Tipo de artículo: Artículo original

Detección de amenazas de seguridad en una red corporativa utilizando algoritmos de machine learning

Detection of security threats in a corporate network using machine learning algorithms

Alfonso A. Guijarro Rodríguez[1] *, https://orcid.org/0000-0001-6046-426X
Gladys C. Jácome-Morales[2] https://orcid.org/0000-0003-1922-7988
Viviana Gonzalez-Mestanza[3], https://orcid.org/0000-0001-5428-7025
Elein Terán-Zurita[4], https://orcid.org/0000-0002-4072-3452
Dennisse E. Torres-Martínez[5], https://orcid.org/0000-0001-7756-7943

[1] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, alfonso.guijarror@ug.edu.ec
[2] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Gladys.jacomem@ug.edu.ec
[3] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, viviana.gonzalezm@ug.edu.ec
[4] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, elein.teranz@ug.edu.ec
[5] Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, dennisse.torresm@ug.edu.ec

*** Autor para correspondencia: alfonso.guijarror@ug.edu.ec**

Resumen

El aumento de los accesos no autorizados que presentan los sistemas informáticos es generado por programas maliciosos y en los últimos años en Ecuador, se han presentado violaciones a la seguridad de las redes corporativas, siendo el ciberataque más común ransomware, por lo que se plantea mejoras a las prácticas de seguridad aplicadas a las redes empleando algoritmos de inteligencia artificial para predecir y mitigar estas amenazas. En este estudio se propone diseñar un modelo de seguridad para detectar las amenazas en una red corporativa con el uso de algoritmos de machine learning. Para esto se realizó una revisión de la literatura en revistas científicas y libros para detallar los conceptos relacionados a la ciberseguridad en ambientes empresariales, además se utilizó la metodología CRISP-DM para la elaboración del modelo computacional, el cual consta de cinco fases, comprensión del negocio, comprensión de datos, preparación de datos, modelado y evaluación. Al finalizar el trabajo el modelo considera que el algoritmo de Random Forest presenta un mayor porcentaje en precisión comparado con otros algoritmos planteados en este estudio, logrando un valor de rendimiento del 99,74%. Por lo cual se concluye que es factible detectar anomalías de seguridad del tráfico de una red corporativa con el uso de algoritmos de machine learning.

Palabras clave: algoritmos, aprendizaje automático, CRISP-DM, detección, tráfico de red.

Abstract

The increase in unauthorized access to computer systems is generated by malicious programs and in recent years in Ecuador, there have been violations to the security of corporate networks, being the most common cyberattack ransomware, so it is proposed improvements to security practices applied to networks using artificial intelligence algorithms to predict and mitigate these threats. This study proposes to design a security model to detect threats in a corporate network with the use of machine learning algorithms. For this, a literature review was conducted in scientific journals and books to detail the concepts related to cybersecurity in enterprise environments, in addition, the CRISP-DM methodology was used to develop the computational model, which consists of five phases, business understanding, data understanding, data preparation, modeling and evaluation. At the end of the work, the

model considers that the Random Forest algorithm presents a higher percentage in accuracy compared to other algorithms proposed in this study, achieving a performance value of 99.74%. Therefore, it is concluded that it is feasible to detect security anomalies in the traffic of a corporate network with the use of machine learning algorithms.

Keywords: *algorithms, machine learning, CRISP-DM, detection, network traffic.*

Recibido: 25/08/2022
Aceptado: 10/12/2022
En línea: 12/12/2022

Introducción

La evolución de la tecnología y el desarrollo del internet ha permitido a las organizaciones optimizar sus recursos y a las personas a establecer una mejor comunicación, con el paso de los años han surgido también nuevas técnicas de inteligencia artificial que permiten mejorar los servicios que ofrecen las organizaciones a usuarios. Sin embargo, el mal uso del internet ha provocado que usuarios ejecuten ataques a sistemas informáticos con programas maliciosos; como malware, phishing o ransomware, que han aumentado significativamente en los últimos años.

Según los datos de la compañía de software ESET, en el año 2020 en Ecuador se registraron más de 51 mil ataques por cryptominers (programa malicioso para la minería de criptomonedas), 140 mil de exploits (aprovecha vulnerabilidades en el sistema), 6 mil detecciones de ransomware y casi 8 mil detecciones de spyware. En ese mismo año, Ecuador ocupó el sexto lugar dentro de los países latinoamericanos con más detecciones de programas maliciosos y sétima posición en detección de phishing, debido a la falta de personal capacitado y la mínima cantidad de programadores expertos en seguridad informática (Abril, 2021).

A raíz de confinamiento social en los años 2020-2021, en las organizaciones aumentó el uso de la tecnología, para realizar las actividades laborales desde casa. Esto benefició la continuidad de los procesos en las organizaciones, sin embargo, se pudo observar un incremento exponencial de los ataques informáticos con referencia a los años anteriores. Durante el año 2021 en Ecuador se registró un aumento del 75% de ciberataques, liderando la lista de los países latinoamericanos con mayores ataques, con un promedio de 89 ataques por minuto (Diazgranados, 2021).

Para el desarrollo de esta investigación el enfoque se realizó en la Carrera de Software, adscrita a la Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil, ya que, siendo una institución de educación superior, no está exenta de ser afectada por ataques informáticos, por lo que tenemos como objetivo el diseño de un modelo de seguridad informática mediante el uso de algoritmos de machine learning, para la detección de amenazas en el tráfico de red.

En el trabajo realizado por Recordon y Ruiz (2020) evalúan la efectividad de técnicas relacionadas al machine learning y data mining para la detección y clasificación de Zero-Day Malware, la metodología utilizada constó de 4 fases;

obtención de los datos, construcción de un dataset, depuración del dataset y finalmente construcción y ajuste de los modelos de clasificación. Una vez elaborados los modelos de machine learning para clasificar y poner a prueba la eficacia de los datos, se obtuvo un porcentaje de 98% de precisión para la detección de malware para los algoritmos de K-Nearest Neighbors, Random Forest y XGboots, mientras que el de Artificial Neural Network (ANN) obtuvo una precisión del 99%. En esta investigación se utilizaron las técnicas mencionadas para identificar el funcionamiento de los algoritmos en la detección de malwares.

De igual forma, Estévez (2020) en su trabajo de titulación, como tema de estudio realizó una comparación de diferentes modelos de machine learning, pero con un dataset etiquetado enfocado en el tráfico de una red corporativa para detectar anomalías; una vez elegido el conjunto de datos, combinó los parámetros correspondientes de cada modelo para optimizar el rendimiento de los algoritmos, compararlos entre sí y seleccionando el de mejor resultado. Utilizó la metodología, aplicada también este trabajo de investigación, CRISP-DM, la cual consiste en seis fases; comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Fue útil dado que para la implementación del modelo de detección mencionado en el artículo se siguió la metodología CRISP-DM.

Otra investigación, con respecto al uso de algoritmos de machine learning Vera (2020), utiliza estos algoritmos para detectar anomalías en la red LAN de la empresa NewOffice, recolectó información mediante el sistema de detección de intrusos (IDS), aplicando una modalidad bibliográfica para su desarrollo y el tipo de investigación descriptiva para establecer la base de análisis de los datos. Obtenidos los resultados, concluyó que el algoritmo implementado realiza dos tipos de predicciones con respecto a los datos obtenidos del IDS, y la herramienta Suricata, que la utilizó para mostrar cuando y como se produce un ataque, tuvo mayor precisión con sus datos debido a que presentó menor número de clúster. En este artículo se utilizó la investigación antes mencionada para determinar el funcionamiento de un modelo de seguridad y clasificación de las anomalías.

Para una mejor comprensión de este trabajo se lo detalla las secciones en forma clásicas, materiales y métodos presenta la metodología CRISP-DM con sus cinco fases comprensión del negocio, comprensión de los datos, preparación de los datos. modelado y evaluación, la sección resultados muestra los valores de rendimiento generados por los algoritmos de machine Learning para el modelo computacional y finalmente se presentan las conclusiones del trabajo y las contribuciones de los autores y financiamiento derivadas de este trabajo.

Materiales y métodos

Para la elaboración del modelo de seguridad se utilizó la metodología CRISP-DM (ver figura 1) que se enfoca principalmente en la minería de datos y está formada por 5 fases, comprensión del negocio, comprensión de los datos, preparación de los datos. modelado y evaluación. Cabe indicar que las herramientas utilizadas en esta metodología fueron: Wireshark para capturar el tráfico de red, la página web de Malware-Traffic-Analysis para la obtención de un dataset con malware, la página web de Virus total para comprobar si los datasets tienen archivos o direcciones maliciosas. Por otro lado, el escenario para la construcción del modelo fue la interfaz gráfica de Jupyter Notebook de Anaconada que mediante el lenguaje de programación de Python permitió la codificación, entrenamiento y prueba de los algoritmos, las principales librerías utilizadas en Python fueron matplotlib para generar gráficos, numpy y panda para el tratamiento de los datos y sklearn para el procesado de los algoritmos.

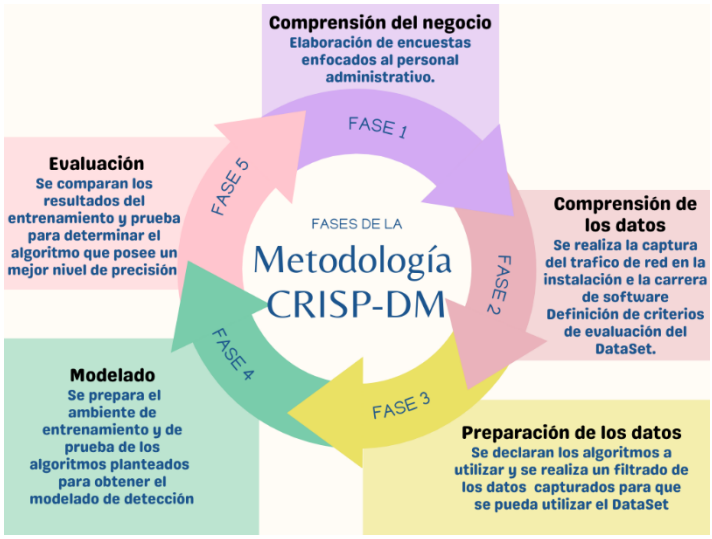

Figura 1. Fases de la metodología CRISP-DM

A continuación, se detallan las fases:

Fase de comprensión de negocio

Para esta fase de la metodología se llevó a cabo la técnica de la encuesta al personal administrativo de la carrera de software con el propósito de comprobar qué tan familiarizados están con el tema de seguridad informática. Además, se realizó la petición para hacer la captura del tráfico de red, al encargado de departamento de sistemas en las instalaciones de la carrera de software. Antes de pasar la siguiente fase fue necesario conocer como los paquetes transitaban en la

red, que campos se obtuvieron de una captura, como se identificó un malware mediante la herramienta de Wireshark y así poder evaluar si el comportamiento de tráfico de red fue normal o no.

Fase de comprensión de los datos

En esta fase se utilizó la herramienta de Wireshark para capturar el tráfico de red en las instalaciones de la carrera de Software de la Universidad de Guayaquil durante un periodo de tiempo de 20 minutos al azar durante una semana. Para poder utilizar los datos del tráfico de red en la fase de preparación se exportaron los datasets obtenidos en pcap de Wireshark en un archivo con extensión (.csv) para su manipulación. Además, se descargó un datasets con malwares para analizar el comportamiento en el tráfico de red capturado.

Fase de preparación de los datos

A la data obtenida se le realiza un procesado, limpieza y extracción en base a las técnicas de minería de datos. Se realiza un filtrado de los datasets por el protocolo TCP. Luego se utilizó la herramienta Anaconda con la interfaz de Jupyter Notebook y el lenguaje de programación de Python para implementar la codificación de los algoritmos. En Python se realizó una limpieza de los datos, una exclusión de los campos que contengan cadena de caracteres y convertimos las direcciones IP a datos numéricos. Además, se realizó una categorización de datos aplicando el método *predict()* del algoritmo Isolation Forest que tiene como objetivo identificar anomalías (*outliers*), este algoritmo clasifica los paquetes del tráfico de red con (1) si los datos son normales y con (-1) los datos anómalos. Por último, se realiza un análisis del comportamiento de un malware y se comprobó que al procesar el dataset en el algoritmo de Isolation Forest sí identifica paquetes con direcciones o archivos maliciosos, mientras que, con el dataset de las capturas de red realizadas, si bien no posee un malware, categorizó como anómalos los datos que presentan inconsistencia en el flujo de la red. Debido a los pocos datos categorizados como anómalos, se realizó una unión entre los datos capturados de la red con un dataset de malware para evitar un desbalance. Por consiguiente, obtener un aprendizaje equitativo entre los datos al momento de implementar el entrenamiento de los algoritmos de clasificación.

Fase de modelado

Para la fase del modelado se utilizó la librería Sklearn de Python que sirve para el preprocesamiento, modelado y clasificación de algoritmos de inteligencia artificial. En trabajo de investigación se seleccionaron cuatro algoritmos: Random Forest, Decision Tree, Logistic Regression y SVM, de acuerdo con la revisión bibliografía realizada. Además,

se elaboró una división de los datos en 70% para el entrenamiento y 30% para la prueba mediante el método *train_test_split*.

Fase de evaluación

En esta fase de la investigación se evaluó la eficiencia al detectar anomalías en el tráfico de red de la carrera de Software con los cuatros modelos seleccionados, cada uno de los algoritmos fue entrenado y evaluado 5 veces para una mayor precisión. En la fase anterior los datos fueron categorizados y divididos para el entrenamiento y prueba, por ende, en esta fase se realizó una tabla comparativa con los resultados obtenidos.

Resultados

En este apartado se detallan los resultados de las métricas de precisión, recall y f1 score por medio de un reporte de clasificación y de la misma forma el porcentaje de evaluación al ejecutar el modelado de cada algoritmo estudiado, obtenidos de la fase de evaluación de la metodología CRISP-DM.

Uno de los algoritmos que se empleó en el modelo de detección fue el de Random Forest (ver figura 2), que forma parte de la clasificación de algoritmos de aprendizaje supervisado basándose en la creación de árboles de decisión para la predicción; a continuación, se muestra la codificación en Python.

```
#Creación del modelo por random forest
Forestmodel = RandomForestClassifier()

#Ajusta el modelo para el entrenamiento
ran_forst = Forestmodel.fit(X_train, y_train)

#TEST DE PREDICCIÓN
y_pred = ran_forst.predict(X_test)

#MATRIZ DE CONFUSIÓN
result = confusion_matrix(y_test, y_pred)

print('\nMATRIZ DE CONFUSIÓN:')
plt.figure(figsize= (6,4))
sns.heatmap(result, annot = True,fmt='d',cmap="YlGnBu")
```

Figura 2. Algoritmo de Random Forest

Para el funcionamiento del modelado se importa la librería del algoritmo para su creación, se ajusta el modelo para el entrenamiento definiendo sus variables X y Y, luego se realiza la predicción y por último se genera una matriz de confusión. Este proceso se realiza por cada algoritmo de clasificación con sus variables respectivas, y con los datos

obtenido de la matriz de confusión se realiza un reporte de clasificación donde se muestran los resultados de las métricas de cada algoritmo utilizado.

Luego, se elaboró una tabla comparativa para analizar y mostrar los resultados de las métricas en base a uno de las datasets obtenidos en la recolección de datos aplicados al modelo y poder determinar cuál de los algoritmos evaluados es la mejor para la predicción de anomalías en un tráfico de red.

Tabla 1. Tabla de resultados de las métricas de cada algoritmo

Algoritmo	**Precisión de evaluación (%) Captura 1**	**Precisión evaluación (%) Captura 2**	**Precisión evaluación (%) Captura 3**	**Precisión (%) Captura 4**	**Promedio (%)**
Random Forest	99,95%	99,28%	100%	99,75%	99,74%
Decision Tree	99,95%	99,28%	99,71%	100%	99,73%
Logistic Regresion	87,81%	99,04%	100%	99,51%	96,59%
SVM	89,94%	99,52%	100%	99,51%	97,24%

Como se muestra en la tabla 1 el dataset es evaluado por los 4 modelos de predicción, donde se puede observar que el algoritmo de Random forest y Decision Tree tienen un promedio de precisión de 99,5%, correspondiente a la métrica con la cual predice los datos correctamente, un recall con promedio de un 80% con el que identifica las anomalías, un f1 promedio de 87,5%. En el modelo de Logistic Regression los promedios resultantes de acurracy es 88%, precision es 88,5%, recall es de 85,5% y f1 de 87%. Mientras que en el último modelo de SVM, el resultado de acurracy es de 90%, con un promedio para métrica de precisión de un 89%, de recall con un 88,5% y un f1 con un 89%.

Tabla 2. Resultados de evaluación

Algoritmos	Random Forest acc = 99%			Decision Tree acc = 99%			Logistic Regression acc = 88%			SVM acc = 89%		
	p(%)	r(%)	f1(%)	p(%)	r(%)	f1(%)	p(%)	r(%)	f1(%)	p(%)	r(%)	f1(%)
Anómalos	100%	60%	75%	100%	60%	75%	91%	76%	83%	91%	82%	86%
No anómalos	99%	100%	100%	99%	100%	100%	86%	95%	91%	89%	95%	92%
Promedio	99,5%	80%	87,5%	99,5%	80%	87,5%	88,5%	85,5%	87%	89%	88,5%	89%

Como último punto se elaboró una comparación con los valores de la evaluación de cada algoritmo por cada captura realizada, para determinar cuál de estos tiene un mayor nivel de precisión, se calculó un promedio de evaluación por las cuatro capturas recolectadas en el periodo de 5 días de la semana. En la Tabla 2, se puede observar que el algoritmo con un mayor promedio de evaluación es el de Random Forest con un 99,74%, sin embargo, el promedio obtenido del Decisión tree no es tan alejado al que se menciona anteriormente ya que tiene un promedio del 99,73%. En contra parte, el algoritmo con el menor promedio fue el de Logistic regression con un valor de 96,59%. Por lo que se puede concluir que los mejores algoritmos para predecir una anomalía en el tráfico de red son los de Random Forest y Decision Tree.

Discusión

Con el propósito de diseñar un modelo de seguridad para la detección de amenazas en el tráfico de red aplicando algoritmos de machine learning, se demostró que con los algoritmos de aprendizaje supervisado de Random Forest y Decision Tree se obtuvo un valor de precisión del 99,74% y 99,73% respectivamente siendo estos los porcentajes de precisión más altos con respecto a los algoritmos de Support Vector Machine (97,24%) y Logistic Regression (96,59%). Los resultados mencionados guardan relación con lo obtenido por Estévez (2020) que al evaluar los algoritmos determinó que "RF es el algoritmo que mejores resultados obtuvo en la fase de modelado. Por ello, es el candidato a ser desplegado y distribuido" (pág. 72). Por otra la investigación de Vásquez (2021) se centró en evaluar los algoritmos supervisado y no supervisado para la detección de activades anómalas en el tráfico de red, indicado que los algoritmos como K-Means o Random Forest, dan resultados superiores al 90%, que demuestra la fiabilidad de los algoritmos.
Galmés (2019) describe que el aprendizaje automático o machine learning "se centra en el diseño de sistema, que permite aprender y hacer predicciones basadas en cierta experiencia, que serían los datos de entradas que recibe la máquina" (pág. 7).

Conclusiones

De acuerdo con la revisión bibliográfica realizada, se identificó que las amenazas más frecuentes en las organizaciones del Ecuador se deben a malwares como: ransomware, spyware y spam donde los atacantes provocan suspensión de servicios, fuga de información confidencial y en varios casos solicitan remuneración para recuperar los datos. También se identificó los algoritmos de machine learning, que son utilizados para la predicción de anomalías en un tráfico de red. Además, se seleccionaron los algoritmos de acuerdo con la revisión bibliográfica, que fueron los siguientes: Isolation Forest, Decision Tree, Logistic Regression y Support Vector Machine.

Se estableció el modelo de seguridad utilizando los algoritmos previamente mencionados y la metodología de desarrollo CRISP-DM, para su creación se utilizó la herramienta Anaconda con la interfaz de Jupyter Notebook, para obtener los dataset con el tráfico de red de la Carrera de Software se empleó el software de Wireshark.

Se evaluó el modelo de seguridad con los resultados obtenidos de la precisión de entrenamiento y prueba de cada algoritmo con el uso de la metodología CRISP-DM en la fase de evaluación. Se elaboró una tabla con los resultados de precisión de cada algoritmo, con él que se concluyó que, el de mejor porcentaje fue Random Forest con un 99,63%., esta tabla No.1 que se encuentra en la sección de resultados.

Con los resultados del modelo de seguridad se concluyó que es factible detectar amenazas debido a que sus valores de precisión fueron óptimos por lo que se pudo resolver la pregunta científica planteada en la investigación. Cabe recalcar que se realizaron pruebas con diferentes capturas del tráfico de red, en distintos días durante una semana, de este modo se obtuvieron valores del entrenamiento variados para validar la viabilidad del modelo.

Conflictos de intereses

En la elaboración de este articulo no hubo conflicto de intereses entre los autores.

Contribución de los autores

1. Conceptualización: Elein Ivette Terán Zurita
2. Curación de datos: Viviana Stefany Gonzalez Mestanza
3. Análisis formal: Elein Ivette Terán Zurita
4. Investigación: Elein Ivette Terán Zurita
5. Metodología: Viviana Stefany Gonzalez Mestanza
6. Administración del proyecto: Alfonso Guijarro Rodríguez
7. Recursos: Viviana Stefany Gonzalez Mestanza y Elein Ivette Terán Zurita
8. Software: Viviana Stefany Gonzalez Mestanza
9. Supervisión: Alfonso Guijarro Rodríguez
10. Validación: Alfonso Guijarro Rodríguez y Gladys Jácome Morales
11. Visualización: Alfonso Guijarro Rodríguez y Gladys Jácome Morales
12. Redacción – borrador original: Viviana Stefany Gonzalez Mestanza y Elein Ivette Terán Zurita
13. Redacción – revisión y edición: Alfonso Guijarro Rodríguez y Gladys Jácome Morales

Financiamiento

Es importante mencionar que los investigadores cubrieron los costos de financiamiento de la investigación como gastos directos, y como gastos indirectos están los recursos con los que contribuye la Universidad de Guayaquil, internet, energía eléctrica, equipos de computación entre otros.

Referencias

Abril, L. (29 de julio de 2021). Ecuador está entre los países con más ciberataques en América Latina. *Diario EL COMERCIO*. Obtenido de https://www.elcomercio.com/tendencias/ecuador-ciberataques-america-latina-hacker.html

Diazgranados, H. (31 de agosto de 2021). *Ciberataques en América Latina crecen un 24% durante los primeros ocho meses de 2021*. Obtenido de Ciberataques en América Latina crecen un 24% durante los primeros ocho meses de 2021: https://latam.kaspersky.com/blog/ciberataques-en-america-latina-crecen-un-24-durante-los-primeros-ocho-meses-de-2021/22718/

Estévez, J. (2020). *Detección de anomalías en la red empleando técnicas de machine learning.* Universidad de Coruña. Obtenido de https://ruc.udc.es/dspace/bitstream/handle/2183/26827/J.J_Estévez_Pereira_2020_Detecció_de_anomalías_en_la_red.pdf?sequence=3&isAllowed=y

Galmés, E. (2019). *Machine Learning aplicado a la segurida*. Universidad Oberta de Catalunya. Obtenido de https://openaccess.uoc.edu/bitstream/10609/111586/6/egalmesgTFM1219memoria.pdf

Recordon, A., & Ruiz, S. (2020). *Detección y Clasificación Zero-Day Malware a través de Data Mining y Machine Learning.* Universidad Nacional de la Plata. Obtenido de http://sedici.unlp.edu.ar/bitstream/handle/10915/117193/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y

Vásquez, M. (2021). *Análisis de amenazas de seguridad basado en la detección de anomalías en el tráfico de red de la infraestructura tecnológica de instituciones de educación superior mediante el uso de técnicas de machine learning.* Cuenca. Obtenido de http://dspace.ucuenca.edu.ec/handle/123456789/37017

Vera, C. (2020). *Detección de anomalías.* Universidad de Guayaquil. Obtenido de http://repositorio.ug.edu.ec/bitstream/redug/55096/1/TESIS%20-%20VERA%20CORDOVA%20ANA%20LUISA.pdf