

## ANEXO I

Evidencia de documentos generados durante la etapa de planificación de la investigación realizada.

### DOCUMENTO 1: Alcance de los antecedentes del estudio.

#### 1. Background

Esta sección describe tres conceptos fundamentales para comprender el enfoque de este estudio: las comunicaciones ópticas, técnicas de aprendizaje automático y áreas de trabajo en las que se podría aplicar técnicas de aprendizaje automático en Redes Ópticas para mejorar su funcionamiento y desempeño.

##### 1.1 Comunicaciones Ópticas

La comunicación óptica es cualquier tipo de comunicación en la que se utiliza luz para llevar la señal a su destino, en lugar de corriente eléctrica. Al encontrarse basado en el principio de reflexión interna total, la señal de luz se transmite a través de fibra óptica. Un sistema de comunicación óptica utiliza un transmisor, que codifica un mensaje en una señal óptica, un canal, que lleva la señal a su destino, y un receptor, que reproduce el mensaje de la señal óptica recibida [1], [2].

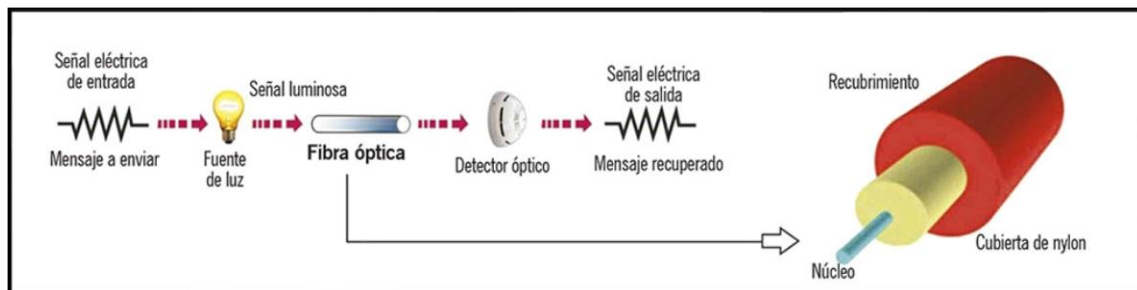


Figura 1. Sistema básico de Comunicación Óptica [3]

Los principales beneficios de la comunicación óptica incluyen un gran ancho de banda, pérdidas excepcionalmente bajas, gran rango de transmisión, ausencia de interferencias electromagnéticas, entre otras. Sin embargo, entre sus desventajas se puede mencionar, que existe un alto costo del cable, transmisor/receptor y otros equipos de apoyo, y la habilidad y experiencia requeridas durante la instalación e interconexión del sistema son complejas [2].

##### 1.2 Técnicas de aprendizaje automático

ML es un área de investigación que se encarga de enseñar a las máquinas a realizar una actividad similar a la mente humana. Aunque su capacidad cognitiva es mucho más limitada que la del ser humano pueden procesar grandes cantidades de información en corto tiempo y obtener información útil [4]. A continuación, se presenta de manera breve una descripción de los 5 algoritmos de aprendizaje automático más populares: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semi-supervisado, aprendizaje por refuerzo y finalmente el aprendizaje auto-supervisado.

- **Aprendizaje supervisado:** El aprendizaje supervisado es una subcategoría de AI/ML, que emplea conjuntos de datos etiquetados para entrenar algoritmos que clasifican información o predicen resultados con precisión. A medida que los datos de entrada ingresan en el modelo, este ajusta sus pesos hasta que el algoritmo opere de manera correcta [5]
- **Aprendizaje no supervisado:** El aprendizaje no supervisado es una subcategoría de AI/ML, que permite que una máquina explore un conjunto de datos sin la necesidad de supervisar el modelo, por lo que es mucho más sencillo de implementar en comparación con el aprendizaje supervisado, ya que el modelo funciona de forma independiente [4].
- **Aprendizaje semi-supervisado:** El aprendizaje semi-supervisado es una subcategoría de AI/ML, que combina técnicas de aprendizaje no supervisado y supervisado. Durante la etapa de entrenamiento, en lugar de agregar etiquetas a todo el conjunto de datos analizados, se revisa y se etiqueta a mano solo una pequeña parte de los datos para entrenar un modelo [6].
- **Aprendizaje reforzado:** El aprendizaje reforzado es una subcategoría de AI/ML, en la que un agente aprende a comportarse en un medio realizando las acciones y viendo los resultados de las acciones. Por cada buena acción, el agente recibe una retroalimentación positiva, y por cada mala acción, el agente recibe una retroalimentación negativa o penalización [4].

### 1.3 Áreas de trabajo en las que se podría aplicar técnicas de aprendizaje automático en Redes Ópticas

Según [8]–[10], las áreas de trabajos en las que se podría aplicar técnicas de ML para solventar problemas relacionados al funcionamiento y desempeño de Redes Ópticas son:

- **Gestión de Recursos (Resource Management):** Crea un método efectivo para analizar el tráfico que recorre a través de una red, reduciendo así el tiempo de inactividad, errores y otros problemas que se puedan suscitar, mientras que permite satisfacer las necesidades de los usuarios y dispositivos al mismo tiempo [11], [12].
- **Asignación de Recursos (Resource Allocation):** Proceso mediante el cual los elementos de una red intentan satisfacer las demandas competitivas que tienen las aplicaciones, principalmente el ancho de banda del enlace y el espacio de búfer en los enrutadores o conmutadores [13].
- **Supervisión de rendimiento (Performance Monitoring):** Implica la gestión de los enlaces entre dispositivos, estaciones de trabajo, servidores, dispositivos virtuales y dispositivos móviles con el fin de identificar la congestión, maximizar y mejorar el rendimiento de la red para el usuario [14], [15].
- **Predicción de tráfico (Traffic Prediction):** Predice el tráfico de red posterior utilizando los datos de tráfico de red anteriores. Esto puede servir como un enfoque proactivo para la administración de redes, tareas de planificación, procesos de seguridad, evitar la congestión y aumentar la velocidad de las redes [16], [17].

- **Clasificación de tráfico (Traffic Classification):** Ofrece la capacidad de reconocer automáticamente la aplicación que ha generado un determinado flujo de paquetes a partir de la observación directa y pasiva de los paquetes individuales, o flujo de paquetes, que fluyen en la red. Cada clase de tráfico identificada puede tratarse de manera diferente, lo que resulta fundamental para una serie de actividades que son de interés para los operadores, proveedores de servicios de Internet y administradores de red en general [18], [19].
- **Control de congestión (Congestion Control):** Se emplea para describir los esfuerzos realizados por los nodos de la red para prevenir o responder a condiciones de sobrecarga [13].
- **Gestión de Fallas (Fault Management):** Proceso de encontrar, aislar y solucionar errores de red en el menor tiempo posible [20].
- **Gestión de QoS y QoE (QoS and QoE Management):** QoS se relaciona con el proceso de gestión de los recursos de la red para reducir la pérdida de paquetes, así como la latencia y la inestabilidad de la red, asignando recursos en función de diversos niveles de prioridad a los distintos tipos de datos que transitan por la red. Mientras que, QoE, es una medida del nivel general de satisfacción del cliente con un proveedor de red [21].
- **Calidad de la Transmisión (QoT):** La calidad de la transmisión es aquel parámetro que depende tanto de la naturaleza del medio que lo rodea como de la naturaleza de la señal. Lograr predecir la calidad de transmisión (QoT) de un camino de luz antes de su despliegue es un paso de fundamental importancia para un diseño optimizado de redes ópticas [22], [23].

## 2. Alcance del estudio

El objetivo general de este estudio es realizar una revisión sistemática de la literatura existente acerca del uso de técnicas y soluciones de aprendizaje automático aplicados para resolver problemas relacionados al funcionamiento y desempeño de Redes Ópticas. Este objetivo general se ha desglosado en preguntas de investigación concretas que se muestran en la TABLA I, así como su justificación para considerarlas.

**TABLA I.** Preguntas de Investigación

No	Pregunta de investigación (RQs)	Justificación
RQ1	¿Cuáles son las técnicas de ML utilizadas para resolver los problemas reportados en [8]–[10] relacionados con el funcionamiento y operación de las redes ópticas?	Saber qué técnicas de ML son utilizadas para problemas reportados en [8]–[10] relacionados con el funcionamiento y operación de las redes ópticas es la piedra angular de este proyecto, ya que permite proporcionar una visión general de esta área de investigación e identificar sus vacíos.
RQ2	¿Qué desafíos informados en [8]–[10] sido mayoritariamente abordados mediante el uso de las técnicas de ML reportadas?	El objetivo de la investigación es obtener conocimiento acerca de cuáles son los desafíos informados en [8]–[10] relacionados con el funcionamiento y operación de las redes ópticas, que han sido mayoritariamente abordados mediante las técnicas de ML identificadas.
RQ3	¿Los datasets que han sido utilizados por las técnicas de ML	Tener conocimiento acerca de la disponibilidad de los datasets empleados por las técnicas de ML en los artículos revisados,

	en los artículos revisados están disponibles?	permite crear un repositorio conjunto de datos para que puedan ser utilizados por quienes los necesiten y se encuentren en casos similares en futuras investigaciones.
RQ4	¿Cuál es el nivel de madurez de las técnicas de ML identificadas?	Utilizando la literatura existente, se asigna un nivel de madurez a las técnicas de ML identificadas en el estudio, para poder reconocer brechas y tendencias de este tema, cuya información obtenida puede ser valiosa para futuras investigaciones.

### 3. Referencias

- [1] "What is Optical Communication? - Definition from Techopedia." <https://www.techopedia.com/definition/24942/optical-communication> (accessed Sep. 18, 2022).
- [2] "What Is Optical Communication." <https://www.glsun.com/article-p105-what-is-optical-communication.html> (accessed Sep. 18, 2022).
- [3] "Luz y fibra óptica en un sistema de comunicación: tendencias actuales - Saberes y Ciencias | Saberes y Ciencias." <https://saberesyciencias.com.mx/2015/06/01/luz-y-fibra-optica-en-un-sistema-de-comunicacion-tendencias-actuales/> (accessed Sep. 18, 2022).
- [4] "TOP 5 Machine Learning Algorithms For Business Applications." <https://mobidev.biz/blog/5-essential-machine-learning-techniques> (accessed Apr. 13, 2022).
- [5] "What is Supervised Learning? | IBM." <https://www.ibm.com/cloud/learn/supervised-learning> (accessed Apr. 13, 2022).
- [6] "¿Qué es machine learning? - España | IBM." <https://www.ibm.com/es-es/cloud/learn/machine-learning> (accessed May 02, 2022).
- [7] "Self-Supervised Learning: Definition, Tutorial & Examples." <https://www.v7labs.com/blog/self-supervised-learning-guide#h1> (accessed Jul. 14, 2022).
- [8] R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *Journal of Internet Services and Applications*, vol. 9, no. 1, Dec. 2018, doi: 10.1186/s13174-018-0087-2.
- [9] R. Gu, Z. Yang, and Y. Ji, "Machine learning for intelligent optical networks: A comprehensive survey," *Journal of Network and Computer Applications*, vol. 157. Academic Press, May 01, 2020. doi: 10.1016/j.jnca.2020.102576.
- [10] Y. Zhang, J. Xin, X. Li, and S. Huang, "Overview on routing and resource allocation based machine learning in optical networks," *Optical Fiber Technology*, vol. 60, Dec. 2020, doi: 10.1016/j.yofte.2020.102355.
- [11] "Overview of Network Resource Management - Managing Network Virtualization and Network Resources in Oracle® Solaris 11.2." [https://docs.oracle.com/cd/E36784\\_01/html/E36813/gecki.html](https://docs.oracle.com/cd/E36784_01/html/E36813/gecki.html) (accessed Jul. 24, 2022).
- [12] "What Is Network Virtualization and Network Resource Management? - Managing Network Virtualization and Network Resources in Oracle® Solaris 11.2." [https://docs.oracle.com/cd/E36784\\_01/html/E36813/gndac.html](https://docs.oracle.com/cd/E36784_01/html/E36813/gndac.html) (accessed Jul. 24, 2022).
- [13] "6.1 Issues in Resource Allocation — Computer Networks: A Systems Approach Version 6.2-dev documentation." <https://book.systemsapproach.org/congestion/issues.html> (accessed Jul. 24, 2022).

[14] "What is Network Performance Monitoring?" <https://www.techtarget.com/searchnetworking/definition/network-performance-monitoring> (accessed Jul. 24, 2022).

[15] "What is Network Performance Monitoring? Definition and FAQs | HEAVY.AI." <https://www.heavy.ai/technical-glossary/network-performance-monitoring> (accessed Jul. 24, 2022).

[16] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Applying deep learning approaches for network traffic prediction," 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017, vol. 2017-January, pp. 2353–2358, Nov. 2017, doi: 10.1109/ICACCI.2017.8126198.

[17] "Overview of traffic prediction system and its impact on TMS efficiency | Download Scientific Diagram." [https://www.researchgate.net/figure/Overview-of-traffic-prediction-system-and-its-impact-on-TMS-efficiency\\_fig5\\_266673838](https://www.researchgate.net/figure/Overview-of-traffic-prediction-system-and-its-impact-on-TMS-efficiency_fig5_266673838) (accessed Jul. 24, 2022).

[18] M. M. Raikar, S. M. Meena, M. M. Mulla, N. S. Shetti, and M. Karanandi, "Data Traffic Classification in Software Defined Networks (SDN) using supervised-learning," *Procedia Comput Sci*, vol. 171, pp. 2750–2759, Jan. 2020, doi: 10.1016/J.PROCS.2020.04.299.

[19] S. Valenti, D. Rossi, A. Dainotti, A. Pescapè, A. Finamore, and M. Mellia, "Reviewing traffic classification," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7754, pp. 123–147, 2013, doi: 10.1007/978-3-642-36784-7\_6/COVER.

[20] "Network Fault Management and Monitoring Tools - ManageEngine OpManager." <https://www.manageengine.com/network-monitoring/network-fault-management.html> (accessed Jul. 24, 2022).

[21] C. Z. Liu and M. Kavakli, "Data-aware QoE-QoS management," *Proceedings of the 2016 IEEE 11th Conference on Industrial Electronics and Applications, ICIEA 2016*, pp. 1818–1823, Oct. 2016, doi: 10.1109/ICIEA.2016.7603882.

[22] "The quality of transmission depends upon:" <https://www.toppr.com/ask/question/the-quality-of-transmission-depends-upon/> (accessed Sep. 12, 2022).

[23] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-Learning Method for Quality of Transmission Prediction of Unestablished Lightpaths," *Journal of Optical Communications and Networking*, Vol. 10, Issue 2, pp. A286–A297, vol. 10, no. 2, pp. A286–A297, Feb. 2018, doi: 10.1364/JOCN.10.00A286.

## DOCUMENTO 2: Identificación de la cadena de búsqueda.

**TABLA I.** Proceso para definición de cadena de búsqueda.

Composición de cadenas de búsqueda	Se toman en cuenta dos temas principales para la creación de la cadena de búsqueda: "Machine Learning", "Optical Communication", "Optical Network"			
	ML	Optical	Cadena de búsqueda de Scopus	# resultados
1 iter	"Machine Learning"	"Optical Communication" - "Optical Network"	TITLE-ABS-KEY ( ( "Machine Learning" ) AND ( "Optical Communication" OR "Optical Network" ) )	841
Final	"Machine Learning"	"Optical Communication" - "Optical Network"	TITLE-ABS-KEY ( ( "Machine Learning" ) AND ( "Optical Communication" OR "Optical Network" ) )	841

Validación de cadenas de búsqueda	Si	Incluido en los resultados		
	No	No incluido en los resultados.		
	Título	Autores	I iteración	Final
Primer set		Tasa de Éxito:	10/10	10/10
1	A QoT prediction technique based on machine learning and NLSE for QoS and new lightpaths in optical communication networks	Yongfeng FU , Jing CHEN, Weiming WU, Yu HUAN2, Jie HONG, Long CHEN, Zhongbin LI	Si	Si
2	An Overview on Application of Machine Learning Techniques in Optical Networks	Francesco Musumeci , Cristina Rottondi , Avishek Nag , Irene Macaluso, Darko Zibar , Marco Ruffini, Massimo Tornatore	Si	Si
3	A Tutorial on Machine Learning for Failure Management in Optical Networks	Francesco Musumeci , Cristina Rottondi , Giorgio Corani, Shahin Shakhkarami, Filippo Cugini, and Massimo Tornatore	Si	Si
4	Machine learning for intelligent optical networks: A comprehensive survey	Rentao Gu , Zeyuan Yang, Yuefeng Ji	Si	Si
5	Overview on routing and resource allocation based machine learning in optical networks	Yongjun Zhang, Jingjie Xin, Xin Li*,Shanguo Huang	Si	Si
6	Autonomous Operations in Optical Networks	Alba P. Vela, Marc Ruiz, and Luis Velasco	Si	Si
7	Examples of Machine Learning Algorithms for Optical Network Control and Management	Alba P. Vela, Marc Ruiz, and Luis Velasco	Si	Si
8	DeepALM: Holistic Optical Network Monitoring based on Machine Learning	Joo Yeon Cho, Jose-Juan Pedreno-Manresa, Sai Kireet Patri, Khoulood Abdelli, Carsten Tropschug, Jim Zou, Piotr Rydlichowski	Si	Si
9	Machine Learning Applications for Short Reach Optical Communication	Yapeng Xie , Yitong Wang, Sithamparanathan Kandeepan and Ke Wang	Si	Si
10	Machine learning for quality of transmission: a picture of the benefits fairness when planning WDM networks	Matteo Lonardi,Jelena Pesic,Thierry Zami,Emmanuel Seve,Nicola Rossi	Si	Si
Análisis de Resultados				
I Iter - Final				
Conclusión		Todos los articulos forman parte de la cadena de búsqueda definida.		
Cadena de búsqueda final				
TITLE-ABS-KEY ( ( "Machine Learning" ) AND ( "Optical Communication" OR "Optical Network" ) )				

### DOCUMENTO 3: Identificación de criterios para procedimiento de inclusión - exclusión automatizado.

Los siguientes criterios se aplicarán utilizando las funciones de Scopus:

#### Criterios de inclusión

De acuerdo con la documentación de Scopus [1], los artículos en esta base de datos son revisados por pares.

**1. Tipo de documento:** Documento de conferencia y artículos.

**2. Idioma:** inglés.

**CONSULTA ADICIONAL:** Número de citas para selección de artículos (en el caso de ser utilizadas).

**Tabla I.** Número de citas por percentiles como para Thomson Reuters [2].

Engineering												
	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	All years

0.01%	808	961	688	530	463	368	329	300	154	68	18	504
0.10%	329	330	275	224	197	160	130	99	63	29	9	200
1.00%	125	117	103	88	78	66	54	40	26	11	3	72
10.00%	37	35	32	29	26	23	18	14	9	3	1	20
20.00%	22	21	19	18	16	14	12	9	5	2	1	12
50.00%	7	7	7	6	6	5	4	3	2	1	1	3

**Tabla II.** Reglas para filtrar por número de citas.

Año de Publicación	Citas mínimas
Antes 2015	7
2015 o 2016	6
2017	5
2018	4
2019	3
2020	2
2021	1
2022	0

El trabajo debe cumplir con todos los criterios de inclusión y exclusión automatizada definidos antes de pasar a los criterios de inclusión y exclusión manual.

## Referencias

[1] [https://www.elsevier.com/\\_data/assets/pdf\\_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf](https://www.elsevier.com/_data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf)

[2] <https://esi.incites.thomsonreuters.com/BaselineAction.action>.

## DOCUMENTO 4: Identificación de criterios para procedimiento de inclusión - exclusión manual.

### 1. Información General

La siguiente información (Tabla I) permite a los investigadores comprender las generalidades del procedimiento manual de inclusión o exclusión de trabajos (screening manual).

**Tabla I.** Información General

<b>Objetivo</b>	Descartar aquellos trabajos que no sean de interés para este estudio.
<b>Instrucciones Generales</b>	<ul style="list-style-type: none"> <li>▪ El equipo de investigación estará compuesto por 2 screeners y 1 supervisor.</li> <li>▪ El equipo de investigación trabajará en un piloto con el fin de armonizar criterios.</li> <li>▪ El coeficiente de confiabilidad entre codificadores a utilizar es el alfa de Krippendorff [1].</li> </ul>



<b>Criterios de Inclusión</b> (Todo debe cumplirse)	<ul style="list-style-type: none"> <li>▪ El artículo es una <b>contribución primaria</b>.</li> <li>▪ La contribución reportada en el artículo está relacionada con el funcionamiento y operación de las redes ópticas.</li> <li>▪ El artículo incluye al menos una contribución que proponga una técnica de ML aplicada a la resolución de algún problema reportado en [2] - [4].</li> </ul>
<b>Criterio de Exclusión</b> (ninguno debe cumplirse)	<ul style="list-style-type: none"> <li>▪ El trabajo reporta un estudio secundario o terciario.</li> <li>▪ La contribución no se centra en el funcionamiento y operación de la redes ópticas.</li> <li>▪ La contribución reportada aborda algún problema reportado en [2] - [4], pero no emplea una técnica de ML.</li> </ul>
<b>Posibles opciones de marca</b>	<ul style="list-style-type: none"> <li>▪ [I]ncluido -&gt; El trabajo cumple con todos los criterios de inclusión y ninguno de exclusión.</li> <li>▪ [E]xcluido -&gt; El artículo no cumple con un criterio de inclusión o cumple con uno o más criterios de exclusión. <ul style="list-style-type: none"> <li>○ [S]ecundario: El trabajo reporta una contribución secundaria o terciaria.</li> </ul> </li> <li>▪ [U]nclaro -&gt; Hay dudas para clasificar el trabajo, por lo que se hará más adelante.</li> </ul>

### 1. Frase de screening

Habrán dos fases en la etapa de screening: 1) una fase piloto para normalizar los criterios de los evaluadores y 2) una fase principal que consiste en una selección basada en los títulos y resúmenes de los artículos y finalmente una selección basada en el texto completo. La Tabla II presenta los detalles de estas dos fases.

**Tabla II.** Descripción de las fases piloto y principal.

<b>Fase Piloto</b>	<ol style="list-style-type: none"> <li>1. La fase piloto será un proceso n-iterativo hasta alcanzar un coeficiente de fiabilidad de 0,8 (alfa de Krippendorff \autocite{Krip}) entre tutor y estudiante.</li> <li>2. Se analizan 5 artículos: las secciones a leer son el título y el resumen.</li> <li>3. Los resultados serán discutidos por tutor y estudiante para armonizar los criterios de los evaluadores antes de la siguiente iteración.</li> <li>4. Se finaliza el proceso piloto, solo si tutor y estudiante logran obtener un coeficiente de concordancia por encima del umbral de “buena” concordancia, es decir, 0.8.</li> </ol>
<b>Fase Principal</b>	<ul style="list-style-type: none"> <li>▪ Los papeles restantes, es decir, los que no se procesaron en el piloto, se procesan de la siguiente manera: <ul style="list-style-type: none"> <li>• En la primera iteración, los evaluadores (tutor y estudiante) trabajarán individualmente leyendo los títulos y resúmenes de los trabajos y marcándolos como incluidos, excluidos o poco claros. El 20% de los artículos son revisados por el tutor y el estudiante a cargo de la investigación. Mientras que, el 80% de artículos restante son analizados exclusivamente por el estudiante. Durante esta fase se definirán paradas para el cálculo y aseguramiento del coeficiente de fiabilidad entre el equipo. Las divergencias entre los artículos revisados por el tutor y estudiante son discutidas y resueltas por el equipo en reuniones planificadas.</li> <li>• En la segunda iteración, los evaluadores volverán a leer</li> </ul> </li> </ul>



	<p>individualmente los artículos de la primera iteración, esta vez siguiendo una estrategia adaptativa hacia la elección del nivel de detalle, es decir, el título, el resumen y las conclusiones son obligatorios, pero si esta información no es clara, se pueden considerar otras partes del documento (por ejemplo, introducción, títulos de sección y subsección). En este caso, solo se elige entre incluidos o excluidos. El 20\% de los artículos son revisados por el tutor y el estudiante a cargo de la investigación. Mientras que, el 80\% de artículos restante son analizados exclusivamente por el estudiante. Las divergencias entre los artículos revisados por el tutor y estudiante son discutidas y resueltas por el equipo en reuniones planificadas.</p>
--	---

## 2. Procedimiento detallado para la selección de cada trabajo

Este procedimiento (Tabla III) se seguirá para la lectura de cada trabajo.

**Tabla III.** Procedimiento de selección.

Paso	Proceso
1	<p><b>Lea las secciones definidas según sea el procedimiento piloto o principal. Mientras lee, preste atención a los siguientes términos:</b></p> <ul style="list-style-type: none"> <li>▪ Optical Network and Machine Learning terms: optical communication, artificial intelligence.</li> <li>▪ Optical network problems related terms: resource allocation, resource management, optical performance monitoring, traffic prediction, traffic classification, congestion control, fault management, QoS, QoE management and QoT.</li> </ul>
2	<p><b><u>Screening basada en títulos y resumen</u></b></p> <p>¿El título describe una investigación primaria e incluye al menos una contribución que describe una técnica de ML aplicada a la resolución de algún problema reportado en [2]- [4] relacionado con el funcionamiento y operación de las redes ópticas?</p> <p>Seleccione <b>sí, no o no está claro</b> si se queda atascado aquí.</p> <ul style="list-style-type: none"> <li>➤ Si la respuesta es "no", marque todos los criterios de la selección basada en resúmenes como "no" y pase al siguiente trabajo.</li> <li>➤ Si la respuesta es "sí" o "no está claro", vaya al Paso 3</li> </ul>
1	<p><b><u>Screening basada en texto completo</u></b></p> <p>¿El artículo es una investigación primaria?</p> <p>Seleccione <b>sí, no o no está claro</b> si se queda atascado aquí</p> <ul style="list-style-type: none"> <li>➤ Si la respuesta es "no", marque los criterios restantes como "no" y pase al siguiente trabajo.</li> <li>➤ Si la respuesta es "sí" o "no está claro", vaya al Paso 4</li> </ul>
2	<p>¿El artículo propone una contribución que está relacionada con el funcionamiento y operación de las redes ópticas?</p> <p>Seleccione <b>sí, no o no está claro</b> si se queda atascado aquí</p> <ul style="list-style-type: none"> <li>➤ Si la respuesta es "no", marque los criterios restantes como "no" y pase al siguiente trabajo.</li> <li>➤ Si la respuesta es "sí" o "no está claro", vaya al Paso 5</li> </ul>

3	¿El artículo incluye al menos una contribución que proponga una técnica de ML aplicada a la resolución de algún problema reportado en [2]- [4]? Selecione sí, no o no está claro si se queda atascado aquí. Ir al siguiente trabajo.
---	---

### 3. Referencias

[1] K. Krippendorff, "Testing the reliability of content analysis data: What is involved and why," in *The Content Analysis Reader*, USA: Sage Publications, 2009, pp. 350–357.

[2] R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *Journal of Internet Services and Applications*, vol. 9, no. 1, Dec. 2018, doi: 10.1186/s13174-018-0087-2.

[3] R. Gu, Z. Yang, and Y. Ji, "Machine learning for intelligent optical networks: A comprehensive survey," *Journal of Network and Computer Applications*, vol. 157. Academic Press, May 01, 2020. doi: 10.1016/j.jnca.2020.102576.

[4] Y. Zhang, J. Xin, X. Li, and S. Huang, "Overview on routing and resource allocation based machine learning in optical networks," *Optical Fiber Technology*, vol. 60, Dec. 2020, doi: 10.1016/j.yofte.2020.102355.

### DOCUMENTO 5: Esquema de clasificación y extracción de datos.

Este documento presenta los datos a ser extraídos, el procedimiento de extracción y el esquema de clasificación utilizado en dicho proceso. La Tabla I muestra información general para este estudio.

**Tabla I.** Información General.

<b>Recursos Humanos</b>	Consejeros:	1 experto en investigación de ML y Comunicaciones Ópticas.
	Número de codificadores:	2 (número mínimo de codificadores mencionado en [1], [2]).
	Calificación del codificador:	Estudiante de pregrado investigando en el dominio óptico.
<b>Unidad de Análisis</b>	Unidad de análisis:	Artículo.
	Cantidad total de artículos para leer:	Aquellos trabajos incluidos en la etapa de selección. Cada artículo será leído por dos codificadores.
	Partes del documento para leer:	Obligatorio y una lectura profunda adaptativa para la clasificación, es decir, <b>el título, el resumen, la introducción y las conclusiones son obligatorios</b> , pero si esta información no es clara, lea las secciones en las que se explica la técnica.
<b>Herramientas</b>	Herramienta para codificar:	Hojas de calculo

#### 1. Preguntas de investigación

Las preguntas de investigación que guían el proceso son:

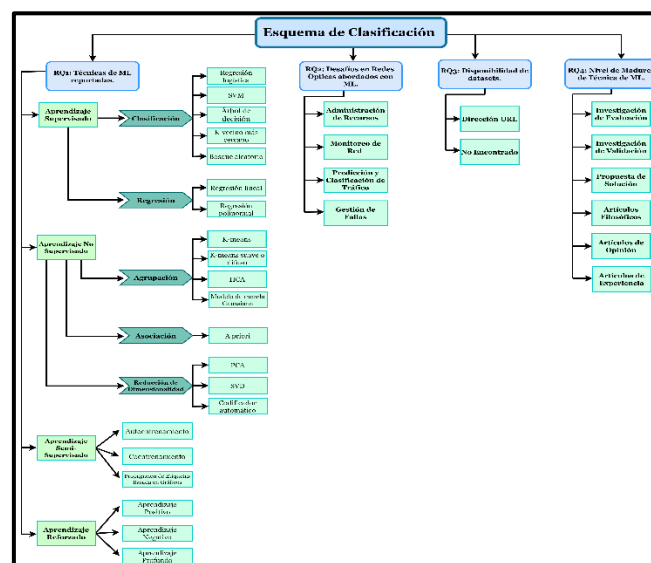
- **RQ1. - Tipo de técnica:** ¿Cuáles son las técnicas de ML utilizadas para resolver los problemas reportados en [3]–[5] relacionados con el funcionamiento y operación de las redes ópticas?
- **RQ2. – Tipo de desafíos:** ¿Qué desafíos informados en [9], [32], [33] han sido mayoritariamente abordados mediante el uso de las técnicas de ML reportadas?

- **RQ3. - Disponibilidad de datasets:** ¿Los datasets que han sido utilizados por las técnicas de ML en los artículos revisados están disponibles?
- **RQ4. - Nivel de madurez:** ¿Cuál es el nivel de madurez de las técnicas de ML identificadas?

Los datos se extraerán teniendo en cuenta que necesitamos responder a las preguntas de investigación. Estos datos incluyen la identificación, datos para la lectura y otros útiles. Los datos que se muestran en la Tabla II se extraerán automáticamente de la metainformación de los artículos.

Información	Datos a extraer del artículo
Información de identificación	ID (Asignado después de las tareas de inclusión-exclusión automatizadas en Scopus).
	DOI (Obtenido de Scopus)
	URL
	Autores
Datos para leer	El título, resumen, introducción y conclusiones son obligatorios, pero si esta información no es clara, <b>lea la sección y figuras en las que se explica la técnica.</b>
Información útil	Número de citas
	Año
	Tipo de documento (documento de conferencia o artículo)
	Título de la fuente (Indica el nombre de la revista)
	Nombre de la conferencia
	Afiliaciones

### 3. Clasificaciones para codificar



### 3.1 Técnicas de Aprendizaje Automático reportadas

El aprendizaje automático es un área de investigación que se encarga de enseñar a las máquinas a realizar una actividad similar a la mente humana. Aunque su capacidad cognitiva es mucho más limitada que la del ser humano pueden procesar grandes cantidades de información en corto tiempo y obtener información útil [3]. Según [4]–[6] las técnicas de aprendizaje automático que son comúnmente utilizadas en Redes de Comunicaciones Ópticas son Aprendizaje Supervisado, Aprendizaje No Supervisado, Aprendizaje Semi-Supervisado, Aprendizaje Reforzado y Aprendizaje Auto-Supervisado.

**Tabla III.** Técnicas de Aprendizaje Automático aplicadas a Redes Ópticas.

Técnica de Aprendizaje Automático	Descripción
Aprendizaje Automático Supervisado	Es una subcategoría de AI/ML que emplea conjuntos de datos etiquetados para entrenar algoritmos que clasifican información o predicen resultados con precisión. A medida que los datos de entrada ingresan en el modelo, este ajusta sus pesos hasta que el algoritmo opere de manera correcta [7].
Aprendizaje Automático No Supervisado	Es una subcategoría de AI/ML que permite que una máquina explore un conjunto de datos sin la necesidad de supervisar el modelo, por lo que es mucho más sencillo de implementar en comparación con el aprendizaje supervisado, ya que el modelo funciona de forma independiente [3].
Aprendizaje Automático Semi-Supervisado	Es una subcategoría de AI/ML, que combina técnicas de aprendizaje no supervisado y supervisado. Durante la etapa de entrenamiento, en lugar de agregar etiquetas a todo el conjunto de datos analizados, se revisa y se etiqueta a mano solo una pequeña parte de los datos para entrenar un modelo [8].
Aprendizaje Automático Reforzado	Es una subcategoría de AI/ML en la que un agente aprende a comportarse en un medio realizando las acciones y viendo los resultados de las acciones. Por cada buena acción, el agente recibe una retroalimentación positiva, y por cada mala acción, el agente recibe una retroalimentación negativa o penalización [3].
Aprendizaje Automático Auto-Supervisado	Es una subcategoría de AI/ML que permite aprender patrones complejos a partir de datos no etiquetados, es decir, cuando el modelo se alimenta con datos no estructurados como entrada, genera etiquetas de datos automáticamente, que se utilizan en iteraciones posteriores como verdades básicas [9].

#### 3.1.1 Subclasificación de las Técnicas de Aprendizaje Automático

Cada técnica de aprendizaje automático a su vez posee su propia subclasificación, que se describe en la Tabla IV.

**Tabla IV.** Subclasificación de Técnicas de Aprendizaje Automático aplicadas a Redes Ópticas

Técnica de Aprendizaje Automático	Subclasificación de Técnica de Aprendizaje Automático	Descripción	Algoritmos	Descripción
Aprendizaje Automático Supervisado	Clasificación	Los algoritmos de clasificación pueden explicar o predecir un valor de clase, es decir, utilizan un algoritmo para asignar con precisión los datos de prueba en categorías específicas [3].	Regresión Logística	La regresión logística estima la probabilidad de que ocurra un evento en función de una o más entradas [10]. La regresión logística se utiliza cuando la variable dependiente tiene dos salidas, es decir, "verdadero" y "falso" o "sí" y "no". Se usa principalmente para resolver problemas de clasificación binaria, como la identificación de spam [7].

			<b>SVM</b>	Máquina de vectores de soporte (del inglés Support-Vector Machines, SVM) es un modelo de aprendizaje supervisado que se utiliza tanto para la clasificación como para la regresión de datos. Este algoritmo construye un hiperplano donde la distancia entre dos clases de puntos de datos es máxima. El hiperplano se conoce como "límite de decisión" y separa las clases de puntos de datos a cada lado del plano [7].
			<b>Árbol de Decisión</b>	Es un modelo de aprendizaje supervisado en el que los datos analizados se dividen continuamente de acuerdo con un parámetro determinado. El árbol de decisión está constituido por dos entidades: nodos de decisión y hojas. Las hojas son las decisiones o los resultados finales, mientras que, los nodos de decisión son en donde se dividen los datos [11].
			<b>K-vecino más cercano</b>	Es un modelo de aprendizaje supervisado que clasifica los puntos de datos en función de su proximidad y asociación con otros datos disponibles. Este algoritmo busca calcular la distancia entre los puntos de datos, generalmente a través de la distancia euclidiana, y luego asigna una clase en función de la categoría o promedio más frecuente [7].
			<b>Bosque Aleatorio</b>	Es un modelo de aprendizaje supervisado que se utiliza tanto para la clasificación como para la regresión de datos. El "bosque" es la representación de una colección de árboles de decisión no correlacionados, que luego se fusionan para reducir la varianza y crear predicciones de datos más precisas [7].
	<b>Regresión</b>	Los métodos de regresión se definen como un proceso para encontrar las correlaciones entre las variables dependientes e independientes [12]. Permite predecir o explicar un valor numérico particular en función de un conjunto de datos anteriores.	<b>Regresión Lineal</b>	Es el algoritmo de regresión más simple. Se lo utiliza para para identificar la relación entre una variable dependiente y una o más variables independientes, permitiendo realizar predicciones sobre resultados futuros. Para cada tipo de regresión lineal se busca trazar una línea de mejor ajuste, la cual se calcula mediante el método de mínimos cuadrados, permitiendo de esta manera realizar predicciones para nuevos datos no vistos [7].
			<b>Regresión Polinomial</b>	Es un algoritmo de regresión en la que la relación entre las variables independientes y las variables dependientes se modelan a través de un polinomio de grado n. Es un caso especial de regresión lineal donde se ajusta la ecuación polinomial en los datos, con una relación curvilínea entre las variables dependientes e independientes [13].
	<b>Agrupación</b>	Los algoritmos de Agrupación o <i>Clustering</i> son técnicas que tratan principalmente de encontrar una estructura o patrón en una colección de datos no etiquetados en función de sus similitudes o diferencias [14].	<b>K-means</b>	Es un algoritmo que divide el conjunto de datos analizados en subgrupos distintos no superpuestos (clústeres) predefinidos por K, donde cada punto de datos pertenece a un solo grupo. K-means intenta hacer que los puntos de datos dentro del clúster sean lo más similares posible y, al mismo tiempo, trata de mantener los clústeres lo más diferentes posible [15].
			<b>K-means suave o difuso</b>	Es un algoritmo similar a K-means. Su única diferencia es que, en lugar de asignar un punto exclusivamente a un solo grupo (clúster), puede haber algún tipo de borrosidad o superposición entre dos o más grupos [14], es decir, los puntos de datos pueden pertenecer a múltiples grupos con distintos grados de pertenencia [17].
			<b>HCA</b>	Análisis de Agrupación jerárquica (del inglés Hierarchical Cluster Analysis, HCA) es un algoritmo que tiene por objetivo asociar elementos o registros que están "cerca" entre sí formando agrupaciones jerárquicas. Puede trabajar de dos formas: aglomerante o divisiva. En la agrupación aglomerante, sus puntos de datos se aíslan inicialmente y luego se fusionan iterativamente según su similitud hasta que se logra un grupo. La agrupación divisiva se puede definir como lo opuesto a la agrupación aglomerante. En este caso, un solo grupo de datos se divide en función de las diferencias entre los puntos de datos analizados [17].
			<b>Modelo de Mezcla Gaussiana</b>	Es un algoritmo que se emplea para clasificar los datos en diferentes categorías según la distribución de probabilidad. Supone que todos los puntos de datos se generan a partir de una combinación de distribuciones gaussianas con parámetros desconocidos [18].
<b>Aprendizaje Automático No Supervisado</b>	<b>Asociación</b>	Los Algoritmos de Asociación son técnicas basadas en reglas que permiten encontrar relaciones no obvias entre variables en un conjunto de datos determinado, por lo que puede considerarse como una herramienta de análisis de datos [14].	<b>A priori</b>	Son técnicas que siguen una secuencia de pasos para encontrar el conjunto de elementos más frecuente en la base de datos dada. Se utilizan dentro de conjuntos de datos transaccionales para identificar conjuntos de elementos frecuentes, o colecciones de elementos, para identificar la probabilidad de consumir un producto dado el consumo de otro producto [17].
	<b>Reducción de la</b>	Los algoritmos de Reducción de Dimensionalidad son técnicas que se emplean cuando el número de	<b>PCA</b>	Análisis de Componentes Principales (del inglés Principal Component Analysis, PCA) es un algoritmo que se utiliza para reducir redundancias y comprimir conjuntos de datos a través de la extracción de características. Utiliza

	<b>Dimensionalidad</b>	características en un conjunto de datos determinado es muy alto. Durante su funcionamiento, estos algoritmos reducen la cantidad de entradas de datos a un tamaño adecuado y al mismo tiempo preservan la integridad del conjunto de datos [17].		transformaciones lineales para crear nuevas representaciones de datos, generando conjuntos de "componentes principales" [17].
			<b>SVD</b>	Descomposición en Valores Singulares (del inglés, Singular Value Decomposition, SVD) es un algoritmo que factoriza una matriz principal en tres matrices de bajo rango. Se usa comúnmente para reducir el ruido y comprimir datos, como archivos de imagen [17].
			<b>Codificador Automático</b>	Son algoritmos que aprovechan las redes neuronales para comprimir datos y luego recrear una nueva representación de la entrada de datos original [17].
<b>Aprendizaje Automático Semi-Supervisado</b>	<b>Autoentrenamiento</b>	El autoentrenamiento (Self-training) es una de las técnicas más simples de aprendizaje semi-supervisado. Se puede escoger cualquier técnica de aprendizaje supervisado de clasificación o regresión y modificarla para que funcione de forma semi-supervisada [19].		
	<b>Co-entrenamiento</b>	El co-entrenamiento (Co-training) es una técnica de aprendizaje semi-supervisado derivada de los algoritmos de aut-entrenamiento, que se utiliza cuando solo se dispone de un pequeño conjunto de datos etiquetados [19].		
	<b>Propagación de Etiquetas basada en Gráficos</b>	La propagación de etiquetas es una de las formas más populares de emplear técnicas de aprendizaje, ya que, representar datos etiquetados y no etiquetados en forma gráfica y luego aplica un algoritmo de propagación de etiquetas que difunde la información a través de toda la red de datos [19].		
<b>Aprendizaje Automático Reforzado</b>	<b>Positivo</b>	El aprendizaje por refuerzo positivo se define como un evento que ocurre debido a un comportamiento específico. Aumenta la fuerza y la frecuencia del comportamiento e impacta de manera positiva en la acción tomada por el agente. Este tipo de algoritmo ayuda a maximizar el rendimiento y mantener el cambio durante un período prolongado [20].		
	<b>Negativo</b>	El aprendizaje por refuerzo negativo se define como un comportamiento que se produce debido a una condición negativa que debería haberse evitado o detenido. Este tipo de algoritmo ayuda a determinar el estándar mínimo de rendimiento. El refuerzo negativo puede ser una herramienta eficaz cuando se usa correctamente [20].		
	<b>Profundo</b>	El aprendizaje reforzado profundo está conformado por un agente inteligente que aprende a optimizar un proceso de decisión. Si el resultado de esa decisión es favorable, el agente automáticamente repetirá esa decisión en el futuro, mientras que, si el resultado es perjudicial, evitará volver a tomar la misma decisión [21].		

### 3.2 Desafíos en Comunicaciones Ópticas abarcados con Aprendizaje Automático.

Según [4]–[6], las áreas de trabajos en las que se podría aplicar técnicas de ML para solventar problemas relacionados al funcionamiento y desempeño de Redes Ópticas son descritas en la Tabla V.

**Tabla V.** Desafíos en Comunicaciones Ópticas.

<b>Desafíos</b>	<b>Descripción</b>
<b>Gestión de Recursos</b>	La gestión de recursos es el proceso de administración y asignación de recursos de una red. Permite crear un método efectivo para analizar el tráfico que recorre a través de una red, reduciendo así el tiempo de inactividad, errores y otros problemas que se puedan suscitar, mientras que permite satisfacer las necesidades de los usuarios y dispositivos al mismo tiempo [22], [23].
<b>Monitoreo de Red</b>	El monitoreo de una red implica la gestión de los enlaces entre dispositivos, estaciones de trabajo, servidores, dispositivos virtuales y dispositivos móviles con el fin de identificar la congestión, maximizar y mejorar el rendimiento de la red para el usuario. Con el uso de herramientas de monitoreo de red automatizadas, los analistas pueden recopilar datos, identificar y medir variables de rendimiento, realizar evaluaciones de rendimiento y diagnosticar problemas de rendimiento de una determinada red [24], [25].
<b>Predicción y Clasificación de Tráfico</b>	La predicción del tráfico de red tiene como objetivo predecir el tráfico de red posterior utilizando los datos de tráfico de red anteriores. Esto puede servir como un enfoque proactivo para la administración de redes, tareas de planificación, procesos de seguridad, evitar la congestión y aumentar la velocidad de las redes [26], [27]. Por otra parte, la clasificación de tráfico ofrece la capacidad de reconocer automáticamente la aplicación que ha generado un determinado flujo de



	paquetes a partir de la observación directa y pasiva de los paquetes individuales, o flujo de paquetes, que fluyen en la red. Cada clase de tráfico identificada puede tratarse de manera diferente, lo que resulta fundamental para una serie de actividades que son de interés para los operadores, proveedores de servicios de Internet y administradores de red en general [28], [30].
<b>Gestión de Fallas</b>	La gestión de fallas de red es el proceso de encontrar, aislar y solucionar errores de red en el menor tiempo posible. La gestión de fallas es un componente de fundamental importancia en la administración de una red, debido a que minimiza el tiempo de inactividad y evita contratiempos en los dispositivos al resolver las fallas detectadas en corto tiempo, garantizando así una disponibilidad óptima de la red y evitando pérdidas comerciales [31].

### 3.3 Disponibilidad de dataset

En la tabla VI, permite identificar si existe la disponibilidad o no de datasets en los papers que serán codificados.

**Tabla VI.** Tipos de disponibilidad de datasets.

<b>Tipo de Disponibilidad</b>	<b>Descripción</b>
Dirección URL	URL (Uniform Resource Locator) es la dirección única y específica que se asigna a cada uno de los recursos (datasets) disponibles en la World Wide Web para que puedan ser localizados por el navegador, visitados y descargados por los usuarios [32].
No Encontrado	Se refiere a que no existe la disponibilidad de la dirección URL de los datasets para que puedan ser localizados por el navegador, visitados y descargados por los usuarios.

### 3.4 Tipo de Investigación

La Tabla VII tiene una descripción de cada tipo de investigación [33]. La Figura 2 proporciona más ayuda para identificar el tipo de investigación. Para utilizar la Figura 2, el codificador debe considerar el cumplimiento de las condiciones (aquellas con V o F) en una columna completa y luego elegir el tipo de investigación (la marcada) en la misma columna. Por ejemplo, si un artículo propone una solución novedosa y no incluye una evaluación empírica u opinión sobre algo, será una Propuesta de solución.

**Tabla VII.** Tipos de investigación propuestos en [33].

<b>Tipo de Investigación</b>	<b>Descripción</b>
Investigación de Evaluación	“Las técnicas se implementan en la práctica y se realiza una evaluación de la técnica. Es decir, se muestra cómo se implementa la técnica en la práctica (implementación de la solución) y cuáles son las consecuencias de la implementación en términos de beneficios y desventajas (evaluación de la implementación). Esto también incluye identificar problemas en la industria”.
Investigación de Validación	“Las técnicas investigadas son novedosas o una actualización significativa y aún no se han implementado en la práctica. Las técnicas utilizadas son, por ejemplo, experimentos, es decir, trabajos realizados en el laboratorio”.
Propuesta de solución	“Se propone una solución para un problema, la solución puede ser novedosa o una extensión significativa de una técnica existente. Los beneficios



	potenciales y la aplicabilidad de la solución se muestran con un pequeño ejemplo o una buena línea de argumentación”.
Artículos filosóficos	“Estos artículos esbozan una nueva forma de ver las cosas existentes al estructurar el campo en forma de taxonomía o marco conceptual”.
Artículos de Opinión	“Estos trabajos expresan la opinión personal de alguien sobre si una determinada técnica es buena o mala, o cómo se deben hacer las cosas. No se basan en metodologías de trabajo e investigación relacionadas”.
Documentos de experiencia	“Los documentos de experiencia explican qué y cómo se ha hecho algo en la práctica. Tiene que ser la experiencia personal del autor”.

Research type classification (T = True, F = False, • = irrelevant or not applicable, R1–R6 refer to rules).						
	R1	R2	R3	R4	R5	R6
<i>Conditions</i>						
Used in practice	T	•	T	F	F	F
Novel solution	•	T	F	•	F	F
Empirical evaluation	T	F	F	T	F	F
Conceptual framework	•	•	•	•	T	F
Opinion about something	F	F	F	F	F	T
Authors' experience	•	•	T	•	F	F
<i>Decisions</i>						
Evaluation research	✓	•	•	•	•	•
Solution proposal	•	✓	•	•	•	•
Validation research	•	•	•	✓	•	•
Philosophical papers	•	•	•	•	✓	•
Opinion papers	•	•	•	•	•	✓
Experience papers	•	•	✓	•	•	•

**Figura 2.** Condiciones para identificar los tipos de investigación propuestos en [34].

Al codificar el tipo de investigación, tenga en cuenta también los siguientes consejos:

- **Filosófico** -> Si muestra ideas de alto nivel **sin dar detalles**.
- **Propuesta de solución** -> Si **da detalles** aportando una **línea de argumentación** o un **ejemplo**.
  - Si hay una **herramienta involucrada**, codifique el documento **desde "propuesta de solución" hacia arriba**.
- **Investigación de validación** -> Si una técnica se aplica en un **entorno experimental**, **incluso utilizando un sistema de software real**, y la sección de resultados sugiere que es una **técnica en evolución** (por ejemplo, mostrando su eficacia en comparación con otras técnicas, etc.).
- **Investigación de evaluación** -> Si una técnica se **aplica en sistemas de software reales** y presenta una **sección explícita** (por ejemplo, discusión, lección aprendida, etc.) que evalúa la técnica y **sugiere que es una técnica madura**.
  - Si el documento solo presenta los resultados de un sistema de software real evaluado, pero **no presenta explícitamente una sección que discuta la evaluación de la técnica**, debe codificarse como **investigación de validación**.
  - Tenga cuidado con la palabra clave **"mundo real"**, ya que podría ser solo un "ejemplo" de escenarios del mundo real para "demostrar" una técnica. En este caso, el documento debe codificarse como una **propuesta de solución** en lugar de una investigación de evaluación.

#### 4. Procedimiento de codificación

Este procedimiento lo realizará cada codificador para clasificar las técnicas de Aprendizaje Automático aplicados a resolver problemas relacionados al funcionamiento y desempeño de Redes Ópticas, la disponibilidad de datasets y el nivel de madurez de cada técnica empleada (tipo de investigación) de cada trabajo dentro de un grupo proporcionado por el investigador a cargo.

#### 4.1 Configuración de materiales para la ejecución de la codificación.

El **material** a proporcionar a los codificadores incluye:

- Este **archivo de codificación**.
- Una **hoja de cálculo** con los datos necesarios para codificar el grupo de artículos, incluye los datos para su identificación y lectura. Adicionalmente, se incluirá una copia de los trabajos para la lectura de la sección de conclusiones.
- Un **formulario electrónico** para codificar cada trabajo.

#### 4.2 Ejecución de codificación

##### 4.1.1. Revisión de la hoja de cálculo

El codificador verificará que los datos estén completos para cada papel del grupo que se está procesando.

##### 4.1.2. Procedimiento de lectura y codificación (por cada artículo)

El codificador:

- A. Escribirá su nombre, cédula y título del trabajo en el formulario
- B. Realizará una **lectura de profundidad tanto obligatoria como adaptativa** para la clasificación, es decir, el **título, el resumen, la introducción y las conclusiones son obligatorios**, pero si esta información no es clara, lea las secciones en las que se explica la técnica. Cuando los artículos no tienen conclusiones, el codificador debe leer algunas secciones equivalentes como resumen o pensamientos finales.
- C. Codificará el trabajo de acuerdo con los elementos mencionados (Tabla III – VII). Se proporcionará un formulario electrónico para guiar esta tarea.
- D. Repetirá el proceso para todos los papeles.

#### 4.3 Reunión de codificadores para discutir resultados

Después de codificar el grupo de artículos, los codificadores se reunirán para revisar los resultados. Para cada artículo, deben:

- A. Compara los resultados.
  - a. Si los valores están igualmente codificados, se mantienen
  - b. Si los valores no están codificados por igual, los codificadores deben estar de acuerdo en el valor final.
- B. Los codificadores registrarán un nuevo registro para cada trabajo. En este caso, el autor es el “equipo”.

#### 5. Referencias

[1] K. Krippendorff, “Testing the Reliability of Content Analysis Data: What is Involved and Why,” in *The Content Analysis Reader*, 2009, pp. 202–208.

[2] M. Lombard, J. Snyder-Duch, and C. C. Bracken, “Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects.” 2010.

- [3] "TOP 5 Machine Learning Algorithms For Business Applications." <https://mobidev.biz/blog/5-essential-machine-learning-techniques> (accessed Apr. 13, 2022).
- [4] R. Boutaba et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *Journal of Internet Services and Applications*, vol. 9, no. 1, Dec. 2018, doi: 10.1186/s13174-018-0087-2.
- [5] R. Gu, Z. Yang, and Y. Ji, "Machine learning for intelligent optical networks: A comprehensive survey," *Journal of Network and Computer Applications*, vol. 157. Academic Press, May 01, 2020. doi: 10.1016/j.jnca.2020.102576.
- [6] Y. Zhang, J. Xin, X. Li, and S. Huang, "Overview on routing and resource allocation based machine learning in optical networks," *Optical Fiber Technology*, vol. 60, Dec. 2020, doi: 10.1016/j.yofte.2020.102355.
- [7] "What is Supervised Learning? | IBM." <https://www.ibm.com/cloud/learn/supervised-learning> (accessed Apr. 13, 2022).
- [8] "¿Qué es machine learning? - España | IBM." <https://www.ibm.com/es-es/cloud/learn/machine-learning> (accessed May 02, 2022).
- [9] "Self-Supervised Learning: Definition, Tutorial & Examples." <https://www.v7labs.com/blog/self-supervised-learning-guide#h1> (accessed Jul. 14, 2022).
- [10] "10 Machine Learning Methods that Every Data Scientist Should Know | by Jorge Castañón | Towards Data Science." <https://towardsdatascience.com/10-machine-learning-methods-that-every-data-scientist-should-know-3cc96e0e0000> (accessed Apr. 13, 2022).
- [11] "Xoriant." <https://www.xoriant.com/blog/product-engineering/decision-trees-machine-learning-algorithm.html> (accessed Apr. 13, 2022).
- [12] "Difference Between Classification and Regression in Machine Learning." <https://machinelearningmastery.com/classification-versus-regression-in-machine-learning/> (accessed Apr. 13, 2022).
- [13] "Understanding Polynomial Regression!!! | by Abhigyan | Analytics Vidhya | Medium." <https://medium.com/analytics-vidhya/understanding-polynomial-regression-5ac25b970e18> (accessed Apr. 13, 2022).
- [14] "Unsupervised Machine Learning: Algorithms, Types with Example." <https://www.guru99.com/unsupervised-machine-learning.html> (accessed May 02, 2022).
- [15] "K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks | by Imad Dabbura | Towards Data Science." <https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a> (accessed Oct. 06, 2022).
- [16] "Fuzzy K-Means Clustering in Mahout." [https://www.edureka.co/blog/fuzzy\\_k-means/](https://www.edureka.co/blog/fuzzy_k-means/) (accessed Oct. 06, 2022).
- [17] "What is Unsupervised Learning? | IBM." <https://www.ibm.com/cloud/learn/unsupervised-learning> (accessed May 02, 2022).
- [18] "Gaussian Mixture Models: What are they & when to use? - Data Analytics." <https://vitalflux.com/gaussian-mixture-models-what-are-they-when-to-use/> (accessed Oct. 06, 2022).
- [19] "Semi-Supervised Learning, Explained | AltexSoft." <https://www.altexsoft.com/blog/semi-supervised-learning/> (accessed May 02, 2022).

- [20] "What are the types of Reinforcement learning algorithms? - FinsliQ Blog." <https://www.finsliqblog.com/ai-and-machine-learning/what-are-the-types-of-reinforcement-learning-algorithms/> (accessed May 02, 2022).
- [21] "Aprendizaje profundo por refuerzo IIC." <https://www.iic.uam.es/aprendizaje-profundo-por-refuerzo/> (accessed May 22, 2022).
- [22] "Overview of Network Resource Management - Managing Network Virtualization and Network Resources in Oracle® Solaris 11.2." [https://docs.oracle.com/cd/E36784\\_01/html/E36813/gecki.html](https://docs.oracle.com/cd/E36784_01/html/E36813/gecki.html) (accessed Jul. 24, 2022).
- [23] "What Is Network Virtualization and Network Resource Management? - Managing Network Virtualization and Network Resources in Oracle® Solaris 11.2." [https://docs.oracle.com/cd/E36784\\_01/html/E36813/gndac.html](https://docs.oracle.com/cd/E36784_01/html/E36813/gndac.html) (accessed Jul. 24, 2022).
- [24] "What is Network Performance Monitoring?" <https://www.techtarget.com/searchnetworking/definition/network-performance-monitoring> (accessed Jul. 24, 2022).
- [25] "What is Network Performance Monitoring? Definition and FAQs | HEAVY.AI." <https://www.heavy.ai/technical-glossary/network-performance-monitoring> (accessed Jul. 24, 2022).
- [26] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Applying deep learning approaches for network traffic prediction," 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017, vol. 2017-January, pp. 2353–2358, Nov. 2017, doi: 10.1109/ICACCI.2017.8126198.
- [27] "Overview of traffic prediction system and its impact on TMS efficiency | Download Scientific Diagram." [https://www.researchgate.net/figure/Overview-of-traffic-prediction-system-and-its-impact-on-TMS-efficiency\\_fig5\\_266673838](https://www.researchgate.net/figure/Overview-of-traffic-prediction-system-and-its-impact-on-TMS-efficiency_fig5_266673838) (accessed Jul. 24, 2022).
- [28] M. M. Raikar, S. M. Meena, M. M. Mulla, N. S. Shetti, and M. Karanandi, "Data Traffic Classification in Software Defined Networks (SDN) using supervised-learning," *Procedia Comput Sci*, vol. 171, pp. 2750–2759, Jan. 2020, doi: 10.1016/J.PROCS.2020.04.299.
- [29] S. Valenti, D. Rossi, A. Dainotti, A. Pescapè, A. Finamore, and M. Mellia, "Reviewing traffic classification," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7754, pp. 123–147, 2013, doi: 10.1007/978-3-642-36784-7\_6/COVER.
- [30] "Network Fault Management and Monitoring Tools - ManageEngine OpManager." <https://www.manageengine.com/network-monitoring/network-fault-management.html> (accessed Jul. 24, 2022).
- [31] "6.1 Issues in Resource Allocation — Computer Networks: A Systems Approach Version 6.2-dev documentation." <https://book.systemsapproach.org/congestion/issues.html> (accessed Jul. 24, 2022).
- [32] "¿Qué es una URL? Aquí te lo contamos." <https://www.edix.com/es/instituto/que-es-url/> (accessed Oct. 02, 2022).
- [33] "Requirements engineering paper classification and evaluation criteria: a proposal and a discussion", doi: 10.1007/s00766-005-0021-6.
- [34] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf Softw Technol*, vol. 64, pp. 1–18, Aug. 2015, doi: 10.1016/J.INFSOF.2015.03.007.