Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2026,
Volumen 10, Número 1.

https://doi.org/10.37811/cl_rcm.v10i1

MODELO DE CLASIFICACIÓN BASADO EN PERCEPTRÓN MULTICAPA PARA LA DETECCIÓN DE INTRUSIONES EN REDES DE COMPUTADORAS

CLASSIFICATION MODEL BASED ON MULTILAYER PERCEPTRON FOR INTRUSION DETECTION IN COMPUTER NETWORKS

Paquita Alejandra Cuadros García
Universidad Autónoma de Coahuila, México

Cristian Stalin Sancho López
University Institute of Materials Technology, Spain

Jessica Pilar Alejandro Becerra
Instituto Tecnológico de Ciudad Guzmán, México

Alexander Patricio Sócola Coronel
University Institute of Materials Technology, Spain

Jonathan Enrique Sanmartin Pangay
Instituto Tecnológico de Ciudad Guzmán, México

DOI: https://doi.org/10.37811/cl_rcm.v10i1.22806

Modelo de clasificación basado en perceptrón multicapa para la detección de intrusiones en redes de computadoras

Paquita Alejandra Cuadros García [1]
cuadros.paquita@itred.edu.ec
p.cuadros.garcia@posgradountumbes.edu.pe
https://orcid.org/0000-0002-0734-6114
Instituto Superior Tecnológico Rey David, Universidad Nacional de Tumbes.
Ecuador

Cristian Stalin Sancho López
cssancho@isthuaquillas.edu.ec
c.sancho.lopez@posgradountumbes.edu.pe
https://orcid.org/0000-0002-2974-5896
Instituto Superior Tecnológico Huaquillas, Universidad Nacional de Tumbes.
Ecuador

Jessica Pilar Alejandro Becerra
jpalejandro@isthuaquillas.edu.ec
https://orcid.org/0000-0003-1612-2988
Instituto Superior Tecnológico Huaquillas
Ecuador

Alexander Patricio Sócola Coronel
apsocola@isthuaquillas.edu.ec
https://orcid.org/0000-0002-4816-2901
Instituto Superior Tecnológico Huaquillas
Ecuador

Jonathan Enrique Sanmartin Pangay
jesanmartin@isthuaquillas.edu.ec
https://orcid.org/0009-0008-3179-1518
Instituto Superior Tecnológico Huaquillas
Ecuador

RESUMEN

La detección de intrusiones en redes de computadoras constituye uno de los principales desafíos en el ámbito de la ciberseguridad, debido al crecimiento exponencial del tráfico digital y a la sofisticación constante de los ataques informáticos. En este estudio se propone el diseño e implementación de un modelo de clasificación supervisada basado en redes neuronales artificiales, específicamente un perceptrón multicapa (MLP), orientado a la detección de intrusiones en redes de computadoras. El estudio adopta un enfoque cuantitativo experimental-descriptivo, siguiendo un proceso sistemático que incluyó la selección del conjunto de datos, el preprocesamiento de la información, el diseño de la arquitectura del modelo, el entrenamiento, la evaluación y la validación de los resultados obtenidos. Los resultados obtenidos muestran que el modelo MLP alcanza una precisión superior al 85 % en la clasificación de conexiones normales e intrusivas, evidenciando un desempeño satisfactorio y estable. Estos hallazgos confirman la efectividad del aprendizaje profundo aplicado a la detección de intrusiones y destacan su potencial como herramienta de apoyo para fortalecer la seguridad en infraestructuras de red. Finalmente, se concluye que el modelo propuesto constituye una solución viable para entornos de ciberseguridad, con posibilidades de mejora mediante la incorporación de nuevas características y la evaluación de arquitecturas alternativas.

Palabras clave: ciberseguridad; detección de intrusiones; aprendizaje automático; redes neuronales

[1] Autor principal
Correspondencia: cuadros.paquita@itred.edu.ec

Classification model based on multilayer perceptron for intrusion detection in computer networks

ABSTRACT

Detecting intrusions in computer networks is one of the main challenges in the field of cybersecurity, due to the exponential growth of digital traffic and the constant sophistication of cyberattacks. This study proposes the design and implementation of a supervised classification model based on artificial neural networks, specifically a multilayer perceptron (MLP), aimed at detecting intrusions in computer networks. The study adopts a quantitative experimental-descriptive approach. A preprocessing, training, and evaluation pipeline was designed for a supervised classification model with MLP (Ajagbe et al., 2024; Kohli, 2025), following a systematic process that included data set selection, information preprocessing, model architecture design, training, evaluation, and validation of the results obtained. The results obtained show that the MLP model achieves an accuracy of over 85% in the classification of normal and intrusive connections, demonstrating satisfactory and stable performance. These findings confirm the effectiveness of deep learning applied to intrusion detection and highlight its potential as a support tool for strengthening security in network infrastructures. Finally, it is concluded that the proposed model constitutes a viable solution for cybersecurity environments, with potential for improvement through the incorporation of new features and the evaluation of alternative architectures.

***Keywords*:** cybersecurity; intrusion detection; machine learning; neural networks.

Artículo recibido 02 enero 2026
Aceptado para publicación: 30 enero 2026

INTRODUCCIÓN

En este estudio se explora un Modelo de clasificación basado en perceptrón multicapa entrenado con UNSW-NB15, dataset realista y diverso, para la detección de intrusiones en redes de computadoras. Considerando que el acelerado crecimiento de las redes de datos, la virtualización de servicios y la adopción masiva de tecnologías conectadas han transformado los ecosistemas digitales en infraestructuras críticas para instituciones, empresas y gobiernos. Este desarrollo ha generado un incremento proporcional de amenazas que aprovechan vulnerabilidades en los flujos de red, superando la capacidad de detección de mecanismos estáticos (García et al., 2025; Alqithami et al., 2024). Los incidentes de intrusión, en especial los ataques de repetición continua como Denegación de Servicio (DoS), exploits automatizados y fuerza bruta, afectan no solo la disponibilidad de los servicios, sino también la confianza en la gestión tecnológica interna (Kohli & Verma, 2023; Mendoza et al., 2025).

Los Sistemas de Detección de Intrusiones (IDS) tradicionales, basados en firmas, muestran limitaciones frente a amenazas nuevas o variantes de ataques, ya que dependen de bases de datos preconstruidas que no siempre reflejan la dinámica real del tráfico moderno (Zhang et al., 2025; Torres et al., 2024). Esto ha impulsado el uso de Machine Learning (ML) y Deep Learning (DL) como alternativas capaces de aprender relaciones no lineales complejas y patrones anómalos en grandes volúmenes de datos (Butt et al., 2022; Chinnasamy et al., 2025).

Dentro de DL, los modelos neuronales como el Perceptrón Multicapa (MLP) han demostrado ser eficaces para clasificar tráfico normal frente a intrusivo, especialmente cuando se combina con selección estratégica de variables y optimizadores adaptativos para lograr convergencia eficiente (Yin et al., 2025; Mebawondu et al., 2025).

La importancia de este trabajo de investigación radica en el crecimiento constante de las redes de datos y el ofrecer seguridad de los mismos es un servicio mínimo que debe ofrecer cualquier red de datos, al implementar el Modelo de clasificación basado en perceptrón multicapa se puede detectar y disminuir las intrusiones en las redes de datos.

Deep Learning (DL) procesa datos jerárquicamente, aprende relaciones complejas y es eficaz en tráfico de red de alta dimensionalidad (Butt et al., 2022; Mendoza et al., 2025). Se recomienda el uso de optimizadores adaptativos como RMSprop o Adam para acelerar la convergencia en modelos MLP

compactos (Vibhute et al., 2024; Mebawondu et al., 2025).

Las redes neuronales artificiales han demostrado un alto rendimiento en tareas de clasificación, reconocimiento de patrones y análisis de grandes volúmenes de datos. En el contexto de la ciberseguridad, el aprendizaje profundo permite analizar tráfico de red de alta dimensionalidad y detectar patrones sutiles asociados a comportamientos maliciosos que podrían pasar desapercibidos para métodos tradicionales.

Finalmente se puede destacar que en este estudio se explora un modelo MLP compacto entrenado con UNSW-NB15, dataset realista y diverso, con el objetivo de contribuir al fortalecimiento de NIDS inteligentes aplicables a contextos académicos y de transformación digital interna.

METODOLOGÍA

El estudio adopta un enfoque cuantitativo experimental-descriptivo. Se diseñó un pipeline de preprocesamiento, entrenamiento y evaluación para un modelo de clasificación supervisada con MLP (Ajagbe et al., 2024; Kohli, 2025). El estudio se centra en el desarrollo de un modelo supervisado capaz de identificar intrusiones en redes de computadoras a partir de datos previamente etiquetados, permitiendo evaluar el desempeño del modelo mediante métricas objetivas.

El diseño de la investigación se basa en la implementación de un modelo de clasificación supervisada utilizando redes neuronales artificiales. Se siguió un proceso sistemático que incluyó la selección del conjunto de datos, el preprocesamiento de la información, el diseño de la arquitectura del modelo, el entrenamiento, la evaluación y la validación de los resultados obtenidos.

Para el desarrollo del modelo se utilizó el conjunto de datos UNSW-NB15, ampliamente reconocido en estudios de detección de intrusiones. Este dataset contiene registros de tráfico de red capturados en un entorno de laboratorio controlado, incluyendo tanto conexiones normales como diferentes tipos de ataques informáticos. El conjunto de datos está compuesto por múltiples características relacionadas con el comportamiento del tráfico de red, así como una variable objetivo denominada label, donde el valor 0 representa conexiones normales y el valor 1 corresponde a conexiones intrusivas. Para el análisis se emplearon los archivos de entrenamiento y prueba, los cuales fueron integrados en un único dataset con el fin de optimizar el proceso de aprendizaje del modelo. Los archivos CSV de training y testing del conjunto UNSW-NB15 fueron unificados para entrenar correctamente el algoritmo. Se mantuvo la

columna label como variable objetivo (0 = normal, 1 = intrusión) y se trataron valores nulos (Chila et al., 2023; Vibhute et al., 2024).

Del conjunto total de características disponibles en el dataset UNSW-NB15, se seleccionaron cinco variables relevantes para el entrenamiento del modelo, debido a su alta correlación con patrones de ataques frecuentes:

- DUR_CON: Duración de la conexión
- SPKTS: Número de paquetes de origen a destino
- DPKTS: Número de paquetes de destino a origen
- SBYTES: Bytes transmitidos de origen a destino
- RATE: Tasa de transferencia del paquete

Estas características presentan un peso significativo en ataques de tipo Ataque de Denegación de Servicio (DoS) y fuerza bruta, debido a la repetición continua y envío constante de paquetes. La selección coincide con evidencia reciente que prioriza atributos temporales, volumétricos y de tasa de paquetes como predictores fuertes en DoS y fuerza bruta (Yin et al., 2025; García et al., 2025; Mendoza et al., 2025).

El preprocesamiento de los datos constituyó una etapa fundamental para garantizar la calidad de la información de entrada al modelo. Las actividades realizadas incluyeron:

- Integración de los archivos de entrenamiento y prueba en un solo conjunto de datos.
- Conversión de las variables seleccionadas a formato numérico.
- Identificación y tratamiento de valores nulos, reemplazándolos por valores adecuados para evitar sesgos en el entrenamiento.
- Normalización de las variables con el fin de homogeneizar las escalas y mejorar la convergencia del modelo.

Estas etapas se sustentan como determinantes para la convergencia y estabilidad en modelos neuronales compactos para NIDS (Kohli & Verma, 2023; Chinnasamy et al., 2025; Vibhute et al., 2024). Además, estas acciones permitieron reducir el ruido en los datos y facilitar el aprendizaje del modelo de clasificación.

Los datos fueron convertidos a formato numérico, se corrigieron valores nulos y se dividieron en conjuntos de entrenamiento (80%) y prueba (20%), garantizando una evaluación objetiva del modelo.

Una vez preprocesados los datos, se procedió a dividir el conjunto en datos de entrenamiento (80%) y prueba (20%), utilizando el método train-test split. Esta división permitió entrenar el modelo con la mayor cantidad de datos posibles y evaluar su capacidad de generalización sobre datos no vistos.

El modelo de clasificación fue implementado utilizando la biblioteca Keras, bajo una arquitectura secuencial de tipo perceptrón multicapa (MLP). La red neuronal se estructuró de la siguiente manera:

- Capa de entrada con cinco neuronas, correspondientes a las variables seleccionadas.
- Tres capas ocultas densas con funciones de activación ReLU y tanh.
- Capa de salida con función de activación softmax, orientada a la clasificación binaria del tráfico de red.
- Para el proceso de entrenamiento se empleó el optimizador RMSprop y la función de pérdida categorical crossentropy, adecuados para problemas de clasificación supervisada.

El modelo se entrenó por 10 épocas mediante optimización de pesos para minimizar la pérdida. Se implementó un Perceptrón Multicapa (MLP) en Keras, con capas densas y una salida softmax para clasificación binaria. Se utilizó la función de pérdida categorical crossentropy y el optimizador RMSprop, reconocido por su eficiencia en redes neuronales compactas aplicadas a la detección de intrusiones. Durante el entrenamiento, el modelo evidenció convergencia adecuada, reflejada en la disminución progresiva de la pérdida y un incremento constante de la precisión, alcanzando valores superiores al 85% en las últimas épocas, lo que confirma un aprendizaje estable sin signos marcados de sobreajuste o subajuste. La literatura respalda el uso de RMSprop en arquitecturas MLP compactas para minimizar la pérdida y mejorar la precisión en la clasificación de tráfico anómalo de red, especialmente en escenarios NIDS basados en características temporales y volumétricas (Butt et al., 2022; Mebawondu et al., 2025).

La evolución de la precisión del modelo (Accuracy) a lo largo de 10 épocas, comparó el desempeño en los conjuntos train y test. Se observó un incremento acelerado de la precisión durante las primeras épocas, seguido de una estabilización gradual conforme avanza el entrenamiento, alcanzando valores cercanos a 0.85 en train y >0.80 en test hacia la época final. La proximidad entre ambas curvas indica

una capacidad adecuada de generalización, sin evidencia marcada de sobreajuste. Fluctuaciones leves en el conjunto de prueba entre las épocas 4 y 5 reflejan variabilidad normal en la convergencia del optimizador adaptativo, comportamiento reportado en modelos compactos MLP aplicados a NIDS (Butt et al., 2022; Mebawondu et al., 2025).

La evaluación del modelo se realizó utilizando el conjunto de datos de prueba, calculando la precisión como métrica principal para medir su capacidad de clasificación. Adicionalmente, se analizó el comportamiento del modelo frente a nuevos datos, permitiendo validar su efectividad para identificar conexiones normales e intrusivas.

Métrica principal: Precisión (Accuracy). Se reportó clasificación correcta alta en tráfico normal e intrusivo, sin sobreajuste significativo y con generalización adecuada en datos no vistos (Butt et al., 2022; Kohli, 2025; Zhang et al., 2025).

RESULTADOS Y DISCUSIÓN

El modelo de clasificación basado en un perceptrón multicapa (MLP) fue evaluado utilizando el conjunto de datos UNSW-NB15, con el fin de analizar su desempeño en la detección de intrusiones en redes de computadoras. Los resultados obtenidos evidencian un rendimiento satisfactorio tanto en la fase de entrenamiento como en la fase de prueba.

Durante el entrenamiento, el modelo mostró una convergencia progresiva, reflejada en el aumento de la precisión y la disminución de la función de pérdida a lo largo de las épocas. Al finalizar el proceso de entrenamiento, se alcanzó una precisión superior al 85 %, lo que indica una adecuada capacidad de aprendizaje de los patrones del tráfico de red.

El modelo alcanzó una precisión final >85% tras 10 épocas, mostrando convergencia estable y disminución progresiva de la función de pérdida, lo cual es competitivo frente a estudios neuronales recientes evaluados en UNSW-NB15 (Butt et al., 2022; Mebawondu et al., 2025; Vibhute et al., 2024).

Tabla 1

Distribución del conjunto de datos utilizado

Conjunto de datos	Número de registros	Porcentaje (%)
Entrenamiento	80 % del total	80 %
Prueba	20 % del total	20 %
Total	100 %	100 %

Nota. El conjunto de datos UNSW-NB15 fue dividido en datos de entrenamiento y prueba con una proporción 80/20, garantizando una evaluación objetiva del modelo.

Tabla 2

Variables seleccionadas para el entrenamiento del modelo

Variable	Descripción
DUR_CON	Duración de la conexión
SPKTS	Paquetes de origen a destino
DPKTS	Paquetes de destino a origen
SBYTES	Bytes transmitidos de origen a destino
RATE	Tasa de transferencia del paquete

Nota. Estas variables fueron seleccionadas por su alta relevancia en la detección de ataques de tipo DoS y fuerza bruta.

Tabla 3

Resultados del entrenamiento del modelo MLP

Métrica	Valor obtenido
Número de épocas	10
Precisión final	> 85 %
Función de pérdida	Disminución progresiva
Optimizador	RMSprop

Nota. El modelo mostró estabilidad durante el entrenamiento, sin evidencias significativas de sobreajuste.

Tabla 4

Desempeño del modelo en el conjunto de prueba

Clase evaluada	Clasificación correcta
Conexiones normales	Alta
Conexiones intrusivas	Alta

Nota. El modelo mantuvo un desempeño consistente al clasificar datos no vistos durante el entrenamiento, lo que confirma su capacidad de generalización.

Tabla 5

Resultado de la predicción del modelo

Registro evaluado	Resultado del modelo
Registro 1	No Ataque
Registro 2	Ataque

Nota. El modelo clasifica el tráfico de red de forma binaria, permitiendo identificar conexiones normales e intrusivas en escenarios simulados.

Los resultados confirman que los modelos MLP compactos, cuando se alimentan con características temporales y de tasa volumétrica, logran precisión alta en detección de intrusiones sin requerir arquitecturas excesivamente profundas, siempre que se aplique normalización y tratamiento adecuado de nulos para reducir ruido (Kohli & Verma, 2023; Chinnasamy et al., 2025). La precisión obtenida (>85%) es competitiva frente a baselines recientes de ANN/NIDS en UNSW-NB15, donde el desempeño mejora cuando se optimiza la selección de variables y se emplean optimizadores adaptativos como RMSprop o Adam para acelerar convergencia y estabilizar la pérdida (Butt et al., 2022; Mebawondu et al., 2025).

Estudios recientes señalan que la correlación entre duración de conexión, tasa de paquetes y volumen de bytes mantiene fuerte asociación con la etiqueta de intrusión, lo cual coincide con la selección realizada en esta investigación y es respaldado por análisis exploratorios de correlación como mapas de calor previos al entrenamiento (Yin et al., 2025; García et al., 2025). No obstante, la literatura también identifica retos persistentes como el desbalance de clases, interpretabilidad de modelos neuronales y costo computacional en entornos edge/IoT, lo que motiva la exploración de técnicas como GAN para aumento de datos, GNN con mecanismos de atención y aprendizaje federado como líneas futuras para mejorar recall y F1-score sin comprometer la simplicidad del MLP (Alqithami et al., 2025; Ajagbe et al., 2024; Portmann et al., 2024).

CONCLUSIONES

El modelo de clasificación basado en perceptrón multicapa desarrollado en este estudio demuestra una alta capacidad para detectar intrusiones en redes de computadoras.

El uso del conjunto de datos UNSW-NB15 y técnicas de aprendizaje profundo permitió obtener resultados satisfactorios y robustos.

Como trabajo futuro, se recomienda incorporar más características, evaluar otros algoritmos de aprendizaje automático y realizar comparaciones con modelos tradicionales de detección de intrusiones.

REFERENCIAS BIBLIOGRÁFICAS

Ajagbe, S. A., et al. (2024). Intrusion detection: A comparison study of machine learning models using the UNSW-NB15 dataset. SN Computer Science, 5(1028). https://link.springer.com/article/10.1007/s42979-024-03369-0.

Alqithami, S., Stiawan, D., & Budiarto, R. (2025). On-the-fly, memory-aware unsupervised learning for network anomaly detection: A systematic literature review. PRIME Journal. https://www.sciencedirect.com/science/article/pii/S2772671125002621.

Baidar, R., Maric, S., & Abbas, R. (2025). *Hybrid deep learning–federated learning powered intrusion detection system for IoT/5G advanced edge computing network*. *arXiv*. https://arxiv.org/abs/2509.15555.

Biradar, J., Shah, S., & Naik, T. (2025). Attention augmented GNN RNN-attention models for advanced cybersecurity intrusion detection. arXiv. https://arxiv.org/abs/2510.25802.

Chinnasamy, R., Subramanian, M., & Easwaramoorthy, S. V. (2025). Deep learning-driven methods for network-based intrusion detection systems: A systematic review. ICT Express. https://www.sciencedirect.com/science/article/pii/S2405959525000050.

Lorenzo, D. (2025). Overview on intrusion detection systems for computers and deep learning approaches. Computational Intelligence and Neuroscience. https://www.mdpi.com/2073-431X/14/3/87.

Effectiveness of Machine Learning Models in Intrusion Detection Systems. (2024). LEIRD Virtual Edition. https://laccei.org/LEIRD2024-VirtualEdition/papers/Contribution_237_a.pdf.

Hybrid Machine Learning–Based Intrusion Detection System. (2025). International Journal of Security and Safety Engineering. https://iieta.org/journals/ijsse/paper/10.18280/ijsse.150815.

Kohli, M. (2025). A comprehensive survey on techniques, challenges, evaluation metrics and applications of deep learning models for anomaly detection. SN Applied Sciences. https://link.springer.com/article/10.1007/s42979-024-03369-0.

Lopez-Chila, et al. (2024). Revisión sobre el uso de machine learning para mejorar la seguridad informática. Universidad Politécnica Salesiana. https://dspace.ups.edu.ec/bitstream/123456789/29275/1/UPS-GT005918.pdf.

Lotfi, S., et al. (2022). Network intrusion detection with limited labeled data using self-supervision. arXiv. https://arxiv.org/abs/2209.03147

Modirrousta, M. H., et al. (2022). Analysis of anomalous behavior in network systems using deep reinforcement learning with CNN architecture. arXiv. https://arxiv.org/abs/2211.16304?utm_source=chatgpt.com

Muhammad, S. H., & Saeed, A. (2025). Network intrusion detection using the UNSW-NB15 dataset and a conditional GAN-augmented CNN. Computer Science & Information Technology (CS & IT). https://www.researchgate.net/publication/395826612_NETWORK_INTRUSION_DETECTION_USING_THE_UNSW-NB15_DATASET_AND_THE_CONDITIONAL_GANAUGMENTED_CNN.

Portmann, M., & Seneviratne, A. (2024). Towards explainable intrusion detection using neural networks. IEEE/ACM Conference on Big Data Computing, 1–8. https://doi.org/10.1109/BDCAT.2024.00012.

Tafreshian, B., & Zhang, S. (2025). A defensive framework against adversarial attacks on ML-based intrusion detection systems. arXiv. https://arxiv.org/abs/2502.15561

Vibhute, A. D., Khan, M., Patil, C. H., Gaikwad, S. V., & Mane, A. V. (2024). Network anomaly detection and performance evaluation of CNN on UNSW-NB15 dataset. Procedia Computer Science. https://www.sciencedirect.com/science/article/pii/S1877050924008871.

Waghmode, P. (2025). Intrusion detection system based on machine learning using multiple datasets including UNSW-NB15. Scientific Reports. https://www.nature.com/articles/s41598-025-95621-7.

Yagual, D. I. Q. (2022). Una revisión del aprendizaje profundo aplicado a la ciberseguridad. Revista RCTU. https://www.revistas.upse.edu.ec/index.php/rctu/article/view/671/556.

Zhang, Y., et al. (2025). A review of deep learning applications in intrusion detection systems. Applied Sciences, 15(3), 1552. https://www.mdpi.com/2076-3417/15/3/1552.