Machine learning para administración de redes y sistemas

Машинное обучение для управления сетями и системами.

Alexis Paleta Osorio, Luis Palalia Mani, Brandon Domínguez León, Kin Nahomi Pérez Flores, María del Carmen Santiago Díaz, Ana Claudia Zenteno Vázquez, Gustavo Trinidad Rubín Linares

Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla, Av. San Claudio, s/n. 72570 Puebla, Puebla

{alexis.paletao, luis.palalia, brandon.dominguez, kin.peresfl}@alumno.buap.mx, {marycarmen.santiago, ana.zenteno, gustavo.rubin}@correo.buap.mx

Resumen: A la hora de realizar verificaciones, evaluaciones y estimaciones, los procesos de machine learning se han destacado por su versatilidad y gran variedad de aplicaciones para resolver todo tipo de problemas, incluyendo la administración de redes y sistemas. Una gran cantidad de procesos pueden agilizarse, mejorarse y crear en base al aprendizaje automático que servirá para llevar al siguiente nivel al sistema que se esté trabajando. En este trabajo se exploran los conceptos fundamentales para poder saber qué es y cómo puede aplicarse el machine learning en muchos enfoques distintos en el área de redes, como el monitoreo, control de acceso y más.

Аннотация: области верификации, оценки и прогнозирования процессы машинного обучения зарекомендовали себя благодаря своей универсальности и широкому спектру применений для решения самых разных задач, включая управление сетями и системами. На основе машинного обучения можно оптимизировать, улучшить и создать множество процессов, что позволит вывести разрабатываемую систему на новый уровень. В данной статье рассматриваются фундаментальные концепции, позволяющие понять, что такое машинное обучение и как его можно применять в различных областях сетевых технологий, таких как мониторинг, контроль доступа и многое другое.

Frases y palabras clave: Machine Learning, Redes, SysAdmin, Sistemas, Administración de redes.

Ключевые слова и фразы: **Машинное обучение**, **Сетевые технологии**, **Верификация и прогнозирование**, **Оптимизация процессов**, **Мониторинг и контроль доступа**

---

1 Introducción

La inteligencia artificial ha tenido un crecimiento sorprendente en estos últimos años, demostrando las muchas aplicaciones y beneficios que se pueden dar en muchas áreas distintas, incluso en las que no están enfocadas al cien por ciento en la tecnología. La utilización de distintas áreas, como lo son el machine learning (subconjunto de la inteligencia artificial) y las redes de computadoras, dan resultados y procesos que por sí solos no serían posibles de alcanzar. Para poder presentar los ejemplos de aplicación de estas áreas del conocimiento primero se necesitan sentar las bases con sus definiciones, y en secciones posteriores, se explorarán ejemplos en los que el uso conjunto de estos enfoques ha podido desarrollarse a un punto altamente destacable que tiene una gran relevancia para el manejo de redes y su administración.

2 Conceptos fundamentales

2.1 Administración de redes

Es la gestión de una red mediante habilidades, procesos y herramientas para garantizar que los recursos de red, como el hardware, el almacenamiento, la memoria, el ancho de banda, los datos y la potencia de procesamiento disponibles en la red, sean fácilmente accesibles para los usuarios y servicios de la manera más eficiente y segura posible [1].

2.2 Machine learning

Existen bastantes definiciones sobre el machine learning, destacando una de ellas se define como el conjunto de métodos que pueden detectar patrones de manera automática en los datos y después usar los descubrimientos obtenidos para realizar predicciones en datos futuros, o desarrollar procesos de toma de decisiones en base a la información entrante [2]. El machine learning es comúnmente dividido en dos tipos, siendo estos el aprendizaje supervisado (clasificación, regresión) y el no supervisado (clustering), dependiendo de la problemática que se esté tratando de implementar se utiliza un tipo de solución para lograr el cometido de los datos que se tienen presentes.

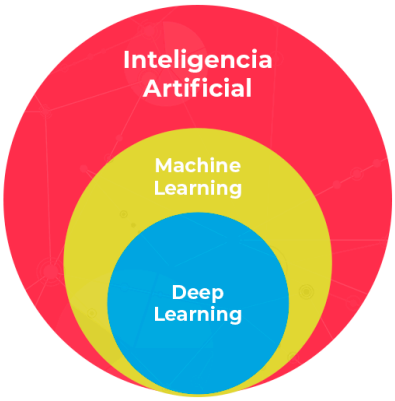

Fig. 1. Relación del machine learning con la inteligencia artificial [3].

3 Aplicaciones del machine learning en la administración de redes

Explorando los procesos que se han llevado a cabo con la utilización de machine learning, se listan los que han tenido un desempeño destacable según los administradores de sistemas/redes, es decir, que han tenido un beneficio importante que mejora en gran medida los procesos convencionales que tenían en un inicio.

3.1 Machine learning en políticas basadas en el control de acceso

Cuando el estado del sistema de control de acceso es complejo, el machine learning ha demostrado ventajas en el rendimiento y generalización para su administración sobre métodos convencionales [4]. El Machine learning (ML) tiene distintas utilidades en el campo de control de acceso, para comprender su capacidad y potencial se realiza el siguiente ejemplo, un modelo de machine learning tomará las decisiones de control de acceso a un servicio, para esto el modelo realiza predicciones de autorización o rechazo para las solicitudes entrantes, este es un problema de aprendizaje supervisado, específicamente de clasificación binaria debido a que cada solicitud puede adoptar únicamente dos estados, uno en el que su acceso es concedido y otro en el que no se le concedió al solicitante. Para la construcción del modelo se necesitará información de usuarios para poder tener una referencia de las características con las cuales las solicitudes deberán ser aceptadas, la información puede variar bastante en cuanto a que se utilizará para la toma de decisiones, un ejemplo puede ser la frecuencia con la que se ha solicitado un servicio en un lapso de tiempo, el abuso de solicitudes podría ser un factor determinante para rechazar una solicitud de acceso, otro ejemplo podría ser el propósito de la solicitud, puede que se hayan hecho varias solicitudes por un rol en específico que debe desempeñar el solicitante y por ende no debe haber inconveniente en permitirle el uso del servicio solicitado. La suma de características, su relevancia, la cantidad y calidad de ejemplos con los que el modelo se entrena determinarán la eficacia de las predicciones, y para este enfoque en particular además del modelo de machine learning encargado de realizar la clasificación de acceso o rechazo en las solicitudes se tiene al administrador del sistema que proporcionará un feedback para la mejora del modelo con la ayuda de un experto, este podría establecer reglas que se necesitan respetar para la decisión del control de acceso y ayudar el modelo a realizar su trabajo de una mejor manera.

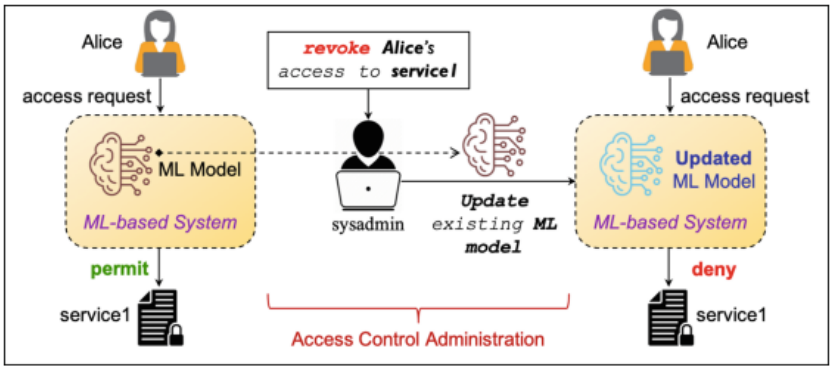

Fig. 2. Problema de administración del control de acceso solventado con ML [4].

3.2 Monitoreo

Existen varios enfoques con los que se podría tratar el monitoreo del sistema, se podría entender como un problema de clasificación para de manera simplista etiquetar los datos con un comportamiento “bueno” o “malo” por ejemplo, otro enfoque es diferenciar a los datos unos con otros para crear grupos de registros similares y en base a la identificación de qué hace a un ejemplo entrar en un grupo u otro poder tomar acción en un clúster determinado.

¿Qué es lo que se puede monitorear en el sistema? Desde procesos, red, servicios, actividades realizadas por un usuario y mucho más, la idea de este ejemplo es determinar si estos casos se comportan de manera “normal”. Para lograr esto, y tomando como idea para resolver el problema la clasificación, se tomaron rangos para lograr diferenciar entre el comportamiento esperado y uno que se aleja de ello. Se muestra la siguiente imagen en la que el área sombreada es la correspondiente a como normalmente es el comportamiento de cierto elemento, los datos que se encuentran fuera de esta zona podrían requerir intervención por la diferencia que hay entre lo que está realizando y lo que se espera que debiera poder hacer.

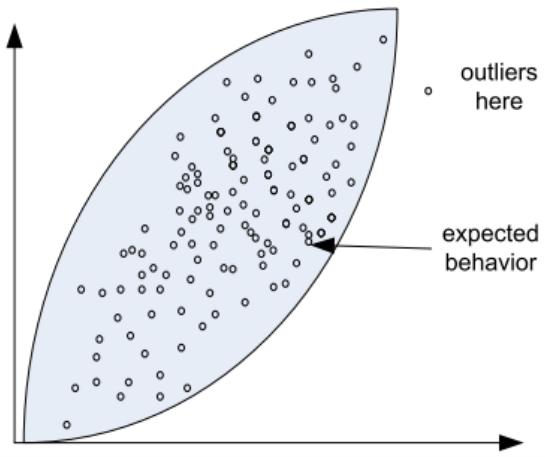

Fig. 3. Representación en dos dimensiones del comportamiento esperado de los datos [5]

3.3 Control de congestión

El control de congestión es un tema que ha sido estudiado desde hace bastante tiempo, se han propuesto varias técnicas para solventar este problema, aunque muchas de estas se basan en su mayoría en el conocimiento experto de los humanos, la realidad es que existen limitaciones para los métodos convencionales debido a que hacen suposiciones basadas en la literatura que muchas veces no llegan a ser completamente realistas.

El machine learning entra justamente para tratar de solventar estas deficiencias, además del comportamiento altamente funcional que pueden arrojar los modelos, la suma del conocimiento en el tema reforzará de gran manera la solución para crear mejores aprendices [6]. Para comprender mejor la aplicación del machine learning en el control de congestión se explica Remy. Remy es un programa de computadora que descubre cómo las computadoras pueden cooperar mejor para compartir una red, crea algoritmos integrales de control de congestión que se integran con el Protocolo de Control de Transmisión (TCP). Estos algoritmos generados por computadora pueden lograr un mayor rendimiento y equidad que los esquemas más sofisticados diseñados por humanos.

Con Remy, los diseñadores de protocolos humanos especifican explícitamente sus conocimientos previos y suposiciones sobre la red, así como un objetivo al que aspirar. Remy crea entonces un algoritmo para ejecutarse en los endpoints, posiblemente complejo, pero con un comportamiento emergente simple: lograr el objetivo de la mejor manera posible en las redes descritas por las suposiciones establecidas [7].

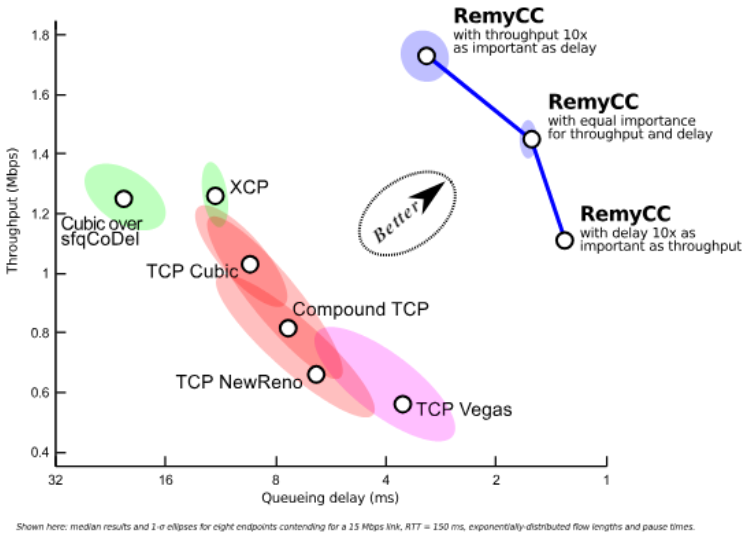

Fig. 4. Rendimiento de los algoritmos generados por Remy contra algoritmos de control de congestión de TCP creados por humanos [7].

3.4 Predicción del tráfico

Los datos pueden presentarse de muchas maneras distintas, la conversión a un formato apropiado para que se les pueda aplicar una técnica de machine learning es un paso importante para tener en consideración. El análisis de series temporales, es decir, datos ordenados en el tiempo, es clave en

múltiples dominios como clima, salud, finanzas, sensores industriales y streaming. Dentro de las redes, una aplicación crítica es la predicción del tráfico, donde el crecimiento de usuarios exige una gestión eficiente de recursos para evitar subutilización o sobrecarga [8]. El poder contar con predicciones sobre el tráfico ayudará a redirigir los recursos en donde y cuando más se necesite, además de que el constante monitoreo de la actividad en la red podría resultar en la identificación de patrones interesantes que beneficien en el entendimiento del uso de la red, además de las predicciones de la cantidad estimada de demanda que se presentará en un momento dato, se puede analizar el comportamiento del tráfico para realizar clasificaciones sobre el bajo o alto intercambio de información para poder identificar las causas de ciertos comportamientos observados en la red.

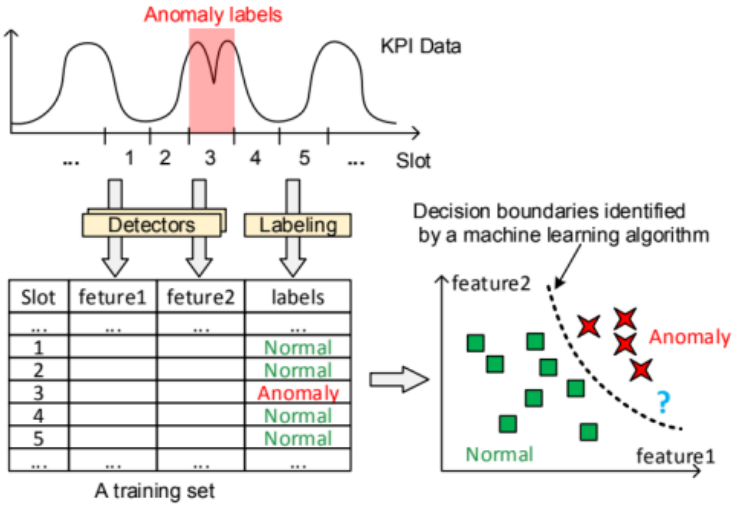

Fig. 5. Ejemplo de detección de anomalías en el tráfico [8].

3.5 Protección a ciberataques

Con el aumento del uso de tecnologías de comunicación, también han aumentado el aumento del número de ciberataques sufridos, y por lo tanto el costo y peligro que estos representan también lo han hecho. Debido a esto la necesidad de métodos y formas de evitar estos ataques ha aumentado junto con la cantidad de estudios realizados para buscar modelos que permitan esto.[9]

Actualmente existen dos enfoques para poder tratar con estos ataques, reactivos y proactivos. Los primeros se tratan de métodos que buscan patrones irregulares en una red que se puedan considerar sospechosos. Estos modelos solamente pueden detectar ataques ya realizados. Los proactivos están basados, en su mayoría, en modelos estadísticos mediante la predicción de series de tiempo.

Uno de los acercamientos que existen al tema de tratar la seguridad de un sistema de red es IDS, o Intrusion Detection System. Se tratan de sistemas que deben ser capaces de recolectar datos de la red relacionados con comportamientos asociados a ataques, almacenarlos, analizarlos, y mandar alertas. [10]

Debido a las características necesarias para un IDS, la implementación de algoritmos de machine learning es conveniente para poder mejorar su desempeño, ya que estos son capaces de manejar grandes y complejos volúmenes de datos, además de poder detectar patrones dentro de estos datos. Todo esto hace que la implementación de machine learning en esta forma de ciberseguridad sea común.

Dentro de esta categoría existen muchas más opciones para implementar un sistema IDS, cada una usando diferentes algoritmos de machine learning, cada una con su propias ventajas y desventajas, por lo que sería imposible dar un formato general, o algún tipo de aprendizaje que sea más usado que otro.

Algo que muchos de estos modelos tienen en común es la forma en que obtienen los datos sobre los cuáles realizan sus predicciones, particularmente una que se ha visto en aumento de uso últimamente son los llamados "honey pots", aplicados por diversos modelos de IDS, y sobre los cuáles se busca mejorar el rendimiento que implementan machine learning. Un ejemplo de un conjunto de datos usado comúnmente es el dataset KDD99, sobre el cuál se han creado varios modelos que implementan K-NN, SVM, Naive Bayes, entre muchos otros, que buscaban poder distinguir conexiones buenas de maliciosas [9].

4 Conclusiones

El avance continuo de las distintas áreas tecnológicas permite poder realizar tareas cada vez más avanzadas, las redes en conjunto con la inteligencia artificial y particularmente el machine learning, han logrado obtener resultados y realizar procesos con un gran impacto para la solución de problemas y necesidades que día con día se componen de una mayor complejidad y requerimientos. Existen una gran cantidad de enfoques con los que se pueden fusionar estas áreas como la seguridad de un sistema, identificar patrones en el uso de ciertos servicios o recursos, o el monitoreo de la actividad que existe en la red, entre muchas otras. El continuo progreso en la implementación de técnicas y propuestas traen consigo un futuro lleno de posibilidades y distintas aplicaciones para poder aprovechar todos estos avances.

Referencias

[1]IBM. (2024, 16 mayo). Gestión de redes. *¿Qué es la gestión de redes?* Recuperado el 30 de abril de 2025, de https://www.ibm.com/mx-es/topics/network-management

[2] Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.

[3] *¿Qué es el Machine Learning? | Aprendizaje Automático | UCM*. (s. f.). Master de Data Science de la Universidad Complutense de Madrid. Recuperado el 30 de abril de 2025, de https://www.masterdatascienceucm.com/que-es-machine-learning/

[4] Nobi, M. N., Krishnan, R., Huang, Y., & Sandhu, R. (2022, September). Administration of machine learning based access control. In *European Symposium on Research in Computer Security* (pp. 189-210). Springer Nature Switzerland.

[5] Chiarini, M., & Couch, A. (2007). *Machine learning for the system administrator*.

[6] Kanakis, M. E., Khalili, R., & Wang, L. (2022). Machine learning for computer systems and networking: A survey. *ACM Computing Surveys*, *55*(4), 1-36.

[7] Balakrishnan, K. W. A. H. (s. f.). *TCP ex Machina*. https://web-mit-edu.translate.goog/remy/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=wa

[8] Jamshidi, S. (2019). The Applications of Machine Learning Techniques in Networking.

[9] Oluoha, O. U., Yange, T. S., Okereke, G. E., & Bakpo, F. S. (2021). Cutting Edge Trends in Deception Based Intrusion Detection Systems—A survey. Journal of Information Security, 12(04), 250–269. https://doi.org/10.4236/jis.2021.124014