultados [Nakamura and Dobashi, 2019] tienen una precisión de menos del 5 % y [Ya et al., 2019] tienen una precisión de más del 90 %, pero en [Nakamura and Dobashi, 2019] nadie tiene que caer en el fraude, mientras que la precisión de [Ya et al., 2019] puede cubrir un una

gama más amplia de marcas sin embargo no pueden cuantificarse el numero de personas cuya información se vio expuesta antes de que el phishing se detecte.
Con eso en mente, se propone la siguiente clasificación en etapas de detección de phishing:

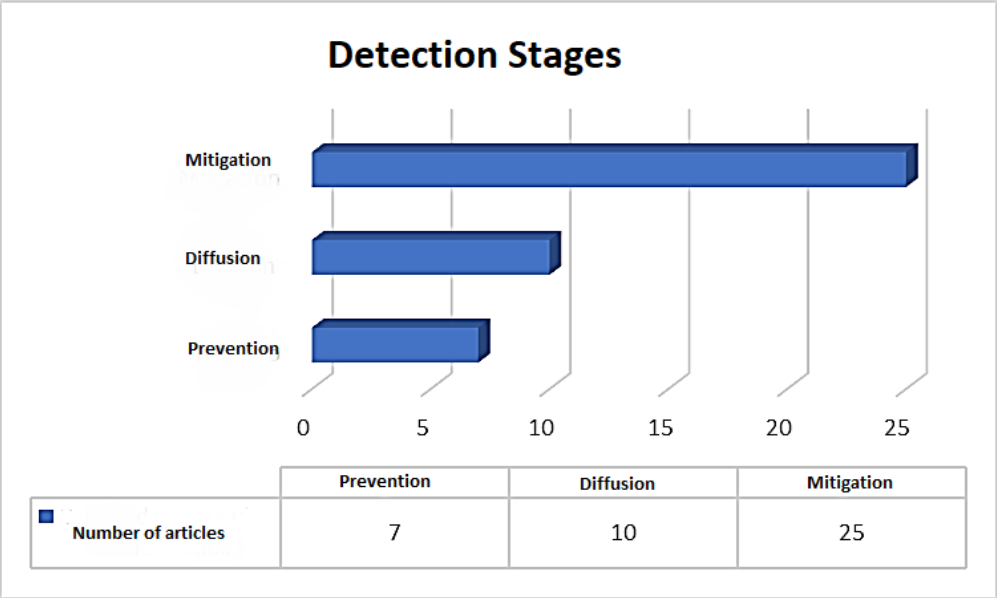


Figura 2-2.: Descripción de la etapa de detección en la que se encuentran los trabajos revisados

2.2.2. Etapa de prevención

Esta etapa actúa sobre el dominios recientemente registrados e incluso dominios generados a partir de palabras clave, como [Nakamura and Dobashi, 2019], [Buber et al., 2017], [Adil et al., 2020], [Spaulding et al., 2016] , [Starov et al., 2019], [Ginsberg and Yu, 2018] y [Li et al., 2016] Según las características extraídas, esta etapa es perfecta para las aplicaciones que realmente evitan que el phishing se propague antes de que llegue a los usuarios.

2.2.3. Etapa de difusión

Esta etapa incluye [Ya et al., 2019], [Li and Wang, 2017], [Li et al., 2020], [Eshmawi and Nair, 2019], [Balim and Gunal, 2019], [Dalgic et al., 2018], [Yan et al., 2020], [Sahoo, 2018], [Baykara and Gürel, 2018], [Lingam et al., 2018] y Lingam 2019 [Lingam et al., 2019] está relacionado con el mecanismo por el cual el phishing llega a los usuarios finales; así es como se presentan dispersados en las redes sociales o en los correos electrónicos.

2.2.4. Etapa de mitigación

En esta etapa actúan el 83 % de los artículos estudiados, cuyos modelos se basan en bases de datos comunitarias o URL informadas, es decir actúan sobre URLs ya desplegadas o reportadas por usuarios. Actualmente, 35 artículos estudiados mostrados en **2-2** se centran en las dos últimas fases, tratando de identificar y estudiar cómo se propaga el phishing, o analizar las URL finales donde los usuarios han caido.

En el resto del escrito, para efectos prácticos, se referirá a estas 3 fases anteriores mencionadas como prevención, transmisión y mitigación, como se muestra en la figura **2-2**

2.2.5. Métodos de detección

Se considera necesario comenzar identificando los métodos utilizados actualmente de detección de phishing que ya proporcionan un panorama general de lo que se está implementando y de las posibilidades técnicas que se pueden utilizar en la detección de Phishing. En **2-3** se observa un compilado de los métodos encontrados en la literatura. Sin embargo aunque se encontraron múltiples métodos de detección se pueden agrupar en 3 grandes grupos excluyendo el conocimiento del usuario para identificar el phishing, estos son: Basados en listas, Heurísticos y métodos basados en machine Learning. Estos tres grupos pueden actuar en conjunto para crear modelos híbridos. La distribución de uso en la literatura de los 3 grandes grupos se observa en **2-4** donde se muestra una clara tendencia en aplicar modelos de machine Learning a este problema de detección con un 82 %.

2.2.5.1. Validación por listas

Al igual que en Buber 2017 [Buber et al., 2017], Mondal 2019 [Mondal et al., 2019] y Patil 2018 [Patil et al., 2018] se consideran características importantes a tener en cuenta en sistemas más estructurados, pueden ayudar a mitigar los ataques porque si en presencia de un sistema interconectado sistema de navegación de información, el hecho de que este modelo sea el más relevante tiene un potencial de mitigación muy efectivo.

Este tipo de tácticas siguen siendo útiles porque permiten actuar donde otros sistemas de verificación fallen, aunque en menor número, pueden actuar en una etapa temprana en función de precedentes notorios, como reactivaciones de ataques anteriores o identificación de IP maliciosas.

2.2.5.2. Heurísticos

La heurística depende de la calidad de las características extraídas de ciertos artículos, por ejemplo, [Ali and Ahmed, 2019], [Huang et al., 2019], [Nakamura and Dobashi, 2019],

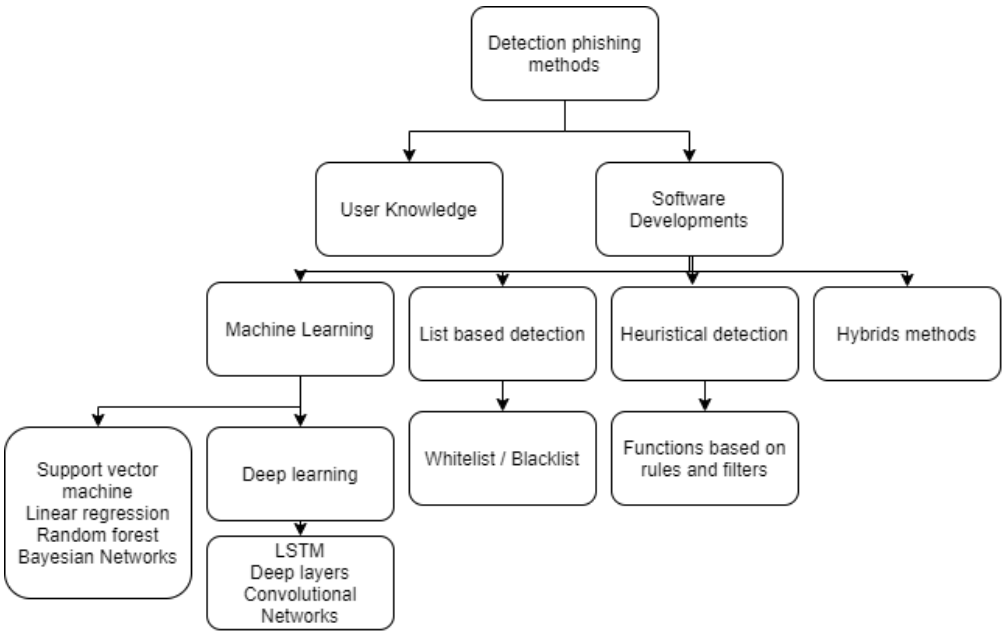


Figura 2-3.: Métodos utilizados en la literatura.

[Nathezhtha et al., 2019], [**?**]. En estos trabajos, los resultados permiten visualizar el comportamiento esperado, es decir, en el caso de estas implementaciones, es necesario saber qué se está buscando, y basado en esto construir algoritmos que permitan identificar estas características esperadas. Este tipo de implementación puede funcionar en cualquiera de las tres etapas, dependiendo de cómo esté diseñado en [Nakamura and Dobashi, 2019] , que se utiliza para la detección temprana. Pero también se puede usar en la fase de mitigación, ya que se puede basar en las capacidades de las URL de phishing implementadas. En este tipo de enfoque, es importante tener información sobre palabras clave y patrones que son efectivos durante la fase de prevención.

2.2.5.3. Machine Learning

Otro mecanismo es el uso de herramientas que monitorean las URL y tratan de dar riesgo en función de las características detectadas en una página web en particular. Resulta que las herramientas que utilizan el aprendizaje automático han logrado los mejores resultados en los últimos años [Anand et al., 2018]. Con base en este riesgo, se pueden tomar decisiones para mitigar el impacto del fraude [Megha et al., 2019]. Si bien los algoritmos de detección de aprendizaje automático brindan los mejores resultados, todavía existe una brecha en la distinción entre detección y verificación de phishing. La diferencia entre estos dos conceptos radica en la fase de implementación de pruebas de estos modelos. De los 25 modelos estudiados, las pruebas se realizaron sobre URLs desplegadas, por lo que no fueron detecciones, sino sistemas de verificación de phishing que funcionaron durante la fase de mitigación. Existe una falta general de resultados de evaluación, implementados en entornos del mundo real y

durante períodos prolongados de tiempo, para evaluar algoritmos de aprendizaje automático en relación con la evolución del phishing y los nuevos métodos que los atacantes usan para implementarlos.

2.2.6. Fuentes de detección

En estudios como [Li and Wang, 2017], [Sharma et al., 2017] y [Pande and Voditel, 2017], se han encontrado diferentes fuentes, ya sea para estudiar las firmas de phishing por sí solas o para verificar resultados como [Adil et al., 2020] y [Li et al., 2016]. Para hacer esto, se deben contar las fuentes de información que vinculan a sitios de phishing o que al menos permiten la extracción de URL relacionadas con funciones de phishing. Aquí hay algunas fuentes que se consideran útiles en diferentes etapas de detección:

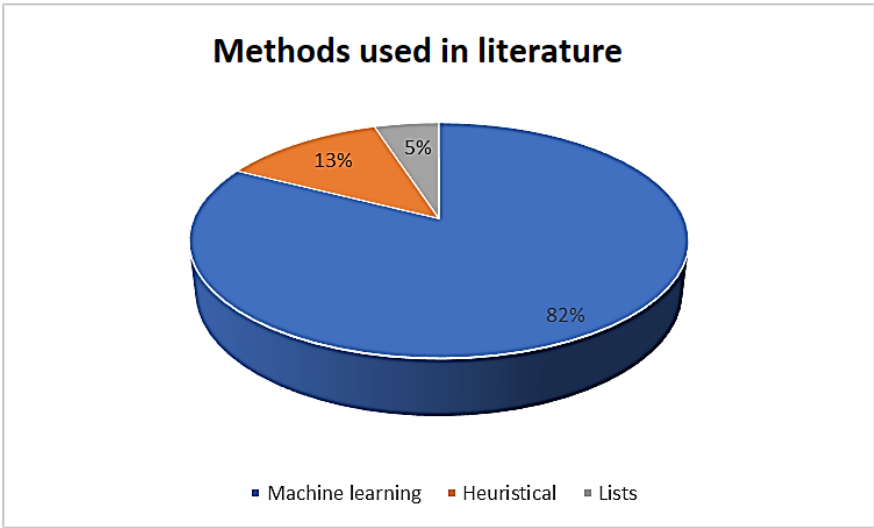


Figura 2-4.: Distribución de métodos usados en la literatura

Nombre de la fuente	Descripción	Referencia
Phishtank	PhishTank es una grupo colaborativa para obtener datos e información sobre phishing en Internet.	[Li et al., 2016] [Sharma et al., 2017] [Pande and Voditel, 2017]
APWG / ecx APWG	Diferentes tipos de grupos de trabajo antiphishing herramientas enfocadas a la detección y centralización de información sobre phishing	[Das et al., 2020]
Openphish	OpenPhish proporciona inteligencia sobre ciberamenazas y servicios	

Tabla 2-1.: Algunas fuentes usadas en la búsqueda de phishing

La tabla **2-1** muestra algunas de las fuentes encontradas en la web para sitios informados

por comunidades o equipos profesionales, a menudo los investigadores pueden reunir sus conjuntos de datos y analizar las características de los ataques de phishing de forma preliminar. Sin embargo, vale la pena mencionar que este tipo de fuentes brindan diferentes tipos de utilidades para diferentes usuarios involucrados en la detección. Pero si bien esta es una excelente manera de recopilar información sobre el phishing activo, el uso de estos datos para detectar el phishing solo ayudará durante la fase de mitigación.

Por lo tanto, es posible resaltar en qué etapa ciertas investigaciones de detección se basan en la elección de la fuente de datos. Las investigaciones que quieran tomar medidas durante la fase de difusión buscarán fuentes relacionadas con las redes sociales, el correo electrónico o la publicidad online, mientras que la fase de mitigación utilizará fuentes relevantes a la tabla **2-1**, lo que permitiría automatizar el proceso de categorización, lo que facilitaría un proceso más sistemático para categorizar URLs con más elementos; aunque en una etapa temprana, sería mejor que el propio registrador de dominios actuara, momento en el que comienza a dar pistas de que el dominio está enfocado a simular otra página web.

Nombre de la fuente	Descripción	Ref
Domainwatch	Herramienta útil para buscar información de un dominio a partir de palabras clave. Proporciona histórico e información actual.	[DomainWatch,]
Urlscan	urlscan.io es un servicio de escaneo análisis de sitios web usando palabras clave.	[urlscan,]

Tabla 2-2.: Fuentes que pueden se usadas en etapas de detección tempranas

La tabla **2-2** muestra algunas fuentes enfocadas en la detección de dominios desde la fase de prevención. Son fuentes basadas en palabras clave que pueden permitir búsquedas de dominios registrados recientemente, así como mostrar whois, certificados de seguridad asociados y otros dominios.

2.2.7. Características utilizadas en detección de phishing

Dentro los estudios de carcateristicas encontrados se destacan [Aung and Yamana, 2019], [Eshmawi and Nair, 2019], [McGahagan et al., 2019], [Yazhmozhi and Janet, 2019] y [Yuan et al., 2018] como alarma de phishing presentan una gran variedad de opciones pasando por características asociadas solo a la URL, como consideraciones adicionales asociadas al registro(WHOIS) o contenido(DOM), se consideran podrían actuar en distintas etapas, debido a que algunas dependen en gran medida de que la URL ya este desplegada con el ataque de phishing , mientras que otras como en los extraídos en métodos heurísticos, saben lo que están buscando específicamente y esto les podría permitir identificar estas características en etapas

tempranas, en la tabla **2-3** se muestran algunas características extraídas en los artículos y su dependencia para ilustrar mejor el ejemplo.

Característica	**Dependencia**
Dominios embebidos dentro de la propia URL	URL ya desplegada, ataque en curso.
No tiene dominio, actúa sobre sobre una IP	Puede extraerse desde su creación y también mediante patrones en IP.
Cantidad de puntos en la URL	URL ya desplegada, ataque en curso.
Numero de palabras en blacklist dentro de la URL	Se debe tener un estudio de las palabras que se buscan
Posición de (TLD) fuera de lugar	URL ya desplegada, ataque en curso.
Formularios dentro del DOM HTML	Ataque desplegado sobre la URL
Los demás enlaces de dentro del DOMHTML tienen distintos dominios	URL ya desplegada, ataque en curso.
La marca a la cual se ataca no se encuentra en el dominio sino en las carpetas o subdominios	Se debe tener información de palabras clave de la marca
Tiempo de registro del dominio (entre más reciente es más probable que sea un phishing)	Se puede conocer desde la creación del dominio
PageRank – la página no está indexada por los buscadores	Característica general de los ataques de phishing debido a su corta duración

Tabla 2-3.: Características comunes extraídas en las investigaciones

las características ejemplificadas en la tabla **2-3** tienen diferentes mecanismos de extracción y diferentes requerimientos obtener su cuantificación. Es por esto que se pueden utilizar en diferentes etapas y utilizar dependiendo de las condiciones del dataset.

2.2.8. El desafío de la selección de características

En estudios como [Zhu et al., 2018] y [Yang et al., 2019], los autores buscan diversificar las características utilizadas para buscar una mayor precisión de detección. Sin embargo, el phishing caracterizado en la actualidad es susceptible a cambios en futuros ataques. Entre las características mencionadas en la literatura, como [Aung and Yamana, 2019], [Eshmawi and Nair, 2019], [McGahagan et al., 2019] y [Yuan et al., 2018] como alertas de phishing, existen múltiples opciones y las condiciones del dataset y . Se piensa que actúan en diferentes etapas, ya que algunas dependen de si la URL se ha desplegado con un ataque de phishing, mientras que otras, como las extraídas en heurística, saben exactamente lo que están buscando, lo que les permite identificar estas características. En los primeros días, estas características tenían diferentes mecanismos de extracción y diferentes requisitos para cuantificarlas.

Idealmente, un sistema completo debería contener la extracción de características para todas las etapas posibles, aunque la identificación se realiza idealmente en la etapa de prevención, por lo que la primera etapa idealmente tendría la mayor cantidad posible de características. Sin embargo, las características de este tipo deben buscarse en las fuentes que se muestran en la tabla **2-2**. Es difícil buscar en estas fuentes si no está seguro de lo que está buscando, las características relacionadas con la marca en este punto ayudan a buscar a través de la gran cantidad de dominios registrados por segundo y pueden ayudar a verificar el phishing antes de que se difunda.

2.3. Desafíos en detección de phishing

Un aspecto importante a considerar son los patrones encontrados en la mayoría de los sistemas de reconocimiento estudiados. Es importante recordar que existen estudios [Das et al., 2020] que han identificado diferentes desafíos en cada uno de estos procesos, y los casos para este trabajo se consideran menciones apropiadas.

1. Fase de Extracción de Fuente: Esta fase presenta el desafío de obtener información que no esté sesgada hacia un determinado conjunto de amenazas, así como obtener información en tiempo real, en general cerrando las brechas que limitan la calidad de los datos y el desarrollo de la investigación.

2. Análisis de datos y extracción de datos relacionados: Aún quedan muchos desafíos por explorar en esta etapa. [Das et al., 2020] plantea dos problemas principales relacionados con la calidad de los datos de los que se extraen estas características y la escala de tiempo desde la campaña de ataque.

3. Entrenamiento y/o puesta a punto del sistema: En esta etapa, se presenta el desafío de configurar las características apropiadas para que mientras el sistema está aprendiendo, no solo pueda adaptarse a los datos con los que fue entrenado, sino también ayudar a futuros ataques.

4. Evaluación del sistema: El desafío es que se deben buscar parámetros de evaluación adecuados, que dependan no solo de la correcta interpretación de los resultados, sino también de la calidad de las fases pasadas, para brindar información que haga más que "verificar la efectividad de suplantación de identidad".

2.3.1. Incluir información de la marca para mejorar el rendimiento del phishing

Es necesario identificar la gran mayoría de las características asociadas con el phishing en una etapa temprana y considerar factores como los cambios en estas características a lo largo del tiempo, los cambios en la tecnología y las credenciales de seguridad. Deben tenerse en cuenta al analizar la selección de características para cambiar la comprensión global del problema para proteger marcas específicas y, por lo tanto, proteger a los usuarios antes de que los ataques se generalicen. En este sentido, la industria debe asumir el rol de proteger los servicios que brinda a los usuarios e implementar sistemas personalizados que aborden el problema desde las características de su amenaza de fraude. Como tal, se ha descubierto que algunos investigadores utilizan un enfoque diferente, que se puede utilizar bajo el concepto de características de marca, como [Zuraiq and Alkasassbeh, 2019], [Ginsberg and Yu, 2018] y [Concone et al., 2019] . Se descubrió que el trabajo relacionado es un ejemplo de phishing de correo electrónico muy similar, que demuestra que las redes neuronales recurrentes superan el 98 % [Ya et al., 2019]. La figura **2-5** muestra las características que se pueden extraer en cada etapa.

Además, [Yazhmozhi and Janet, 2019] muestra cómo ejecutar modelos implementables en tiempo real a partir de la extracción de características basada en NLP-W2V. Sin embargo, no se hace ninguna conexión con la etapa de prevención de phishing, por lo que no se puede determinar si funcionará en la etapa anterior, pero muestra cómo se puede utilizar para extraer posibles características de marca y sus representaciones vectoriales.
Otro método [Yao et al., 2018] proporciona resultados utilizando las características extraídas del logotipo, mostrando una precisión del 97 %. Por lo tanto, la verificación de imágenes debe considerarse desde una etapa temprana.

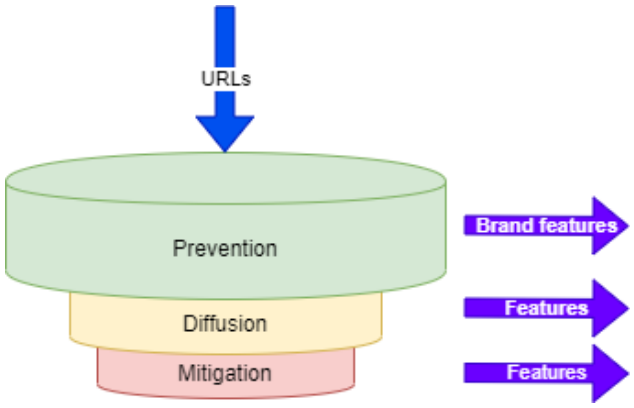


Figura 2-5.: Diagrama de extracción de características y su importancia en las diferentes etapas

2.3.2. Cómo usar patrones y características de marca en la fase de prevención

Según la Figura **2-2**, la mejora más valiosa que pueden desarrollar las nuevas investigaciones es la capacidad de aumentar la detección de la fase preventiva. Además, estos deben mejorar la precisión de esta etapa, la mayoría de los métodos utilizados en esta etapa están relacionados con los métodos heurísticos y de lista, pero menos con los métodos de aprendizaje automático. Los métodos de aprendizaje automático necesitan algunas características relevantes para enseñar al modelo a aprender las características de un ataque de phishing. Pero, ¿qué tipo de características puede obtener un modelo de los dominios registrados recientemente? El enfoque actual puede no tener la respuesta. El dominio más cercano tiene solo una serie de caracteres asociados y puede proporcionar información a través de una solicitud de WHOIS que proporcione información de registro del dominio.

Dados los patrones y el reconocimiento de las marcas, aquí podría surgir un nuevo enfoque y, si bien esto puede sonar como el comienzo de un enfoque heurístico, los modelos de aprendizaje automático que pueden aprender patrones y características asociados con una marca pueden ser una poderosa herramienta de reconocimiento para marcas específicas. Suplantación de identidad. Por tanto, la importancia de estudiar estas características es crucial para desarrollar sistemas de detección de phishing en la fase de prevención. A continuación, se describen las características globales a considerar para esta propuesta:

2.3.2.1. Caracterización del atacante

Los dominios registrados recientemente pueden detectar muchas características relacionadas con los registros de WHOIS, como los datos del registrante, el registrador, el alojamiento, los servicios de país y eventos como posibles registros MX, así como la certificación SSL y sus respectivas organizaciones.

2.3.2.2. Análisis de caracteres en la URL

Por lo general, el dominio puede contener el nombre de la marca afectada o caracteres similares [Nakamura and Dobashi, 2019] que se pueden procesar solo con el dominio, teniendo en cuenta características adicionales como [Li et al., 2016] e involucrando métodos como [Ya et al., 2019] , [McGahagan et al., 2019] y [Xiang et al., 2011]

2.3.2.3. Características de la marca

Como se describe en la Sección 5.1, el conocimiento sobre marcas puede proporcionar sistemas de detección, brindando seguridad a los usuarios que desean acceder a los servicios ofrecidos, conocimiento sobre colores comunes utilizados, dominios oficiales o IP relacionadas, idiomas

utilizados en páginas oficiales, así mismo, patrones comúnmente utilizados en ataques de phishing como palabras clave y patrones en la ruta pueden proporcionar características de alta calidad para sistemas de detección robustos.

3. Análisis fuentes de datos y selección de características a considerar en detección de phishing

3.1. Descripción de dataset

En esta sección se estudiará el dataset disponible para este estudio, inicialmente se abordará el dataset de URLs phishing, se explicará su obtención y basado en este, se escogerán 3 marcas de estudio con las que se realizará el análisis exploratorio para estos datasets obteniendo las palabras clave con los que se realizarán las búsquedas de URLs no maliciosas. Posteriormente basado en lo observado previamente en la literatura y en el análisis exploratorio se escogerán las características a utilizar en el modelo.

3.1.1. Dataset de URLs de phishing

Se recolectaron un total de 81.375 URLs de phishing únicas, dentro del dataset obtenido de un conjunto de fuentes como phishtank, openphish, y cuentas de twitter que reportan phishing, esta recolección incluyó una gran cantidad de trabajo manual para lograr identificar y verificar el estado del phishing para asociarlo con una marca especifica. Este dataset cuenta con URLs phishing reportadas desde el 1 de enero del 2019 hasta noviembre del 2022 lo que proporciona un rango de tiempo de casi 3 años para estudiar diferentes comportamientos en los ataques de phishing en ese periodo de tiempo.

En una primera exploración se indagó por los dominios relacionados con las URLs, encontrando una gran cantidad de dominios repetidos, al indagar por la naturaleza de los mismos se encontró que pertenecían a servicios de hosting bajo un mismo dominio, la tabla **3-1** relaciona el top 20 de estos dominios usados para cometer fraude.

También se profundizó en las marcas mayormente afectadas de nuestro dataset para así poder saber qué marcas son las que se están estudiando y enfocar el estudio hacia estas, encontrando que en el dataset recolectado se presentaba la siguiente distribución en el TOP10 **3-2**.

Aún dentro del top 10 de marcas se observa una distribución dispersa entre las marcas

Dominio	Cantidad en dataset
repl.co	5.601
000webhostapp.com	2.809
webcindario	1.450
umbler.net	596
atwebpages.com	378
ddns.net	325
sslblindado.com	183
tonohost.com	154
azure.com	151
gotdns.ch	150
teste.website	139
webnode.es	131
125mb.com	131
loca.lt	122
securewebsession.com	105
sytes.net	104
presse.ci	102
duckdns.org	99
amazonaws.com	96
joomla.com	93

Tabla 3-1.: Top 20 Dominios usados en fraude

mayormente afectadas **3-1**, es así como se pretende no solo analizar cantidad sino también constancia en el tiempo para determinar las marcas que finalmente se tomarán como referencia en el estudio.

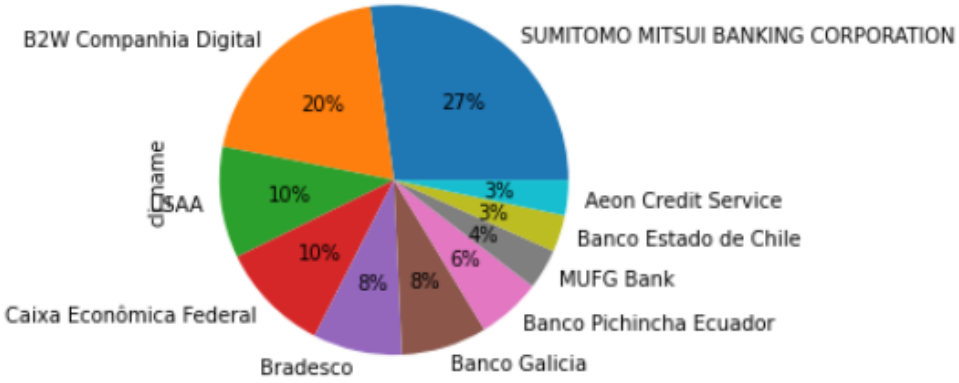


Figura 3-1.: Distribución top marcas

Marca	Cantidad de URLs phishing
Sumitomo Mitsubishi Banking Corporation	16.454
B2W Companhia digital	11.954
USAA	6.237
CaixaEconomical Federal	6.199
Bradesco	4.997
Banco Galicia	4.806
Banco Pichincha Ecuador	3.585
MUFG Bank	2.263
Banco estado de Chile	2.062
AEON Credit Service	1.971

Tabla 3-2.: Top 10 Marcas afectadas

Considerando los datos a través del tiempo se escoge utilizar 3 marcas que mantienen un numero de ataques de phishing continuo en el tiempo, como se muestra en la figura **3-2**.

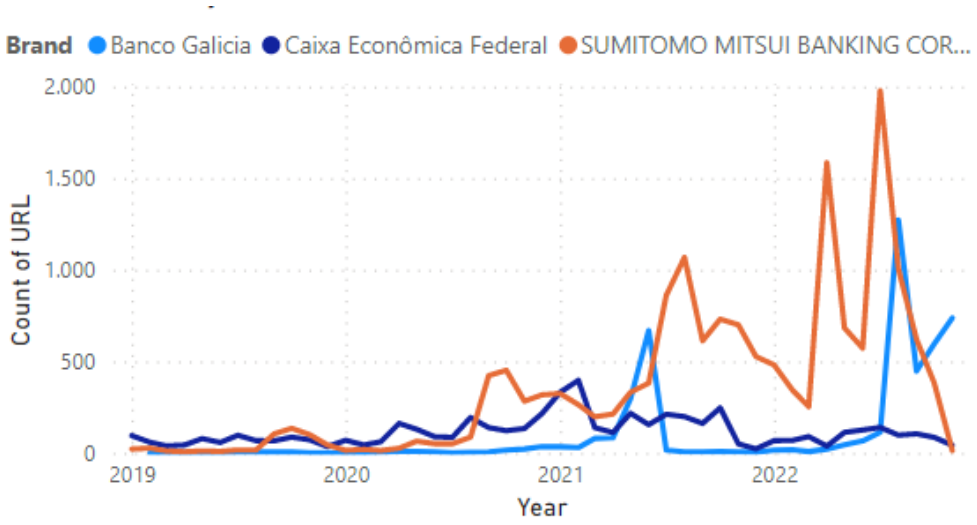


Figura 3-2.: Distribución de ataques de phishing por marca en el tiempo

Considerando **3-2** donde se observa una cantidad de muestras suficientes y constantes en el tiempo para 3 marcas y **3-1** donde esas 3 marcas se encuentran dentro de las 10 marcas con mayor ataques de phishing en el dataset, se tiene que las marcas a estudiar serán las mostradas en **3-2**, de ahora en adelante conocidas como A, B Y C .

3.1.1.1. Extracción de palabras clave

Teniendo identificadas las marcas a estudiar se procede con el análisis de las URLs para así determinar las palabras clave de cada una de las marcas. Para extraer las palabras clave de

Tabla 3-3.: Descripción del dataset URLs phishing

Dataset por Marca	Cantidad de URLs
Dataset phishing marca A	5.469
Dataset phishing marca B	4.723
Dataset phishing marca C	16.448

las marcas afectadas se consideraron las URLs de phishing asociadas a cada marca dentro del dataset, se realizó una extracción de n-caracteres a partir de 4 caracteres, posteriormente se realizó una cuantificación de los caracteres y se escogieron los que presentaron mayor recurrencia y se presentan en **3-4**.

Marca A	Marca B	Marca C
licia	caixa	smbc
galic	acess	smbc-ca
galicia	caiixa	bc-card
alici	acesso	paypay

Tabla 3-4.: palabras clave relacionadas por Marca

3.1.2. Dataset de URLs no Maliciosas

Se recolectaron un total de 3'227.761 que fueron clasificadas como no maliciosas basadas en el mismo criterio de búsqueda en phishtank , Openphish y URL reportadas por Twitter teniendo especial cuidado en relacionar las URLs a través de búsquedas de palabras clave **3-4** relacionadas a las marcas afectadas se conformó el dataset de URLS no maliciosas para cada marca.

Tabla 3-5.: Descripción del dataset URLs no phishing

Dataset por Marca	Cantidad de URLs
Dataset no-phishing marca A	53.125
Dataset no-phishing marca B	91.679
Dataset no-phishing marca C	20.770

3.2. Características

3.2.1. Características descriptivas de la URL

Dentro del análisis exploratorio se evidenció la variedad de URLs que se tienen para cada marca, es así como se necesita identificar características que puedan ayudar a caracterizar las URLs propias de la marca , para esto se tuvieron en cuenta las escogidas en la literatura como en [Xiang et al., 2011] dentro de las que se escogieron la longitud de la URL, el identificar caracteres no comunes, el numero de puntos de la URL, así como también si la URL tiene protocolo http o https (certificado ssl).

Tabla 3-6.: Descripción de características descriptivas de las URLs phishing marca A

descripción	etiqueta	puntos	caracteres	certificado ssl	longitud
conteo	58.594	58.594	58.594	58.594	58.594
promedio	0,09	1,80	2,50	0,02	63,85
std	0,29	1,20	1,37	0,15	104,54
min	0,00	0,00	0,00	0,00	11,00
25 %	0,00	1,00	2,00	0,0	25,00
50 %	0,00	1,00	2,00	0,0	31,00
75 %	0,00	2,00	2,00	0,0	47,00
max	1,00	12,00	11,00	1,0	1.234

Tabla 3-7.: Descripción de características descriptivas de las URLs phishing marca B

descripción	etiqueta	puntos	caracteres	certificado ssl	longitud
conteo	96.402	96.402	96.402	96.402	96.402
mean	0,05	2,63	3,59	0,02	222,61
std	0,22	1,43	1,92	0,14	187,07
min	0,00	0,00	2,00	0,00	11,00
25 %	0,00	2,00	2,00	0,00	41,00
50 %	0,00	2,00	3,00	0,00	140,00
75 %	0,00	3,00	6,00	0,00	323,00
max	1,00	21,00	11,00	1,00	1.941

En las tablas **3-6**, **3-7**, **3-8** se puede observar en detalle cada dataset , mostrando a través del promedio de su etiqueta que los tres corresponden a dataset desbalanceados, además de ser todos también de diferentes tamaños siendo el de la marca B el dataset de mayor

Tabla 3-8.: Descripción de características descriptivas de las URLs phishing marca C

descripción	etiqueta	puntos	caracteres	certificado ssl	longitud
conteo	37.218	37.218	37.218	37.218	37.218
mean	0,44	2,19	2,02	0,00	31,06
std	0,50	1,48	0,32	0,05	23,20
min	0,00	1,00	0,00	0,00	8,00
25 %	0,00	2,00	2,00	0,00	24,00
50 %	0,00	2,00	2,00	0,00	26,00
75 %	1,00	2,00	2,00	0,00	34,00
max	1,00	30,00	8,00	1,00	565,00

tamaño. Adicional a esto se nos muestra la característica relacionada a certificado ssl una característica que tienen la minoría de URLs.

3.2.2. Caracterización por patrón en dominio y TLD

Teniendo en cuenta la tabla **3-1** se decide incluir como característica una etiqueta que permita identificar si el dominio asociado a la URL está asociado a los dominios que son abusados más de 5 veces en el dataset, esto permite abarcar una serie de servicios que ofrecen hostear gratuitamente o que facilita herramientas para exponer un ataque de phishing. Adicionalmente se hizo un análisis en el TLD de cada URL, esto permitió identificar que el 45 % de las URLs contienen .com como TLD, debido a esto se decide agregar otra etiqueta donde haya TLDs diferentes a .com.

3.2.3. Caracterización de marca por caracteres

Se decidió incluir como característica el conteo de cada letra del abecedario de la a-z, aunque en un inicio se decidió tomar en cuenta ser sensible a mayúsculas y minúsculas e incluir también números, después de realizar el análisis de correlación entre características mostrado en A, se encontró que los números presentaban alta correlación con muchas características que se considera pueden confundir al modelo más que marcar un claro patrón. Se espera que estas características puedan marcar una tendencia asociada a la propia marca. Para esto se realizó un análisis de correlación entre la letra y la etiqueta de phishing, para lo que directamente no se encontró una relación directa ni en la marca A , ni tampoco en la B, siendo la C la que mayor correlación presenta en algunas letras. También fue llamativo el hecho de que a primera vista en las figuras **3-3**,**3-4** Y**3-5** muestran un mapa de correlación completamente diferente para cada marca.

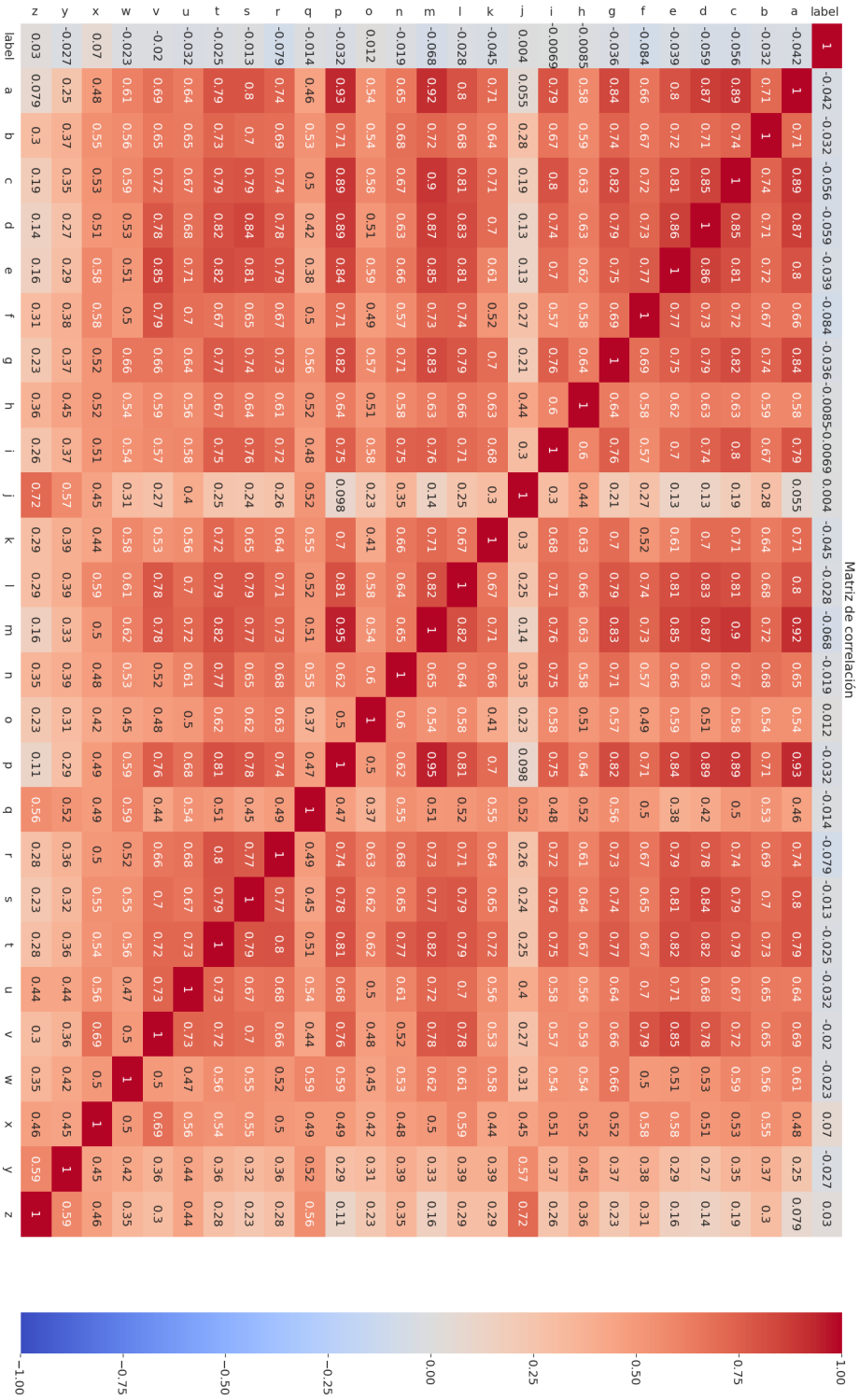


Figura 3-3.: Tabla de correlación entre caracteres y etiqueta de phishing para la marca A

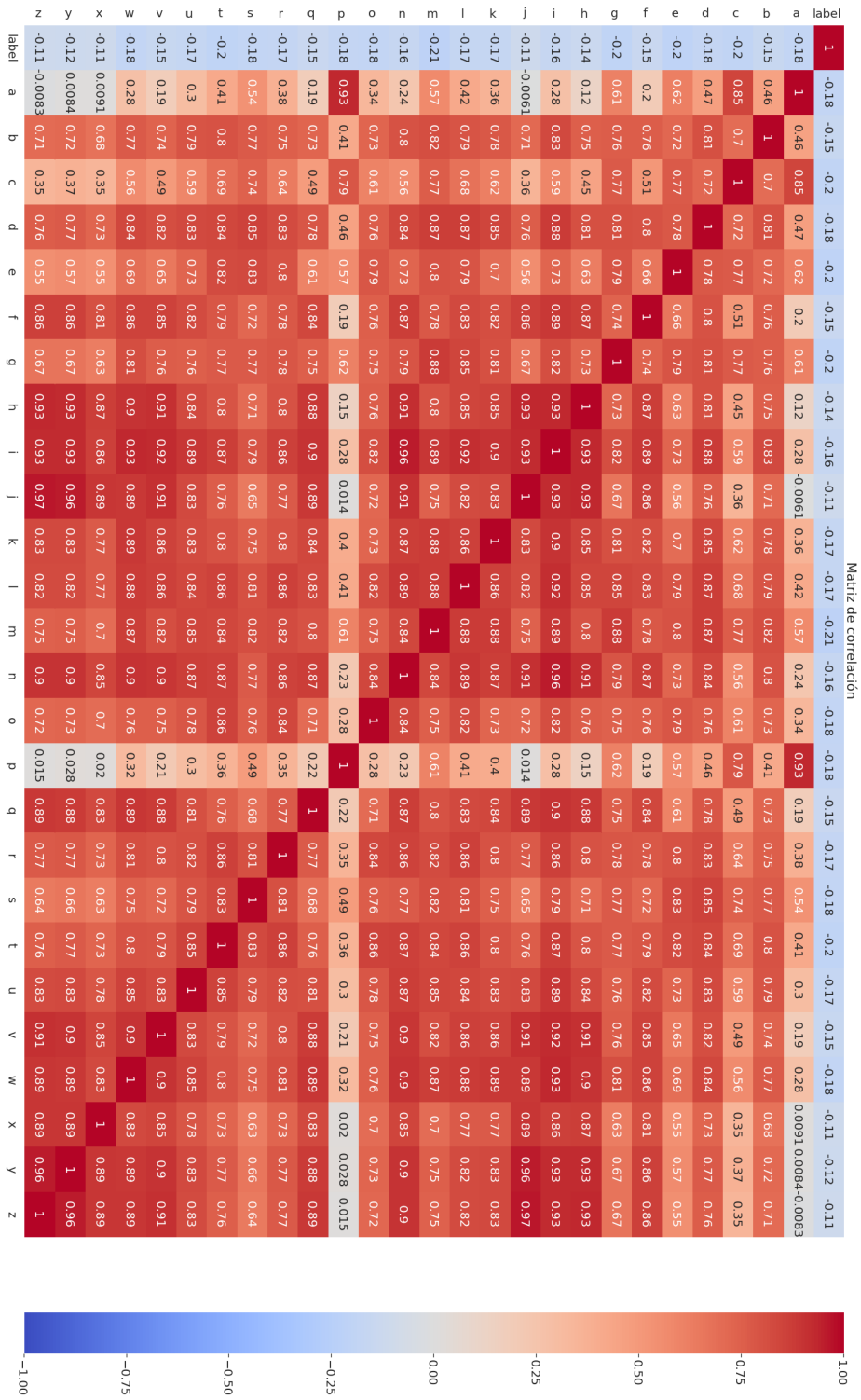


Figura 3-4.: Tabla de correlación entre caracteres y etiqueta de phishing para la marca B

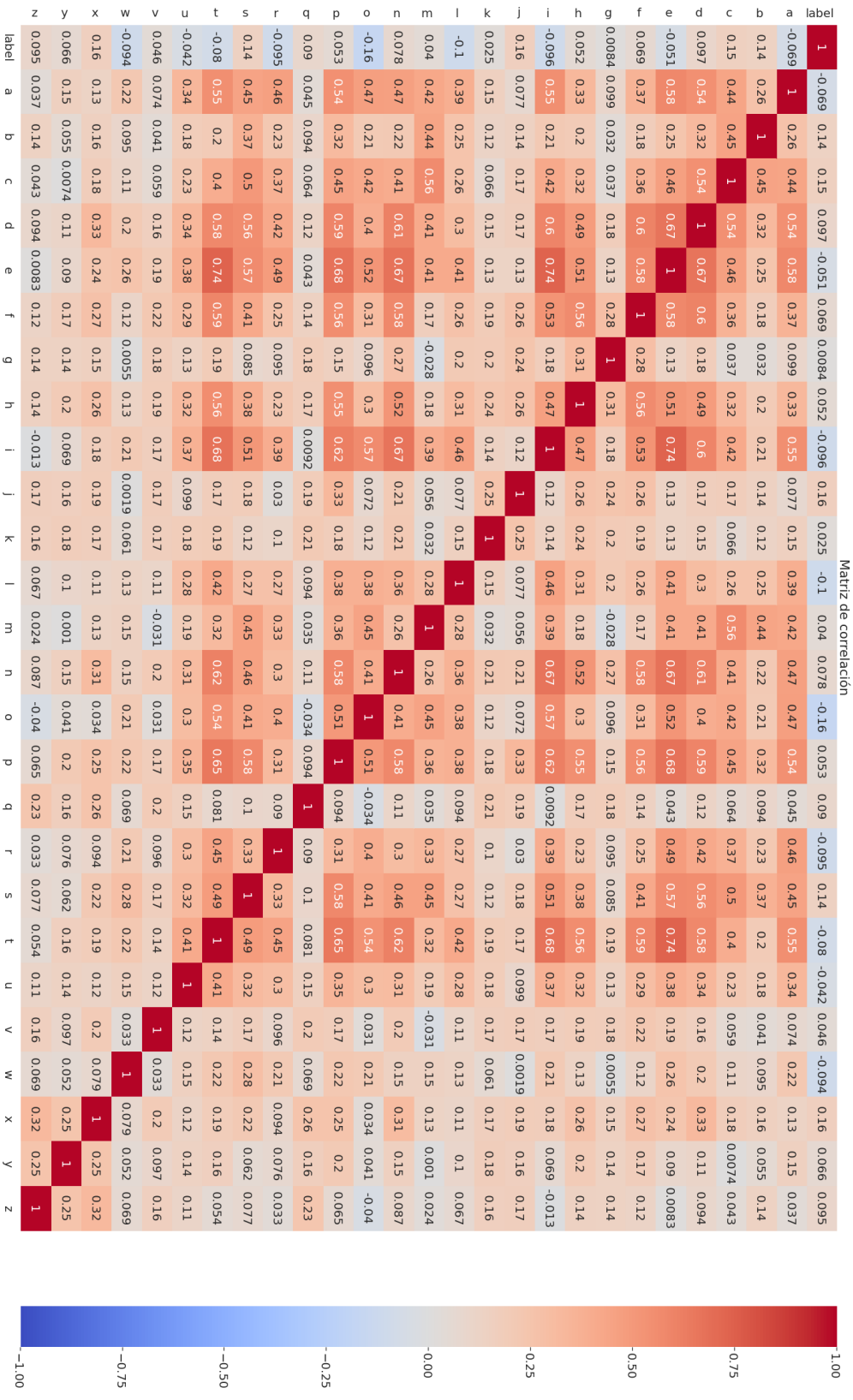


Figura 3-5.: Tabla de correlación entre caracteres y etiqueta de phishing para la marca C

Características	Descripción
Puntos	Numero de puntos encontrados en la URL
caracteres especiales	Conteo de caracteres especiales
ssl	Hace uso del protocolo https
longitud	longitud de la URL
Dominios abusados	hace parte de dominios abusados en más de 5 ocasiones
TLDs abusados	Hace parte de TLDs abusados frecuentemente excluyendo .com y .net
a..z	Conteo de caracteres en la URL

Tabla 3-9.: Características seleccionadas

En **3-9** se compilan las características seleccionadas finales a utilizar en el modelo

4. Diseño e implementación de modelo de detección de phishing en etapas tempranas

En el presente capítulo se abordará el diseño del modelo completo, el cual consta de dos partes; el primero relacionado con un modelo heurístico para clasificar y filtrar URLs relacionadas a las marcas estudiadas y una segunda parte donde se abordará un modelo clasificador de machine machine learning que permita diferenciar las URLs phishing de las que no lo son. Para el modelo heurístico se establecerá un proceso basado en la caracterización de marca que se realizó en el **capitulo tres**. Para el clasificador se realizará una exploración por los principales modelos utilizados en la literatura: Regresión logística, maquinas de soporte vectorial, bosque aleatorio y redes neuronales. Esto con el fin de seleccionar el modelo adecuado para cada marca en esta tarea. Es decir que se implementarán un total de 4 modelos por marca y adicionalmente se implementará un modelo genérico donde no se realiza distinción del dataset por marca y se utilizará clasificando su propio dataset y el dataset de cada marca para realizar la comparación. Finalmente se mostrarán los resultados de cada modelo.

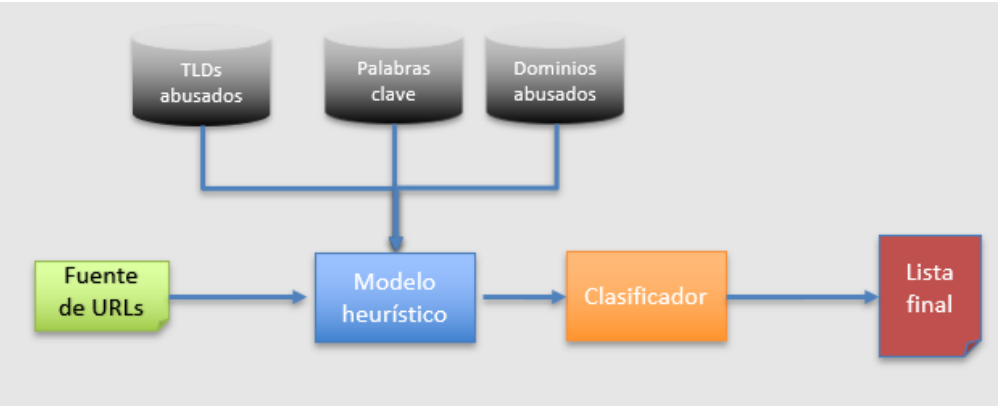


Figura 4-1.: Esquema modelo completo

4.1. Modelo heurístico

Después de reunir las fuentes de URLs de las que podemos obtener URLs maliciosas, se vuelve necesario filtrar la información de tal manera que para el modelo ya entrenado podamos obtener URLs asociadas a la marca afectada, para lo que en esta propuesta se propone establecer un modelo heurístico previo al clasificador para identificar URLs asociados a la marca, en esto se utiliza el mismo análisis hecho en **3.1.1** para dominios comúnmente abusados, **3.1.1.1** para palabras clave y **3.2.2** para TLDs abusados.

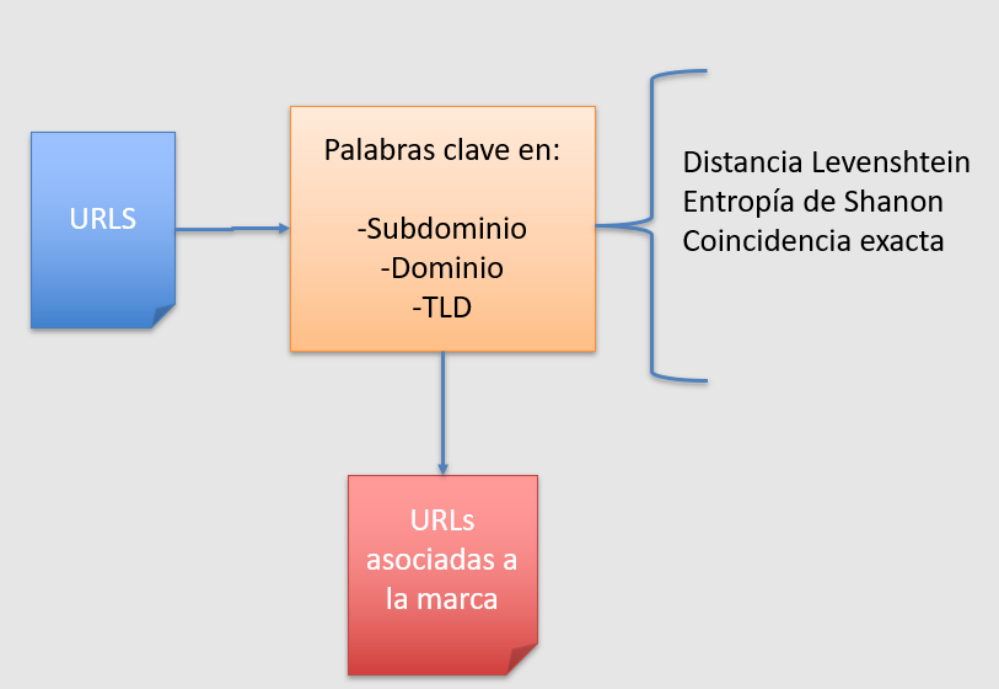


Figura 4-2.: Diagrama modelo Heurístico

El análisis de marca realizado en el **capitulo 3**, proporciona datos suficientes para realizar un primer filtro a través del modelo heurístico mostrado en **4-3** donde mediante la validación de coincidencia de palabras clave, TLDs y dominios asociados a la marca se establecen reglas para que se decir si la URL analizada por el modelo heurístico debe continuar al clasificador.

Para ello se establecieron 4 reglas, la primera buscando coincidencia exacta por palabra clave, el segundo enfocado a analizarlo a través de distancia levenstain con las palabras claves extraídas en **3-4** , el tercero coincidencia de dominio comúnmente abusado teniendo en cuenta el análisis hecho en **3-1** y por ultimo el ultimo filtro enfocado a observar la entropía de Shanon propia de la URL donde si el cálculo en la entropía es mayor a 3 pasa por el filtro teniendo en cuenta que a mayor entropía se incrementa el grado de sospecha en la URL. Las validaciones de cada método se muestra en la tabla **4-1**

Método	Validación
Coincidencia exacta palabra clave	Palabra clave dentro de la URL.
Coincidencia de dominio comúnmente abusado	Dominio de URL coincide con lista de dominios sección**3.2.2**
Cantidad de puntos en la URL	URL ya desplegada, ataque en curso.
Distancia Levenshtein	distancia levenhstein de palabra clave en URL menor a 2
Entropía de shanon	Cálculo de entropía de shanon de una cadena de carcateres mayor a 2.

Tabla 4-1.: Validaciones del métodos heurísticos

Si la URL supera alguno de esos métodos se convierte en sospechosa para pasar a ser validada por el clasificador,el diagrama del modelo se muestra en **4-3**

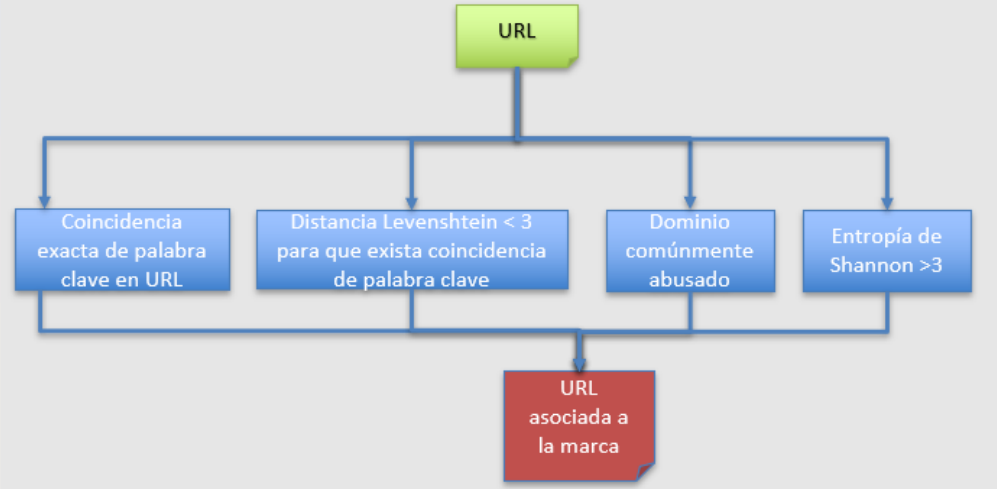


Figura 4-3.: Modelo Heurístico

4.2. Modelo Clasificador

Para el diseño del clasificador es necesario escoger uno de los tantos modelos disponibles para la tarea y observar cual puede ser el mejor modelo para la solución del problema. Para esto se exploraran los siguientes modelos para cada uno de las marcas : Regresión logística, máquina de soporte vectorial, bosque aleatorio y finalmente red neuronal.

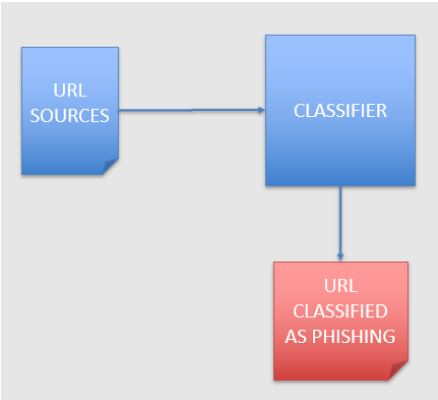
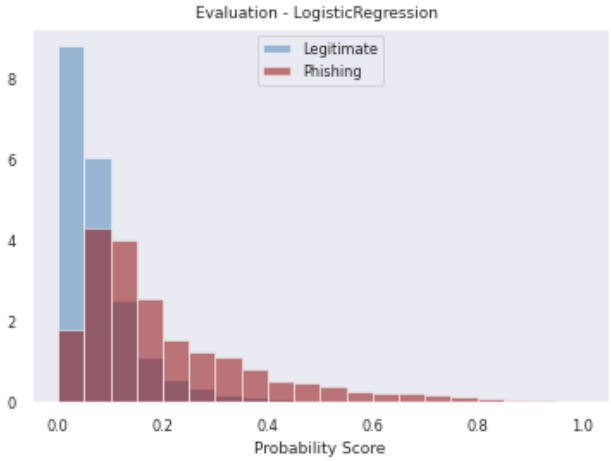


Figura 4-4.: Diagrama de clasificador.

4.2.1. Regresión logística

Para este modelo se utilizan la matriz de características construida en la sección anterior. Como parámetro de entrada del modelo, se utiliza una distribución 70-30 para entrenamiento y pruebas respectivamente. Para la construcción del modelo se hace uso de la librería sklearn donde dentro de sus librerías proporciona en la sección de modelos lineales una manera fácil de implementar un modelo de regresión lineal donde se configura una fuerza de regularización inversa de 1.1, un máximo de iteraciones de 1.000.

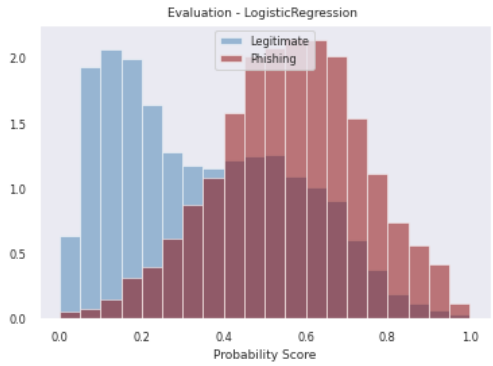


threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,74	0,69	0,22	0,74
0,20	0,92	0,37	0,33	0,87
0,30	0,97	0,23	0,43	0,90
0,40	0,99	0,13	0,50	**0,91**
0,50	0,99	0,08	0,56	0,91
0,60	1,00	0,05	0,63	0,91
0,70	1,00	0,02	0,64	0,91
0,80	1,00	0,01	0,65	0,91
0,90	1,00	0,00	0,75	0,91
1,00	1,00	0,00	1,00	0,91

Figura 4-5.: Resultados de regresión logística para la marca A

threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,74	0,69	0,22	0,74
0,20	0,92	0,37	0,33	0,87
0,30	0,97	0,23	0,43	0,90
0,40	0,99	0,13	0,50	**0,91**
0,50	0,99	0,08	0,56	0,91
0,60	1,00	0,05	0,63	0,91
0,70	1,00	0,02	0,64	0,91
0,80	1,00	0,01	0,65	0,91
0,90	1,00	0,00	0,75	0,91
1,00	1,00	0,00	1,00	0,91

Figura 4-6.: Resultados de regresión logística para la marca B

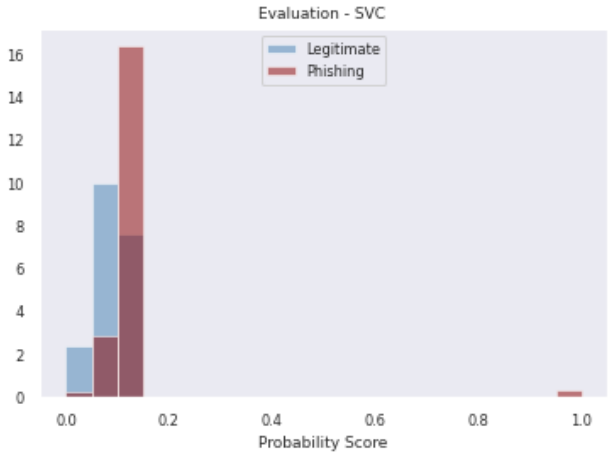


threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,13	0,99	0,47	0,51
0,20	0,33	0,97	0,53	0,61
0,30	0,48	0,92	0,58	0,67
0,40	0,59	0,82	0,62	**0,70**
0,50	0,72	0,64	0,64	0,68
0,60	0,84	0,43	0,67	0,66
0,70	0,93	0,22	0,72	0,62
0,80	0,98	0,09	0,78	0,59
0,90	0,99	0,03	0,80	0,57
1,00	1,00	0,00	1,00	0,56

Figura 4-7.: Resultados de regresión logística para la marca C

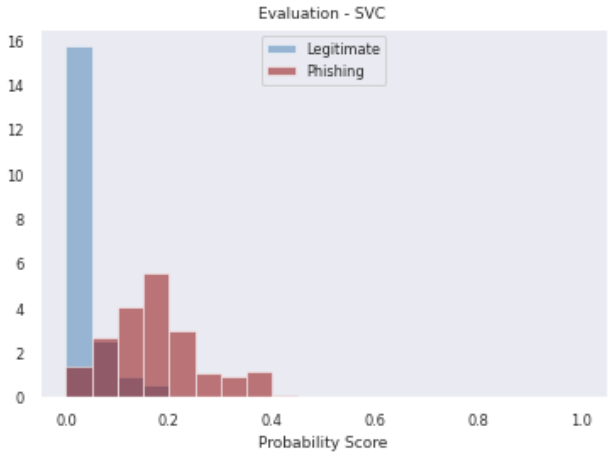
4.2.2. Máquinas de soporte vectorial

Para este modelo se utiliza la misma distribución de datos para entrenamiento y pruebas. Utilizando sklearn su función especifica que permite crear un modelo de maquina de soporte vectorial bastante sencilla de implementar.



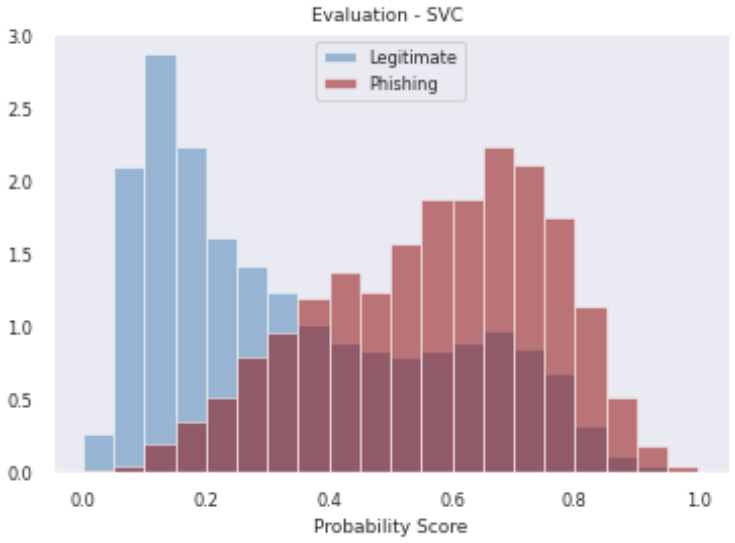
threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,62	0,84	0,18	0,64
0,20	1,00	0,02	0,60	0,91
0,30	1,00	0,02	0,65	0,91
0,40	1,00	0,02	0,68	0,91
0,50	1,00	0,02	0,77	0,91
0,60	1,00	0,02	0,81	0,91
0,70	1,00	0,02	0,87	0,91
0,80	1,00	0,02	0,89	0,91
0,90	1,00	0,02	0,94	0,91
1,00	1,00	0,00	1,00	**0,91**

Figura 4-8.: Resultados de maquina de soporte vectorial para la marca A



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,05	0,05
0,10	0,92	0,80	0,33	0,91
0,20	0,99	0,32	0,64	0,96
0,30	1,00	0,11	0,79	**0,96**
0,40	1,00	0,00	0,50	0,95
0,50	1,00	0,00	0,00	0,95
0,60	1,00	0,00	0,00	0,95
0,70	1,00	0,00	0,00	0,95
0,80	1,00	0,00	0,00	0,95
0,90	1,00	0,00	1,00	0,95
1,00	1,00	0,00	1,00	0,95

Figura 4-9.: Resultados de maquina de soporte vectorial para la marca B

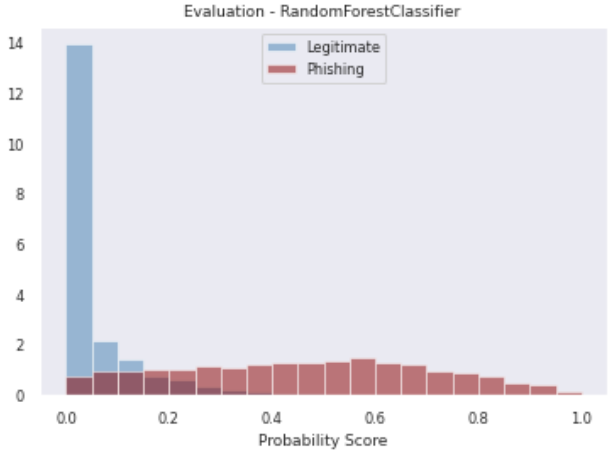


threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,12	1,00	0,47	0,51
0,20	0,37	0,97	0,55	0,64
0,30	0,53	0,90	0,60	0,69
0,40	0,64	0,80	0,64	**0,71**
0,50	0,73	0,65	0,66	0,70
0,60	0,81	0,49	0,67	0,67
0,70	0,90	0,29	0,69	0,63
0,80	0,97	0,09	0,75	0,59
0,90	1,00	0,01	0,73	0,56
1,00	1,00	0,00	1,00	0,56

Figura 4-10.: Resultados de maquina de soporte vectorial para la marca C

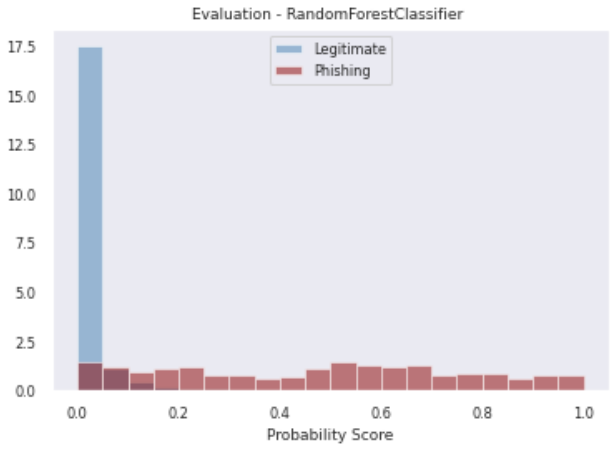
4.2.3. Bosque aleatorio

Para la implementación del modelo de bosque aleatorio se utilizó la librería sklearn que proporciona una clase de clasificador de bosque aleatorio lista para configurar, donde se estableció un estimador-n de 150 y se procedió a evaluar el modelo de la misma forma que se hizo previamente.



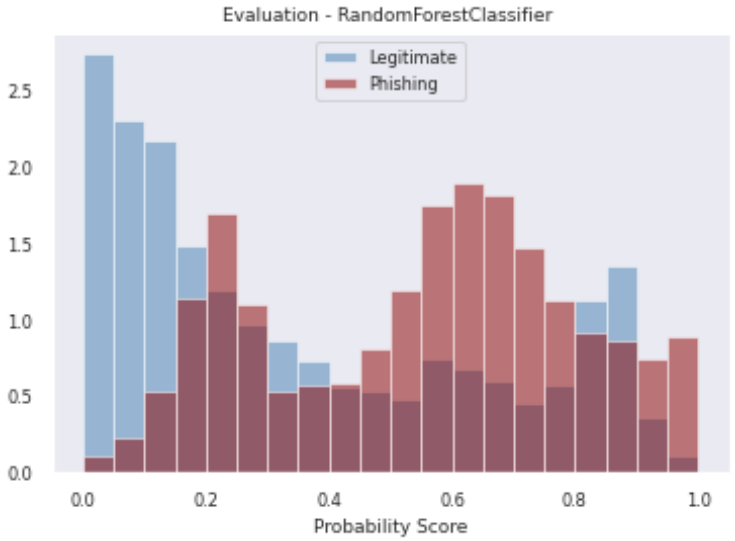
threshold	specificity	recall	precision	accuracy
0,00	0,34	0,99	0,13	0,40
0,10	0,82	0,91	0,34	0,82
0,20	0,92	0,81	0,51	0,91
0,30	0,97	0,70	0,68	0,94
0,40	0,99	0,58	0,81	**0,95**
0,50	0,99	0,45	0,89	0,94
0,60	1,00	0,32	0,91	0,93
0,70	1,00	0,19	0,92	0,92
0,80	1,00	0,09	0,95	0,91
0,90	1,00	0,03	0,98	0,91
1,00	1,00	0,00	1,00	0,91

Figura 4-11.: Resultados de bosque aleatorio para la marca A



threshold	specificity	recall	precision	accuracy
0,00	0,69	0,99	0,14	0,71
0,10	0,93	0,86	0,40	0,93
0,20	0,96	0,76	0,51	0,95
0,30	0,97	0,66	0,57	0,96
0,40	0,98	0,59	0,59	**0,96**
0,50	0,98	0,50	0,59	0,96
0,60	0,99	0,36	0,57	0,96
0,70	0,99	0,23	0,56	0,95
0,80	0,99	0,15	0,55	0,95
0,90	1,00	0,08	0,70	0,95
1,00	1,00	0,00	1,00	0,95

Figura 4-12.: Resultados de bosque aleatorio para la marca B



threshold	specificity	recall	precision	accuracy
0,00	0,02	1,00	0,45	0,45
0,10	0,26	0,98	0,51	0,58
0,20	0,44	0,90	0,56	0,64
0,30	0,54	0,76	0,57	0,64
0,40	0,62	0,70	0,60	0,66
0,50	0,68	0,63	0,61	**0,66**
0,60	0,74	0,49	0,60	0,63
0,70	0,80	0,30	0,55	0,58
0,80	0,85	0,17	0,48	0,55
0,90	0,98	0,08	0,74	0,58
1,00	1,00	0,00	1,00	0,56

Figura 4-13.: Resultados de bosque aleatorio para la marca C

De este modelo podemos obtener información de la relevancia de las características por medio de una análisis de importancia de características que proporciona la misma librería, es interesante observar de que manera están impactando las características en la clasificación para así poder observar patrones asociados a la marca.

pos	característica	nombre	porcentaje
1	feature 3	len	9,1 %
2	feature 0	points	7,1 %
3	feature 24	s	4,8 %
4	feature 6	a	4,7 %
5	feature 21	p	4,5 %
6	feature 8	c	4,1 %

Sigue en la página siguiente.

pos	característica	nombre	porcentaje
7	feature 23	r	4,1 %
8	feature 10	e	4,1 %
9	feature 20	o	3,9 %
10	feature 14	i	3,7 %
11	feature 25	t	3,5 %
12	feature 18	m	3,4 %
13	feature 19	n	3,4 %
14	feature 28	w	3,1 %
15	feature 17	l	3,0 %
16	feature 7	b	2,9 %
17	feature 9	d	2,9 %
18	feature 5	abusedtld	2,8 %
19	feature 29	x	2,6 %
20	feature 12	g	2,5 %
21	feature 26	u	2,3 %
22	feature 11	f	2,3 %
23	feature 13	h	2,1 %
24	feature 4	freehost	2,0 %
25	feature 27	v	2,0 %
26	feature 1	characters	1,7 %
27	feature 16	k	1,5 %
28	feature 31	z	1,2 %
29	feature 30	y	1,1 %
30	feature 22	q	0,97 %
31	feature 15	j	0,82 %
32	feature 2	ssl	0,02 %

Tabla 4-2.: Ranking de carcateristicas marca A.

pos	característica	nombre	porcentaje
1	feature 3	len	12,6 %
2	feature 0	points	6,8 %
3	feature 1	m	6,6 %
4	feature 21	p	5,6 %
5	feature 28	w	4,8 %
6	feature 17	l	4,4 %
7	feature 10	e	4,4 %

Sigue en la página siguiente.

pos	característica	nombre	porcentaje
8	feature 25	t	4,3 %
9	feature 6	a	4,0 %
10	feature 14	i	3,9 %
11	feature 23	r	3,8 %
12	feature 24	s	3,7 %
13	feature 20	o	3,7 %
14	feature 8	c	3,6 %
15	feature 12	g	3,3 %
16	feature 19	n	3,3 %
17	feature 4	freehost	2,6 %
18	feature 9	d	2,4 %
19	feature 7	b	2,3 %
20	feature 13	h	1,9 %
21	feature 26	u	1,7 %
22	feature 11	f	1,5 %
23	feature 27	v	1,3 %
24	feature 16	k	1,2 %
25	feature 5	abusedtld	1,0 %
26	feature 30	y	0,95 %
27	feature 1	characters	0,79 %
28	feature 29	x	0,74 %
29	feature 15	j	0,66 %
30	feature 31	z	0,55 %
31	feature 22	q	0,35 %
32	feature 2	ssl	0,01 %

Tabla 4-3.: Ranking de carcateristicas marca B.

pos	característica	nombre	porcentaje
1	feature 3	len	8,9 %
2	feature 24	s	7,1 %
3	feature 8	c	5,4 %
4	feature 20	o	4,8 %
5	feature 0	points	4,2 %
6	feature 14	i	4,0 %
7	feature 7	b	3,8 %
8	feature 18	m	3,7 %

Sigue en la página siguiente.

pos	característica	nombre	porcentaje
9	feature 10	e	3,7 %
10	feature 28	w	3,7 %
11	feature 6	a	3,6 %
12	feature 15	j	3,5 %
13	feature 19	n	3,5 %
14	feature 9	d	3,3 %
15	feature 21	p	3,2 %
16	feature 25	t	3,1 %
17	feature 23	r	3,0 %
18	feature 29	x	2,7 %
19	feature 13	h	2,6 %
20	feature 17	l	2,4 %
21	feature 30	y	2,4 %
22	feature 11	f	2,3 %
23	feature 26	u	2,3 %
24	feature 12	g	2,3 %
25	feature 27	v	2,1 %
26	feature 16	k	1,9 %
27	feature 31	z	1,8 %
28	feature 22	q	1,5 %
29	feature 5	abusedtld	1,4 %
30	feature 1	characters	0,2 %
31	feature 4	freehost	0,1 %
32	feature 2	ssl	0,03 %

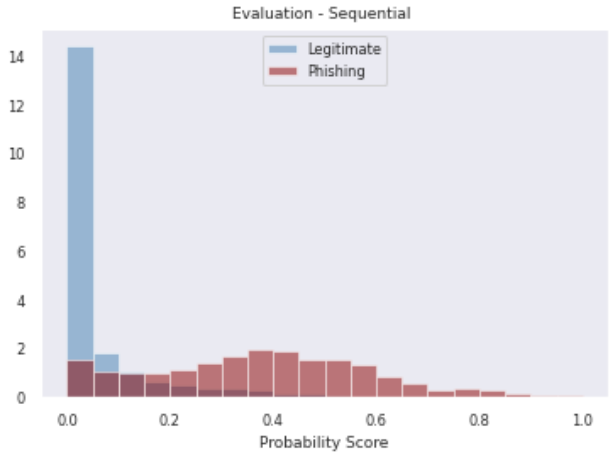
Tabla 4-4.: Ranking de características marca C.

4.2.4. Red Neuronal

Finalmente el último modelo a implementar es una red neuronal multicapa, el cual se implementa utilizando la librería tensorflow, siendo un modelo secuencial. La arquitectura se muestra en la **4-14**, con una salida sigmoide. Este mismo modelo se implementa para las tres marcas.

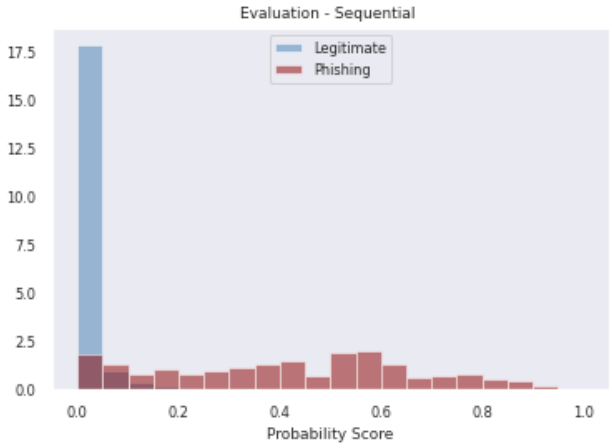
```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_24 (Dense)            (None, 32)                1056

 batch_normalization_12 (Bat  (None, 32)               128
 chNormalization)

 dense_25 (Dense)            (None, 20)                660

 batch_normalization_13 (Bat  (None, 20)               80
 chNormalization)

 dropout_12 (Dropout)        (None, 20)                0

 dense_26 (Dense)            (None, 10)                210

 dropout_13 (Dropout)        (None, 10)                0

 dense_27 (Dense)            (None, 1)                 11

=================================================================
Total params: 2,145
Trainable params: 2,041
Non-trainable params: 104
```

Figura 4-14.: Arquitectura red neuronal



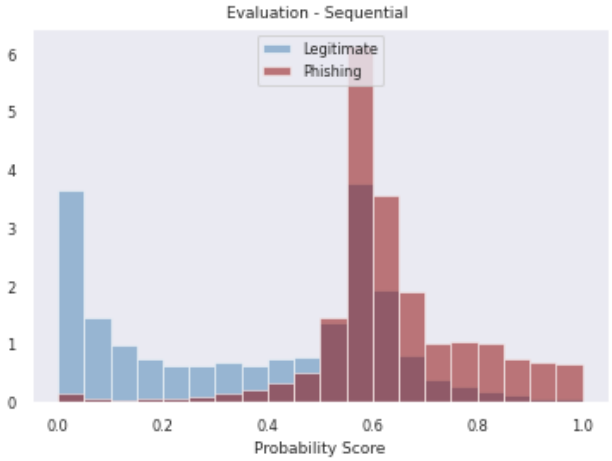
threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,81	0,87	0,32	0,82
0,20	0,90	0,77	0,44	0,89
0,30	0,94	0,64	0,54	0,91
0,40	0,98	0,45	0,66	**0,93**
0,50	0,99	0,28	0,76	0,92
0,60	1,00	0,14	0,85	0,92
0,70	1,00	0,06	0,90	0,91
0,80	1,00	0,03	0,98	0,91
0,90	1,00	0,01	1,00	0,91
1,00	1,00	0,00	1,00	0,91

Figura 4-15.: Resultados de red neuronal multicapa para la marca A



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,05	0,05
0,10	0,94	0,85	0,41	0,93
0,20	0,97	0,75	0,53	0,96
0,30	0,98	0,67	0,60	**0,96**
0,40	0,98	0,55	0,64	0,96
0,50	0,99	0,43	0,67	0,96
0,60	1,00	0,24	0,71	0,96
0,70	1,00	0,14	0,74	0,96
0,80	1,00	0,07	0,76	0,95
0,90	1,00	0,01	0,75	0,95
1,00	1,00	0,00	1,00	0,95

Figura 4-16.: Resultados de red neuronal multicapa para la marca B



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,26	0,99	0,51	0,58
0,20	0,34	0,98	0,54	0,63
0,30	0,41	0,97	0,57	0,66
0,40	0,47	0,95	0,59	0,69
0,50	0,55	0,91	0,62	**0,71**
0,60	0,81	0,53	0,68	0,68
0,70	0,94	0,26	0,78	0,64
0,80	0,98	0,15	0,84	0,61
0,90	0,99	0,07	0,88	0,58
1,00	1,00	0,00	1,00	0,56

Figura 4-17.: Resultados de red neuronal multicapa para la marca C

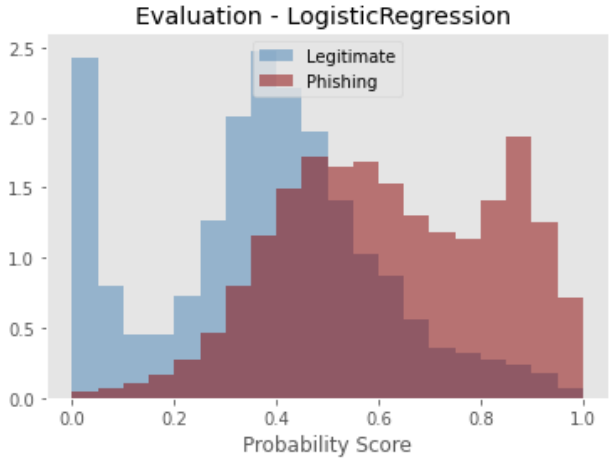
4.3. Clasificador genérico

Para comparar los modelos por marca contra un modelo genérico, se realizó el mismo ejercicio sin distinguir la marca, es decir el mismo dataset inicial con la diferencia que no se separa el dataset por marca. Se entrena el modelo genérico con un dataset balanceado de URLs phishing y no phishing asociadas a diferentes marcas **3-1**. y se realizarán 2 experimentos; el primero consiste en evaluar el modelo con un muestreo de su propio dataset y el segundo experimento consiste en evaluar el modelo con el dataset de cada marca A, B Y C.

4.3.1. Regresión logística modelo genérico

Para este modelo se utilizan las mismas características de los modelos anteriores, con una distribución 70-30 para entrenamiento y pruebas respectivamente. para la construcción del modelo se hace uso de la librería sklearn para implementar un modelo de regresión lineal donde se configura una fuerza de regularización inversa de 1.1 , un máximo de iteraciones de 1000.

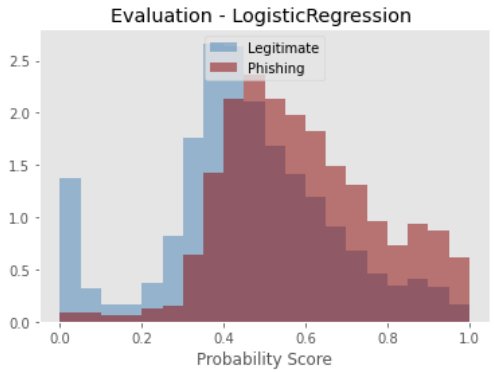
4.3.1.1. Evaluación con dataset propio



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,50	0,50
0,10	0,16	0,99	0,54	0,58
0,20	0,21	0,98	0,55	0,59
0,30	0,31	0,94	0,58	0,63
0,40	0,53	0,85	0,64	0,69
0,50	0,74	0,69	0,72	**0,71**
0,60	0,86	0,52	0,78	0,69
0,70	0,93	0,38	0,84	0,65
0,80	0,96	0,26	0,87	0,61
0,90	0,99	0,10	0,89	0,54
1,00	1,00	0,00	1,00	0,50

Figura 4-18.: Resultados de regresión logística para Modelo genérico

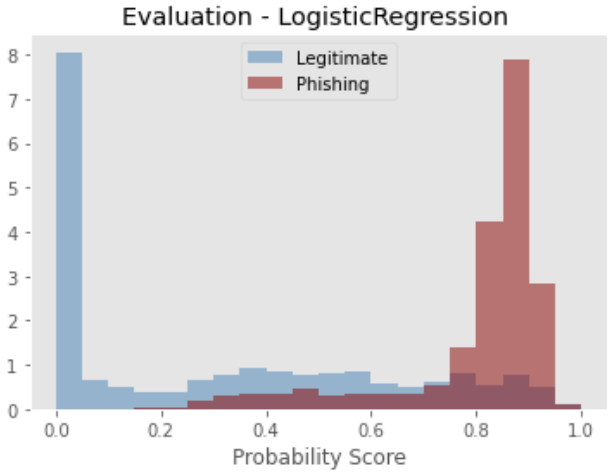
4.3.1.2. Evaluación modelo genérico con dataset marca A



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,08	0,99	0,10	0,17
0,20	0,10	0,99	0,10	0,18
0,30	0,16	0,97	0,11	0,24
0,40	0,38	0,87	0,13	0,43
0,50	0,62	0,64	0,15	0,62
0,60	0,78	0,44	0,17	0,74
0,70	0,88	0,27	0,19	0,82
0,80	0,94	0,16	0,21	0,86
0,90	0,98	0,07	0,24	0,89
1,00	1,00	0,00	1,00	**0,91**

Figura 4-19.: Resultados de regresión logística para Modelo genérico con dataset marca A

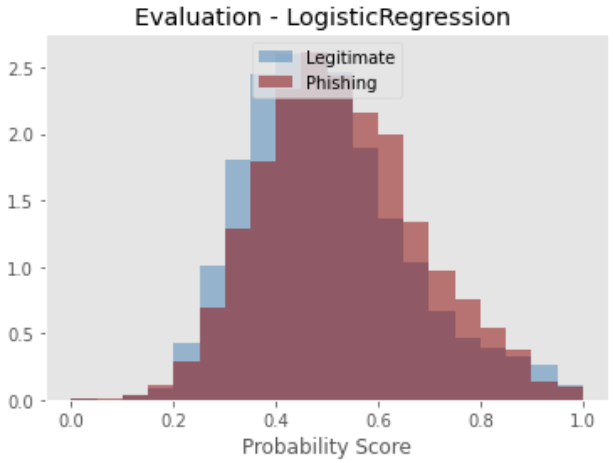
4.3.1.3. Evaluación modelo genérico con dataset marca B



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,05	0,05
0,10	0,43	1,00	0,08	0,46
0,20	0,48	1,00	0,09	0,50
0,30	0,53	0,99	0,10	0,55
0,40	0,62	0,95	0,11	0,63
0,50	0,70	0,92	0,13	0,71
0,60	0,78	0,88	0,17	0,79
0,70	0,83	0,85	0,21	0,83
0,80	0,90	0,75	0,29	0,90
0,90	0,97	0,15	0,20	0,93
1,00	1,00	0,00	1,00	**0,95**

Figura 4-20.: Resultados de regresión logística para Modelo genérico con dataset marca B

4.3.1.4. Evaluación modelo genérico con dataset marca C



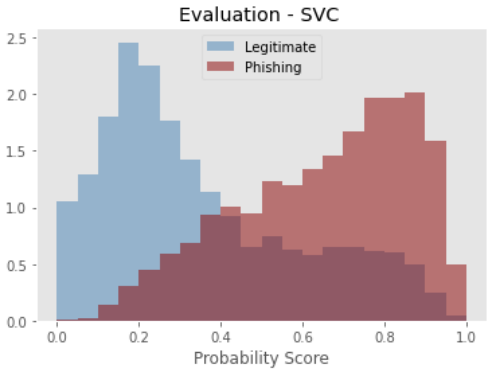
threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,00	1,00	0,44	0,44
0,20	0,01	0,99	0,44	0,44
0,30	0,08	0,94	0,45	0,46
0,40	0,29	0,79	0,47	0,51
0,50	0,55	0,54	0,49	0,55
0,60	0,77	0,31	0,52	**0,57**
0,70	0,89	0,14	0,51	0,56
0,80	0,95	0,06	0,46	0,55
0,90	0,98	0,01	0,34	0,55
1,00	1,00	0,00	1,00	0,56

Figura 4-21.: Resultados de regresión logística para Modelo genérico con dataset marca C

4.3.2. Máquinas de soporte vectorial modelo genérico

Para este modelo se utiliza la misma distribución de datos para entrenamiento y pruebas.Utilizando sklearn su función especifica que permite crear un modelo de maquina de soporte vectorial bastante sencilla de implementar.

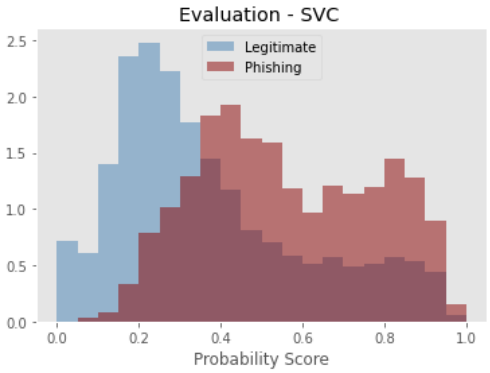
4.3.2.1. Evaluación dataset propio



threshold	specificity	recall	precision	accuracy
0,0	0,00	1,00	0,49	0,49
0,1	0,11	0,99	0,53	0,55
0,2	0,33	0,97	0,59	0,65
0,3	0,53	0,92	0,66	0,72
0,4	0,65	0,84	0,719	**0,75**
0,5	0,74	0,73	0,74	0,73
0,6	0,80	0,62	0,76	0,71
0,7	0,86	0,48	0,78	0,67
0,8	0,93	0,30	0,81	0,61
0,9	0,98	0,10	0,87	0,54
1,0	1,00	0,00	1,00	0,50

Figura 4-22.: Resultados de maquina de soporte vectorial modelo genérico

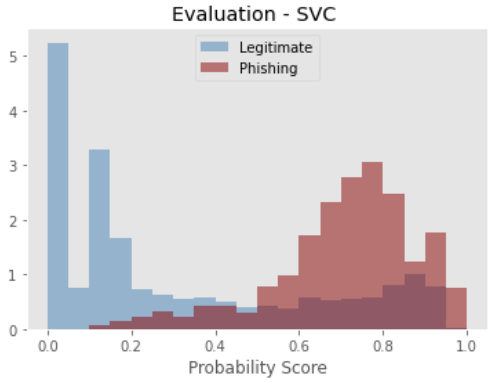
4.3.2.2. Evaluación modelo genérico con dataset marca A



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,07	1,00	0,10	0,15
0,20	0,25	0,98	0,12	0,32
0,30	0,49	0,89	0,15	0,53
0,40	0,65	0,73	0,18	0,66
0,50	0,76	0,54	0,19	0,74
0,60	0,82	0,41	0,19	0,78
0,70	0,87	0,31	0,19	0,82
0,80	0,92	0,19	0,19	0,85
0,90	0,98	0,05	0,18	0,89
1,00	1,00	0,00	1,00	**0,91**

Figura 4-23.: Resultados de maquina de soporte vectorial modelo genérico con dataset marca A

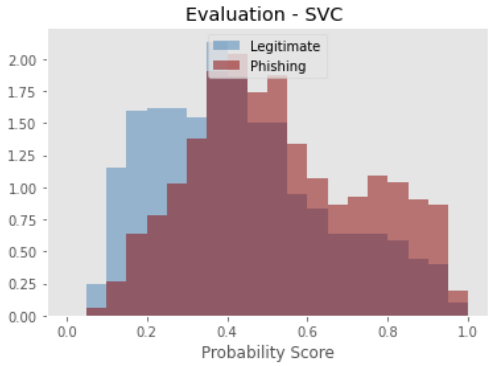
4.3.2.3. Evaluación modelo genérico con dataset marca B



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,05	0,05
0,10	0,30	1,00	0,07	0,33
0,20	0,55	0,99	0,10	0,57
0,30	0,62	0,96	0,11	0,63
0,40	0,67	0,93	0,13	0,69
0,50	0,72	0,89	0,14	0,73
0,60	0,76	0,81	0,15	0,76
0,70	0,81	0,60	0,14	0,80
0,80	0,87	0,31	0,11	0,84
0,90	0,96	0,13	0,14	0,92
1,00	1,00	0,00	1,00	**0,95**

Figura 4-24.: Resultados de maquina de soporte vectorial modelo genérico con dataset marca B

4.3.2.4. Evaluación modelo genérico con dataset marca C



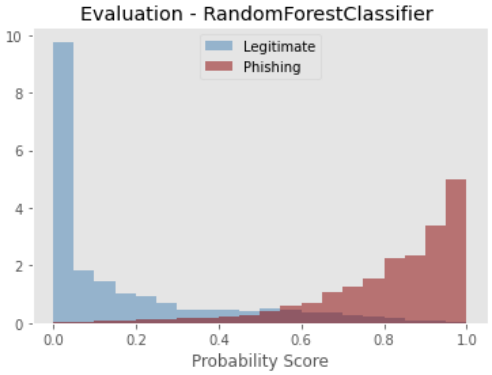
threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,01	1,00	0,44	0,45
0,20	0,15	0,95	0,47	0,50
0,30	0,31	0,86	0,50	0,55
0,40	0,49	0,70	0,52	0,58
0,50	0,68	0,49	0,55	**0,60**
0,60	0,79	0,35	0,56	0,59
0,70	0,86	0,25	0,58	0,59
0,80	0,92	0,15	0,61	0,58
0,90	0,97	0,05	0,62	0,57
1,00	1,00	0,00	1,00	0,56

Figura 4-25.: Resultados de maquina de soporte vectorial modelo genérico con dataset marca C

4.3.3. Bosque aleatorio modelo genérico

Para la implementación del modelo de bosque aleatorio se utilizó la librería sklearn que proporciona una clase de clasificador de bosque aleatorio lista para configurar, donde es estableció un estimador-n de 150 y se procedió a evaluar el modelo de la misma forma que se hizo previamente.

4.3.3.1. Evaluación con dataset propio



threshold	specificity	recall	precision	accuracy
0,0	0,29	0,99	0,58	0,64
0,1	0,58	0,99	0,70	0,79
0,2	0,70	0,98	0,77	0,84
0,3	0,78	0,97	0,81	0,88
0,4	0,83	0,95	0,85	0,89
0,5	0,87	0,92	0,88	**0,90**
0,6	0,92	0,87	0,91	0,89
0,7	0,95	0,78	0,94	0,87
0,8	0,98	0,63	0,97	0,81
0,9	0,99	0,40	0,98	0,69
1,0	1,00	0,00	1,00	0,50

Figura 4-26.: Resultados de bosque aleatorio modelo genérico

pos	característica	nombre	porcentaje
1.	feature 3	len	10,4 %
2.	feature 0	points	6,7 %
3.	feature 1	characters	5,2 %
4.	feature 6	a	4,6 %
5.	feature 18	m	4,5 %
6.	feature 20	o	4,4 %
7.	feature 8	c	4,0 %
8.	feature 4	freehost	3,8 %
9.	feature 10	e	3,6 %
10.	feature 21	p	3,6 %
11.	feature 24	s	3,4 %

Sigue en la página siguiente.

pos	característica	nombre	porcentaje
12.	feature 16	k	3,2 %
13.	feature 19	n	3,1 %
14.	feature 14	i	3,0 %
15.	feature 25	t	2,9 %
16.	feature 23	r	2,8 %
17.	feature 26	u	2,7 %
18.	feature 17	l	2,7 %
19.	feature 28	w	2,5 %
20.	feature 7	b	2,4 %
21.	feature 9	d	2,4 %
22.	feature 12	g	2,2 %
23.	feature 13	h	2,1 %
24.	feature 15	j	2,0 %
25.	feature 30	y	1,7 %
26.	feature 5	abusedtld	1,7 %
27.	feature 11	f	1,6 %
28.	feature 29	x	1,4 %
29.	feature 27	v	1,4 %
30.	feature 22	q	1,3 %
31.	feature 31	z	1,2 %
32.	feature 2	ssl	0,03 %

Tabla 4-5.: Ranking de características modelo genérico.

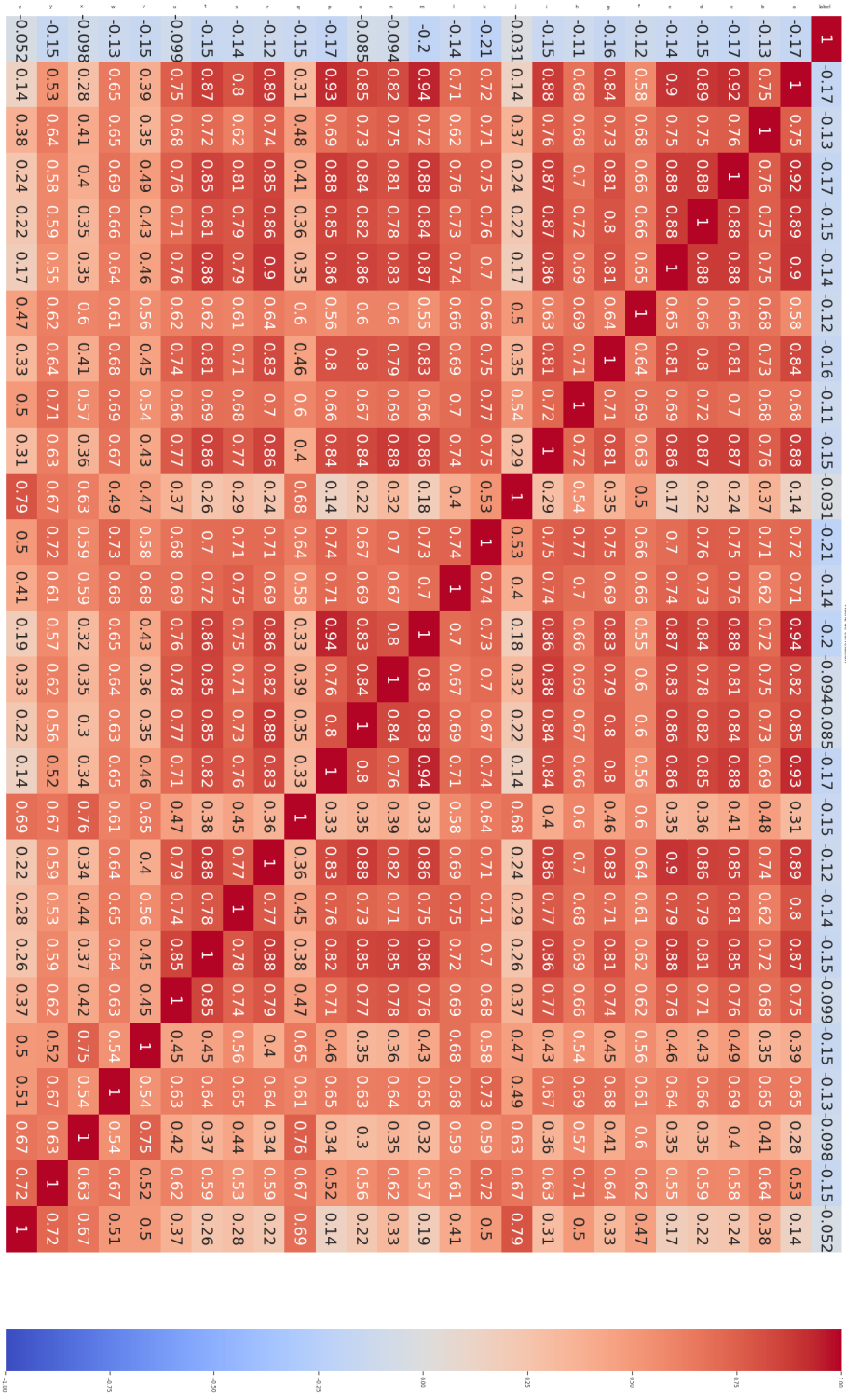
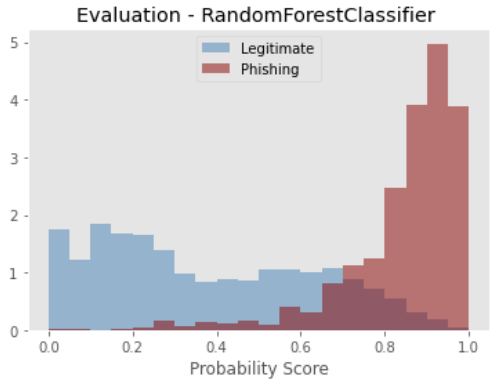


Figura 4-27.: Correlación caracteres a ataques de phishing genérico

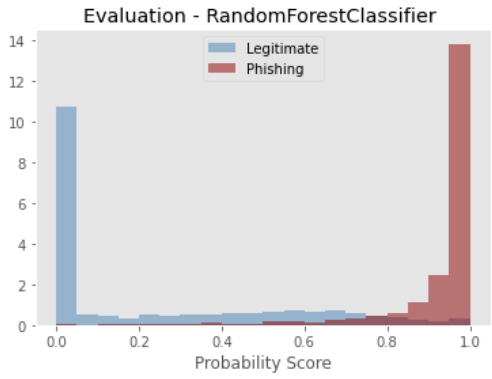
4.3.3.2. Evaluación modelo genérico con dataset marca A



threshold	specificity	recall	precision	accuracy
0,00	0,05	1,00	0,10	0,14
0,10	0,16	1,00	0,11	0,24
0,20	0,33	1,00	0,13	0,40
0,30	0,48	0,99	0,16	0,52
0,40	0,57	0,98	0,19	0,61
0,50	0,66	0,96	0,23	0,69
0,60	0,76	0,94	0,29	0,78
0,70	0,87	0,88	0,40	0,87
0,80	0,95	0,75	0,60	0,93
0,90	0,99	0,42	0,82	**0,94**
1,00	1,00	0,00	1,00	0,91

Figura 4-28.: Resultados de bosque aleatorio modelo genérico con dataset marca A

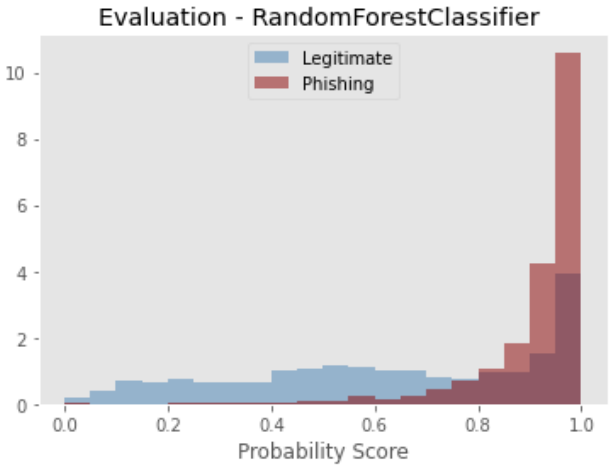
4.3.3.3. Evaluación modelo genérico con dataset marca B



threshold	specificity	recall	precision	accuracy
0,00	0,36	1,00	0,07	0,39
0,10	0,56	1,00	0,10	0,58
0,20	0,60	0,99	0,11	0,62
0,30	0,65	0,99	0,13	0,66
0,40	0,70	0,98	0,14	0,71
0,50	0,76	0,97	0,17	0,77
0,60	0,82	0,95	0,22	0,83
0,70	0,89	0,94	0,30	0,89
0,80	0,94	0,90	0,45	0,94
0,90	0,98	0,81	0,64	**0,97**
1,00	1,00	0,00	1,00	0,95

Figura 4-29.: Resultados de bosque aleatorio modelo genérico con dataset marca B

4.3.3.4. Evaluación modelo genérico con dataset marca C



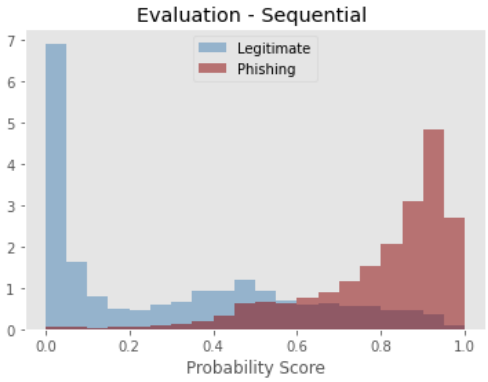
threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,03	1,00	0,45	0,46
0,20	0,10	1,00	0,47	0,50
0,30	0,17	1,00	0,49	0,53
0,40	0,24	0,99	0,51	0,57
0,50	0,34	0,98	0,54	0,63
0,60	0,45	0,97	0,58	0,68
0,70	0,55	0,94	0,63	0,73
0,80	0,64	0,88	0,66	**0,75**
0,90	0,73	0,73	0,68	0,73
1,00	1,00	0,00	1,00	0,56

Figura 4-30.: Resultados de bosque aleatorio modelo generico con dataset marca C

4.3.4. Red Neuronal modelo genérico

Finalmente el ultimo modelo a implementar es una red neuronal multicapa se implementa con las mismas especificaciones descritas para los modelos anteriores descritas en: **4-14**, con una salida sigmoide.

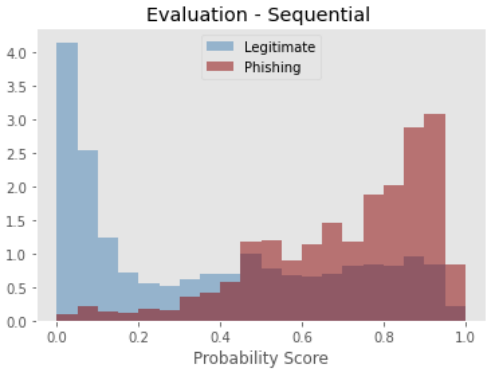
4.3.4.1. Evaluación modelo genérico con dataset propio



threshold	specificity	recall	precision	accuracy
0,0	0,00	1,00	0,49	0,49
0,1	0,42	0,99	0,63	0,71
0,2	0,49	0,98	0,66	0,74
0,3	0,54	0,98	0,68	0,76
0,4	0,62	0,96	0,71	0,79
0,5	0,73	0,91	0,77	0,82
0,6	0,81	0,85	0,81	**0,83**
0,7	0,87	0,76	0,85	0,82
0,8	0,92	0,63	0,90	0,78
0,9	0,97	0,37	0,94	0,67
1,0	1,00	0,00	1,00	0,50

Figura 4-31.: Resultados de red neuronal modelo genérico

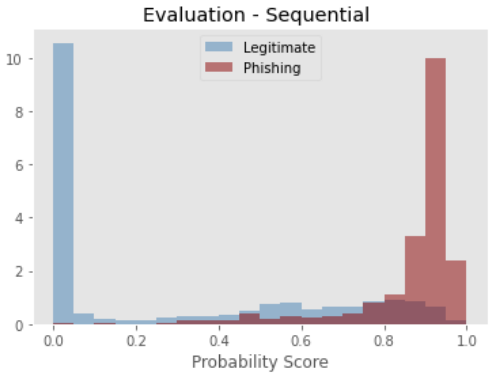
4.3.4.2. Evaluación modelo genérico con dataset marca A



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,09	0,09
0,10	0,33	0,98	0,13	0,40
0,20	0,43	0,97	0,15	0,48
0,30	0,49	0,96	0,16	0,53
0,40	0,55	0,92	0,17	0,58
0,50	0,64	0,83	0,19	0,65
0,60	0,71	0,72	0,20	0,71
0,70	0,78	0,59	0,22	0,76
0,80	0,86	0,44	0,24	0,82
0,90	0,95	0,20	0,28	0,88
1,00	1,00	0,00	1,00	**0,91**

Figura 4-32.: Resultados de red neuronal modelo genérico dataset marca A

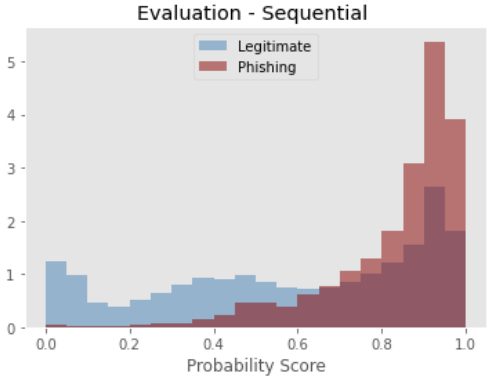
4.3.4.3. Evaluacion modelo genérico con dataset marca B



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,05	0,05
0,10	0,55	1,00	0,10	0,57
0,20	0,56	1,00	0,11	0,59
0,30	0,58	0,99	0,11	0,60
0,40	0,61	0,98	0,12	0,63
0,50	0,66	0,95	0,13	0,67
0,60	0,73	0,93	0,15	0,74
0,70	0,79	0,90	0,18	0,80
0,80	0,87	0,84	0,25	0,87
0,90	0,96	0,62	0,44	0,94
1,00	1,00	0,00	1,00	**0,95**

Figura 4-33.: Resultados de red neuronal modelo genérico dataset marca B

4.3.4.4. Evaluación modelo genérico con dataset marca C



threshold	specificity	recall	precision	accuracy
0,00	0,00	1,00	0,44	0,44
0,10	0,11	1,00	0,47	0,50
0,20	0,15	0,99	0,48	0,52
0,30	0,21	0,99	0,50	0,55
0,40	0,30	0,97	0,52	0,60
0,50	0,39	0,94	0,55	0,63
0,60	0,47	0,90	0,57	0,66
0,70	0,55	0,83	0,59	**0,67**
0,80	0,64	0,71	0,61	0,67
0,90	0,78	0,46	0,62	0,64
1,00	1,00	0,00	1,00	0,56

Figura 4-34.: Resultados de red neuronal modelo genérico dataset marca C

5. Análisis de resultados

5.1. Efectividad

Teniendo los modelos 4 modelos con sus respectivas métricas se obtiene la tabla mostrada en **5-1** con la mejor configuración en efectividad de cada modelo. Teniendo en cuenta que son dataset desbalanceados, en la escogencia del mejor modelo para cada marca se deben tener en cuenta parámetros como recall, especificidad y precisión. Excluyendo métricas de efectividad altas como 4.33(a), 4.32(a),4.24(a), 4.23(a), 4.20(a), 4.19(a) y 4.8(a) donde en su mayoría se trata de modelos genéricos y relacionados a las marcas A y B que son los dataset mayormente desbalanceados. Se observa refiriéndose a efectividad en detección los modelos A Y B enfocados a una marca especifica muestran un 5 % y 6 % de mayor efectividad que el modelo genérico.

Tabla 5-1.: Tabla comparativa de efectividad en detección

Marca	Regresion	SVM	Bosque aleatorio(BA)	Red Neuronal(RN)	Mejor modelo
Marca A	0,91	0,91	**0,95**	0,93	Bosque Aleatorio
Marca B	0,91	0,96	0,96	**0,96**	Red Neuronal
Marca C	0,7	0,71	0,66	**0,71**	Red Neuronal
Genérico	0,71	0,75	**0,9**	0,83	Bosque aleatorio

Si se detalla cada marca y sus resultados específicos se puede sugerir por ejemplo que para la marca A y B existe una facilidad en identificar patrones asociados a la marca en la propia URL. Ambos modelos muestran efectividad alta en todos sus modelos siempre estando por encima incluso de mejor modelo genérico. Esto puede sugerir que encontrar características propias a la propia marca en sus ataques de phishing es sencillo y se pueden explorar otras características asociadas por ejemplo al HTML para aumentar efectividad.

En cuanto al modelo C si bien presenta baja efectividad en los modelos planteados con estas características, puede ser debido a que solo se están tomando características referidas a la URL; pueden explorarse otro tipo de características asociadas al HTML o WHOIS para encontrar patrones relativos a la marca. Ya que incluso el modelo genérico presenta mayor efectividad con el modelo bosque aleatorio con la marca C.

5.2. Características seleccionadas

Teniendo en cuenta las tablas **4-2**, **4-3**, **4-4** y **4-5** se observa un común denominador asociado a una característica que muestra relevancia en la identificación del phishing, esta es la longitud de la URL y la cantidad de puntos, que a la vez está asociado a la cantidad de subdominios asociados a un ataque de phishing, esto se logra explicar desde la existencia de dominios comúnmente abusados donde el ataque no se encuentra en el dominio, sino que se le asocia un subdominio. Siguiente a estas dos características se encuentran las relacionadas con caracteres posiblemente relacionados a la marca afectada. En **5-2** para características genéricas y **5-3** para caracteres, se utiliza información tomada del modelo Bosque Aleatorio para identificar las características de mayor influencia en el clasificador.

Tabla 5-2.: Tabla comparativa de efectividad en detección

Marca	Característica genéricas más influyentes	Porcentaje de influencia
Marca A	longitud	9,1 %
	puntos	7,1 %
	abusedtld	2,8 %
	abusedhost	2 %
	caracteres especiales	1,7 %
Marca B	longitud	12,6 %
	puntos	6,8 %
	abusedhost	2,6 %
	abusedtld	1 %
Marca C	longitud	8,9 %
	puntos	4,2 %
	abusedhost	1,4 %
Genérico	longitud	10,4 %
	puntos	6,7 %
	abusedhost	3,8 %
,	caracteres especiales	5,2 %

de la tabla **5-3** se puede notar que cada marca tiene asociada una combinación única de caracteres, si bien comparten algunos como s, m y a, tienen otros caracteres que pueden sugerir patrones en las URLs de los ataques de phishing de su propia marca. Hay una característica genérica que vale la pena mencionar porque se encuentra en la marca A y B en mayor medida que en la marca C o en el modelo genérico y es el TLD comúnmente abusado, esto puede tener una explicación de geolocalización que nos sirve para entender que para las marcas A y B se suelen incluir patrones en el TLD propios del país de la marca.

Tabla 5-3.: Tabla comparativa de efectividad en detección

Marca	Caracteres con mayor influencia	Porcentaje de influencia
Marca A	s	4,8 %
	a	4,7 %
	p	4,5 %
	r	4,1 %
	e	4,1 %
	c	4,1 %
Marca B	m	6.6 %
	p	5,6 %
	w	4,8 %
	l	4,4 %
	e	4,4 %
	t	4,3 %
Marca C	s	7 %
	c	5 %
	o	4,8 %
	i	4 %
Genérico	a	4,6 %
	m	4,5 %
	o	4,4 %
	c	4 %

Adicionalmente la gráfica de correlación **3-3**, **3-4** y **3-5** dan muestra de que existe un patrón de marca asociado a la caracterización por caracteres. Siendo la marca C la que muestra una fuerte correlación con las características genéricas, lo que puede explicar su semejanza con la gráfica de **4-27**, que reitera la tendencia de las URLs asociadas a C a no presentar patrones de marca en la URL.

5.3. Modelo de detección de phishing en etapas de prevención

Teniendo en cuenta los resultados obtenidos y el proceso realizado para cada una de las marcas se procede a consolidar la propuesta final del modelo **4-1**, el cual se compone de tres partes principales: la primera referente a la selección de fuentes, como se describió en la sección **2.2.6**, permite realizar la selección de fuentes para enfocarse en etapa de prevención, sin embargo el modelo propuesto permite vincular tantas fuentes en diferentes etapas como

sea necesario teniendo particular cuidado de incluir dominios o subdominios recientemente registrados.

Posteriormente a través del análisis de marca se hará uso del modelo heurístico para filtrar las URLs que provienen de las fuentes añadidas al modelo, usando tres fuentes de datos principales: Palabras claves, dominios comúnmente abusados y TLDs abusados. Este filtrado además de proporcionar entrada a nuestro modelo entrenado para las marcas seleccionadas, también funciona como primer filtro reduciendo las URLs de entrada a URLs de interés particular para la marca afectada.

Finalmente se sugiere el uso del clasificador bosque aleatorio para la marca A y de red neuronal para B y C basado en su efectividad. Se sugiere aumentar las características asociadas a la marca para la marca C para aumentar su efectividad, considerando no solo la URL sino también información relacionada al contenido de la pagina web he incluso información relacionada a su registro (WHOIS).

6. Conclusiones y recomendaciones

Se analizaron y recolectaron URLs phishing de diferentes fuentes de información como o phishtank, openphish, y cuentas de twitter, identificando la marca afectada y mediante el análisis de palabras clave, se buscaron otras URLs no phishing logrando construir datasets específicos para cada marca mediante buscadores y otros proveedores como **2-1** y **2-2**. Esto solo fue posible conociendo información de las palabras claves de la marca que permitiera relacionar las URLs obtenidas en la búsqueda a una marca. Hay que mencionar que este mismo proceso se propone aplicarse en etapas de detección tempranas para la búsqueda de posibles phishing en dominios recientemente registrados asociados a la marca por palabras clave. Posterior a esto se realizó la ingeniería de características para determinar qué características serían seleccionadas para utilizarse en los diferentes modelos identificando en el camino patrones en TLDs y dominios que junto a las palabras clave sirven de suministro para aplicar a modelos heurísticos.

Se diseñó un modelo de detección de phishing que pudiera aprovechar el proceso anteriormente descrito y que mediante el ingreso de URLs sospechosas pudiera ser asociado a la marca mediante 3 fuentes de datos recolectadas previamente: TLDs comúnmente abusados, Palabras clave asociadas a la marca y dominios abusados. Posteriormente a el modelo heurístico se adjuntaba un clasificador para el cual se entrenaron 4 modelos para cada una de las marcas A,B y C entrenados con su dataset específico y adicionalmente se entrenaron 4 modelos con un dataset genérico es decir sin considerar agrupar las URLs por marca. Esto permitió escoger el mejor modelo para cada una de las marcas y comparar su comportamiento frente a las otras.

Se evaluaron los modelos de clasificación y se escogió el mejor modelo clasificador para cada marca así completando la implementación del modelo completo. En la escogencia del mejor modelo para cada marca se tuvo en cuenta que se trataba con dataset desbalanceados por lo que métricas como la especificidad , el recall y la precisión fueron tomados en cuenta a la hora de esta selección en **5-1**.

De los resultados **5-1** de los 3 modelos se puede observar que las URL no maliciosas obtenidas a partir de palabras clave son identificables por los dos primeros modelos entrenados para identificar a la marca A y B, sin embargo en el tercero marca C, existe una efectividad baja comparada con los otros dos modelos, lo que puede relacionarse a una mayor entropía

dentro de las URLs asociadas a esa marca.

Al revisar la literatura, aunque hay buenos resultados para la "verificación"de phishing a partir de muestras de URL, con una precisión de detección de más del 90 %, en la mayoría de los casos, todavía hay muchos desafíos, y teniendo en cuenta resultados como los mostrados en **5-1**, se recomienda probar este tipo de enfoques para lograr la acción integrada de tres fases descrita en este documento.

Si bien algunas aproximaciones pueden aplicarse a grupos de amenazas específicos, es bueno revisar los métodos utilizados para buscar de manera efectiva los beneficios que se aplican a la realidad. Hay problemas para proporcionar modelos que funcionen en una etapa temprana cuando el objetivo no es la mitigación sino la prevención. No hay forma de tomar medidas desde el propio registrador de dominios sin saber qué buscar en el dominio registrado o qué contenido del dominio ayuda a identificar posibles amenazas. Este tipo de aproximaciones permiten focalizar los esfuerzos para actuar en etapas tempranas mediante caracterización de los ataques de phishing de la marca y permite actuar desde la misma búsqueda de URLs a analizar.

Además, el modelo implementado necesita ser adaptable en el tiempo, aspecto que no se encuentra en los resultados de los trabajos relacionados. La industria debe entender que debe proteger los servicios prestados a los usuarios mediante la implementación de sistemas personalizados que aborden su propia funcionalidad. Por esta razón, se debe considerar un enfoque especial para evitar el phishing de las marcas afectadas. La identificación de características específicas de la marca ayuda a centrarse en la detección temprana. Es así como extraer características asociadas a la marca , permite que mientras la marca no cambie de nombre o no presente alteraciones en sus palabras clave perdure en el tiempo.

Los modelos entrenados funcionan para detectar URLs relacionadas a la marca a partir de palabras clave, lo que permite hacer una búsqueda proactiva de las URLs potencialmente peligrosas para esta. Adicionalmente presentan una alta efectividad en clasificación en 2 de los 3 modelos para esto se considera clave la asociación de caracteres propios de la marca y las características asociadas a dominios frecuentemente abusados y TLDs abusados. Finalmente, frente al modelo que presenta menor efectividad, la entropía propia de las URLs que afectan a esta marca en específico limita la asociación de la marca a las características extraídas, por lo que buscar otro tipo de características que puedan asociarse a la marca, son necesarias para aumentar su efectividad en detección.

A. Anexo:A Analisis correlación de carcateristicas extendida

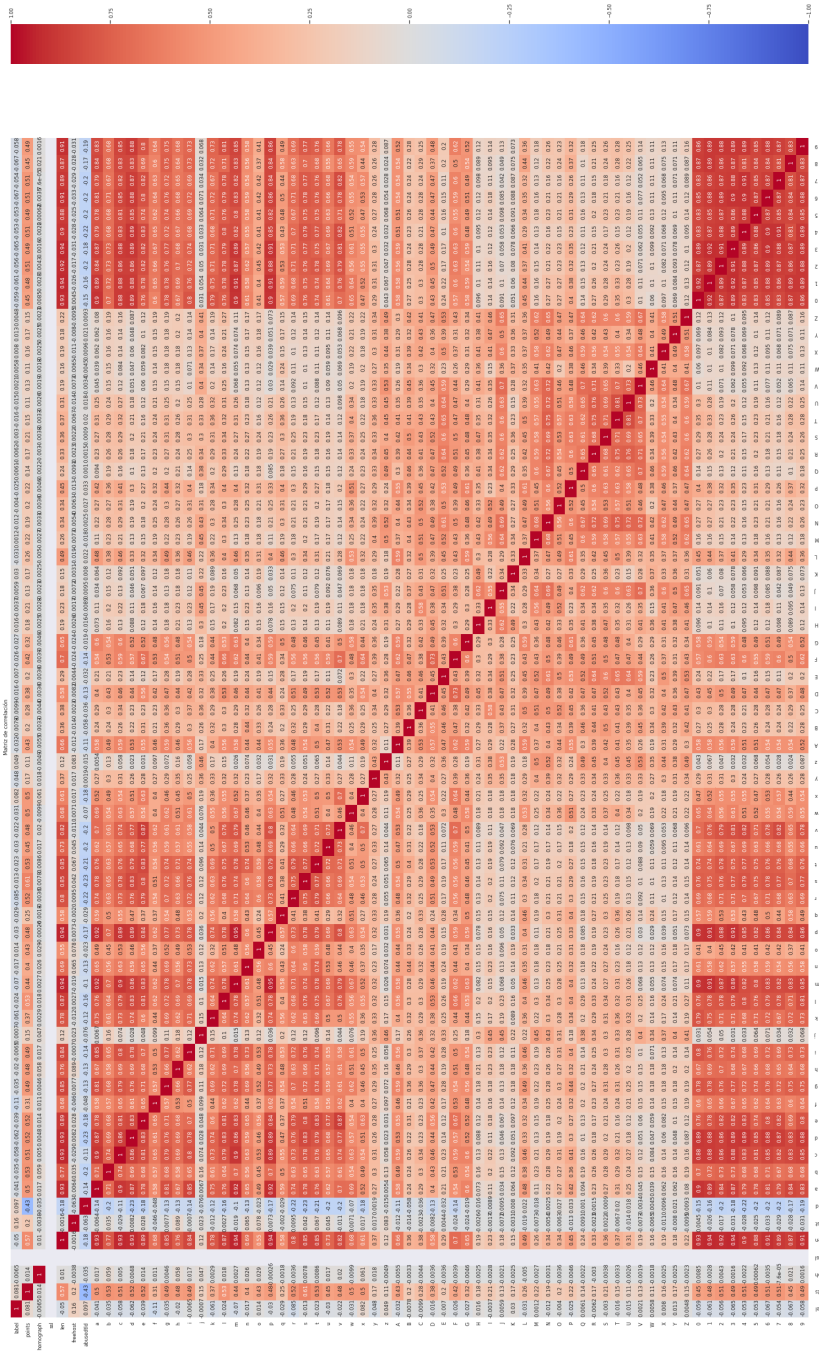


Figura A-1.: Analisis correlativo de carcateristicas extendido

Bibliografía

[A, 2020] A, A. A. (2020). *Towards the Detection of Phishing Attacks Praveen K TIFAC-CORE in Cyber Security Amrita Vishwa Vidyapeetham.*

[Adil et al., 2020] Adil, M., Khan, R., and Ghani, M. A. N. U. (2020). Preventive Techniques of Phishing Attacks in Networks. In *2020 3rd International Conference on Advancements in Computational Sciences (ICACS)*, pages 1–8.

[Ali and Ahmed, 2019] Ali, W. and Ahmed, A. A. (2019). Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting. *IET Information Security*, 13(6):659–669.

[Anand et al., 2018] Anand, A., Gorde, K., Moniz, J. R. A., Park, N., Chakraborty, T., and Chu, B. (2018). Phishing URL Detection with Oversampling based on Text Generative Adversarial Networks. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 1168–1177.

[apwg, 2022] apwg (2022). PHISHING ACTIVITY TRENDS REPORT Q4 2021.

[Aung and Yamana, 2019] Aung, E. S. and Yamana, H. (2019). URL-Based Phishing Detection Using the Entropy of Non-Alphanumeric Characters. In *Proceedings of the 21st International Conference on Information Integration and Web-Based Applications & Services*, iiWAS2019, page 385–392, New York, NY, USA. Association for Computing Machinery.

[Baig et al., 2021] Baig, M. S., Ahmed, F., and Memon, A. M. (2021). Spear-phishing campaigns: Link vulnerability leads to phishing attacks, spear-phishing electronic/uav communication-scam targeted. In *2021 4th International Conference on Computing Information Sciences (ICCIS)*, pages 1–6.

[Balim and Gunal, 2019] Balim, C. and Gunal, E. S. (2019). Automatic Detection of Smishing Attacks by Machine Learning Methods. In *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, pages 1–3.

[Barreiro and Camargo, 2022] Barreiro, D. A. and Camargo, J. E. (2022). A systematic review on phishing detection: A perspective beyond a high accuracy in phishing detection. pages 173–188.

[Baykara and Gürel, 2018] Baykara, M. and Gürel, Z. Z. (2018). Detection of phishing attacks. In *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*, pages 1–5.

[Buber et al., 2017] Buber, E., Demir, , and Sahingoz, O. K. (2017). Feature selections for the machine learning based detection of phishing websites. In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, pages 1–5.

[Concone et al., 2019] Concone, F., Re, G. L., Morana, M., and Ruocco, C. (2019). Assisted Labeling for Spam Account Detection on Twitter. In *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 359–366.

[Dalgic et al., 2018] Dalgic, F. C., Bozkir, A. S., and Aydos, M. (2018). Phish-IRIS: A New Approach for Vision Based Brand Prediction of Phishing Web Pages via Compact Visual Descriptors. In *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pages 1–8.

[Das et al., 2020] Das, A., Baki, S., Aassal, A. E., Verma, R., and Dunbar, A. (2020). SoK: A Comprehensive Reexamination of Phishing Research From the Security Perspective. *IEEE Communications Surveys & Tutorials*, 22(1):671–708.

[DomainWatch,] DomainWatch. DomainWatch - Domain WHOIS Search, Website Information.

[Eshmawi and Nair, 2019] Eshmawi, A. and Nair, S. (2019). The Roving Proxy Framewrok for SMS Spam and Phishing Detection. In *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, pages 1–6.

[Ginsberg and Yu, 2018] Ginsberg, A. and Yu, C. (2018). Rapid Homoglyph Prediction and Detection. In *2018 1st International Conference on Data Intelligence and Security (ICDIS)*, pages 17–23.

[Huang et al., 2019] Huang, Y., Qin, J., and Wen, W. (2019). Phishing URL Detection Via Capsule-Based Neural Network. In *2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, pages 22–26.

[JAMES, 2005] JAMES, L. (2005). *Phishing Exposed*.

[Li and Wang, 2017] Li, J. and Wang, S. (2017). PhishBox: An Approach for Phishing Validation and Detection. In *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, pages 557–564.

[Li et al., 2020] Li, Q., Cheng, M., Wang, J., and Sun, B. (2020). LSTM based Phishing Detection for Big Email Data. *IEEE Transactions on Big Data*, page 1.

[Li et al., 2016] Li, X., Geng, G., Yan, Z., Chen, Y., and Lee, X. (2016). Phishing detection based on newly registered domains. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 3685–3692.

[Lingam et al., 2018] Lingam, G., Rout, R. R., and Somayajulu, D. V. L. N. (2018). Detection of Social Botnet using a Trust Model based on Spam Content in Twitter Network. In *2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*, pages 280–285.

[Lingam et al., 2019] Lingam, G., Rout, R. R., and Somayajulu, D. V. L. N. (2019). Deep Q-Learning and Particle Swarm Optimization for Bot Detection in Online Social Networks. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–6.

[McGahagan et al., 2019] McGahagan, J., Bhansali, D., Gratian, M., and Cukier, M. (2019). A Comprehensive Evaluation of HTTP Header Features for Detecting Malicious Websites. In *2019 15th European Dependable Computing Conference (EDCC)*, pages 75–82.

[Megha et al., 2019] Megha, N., Babu, K. R. R., and Sherly, E. (2019). An Intelligent System for Phishing Attack Detection and Prevention. In *2019 International Conference on Communication and Electronics Systems (ICCES)*, pages 1577–1582.

[Mondal et al., 2019] Mondal, S., Maheshwari, D., Pai, N., and Biwalkar, A. (2019). A Review on Detecting Phishing URLs using Clustering Algorithms. In *2019 International Conference on Advances in Computing, Communication and Control (ICAC3)*, pages 1–6.

[Nakamura and Dobashi, 2019] Nakamura, A. and Dobashi, F. (2019). Proactive Phishing Sites Detection. In *IEEE/WIC/ACM International Conference on Web Intelligence*, WI '19, page 443–448, New York, NY, USA. Association for Computing Machinery.

[Nathezhtha et al., 2019] Nathezhtha, T., Sangeetha, D., and Vaidehi, V. (2019). WC-PAD: Web Crawling based Phishing Attack Detection. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–6.

[Pande and Voditel, 2017] Pande, D. N. and Voditel, P. S. (2017). Spear phishing: Diagnosing attack paradigm. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 2720–2724.

[Patil et al., 2018] Patil, V., Thakkar, P., Shah, C., Bhat, T., and Godse, S. P. (2018). Detection and Prevention of Phishing Websites Using Machine Learning Approach. In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pages 1–5.

[Sahoo, 2018] Sahoo, P. K. (2018). Data mining a way to solve Phishing Attacks. In *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, pages 1–5.

[Sharma et al., 2017] Sharma, H., Meenakshi, E., and Bhatia, S. K. (2017). A comparative analysis and awareness survey of phishing detection tools. In *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pages 1437–1442.

[Spaulding et al., 2016] Spaulding, J., Upadhyaya, S., and Mohaisen, A. (2016). The Landscape of Domain Name Typosquatting: Techniques and Countermeasures. In *2016 11th International Conference on Availability, Reliability and Security (ARES)*, pages 284–289.

[Starov et al., 2019] Starov, O., Zhou, Y., and Wang, J. (2019). Detecting Malicious Campaigns in Obfuscated JavaScript with Scalable Behavioral Analysis. In *2019 IEEE Security and Privacy Workshops (SPW)*, pages 218–223.

[urlscan,] urlscan. URL and website scanner.

[Xiang et al., 2011] Xiang, G., Hong, J., Rose, C. P., and Cranor, L. (2011). CANTINA+: A Feature-rich Machine Learning Framework for Detecting Phishing Web Sites.

[Ya et al., 2019] Ya, J., Liu, T., Zhang, P., Shi, J., Guo, L., and Gu, Z. (2019). NeuralAS: Deep Word-Based Spoofed URLs Detection Against Strong Similar Samples. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7.

[Yan et al., 2020] Yan, X., Xu, Y., Xing, X., Cui, B., Guo, Z., and Guo, T. (2020). Trustworthy Network Anomaly Detection Based on an Adaptive Learning Rate and Momentum in IIoT. *IEEE Transactions on Industrial Informatics*, page 1.

[Yang et al., 2019] Yang, P., Zhao, G., and Zeng, P. (2019). Phishing Website Detection Based on Multidimensional Features Driven by Deep Learning. *IEEE Access*, 7:15196–15209.

[Yao et al., 2018] Yao, W., Ding, Y., and Li, X. (2018). LogoPhish: A New Two-Dimensional Code Phishing Attack Detection Method. In *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*, pages 231–236.

[Yazhmozhi and Janet, 2019] Yazhmozhi, V. M. and Janet, B. (2019). Natural language processing and Machine learning based phishing website detection system. In *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pages 336–340.

[Yuan et al., 2018] Yuan, H., Chen, X., Li, Y., Yang, Z., and Liu, W. (2018). Detecting Phishing Websites and Targets Based on URLs and Webpage Links. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 3669–3674.

[Zhu et al., 2018] Zhu, E., Ye, C., Liu, D., Liu, F., Wang, F., and Li, X. (2018). An Effective Neural Network Phishing Detection Model Based on Optimal Feature Selection. In *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*, pages 781–787.

[Zuraiq and Alkasassbeh, 2019] Zuraiq, A. A. and Alkasassbeh, M. (2019). Review: Phishing Detection Approaches. In *2019 2nd International Conference on new Trends in Computing Sciences (ICTCS)*, pages 1–6.