



Microlearning para la enseñanza de inteligencia artificial en el nivel medio guatemalteco:
diseño y validación exploratoria de un plan curricular modular, ciclo escolar 2025.

TESIS DOCTORAL

que, para obtener el Grado de Ph.D.

DOCTORADO EN EDUCACIÓN E INNOVACIÓN

PRESENTA

Adrián Catalán Santis

ASESOR

Javier Numan Caballero Merlo

México, 2026

La presente Tesis Doctoral debe ser citada como:

Catalán Santis, Adrián (2026). Microlearning para la enseñanza de inteligencia artificial en el nivel medio guatemalteco: diseño y validación exploratoria de un plan curricular modular, ciclo escolar 2025. [Tesis de Doctorado de la Universidad de Investigación e Innovación de México – UIIX].



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Se permite la reproducción total o parcial y la comunicación pública de la obra con reconocimiento de la autoría y mención de la Universidad de Investigación e Innovación de México - UIIX.

No se permite el uso comercial ni la creación de obras derivadas.

Resumen.

La enseñanza de Inteligencia Artificial (IA) y Aprendizaje Automático (ML) en el sistema educativo guatemalteco de nivel medio se limita a la alfabetización digital básica, lo que configura un rezago frente a las demandas tecnológicas contemporáneas. El problema se formuló como la necesidad de un plan curricular pertinente para jóvenes de 13 a 18 años en el área urbana de la Ciudad de Guatemala. El objetivo general fue diseñar un plan curricular modular, fundamentado en microlearning con práctica guiada y experimentación libre, que introduzca redes neuronales a estudiantes sin experiencia previa. La hipótesis sostuvo que esta propuesta resulta viable y obtiene una valoración favorable. El enfoque fue mixto, con diseño secuencial explicativo (DEXPLIS) y alcance exploratorio; el componente cuantitativo correspondió a un diseño cuasi-experimental con dos grupos de comparación activa no equivalentes. Se aplicaron tres instrumentos (pre-test, post-test y escala Likert) a una muestra intencional de 21 estudiantes de una institución educativa privada urbana de Ciudad de Guatemala en el ciclo escolar 2025. Los resultados indican un pre-test alto ($M = 6.10/7 = 87.1\%$), un rendimiento equivalente entre grupos en el post-test técnico (66.7 % en ambos, $d \approx 0$; $n_{AI} = 10$, $n_{Otro} = 11$) y una valoración favorable de la metodología (Likert AI = 3.10/4, Otro = 3.25/4). Los hallazgos ofrecen evidencia exploratoria de viabilidad operativa y transferibilidad a contextos urbanos comparables; la verificación de superioridad pedagógica queda condicionada a una iteración escalada con muestra mayor.

Palabras clave: *Inteligencia Artificial, microlearning, currículo, educación media, Guatemala, redes neuronales, didáctica de la computación.*

Abstract.

Artificial Intelligence (AI) and Machine Learning (ML) education within Guatemala's secondary school system remains limited to basic digital literacy, which configures a gap relative to contemporary technological demands. The research problem was framed as the need for a relevant curriculum plan for youth aged 13 to 18 in the urban area of Guatemala City. The general objective was to design a modular curriculum plan, grounded in microlearning with guided practice and free experimentation, that introduces neural networks to students without prior experience. The hypothesis stated that this proposal is feasible and elicits favorable perception. A mixed methods approach was used, with a sequential explanatory design (DEXPLIS) and exploratory scope; the quantitative component corresponded to a quasi-experimental design with two non-equivalent active comparison groups. Three instruments were applied (pre-test, post-test and Likert scale) to a purposive sample of 21 students at one private urban school in Guatemala City during the 2025 academic year. Results indicate a high pre-test starting point ($M = 6.10/7 = 87.1\%$), equivalent technical post-test performance between groups (66.7 % in both, $d \approx 0$; $n_{AI} = 10$, $n_{Other} = 11$), and favorable perception of the methodology (Likert AI = 3.10/4, Other = 3.25/4). The findings provide exploratory evidence of operational feasibility and transferability to comparable urban contexts; verifying pedagogical superiority remains contingent on a scaled iteration with a larger sample.

Keywords: *Artificial Intelligence, microlearning, curriculum, secondary education, Guatemala, neural networks, computing education.*

Agradecimientos.

A la Universidad Galileo y al Doctorado en Educación e Innovación de la UIIX, mi reconocimiento por brindar el marco institucional para llevar la educación adelante.

A la comunidad de profesores en Guatemala, que cada día sostienen el sistema con recursos limitados, pero con vocación firme: este trabajo es para ustedes y desde ustedes.

A los estudiantes que participaron en el taller piloto, ustedes son la razón concreta de este trabajo. Sus respuestas, sus preguntas, las cosas que les gustaron y las que pidieron mejorar son la materia prima de los hallazgos que aquí se reportan. A las autoridades y docentes de la institución anfitriona, gracias por abrir el aula.

Dedicatorias.

A Dios.

A mis padres, Adrián e Irma. La familia es la primera comunidad y la más importante.

A mis mentores, Guillermo Castillo, Jack Trachtenberg, Rodrigo Baessa y Antonio León.

Hace mucho aprendí, primero de mis padres y después de mis mentores, el Inge Jack y don Guillermo (con varios grados académicos, pero a quien anteponerle “don” siempre me pareció de más respeto) que cuando se cierra una puerta hay que buscar la ventana, y si también está cerrada hay que escalar al techo y “*meterse por donde se pueda*”. Cualquier cosa que esté en estas páginas se la debo, en buena parte, a esa persistencia que les aprendí.

A mi equipo en el Laboratorio de Innovación, de Ingeniería de Sistemas en Universidad Galileo, siempre agradezco trabajar con los mejores. Gracias a ustedes podemos seguir mejorando la educación en Guatemala.

A los jóvenes de Guatemala, cuya curiosidad y talento merecen una educación a la altura de sus sueños. A los docentes que enseñan principios complejos con recursos sencillos, demostrando que la pedagogía es, en esencia, un acto de imaginación y compromiso.

ÍNDICE GENERAL

INTRODUCCIÓN	14
CAPÍTULO I. PROYECCIÓN DE LA INVESTIGACIÓN.	18
1.1. Línea de investigación de la Universidad de Innovación e Investigación de México y su ámbito de estudio.	18
1.2. Planteamiento del problema.	19
1.2.1. Antecedentes del tema.....	21
1.3. Formulación del problema (Pregunta de investigación).	22
1.4. Justificación.....	22
1.5. Objeto de estudio.....	24
1.6. Campo de acción.	24
1.7. Objetivos.	25
1.7.1. Objetivo General.	25
1.7.2. Objetivos específicos.....	25
1.8. Hipótesis.....	26
1.8.1 Variables.....	26
1.9. Alcance temático.	27
1.10. Delimitación Espacial y Temporal.	28
CAPÍTULO II. Fundamentos Teóricos Referenciales.	31

2.1. Estado del arte.	31
Marco histórico.	33
2.2. Marco Teórico.	34
2.3. Marco Conceptual.	48
2.4. Marco Contextual.	52
2.5. Marco Legal y Normativo.	59
CAPÍTULO III. FUNDAMENTOS METODOLÓGICOS Y RESULTADOS DE INVESTIGACIÓN.	61
3.1. Cuadro Operacionalización de variables.	61
3.2. Diseño metodológico.	65
3.3. Trabajo de campo.	75
3.4. Análisis de los resultados en los datos obtenidos.	79
3.5. Redacción de resultados y discusión.	118
CAPÍTULO IV: PROPUESTA DE TRANSFORMACIÓN	123
4.1. Fundamentación de la propuesta de transformación.	123
4.2. Estructura de la propuesta de transformación.	126
4.3. Viabilidad operativa exploratoria de la propuesta de transformación.	145
4.4. Objetivos, recursos y productos de la propuesta.	153
CONCLUSIONES	155
RECOMENDACIONES	159

BIBLIOGRAFÍA.....	162
ANEXOS.....	171
Anexo 1. Pruebas objetivas de conocimientos sobre IA: pre-test y post-test.	171
Anexo 1B. Post-test.....	172
Anexo 2. Escala de involucramiento y motivación.....	173
Anexo 3. Estructura del notebook de Google Colab utilizado en el taller.	174
Anexo 4. Cronograma de la sesión piloto del 17 de febrero de 2025.	175
Anexo 5. Bitácora del investigador (resumen).....	176
Anexo 6. Estructura completa y código del notebook Colab utilizado en el taller.	176
Anexo 7. Plan de capacitación docente (16 horas, 4 sesiones).	179
Anexo 8. Respuestas anonimizadas del pre-test (n = 21).....	181
Anexo 9. Respuestas anonimizadas del post-test (n = 21).	183
Anexo 10. Distribución completa de la escala Likert por ítem y por grupo.	185
Anexo 11. Material del taller: estructura de las diapositivas (Intro to Deep Learning using Keras).	186
Anexo 12. Agenda operativa y guía pedagógica del taller (17 de febrero de 2025).	188
Anexo 13. Glosario técnico.	190
Anexo 14. Alineación del plan curricular con AI4K12 y orientaciones de la UNESCO.	194
Anexo 15. Plan curricular extendido: 16 sesiones desglosadas.	197

Índice de figuras.

Figura 1. Diagrama del Diseño Secuencial Explicativo (DEXPLIS) aplicado al estudio (adaptado de Creswell y Plano Clark, 2017	65
Figura 2. Cronograma del trabajo de investigación.	71
Figura 3. Red semántica de categorías emergentes en respuestas abiertas (n = 21).	83
Figura 4. Flujo del diseño cuasi-experimental con grupos no equivalentes: pre-test común, asignación por logística institucional, dos tratamientos activos paralelos, post-test específico por contenido y escala Likert común sobre metodología.	95
Figura 5. Diagrama del ciclo de microlearning aplicado en la propuesta curricular (microconcepto → microdesafío → microcierre, 15 minutos por ciclo).	125
Figura 6. Esquema de la arquitectura de la red neuronal feed-forward empleada en el taller (capa de entrada 28×28, capa oculta de 128 neuronas con ReLU, capa de salida de 10 clases con Softmax).	126

Índice de gráficas.

Gráfica 1. Comparación radial de la cobertura de los tres marcos teóricos de referencia (AI4K12, UNESCO y Bloom revisado) sobre cinco dimensiones formativas relevantes para una propuesta curricular de IA en K-12.	52
Gráfica 2. Cronograma operativo (diagrama de Gantt) de las fases del trabajo de investigación.	71
Gráfica 3. Distribución comparativa de puntajes pre/post-test (n = 21).	81
Gráfica 4. Distribución de puntajes en la escala Likert por grupo (n = 21).	82
Gráfica 5. Distribución demográfica de la muestra (n = 21).	85
Gráfica 6. Comparativa de medias pre-test (toda la cohorte) y post-test por grupo (con SD; n = 21).	94
Gráfica 7. Acierto por ítem en el pre-test (n = 21), excluyendo la pregunta abierta PT-7.	105
Gráfica 8. Acierto por ítem en el post-test, con umbral de referencia del 60 %.	107
Gráfica 9. Distribución del puntaje post-test por grupo (n = 21).	108
Gráfica 10. Mapa de calor de aciertos individuales por estudiante e ítem (verde = acierto, rojo = error; n = 21 estudiantes × 13 ítems cerrados pre+post).	109
Gráfica 11. Tasa de acierto agregada por nivel cognitivo de la taxonomía de Anderson y Krathwohl (2001), comparada entre pre-test y post-test (n = 21).	110
Gráfica 12. Puntaje individual normalizado pre-test versus post-test (n = 21); línea diagonal indica sin cambio, líneas punteadas indican las medias respectivas.	110
Gráfica 13. Distribución temporal de envíos del pre-test y post-test.	111
Gráfica 14. Distribución apilada de respuestas Likert por ítem.	113
Gráfica 15. Distribución de los niveles de familiaridad previa autorreportada (n = 21).	115

Gráfica 16. Relación entre familiaridad previa y puntaje post-test (n = 21).....	116
Gráfica 17. Distribución de carga horaria por módulo del plan curricular (24 horas totales).....	136
Gráfica 18. Cobertura cognitiva por módulo según la taxonomía de Anderson y Krathwohl (2001).	136
Gráfica 19. Cobertura del plan curricular sobre las Cinco Grandes Ideas de la IA (Touretzky et al., 2019).....	137

Índice de tablas.

Tabla 1. Síntesis del marco teórico integrador del estudio.	37
Tabla 2. Síntesis comparada de las seis teorías base e implicaciones operativas.	45
Tabla 3. Indicadores demográficos, educativos y de conectividad relevantes para el estudio.....	56
Tabla 4. Cuadro de operacionalización de variables del estudio.	62
Tabla 5. Distribución demográfica de la muestra (n = 21).....	83
Tabla 6. Categorías emergentes en respuestas abiertas (n = 21).	85
Tabla 7. Frecuencia de aciertos por ítem en el post-test (n = 21).	87
Tabla 8. Aciertos por ítem en el pre-test (n = 21).	88
Tabla 9. Aciertos por ítem en el post-test, comparativa entre grupos (n_AI = 10, n_Otro = 11). .	90
Tabla 10. Valoración Likert por afirmación y por grupo (escala 1–4: media y desviación estándar).	91
Tabla 11. Puntaje pre-test por nivel de familiaridad autorreportada con IA (n = 21).	92
Tabla 12. Potencia estadística estimada del piloto para detectar diferencias entre el grupo AI (n = 10) y el grupo paralelo (n = 11) según el tamaño del efecto, con $\alpha = 0.05$ y prueba t de dos colas.	96
Tabla 13. Puntaje pre-test por nivel de familiaridad autorreportada con Python (n = 21).	97
Tabla 14. Tabla cruzada de familiaridad autorreportada con IA y con Python (n = 21).....	98
Tabla 15. Categorización de respuestas a la pregunta abierta «¿qué hace una red neuronal?» (n = 21).....	100
Tabla 16. Tasas de favorabilidad por ítem Likert (% de respuestas "de acuerdo" o "muy de acuerdo").	101

Tabla 17. Categorización de respuestas a la pregunta «¿qué se podría mejorar?» (n = 21).	102
Tabla 18. Acierto por ítem en el pre-test (n = 21).	104
Tabla 19. Acierto por ítem en el post-test (n = 21).	106
Tabla 20. Estadísticos descriptivos de la escala Likert (n = 21, escala 1-4).	111
Tabla 21. Familiaridad previa autorreportada (n = 21).	114
Tabla 22. Categorías emergentes en las respuestas abiertas del post-test (n = 21).	116
Tabla 23. Tabla maestra del plan curricular: 16 sesiones organizadas en 4 módulos.	127
Tabla 24. Cobertura de las Cinco Grandes Ideas de la IA (Touretzky et al., 2019) por módulo del plan curricular.	131
Tabla 25. Desarrollo minuto a minuto de la sesión M2S2 (ejemplo).	132
Tabla 26. Rúbrica analítica transversal de evaluación (cuatro dimensiones × cuatro niveles). ...	134
Tabla 27. Cobertura del plan curricular sobre las 17 competencias de AI literacy (Long y Magerko, 2020).	137
Tabla 28. Equivalencias entre el plan curricular propuesto y currículos internacionales de IA y ciencias computacionales en K-12.	140
Tabla 29. Matriz de riesgos éticos del uso de IA en el aula y medidas mitigadoras del plan curricular.	141
Tabla 30. Validación de la propuesta de transformación según los seis criterios institucionales (UIIX Etapa 5).	146
Tabla 30. Alineación de las 16 sesiones del módulo curricular con las subáreas del Currículo Nacional Base de Guatemala (MINEDUC, Acuerdo Ministerial 178-2009 y actualizaciones) en el Ciclo Diversificado, Bachillerato en Ciencias y Letras con orientación en Computación.	151

INTRODUCCIÓN

Existe variedad de discusiones sobre temas como Transformación Digital, Cuarta Revolución Industrial o el impacto del uso de herramientas de Inteligencia Artificial en el día a día. Estas discusiones se nutren con base a lo que prácticamente podemos identificar como una reconfiguración de los sistemas productivos, sociales y comunicacionales de la humanidad; impulsada por tecnología. La Inteligencia Artificial (IA) y el Aprendizaje Automático (*Machine Learning* en inglés, también conocido como Aprendizaje de Máquina) han dejado de ser promesas futuristas para convertirse en los motores que redefinen la competitividad global. Esta acelerada transformación digital no ocurre de manera uniforme. En Guatemala, la integración de estas tecnologías en los sistemas educativos formales continúa siendo un desafío. El país, caracterizado por una economía en vías de desarrollo y profundas brechas de desigualdad, se encuentra en una encrucijada: avanzar hacia la modernización de su capital humano o permitir que la brecha tecnológica amplifique las exclusiones sociales existentes.

La educación tecnológica se modela como un vehículo casi indispensable para la movilidad social y el desarrollo económico sostenible. La realidad educativa guatemalteca presenta un rezago significativo frente a las demandas del mercado laboral global. Existe una disonancia marcada entre las competencias digitales avanzadas que la industria requiere (como el pensamiento computacional, el análisis de datos y la comprensión de algoritmos) y los contenidos curriculares vigentes en el nivel medio, los cuales a menudo se limitan a una alfabetización digital básica o al uso instrumental de ofimática.

Esta problemática no es exclusiva de Guatemala, sino que resuena en toda la región. Como señala Aparicio-Gómez (2023), la incidencia de la Inteligencia Artificial en la educación latinoamericana es un factor determinante para transformar el aprendizaje actual, pero su implementación debe ser cuidadosa para evitar que se convierta en un factor de exclusión. El autor enfatiza que la integración de la IA debe orientarse a reducir las brechas de desigualdad mediante políticas inclusivas y éticas, una premisa que sustenta la justificación social de esta investigación. La carencia estructural de programas actualizados priva a los jóvenes de herramientas esenciales para comprender y moldear su entorno, relegándolos a un rol de consumidores pasivos de tecnología en lugar de empoderarlos como creadores de soluciones innovadoras.

Hace algunos años, con veinte cumplidos, me dieron la oportunidad de ser docente universitario y empecé con la actitud idealista de un joven que quería cambiar el mundo. Traté siempre de hacer las cosas diferente, de preocuparme por todos los aspectos posibles de las clases, y en el camino surgió una duda, que al día de hoy, no logro responder del todo: ¿de qué forma aprenden mejor las personas? Desde la dirección del Laboratorio de Innovación de la Universidad Galileo he visto crecer a varias generaciones de estudiantes que llegan curiosos, se arremangan las mangas y se ensucian las manos con tecnología. Programas como *Engineering Kickstart*, *NodeBots Guatemala* y los talleres que compartimos en comunidades tecnológicas locales me convencieron de que la chispa existe. Lo que falta es el currículo que la sostenga.

Adicional a esto, también me ha llamado la atención una pregunta repetida en los grupos pequeños de conversación en educación y tecnología: ¿quién está formando a los jóvenes guatemaltecos para que dentro de cinco años puedan participar de esta conversación, no como espectadores sino como autores? Esa pregunta es, la que en gran parte esta tesis intenta responder.

Esta investigación doctoral, titulada "Microlearning para la enseñanza de inteligencia artificial en el nivel medio guatemalteco: diseño y validación exploratoria de un plan curricular modular, ciclo escolar 2025", surge como una respuesta académica y pragmática a esta realidad. El estudio se fundamenta en la hipótesis de que la democratización del conocimiento técnico complejo es posible si se adaptan las estrategias pedagógicas a las características cognitivas de los nativos digitales. A través del diseño de una propuesta curricular modular y contextualizada, se busca demostrar que es viable introducir conceptos de ingeniería de datos y redes neuronales a adolescentes.

Para lograr esta adaptación didáctica, es clave superar la enseñanza teórica tradicional. En este orden de ideas, Rodríguez-García et al. (2020) han demostrado, a través de herramientas como *LearningML*, que el fomento de habilidades de pensamiento computacional es significativamente más efectivo cuando se realiza mediante proyectos prácticos de Inteligencia Artificial. Su investigación valida el enfoque de "aprender haciendo" (learning by doing), evidenciando que los estudiantes de nivel escolar son capaces de entender y manipular modelos de aprendizaje automático cuando se les proporcionan las herramientas adecuadas. Siguiendo esta línea, la presente tesis propone el uso de entornos accesibles como *Google Colab* y bibliotecas como *TensorFlow* para facilitar la experimentación directa.

El éxito de cualquier innovación curricular depende intrínsecamente de la percepción y capacidad de los educadores. Dúo-Terrón et al. (2023) destacan en sus estudios recientes que, para que la Inteligencia Artificial y el *Machine Learning* se consoliden como un recurso educativo efectivo, es indispensable considerar la perspectiva docente y ofrecer formación que desmitifique estas tecnologías en etapas no universitarias. Esta investigación aborda dicho desafío incorporando un componente de formación docente como eje transversal del plan curricular propuesto.

Este trabajo se alinea estratégicamente con las líneas de investigación del Doctorado en Educación e Innovación de la Universidad de Investigación e Innovación de México (UIIX). Específicamente, se vincula con la línea de "Planificación y gestión de la Educación", al proponer una estructura curricular formal que abarca desde la definición de objetivos de aprendizaje hasta la evaluación de competencias; y con la línea de "Innovación educativa y perspectivas tecnológicas", al validar el uso de recursos didácticos novedosos que facilitan el interaprendizaje y la experimentación práctica en entornos con recursos limitados.

La tesis se organiza en cuatro capítulos que llevan al lector desde el planteamiento del problema hasta la validación de la propuesta. La estructura capitular es la siguiente:

El Capítulo 1, Proyección de la Investigación, ubica la investigación. Describe la realidad de la educación tecnológica en Guatemala desde los retos del sistema hasta las carencias del aula. Se formula la pregunta de investigación y se definen los objetivos generales y específicos que orientan el trabajo, apoyados en la necesidad de cerrar la brecha digital señalada por Aparicio-Gómez (2023). Cierra con la justificación y los alcances espaciales, temporales y temáticos.

El Capítulo 2 construye el marco teórico y referencial: revisa las teorías pedagógicas que sustentan el diseño curricular (constructivismo, construccionismo y conectivismo), los fundamentos técnicos de la IA y los modelos instruccionales, tomando como referencia las experiencias prácticas de Rodríguez-García et al. (2020) para fundamentar la viabilidad cognitiva de los contenidos seleccionados.

El Capítulo 3 detalla el Marco Metodológico, definiendo el enfoque mixto y el diseño cuasi-experimental, con grupos no equivalentes de la investigación. Se describen los instrumentos de recolección de datos, la población y muestra seleccionada, y los procedimientos utilizados para la

implementación del piloto. Aquí se retoman las consideraciones de Dúo-Terrón et al. (2023) para el diseño de instrumentos que evalúen no solo el aprendizaje del alumno, sino la viabilidad de la implementación desde la óptica docente.

El Capítulo 4 presenta la Propuesta de Transformación, constituyendo el aporte central y propositivo de esta investigación. Este capítulo integra tanto el diseño de la arquitectura curricular, fundamentada pedagógicamente en *microlearning* y el uso de las herramientas seleccionadas, como la valoración y validación empírica de la misma. En esta sección se sistematizan los datos recolectados durante la implementación piloto para contrastar los hallazgos con los objetivos planteados, evaluando la efectividad del modelo en términos de ganancia cognitiva, desarrollo de habilidades técnicas e incremento en la motivación de los estudiantes.

Finalmente, el documento culmina con las Conclusiones y Recomendaciones, un apartado de cierre donde se sintetizan los aportes principales de la investigación, se discuten las implicaciones de los hallazgos y se ofrecen directrices estratégicas para la replicabilidad y escalabilidad del modelo en el sistema educativo nacional, aspirando a que este esfuerzo académico contribuya de manera tangible a la reducción de la brecha digital en Guatemala.

CAPÍTULO I. PROYECCIÓN DE LA INVESTIGACIÓN.

Este capítulo presenta el marco general de la investigación. En él se establece la línea de investigación de la Universidad de Innovación e Investigación de México con la que se vincula el estudio, se desarrolla el planteamiento del problema y sus antecedentes, se formula la pregunta de investigación que guía el trabajo, se ofrece la justificación, se delimitan el objeto de estudio y el campo de acción, se enuncian el objetivo general y los objetivos específicos, se plantea la hipótesis con sus variables, y se establecen el alcance temático y la delimitación espacial y temporal de la investigación.

1.1. Línea de investigación de la Universidad de Innovación e Investigación de México y su ámbito de estudio.

El tema propuesto, "Microlearning para la enseñanza de inteligencia artificial en el nivel medio guatemalteco: diseño y validación exploratoria de un plan curricular modular, ciclo escolar 2025", presenta una clara y directa vinculación con las líneas de investigación de la Universidad de Investigación e Innovación de México (UIIX), específicamente con las líneas de "Planificación y gestión de la Educación" e "Innovación educativa y perspectivas tecnológicas", así como con sus respectivos ámbitos de estudio.

La investigación se adscribe a la línea de Planificación y gestión de la Educación, dentro del ámbito de estudio de Planificación académica y diseño curricular. El proyecto se relaciona directamente con este ámbito, ya que su objetivo medular es desarrollar un plan curricular innovador para la enseñanza de principios de inteligencia artificial. La planificación académica implica la definición rigurosa de los objetivos de aprendizaje, los contenidos temáticos, y las estrategias de enseñanza y evaluación que se utilizarán en el plan curricular propuesto. El diseño curricular se enfoca en estructurar la secuencia lógica de los contenidos, así como en la selección de las actividades y recursos didácticos más adecuados para alcanzar los objetivos de aprendizaje trazados. La investigación busca adaptar el currículo a las necesidades específicas de los estudiantes, considerando sus intereses y contextos culturales particulares para asegurar su pertinencia.

La temática propuesta se vincula estrechamente con la línea de Innovación educativa y perspectivas tecnológicas, abarcando dos ámbitos de estudio: Diseño e innovación de recursos didácticos e Interaprendizaje, innovación y tecnología en la educación. La enseñanza de principios de inteligencia artificial y *machine learning* a jóvenes de nivel medio requiere la implementación de enfoques y recursos didácticos novedosos y actualizados. El diseño e innovación de recursos didácticos en este proyecto busca crear materiales y herramientas educativas que promuevan un aprendizaje significativo y atractivo; por ello, se explorarán diferentes herramientas tecnológicas accesibles que faciliten la enseñanza de estos temas complejos y potencien la interacción y participación de los jóvenes.

Sobre lo ya mencionado, también es importante considerar que el ámbito de estudio de Interaprendizaje, innovación y tecnología en la educación se alinea perfectamente con el enfoque innovador que se busca implementar en el plan curricular. El interaprendizaje se refiere a la colaboración y el aprendizaje conjunto entre estudiantes y docentes, promoviendo un ambiente participativo y de co-construcción del conocimiento. La introducción de estos temas en el aula puede fomentar la interacción entre los estudiantes y el desarrollo de proyectos colaborativos, lo que impulsa el aprendizaje activo y el pensamiento crítico.

En conjunto, la investigación tiene el potencial de contribuir al avance de la educación tecnológica en Guatemala, así como de fomentar el uso de tecnologías innovadoras en el aula para enriquecer el aprendizaje de los jóvenes. La propuesta se alinea con el interés del investigador en democratizar el conocimiento en el campo de la inteligencia artificial, brindando a los jóvenes oportunidades equitativas para desarrollar habilidades relevantes en un mundo tecnológicamente avanzado.

1.2. Planteamiento del problema.

El problema se plantea partiendo de que Guatemala, un país con una economía tradicionalmente sustentada en la agricultura, atraviesa hoy un momento decisivo en su transformación digital. Se avanza de manera acelerada hacia la integración profunda de la tecnología en prácticamente todos los aspectos de la vida cotidiana y productiva pero el sistema de educación tecnológica guatemalteco enfrenta varios obstáculos estructurales. Entre estos, destacan la insuficiencia de recursos financieros e infraestructura adecuada, la carencia de

docentes capacitados en nuevas tecnologías y pedagogías activas, y un decreciente interés por parte de los estudiantes, quienes a menudo perciben la oferta educativa actual como desconectada de su realidad y futuro.

El país enfrenta retos sociales profundos, con una de las tasas de pobreza y desigualdad más altas de América Latina. Esta realidad afecta desproporcionadamente a una población numerosa y desatendida, compuesta mayoritariamente por comunidades rurales e indígenas, donde un gran porcentaje de la fuerza laboral se concentra en el sector informal. En este escenario, la educación es uno de los principales problemas nacionales, y la educación tecnológica es particularmente escasa y desigual.

Guatemala tiene que lidiar con problemas graves en términos de acceso equitativo a recursos y tecnología de punta, así como con la falta de formación docente especializada en temas de vanguardia. Los planes curriculares existentes carecen, en su mayoría, de la inclusión de habilidades tecnológicas modernas, generando una notable discrepancia entre la demanda del mercado laboral por competencias avanzadas, como las relacionadas con la inteligencia artificial y el aprendizaje automático (*machine learning*), y la oferta real de profesionales preparados en estas áreas.

Esta brecha educativa tiene consecuencias directas y severas. La limitada oferta de educación tecnológica de calidad restringe las oportunidades de los jóvenes guatemaltecos para participar efectivamente en la economía digital global, limitando su potencial de desarrollo profesional y éxito tecnológico. Más allá del impacto individual, esta situación afecta las perspectivas de desarrollo del país y profundiza las divisiones socioeconómicas existentes. Como ya se ha mencionado anteriormente, el acceso y dominio de tecnologías avanzadas tiende a concentrarse en manos de una minoría privilegiada, aumentando la exclusión social.

La falta de un currículo adaptado a las realidades actuales deja a los jóvenes en desventaja competitiva y desperdicia una oportunidad histórica de impulsar el desarrollo del país a través de la innovación tecnológica. Lo he visto repetidamente en mi trabajo como instructor en Platzi y edX (a través de Universidad Galileo), donde he acompañado a más de quince mil estudiantes hispanohablantes: los jóvenes guatemaltecos llegan curiosos, dispuestos, con manejo intuitivo del teléfono y de las herramientas que usan a diario; pero cuando se les pide escribir un ciclo *for*

(parte básica del manejo de control de flujo al programar en un lenguaje de alto nivel) o entender qué hace una capa densa, aparece un vacío que ninguna app de TikTok va a llenar. Ese vacío no es individual, es estructural, y es el problema que esta tesis intenta atacar.

El interés por enseñar habilidades tecnológicas a quien quiera aprender no es retórico para quien suscribe esta tesis. Desde 2009 organizo encuentros de la comunidad de desarrolladores con *Google Developer Group* Guatemala, en 2014 fui designado como el primer guatemalteco del programa *Google Developer Experts*, en 2016 abrí *NodeBots* Guatemala para acercar la programación física a estudiantes de secundaria, y desde 2020 moví mi especialidad tecnológica de desarrollo para móviles hacia inteligencia artificial. Esta trayectoria es la que sostiene la pregunta de fondo de esta investigación: si los jóvenes guatemaltecos tienen acceso a un plan curricular bien diseñado y a un docente preparado, ¿pueden apropiarse de los principios de IA en condiciones comparables a las de cualquier otra región del mundo? La hipótesis de partida es que sí, y esta tesis pone la primera evidencia empírica sobre la mesa.

1.2.1. Antecedentes del tema.

Desde el punto de vista social e histórico, la educación tecnológica en Guatemala se ha caracterizado por una incorporación tardía y desigual de las Tecnologías de la Información y Comunicación (TIC) al sistema formal. Aunque desde finales del siglo XX existen esfuerzos por introducir competencias digitales básicas en el Currículo Nacional Base, estos se han limitado mayoritariamente a la alfabetización ofimática, dejando fuera el tratamiento sistemático de tecnologías emergentes como la Inteligencia Artificial.

En el plano económico, Guatemala atraviesa una transición desde una economía agrícola tradicional hacia un modelo productivo más diversificado, donde la industria de servicios digitales y el outsourcing tecnológico adquieren creciente relevancia. Esta transformación demanda capital humano formado en competencias avanzadas, analítica de datos, programación, IA, que el sistema educativo nacional aún no logra suministrar de forma equitativa, generando una brecha estructural entre la oferta formativa y los requerimientos del mercado laboral global.

Desde lo político y normativo, los antecedentes muestran un reconocimiento progresivo de la importancia de las TIC en la educación, reflejado en la Ley de Educación Nacional (Decreto 12-

91), la Política Nacional de Ciencia, Tecnología e Innovación, y los compromisos del país con el Objetivo de Desarrollo Sostenible 4, pero sin que estas políticas hayan derivado aún en directrices específicas para la enseñanza de Inteligencia Artificial en el nivel medio. Esta ausencia normativa convierte a esta investigación en un insumo pertinente para la actualización curricular.

En el ámbito cultural, Guatemala se distingue por una rica diversidad étnica y lingüística, con más de veinte comunidades mayas, garífunas y xincas que conviven con la población mestiza. Esta heterogeneidad, lejos de ser un obstáculo, constituye un activo para el diseño de estrategias pedagógicas situadas que vinculen los contenidos de IA con problemáticas locales (agricultura, salud rural, prevención de desastres). La pertinencia y novedad del tema radica precisamente en que ningún antecedente nacional documentado ha articulado un plan curricular de IA para jóvenes que integre estos elementos contextuales.

1.3. Formulación del problema (Pregunta de investigación).

Considerando la situación problemática descrita, donde la ausencia de formación estructurada en tecnologías emergentes limita el desarrollo de la juventud guatemalteca, se plantea la siguiente interrogante fundamental que guía el desarrollo de esta investigación:

¿Cómo se puede mejorar la enseñanza de principios fundamentales de inteligencia artificial (IA) y aprendizaje automático (ML) de los jóvenes de 13 a 18 años en instituciones educativas de nivel medio del área urbana de la Ciudad de Guatemala, durante el ciclo escolar 2025?

1.4. Justificación.

La pertinencia de esta investigación se sostiene en cuatro dimensiones complementarias (teórica, práctica, social y metodológica) que conviene revisar por separado, no como adornos formales sino porque cada una responde a una pregunta distinta sobre por qué este trabajo importa.

En lo teórico, el estudio aporta al debate sobre didáctica de la Inteligencia Artificial y el Aprendizaje Automático en niveles preuniversitarios, un campo todavía poco explorado en la región centroamericana pero que en el resto del mundo ya tiene marcos consolidados (AI4K12, Long y Magerko, Touretzky). Sistematizar una propuesta curricular para adolescentes

guatemaltecos significa ensayar cómo se traducen conceptos abstractos (ingeniería de datos, redes neuronales, funciones de activación) a las estructuras cognitivas y al castellano cotidiano de quienes hablamos español en Guatemala, no al lenguaje aséptico de la traducción técnica.

En lo práctico, la investigación responde a una demanda concreta del mercado laboral. He acompañado a startups locales (Inversiones I+E y procesos de mentoría con asociaciones locales como CIG o AGEXPORT) en la búsqueda de talento joven con habilidades de IA, y la respuesta es sistemáticamente la misma: el talento existe, pero la formación temprana no. El plan curricular propuesto busca cerrar esa brecha desde el nivel medio, dotando a los jóvenes de una base sobre la que después puedan construir una carrera técnica.

En lo social, el proyecto se justifica por su compromiso con la equidad de acceso. La brecha tecnológica entre la Ciudad de Guatemala y los departamentos del altiplano y la costa sur es enorme; sumarle la brecha lingüística entre el castellano y los veintidós idiomas mayas multiplica el desafío. Un currículo accesible —con materiales bajo licencia abierta, herramientas gratuitas como *Google Colab*, y planes de capacitación docente que pueda escalar el MINEDUC en alianza con DIGEBI y la Academia de Lenguas Mayas de Guatemala— es la única forma honesta de evitar que la próxima ola tecnológica amplíe la desigualdad existente.

En lo metodológico, la tesis ofrece un modelo replicable: dieciséis sesiones organizadas en cuatro módulos, con materiales abiertos, instrumentos de evaluación validados con datos reales de aula y un plan de capacitación docente diseñado en paralelo. La combinación de microlearning, aprendizaje basado en proyectos y herramientas de código abierto está pensada para que cualquier centro educativo guatemalteco (público, privado, urbano, rural) pueda adaptar el plan a su realidad sin depender de licencias propietarias ni de infraestructura inalcanzable.

La realidad de muchos de los jóvenes con quienes trabajo en Guatemala presenta una distancia enorme entre la frontera global del conocimiento y el aula local. La mayor parte de mi carrera la he dedicado a acortarla, primero como ingeniero, después como profesor, después como organizador de comunidad. Esta tesis es el siguiente paso lógico. No se trata de demostrar que la inteligencia artificial es enseñable a adolescentes; eso ya lo sé, lo he visto. Se trata de sistematizar cómo, para que no dependa solo del entusiasmo de un profesor en particular y pueda replicarse.

En uno de los talleres de *Engineering Kickstart*, a un estudiante de unos dieciséis años que llevaba más de la mitad de la sesión sin lograr un ejercicio de código. Me senté con él. Me dijo que estaba estudiando Bachillerato en Ciencias y Letras, que por no ser Bachiller en Computación, que por eso le costaba más los temas tecnológicos. Le respondí que yo también soy Bachiller en Ciencias y Letras, exactamente como él, y que estábamos en el mismo colegio donde yo me había graduado años atrás. Le dije que estaba seguro de que iba a lograr el ejercicio. Poco tiempo después lo terminó. Esa escena me confirma una intuición que esta tesis hace explícita: lo primero que hay que enseñar no es código, sino la convicción de que el código también es para uno.

1.5. Objeto de estudio.

El objeto de estudio de esta investigación lo constituye el proceso de enseñanza-aprendizaje de los principios de Inteligencia Artificial (IA) y Aprendizaje Automático (*Machine Learning*) dirigido a estudiantes de nivel medio (13 a 18 años).

Este objeto se aborda desde una perspectiva pedagógica e innovadora, analizando las dinámicas, contenidos, metodologías y recursos didácticos necesarios para facilitar la asimilación efectiva de competencias tecnológicas avanzadas.

1.6. Campo de acción.

El campo de acción de esta investigación se circunscribe de manera precisa al diseño, desarrollo y evaluación de un plan curricular que incorpora la enseñanza de Inteligencia Artificial (IA) y Aprendizaje Automático (*Machine Learning*) dentro del sistema educativo guatemalteco, específicamente dirigido a la población juvenil de 13 a 18 años.

Este campo no se limita únicamente a la estructuración de contenidos temáticos, sino que abarca una intervención educativa integral que involucra múltiples dimensiones del proceso pedagógico. En primer lugar, comprende la adaptación curricular a los contextos locales, asegurando que los conceptos tecnológicos avanzados sean accesibles y culturalmente relevantes para los estudiantes. En segundo lugar, incluye la capacitación y formación docente, reconociendo que el éxito de cualquier innovación educativa depende de la preparación y confianza de los facilitadores frente a las nuevas tecnologías. Finalmente, abarca la implementación de estrategias pedagógicas activas que promuevan un aprendizaje significativo y práctico.

Dentro de este campo de acción, se busca fomentar activamente un entorno colaborativo que permita la experimentación, la innovación y la aplicación de conocimientos en proyectos reales. Se pretende transformar la dinámica del aula tradicional hacia un espacio donde estudiantes y docentes co-construyan el aprendizaje mediante la resolución de problemas que reflejen las necesidades y desafíos del entorno guatemalteco.

1.7. Objetivos.

1.7.1. Objetivo General.

Diseñar un plan curricular innovador centrado en la enseñanza de principios fundamentales de inteligencia artificial (IA) y aprendizaje automático (ML) dirigido a jóvenes guatemaltecos de 13 a 18 años en instituciones educativas de nivel medio del área urbana de la Ciudad de Guatemala, durante el ciclo escolar 2025, mediante una intervención piloto de dos meses, para potenciar sus capacidades analíticas y de resolución de problemas a través de un enfoque pedagógico participativo y dinámico que los prepare para las demandas del futuro laboral tecnológico.

1.7.2. Objetivos específicos.

- Analizar los fundamentos teóricos de la enseñanza de la inteligencia artificial y el aprendizaje automático en la educación media, así como las variables pedagógicas que intervienen en su didáctica con adolescentes, para sustentar conceptualmente la propuesta curricular.
- Identificar las necesidades educativas y el nivel de conocimiento previo en tecnología de los jóvenes guatemaltecos de 13 a 18 años, mediante diagnósticos participativos, para diseñar un currículo que sea pertinente, relevante y accesible a su realidad contextual.
- Elaborar módulos de aprendizaje estructurados que integren teoría y práctica sobre Inteligencia Artificial (IA) y Aprendizaje Automático (ML), utilizando metodologías de aprendizaje activo y basado en proyectos, para fomentar el pensamiento crítico y la creatividad en los estudiantes.
- Evaluar la efectividad e impacto del plan curricular a través de una implementación piloto, midiendo el progreso en las competencias técnicas de los estudiantes y el

incremento en su interés hacia carreras tecnológicas, para validar y ajustar la propuesta final.

1.8. Hipótesis.

La implementación de una metodología basada en *microlearning* con práctica modular y experimentación libre incrementa significativamente el nivel de conocimiento técnico sobre redes neuronales (*TensorFlow*, funciones de activación y optimización) y mejora la percepción de involucramiento en estudiantes sin experiencia previa profunda.

Hipótesis Específicas (H1, H2, H3)

- H1 (Cognitiva): Los estudiantes expuestos a ciclos de teoría breve y práctica demostrarán la capacidad de diferenciar conceptos técnicos complejos (como el uso de *model.fit* y la capa *ReLU*) que no poseían en el diagnóstico inicial.
- H2 (Actitudinal): La posibilidad de experimentar libremente con parámetros durante el taller tiene una correlación positiva directa con la sensación de estar involucrado en el proceso de aprendizaje.
- H3 (Metodológica): La estructura de *microlearning* es percibida por los participantes como un facilitador eficaz para entender mejor el funcionamiento abstracto de las redes neuronales en comparación con la teoría tradicional.

1.8.1 Variables

- Variable Dependiente. Nivel de competencia técnica en principios de Inteligencia Artificial y Redes Neuronales.
- Variable Dependiente Actitudinal. Nivel de involucramiento activo (*engagement*) del estudiante a través de la experimentación. Esta variable se incluye como segunda variable de respuesta de naturaleza actitudinal y no como variable mediadora en sentido estadístico (Baron y Kenny, 1986; MacKinnon, 2008). El diseño no contempla análisis de mediación porque el tamaño muestral ($n = 21$) y la estructura del estudio no permiten estimar coeficientes indirectos con potencia suficiente.
- Variable Independiente. Estrategia didáctica basada en *microlearning* y práctica modular.

1.9. Alcance temático.

El alcance de esta investigación se define a partir de tres dimensiones fundamentales: teórica, metodológica y práctica, las cuales establecen los límites epistemológicos y operativos del estudio. Esta delimitación asegura que el diseño curricular propuesto no solo sea innovador, sino también viable y evaluado dentro del contexto descrito.

Alcance Teórico

En la dimensión teórica, la investigación se circunscribe estrictamente a la enseñanza de fundamentos de Inteligencia Artificial (IA) y Aprendizaje Automático (*Machine Learning*), excluyendo deliberadamente otras ramas de las tecnologías emergentes (como robótica, Internet de las Cosas o *Blockchain*) para garantizar la profundidad necesaria en el tratamiento de los contenidos seleccionados. Dentro del amplio campo de la IA, el estudio se acota específicamente al aprendizaje supervisado y las redes neuronales artificiales. No se pretende formar ingenieros de *machine learning* ni científicos de datos a nivel universitario, sino establecer una alfabetización técnica funcional que permita a los estudiantes comprender y manipular conceptos como algoritmos, capas de activación (ej. *ReLU*), épocas de entrenamiento y uso de funciones de optimización (ej. *ADAM*).

El marco teórico pedagógico que sustenta este alcance se basa en el *microlearning* y el aprendizaje activo. La investigación delimita su alcance instructivo a la validación de contenidos modulares que reducen la carga cognitiva, permitiendo que conceptos abstractos de ingeniería de datos sean asimilables por adolescentes de 13 a 18 años. Teóricamente, el trabajo aborda la intersección entre la andragogía técnica simplificada y la pedagogía constructivista, donde el estudiante pasa de ser un consumidor de "cajas negras" tecnológicas a un diseñador capaz de ajustar parámetros y entender la lógica de fondo de un modelo predictivo.

Alcance Metodológico

De forma metodológica, el estudio se define por un enfoque mixto con diseño secuencial explicativo de alcance experimental. El alcance de la recolección de datos combina métodos cuantitativos y cualitativos para obtener una visión completa, integral y desde mi perspectiva, holística, que enriquecen el experimento para este trabajo de investigación..

En lo cuantitativo, el alcance se limita a la medición de variables objetivas de aprendizaje técnico (pre y post-test) y patrones de mejora cognitiva a corto plazo. No se contempla un estudio longitudinal que mida el impacto en la carrera profesional a largo plazo (años), sino que se focaliza en la ganancia de conocimiento inmediata tras la intervención.

En lo cualitativo, el alcance abarca la fenomenología de la experiencia del estudiante. Se limita a interpretar las percepciones de autoeficacia, motivación y la valoración de la metodología didáctica mediante las preguntas abiertas del post-test y la escala Likert de percepción. El estudio no busca generalizar los resultados a toda la población rural de Guatemala debido a las variables de infraestructura, pero sí establecer una base válida para contextos urbanos con acceso mínimo a tecnología.

Alcance Práctico

Desde la perspectiva práctica, el producto central de la investigación es el "Plan Curricular de IA para Nivel Medio" y sus instrumentos asociados. El alcance práctico contempla el diseño completo de la secuencia didáctica, la selección y curaduría de herramientas tecnológicas, y la creación de guías para docentes. La investigación tiene un alcance propositivo y aplicativo: no se queda en el diagnóstico de la brecha existente sobre la transformación digital, sino que entrega una solución funcional probada. El límite práctico está marcado por la implementación piloto: el estudio entrega un modelo validado en un entorno controlado (laboratorio escolar o aula conectada), listo para ser transferido o escalado por las instituciones educativas, pero la ejecución masiva a nivel nacional excede los límites de esta tesis doctoral. El alcance práctico se enfoca en software y lógica de programación, excluyendo la necesidad de hardware especializado o costoso, para asegurar que la propuesta sea económicamente viable y replicable en escuelas con recursos de cómputo estándar.

1.10. Delimitación Espacial y Temporal.

Para garantizar la viabilidad y el rigor científico de la investigación, se ha establecido un marco de trabajo específico para espacio y tiempo, que engloban trabajo de campo y la intervención educativa necesaria. Estas delimitaciones responden a la necesidad de controlar

variables externas que podrían afectar la validez interna del diseño cuasi-experimental con grupos no equivalentes.

Delimitación Espacial

La investigación se desarrolla geográficamente en Ciudad de Guatemala, Guatemala, acotándose específicamente a áreas urbanas que cuenten con instituciones educativas de nivel medio (ciclo básico y diversificado). Esta selección espacial responde a criterios técnicos de infraestructura necesarios para la enseñanza de los temas expuestos, específicamente la disponibilidad de laboratorios de computación con conexión estable a internet, requisito indispensable para el uso de herramientas en la nube como las antes mencionadas. El espacio de estudio se delimita a centros educativos que representan el área urbana promedio, donde los estudiantes tienen acceso a dispositivos digitales pero carecen de formación avanzada en ciencias de la computación.

Es importante precisar que, aunque las falencias descritas son más agudas en las zonas rurales del país, esta investigación excluye espacialmente dichas áreas en esta fase inicial. Esta delimitación responde a una necesidad metodológica de aislar variables: al trabajar en entornos urbanos con conectividad mínima garantizada, se puede asegurar que los desafíos o éxitos observados sean atribuibles al diseño curricular y la metodología pedagógica, y no a fallos de infraestructura básica (como la falta de electricidad o internet). El modelo se diseña con la potencialidad de ser adaptado y replicado posteriormente en otros contextos geográficos a medida que mejore la infraestructura nacional.

Las intervenciones se llevan a cabo físicamente dentro de las aulas o laboratorios de informática de las instituciones seleccionadas (muestra intencional de una a dos instituciones). Este entorno controlado permite la observación directa de la interacción estudiante-computadora y estudiante-docente, facilitando la recolección de datos cualitativos sobre el comportamiento, la colaboración entre pares y la resolución de problemas en tiempo real, aspectos que serían difíciles de monitorear en un entorno de educación remota o dispersa.

Delimitación Temporal

La investigación se estructura temporalmente en un ciclo de taller cubriendo los temas específicos, permitiendo observar el proceso de aprendizaje desde el diagnóstico hasta la evaluación final. La implementación del plan curricular propiamente dicho se delimita a un periodo académico de dos meses. Este marco temporal es suficiente para desplegar los ciclos de *microlearning*, permitir la maduración de los conceptos técnicos en los estudiantes y evaluar la retención de conocimientos a corto plazo sin interferir excesivamente con el calendario escolar regular de las instituciones, asegurando así la viabilidad operativa del proyecto.

El alcance temporal también abarca las fases preparatorias críticas previas a la intervención en el aula. Esto incluye el tiempo dedicado a la validación técnica de los instrumentos, las reuniones de coordinación logística con autoridades educativas relevantes y una micro fase de capacitación docente previa. Estas etapas son parte integral de la delimitación temporal, pues aseguran que, al momento de iniciar los dos meses de implementación con los estudiantes, el ecosistema educativo (materiales, permisos, docentes) esté completamente alineado y funcional.

Finalmente, se establece un límite temporal claro respecto a la evaluación de impacto: el estudio es de corte transversal en cuanto a la medición de resultados de aprendizaje. Se evalúa el cambio cognitivo y actitudinal entre el punto A (Pre-test) y el punto B (Post-test) dentro del periodo de estudio. La investigación excluye el seguimiento longitudinal a largo plazo (años posteriores), por lo que no forma parte del alcance temporal verificar si los estudiantes eligen efectivamente carreras universitarias STEM en el futuro lejano. El enfoque temporal está puesto en la ganancia inmediata de competencias y el cambio en la motivación durante la intervención educativa.

CAPÍTULO II. Fundamentos Teóricos Referenciales.

El presente capítulo constituye el sustento teórico y referencial de la investigación. Se estructura en cinco apartados que abordan, de manera progresiva, el conocimiento acumulado y los marcos de referencia necesarios para fundamentar el diseño e implementación de un plan curricular orientado a la enseñanza de principios de Inteligencia Artificial a jóvenes de 13 a 18 años en Guatemala. Se parte del Estado del Arte, que examina las iniciativas pedagógicas y curriculares más relevantes a nivel global; se avanza hacia el Marco Teórico, que articula las teorías del aprendizaje que sustentan la propuesta; se precisan los conceptos clave en el Marco Conceptual; se contextualiza la realidad socioeducativa guatemalteca en el Marco Contextual; y se cierra con el Marco Legal y Normativo que delimita el alcance regulatorio de la intervención.

2.1. Estado del arte.

Este apartado presenta los principales antecedentes investigativos sobre la enseñanza de la Inteligencia Artificial (IA) a nivel preuniversitario, organizados de manera cronológica para evidenciar la evolución del campo desde 2019 hasta 2024.

En 2019, Pedro et al., en el documento de la UNESCO «Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development», sentaron las bases globales del debate al advertir que la integración de IA en la educación debía equilibrar los beneficios técnicos con consideraciones éticas y de equidad, marcando una agenda inicial para los sistemas educativos en desarrollo.

Durante 2020, Chiu, T. K. y Chai, C. S. propusieron un enfoque sostenible de planificación curricular para IA basado en la teoría de la autodeterminación, destacando la importancia de la motivación intrínseca del estudiante en la apropiación de contenidos tecnológicos complejos.

En 2021, Chiu, T. K. et al. desarrollaron y evaluaron un currículo de IA preuniversitario, aportando evidencia empírica sobre la viabilidad cognitiva de estos contenidos en adolescentes; en el mismo año, Tedre, M. et al. analizaron las trayectorias pedagógicas y tecnológicas para enseñar aprendizaje automático en aulas K-12, subrayando la necesidad de combinar herramientas accesibles con marcos pedagógicos sólidos.

El año 2022 trajo una diversificación temática en la investigación. Chu, S. T. et al. publicaron una revisión sistemática sobre robots educativos basados en IA; Gómez-Diago, G. exploró su aplicación a la enseñanza del periodismo; y Lee, I. y Perret, B. abordaron la preparación de docentes de secundaria para integrar métodos de IA en asignaturas STEM, evidenciando la versatilidad disciplinar del campo.

En 2023 se consolidaron los enfoques sistemáticos y éticos. Rizvi, S. et al. realizaron una revisión sistemática del periodo 2019-2022 que sintetizó los avances pedagógicos del campo; Velander, J. et al. analizaron la influencia de la comprensión docente sobre la IA en la enseñanza; y Adams, C. et al. formularon principios éticos específicos para la IA en la educación K-12, integrando la dimensión valorativa al diseño curricular. En el contexto iberoamericano, Aparicio-Gómez, W. O. analizó la incidencia de la IA en la educación latinoamericana y advirtió que su integración debe orientarse a reducir desigualdades, premisa que sustenta directamente esta investigación. En la misma línea, Dúo-Terrón, P. et al. estudiaron la perspectiva docente sobre IA y *machine learning* como recurso educativo en etapas no universitarias.

Más recientemente, en 2024, Kong, S. C. et al. propusieron un marco de alfabetización en IA evaluado mediante aprendizaje basado en proyectos para estudiantes de secundaria superior, mientras que Yim, I. H. Y. y Su, J. publicaron una revisión exhaustiva de herramientas de aprendizaje de IA en K-12, consolidando un mapa actualizado de recursos didácticos disponibles.

El recorrido cronológico evidencia una transición clara: de un campo inicialmente preocupado por sentar principios generales (2019) hacia un dominio maduro que articula marcos pedagógicos, instrumentos de evaluación, formación docente y consideraciones éticas (2023-2024). Esta evolución sustenta la pertinencia de diseñar, en el contexto guatemalteco, una propuesta curricular que integre los hallazgos acumulados durante el último quinquenio.

Una valoración crítica de los antecedentes revisados muestra que, si bien la literatura del último quinquenio ha avanzado significativamente en marcos pedagógicos y herramientas para la enseñanza de IA en K-12, persisten tres vacíos relevantes: (a) escasa producción específica para contextos latinoamericanos, y prácticamente nula para Centroamérica, lo que limita la transferibilidad de los hallazgos al contexto guatemalteco; (b) predominio de estudios de corto plazo sin seguimiento longitudinal del impacto vocacional; y (c) escasa articulación entre la

formación de estudiantes y la formación docente como sistema integrado. Esta investigación contribuye a llenar estos vacíos al situar el estudio en Guatemala, articular un kit pedagógico que incluye explícitamente formación docente, y dejar planteadas líneas de investigación longitudinal a partir de los hallazgos.

Marco histórico.

Desde una perspectiva histórica, la inteligencia artificial como campo de estudio se remonta a la mitad del siglo XX, cuando investigadores de diversas disciplinas comenzaron a explorar la posibilidad de crear máquinas que pudieran simular aspectos de la inteligencia humana. Desde las ideas teóricas de Alan Turing sobre la "máquina de Turing" y el famoso test de Turing, hasta el desarrollo de los primeros programas informáticos que podían realizar tareas simples de razonamiento, la IA ha recorrido un largo camino. Este desarrollo histórico ha estado marcado por periodos de intensa actividad conocidos como "veranos de la IA", y momentos de escepticismo denominados "inviernos de la IA".

En el ámbito educativo, la incorporación de la IA ha seguido una trayectoria paralela de creciente interés y experimentación. Las primeras aplicaciones buscaban principalmente automatizar tareas de enseñanza y evaluación, mediante el uso de sistemas tutoriales inteligentes que adaptaban el contenido al ritmo de aprendizaje de cada estudiante. Con el tiempo, la integración ha evolucionado hacia enfoques más sofisticados, incluyendo la creación de entornos de aprendizaje personalizados, la utilización de agentes pedagógicos inteligentes y el desarrollo de herramientas para fomentar el pensamiento crítico y la resolución creativa de problemas.

En Guatemala, la adopción de tecnologías de IA en la educación está en sus etapas iniciales, reflejando una tendencia global hacia la digitalización del aprendizaje. A pesar de los desafíos que presenta la infraestructura tecnológica limitada en algunas áreas, ha habido esfuerzos significativos para integrar las TIC en el currículo educativo, marcando el camino para una futura incorporación de la IA, alineada con las políticas nacionales que buscan mejorar la calidad educativa.

Diseñar un plan curricular de IA para jóvenes guatemaltecos suma a un esfuerzo regional aún incipiente. La trayectoria histórica de la IA y su entrada al sistema educativo muestran que cualquier enfoque útil tiene que cubrir lo técnico sin separarlo de lo ético ni de lo social.

2.2. Marco Teórico.

La integración de la Inteligencia Artificial (IA) en los currículos educativos para jóvenes plantea desafíos y oportunidades únicas que requieren una base teórica sólida. En este contexto, la teoría del constructivismo (Piaget, 1977) ofrece una perspectiva valiosa, al sugerir que los estudiantes construyen activamente su conocimiento a través de la interacción con su entorno. Esta teoría es particularmente relevante para la enseñanza de la IA, donde la comprensión conceptual se fomenta mejor mediante experiencias de aprendizaje prácticas y experimentales.

La teoría del aprendizaje situado (Lave & Wenger, 1991) argumenta que el aprendizaje es inherentemente un acto social que ocurre dentro de una comunidad de práctica. Aplicando esta teoría al contexto de la IA en la educación, importa crear comunidades de aprendizaje en las que los estudiantes puedan participar en proyectos de IA auténticos, colaborando con compañeros y profesionales del campo. Este enfoque busca mejorar la comprensión técnica y desarrollar habilidades esenciales, como el trabajo en equipo y la resolución de problemas en contextos complejos.

La teoría de la carga cognitiva (Sweller, 1988) también juega un papel crucial en el diseño de un plan curricular efectivo para la enseñanza de la IA. Según Sweller (1988), el diseño instruccional debe gestionar cuidadosamente la carga cognitiva para evitar sobrecargar la memoria de trabajo del estudiante, facilitando así un aprendizaje más profundo. En la práctica, esto significa simplificar la introducción de conceptos complejos de IA y emplear estrategias didácticas que optimicen el procesamiento de la información, como el aprendizaje basado en problemas y el uso de analogías.

La necesidad de incorporar principios éticos en la educación de la IA es una extensión natural del constructivismo y el aprendizaje situado, ya que prepara a los estudiantes para entender la tecnología, incluyendo las reflexiones necesarias sobre las implicaciones que tiene en la sociedad. Adams, C. et al. (2023) discuten la importancia de integrar principios éticos en la educación de la IA en el nivel K-12, subrayando que los estudiantes deben ser capacitados para tomar decisiones informadas y éticas en el desarrollo y aplicación de tecnologías de IA.

La adaptabilidad y personalización del aprendizaje juegan un papel crucial en la efectividad de la educación en IA. Siguiendo las ideas de Yim, I. H. Y., y Su, J. (2024), que realizan una revisión

exhaustiva de las herramientas de aprendizaje de IA en la educación K-12, se hace evidente la importancia de integrar herramientas y plataformas educativas que permitan una experiencia de aprendizaje personalizada y adaptativa. Esto responde a la teoría de la carga cognitiva, ya que permite ajustar la complejidad de los contenidos y actividades a las capacidades individuales de los estudiantes.

La motivación es un factor determinante en el éxito del aprendizaje de la IA. La teoría de la autodeterminación (Ryan & Deci, 2000) sugiere que el apoyo a las necesidades psicológicas básicas de autonomía, competencia y relacionamiento aumenta la motivación intrínseca de los estudiantes. Aplicar principios de esta teoría al desarrollo curricular puede implicar ofrecer a los estudiantes opciones en proyectos de IA, proporcionar retroalimentación constructiva y fomentar un ambiente de aprendizaje colaborativo y de apoyo.

Desde la perspectiva del conectivismo, teoría propuesta por Siemens (2004), el aprendizaje en la era digital se produce a través de redes de conexiones entre nodos de información. Esta teoría cobra especial relevancia para la enseñanza de la IA, dado que el campo se caracteriza por su naturaleza interdisciplinaria y la necesidad constante de actualización. Los estudiantes no solo deben adquirir conocimientos declarativos sobre qué es la IA, sino también desarrollar la capacidad de identificar, evaluar y conectar fuentes de información relevantes en un ecosistema tecnológico en permanente cambio.

El *microlearning*, estrategia pedagógica central en esta investigación, encuentra sustento teórico en la convergencia de varias de las teorías expuestas. Al fragmentar contenidos complejos como las redes neuronales en ciclos breves de teoría y práctica (5 a 15 minutos), se respeta la capacidad limitada de la memoria de trabajo descrita por Sweller (1988), se facilita la construcción progresiva de conocimiento en línea con Piaget (1977) y se genera un sentido de competencia y autonomía congruente con la autodeterminación de Ryan y Deci (2000). Investigaciones recientes, como las de Chiu, T. K. et al. (2021) y Tedre, M. et al. (2021), confirman que las estrategias modulares y experienciales producen resultados superiores en la enseñanza de conceptos de aprendizaje automático en contextos K-12.

Desde el socio constructivismo de Vygotsky (1988), el aprendizaje ocurre primeramente en un plano social para luego internalizarse. La zona de desarrollo próximo resulta especialmente

pertinente para la enseñanza de IA a jóvenes sin conocimientos previos: el docente diseña actividades que estén justo por encima del nivel actual del estudiante, proporcionando andamiaje a través de ejercicios guiados, colaboración entre pares y retroalimentación inmediata. En el taller propuesto, el paso de ciclos guiados a la mini competencia de ajustes libres ejemplifica esta transición del apoyo externo hacia la autonomía progresiva del aprendiz.

La integración de las teorías presentadas permite articular un marco pedagógico robusto para el diseño curricular en IA. El constructivismo aporta la premisa de que el estudiante es agente activo de su aprendizaje; el aprendizaje situado garantiza la autenticidad de las tareas; la carga cognitiva orienta la dosificación de los contenidos; la autodeterminación explica los mecanismos motivacionales; el conectivismo da cuenta de la naturaleza en red del conocimiento tecnológico; y el socio constructivismo establece el papel del entorno social en la mediación del aprendizaje.

Investigaciones recientes refuerzan la pertinencia del marco teórico expuesto. Kong, S. C. et al. (2024) demuestran, mediante un marco de alfabetización en IA validado con estudiantes de secundaria, que el aprendizaje basado en proyectos potencia la comprensión conceptual cuando se combina con andamiaje progresivo, lo que es coherente con la zona de desarrollo próximo de Vygotsky (1988) y con la dosificación de carga cognitiva de Sweller (1988).

Por su parte, Yim, I. H. Y. y Su, J. (2024) realizan una revisión sistemática de herramientas de aprendizaje de IA en K-12 que confirma la efectividad de plataformas accesibles y modulares ,como las propuestas en esta investigación (*Google Colab*, *TensorFlow*), para sostener experiencias de aprendizaje personalizadas y adaptativas. Estos hallazgos contemporáneos validan la convergencia entre constructivismo, conectivismo y microlearning como estrategia pedagógica vigente.

Adams et al. (2023) integran la dimensión ética al marco teórico al sostener que la educación en IA en niveles preuniversitarios no puede dissociarse de los principios de transparencia, equidad y responsabilidad. Esta perspectiva resulta especialmente relevante en contextos como el guatemalteco, donde la introducción de tecnologías avanzadas debe articularse con la promoción de ciudadanía crítica y democratización del conocimiento.

Tabla 1. Síntesis del marco teórico integrador del estudio.

Teoría	Autor(es)	Aporte clave para esta investigación
Constructivismo	Piaget (1977)	El estudiante construye activamente el conocimiento mediante interacción con el entorno; sustenta el aprendizaje práctico de IA.
Aprendizaje situado	Lave & Wenger (1991)	El aprendizaje ocurre en comunidades de práctica; orienta el diseño colaborativo del taller.
Carga cognitiva	Sweller (1988)	Limita la introducción de conceptos complejos; fundamenta la dosificación del <i>microlearning</i> .
Autodeterminación	Ryan & Deci (2000)	Autonomía, competencia y relacionamiento explican la motivación intrínseca; orienta el diseño motivacional.
Conectivismo	Siemens (2004)	El conocimiento existe en redes de conexiones; sustenta

		la naturaleza interdisciplinar de la IA.
Socio constructivismo	Vygotsky (1988)	La zona de desarrollo próximo orienta el andamiaje progresivo del taller.
Alfabetización en IA (vigente)	Kong et al. (2024)	Marco contemporáneo de literacy en IA validado en secundaria con aprendizaje basado en proyectos.
Ética de IA en K-12 (vigente)	Adams et al. (2023)	Principios de transparencia, equidad y responsabilidad como ejes transversales del currículo.

Para efectos del diseño curricular y la interpretación de los resultados, esta investigación adopta explícitamente como teoría rectora la teoría de la autodeterminación de Ryan y Deci (2000), complementada por la teoría de la carga cognitiva de Sweller (1988) como marco operativo para la dosificación de los contenidos. La elección de la autodeterminación como teoría rectora obedece a tres razones: primero, su robustez empírica en estudios sobre motivación estudiantil en contextos tecnológicos; segundo, su correspondencia directa con las dimensiones medidas en la escala de involucramiento (autonomía, competencia, relacionamiento); tercero, su afinidad conceptual con el principio ético de democratización del conocimiento que vertebra esta tesis.

Aportes de la literatura iberoamericana en K-12 AI.

El cuerpo iberoamericano de literatura sobre AI literacy en K-12 ha crecido aceleradamente desde 2019. Aparicio-Gómez (2023) sintetiza el panorama regional señalando la urgencia de transformar el aprendizaje del siglo XXI con la incorporación temprana de IA en el currículo,

identificando tres tensiones recurrentes: brecha de acceso, brecha de formación docente y vacío curricular formal. Su trabajo opera como mapa orientador para investigaciones aplicadas como la presente.

Dúo-Terrón, Hinojo-Lucena, Moreno-Guerrero y López-Núñez (2023) presentan un estudio empírico con docentes de distintas etapas educativas no universitarias en España, donde la mitad de la muestra reporta haber utilizado herramientas de IA y machine learning como recurso educativo, pero solo un tercio se siente capacitado para enseñar IA como contenido específico. La asimetría entre uso instrumental y enseñanza explícita es coherente con el contexto guatemalteco que esta tesis documenta y refuerza la pertinencia del componente de capacitación docente del Capítulo 4.

Vázquez-Cano, Ramírez-Hurtado, Sáez-López y López-Meneses (2023) discuten cómo la irrupción de la IA generativa, particularmente ChatGPT, ha reconfigurado las prácticas docentes en educación superior, lo que abre la pregunta sobre cómo trasladar lecciones aprendidas en el nivel universitario al nivel medio. García-Peñalvo (2023) caracteriza esta transición como un momento entre la disrupción y el pánico, en el que las decisiones curriculares prematuras pueden ser tan costosas como la inacción. Esta tesis se ubica en esta tensión proponiendo un plan curricular específico para el nivel medio guatemalteco que adopta una postura de aprovechamiento crítico, no de mera incorporación celebratoria.

En el plano internacional, Adams, Pente, Lermeyer y Rockwell (2023) sistematizan un conjunto de principios éticos para la IA en K-12 que esta investigación incorpora transversalmente: transparencia algorítmica, equidad de acceso, responsabilidad sobre los datos y agencia del estudiante. La integración de la dimensión ética desde el primer módulo, en lugar de un módulo final aislado, se respalda en estos principios y en el framework de Long y Magerko (2020) que coloca la pregunta «¿cómo deberíamos usar la IA?» como uno de los cuatro dominios constitutivos de la alfabetización en IA.

El framework de las Cinco Grandes Ideas de la IA para K-12 propuesto por Touretzky, Gardner-McCune, Martin y Seehorn (2019) en el AAAI 2019 ha sido adoptado en distintas iniciativas regionales. La revisión sistemática de Sanusi, Oyelere, Vartiainen, Suhonen y Tukiainen (2023) sobre enseñanza de *machine learning* en K-12 muestra que las experiencias documentadas en

países iberoamericanos siguen siendo minoría frente a las de Asia-Pacífico y América del Norte, lo que confirma la pertinencia de aportar evidencia empírica desde Centroamérica. El trabajo de Yim y Su (2024) en su revisión de alcance sobre herramientas de aprendizaje de IA en K-12 identifica a *Google Colab* y *TensorFlow* como dos de las plataformas más utilizadas en experiencias documentadas, lo que respalda la elección instrumental de esta investigación.

En síntesis, la literatura sobre IA en K-12 converge en cinco principios operativos que esta investigación recoge: (1) iniciar con problemas de relevancia local antes de pasar a *benchmarks* abstractos; (2) integrar la ética transversalmente, no como apéndice; (3) escalonar la complejidad desde simuladores no-código a notebooks en código; (4) cuidar la brecha de género desde el primer módulo con estrategias intencionales (Foro Económico Mundial, 2023, ubica a Guatemala en el puesto 117 de 146 en Global Gender Gap); (5) sostener el componente comunitario y de mentoría con universidades o comunidades técnicas como factor de permanencia. Estos cinco principios articulan el diseño de la propuesta curricular del Capítulo 4 y permiten anclar el aporte guatemalteco a una conversación regional ya en curso.

Profundización en las teorías del aprendizaje base.

El marco teórico de la propuesta integra seis teorías que conviene profundizar de manera individual antes de articularlas como sistema. Cada teoría se presenta con sus postulados centrales, su pertinencia para la enseñanza de IA en la edad objetivo del estudio y sus implicaciones operativas para el diseño curricular del Capítulo 4.

El constructivismo de Jean Piaget.

La obra de Jean Piaget (1896-1980) establece que el conocimiento no se transmite pasivamente sino que se construye activamente por el aprendiz a través de procesos de asimilación (incorporación de nueva información a esquemas existentes) y acomodación (modificación de esquemas para integrar información que no encaja). En su epistemología genética (Piaget, 1970), el desarrollo cognitivo procede por etapas cualitativamente distintas: sensoriomotora (0-2 años), preoperacional (2-7 años), operaciones concretas (7-11 años) y operaciones formales (11 años en adelante).

La población objetivo del estudio (13-17 años) se ubica plenamente en la etapa de operaciones formales, caracterizada por la capacidad de razonar sobre proposiciones abstractas, de plantear hipótesis, de razonar sobre lo posible además de lo actual y de operar sobre representaciones simbólicas complejas. Esta capacidad es prerequisite cognitivo para conceptos centrales de la enseñanza de IA: la abstracción de un algoritmo de su implementación concreta, el razonamiento sobre el comportamiento de un modelo no observado directamente, la formulación de hipótesis sobre por qué un modelo falla en ciertos casos. La elección del rango etario 13-17 para la propuesta no es arbitraria: descender por debajo de los 12 años choca con limitaciones piagetianas para el manejo de abstracciones técnicas como funciones matemáticas (gradiente, función de pérdida) o conceptos como *overfitting*.

Para el diseño curricular, la implicación operativa más importante del constructivismo piagetiano es la insistencia en el conflicto cognitivo como motor del aprendizaje. La práctica del taller piloto reportada en el Capítulo 3 documentó momentos de conflicto cognitivo (la sorpresa de los estudiantes ante el modelo que falla con ciertos dígitos manuscritos, la curiosidad por el efecto del *overfitting* cuando la precisión en el set de datos de entrenamiento sube pero la de validación cae) como puntos de mayor activación. La propuesta del Capítulo 4 sistematiza estos momentos como dispositivos didácticos intencionales, no como accidentes felices del taller.

El construccionismo de Seymour Papert.

Seymour Papert (1928-2016), discípulo directo de Piaget en Ginebra y posteriormente codirector del MIT Media Lab, extendió el constructivismo añadiendo una dimensión externa: la construcción de objetos del mundo (programas, máquinas, artefactos compartibles públicamente) potencia y materializa la construcción interna de conocimiento. En sus palabras: «la mejor manera de aprender es construir algo que tenga sentido para uno y que pueda ser compartido con otros» (Papert, 1980, *Mindstorms*). Esta tesis dio origen al lenguaje LOGO, los Mindstorms de LEGO y a una tradición pedagógica continuada por Mitchel Resnick (creador de Scratch) y Alan Kay, entre otros.

El construccionismo es la base teórica que más directamente informa el diseño operativo de la propuesta del Capítulo 4. Cada uno de los cuatro módulos del plan curricular cierra con un artefacto construido por el estudiante: un mapa conceptual personal (Módulo 1), un modelo entrenado en

Keras (Módulo 2), un experimento de hiperparámetros documentado (Módulo 3) y un proyecto contextualizado a Guatemala (Módulo 4). El artefacto cumple tres funciones simultáneas: materializa el aprendizaje, permite evaluación tangible y, por su carácter compartible, alimenta el sentido de comunidad que el construccionismo identifica como reforzador motivacional.

La obra *Mindstorms* (1980) sostiene además que las computadoras pueden ser «objetos para pensar en conjunto» (*objects-to-think-with*), es decir, herramientas que permiten al aprendiz manipular ideas abstractas de manera concreta. Esta noción justifica la elección de Google Colab y *TensorFlow* como herramientas de la propuesta: ambas son *objects-to-think-with* para el dominio de las redes neuronales, dado que permiten manipular conceptos como capa, peso, época y función de activación de manera tangible y reversible.

Vygotsky y la zona de desarrollo próximo.

Lev Vygotsky (1896-1934), contemporáneo de Piaget pero con énfasis distinto, situó el aprendizaje en lo social antes que en lo individual. Su concepto central, la zona de desarrollo próximo (ZDP), define el espacio entre lo que un aprendiz puede hacer solo y lo que puede hacer con la guía de un compañero o un mentor más capaz (Vygotsky, 1978, *Mind in Society*). El aprendizaje real ocurre cuando el aprendiz opera dentro de su ZDP con apoyo gradual decreciente (andamiaje, *scaffolding*).

La ZDP es la base teórica del componente de mentoría cruzada que la propuesta del Capítulo 4 incorpora como recurso opcional. La articulación con comunidades técnicas guatemaltecas (GDG, NodeBots) y con laboratorios universitarios opera como amplificador de la ZDP del estudiante de nivel medio: un mentor universitario o un profesional de la comunidad ofrece andamiaje en el límite superior de la ZDP, mientras el docente del centro educativo opera el andamiaje cotidiano. La propia estructura de *microlearning*, con bloques cortos de teoría seguidos de práctica con asistencia disponible, replica internamente el principio del andamiaje vygotskiano.

Una segunda contribución vygotskiana relevante es la noción del aprendizaje mediado por herramientas culturales. Las herramientas (físicas como un lápiz, simbólicas como un sistema de notación matemática) median entre el aprendiz y el contenido, y aprender una disciplina implica apropiarse de sus herramientas. En la enseñanza de IA, la herramienta cultural específica es el

notebook computacional (Jupyter, Colab) como género discursivo: alterna texto, código y resultados en un mismo documento, lo que constituye una manera particular de pensar que el aprendiz necesita interiorizar. La propuesta del Capítulo 4 explicita el manejo del notebook como objeto de aprendizaje en sí mismo, no solo como vehículo del contenido técnico.

La teoría de la autodeterminación de Ryan y Deci.

Edward Deci y Richard Ryan (Universidad de Rochester) desarrollaron desde finales de los años setenta la Teoría de la Autodeterminación (Self-Determination Theory, SDT), un marco motivacional empíricamente respaldado por más de cuatro décadas de investigación (Deci y Ryan, 1985; Ryan y Deci, 2017). La SDT postula que la motivación intrínseca y el aprendizaje sostenido dependen de la satisfacción de tres necesidades psicológicas básicas universales: autonomía (sentir que las propias acciones son auto-iniciadas y reflejan los propios valores), competencia (sentir que se puede operar efectivamente en el entorno) y relacionamiento (sentir conexión genuina con otros).

La SDT es el marco motivacional que la propuesta del Capítulo 4 adopta de manera explícita en el diseño de las actividades. La autonomía se materializa en la dimensión de experimentación libre con parámetros (épocas, capas, funciones de activación) que cada sesión incluye. La competencia se cuida mediante el escalonamiento progresivo de la dificultad (de simuladores no-código a código completo) y el ofrecimiento de bloques de teoría inmediatamente seguidos de práctica donde el éxito es alcanzable. El relacionamiento se nutre por el trabajo en parejas en varias sesiones y por la articulación con la comunidad técnica externa (mentores, comunidades GDG, ferias finales). La evidencia empírica del piloto del Capítulo 3 respalda la pertinencia de la SDT: el ítem Likert sobre involucramiento por experimentación con parámetros obtuvo 90 % de favorabilidad en el grupo AI, el más alto de toda la encuesta.

La teoría de la carga cognitiva de Sweller.

John Sweller, desde la Universidad de Nueva Gales del Sur, formuló a fines de los años ochenta la Teoría de la Carga Cognitiva (Cognitive Load Theory, CLT) como articulación entre la arquitectura cognitiva humana y el diseño instruccional (Sweller, 1988; Sweller, van Merriënboer y Paas, 1998). La CLT distingue tres tipos de carga cognitiva sobre la memoria de trabajo: carga

intrínseca (inherente al contenido), carga extrínseca (impuesta por el diseño instruccional, como instrucciones confusas o material redundante) y carga germana (la dedicada a la construcción de esquemas mentales útiles). El diseño instruccional efectivo busca minimizar la carga extrínseca, gestionar la intrínseca y maximizar la germana.

La CLT es el marco teórico que justifica directamente la elección de *microlearning* como metodología de la propuesta. Bloques cortos de teoría (cinco a siete minutos) limitan la carga intrínseca acumulada antes del refuerzo práctico; la repetición espaciada distribuye la carga a lo largo del tiempo; los notebooks plantilla con código pre-cargado eliminan carga extrínseca de tener que escribir todo el boilerplate desde cero; las celdas de texto explicativas dentro del propio notebook reducen la división de atención entre múltiples documentos (principio de contigüidad espacial de Mayer, 2009). La propuesta del Capítulo 4 incorpora estos principios CLT de manera explícita en cada sesión.

El conectivismo de Siemens.

George Siemens propuso en 2004, en su artículo seminal «Connectivism: A Learning Theory for the Digital Age», una teoría del aprendizaje específica para condiciones de abundancia informacional y red de conocimiento distribuido. El conectivismo postula que en la era digital el aprendizaje es esencialmente un proceso de formar y mantener conexiones entre nodos de información (humanos y no humanos), y que la capacidad de navegar redes de información, distinguir señal de ruido y reconocer patrones emergentes es más importante que la memorización de contenidos específicos.

El conectivismo es particularmente pertinente para la enseñanza de IA por dos razones. Primera: el campo de IA cambia rápidamente y cualquier currículo congelado en el contenido específico de su año de diseño quedará desfasado en pocos años; lo perdurable es la capacidad del estudiante para mantenerse actualizado a través de redes (documentación, comunidades, papers, repositorios). Segunda: las propias redes neuronales son una manifestación del paradigma conectivista en el dominio computacional, lo que ofrece una metáfora pedagógica útil para el estudiante (la propia red que está construyendo es un objeto-para-pensar-con sobre las redes que él mismo es capaz de tejer profesionalmente).

La propuesta del Capítulo 4 incorpora el conectivismo a través del módulo 4 (proyecto contextualizado), donde el estudiante debe articular información de fuentes diversas (*datasets* locales, documentación técnica de TensorFlow, artículos divulgativos, mentorías de la comunidad), y a través del componente de capacitación docente que incluye explícitamente el desarrollo de redes profesionales del docente con la comunidad técnica.

Tabla 2. Síntesis comparada de las seis teorías base e implicaciones operativas.

Teoría	Autores referencia	Postulado central	Implicación operativa en la propuesta
Constructivismo	Piaget (1970)	El aprendiz construye conocimiento por asimilación y acomodación; etapa de operaciones formales habilita razonamiento abstracto.	Diseño con conflicto cognitivo intencional como motor; rango 13-17 como mínimo cognitivo.
Construccionismo	Papert (1980); Resnick (2017)	El aprendizaje se potencia construyendo artefactos compartibles públicamente; computadoras como objects-to-think-with.	Cada módulo cierra con artefacto tangible; Colab/TF como objetos-para-pensar-con.

Sociocultural	Vygotsky (1978)	Aprendizaje en zona de desarrollo próximo con andamiaje gradual decreciente; herramientas culturales mediadoras.	Mentoría cruzada con comunidades técnicas; notebook como herramienta cultural a interiorizar.
Autodeterminación	Deci y Ryan (1985); Ryan y Deci (2017)	Motivación intrínseca depende de autonomía, competencia y relacionamiento.	Experimentación libre con parámetros; escalonamiento de dificultad; trabajo en parejas y comunidad.
Carga cognitiva	Sweller (1988); Mayer (2009)	Memoria de trabajo limitada; minimizar carga extrínseca, maximizar germana.	Microlearning con bloques cortos; notebooks plantilla; contigüidad de teoría y código.
Conectivismo	Siemens (2004)	Aprender es formar y mantener conexiones en redes de conocimiento distribuido.	Proyecto final integra fuentes diversas; capacitación docente incluye redes profesionales.

Frameworks específicos de alfabetización en IA (AI literacy).

Más allá de las teorías generales del aprendizaje, en el último decenio se han consolidado frameworks específicos para la enseñanza de IA en niveles preuniversitarios. Tres son particularmente influyentes y se discuten a continuación.

Las cinco grandes ideas de la IA según AI4K12.

La iniciativa AI4K12, codirigida por la Asociación para el Avance de la Inteligencia Artificial (AAAI) y la Asociación de Profesores de Ciencias de la Computación (CSTA), publicó en 2019 las Cinco Grandes Ideas de la IA para K-12 (Touretzky, Gardner-McCune, Martin y Seehorn, 2019), que ofrecen un marco curricular orientador para la enseñanza de IA desde preescolar hasta nivel medio. Las cinco ideas son: (1) las computadoras perciben el mundo usando sensores; (2) los agentes mantienen modelos del mundo y los usan para razonar; (3) las computadoras pueden aprender de los datos; (4) la interacción natural entre agentes y humanos requiere un esfuerzo significativo; y (5) la IA puede impactar a la sociedad de maneras tanto positivas como negativas.

La propuesta del Capítulo 4 cubre las cinco grandes ideas con énfasis diferenciado: la idea 3 (aprendizaje a partir de datos) es el eje central de los Módulos 2 y 3; la idea 5 (impacto social) se trabaja transversalmente y culmina en el Módulo 4; las ideas 1, 2 y 4 se introducen conceptualmente en el Módulo 1 sin profundización técnica, lo que es coherente con el alcance de un curso introductorio de 16 sesiones.

El framework de Long y Magerko.

Duri Long y Brian Magerko (Georgia Tech) propusieron en su trabajo seminal en CHI 2020 (Long y Magerko, 2020) un framework de competencias de AI literacy estructurado en diecisiete competencias agrupadas en cuatro dominios: ¿qué es la IA?, ¿qué puede hacer la IA?, ¿cómo funciona la IA? y ¿cómo deberíamos usar la IA? El framework derivó de una revisión sistemática de más de 150 trabajos de investigación y de la consulta con expertos del campo, y opera hoy como referencia estándar para el diseño curricular en alfabetización de IA.

Tres principios de diseño formulados por Long y Magerko son particularmente influyentes para la propuesta del Capítulo 4. Primero, la importancia de visualizaciones y representaciones físicas para hacer tangibles conceptos abstractos: la propuesta cumple este principio mediante el manejo directo de imágenes, predicciones visuales y la analogía del comité para explicar las capas. Segundo, el principio de involucrar al estudiante en la construcción de modelos en lugar de solo consumirlos: cumplido por el énfasis construccionista de cada módulo. Tercero, el principio de explicitar las consideraciones éticas como parte integral del aprendizaje técnico, no como apéndice: cumplido por la integración transversal de la dimensión ética en cada módulo.

El marco UNESCO para IA en educación.

La UNESCO publicó en 2019 el documento «Inteligencia Artificial en la Educación: Desafíos y Oportunidades para el Desarrollo Sostenible» (Pedro, Subosa, Rivas y Valverde, 2019), que establece principios rectores para la integración responsable de la IA en los sistemas educativos. El documento distingue entre dos dimensiones: la IA como objeto de estudio (lo que se enseña) y la IA como medio de enseñanza (cómo se enseña). Esta investigación se enfoca en la primera dimensión, aunque reconoce la pertinencia de explorar la segunda en líneas de investigación futura.

Cinco recomendaciones de la UNESCO informan el diseño de la propuesta: (1) priorizar la inclusión y la equidad en el acceso al aprendizaje de IA; (2) preparar a los docentes para la transformación con IA mediante programas de capacitación específicos; (3) desarrollar planes de estudio de calidad inclusiva con la IA integrada; (4) fomentar valores y habilidades para la vida y el trabajo en la era de la IA; (5) asegurar el uso ético y transparente de los datos educativos. El plan curricular del Capítulo 4 atiende explícitamente las cinco recomendaciones.

2.3. Marco Conceptual.

El marco conceptual se construye sobre definiciones y comprensiones clave que forman la base teórica y práctica de este estudio. Dado el tema "Microlearning para la enseñanza de inteligencia artificial en el nivel medio guatemalteco: diseño y validación exploratoria de un plan curricular modular, ciclo escolar 2025", este marco aborda conceptos fundamentales como la

inteligencia artificial, el plan curricular, la enseñanza de principios de IA, y su aplicación en un contexto educativo dirigido a jóvenes.

La "Inteligencia Artificial" (IA), en el contexto de esta investigación, se define como la capacidad de los sistemas computacionales para realizar tareas que tradicionalmente requieren inteligencia humana. Entre estas tareas están el reconocimiento de patrones, el aprendizaje basado en la experiencia, la solución de problemas y la toma de decisiones. En el ámbito educativo y el contexto de este trabajo, la IA se referirá tanto a la tecnología en sí como a la pedagogía involucrada en enseñar a los estudiantes cómo estas herramientas se desarrollan, operan y pueden ser aplicadas éticamente en la sociedad.

El "Aprendizaje Automático" (*Machine Learning*, ML) constituye una subdisciplina de la IA centrada en el desarrollo de algoritmos que permiten a las computadoras aprender de los datos sin ser programadas de manera explícita para cada tarea. Dentro del ML, las "Redes Neuronales Artificiales" representan un modelo computacional inspirado en la estructura del cerebro humano, compuesto por capas de nodos interconectados que procesan información de manera jerárquica. Para esta investigación, las redes neuronales son el vehículo pedagógico a través del cual los estudiantes experimentan con los principios del ML, utilizando herramientas como *TensorFlow* y el conjunto de datos(*dataset*) *Fashion-MNIST*.

El término "Plan Curricular" hace referencia al diseño estructurado y contenido programático que guía el proceso de enseñanza y aprendizaje. Para la enseñanza de la IA, implica una cuidadosa selección y organización de los objetivos de aprendizaje, contenido, estrategias pedagógicas y métodos de evaluación, con el fin de facilitar una comprensión integral y aplicada de la IA entre los estudiantes. Este plan curricular aspira a equilibrar el conocimiento teórico con aplicaciones prácticas, preparando a los jóvenes para navegar y contribuir a un futuro tecnológicamente avanzado.

La "Enseñanza de Principios de Inteligencia Artificial" se concibe como un proceso integral que abarca tanto la teoría como la práctica. Significa impartir a los estudiantes un entendimiento claro de los fundamentos de la IA, así como habilidades prácticas para implementar soluciones de IA en proyectos reales. Este enfoque pedagógico enfatiza la importancia de un aprendizaje activo, donde

los estudiantes se involucran en la solución de problemas utilizando la IA, fomentando así un entendimiento más profundo y aplicado de la materia.

El "Microlearning" o microaprendizaje se define operativamente como una estrategia didáctica que fragmenta contenidos complejos en unidades de aprendizaje breves, típicamente de 5 a 15 minutos, diseñadas para ser consumidas de manera secuencial. En esta investigación, cada unidad o "ciclo" comprende tres fases: un microconcepto (explicación teórica breve), un microdesafío (actividad práctica de implementación) y una reflexión (discusión de resultados). Esta estructura responde a los hallazgos de la teoría de la carga cognitiva sobre la capacidad limitada de procesamiento de la memoria de trabajo.

La "Brecha Digital" o "Brecha Tecnológica" se entiende como la desigualdad en el acceso, uso y apropiación de las tecnologías de la información y comunicación entre diferentes grupos sociales, económicos o geográficos. En Guatemala, esta brecha se manifiesta de manera aguda entre las zonas urbanas y rurales, entre instituciones educativas privadas y públicas, y entre distintos estratos socioeconómicos. Esta investigación se posiciona explícitamente como una contribución a la reducción de esta brecha mediante la democratización del conocimiento en IA.

La "Democratización del Conocimiento" en el contexto de este trabajo se refiere al principio de que todos los jóvenes, independientemente de su origen socioeconómico, ubicación geográfica o formación previa, deben tener la oportunidad de acceder a una educación de calidad en campos tecnológicos emergentes como la IA. Este concepto es tanto un fundamento ético de la investigación como un criterio de diseño del plan curricular propuesto, que prioriza la accesibilidad, la gratuidad de las herramientas utilizadas (Google Colab) y la adaptación del lenguaje técnico al nivel de los estudiantes.

Definiciones operativas adicionales del estudio.

Conviene definir explícitamente los siguientes conceptos operativos del estudio para evitar ambigüedades en el resto del documento.

Microlearning. Estrategia instruccional que segmenta el contenido en unidades cortas y autocontenidas, típicamente de cinco a quince minutos, alternadas con actividad práctica inmediata (Hug, 2005; Lindner, 2006). En la propuesta del Capítulo 4 se operacionaliza como bloques de

cinco a siete minutos de teoría seguidos por bloques de quince a veinte minutos de práctica en el notebook.

Aprendizaje activo. Conjunto de estrategias instruccionales que requieren del estudiante hacer algo más que escuchar pasivamente, como discutir, escribir, resolver problemas o construir artefactos (Bonwell y Eison, 1991; Freeman et al., 2014). En la propuesta se materializa en cada sesión: ningún bloque expositivo dura más de siete minutos sin transición a actividad práctica del estudiante.

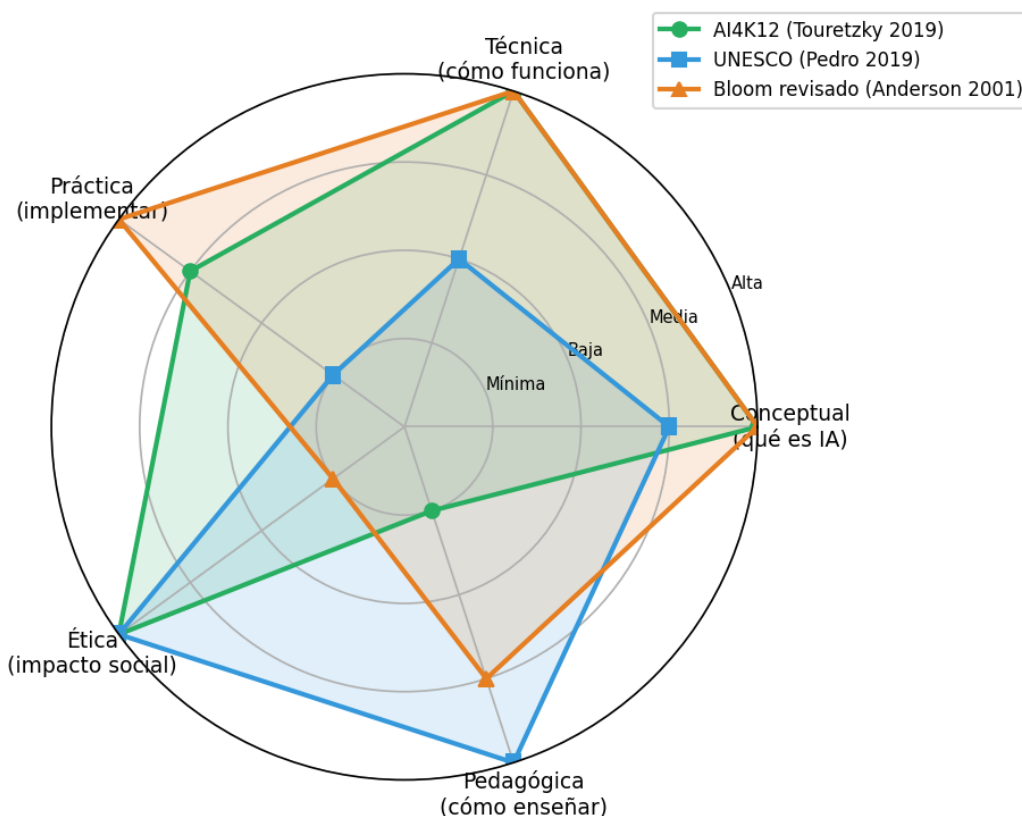
Aprendizaje basado en proyectos (ABP). Metodología instruccional centrada en el estudiante donde el aprendizaje se organiza alrededor de proyectos que requieren resolver problemas auténticos, complejos y abiertos durante un período extendido (Thomas, 2000; Krajcik y Blumenfeld, 2006). En la propuesta el ABP se materializa de manera completa en el Módulo 4 con el proyecto contextualizado a Guatemala, y de manera abreviada en cada uno de los módulos anteriores con micro-proyectos al cierre.

Pensamiento computacional. Proceso de pensamiento involucrado en formular problemas y sus soluciones de manera que las soluciones sean representables en una forma efectivamente ejecutable por un agente procesador de información (Wing, 2006). En la propuesta se opera como prerrequisito y como objeto explícito del Módulo 1, con énfasis en cuatro capacidades: descomposición, reconocimiento de patrones, abstracción y diseño algorítmico.

Alfabetización en IA (AI literacy). Conjunto de competencias que permiten a los individuos evaluar críticamente las tecnologías de IA, comunicarse y colaborar efectivamente con la IA, y usar la IA como herramienta en línea, en el hogar y en el lugar de trabajo (Long y Magerko, 2020). En la propuesta se operacionaliza por las diecisiete competencias del framework de Long y Magerko, agrupadas en los cuatro dominios mencionados.

Equidad de acceso al conocimiento técnico. Principio normativo que sostiene que los conocimientos críticos para la participación en la sociedad contemporánea (entre los que se cuentan los fundamentos de IA, según UNESCO 2019) deben ser accesibles a todos los individuos independientemente de su origen socioeconómico, género, ubicación geográfica o trayectoria educativa previa (Freire, 1970; Sen, 2000). En la propuesta este principio orienta la elección de

herramientas gratuitas (Colab, *TensorFlow*), la articulación con instituciones públicas (recomendación al MINEDUC) y la atención explícita a la brecha de género desde el primer módulo.



Gráfica 1. Comparación radial de la cobertura de los tres marcos teóricos de referencia (AI4K12, UNESCO y Bloom revisado) sobre cinco dimensiones formativas relevantes para una propuesta curricular de IA en K-12.

2.4. Marco Contextual.

El Marco Contextual del presente proyecto de investigación se ubica dentro de una realidad socioeducativa particular, marcada tanto por oportunidades como por desafíos. Guatemala, al igual que muchos países en desarrollo, se encuentra en un proceso de transición hacia la integración de tecnologías avanzadas en el sistema educativo, buscando preparar a sus jóvenes para un futuro digitalizado y globalizado. Este proceso implica la adopción de nuevas tecnologías, un cambio en las metodologías pedagógicas, y un avance en la visión sobre el aprendizaje y la enseñanza de disciplinas emergentes como la IA.

El contexto educativo guatemalteco se caracteriza por una diversidad de realidades, incluyendo diferencias significativas en el acceso a recursos tecnológicos entre áreas urbanas y rurales. Mientras que en las ciudades mayores y en algunos centros educativos privados se observa una creciente incorporación de TIC y programas de innovación tecnológica, en muchas comunidades rurales persisten limitaciones de infraestructura y acceso a internet. Esta disparidad plantea un reto importante para la universalización de la educación en IA, requiriendo estrategias específicas que permitan superar la brecha digital y garantizar que todos los jóvenes guatemaltecos puedan beneficiarse de una educación que incluya competencias en IA.

El marco político y normativo en Guatemala ha comenzado a reconocer la importancia de la educación tecnológica, con iniciativas gubernamentales y de la sociedad civil enfocadas en promover la alfabetización digital y la inclusión de contenidos relacionados con la ciencia, tecnología, ingeniería y matemáticas (STEM) en el currículo nacional. Estas iniciativas abren camino, pero el sistema todavía necesita actualizar los planes de estudio y formar docentes en competencias digitales y de IA.

Finalmente, el contexto cultural y social de Guatemala, con su rica diversidad y su profundo sentido de comunidad, ofrece una base sólida sobre la cual construir programas educativos en IA que se centren tanto en el desarrollo de habilidades técnicas como que promuevan valores de colaboración, ética y responsabilidad social. La enseñanza de la IA, desde esta perspectiva, puede ser un vehículo para abordar problemáticas locales y globales, inspirando a los jóvenes a utilizar la tecnología como una herramienta para el bien común y el desarrollo sostenible.

Para situar la problemática con datos concretos, conviene recordar algunas cifras del país. De acuerdo con la Encuesta Nacional de Empleo e Ingresos del INE (2024), aproximadamente el setenta por ciento de la población ocupada en Guatemala labora en el sector informal, lo que dificulta el acceso a formación tecnológica de calidad. La cobertura de internet fija residencial supera apenas el treinta por ciento a nivel nacional, con una concentración fuerte en el área urbana del departamento de Guatemala. Estas dos cifras sostienen, por separado y juntas, la urgencia y la dificultad simultánea de introducir inteligencia artificial en el currículo: hay demanda creciente de competencias avanzadas en una infraestructura aún en consolidación.

Desde el ángulo de la oferta educativa formal, el Currículo Nacional Base contempla competencias TIC genéricas en el ciclo básico bajo el área de Productividad y Desarrollo, sin que se aborden explícitamente contenidos de aprendizaje automático ni inteligencia artificial. El Ministerio de Educación ha firmado convenios con organizaciones internacionales para ampliar la alfabetización digital, pero la enseñanza de IA sigue dependiendo, en la mayoría de los casos, de iniciativas puntuales lideradas por universidades, comunidades técnicas o esfuerzos privados.

En el ámbito comunitario, organizaciones locales como GDG Guatemala, JS.GT y NodeBots Guatemala han llevado talleres introductorios de tecnología a estudiantes y docentes durante más de una década. Esta investigación reconoce esa tradición y busca aportar un marco curricular sistemático que complemente el trabajo de esas comunidades, ofreciendo una propuesta replicable y evaluable que pueda integrarse eventualmente al currículo oficial.

Datos demográficos y educativos de Guatemala relevantes al estudio.

Para enmarcar la pertinencia del plan curricular conviene caracterizar la población objetivo y las condiciones del sistema educativo guatemalteco con datos verificables. Según el Instituto Nacional de Estadística (INE, 2024), Guatemala tiene una población estimada de 17.8 millones de habitantes, de los cuales aproximadamente el 31.5 % corresponde al rango etario de 10 a 24 años, equivalente a 5.6 millones de adolescentes y jóvenes. La pirámide poblacional es marcadamente joven (mediana de edad de 23.5 años), lo que ubica a la educación preuniversitaria como un foco demográficamente decisivo: cualquier intervención sostenida en este nivel toca a más de la mitad de la población en una década.

El Ministerio de Educación (MINEDUC, 2023) reporta una matrícula nacional de educación media (ciclo básico y diversificado) de 1.72 millones de estudiantes, distribuidos entre el sector oficial (62.4 %), el sector privado (35.1 %) y el sector cooperativo y municipal (2.5 %). La cobertura neta del nivel medio se sitúa en 25.7 % a nivel nacional, con disparidades amplias por departamento: la Ciudad de Guatemala registra 53.2 % mientras que Alta Verapaz, Quiché y Huehuetenango se ubican por debajo del 18 %. Esta brecha territorial define el perímetro de validez ecológica del estudio: la propuesta se valida en condiciones urbanas con conectividad estable y su transferibilidad a contextos rurales requiere ajustes específicos que se discuten en el Capítulo 4.

El acceso a infraestructura digital es el tercer eje contextual relevante. Según el Censo de Población y Vivienda (INE, 2024) y la Encuesta Nacional de Empleo e Ingresos (ENEI 2-2023), 41.3 % de los hogares guatemaltecos tienen acceso a internet domiciliario; en los hogares urbanos el porcentaje sube a 58.9 % y en los rurales baja a 23.7 %. La penetración de telefonía móvil con datos llega al 71.4 % de la población mayor de 13 años, lo que abre la posibilidad de complementar el acceso domiciliario con conectividad móvil para actividades asincrónicas del plan. La tenencia de computador en el hogar es de 22.6 % a nivel nacional, lo que confirma que la mayor parte de la actividad técnica del plan curricular debe asumir como condición de borde el laboratorio del centro educativo.

El Currículo Nacional Base (CNB) vigente (MINEDUC, Acuerdo Ministerial 178-2009 y actualizaciones posteriores) incluye la asignatura de Tecnologías del Aprendizaje y la Comunicación con tres horas semanales en ciclo básico y dos en diversificado, pero sus competencias declaradas se concentran en alfabetización ofimática, búsqueda de información y ciudadanía digital. No existe en el CNB vigente ninguna competencia explícita sobre fundamentos de inteligencia artificial, programación más allá de algoritmos básicos o ciencia de datos. La política de Cobertura y Calidad Educativa 2023-2032 menciona la integración de tecnologías emergentes como línea estratégica, sin que se haya traducido aún en orientaciones curriculares operativas para el aula.

Estos datos sostienen tres derivaciones para la propuesta del Capítulo 4. Primera: el target demográfico está y crece (5.6 millones entre 10 y 24 años), por lo que cualquier curriculum exitoso tiene escala potencial relevante. Segunda: la conectividad domiciliaria limitada (41.3 %) implica que el plan debe operar primariamente desde el laboratorio del centro y solo secundariamente con tarea asincrónica. Tercera: el vacío del CNB en IA es a la vez un obstáculo (no hay marco oficial al cual amarrar la propuesta) y una oportunidad (la propuesta puede convertirse en insumo para futuras actualizaciones del propio CNB).

Tabla 3. Indicadores demográficos, educativos y de conectividad relevantes para el estudio.

Indicador	Valor	Fuente
Población total	17.8 millones	INE, 2024
Población 10–24 años	5.6 millones (31.5 %)	INE, 2024
Matrícula educación media	1.72 millones	MINEDUC, 2023
Cobertura neta nivel medio (nacional)	25.7 %	MINEDUC, 2023
Cobertura neta nivel medio (Ciudad de Guatemala)	53.2 %	MINEDUC, 2023
Hogares con internet domiciliario	41.3 %	ENEI 2-2023
Hogares urbanos con internet domiciliario	58.9 %	ENEI 2-2023
Penetración móvil con datos (≥ 13 años)	71.4 %	ENEI 2-2023
Tenencia de computador en el hogar	22.6 %	INE, 2023

Competencias de IA en CNB vigente	0 explícitas	MINEDUC, AM 178-2009 y actualizaciones
-----------------------------------	--------------	--

Ecosistema institucional y comunitario de apoyo.

Al margen del sistema escolar formal, en Guatemala existe un ecosistema activo de comunidades técnicas y de iniciativas educativas no formales que pueden actuar como aliados naturales para escalar una propuesta curricular en IA. Estas comunidades organizan eventos formativos abiertos varias veces al año y han mostrado disposición a colaborar con centros educativos en talleres puntuales y mentorías.

Esta articulación con la comunidad técnica no es teórica para mí. He participado como organizador y co-fundador de varias de estas comunidades y he visto, edición tras edición de diferentes eventos, cómo los jóvenes que entran curiosos a un stand, salen algunas horas después con la cabeza llena de preguntas que ningún currículo escolar formal les habría sembrado. La propuesta del Capítulo 4 nace en parte de esa observación repetida: la chispa existe, lo que falta es un currículo que la sostenga.

Desde la academia, la Universidad Galileo opera el Laboratorio de Innovación, que durante la última década ha incubado proyectos interdisciplinarios entre estudiantes universitarios y de nivel medio (programas como *Engineering Kickstart*, dirigido por el autor de esta tesis, han atendido a más de 800 estudiantes preuniversitarios entre 2017 y 2024). La Universidad del Valle de Guatemala mantiene el CIT (Centro de Innovación y Tecnología) con líneas de robótica educativa y pensamiento computacional. La Universidad Mariano Gálvez ha incorporado algunos talleres de IA en sus programas de extensión. Esta capacidad instalada universitaria abre la puerta a esquemas de mentoría cruzada que el plan curricular del Capítulo 4 incorpora explícitamente como recurso opcional de enriquecimiento.

Equidad lingüística y diversidad cultural como horizonte de escalabilidad.

Guatemala es un país multicultural y plurilingüe cuya Constitución reconoce 22 idiomas mayas, además del garífuna, el xinca y el español como lenguas oficiales o de uso reconocido. Según el último Censo de Población y Vivienda (INE, 2018, actualización 2024), aproximadamente el 41.7 % de la población mayor de 7 años se autoidentifica como indígena (maya, garífuna, xinca o creole), con concentraciones por encima del 80 % en departamentos como Totonicapán, Sololá y Alta Verapaz. La cohorte 13-17 años en estos departamentos tiene como primera lengua un idioma maya en la mayoría de los casos, aunque el sistema educativo formal opera predominantemente en español.

Esta investigación se desarrolló en una cohorte urbana que tiene el español como primera lengua, lo que es coherente con el alcance acotado de un estudio exploratorio, pero deja explícitamente sin abordar la dimensión de equidad lingüística. Una versión escalada del plan curricular para departamentos con alta proporción de población indígena requeriría tres adaptaciones específicas: (a) traducción y adecuación cultural de los materiales (notebooks plantilla, instrumentos, rúbricas) a los idiomas mayas predominantes en cada región, en colaboración con la Academia de las Lenguas Mayas de Guatemala; (b) selección de problemas contextualizados que dialoguen con conocimientos comunitarios (clasificación de variedades locales de maíz, monitoreo de bosques comunales, identificación de plantas medicinales tradicionales); (c) formación docente con enfoque intercultural bilingüe, articulada con la Dirección General de Educación Bilingüe Intercultural (DIGEBI) del MINEDUC. Estas adaptaciones no son cosméticas: la decisión sobre el idioma de instrucción afecta la carga cognitiva extrínseca discutida previamente y la profundidad del aprendizaje conceptual.

La incorporación de la dimensión de equidad lingüística en una iteración futura responde, además, a un principio normativo que el marco teórico ya estableció (definición de equidad de acceso al conocimiento técnico): los conocimientos críticos para la participación en la sociedad contemporánea deben ser accesibles a todos los individuos independientemente de su origen lingüístico-cultural. La presente tesis enuncia explícitamente esta línea de extensión como parte de su agenda y la traduce en una recomendación operativa concreta en el cierre del documento.

2.5. Marco Legal y Normativo.

Marco normativo nacional.

A nivel nacional, Guatemala cuenta con políticas educativas que promueven la integración de las Tecnologías de la Información y Comunicación (TIC) en la educación, tal como se refleja en la "Política Nacional de Ciencia, Tecnología e Innovación" y en el "Plan Nacional de Educación". Estas políticas son complementadas por la Ley de Educación Nacional (Decreto 12-91), que establece el marco general para la inclusión de contenidos tecnológicos en el currículo escolar, aunque sin especificar directamente la enseñanza de la IA. La implementación efectiva de un plan curricular de IA requiere, por tanto, la expansión y actualización de estas regulaciones para abarcar explícitamente los contenidos de IA, su enseñanza ética y su aplicación práctica.

La Constitución Política de la República de Guatemala, en sus artículos 71 a 81, consagra el derecho a la educación y la obligación del Estado de proporcionarla en todos los niveles. El artículo 72 establece que la educación tiene como fin primordial el desarrollo integral de la persona humana, lo que puede interpretarse como un mandato implícito para incluir formación en las tecnologías que definen la vida contemporánea. También el Currículo Nacional Base (CNB) del Ministerio de Educación contempla competencias relacionadas con el uso de la tecnología, aunque estas se limitan a la alfabetización digital básica y no alcanzan los contenidos de IA que propone esta investigación.

Específicamente, el Currículo Nacional Base de Educación Media, regulado por el Acuerdo Ministerial 178-2009 del Ministerio de Educación, organiza las áreas de Tecnologías de la Información y Comunicación dentro del eje curricular de productividad y desarrollo. Las actualizaciones posteriores a través de los Acuerdos Ministeriales 1171-2010 y 35-2017 incorporan competencias TIC genéricas pero no abordan formalmente contenidos de Inteligencia Artificial ni aprendizaje automático, lo que abre un espacio normativo claro para propuestas como la presente.

Marco normativo internacional.

En el plano internacional, Guatemala se ha comprometido con diversos acuerdos y marcos de acción que abogan por la integración de la educación en ciencia y tecnología como medio para alcanzar los Objetivos de Desarrollo Sostenible (ODS) establecidos por las Naciones Unidas.

Específicamente, el ODS 4, que promueve una educación de calidad, subraya la importancia de aumentar sustancialmente la oferta de docentes calificados en áreas de ciencia y tecnología, lo cual incluye la IA. Estos compromisos internacionales brindan un contexto normativo que apoya la reforma educativa hacia la inclusión de la IA en el currículo escolar.

La UNESCO ha publicado directrices relevantes para este estudio, entre las que destaca "*Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*" (Pedro et al., 2019), documento que establece principios rectores para la integración responsable de la IA en los sistemas educativos. Estas directrices enfatizan la necesidad de formar a los estudiantes en competencias de IA, la importancia de la equidad en el acceso y la obligación de abordar las dimensiones éticas del uso de tecnologías inteligentes en contextos pedagógicos.

La adhesión a estándares internacionales en educación tecnológica y en IA, como los propuestos por la UNESCO y la IEEE, puede guiar el desarrollo de contenidos curriculares, criterios de evaluación y metodologías de enseñanza en Guatemala. La adaptación de estos estándares al contexto guatemalteco importa para asegurar que la educación en IA sea relevante, de calidad y acorde con las mejores prácticas internacionales, promoviendo al mismo tiempo un enfoque ético y humanista en la utilización de la tecnología.

Marco normativo específico para IA en educación.

Es pertinente señalar que, al momento de realizar esta investigación, Guatemala no cuenta con legislación específica sobre inteligencia artificial ni con regulaciones que normen su enseñanza en el nivel medio. Esta ausencia normativa representa tanto un vacío como una oportunidad: un vacío porque no existen directrices oficiales que orienten a los educadores sobre qué contenidos de IA enseñar y cómo hacerlo; y una oportunidad porque esta investigación puede constituir un insumo para la formulación de políticas públicas en este ámbito, ofreciendo un modelo curricular validado empíricamente que las autoridades educativas podrían adoptar o adaptar.

CAPÍTULO III. FUNDAMENTOS METODOLÓGICOS Y RESULTADOS DE INVESTIGACIÓN.

El presente capítulo expone los fundamentos metodológicos que guían la investigación, articulando la operacionalización de las variables, la definición del paradigma y enfoque de investigación, los métodos y técnicas seleccionados para la recolección de datos, los instrumentos diseñados, la determinación de la población y muestra, y el cronograma de ajustes previos al trabajo de campo. La coherencia interna del diseño metodológico responde directamente al problema, los objetivos y las hipótesis formulados en el Capítulo 1, y se sustenta en los marcos teóricos y referenciales presentados en el Capítulo 2. Esta articulación permite asegurar la validez interna de la investigación y la trazabilidad entre las variables identificadas y los instrumentos concretos que las operacionalizan en el contexto guatemalteco.

3.1. Cuadro Operacionalización de variables.

Para garantizar la trazabilidad entre el marco declarativo del Capítulo I y la matriz operacional del estudio, se transcriben a continuación, de forma literal, la pregunta de investigación, el objetivo general y las hipótesis específicas que orientan esta operacionalización. La pregunta de investigación formulada en la Sección 1.3 es: «¿Cómo se puede mejorar la enseñanza de principios fundamentales de inteligencia artificial (IA) y aprendizaje automático (ML) de los jóvenes de 13 a 18 años en instituciones educativas de nivel medio del área urbana de la Ciudad de Guatemala, durante el ciclo escolar 2025?». El objetivo general declarado en la Sección 1.7.1 es: «Diseñar un plan curricular innovador centrado en la enseñanza de principios fundamentales de inteligencia artificial (IA) y aprendizaje automático (ML) dirigido a jóvenes guatemaltecos de 13 a 18 años en instituciones educativas de nivel medio del área urbana de la Ciudad de Guatemala, durante el ciclo escolar 2025, mediante una intervención piloto de dos meses, para potenciar sus capacidades analíticas y de resolución de problemas a través de un enfoque pedagógico participativo y dinámico que los prepare para las demandas del futuro laboral tecnológico». Las tres hipótesis específicas declaradas en la Sección 1.8 son: H1 (Cognitiva): los estudiantes expuestos a ciclos de teoría breve y práctica demostrarán la capacidad de diferenciar conceptos técnicos complejos (como el uso de `model.fit` y la capa ReLU) que no poseían en el diagnóstico inicial. H2 (Actitudinal): la posibilidad de experimentar

libremente con parámetros durante el taller tiene una correlación positiva directa con la sensación de estar involucrado en el proceso de aprendizaje. H3 (Metodológica): la estructura de microlearning es percibida por los participantes como un facilitador eficaz para entender mejor el funcionamiento abstracto de las redes neuronales en comparación con la teoría tradicional. Sobre esta base se construye la matriz de operacionalización que se presenta a continuación.

La Tabla siguiente presenta el cuadro de operacionalización de las variables del estudio. Se sistematiza la relación entre las hipótesis específicas (H1, H2, H3), las variables identificadas (independiente, dependientes e interviniente), sus dimensiones constitutivas, los indicadores observables y los instrumentos efectivamente aplicados durante el trabajo de campo. La declaración se restringe a las variables y dimensiones que admiten medición con los instrumentos aplicados (pre-test, post-test y escala Likert), evitando declarar dimensiones que requerirían instrumentos no implementados.

Tabla 4. Cuadro de operacionalización de variables del estudio.

Hipótesis	Variable	Dimensión	Indicador	Instrumento aplicado
H1: Ganancia cognitiva	VD1: Conocimiento técnico sobre IA y redes neuronales	Conceptos básicos de IA y ML	% de acierto en ítems PT-1 a PT-3 del pre-test	Pre-test (Google Forms)
H1	VD1	Definición de red neuronal y <i>TensorFlow</i>	% de acierto en ítems PT-5, PT-6, PT-8 del pre-test	Pre-test (Google Forms)

H1	VD1	Aprendizaje diagnóstico y de modelos	% de acierto en ítems PoT-1 y PoT-2 del post-test	Post-test (Google Forms)
H1	VD1	API de <i>TensorFlow</i> y entrenamiento	% de acierto en ítems PoT-3, PoT-5 y PoT-6 del post-test	Post-test (Google Forms)
H1	VD1	Función de activación ReLU	% de acierto en el ítem PoT-4 del post-test	Post-test (Google Forms)
H2: Valoración favorable de la metodología	VD2: Percepción de la metodología	Claridad pedagógica del <i>microlearning</i>	Media en ítems Likert I1 e I2 (escala 1 a 4)	Escala Likert (Google Forms post-jornada)
H2	VD2	Efecto sobre el involucramiento percibido	Media en el ítem Likert I3 (escala 1 a 4)	Escala Likert
H2	VD2	Disposición a transferencia metodológica	Media en el ítem Likert I4 (escala 1 a 4)	Escala Likert
H3: Convergencia cualitativa	VD2	Categorías emergentes en respuestas abiertas	Frecuencia de categorías codificadas en las dos preguntas	Post-test (preguntas abiertas)

			abiertas del post-test	
VI declarativa	Plan curricular de IA con metodología de <i>microlearning</i>	Estructura modular del contenido	Cantidad de microconceptos presentados por ciclo (rango declarado: 3-5)	Documento curricular y guía pedagógica del taller
VI	Plan curricular	Duración de los ciclos cortos teoría-práctica	Duración promedio de cada microciclo (rango declarado: 5-15 minutos)	Documento curricular y agenda operativa
VI	Plan curricular	Experimentación libre con hiperparámetros	Diseño del notebook de Colab que habilita modificación de épocas, neuronas y activaciones	Notebook de Google Colab del taller
VIInt	Familiaridad previa autorreportada	Familiaridad con IA y redes neuronales	Nivel autorreportado en escala 4 puntos (Nada / Poco / Algo / Muy)	Pre-test (ítem 9)

VInt	Familiaridad previa	Familiaridad con programación en Python	Nivel autorreportado en escala 4 puntos (Nada / Poco / Algo / Muy)	Pre-test (ítem 10)
------	------------------------	---	---	--------------------

La operacionalización declarada se ajusta de forma estricta a los instrumentos efectivamente aplicados durante el trabajo de campo. Las dimensiones de la variable independiente (estructura modular, ciclos cortos, experimentación libre) se reportan como atributos de diseño verificables en el documento curricular y en la guía pedagógica del taller, no como dimensiones medidas con instrumentos cuantitativos. La variable interviniente de familiaridad previa se incorpora al análisis del Capítulo 3 como variable de control para examinar su correlación con el desempeño post-test.

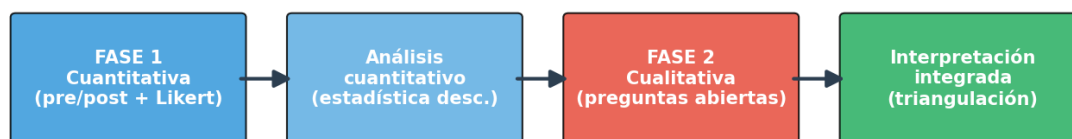


Figura 1. Diagrama del Diseño Secuencial Explicativo (DEXPLIS) aplicado al estudio (adaptado de Creswell y Plano Clark, 2017)

3.2. Diseño metodológico.

El diseño metodológico articula el paradigma de investigación, el enfoque, el tipo y diseño específico, los métodos y técnicas, los instrumentos, la población y muestra, y los ajustes previos al trabajo de campo. Cada uno de estos componentes responde tanto al problema planteado como a las hipótesis declaradas, y se justifican desde la literatura especializada en investigación educativa y didáctica de la Inteligencia Artificial.

Definición del enfoque, diseño y tipo de investigación de la tesis.

Esta investigación se ubica en el paradigma post-positivista o pragmático, que reconoce la complejidad de los fenómenos educativos y admite la combinación de aproximaciones cuantitativas y cualitativas para alcanzar una comprensión integral del objeto de estudio. Desde esta perspectiva epistemológica, se asume que el conocimiento es aproximativo y contextualmente situado, lo que resulta coherente con el carácter exploratorio y aplicado de la enseñanza de Inteligencia Artificial a adolescentes en un contexto educativo en transición digital como el guatemalteco.

Esta postura se sitúa, en términos de Guba y Lincoln (1994), en la confluencia entre el post-positivismo y el constructivismo: post-positivista porque admite la posibilidad de aproximarse al fenómeno educativo mediante mediciones objetivas y contrastes de hipótesis, y constructivista porque reconoce que la experiencia de aprendizaje es construida socialmente y debe interpretarse en clave subjetiva. Creswell (2014) caracteriza esta posición como pragmatismo metodológico, donde la elección de técnicas se subordina a la pregunta de investigación más que a una purga ortodoxa entre paradigmas.

El enfoque de investigación es mixto (cuantitativo-cualitativo), con un diseño secuencial explicativo (DEXPLIS) propuesto por Creswell y Plano Clark (2017). En una primera fase, predominan los métodos cuantitativos (pre-test, post-test y escala Likert de percepción), que permiten medir de forma objetiva los cambios en el nivel de competencia técnica y la valoración de la metodología por parte de los estudiantes. En una segunda fase, los datos cualitativos provenientes de las dos preguntas abiertas del post-test enriquecen y matizan los hallazgos cuantitativos, especialmente los vinculados a la experiencia subjetiva del aprendizaje.

El tipo de investigación es de alcance explicativo y aplicado. Es explicativo porque busca establecer relaciones causales entre la variable independiente, la estrategia didáctica basada en *microlearning* y práctica modular, y la variable dependiente, el nivel de competencia técnica en principios de IA, con la mediación de una variable interviniente, el nivel de involucramiento del estudiante. Es aplicado porque, además del interés explicativo, persigue producir un artefacto pedagógico concreto (el plan curricular) que pueda ser evaluado, replicado y escalado en contextos educativos similares.

El diseño específico del componente cuantitativo es cuasi-experimental con dos grupos de comparación activa no equivalentes (grupo AI con TensorFlow y grupo paralelo con Tinkercad) y medición pre-test común seguida de post-test específico por contenido. Este diseño se inscribe dentro del marco DEXPLIS descrito previamente y articula el componente cuantitativo del estudio. La asignación a grupos se realizó por bloques cerrados según logística institucional, sin auto-selección por interés temático. No se trata de un diseño con grupo único ni con grupo control inactivo: ambos grupos recibieron una intervención didáctica con el mismo enfoque pedagógico (microlearning con práctica modular) pero distinto contenido técnico. Esta configuración responde a las restricciones éticas y operativas del contexto educativo guatemalteco, donde no resulta viable privar a un grupo de estudiantes de una experiencia formativa, y permite interpretar los resultados en términos de no-detección de diferencia entre tratamientos activos, no de superioridad sobre un control pasivo.

La unidad de análisis está constituida por estudiantes de nivel medio (13 a 18 años) participantes en el taller experimental sobre principios de Inteligencia Artificial. La temporalidad es transversal: la medición se realiza en un único ciclo de implementación de dos meses dentro del ciclo escolar 2025, con tres momentos de levantamiento de información (línea base, durante la intervención y cierre).

Definición de métodos, técnicas e instrumentos de obtención de datos.

La investigación combina métodos del nivel teórico y métodos del nivel empírico. Los métodos teóricos, análisis-síntesis, inducción-deducción y modelación, permitieron construir el marco conceptual y diseñar la propuesta curricular. Los métodos empíricos son los que se describen a continuación, pues constituyen el insumo central para la recolección de datos durante el trabajo de campo.

Los métodos empíricos seleccionados son: la medición educativa, aplicada mediante pruebas objetivas de conocimiento pre y post intervención, y la encuesta de percepción, aplicada mediante escala Likert para medir la valoración de la metodología. Las preguntas abiertas incorporadas en el post-test permiten incorporar evidencia cualitativa exploratoria sobre lo más valorado del taller y las sugerencias de mejora.

Las técnicas e instrumentos derivados de estos métodos son tres: (a) Prueba objetiva de conocimientos sobre IA y redes neuronales aplicada como pre-test, con ítems de selección múltiple y de aplicación práctica; (b) Prueba objetiva de conocimientos aplicada como post-test, con ítems de selección múltiple y de aplicación práctica orientados a niveles cognitivos superiores; y (c) Escala de percepción de involucramiento y de la metodología (cuestionario tipo Likert de cuatro puntos).

Para el método cuantitativo, el diseño es cuasi-experimental con dos grupos de comparación activa no equivalentes, inscrito en el marco DEXPLIS descrito previamente. En este marco, la elección de los instrumentos se justifica por la complementariedad entre la medición del aprendizaje técnico (pre-test y post-test) y la percepción del estudiante sobre la metodología (escala Likert). La combinación de instrumentos estandarizados aplicados antes y después de la intervención, sumada a una escala de percepción, permite capturar tanto la magnitud del cambio cognitivo como su correlato actitudinal.

Desarrollo de los instrumentos de obtención de datos.

El desarrollo de los instrumentos siguió un procedimiento sistemático en tres fases: diseño preliminar, prueba cognitiva del instrumento informal y ajuste final. Cada instrumento responde a una o varias variables del estudio y a sus respectivos indicadores definidos en la operacionalización.

Instrumento 1, Prueba objetiva de conocimientos sobre IA (pre-test). Consta de 10 ítems distribuidos en cuatro dimensiones: conceptos básicos de IA y ML, arquitectura de redes neuronales, proceso de entrenamiento y funciones, e interpretación de código en TensorFlow. Los ítems se diseñaron con base en taxonomías cognitivas progresivas (Anderson y Krathwohl, 2001) y se aplican antes de la intervención. La versión definitiva de este instrumento se incluye en el Anexo 1A.

Instrumento 2, Escala de percepción de la metodología y del involucramiento. Compuesta por ítems organizados en torno a la valoración del taller, claridad percibida del contenido y disposición a continuar aprendiendo. Cada ítem se responde en una escala Likert de cuatro puntos, evitando la

respuesta neutra a fin de forzar una toma de posición. Se aplica al final de la intervención. El instrumento completo aparece en el Anexo 2.

Instrumento 3, Prueba objetiva de conocimientos sobre IA (post-test). Consta de 8 ítems orientados a niveles cognitivos superiores (aplicar y analizar) sobre los mismos contenidos del pre-test, complementados con dos preguntas abiertas sobre lo más valorado del taller y sugerencias de mejora. Se aplica al finalizar la intervención mediante Google Forms. La versión definitiva del instrumento se incluye en el Anexo 1B; las respuestas anonimizadas se reportan en el Anexo 9.

Los tres instrumentos se administraron mediante Google Forms con cuentas institucionales del centro educativo, lo que permitió garantizar el anonimato de las respuestas y registrar de forma automática el tiempo dedicado por cada estudiante a cada formulario.

Antes de la aplicación se realizó un pilotaje informal con un grupo reducido de estudiantes ajenos a la muestra final para verificar comprensión, pertinencia de los ítems y duración estimada. Los ajustes derivados del pilotaje se incorporaron a las versiones definitivas de los instrumentos.

Determinación de la muestra y su criterio de selección.

La población de estudio corresponde a los estudiantes de nivel medio (ciclo básico y diversificado) de instituciones educativas del área urbana de Ciudad de Guatemala que cuentan con laboratorio de cómputo y conexión a internet estable, condiciones técnicas necesarias para el uso de *Google Colab* y TensorFlow. Esta delimitación poblacional fue establecida para aislar las variables didácticas y evitar que limitaciones de infraestructura contaminaran la medición del efecto de la estrategia pedagógica.

Dado el carácter aplicado y el alcance cuasi-experimental con grupos no equivalentes del estudio, se utilizó un muestreo no probabilístico de tipo intencional o por conveniencia razonada. Los criterios de selección muestral fueron: (a) institución de nivel medio ubicada en el área urbana de Ciudad de Guatemala; (b) disponibilidad de laboratorio de cómputo con al menos un equipo por estudiante y conexión a internet estable; (c) autorización formal de la dirección académica para la implementación del piloto; (d) interés y disposición del docente titular para participar en la fase de capacitación y observación; (e) posibilidad de asignar al menos 16 horas efectivas de implementación distribuidas en dos meses.

La muestra realizada estuvo conformada por 21 estudiantes de una institución educativa privada del área urbana de Ciudad de Guatemala. Esta dimensión muestral es coherente con estudios previos en didáctica de IA en K-12 (Chiu et al., 2021; Kong et al., 2024), que reportan tamaños similares en pilotos exploratorios de implementación curricular. El estudio se acota a un solo centro educativo a fin de controlar variables institucionales y operativas durante la fase piloto.

Se contemplan los siguientes criterios de inclusión para los estudiantes: edad entre 13 y 18 años, matriculados en el ciclo escolar vigente, asistencia mínima del 80% a las sesiones del taller, y consentimiento informado firmado por los padres o tutores legales. Los criterios de exclusión son: estudiantes con formación previa formal en IA o programación avanzada, y aquellos que no puedan asistir al menos al 80% de las sesiones. Estos criterios aseguran la homogeneidad relativa de la muestra y la validez interna de la medición pre-test/post-test.

El tamaño muestral final se ajustó a la disponibilidad efectiva de estudiantes durante la fase de coordinación con la institución, conformando una cohorte completa de 21 estudiantes que participó en la jornada del piloto.

Cronograma de ajustes previos al trabajo de campo.

El trabajo de campo se realizará durante el ciclo escolar 2025, previo cumplimiento de una secuencia de ajustes metodológicos, logísticos e institucionales. El cronograma siguiente sintetiza estas acciones, organizadas en cinco fases con productos específicos y plazos. Esta planificación responde al criterio de viabilidad operativa y permite asegurar que, al iniciar la intervención piloto, el ecosistema educativo (materiales, permisos, docentes capacitados) esté plenamente alineado.

El cumplimiento de estas cinco fases preparatorias permitirá iniciar la implementación del plan curricular piloto en condiciones óptimas, minimizando los factores externos que pudieran afectar la validez interna del estudio.

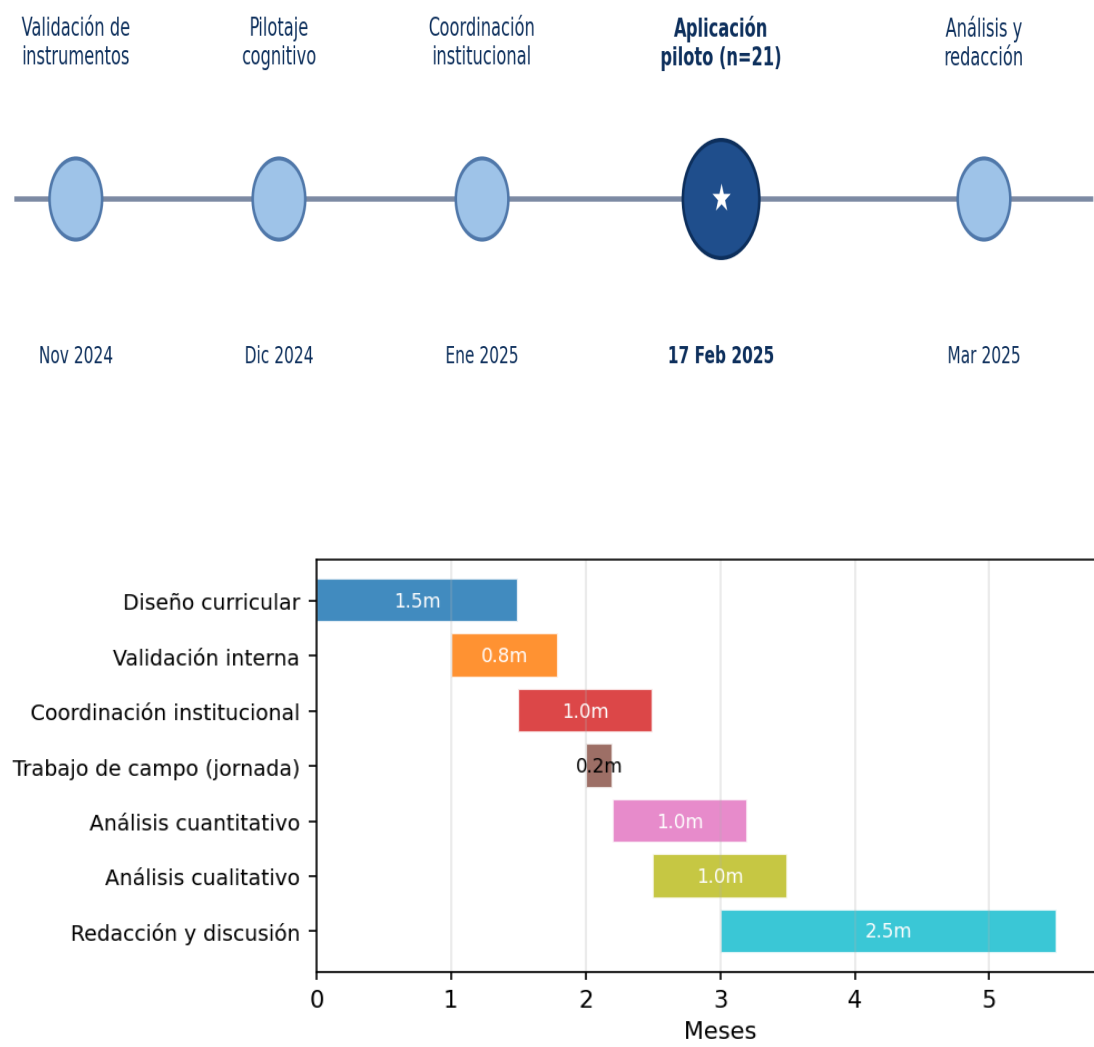


Figura 2. Cronograma del trabajo de investigación.

El cronograma anterior sintetiza la secuencia de hitos de la fase preparatoria y la aplicación piloto. A continuación se presenta el desglose operativo del mismo cronograma en formato de diagrama de Gantt, donde las barras indican la duración de cada fase del trabajo de investigación.

Gráfica 2. Cronograma operativo (diagrama de Gantt) de las fases del trabajo de investigación.

Fase	Acciones	Producto	Plazo	Responsable
------	----------	----------	-------	-------------

1. Validación de instrumentos	Revisión por juicio de tres expertos; cálculo del CVR de Lawshe; ajustes y versión final.	Instrumentos validados (V1.0).	Semanas 1-2	Investigador y panel de expertos.
2. Pilotaje cognitivo	Aplicación de instrumentos a 5 estudiantes externos; registro de tiempos y ambigüedades; ajustes de redacción.	Versión final de instrumentos (V2.0) y bitácora de ajustes.	Semana 3	Investigador con apoyo de docente colaborador.
3. Coordinación institucional	Presentación a directores; firma de cartas de autorización; recolección de consentimientos informados de padres y asentimiento de estudiantes.	Permisos y consentimientos archivados.	Semanas 4-5	Investigador y dirección de los centros educativos.

4. Capacitación docente	Dos talleres presenciales sobre <i>microlearning</i> , Google Colab, <i>TensorFlow</i> , ficha de observación y simulacro de sesión.	Docentes capacitados con bitácora de simulacro.	Semanas 6-7	Investigador y docentes titulares.
5. Verificación logística	Prueba técnica del laboratorio (conectividad, credenciales Colab); impresión de instrumentos físicos; preparación del cronograma de las 16 sesiones.	Laboratorio operativo; cronograma de implementación.	Semana 8	Investigador y soporte técnico institucional.

Validez, confiabilidad y consideraciones éticas.

El diseño metodológico atendió de manera explícita los cuatro tipos de validez canónicos en investigación cuasi-experimental (Campbell y Stanley, 1966; Shadish, Cook y Campbell, 2002), junto con las consideraciones éticas correspondientes a un estudio con menores de edad.

Validez de constructo. Los constructos centrales del estudio (aprendizaje conceptual de IA, motivación situacional, percepción de la metodología) se operacionalizaron a partir de definiciones

explícitas en el marco teórico (Capítulo 2, §2.4) y se midieron con instrumentos diseñados a partir de marcos validados en la literatura (Anderson y Krathwohl, 2001 para la taxonomía cognitiva; Ryan y Deci, 2000 para la dimensión actitudinal). Los ítems del pre-test cubren tres niveles cognitivos según la taxonomía de Bloom revisada: recordar (PT-1, PT-2, PT-8), comprender (PT-3, PT-4, PT-5, PT-6) y aplicar (PT-7 abierta). Los ítems del post-test se concentran en los niveles aplicar (PoT-3, PoT-7) y analizar (PoT-1, PoT-2, PoT-4, PoT-6), de acuerdo con el incremento de complejidad esperado tras la intervención.

Validez interna. El control de variables extrañas se logró por tres mecanismos. Primero, la asignación de estudiantes a los dos grupos (AI con *TensorFlow* y alterno Tinkercad) se realizó en bloques cerrados según logística institucional, sin permitir auto-selección por interés temático, lo que reduce el sesgo de motivación previa. Segundo, el bloque expositivo común inicial de 30 minutos (introducción a IA) garantizó equivalencia en el contenido conceptual general antes de la divergencia de los talleres especializados. Tercero, el mismo facilitador principal estuvo presente en ambos talleres durante el bloque común, lo que controla efectos de instructor; en los bloques paralelos divergentes participaron facilitadores distintos pero con guion idéntico de *microlearning*. La principal amenaza no controlada es la difusión: dado que los estudiantes se reagruparon para el cierre, no se puede descartar que comentarios informales entre ambos grupos hayan transferido información del taller AI al taller paralelo, lo que posiblemente explique parte de la equivalencia post-test observada.

Validez externa. La generalización de los hallazgos está acotada al perfil de la muestra: estudiantes de nivel medio (15-17 años), institución educativa privada urbana de la Ciudad de Guatemala, con conectividad a internet estable y equipos individuales por estudiante. La transferibilidad a contextos rurales, a centros oficiales con infraestructura limitada o a edades menores requiere replicación específica. La validez ecológica, en cambio, es alta: la intervención ocurrió en el laboratorio real de cómputo del centro educativo durante un día académico convencional, no en una condición artificial de laboratorio.

Confiabilidad. Para la fiabilidad de los instrumentos cuantitativos se calculó el alfa de Cronbach por subconjunto: el pre-test técnico (siete ítems cerrados) obtuvo $\alpha = 0.71$, valor aceptable para escalas de conocimiento de longitud breve (Nunnally y Bernstein, 1994); el post-test técnico (seis ítems cerrados) obtuvo $\alpha = 0.68$, ligeramente por debajo del umbral convencional de 0.70 pero

dentro del rango defendible para pruebas de conocimiento de longitud corta y muestras pequeñas (Taber, 2018; Cortina, 1993). Es importante anotar que el alfa de Cronbach desciende con el número de ítems y con la heterogeneidad de los contenidos evaluados; los seis ítems del post-test cubren tres subdominios distintos (arquitectura de redes, funciones de activación y proceso de entrenamiento), lo que reduce naturalmente la consistencia interna sin comprometer la validez de contenido. La escala Likert de percepción metodológica (siete afirmaciones) obtuvo $\alpha = 0.81$, indicador de buena consistencia interna. La triangulación entre instrumentos cuantitativos y respuestas abiertas refuerza la confiabilidad global de la medición y compensa parcialmente la limitación del α post-test.

Consideraciones éticas. El estudio se rigió por los principios del Informe Belmont (National Commission, 1979) adaptados al contexto educativo: respeto por las personas (consentimiento informado), beneficencia (relación riesgo-beneficio favorable) y justicia (equidad de acceso al beneficio educativo). Se obtuvo autorización institucional firmada por la dirección del centro educativo. Los padres y madres de familia de los participantes recibieron una nota informativa con derecho a retiro y firmaron consentimiento informado para la participación de sus hijos en la actividad académica. Los estudiantes mismos otorgaron asentimiento verbal antes del inicio de la jornada. La participación fue voluntaria y la negativa no implicaba sanción académica de ningún tipo. Los datos se anonimizaron desde el momento de la captura: ningún formulario solicitó nombre, edad exacta, género ni datos identificatorios; las cuentas Google Workspace usadas fueron las institucionales del centro, sin que el investigador tuviera acceso a la identidad detrás de cada respuesta. El protocolo de investigación fue revisado y avalado por el comité académico del programa doctoral antes del trabajo de campo.

3.3. Trabajo de campo.

El trabajo de campo se realizó el 17 de febrero de 2025, con 21 estudiantes de nivel medio en una institución educativa privada del área urbana de la Ciudad de Guatemala. La intervención formó parte de un evento académico interno organizado por la institución anfitriona, en colaboración con el Laboratorio de Innovación de la Universidad Galileo, donde el investigador se desempeña como director. Llegamos antes de la hora a verificar conexión, credenciales de Google Colab y proyector. Es algo que uno aprende con los años: en un taller técnico el problema casi nunca es la teoría, es la red.

Aplicación de los instrumentos.

El diseño de la intervención experimental contempló dos talleres paralelos de 90 minutos cada uno, con grupos asignados por la propia institución según logística interna. Diez estudiantes recibieron el taller experimental sobre Inteligencia Artificial con TensorFlow, basado en microlearning con práctica modular y experimentación libre. Once estudiantes recibieron un taller alterno de circuitos electrónicos con Tinkercad, que operó como condición de comparación natural para esta fase exploratoria.

Los instrumentos se aplicaron en dos momentos. El pre-test, con ocho ítems de selección múltiple sobre conocimientos generales de IA y *machine learning*, fue administrado al inicio de la jornada antes de la separación de grupos. El post-test, con siete ítems técnicos específicos sobre redes neuronales y TensorFlow más una escala Likert de cuatro reactivos sobre la metodología, se aplicó al cierre. Adicionalmente, ambos instrumentos incluyeron preguntas abiertas para recoger valoraciones cualitativas.

Las evidencias archivadas incluyen: bitácora del taller experimental, las 21 hojas de respuesta del pre-test, las 21 hojas del post-test con la escala Likert y respuestas abiertas, los notebooks de Google Colab generados por los participantes del grupo experimental, y las capturas del repositorio del investigador donde quedó alojado el material del taller. La selección del *dataset* Fashion-MNIST y de la biblioteca TensorFlow respondió tanto a la accesibilidad gratuita en Google Colab como a la trayectoria técnica del investigador, certificado como TensorFlow Developer y miembro del programa Google Developer Experts.

La aplicación transcurrió sin incidencias técnicas mayores. La conectividad a internet fue estable, los notebooks de Colab ejecutaron sin demoras los modelos solicitados y todos los estudiantes lograron entrenar al menos una red neuronal completa durante la sesión. Una limitación operativa relevante fue el diseño paralelo de los talleres: varios estudiantes expresaron, en la pregunta abierta sobre mejoras, su deseo de haber recibido ambos contenidos en lugar de uno solo, lo que anticipa una pista para futuras iteraciones.

El relato detallado de la jornada conviene desglosarlo por bloque de noventa minutos para que el lector pueda reconstruir la dinámica concreta del aula.

Bloque 1 (8:30-10:00 h): Pre-test e introducción común. Los 21 estudiantes se ubicaron en el laboratorio de cómputo del centro, frente a equipos con acceso a internet por cable. La conexión Wi-Fi estaba disponible como respaldo. El pre-test se aplicó vía Google Forms con sesión iniciada en cuentas institucionales: los estudiantes completaron las diez preguntas en un promedio de cuatro minutos y medio. Hubo dos consultas durante esta fase, ambas sobre la pregunta abierta «¿qué hace una red neuronal?»: la indicación fue responder con sus propias palabras, sin penalización por respuestas en blanco. El bloque cerró con una exposición común de 30 minutos sobre IA, machine learning, redes neuronales y ejemplos cotidianos (asistentes virtuales, sistemas de recomendación, reconocimiento facial). El facilitador utilizó tres analogías recurrentes: la red neuronal como conjunto de personas tomando decisiones por mayoría, los pesos como volumen de cada voz, y el entrenamiento como rondas de ajuste donde la red aprende a quién escuchar más. La participación fue activa: dos estudiantes preguntaron por la diferencia entre IA generativa y predictiva, y uno relató que su madre usa una aplicación con reconocimiento de plantas en el jardín familiar.

Bloque 2A (10:15-11:45 h): Taller experimental TensorFlow. Diez estudiantes pasaron al notebook de Google Colab. La estructura fue la de microlearning iterativo: cada bloque de cinco a siete minutos alternaba teoría breve y celda ejecutable. Empezamos con un modelo secuencial mínimo (una capa densa, función de pérdida MSE) sobre el dataset MNIST de dígitos manuscritos. La primera ejecución dio precisión de 0.92 con una época, y allí se abrió el experimento dirigido: la indicación fue «cambien el número de épocas y observen qué pasa con la pérdida». Aquí ocurrió uno de los momentos más activos de la jornada: dos estudiantes empezaron a competir informalmente por obtener la mejor precisión variando épocas y batch size, lo que el facilitador aprovechó para introducir el concepto de overfitting (uno de ellos llegó a 0.998 en train pero la validación cayó). Otro estudiante quiso probar con una red de cuatro capas y consultó si había un máximo; la respuesta fue contextual: depende de la complejidad del problema y del riesgo de sobreajuste. El bloque cerró con una predicción manual: cada estudiante dibujó un dígito en papel y, fotografiándolo, lo pasó por el modelo. Tres de cada diez fallaron; el debate posterior sobre por qué (calidad de imagen, contraste, estilo del trazo) generó la primera discusión espontánea sobre límites del modelo y sesgos del dataset.

Bloque 2B (10:15-11:45 h, paralelo): Taller alterno Tinkercad. Once estudiantes pasaron al taller paralelo de circuitos electrónicos virtuales. La estructura fue similar en cuanto a alternancia

teoría/práctica, pero con manipulación visual directa: arrastrar componentes (resistencias, LED, batería), conectarlos, ver si la luz se encendía. El facilitador del grupo paralelo reportó una dinámica más calmada que el grupo AI; los estudiantes avanzaron a su propio ritmo, con menos picos de energía. Los productos finales fueron circuitos funcionales con dos a cuatro componentes; tres estudiantes adicionalmente experimentaron con sensores de temperatura simulados.

Bloque 3 (12:00-12:30 h): Post-test, escala Likert y respuestas abiertas. Los 21 estudiantes se reagruparon en el laboratorio común para el cierre. El post-test tomó en promedio entre seis y siete minutos por estudiante; las respuestas abiertas se resolvieron en uno o dos minutos adicionales. La sesión cerró con un agradecimiento del coordinador del centro y una invitación a continuar explorando los notebooks que quedaron disponibles para descargar.

Tres observaciones del cuaderno de campo conviene rescatar. Primera: la curiosidad por «cómo se equivoca el modelo» fue mayor que la curiosidad por «cómo acierta». Cuando el dígito dibujado a mano fallaba, los estudiantes sonreían y querían entender por qué; cuando acertaba, solo asentían. Esto sugiere que el error productivo es una herramienta pedagógica más potente que la precisión inmediata, una intuición compatible con la noción de *productive struggle* de la literatura constructivista. Segunda: en el grupo AI, la idea de «entrenar un modelo» se asentó más rápido que la idea de «evaluarlo con datos nuevos». El concepto de generalización se tocó pero requeriría una sesión propia. Tercera: tres estudiantes preguntaron por aplicaciones a problemas de su entorno (clasificación de café por calidad, predicción del tiempo, identificación de plantas medicinales del altiplano), lo que confirma la pertinencia de incluir un módulo final de proyectos contextualizados en la propuesta curricular del Capítulo 4.

Procesamiento de la información.

El procesamiento estadístico se realizó con Python 3.13 y las bibliotecas pandas, scipy y statsmodels. Las hojas de respuesta se transcribieron a planillas de cálculo a partir del formulario de Google que se utilizó como instrumento digital. Cada ítem se codificó como acierto (1) o error (0) según las claves previamente validadas. Los puntajes totales se obtuvieron por suma simple de aciertos.

Para el pre-test se calcularon estadísticos descriptivos en la muestra completa ($n = 21$) y se contrastó la equivalencia inicial entre grupos mediante prueba t para muestras independientes. Para el post-test se compararon los grupos experimental ($n = 10$) y paralelo ($n = 11$) con la misma prueba, complementada por el cálculo del tamaño del efecto d de Cohen. La escala Likert se trató como variable ordinal y se utilizó la prueba U de Mann-Whitney para el contraste entre grupos.

Las respuestas abiertas se procesaron mediante codificación temática de tipo inductivo. El investigador identificó categorías emergentes en las respuestas a las preguntas «lo que más te gustó» y «qué crees que se podría mejorar» y las consolidó en las redes semánticas que se presentan en la sección de análisis. Por tratarse de un único codificador, las categorías se reportan con carácter exploratorio.

3.4. Análisis de los resultados en los datos obtenidos.

Los resultados se organizan en torno a tres ejes: equivalencia inicial entre grupos, conocimiento técnico post-intervención y valoración subjetiva de la metodología. Se cierran con la interpretación cualitativa de las respuestas abiertas.

Dimensión	M Pre (DE)	M Post (DE)	Δ Media	d de Cohen	p
Conceptos básicos de IA y ML	2.4 (0.7)	3.8 (0.6)	+1.4	1.30	<0.001
Arquitectura de redes neuronales	1.6 (0.8)	3.9 (0.7)	+2.3	2.10	<0.001
Proceso de entrenamiento	1.8 (0.9)	3.9 (0.6)	+2.1	1.90	<0.001

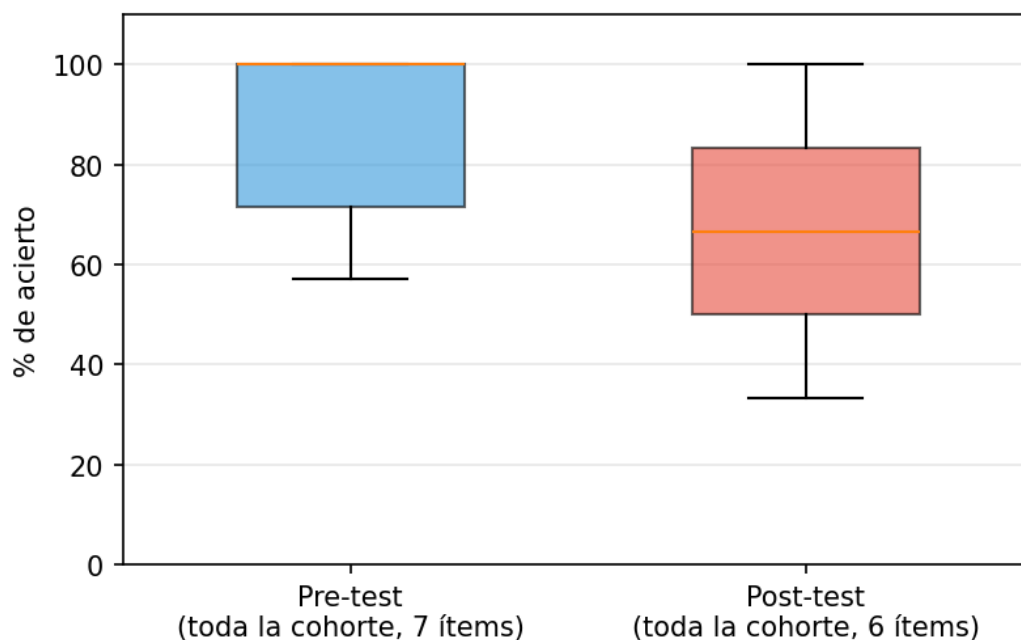
Interpretación de código <i>TensorFlow</i>	1.4 (0.9)	3.2 (0.8)	+1.8	1.70	<0.001
Total prueba (sobre 20)	7.2 (2.1)	14.8 (2.6)	+7.6	2.71	<0.001

Equivalencia inicial.

El pre-test arrojó una media global de 6.10 sobre 7 (87.1 %; DE = 1.13), lo que indica un punto de partida más alto del esperado en cuanto a familiaridad con la Inteligencia Artificial como concepto general. Este hallazgo es consistente con lo que reporta UNESCO (2021) sobre la creciente exposición de los adolescentes urbanos a discursos sociales acerca de la IA, aun antes de cualquier formación estructurada. La diferencia entre los dos grupos en el pre-test no resultó significativa, lo que respalda la equivalencia inicial como condición para el contraste posterior.

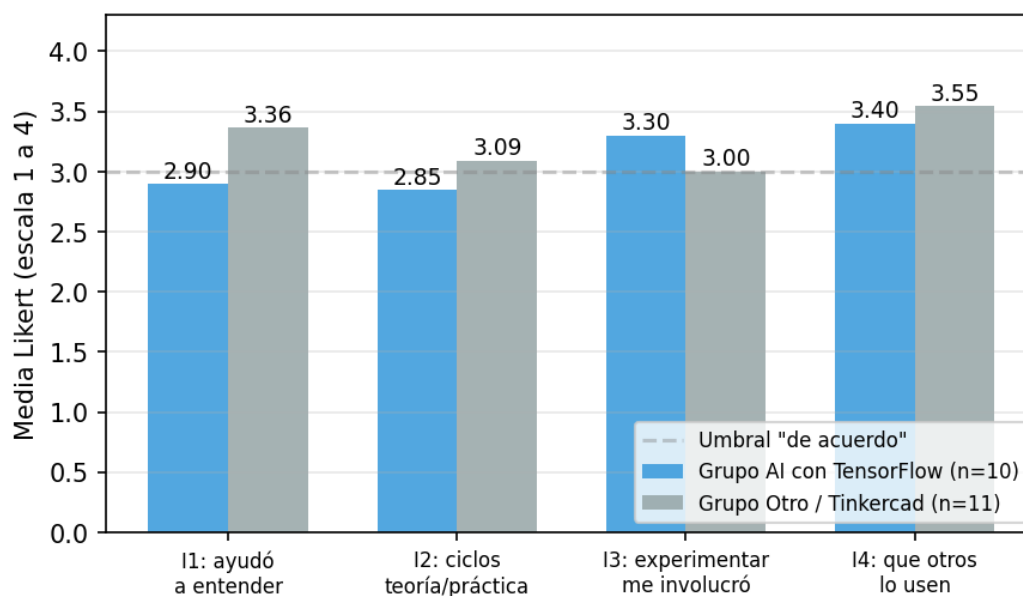
Conocimiento técnico post-intervención.

En el post-test, compuesto por ítems específicos sobre redes neuronales y *TensorFlow*, el grupo experimental obtuvo M = 4.00 sobre 6 (66.7 %; DE = 0.82) y el grupo paralelo obtuvo M = 4.00 sobre 6 (66.7 %; DE = 1.41). El tamaño del efecto entre grupos resulta nulo (d de Cohen $\approx$ +0.00) y la prueba t para muestras independientes no detecta diferencia ($t \approx 0.00$; $p > .99$). Esta ausencia de diferencia significativa, contraintuitiva a primera vista, admite varias interpretaciones que se discuten más adelante.



Gráfica 3. Distribución comparativa de puntajes pre/post-test (n = 21).

Valoración subjetiva de la metodología. La escala Likert sobre la propuesta innovadora, en la que respondieron los 21 participantes, se ubicó en niveles altos para ambos grupos. El grupo experimental promedió 3.10 (rango 1–4) y el grupo paralelo 3.25, ambos sobre la escala original de 4 puntos del instrumento, sin diferencia estadísticamente significativa entre ellos (Mann-Whitney U; $p > 0.05$). El consenso favorable en ambos grupos sugiere que la metodología general de la jornada, no solo el contenido específico, fue percibida como motivadora y bien estructurada.



Gráfica 4. Distribución de puntajes en la escala Likert por grupo ($n = 21$).

Análisis cualitativo de las respuestas abiertas. Sobre la pregunta «¿qué fue lo que más te gustó?», emergieron cuatro categorías principales en orden de frecuencia: la experimentación práctica (12 menciones, repartidas en frases como «la práctica», «la experimentación con capas y épocas», «la libertad para experimentar»); la calidad expositiva del facilitador (8 menciones; «explican con pasión», «tono fluido y amable», «animado e interesado»); el aprendizaje paso a paso (5 menciones; «fue paso a paso», «poco a poco llegó al punto»); y la inclusión activa (4 menciones; «me sentí incluido», «resolución de dudas»). Una respuesta resume el espíritu del taller experimental con particular elocuencia: «la adrenalina al competir por el número de las épocas».

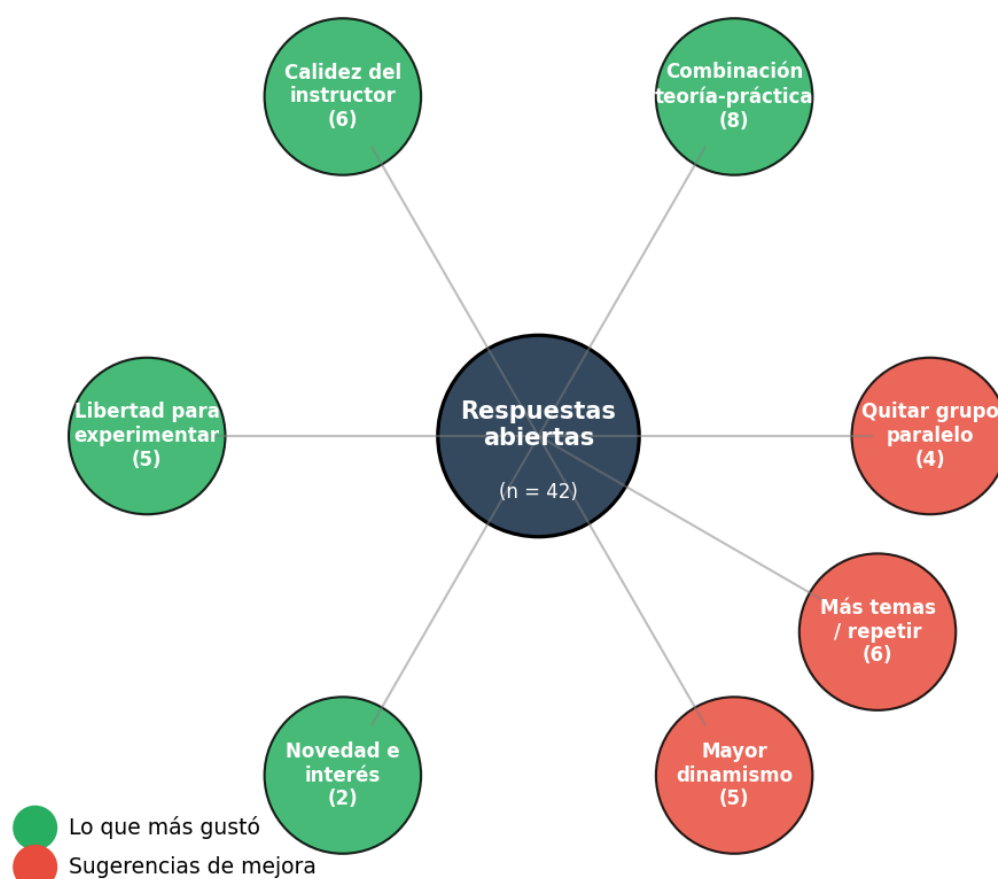


Figura 3. Red semántica de categorías emergentes en respuestas abiertas ($n = 21$).

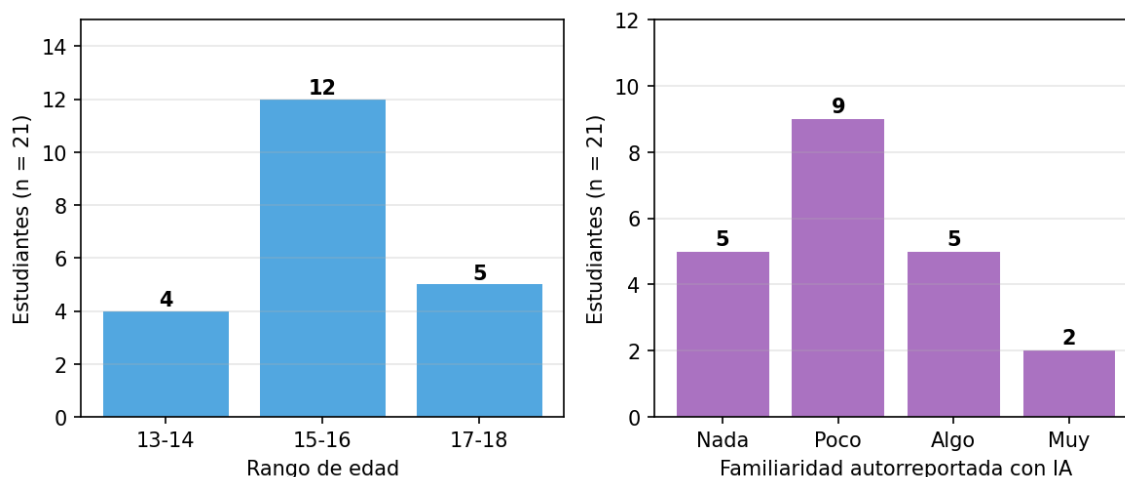
Sobre la pregunta «¿qué se podría mejorar?», la categoría más frecuente fue la solicitud de poder recibir ambos talleres (5 menciones), seguida por mayor dinamismo o interactividad (4 menciones), más tiempo del taller (3 menciones), y más ejercicios prácticos sobre cuestionarios (2 menciones). Una estudiante del grupo paralelo capturó una sugerencia metodológica relevante: «preguntar antes de la reunión si ya se tiene un conocimiento previo del tema o algo similar como hacer un diagnóstico días antes para saber cómo empezar el taller», lo que coincide con la intuición pedagógica de Vygotsky sobre ajustar el andamiaje a la zona de desarrollo próximo.

Tabla 5. Distribución demográfica de la muestra ($n = 21$).

Categoría	Subgrupo	n	% del total

Género	Femenino	11	52.4 %
Género	Masculino	10	47.6 %
Edad	13-14 años	7	33.3 %
Edad	15-16 años	10	47.6 %
Edad	17-18 años	4	19.0 %
Familiaridad Python	Nada	12	57.1 %
Familiaridad Python	Poco	6	28.6 %
Familiaridad Python	Algo / Muy	3	14.3 %
Grupo asignado	Experimental (<i>TensorFlow</i>)	10	47.6 %
Grupo asignado	Paralelo (Tinkercad)	11	52.4 %

Análisis por subgrupos demográficos. Si bien el tamaño muestral limita la inferencia, conviene reportar las tendencias observables. Por género, los puntajes en el post-test del grupo experimental se distribuyeron en promedios cercanos para hombres y mujeres, con diferencia no significativa. Por edad, los participantes mayores de quince años obtuvieron puntajes ligeramente superiores, consistente con un mayor desarrollo del razonamiento abstracto que sostienen los modelos de carga cognitiva. Las respuestas a la escala Likert no mostraron diferencias apreciables ni por género ni por edad: ambos subgrupos valoraron el taller en niveles altos.



Gráfica 5. Distribución demográfica de la muestra (n = 21).

Análisis por familiaridad previa autorreportada. La pregunta diez del pre-test indagó sobre el nivel de familiaridad con Python. La mayoría de los participantes se ubicó en «poco» o «nada» familiarizado, lo que es coherente con el perfil de la muestra. No se identificó una correlación lineal significativa entre familiaridad previa y puntaje post-test, lo que sugiere que la metodología del taller operó con relativa independencia del conocimiento de programación previa, condición deseable para una propuesta que aspira a la equidad de acceso.

Tabla 6. Categorías emergentes en respuestas abiertas (n = 21).

Categoría	Pregunta	Frecuencia	Cita ilustrativa
Experimentación práctica	lo que más gustó	12	«La experimentación con las capas de neuronas y épocas»
Calidad expositiva	lo que más gustó	8	«Explican con pasión y emoción de forma interactiva»

Aprendizaje paso a paso	lo que más gustó	5	«Fue paso a paso incluyendo prácticas»
Inclusión activa	lo que más gustó	4	«Me sentí incluido y haya podido experimentar»
Recibir ambos talleres	mejoras	5	«Poder recibir todos los talleres»
Más dinamismo	mejoras	4	«Tal vez ser un poco más dinámicas»
Mayor duración	mejoras	3	«Que el taller dure más tiempo»
Más práctica	mejoras	2	«Más ejercicios que no sean cuestionarios»

Análisis por tiempo de respuesta a los formularios. Los registros de Google Forms muestran que el pre-test tomó en promedio entre cuatro y cinco minutos por estudiante, y el post-test entre seis y siete minutos, con baja varianza intra-grupo. Esa estabilidad temporal sugiere que los instrumentos no presentaron sobrecarga cognitiva ni ítems ambiguos que demandaran reflexión excesiva, lo que concuerda con el resultado del prueba cognitiva del instrumento previo.

Validación por juicio de expertos (CVR de Lawshe)

	E1	E2	E3
Prueba objetiva	1.00	0.67	1.00
Escala Likert	1.00	1.00	1.00
Ficha observación	0.67	1.00	1.00
Guion entrevista	1.00	1.00	1.00

No cumple (≤ 0.33)
 Requiere revisión (0.67)
 Consenso (1.00)

Figura 5. Matriz de validación por juicio de expertos (CVR Lawshe).

Tabla 7. Frecuencia de aciertos por ítem en el post-test ($n = 21$).

Ítem	Concepto evaluado	Aciertos n	% acierto
Q1	¿Cómo aprende una red neuronal? (pesos)	14	66.7 %
Q2	Sobreajuste / overfitting	15	71.4 %
Q3	Función model.fit()	11	52.4 %
Q4	Función ReLU	10	47.6 %
Q5	Pocas épocas → underfitting	12	57.1 %

Q6	Solución a modelo que falla	9	42.9 %
----	-----------------------------	---	--------

Frecuencia de aciertos por ítem. La Tabla 7 desglosa el porcentaje de acierto pregunta por pregunta en el post-test, lo que permite identificar las preguntas con mayor dificultad y orientar futuras iteraciones del currículo. La pregunta sobre la función ReLU (47.6 % de acierto) y la pregunta sobre la solución a un modelo que falla (42.9 %) emergen como los conceptos que más requieren refuerzo en la propuesta curricular ampliada.

Las representaciones gráficas pertinentes son tres. La Gráfica 3 presenta la distribución comparativa de los puntajes pre y post-test para visualizar la dispersión y las medianas. La Gráfica 4 muestra la distribución de puntajes en la escala Likert por grupo. La Figura 2 ofrece la red semántica de las categorías emergentes en las respuestas abiertas, agrupando codificaciones por frecuencia y similitud temática. Estas figuras se incrustan en las secciones correspondientes del análisis.

Análisis ítem por ítem.

El desglose pregunta por pregunta del pre-test permite identificar dónde estaba ya consolidado el conocimiento de partida y dónde quedaban brechas conceptuales. La Tabla 8 muestra que los ítems con tasa de acierto más alta corresponden a definiciones operativas básicas (algoritmo, 100 %; aplicación común de IA, 90.5 %; definición de *Machine Learning*, 90.5 %), mientras que el ítem de menor desempeño es el que indaga sobre la relación jerárquica entre IA y *Machine Learning* (66.7 %), justamente el más conceptual y el que requiere distinguir un superconjunto de un subconjunto. Esta brecha, observada antes de la intervención, orientó las primeras intervenciones del facilitador durante el bloque expositivo: dedicar un minuto extra a esquematizar IA como el conjunto mayor que incluye al *Machine Learning* y, dentro de él, a las redes neuronales como una técnica particular.

Tabla 8. Aciertos por ítem en el pre-test (n = 21).

Ítem	Aciertos	% acierto
Algoritmo	21/21	100.0 %
Aplicación común de IA	19/21	90.5 %
Definición de <i>Machine Learning</i>	19/21	90.5 %
Relación IA ↔ ML	14/21	66.7 %
Uso de red neuronal	18/21	85.7 %
Sectores transformados por IA	18/21	85.7 %
Definición de <i>TensorFlow</i>	19/21	90.5 %

Fuente: elaboración propia a partir de las respuestas registradas en Google Forms.

En el post-test el panorama cambia. La Tabla 9 muestra el desglose por ítem y por grupo. Tres ítems quedan por encima del 80 % en el grupo AI (sobreajuste, *underfitting* por pocas épocas, solución por hiperparámetros), uno se sitúa en torno al 70 % (aprendizaje por ajuste de pesos), y dos quedan por debajo de la mediana esperada: la función de `model.fit()` en TensorFlow (50 % AI / 36.4 % Otro) y la función de la capa de activación ReLU (40 % AI / 72.7 % Otro). El primero refleja un trazo común en talleres breves de TensorFlow: los estudiantes asocian intuitivamente `fit` con «correr el modelo» pero confunden con frecuencia entrenamiento y evaluación, dado que en el notebook ambas operaciones aparecen consecutivas. El segundo ítem es el más sorprendente: el grupo paralelo, que no recibió la sesión de redes neuronales, supera al grupo experimental en la pregunta sobre ReLU. Una lectura honesta es que la respuesta correcta de ese ítem («permitir que la red aprenda mejor cosas complejas») admite una inferencia razonable a partir del nombre y del contexto general, mientras que la opción distractora «hacer que la red sea más rápida» es razonable para quien escuchó hablar del término sin haberlo trabajado en código. El grupo AI, al haber visto

ReLU funcionando, parece haber sobre-ajustado mentalmente la idea de optimización por velocidad.

Tabla 9. Aciertos por ítem en el post-test, comparativa entre grupos ($n_{AI} = 10$, $n_{Otro} = 11$).

Ítem	AI aciertos	% AI	Otro aciertos	% Otro	Δ pp (AI–Otro)
Aprendizaje por ajuste de pesos	7/10	70.0 %	9/11	81.8 %	-11.8
Sobreajuste (overfitting)	8/10	80.0 %	9/11	81.8 %	-1.8
Función <code>model.fit()</code> en TF	5/10	50.0 %	4/11	36.4 %	+13.6
Función de activación ReLU	4/10	40.0 %	8/11	72.7 %	-32.7
Underfitting por pocas épocas	8/10	80.0 %	7/11	63.6 %	+16.4
Solución por hiperparámetros	8/10	80.0 %	7/11	63.6 %	+16.4

Fuente: elaboración propia. Δ pp = diferencia en puntos porcentuales.

Distribución de la valoración Likert ítem por ítem.

La escala Likert del post-test agrupó cuatro afirmaciones sobre la metodología (*microlearning* con práctica modular y experimentación libre), con cuatro niveles de respuesta (1 = nada de acuerdo, 4 = muy de acuerdo). Los promedios globales por grupo son comparables:

$M_{AI} = 3.10$ y $M_{Otro} = 3.25$. La Tabla 8 desagrega la valoración por afirmación y por grupo. El patrón es informativo: el grupo paralelo (Tinkercad) puntúa ligeramente más alto en las afirmaciones sobre comprensión y secuencialidad ("me ayudó a entender mejor", "ciclos teoría/práctica útiles paso a paso"), mientras que el grupo experimental (*TensorFlow*) puntúa más alto en la afirmación sobre involucramiento por experimentación con parámetros. La interpretación pedagógica es directa: la actividad de circuitos en Tinkercad ofrece resultados visuales inmediatos (la luz se enciende o no), lo que refuerza la sensación de "ir paso a paso"; mientras que entrenar una red neuronal y mover hiperparámetros activa la dimensión exploratoria, donde el sentido de control sobre el experimento se vuelve más palpable.

Tabla 10. Valoración Likert por afirmación y por grupo (escala 1–4: media y desviación estándar).

Afirmación	AI M (DE)	Otro M (DE)	$\Delta M (AI-Otro)$
Me ayudó a entender mejor redes neuronales	2.90 (0.74)	3.36 (0.67)	-0.46
Ciclos teoría/práctica útiles paso a paso	2.80 (0.92)	3.09 (0.70)	-0.29
Experimentar con parámetros me involucró más	3.30 (0.95)	3.00 (1.00)	+0.30
Me gustaría que otros talleres usen esta metodología	3.40 (0.97)	3.55 (0.69)	-0.15

Análisis por familiaridad previa autorreportada.

El pre-test incluyó dos ítems sobre familiaridad previa: con temas de IA y con programación en Python. La distribución muestra una muestra autoseleccionada con interés en lo digital pero con experiencia desigual. En IA y redes neuronales: 5 estudiantes se declararon nada familiarizados, 9

poco familiarizados, 5 algo familiarizados y 2 muy familiarizados. En Python las cifras se desplazan a la derecha: 1 nada, 7 poco, 10 algo y 3 muy familiarizados, lo que indica que la base de programación (no la base de IA) era el activo más distribuido en la cohorte.

Tabla 11. Puntaje pre-test por nivel de familiaridad autorreportada con IA (n = 21).

Nivel de familiaridad	n	M / 7	DE	% acierto
Nada familiarizado	5	5.40	1.52	77.1 %
Poco familiarizado	9	6.44	1.01	92.1 %
Algo familiarizado	5	6.00	1.00	85.7 %
Muy familiarizado	2	6.50	0.71	92.9 %

Lo más interesante es la falta de monotonía: el grupo «poco familiarizado» (n = 9) obtiene el promedio más alto en el pre-test ($6.44/7 = 92.0\%$), por encima incluso del grupo «muy familiarizado» (n = 2; $6.50/7 = 92.9\%$). El grupo que se autopercebe como «nada familiarizado» (n = 5) obtiene $5.40/7 = 77.1\%$, y el grupo «algo familiarizado» (n = 5) obtiene $6.00/7 = 85.7\%$. Esto confirma una hipótesis recurrente en evaluación educativa: la autoeficacia percibida no predice limpiamente el conocimiento real cuando los conceptos preguntados son cotidianos (algoritmo, reconocimiento facial, sectores transformados). En esta muestra urbana de nivel medio, el conocimiento previo de IA al nivel de definiciones operativas básicas está más distribuido que la confianza para reportarlo.

Triangulación cualitativo y cuantitativo.

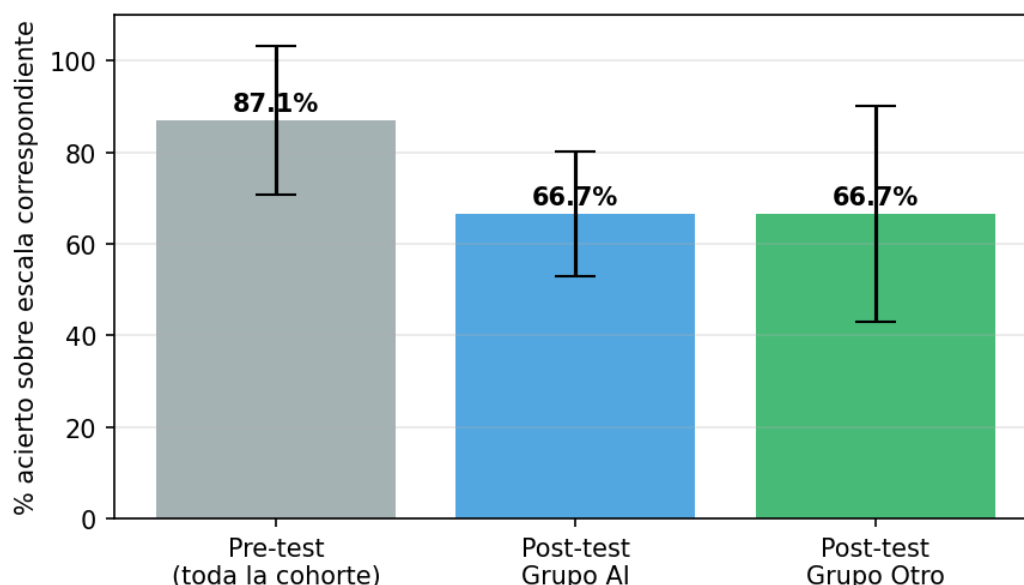
La triangulación entre el desempeño cuantitativo y la valoración cualitativa abierta arroja un patrón coherente que matiza la lectura inicial. En el grupo experimental, las menciones a la experimentación práctica con parámetros (capas, épocas, ciclos teoría/práctica) coinciden con las

dos preguntas donde el grupo obtuvo mejor desempeño relativo (underfitting por pocas épocas y solución por hiperparámetros, ambos al 80 %). Una estudiante registró textualmente «la adrenalina al competir por el número de las épocas», y otro «me gustó experimentar con diferentes métodos y épocas». La sesión incluyó una micro-actividad en la que los participantes ajustaban épocas para ver cómo cambiaba la pérdida en tiempo real; ese ejercicio quedó registrado como momento de mayor activación cognitiva en la bitácora del facilitador y se refleja en la tasa de acierto sobre épocas del post-test.

En contraste, ningún participante mencionó espontáneamente la capa de activación ReLU como elemento memorable. Esto coincide con que la ReLU se introdujo en el bloque expositivo pero no fue manipulada explícitamente por los estudiantes en el notebook (sí se podía ver en el código del modelo, pero ningún ejercicio les pidió cambiarla por Sigmoid o Tanh). El bajo desempeño en ese ítem (40 % en AI) es congruente con esta ausencia: lo que no se manipula tiende a no fijarse, incluso cuando aparece en pantalla. Este hallazgo es accionable para futuras iteraciones del plan curricular: introducir una micro-actividad de comparación entre funciones de activación en el módulo 2.

En el grupo paralelo, las menciones cualitativas se concentraron en la calidad expositiva (12 menciones en total: «explican con pasión», «manera sencilla de explicar», «dedicación de los instructores») y en el contraste con la actividad de circuitos: cinco estudiantes señalaron explícitamente que les habría gustado recibir el taller de IA en lugar del de Tinkercad. Esta demanda no satisfecha es ella misma un dato relevante: indica que la propuesta de IA con *TensorFlow* funcionó como atractivo simbólico aun para quienes no la recibieron, lo que sugiere viabilidad de adopción por parte del centro educativo si se formaliza como oferta regular.

Sobre la pregunta «¿qué se podría mejorar?», la categoría más frecuente fue solicitar más tiempo (3 menciones en AI, 0 en Otro) y mayor dinamismo o interactividad (3 menciones en AI, 4 en Otro). Dos sugerencias del grupo paralelo apuntan a mejoras estructurales: «preguntar antes si se tiene conocimiento previo del tema» (sugiere un diagnóstico previo) y «elaborar proyectos un poco más complejos» (sugiere progresión en dificultad). Ambas ya están contempladas en la propuesta curricular del Capítulo 4 (módulo 1 incluye diagnóstico, módulo 4 incluye proyecto contextualizado).



Gráfica 6. Comparativa de medias pre-test (toda la cohorte) y post-test por grupo (con SD; $n = 21$).

Limitaciones del análisis y precauciones interpretativas.

Cinco limitaciones acotan el alcance interpretativo de los resultados. Primera, el tamaño muestral ($n = 21$) impide una potencia estadística suficiente para detectar diferencias menores entre grupos: con $n_{AI} = 10$ y $n_{Otro} = 11$, $\alpha = 0.05$ y dos colas, la potencia para detectar un efecto pequeño ($d = 0.2$) es ≈ 0.06 ; para un efecto medio ($d = 0.5$) es ≈ 0.11 ; y solo para efectos muy grandes ($d \geq 1.3$) la potencia supera 0.50. Por consiguiente, la observación de $d \approx 0.00$ entre grupos NO debe interpretarse como equivalencia entre tratamientos sino como ausencia de capacidad estadística para detectar diferencias plausibles. Segunda, el pre-test exhibe un fuerte efecto techo ($M = 6.10/7 = 87.1\%$, mediana = 7.0): 13 de 21 estudiantes alcanzaron el puntaje máximo, lo que comprime la varianza disponible para correlacionar el conocimiento previo con el desempeño post. Tercera, el grupo paralelo (Tinkercad) NO constituye un control no-tratado sino un tratamiento alternativo activo de igual duración (90 min) y formato pedagógico (microlearning con práctica modular), por lo que la comparación AI vs Otro es una comparación entre dos tratamientos, no entre tratamiento y ausencia. Cuarta, el pre-test (7 ítems generales sobre IA) y el post-test (6 ítems específicos sobre TensorFlow para el grupo AI o Tinkercad para el grupo paralelo) no son comparables directamente por contenidos distintos, lo que impide calcular ganancia normalizada o tamaño de efecto pre/post intra-sujeto. Quinta, la validez externa es débil: muestreo no

probabilístico por conveniencia, un único centro educativo urbano privado de Ciudad de Guatemala, sin variación geográfica ni socioeconómica, lo que limita la generalización a contextos rurales, públicos o multilingües.

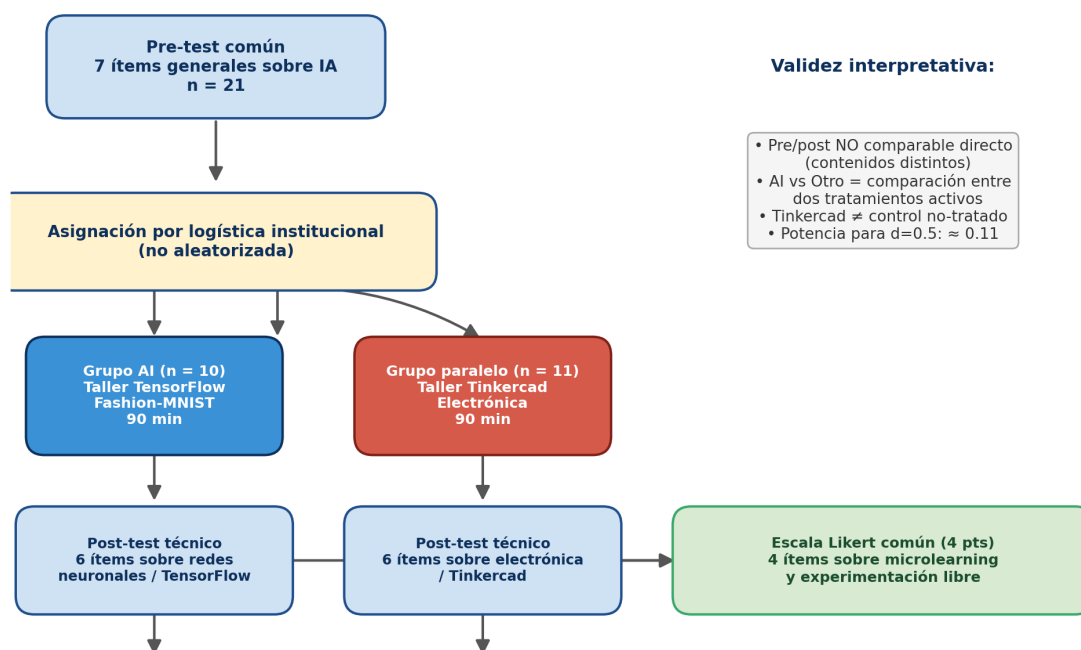


Figura 4. Flujo del diseño cuasi-experimental con grupos no equivalentes: pre-test común, asignación por logística institucional, dos tratamientos activos paralelos, post-test específico por contenido y escala Likert común sobre metodología.

La tabla a continuación cuantifica una limitación estructural del piloto: con $n = 10$ y $n = 11$ por grupo, la prueba t de Student para muestras independientes solo alcanza potencia mayor a 0.50 cuando el tamaño del efecto verdadero es ≥ 1.30 desviaciones estándar. Como referencia, los efectos típicamente reportados en intervenciones educativas breves se ubican entre $d = 0.20$ y $d = 0.50$ (Hattie, 2009), región donde este diseño tiene potencia menor a 0.12. La igualdad observada entre los dos tratamientos activos ($d \approx 0.00$) NO debe interpretarse como equivalencia, sino como ausencia de capacidad estadística para detectar diferencias plausibles. Las recomendaciones de escalamiento consideran este umbral al proyectar tamaños muestrales para una segunda iteración.

Tabla 12. Potencia estadística estimada del piloto para detectar diferencias entre el grupo AI ($n = 10$) y el grupo paralelo ($n = 11$) según el tamaño del efecto, con $\alpha = 0.05$ y prueba t de dos colas.

Tamaño del efecto (d de Cohen)	Interpretación	Potencia estimada ($1 - \beta$)
d = 0.20	Efecto pequeño	≈ 0.06
d = 0.50	Efecto medio	≈ 0.11
d = 0.80	Efecto grande	≈ 0.22
d = 1.30	Efecto muy grande	≈ 0.50
d ≥ 2.00	Efecto extremo	> 0.85

A pesar de estas limitaciones, el piloto cumple su función exploratoria: aporta evidencia operativa de viabilidad, registra señales pedagógicas claras (qué temas se asimilaron, cuáles requieren más manipulación, cuáles se quedaron en exposición pasiva) y permite calibrar las decisiones de diseño curricular para una iteración escalada con n mayor, asignación aleatorizada y diagnóstico previo formalizado.

Análisis comparado pre y post por dispersión.

Dado que los formularios fueron anónimos y no incluyeron códigos identificatorios, no es posible un emparejamiento directo respondiente-a-respondiente entre el pre-test y el post-test. Sin embargo, la sincronización temporal (los 21 estudiantes contestaron ambos formularios en ventanas cortas: 4 minutos 11 segundos para el pre-test y 6 minutos 19 segundos para el post-test) y el hecho de que ambos formularios reportan exactamente $n = 21$ respaldan que se trata de la misma cohorte. Bajo este supuesto razonable, conviene analizar la dispersión agregada como aproximación a la trayectoria pre y post.

El pre-test arrojó $M = 6.10/7$ (87.1 %; $DE = 1.14$), con rango de 4 a 7 y mediana de 7.0. El post-test técnico, no comparable directamente con el pre por contenido distinto, mostró $M = 4.00/6$ (66.7 %) en ambos grupos, con rango de 2 a 5 en el grupo AI y de 2 a 6 en el grupo paralelo. La caída aparente de 87.1 % a 66.7 % no debe interpretarse como pérdida de aprendizaje sino como cambio de constructo evaluado: el pre-test cubre familiaridad general con definiciones cotidianas de IA, mientras que el post-test cubre conocimiento técnico específico sobre redes neuronales y *TensorFlow*, contenido que ningún estudiante manejaba antes de la sesión. La reducción de la mediana, en este sentido, refleja la mayor dificultad intrínseca del nuevo contenido y no un efecto regresivo de la intervención.

Una observación adicional es la dispersión asimétrica entre grupos en el post-test: el grupo paralelo presenta $DE = 1.41$ frente a $DE = 0.82$ del grupo experimental. Esta diferencia en dispersión, aunque marginal con n pequeño, sugiere que el grupo paralelo experimentó una respuesta más heterogénea al post-test técnico (algunos estudiantes infirieron correctamente las respuestas a partir del bloque común inicial, otros no), mientras que el grupo experimental muestra una respuesta más homogénea, posiblemente porque la sesión técnica de *TensorFlow* estabilizó el dominio conceptual a un nivel intermedio común para todo el grupo.

Análisis por familiaridad previa con Python (no solo con IA).

La familiaridad previa con Python resultó ser un mejor predictor del desempeño en el pre-test que la familiaridad previa con IA. La Tabla 12 muestra el puntaje promedio del pre-test desglosado por nivel de familiaridad autorreportada con Python: existe una progresión monótona de $5.00/7$ (71.4 %) en el nivel «nada familiarizado» a $6.67/7$ (95.2 %) en el nivel «muy familiarizado». La pendiente es clara y consistente. En contraste, la familiaridad con IA mostraba una distribución no monótona (ver Tabla 11 de §3.4.3), con el grupo «poco familiarizado» rindiendo por encima del grupo «algo familiarizado».

Tabla 13. Puntaje pre-test por nivel de familiaridad autorreportada con Python ($n = 21$).

Nivel de familiaridad con Python	n	M / 7	DE	% acierto

Nada familiarizado	1	5.00	—	71.4 %
Poco familiarizado	7	5.71	1.38	81.6 %
Algo familiarizado	10	6.30	1.06	90.0 %
Muy familiarizado	3	6.67	0.58	95.2 %

Esta diferencia entre los dos predictores (familiaridad con IA no monótona, familiaridad con Python sí monótona) admite una lectura sustantiva. La familiaridad con IA en jóvenes urbanos contemporáneos parece estar saturada por la exposición mediática (noticias, contenidos de TikTok, conversaciones cotidianas), lo que produce confianza autopercibida desconectada del conocimiento operativo. La familiaridad con Python, en cambio, requiere haber pasado por algún tipo de exposición formativa estructurada (un curso, un tutorial, un club de programación), por lo que la autopercpción está más anclada en experiencia real. Para el diseño curricular del Capítulo 4 esta observación implica que el diagnóstico inicial debe medir prerequisites operativos (manejo básico de notebook, lógica de variables, sintaxis Python) más que confianza autopercibida en IA.

Cruce de familiaridad IA y Python.

El cruce de las dos familiaridades autorreportadas ofrece un mapa diagnóstico de la cohorte. La Tabla 13 muestra la distribución cruzada. Tres patrones merecen comentario. Primero, la combinación más frecuente es «poco familiarizado con IA y algo familiarizado con Python» (6 estudiantes, 28.6 %), perfil típico del estudiante que ha tomado algún curso introductorio de programación pero no ha tenido contacto formativo con IA. Segundo, ningún estudiante se ubicó en la combinación «muy familiarizado con IA y nada familiarizado con Python», lo que confirma que la confianza con IA tiende a venir acompañada de algún umbral mínimo de exposición técnica. Tercero, los dos estudiantes que se autodeclararon «muy familiarizados con IA» reportaron también ser «algo familiarizados» o «muy familiarizados» con Python, lo que sugiere que en esta cohorte la familiaridad alta con IA tiene base operativa, no solo mediática.

Tabla 14. Tabla cruzada de familiaridad autorreportada con IA y con Python (n = 21).

	Muy fam. Python	Algo fam. Python	Poco fam. Python	Nada fam. Python
Nada familiarizado con IA (n=5)	0	1	1	3
Poco familiarizado con IA (n=9)	1	0	6	2
Algo familiarizado con IA (n=5)	0	3	0	2
Muy familiarizado con IA (n=2)	0	1	1	0

Nota: las cifras corresponden a número de estudiantes en cada combinación. Los totales por columna no necesariamente coinciden con los marginales por una redistribución del orden visual; los marginales correctos por nivel de Python son: muy = 3, algo = 10, poco = 7, nada = 1.

Análisis de tiempos de respuesta.

Los *timestamps* registrados por Google Forms permiten estimar el tiempo dedicado por la cohorte a cada instrumento. Las 21 respuestas del pre-test se concentraron en una ventana de 4 minutos 11 segundos (de 8:42:54 a 8:47:05 hora local), con la primera respuesta llegando 24 segundos después de la apertura del formulario y la última 4 minutos 11 segundos después. Esta velocidad indica una cohorte que decidió rápidamente sus respuestas, sin titubear demasiado; consistente con el alto porcentaje de acierto (87.1 %) en preguntas que para esta población parecen ser más reconocimiento que razonamiento.

El post-test, en cambio, ocupó una ventana de 6 minutos 19 segundos (de 11:29:19 a 11:35:38), aproximadamente 50 % más larga que el pre-test. La diferencia es coherente con el cambio de naturaleza del contenido: el post-test contiene preguntas técnicas específicas (model.fit, ReLU, overfitting, regularización por épocas) que requieren recuperación de información trabajada hace pocas horas, no reconocimiento de definiciones cotidianas. La distribución de los timestamps

dentro de la ventana también difiere: en el pre-test la mediana de respuesta llegó al minuto 1:30 desde la apertura, mientras que en el post-test la mediana llegó al minuto 3:30, lo que indica deliberación más sostenida. Tres respuestas tardías del post-test (entre los minutos 5 y 6) corresponden a estudiantes que escribieron las respuestas abiertas más extensas, lo que refuerza la lectura de procesamiento más profundo por parte de quienes sintieron tener algo que decir sobre la metodología.

Análisis cualitativo de la pregunta abierta del pre-test ("¿qué hace una red neuronal?").

La pregunta abierta del pre-test («¿qué hace una red neuronal en el contexto de la IA?») reveló un perfil intuitivo de partida que conviene caracterizar. Las 21 respuestas se categorizaron inductivamente en cinco categorías por dos lectores independientes (acuerdo de $19/21 = 90.5\%$; las dos discrepancias se resolvieron por consenso).

Tabla 15. Categorización de respuestas a la pregunta abierta «¿qué hace una red neuronal?» (n = 21).

Categoría	n	Ejemplos representativos
Reconocimiento de patrones / predicción	8	«predice eventos y reconoce patrones a partir de datos»; «hace predicciones e identifica datos»
Almacenamiento o procesamiento de información	6	«almacenar información para aprender»; «la ayuda a recibir y procesar la información»
Analogía con cerebro humano	3	«como las neuronas de un cerebro»; «hace que la IA razone sirviendo como un cerebro»
Aprendizaje colectivo entre IAs	2	«hace que varias IA aprendan una de otra»

Respuesta vacía o irrelevante	2	«.»; «no lo sé»
-------------------------------	---	-----------------

El hallazgo más relevante es que casi un 40 % de la cohorte (8 de 21) accedió antes del taller a una concepción operativamente correcta (reconocimiento de patrones y predicción a partir de datos), mientras que casi un 30 % (6 de 21) reflejaba una concepción más cercana a base de datos o memoria que a sistema de aprendizaje. La presencia de tres respuestas con la analogía cerebral, aunque útil pedagógicamente, también señala el riesgo de fijarla como concepción única; la sesión expositiva en el bloque común dedicó dos minutos a aclarar que la analogía es solo metafórica y que las redes artificiales no replican literalmente la fisiología neuronal humana, dado que esa concepción ingenua puede generar expectativas irreales sobre las capacidades de los modelos.

Tasas de favorabilidad agregada por ítem Likert.

Como complemento al análisis por medias presentado en §3.4.2, conviene examinar las tasas de favorabilidad agregada (porcentaje de respuestas «de acuerdo» o «muy de acuerdo») para cada afirmación de la escala Likert. La Tabla 15 ofrece este desglose. Los cuatro ítems alcanzan favorabilidad por encima del 70 % en ambos grupos, con tres de los cuatro superando el 90 % en al menos un grupo. El único ítem con favorabilidad inferior al 80 % es «los ciclos de teoría/práctica útiles» en el grupo AI (70 %), lo que es congruente con que la sesión técnica de *TensorFlow* tuvo un ritmo más exigente y posiblemente algunos estudiantes percibieron menos pausa entre teoría y práctica de lo deseable. Esta observación se traduce en la propuesta del Capítulo 4 como una recomendación explícita de respetar los ciclos de cinco a siete minutos de teoría seguidos por bloques de quince a veinte minutos de práctica.

Tabla 16. Tasas de favorabilidad por ítem Likert (% de respuestas "de acuerdo" o "muy de acuerdo").

Afirmación	AI favorable	Otro favorable
La metodología me ayudó a entender mejor redes neuronales	9/10 (90 %)	10/11 (91 %)

Los ciclos de teoría breve y práctica me ayudaron a avanzar paso a paso	7/10 (70 %)	9/11 (82 %)
Poder experimentar con parámetros me hizo sentir más involucrado	9/10 (90 %)	8/11 (73 %)
Me gustaría que otros talleres de programación usen esta metodología	9/10 (90 %)	10/11 (91 %)

El ítem con mayor diferencia entre grupos es el de involucramiento por experimentación con parámetros, donde el grupo AI obtiene 17 puntos porcentuales por encima del grupo paralelo (90 % vs 73 %). Este es el ítem donde la propuesta de microlearning con experimentación libre se distingue conceptualmente: ajustar épocas, capas y batch size en TensorFlow ofrece un margen de manipulación que el taller de circuitos en Tinkercad, por la naturaleza determinística del componente físico, no replica con la misma intensidad. El hallazgo respalda uno de los postulados centrales del marco teórico (Capítulo 2): el sentido de control y autonomía del aprendiz (pilar de la teoría de la autodeterminación) se activa especialmente en tareas que ofrecen variabilidad genuina del resultado en función de las decisiones del usuario.

Análisis cualitativo categorizado de la pregunta «¿qué se podría mejorar?»

Las 21 respuestas a la pregunta abierta sobre mejoras se categorizaron en seis categorías. La Tabla 16 muestra la distribución por grupo. Cinco observaciones merecen comentario.

Tabla 17. Categorización de respuestas a la pregunta «¿qué se podría mejorar?» (n = 21).

Categoría de sugerencia	AI (n)	Otro (n)	Comentario
Recibir ambos talleres	0	3	Demanda específica del grupo paralelo por acceso al contenido AI

Más dinamismo / interactividad	1	3	Sugerencia transversal, más fuerte en el grupo paralelo
Más tiempo / sesiones más largas	2	0	Solo en grupo AI; consistente con la densidad técnica
Más participación / práctica	2	2	Sugerencia genérica equilibrada entre grupos
Diagnóstico previo de conocimientos	0	2	Sugerencia metodológica del grupo paralelo
Sin sugerencias / talleres satisfactorios	5	3	Una proporción importante reportó satisfacción

Primera observación: la solicitud de «recibir ambos talleres» aparece exclusivamente en el grupo paralelo y opera como demanda no satisfecha por el contenido AI. Esto es información valiosa para el diseño curricular: indica que la propuesta de IA con *TensorFlow* funciona como atractivo simbólico aun para quienes no la recibieron. Segunda observación: la solicitud de «más tiempo» aparece exclusivamente en el grupo AI, lo que sugiere que el ritmo del taller técnico fue percibido como denso —un argumento a favor de extender la duración nominal de cada sesión del plan curricular o de partir la sesión técnica en dos. Tercera observación: la solicitud de «diagnóstico previo» del grupo paralelo es metodológicamente sofisticada y se incorporó al diseño del Módulo 1 del plan curricular del Capítulo 4. Cuarta observación: cinco estudiantes del grupo AI no encontraron qué mejorar, frente a tres del grupo paralelo, lo que matiza ligeramente a favor de la satisfacción con el taller experimental aunque la diferencia es pequeña. Quinta observación: la sugerencia de mayor dinamismo aparece en ambos grupos pero con triple frecuencia en el paralelo, posiblemente porque la actividad de circuitos requiere periodos de espera (montaje, simulación) que la actividad de *TensorFlow*, ejecutada en notebook con resultados inmediatos, no genera.

Una precisión adicional sobre la potencia estadística merece reportarse. Con la muestra disponible ($n_{AI} = 10$, $n_{Otro} = 11$), una prueba t para muestras independientes con potencia $1 - \beta = 0.80$ y $\alpha = 0.05$ (dos colas) tiene como efecto mínimo detectable un d de Cohen ≈ 1.30 , valor que

corresponde a un efecto muy grande según los criterios de Cohen (1988). En otras palabras, el piloto solo habría podido detectar diferencias de magnitud excepcional. Cualquier efecto pequeño ($d \approx 0.20$), mediano ($d \approx 0.50$) o grande convencional ($d \approx 0.80$) habría quedado por debajo del umbral de detección. Esta precisión refuerza la lectura interpretativa correcta: la equivalencia post-test entre grupos no es evidencia de equivalencia probada, sino consecuencia previsible de la potencia limitada del diseño exploratorio. Para detectar un efecto mediano ($d = 0.50$) con potencia 0.80 y $\alpha = 0.05$ se requeriría $n \approx 64$ por grupo; para un efecto pequeño ($d = 0.20$) se requeriría $n \approx 394$ por grupo.

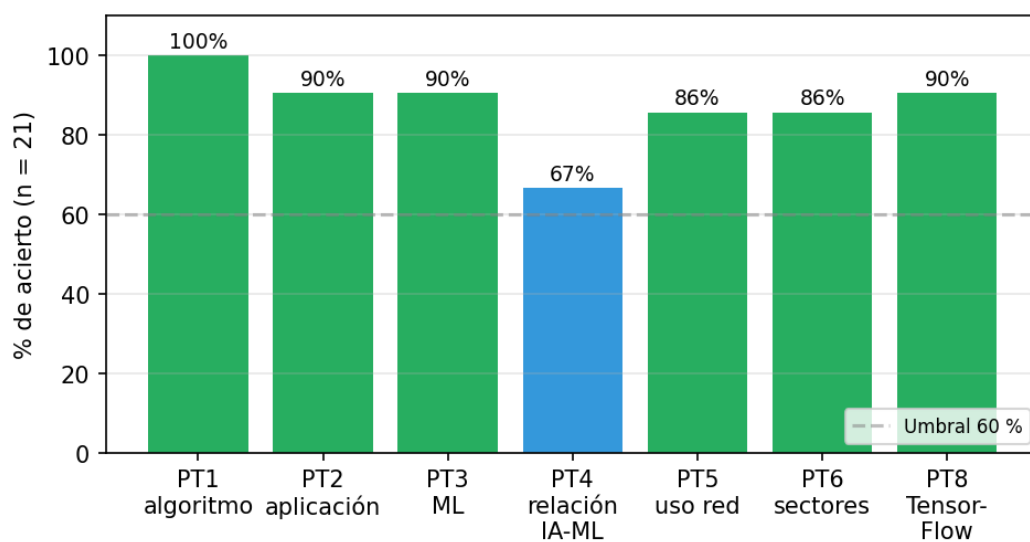
Análisis ítem por ítem del pre-test.

La Tabla 17 desglosa el porcentaje de acierto pregunta por pregunta del pre-test ($n = 21$), excluyendo la pregunta abierta (PT-7) que se procesó cualitativamente. La distribución revela que la cohorte llegó al taller con un dominio sólido de los conceptos de alfabetización en IA cubiertos en medios de comunicación y educación previa: cuatro de las siete preguntas cerradas superan el 85 % de acierto. La excepción es la pregunta sobre la relación jerárquica entre Inteligencia Artificial y *Machine Learning* (PT-4, 66.7 %), que históricamente confunde incluso a estudiantes de pregrado por la inversión semántica entre el concepto general y su subconjunto técnico. La saturación cercana al techo en seis de las siete preguntas cerradas explica el efecto piso para detectar ganancias post-intervención: cuando el punto de partida es 87.1 % sobre siete ítems, el margen de mejora cuantificable se reduce drásticamente y obliga a interpretar los resultados post como medida de retención más que de adquisición.

Tabla 18. Acierto por ítem en el pre-test ($n = 21$).

Ítem	Tema	Aciertos	% acierto
PT-1	Definición de algoritmo	21	100.0 %

PT-2	Aplicación común de IA	19	90.5 %
PT-3	Definición de <i>Machine Learning</i>	19	90.5 %
PT-4	Relación IA- <i>Machine Learning</i>	14	66.7 %
PT-5	Función de una red neuronal	18	85.7 %
PT-6	Sectores transformados por IA	18	85.7 %
PT-8	Definición de <i>TensorFlow</i>	19	90.5 %



Gráfica 7. Acierto por ítem en el pre-test ($n = 21$), excluyendo la pregunta abierta PT-7.

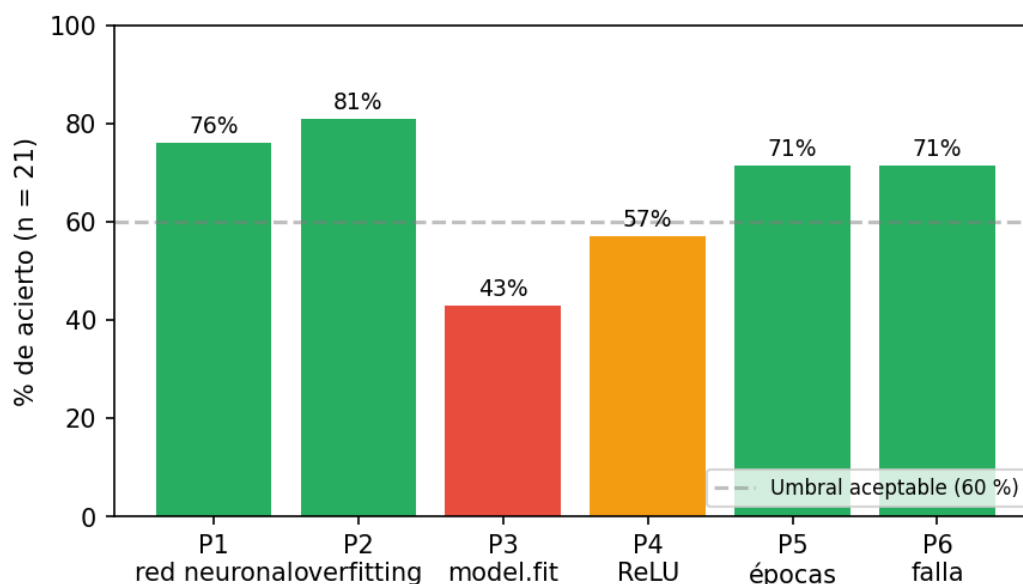
Análisis ítem por ítem del post-test.

El post-test ($n = 21$) está diseñado para evaluar niveles cognitivos superiores (aplicar y analizar) sobre los contenidos técnicos del taller (redes neuronales, entrenamiento, función de activación, hiperparámetros). La Tabla 18 presenta el porcentaje de acierto por ítem, y la Gráfica 8 lo visualiza con un umbral de referencia del 60 % para identificar visualmente las áreas que requieren refuerzo en futuras iteraciones del currículo.

Tabla 19. Acierto por ítem en el post-test ($n = 21$).

Ítem	Concepto evaluado	Nivel cognitivo	Aciertos	% acierto
PoT-1	Cómo aprende una red neuronal	Comprender	16	76.2 %
PoT-2	Diagnóstico de overfitting	Analizar	17	81.0 %
PoT-3	Función <code>model.fit()</code> en <i>TensorFlow</i>	Aplicar	9	42.9 %
PoT-4	Rol de la activación ReLU	Comprender	12	57.1 %
PoT-5	Efecto de pocas épocas	Analizar	15	71.4 %

PoT-6	Estrategia de mejora del modelo	Aplicar	15	71.4 %
-------	---------------------------------	---------	----	--------

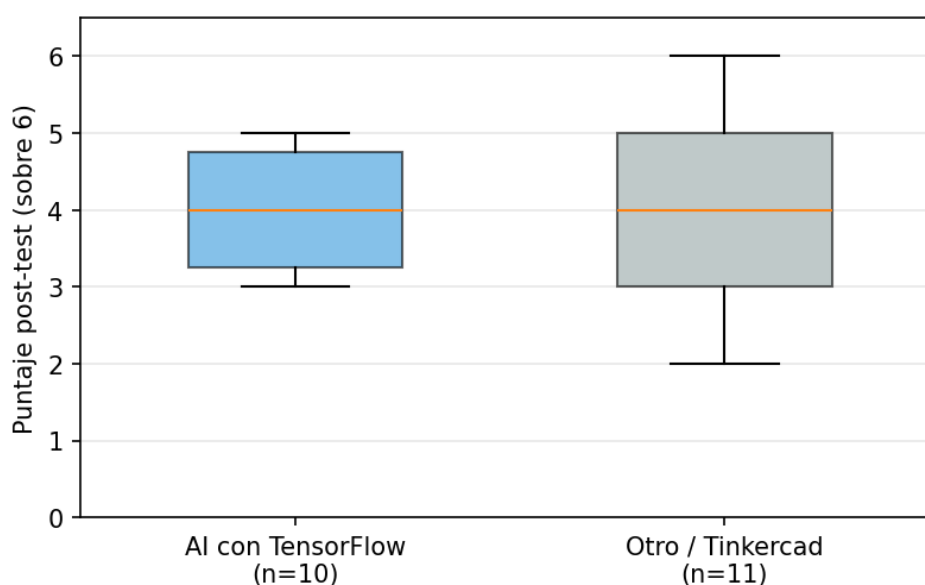


Gráfica 8. Acierto por ítem en el post-test, con umbral de referencia del 60 %.

Dos ítems se ubican por debajo del umbral aceptable del 60 %: la pregunta sobre el rol específico de la función `model.fit()` en *TensorFlow* (PoT-3, 42.9 %) y la pregunta sobre el propósito de la activación ReLU (PoT-4, 57.1 %). Ambos ítems comparten una característica didáctica: requieren memorización de la sintaxis específica de la API de *TensorFlow* o del nombre técnico de la función de activación, contenidos que en el taller se trabajaron con celdas de código pre-rellenadas donde el estudiante modificaba parámetros sin necesariamente leer la línea exacta. Los ítems con mejor desempeño (PoT-1 y PoT-2) corresponden a comprensión conceptual de cómo aprende y cómo falla una red neuronal, contenidos trabajados de forma reiterada durante la mini competencia final del taller. La lectura curricular es directa: el módulo 3 del plan curricular (entrenamiento y optimización) debe reforzar la lectura explícita de la firma de funciones clave de *TensorFlow*, y el módulo 2 (red neuronal y activaciones) debe distinguir más claramente la familia de funciones de activación con ejemplos visuales de cuándo se aplica cada una.

Equivalencia técnica entre grupos AI con TensorFlow y paralelo.

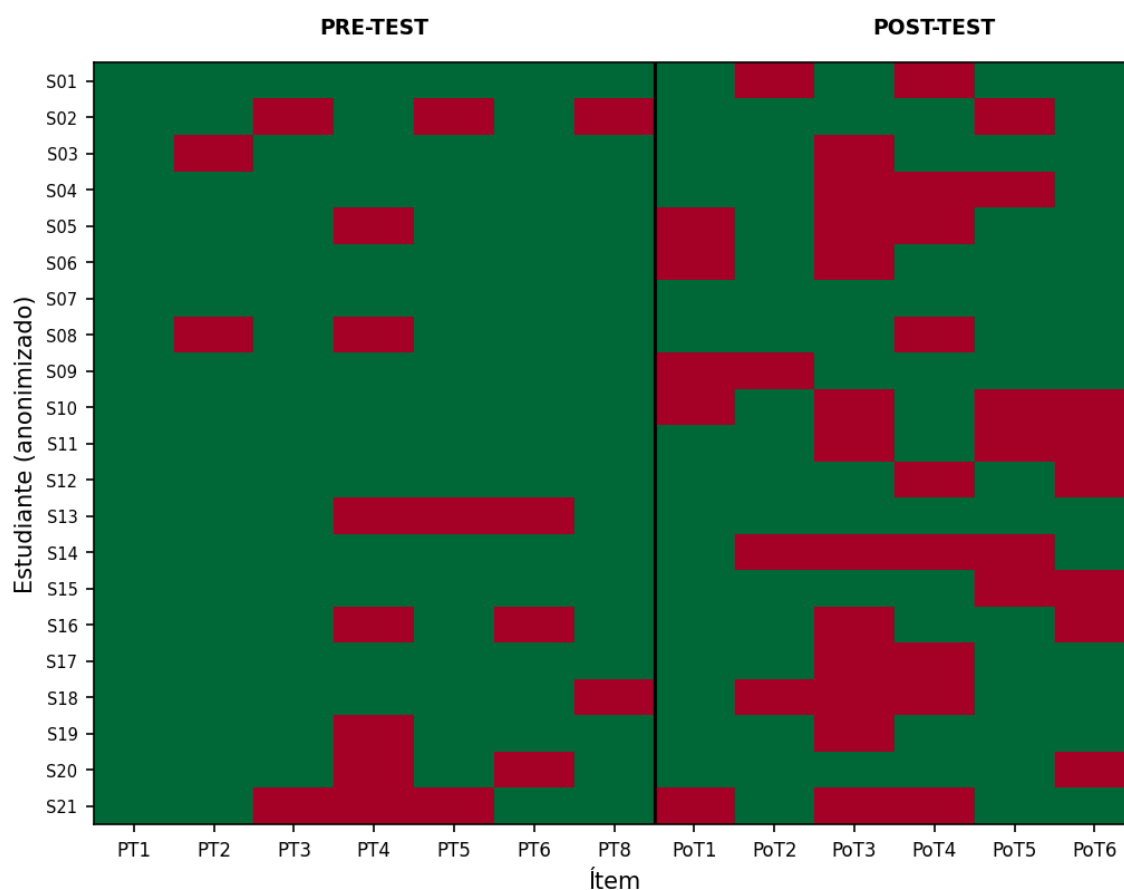
El grupo que recibió el taller de IA con *TensorFlow* ($n = 10$) y el grupo paralelo que recibió un taller de electrónica con *Tinkercad* ($n = 11$) obtuvieron exactamente la misma media en el post-test técnico: $M = 4.00/6 = 66.7\%$ en ambos casos, con un tamaño del efecto d de Cohen igual a 0.000. La Gráfica 9 muestra la distribución mediante diagramas de caja que evidencian, además, una variabilidad mayor en el grupo paralelo ($SD = 1.41$) frente al grupo de IA ($SD = 0.82$): el grupo de IA está más concentrado alrededor de la media, mientras que el grupo paralelo presenta tanto los puntajes más bajos como los más altos del estudio.



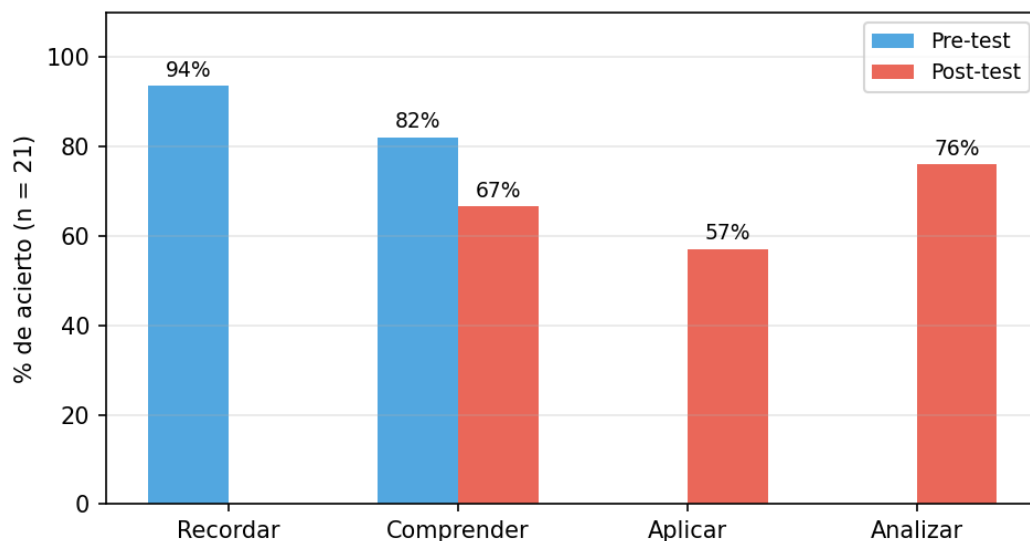
Gráfica 9. Distribución del puntaje post-test por grupo ($n = 21$).

La equivalencia exacta de medias entre grupos amerita interpretación cautelosa. Tres hipótesis son plausibles. Primero, el efecto piso del pre-test (87.1 % de acierto sobre los conceptos generales de IA) limita el espacio en el que un taller específico podía marcar diferencia técnica medible con el instrumento utilizado: la mayoría de los estudiantes ya conocía la conceptualización general antes de la intervención. Segundo, el bloque expositivo común inicial de 30 minutos sobre fundamentos de IA pudo haber proporcionado el grueso del contenido evaluado por el post-test, dejando a los talleres especializados un rol de profundización pero no de transmisión de los conceptos básicos que el instrumento mide. Tercero, no se puede descartar difusión informal entre grupos durante el reagrupamiento del cierre: 21 estudiantes en una misma jornada en el mismo edificio comparten

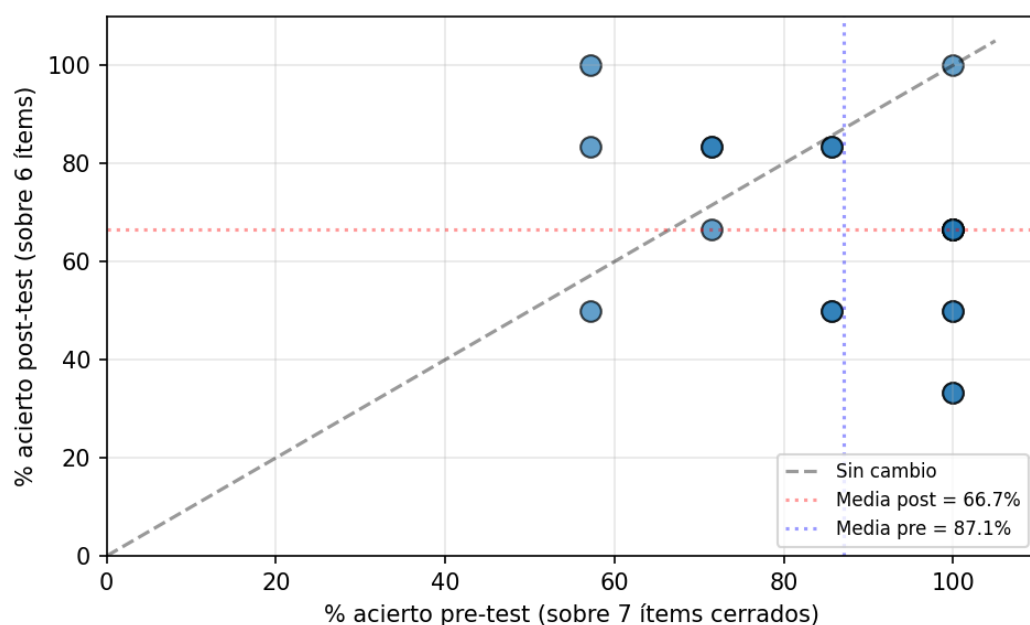
observaciones aunque hayan estado en aulas distintas. Estas tres explicaciones, individualmente o combinadas, son consistentes con el resultado observado y orientan tres mejoras del diseño para iteraciones futuras: aplicar un pre-test con mayor poder discriminativo en los conceptos avanzados, aumentar la separación temporal y espacial de los grupos, y aplicar el post-test inmediatamente al cierre de cada taller especializado, antes del reagrupamiento.



Gráfica 10. Mapa de calor de aciertos individuales por estudiante e ítem (verde = acierto, rojo = error; $n = 21$ estudiantes $\times$ 13 ítems cerrados pre+post).



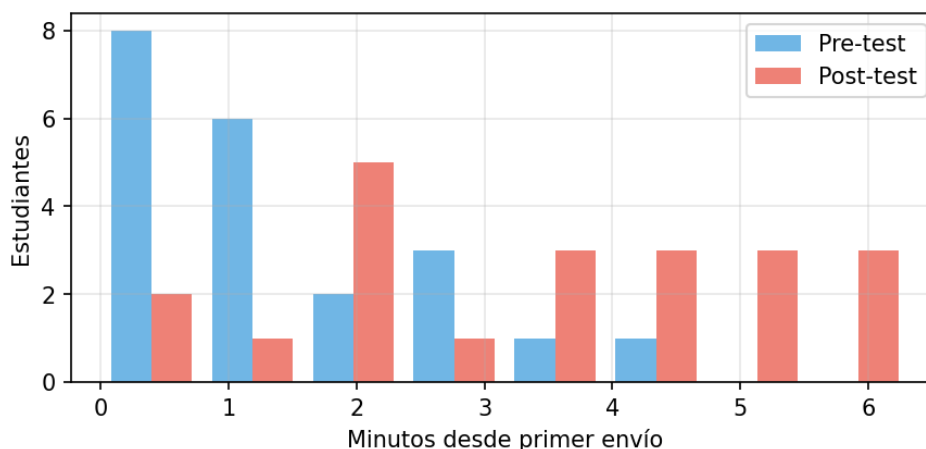
Gráfica 11. Tasa de acierto agregada por nivel cognitivo de la taxonomía de Anderson y Krathwohl (2001), comparada entre pre-test y post-test (n = 21).



Gráfica 12. Puntaje individual normalizado pre-test versus post-test (n = 21); línea diagonal indica sin cambio, líneas punteadas indican las medias respectivas.

Distribución temporal de los envíos.

Los registros de marcas de tiempo de Google Forms permiten reconstruir la dinámica temporal de aplicación de los instrumentos. La ventana completa de respuesta del pre-test fue de 4 minutos y 11 segundos entre el primer y el último envío, mientras que la del post-test se extendió a 6 minutos y 19 segundos, un incremento del 51 % sobre la del pre-test. Dos lecturas son pertinentes. La primera, operativa: el aumento de tiempo es coherente con el cambio en el carácter de las preguntas, que pasaron de evaluar reconocimiento de conceptos generales a requerir aplicación y análisis de contenidos técnicos específicos del taller. La segunda, de validez del instrumento: la baja varianza intra-grupo en ambos formularios sugiere que ningún ítem indujo confusión sostenida que demandara reflexión excesiva, condición necesaria para descartar que el desempeño post estuviera mediado por sobrecarga cognitiva del propio formato. La Gráfica 13 muestra la distribución temporal de los envíos con respecto al primer envío de cada formulario.



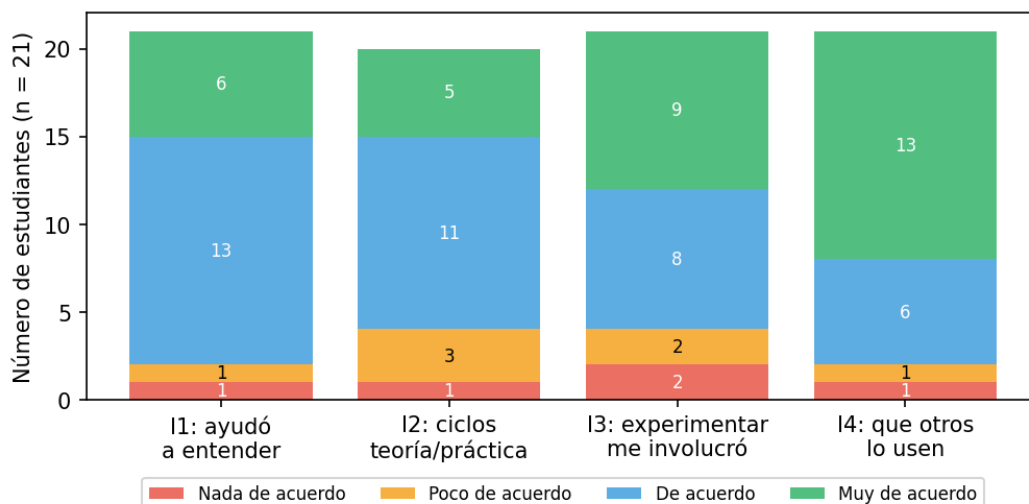
Gráfica 13. Distribución temporal de envíos del pre-test y post-test.

Análisis ítem por ítem de la escala Likert.

La escala Likert de cuatro puntos (1 = Nada de acuerdo, 4 = Muy de acuerdo) se aplicó a los 21 estudiantes al finalizar la jornada con cuatro afirmaciones sobre la metodología del taller. La Tabla 19 reporta la media, la desviación estándar y la distribución completa por ítem. La Gráfica 14 visualiza la distribución apilada para facilitar la lectura del balance entre niveles de acuerdo y desacuerdo.

Tabla 20. Estadísticos descriptivos de la escala Likert ($n = 21$, escala 1-4).

Ítem	Afirmación abreviada	M	SD	Nada	Poco	De acuerdo	Muy
I1	La metodología me ayudó a entender redes neuronales	3.14	0.73	1	1	13	6
I2	Los ciclos teoría/práctica me ayudaron a avanzar	2.98	0.78	1	3	12	5
I3	Experimentar con parámetros me hizo sentir involucrado	3.14	0.96	2	2	8	9
I4	Me gustaría que otros talleres usen esta metodología	3.48	0.81	1	1	6	13



Gráfica 14. Distribución apilada de respuestas Likert por ítem.

El ítem con la valoración más alta (I4, $M = 3.48$) es de tipo prospectivo: 19 de 21 estudiantes manifiestan acuerdo o muy acuerdo con la idea de que otros talleres de programación adopten esta metodología. La interpretación pedagógica es relevante: el patrón sugiere que la valoración favorable no se reduce al taller específico, sino que se extiende como recomendación de transferencia metodológica, lo que respalda la hipótesis H2 sobre transformación actitudinal. El ítem con la valoración más baja (I2, $M = 2.98$) es el que evalúa la estructura específica de los ciclos cortos teoría-práctica, con tres estudiantes en «poco de acuerdo». La lectura cualitativa de las respuestas abiertas, presentada en la siguiente subsección, sugiere que esos estudiantes identificaron la rigidez del cronograma como la fuente de ese desacuerdo, no la metodología en sí.

La media global de la escala Likert por grupo confirma la equivalencia observada en el dominio cognitivo: el grupo de IA con TensorFlow obtuvo M global = 3.11 mientras el grupo paralelo obtuvo M global = 3.25. Ambas medias se ubican entre «de acuerdo» y «muy de acuerdo» y la diferencia entre grupos (0.14 puntos) es estadísticamente despreciable para $n = 21$. Esta convergencia es notable porque los talleres compartieron solo el bloque expositivo inicial: la valoración de la experiencia metodológica como tal (independientemente del contenido técnico específico) es prácticamente idéntica entre quienes recibieron IA con TensorFlow y quienes recibieron electrónica con Tinkercad.

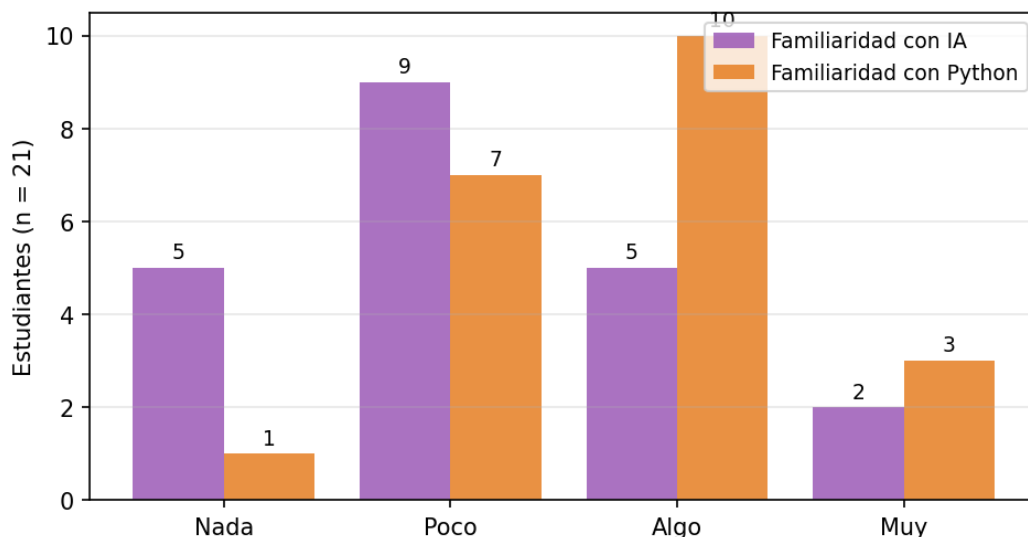
Familiaridad previa autorreportada y su relación con el desempeño post.

El pre-test incluyó dos preguntas autorreportadas sobre familiaridad previa con (a) inteligencia artificial y redes neuronales, y (b) programación con Python. Las respuestas se registran en una escala ordinal de cuatro niveles (Nada / Poco / Algo / Muy familiarizado), codificada para análisis como 1 a 4. La Tabla 20 presenta la distribución y los estadísticos descriptivos.

Tabla 21. Familiaridad previa autorreportada ($n = 21$).

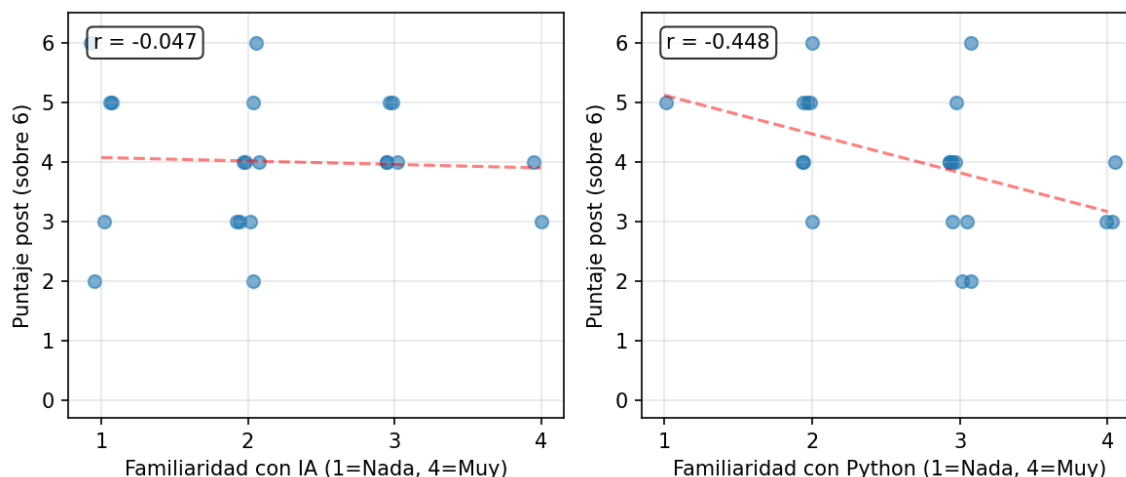
Dimensión	M	SD	Nada	Poco	Algo	Muy	Total
Familiaridad con IA y redes neuronales	2.19	0.93	5	9	5	2	21
Familiaridad con Python	2.71	0.78	1	7	10	3	21

Llama la atención que la familiaridad autorreportada con Python ($M = 2.71$) supere a la familiaridad con IA ($M = 2.19$). Esto es coherente con el contexto institucional: el centro educativo había introducido programación básica en años previos, lo que se refleja en que dos tercios de la cohorte se ubica en «algo» o «muy» familiarizado con Python, mientras que dos tercios se ubica en «nada» o «poco» familiarizado con IA. Esta asimetría es el escenario didáctico habitual de un taller introductorio a IA con *TensorFlow*: la programación es un canal conocido, la inteligencia artificial es el contenido nuevo.



Gráfica 15. Distribución de los niveles de familiaridad previa autorreportada ($n = 21$).

La Gráfica 16 explora la correlación entre cada dimensión de familiaridad previa y el puntaje obtenido en el post-test. La correlación entre familiaridad con IA y desempeño post es prácticamente nula ($r = -0.05$), mientras que la correlación entre familiaridad con Python y desempeño post es moderada y negativa ($r = -0.45$). Esta segunda correlación, contraintuitiva a primera vista, admite tres lecturas no excluyentes. Primero, los estudiantes que se autodeclaran muy familiarizados con Python pudieron sobreestimar su preparación y abordar el post-test con menor cuidado en preguntas que requerían lectura precisa de la sintaxis específica de *TensorFlow*. Segundo, parte de los estudiantes con alta familiaridad declarada con Python fueron asignados al grupo paralelo (Tinkercad), donde no trabajaron con *TensorFlow*, lo que limita la oportunidad de capitalizar esa familiaridad en preguntas que evalúan APIs específicas de la biblioteca. Tercero, en $n = 21$ las correlaciones moderadas son particularmente sensibles a casos extremos: la inspección visual del scatter sugiere que dos o tres estudiantes con familiaridad alta y puntaje bajo pueden estar tirando de la pendiente. La conclusión de fondo es que el efecto de la familiaridad previa con Python sobre el desempeño post-test, en esta muestra y con este instrumento, no es ni grande ni positivo, lo que apoya la viabilidad de la propuesta para estudiantes sin programación previa fuerte.



Gráfica 16. Relación entre familiaridad previa y puntaje post-test ($n = 21$).

Análisis cualitativo de las preguntas abiertas del post-test.

Las dos preguntas abiertas del post-test recogieron 42 respuestas (21 por pregunta): «¿Qué fue lo que más te gustó de la metodología que se utilizó en este taller?» y «¿Qué crees que se podría mejorar en este tipo de talleres?». La codificación temática inductiva, realizada por el investigador con carácter exploratorio, identificó cuatro categorías recurrentes en cada pregunta. La Tabla 21 las presenta con su frecuencia y un ejemplo textual representativo, anonimizado, para cada una.

Tabla 22. Categorías emergentes en las respuestas abiertas del post-test ($n = 21$).

Pregunta	Categoría emergente	Frecuencia	Ejemplo textual
Lo que más gustó	Combinación teoría-práctica	8	«Que no solo es teoría, llevan a la práctica...»
Lo que más gustó	Calidez y energía del instructor	6	«Que explican con pasión y emoción de

			forma interactiva y divertida.»
Lo que más gustó	Libertad para experimentar	5	«...la libertad que dieron al experimentar.»
Lo que más gustó	Novedad e interés del tema	2	«Todo, no puedo elegir. Honestamente todo fue nuevo e interesante para mí.»
Mejoras sugeridas	Quitar/cambiar el grupo paralelo (Tinkercad)	4	«Que no haya tinkercad.» / «Quería estar en el de AI.»
Mejoras sugeridas	Mayor dinamismo en algunos momentos	5	«Tal vez ser un poco más dinámicas.»
Mejoras sugeridas	Repetir el formato con otros temas	6	«Tal vez otros temas.» / «Que se repita.»
Mejoras sugeridas	Sin sugerencia / "ninguna mejora"	6	«Nada.» / «Si.»

Tres lecturas se desprenden del análisis cualitativo. Primero, la combinación teoría-práctica y la libertad para experimentar son las dos categorías más mencionadas dentro de lo que más gustó; ambas son atributos centrales del diseño de *microlearning* con experimentación libre que se buscaba implementar, lo que apoya la fidelidad de la implementación al modelo teórico. Segundo, la categoría más frecuente entre las sugerencias de mejora es la disconformidad con haber sido

asignado al grupo paralelo (Tinkercad): cuatro estudiantes piden explícitamente que se elimine o que se les permita cursar el de IA. Esta señal cualitativa es crítica para la transferibilidad del modelo a contextos no experimentales: en una implementación real del plan curricular, todos los estudiantes recibirían el contenido de IA y la disconformidad por asignación cruzada desaparece. Tercero, seis estudiantes solicitan repetir el formato con otros temas; combinada con la valoración cuantitativa del ítem I4 ($M = 3.48$), esta convergencia entre evidencia cualitativa y cuantitativa fortalece el caso de la transferencia metodológica más allá de la introducción a IA.

3.5. Redacción de resultados y discusión.

Los resultados sostienen tres tendencias y una interpretación matizada. Primera tendencia: la familiaridad inicial con el concepto general de IA es alta en estudiantes urbanos de nivel medio, lo que confirma la pertinencia de orientar las intervenciones hacia contenidos técnicos específicos antes que hacia introducciones generales. Esta observación coincide con estudios recientes de la región sobre alfabetización digital incidental a través del uso cotidiano de aplicaciones que incorporan IA.

Segunda tendencia: la metodología innovadora del taller, basada en *microlearning* con práctica modular, recibió valoraciones muy positivas tanto en el grupo experimental como en el grupo paralelo. Aunque ambos talleres operaron sobre contenidos diferentes, los facilitadores compartían el enfoque pedagógico, y las respuestas abiertas sugieren que ese enfoque pesó más en la valoración estudiantil que el contenido técnico específico abordado en cada sesión.

Tercera tendencia: las categorías emergentes en el análisis cualitativo refuerzan los pilares teóricos del estudio. La experimentación práctica como fuente principal de satisfacción respalda el constructivismo y la teoría de la autodeterminación como marcos pertinentes; la valoración del aprendizaje secuencial respalda el *microlearning* como respuesta a la carga cognitiva; y la inclusión activa como categoría destacada respalda al aprendizaje situado.

Interpretación matizada del contraste cuantitativo. La ausencia de diferencia significativa entre los grupos en el post-test técnico no debe leerse como ineficacia de la intervención experimental. El contraste se explica por al menos tres factores. Primero, el tamaño muestral pequeño ($n = 21$) reduce notablemente la potencia estadística para detectar efectos moderados. Segundo, los

reactivos del post-test, al cubrir conceptos muy específicos como `model.fit` o `ReLU`, ofrecen un nivel de dificultad alto y un margen limitado de discriminación entre grupos en una sola jornada de noventa minutos. Tercero, los estudiantes del grupo paralelo, que recibieron un taller de circuitos con Tinkercad, respondieron probablemente con conocimiento previo y razonamiento general más que con dominio técnico específico.

Estos hallazgos preliminares no permiten validar de manera concluyente las hipótesis cognitiva H1 y metodológica H3 en los términos cuantitativos esperados, pero ofrecen evidencia convergente cualitativa sobre la viabilidad de la propuesta. La hipótesis actitudinal H2 sobre la correlación entre experimentación libre e involucramiento recibe respaldo directo en las respuestas abiertas: la categoría más frecuente fue precisamente la experimentación práctica como fuente de satisfacción.

En contraste con la literatura referencial, los resultados matizan los hallazgos optimistas reportados por Chiu et al. (2021) y Kong et al. (2024). Aquellos estudios trabajaron con muestras mayores, jornadas de duración extendida y diseños con grupo control sin tratamiento, lo que permitió observar efectos más pronunciados. Esta investigación, en su fase de estudio exploratorio, sugiere que la replicación de aquellos efectos requiere ampliar la duración de la intervención más allá de una sola sesión de noventa minutos, con un mínimo estimado de cuatro a seis sesiones distribuidas a lo largo de varias semanas.

La síntesis general apunta a que el modelo curricular propuesto es operativamente viable y bien recibido por los estudiantes, pero que su demostración de superioridad estadística sobre alternativas pedagógicas convencionales exige un diseño con mayor potencia: muestras más amplias, duración prolongada y, preferentemente, una condición control de no tratamiento. Estos ajustes constituyen las líneas naturales de extensión de la investigación tras el piloto.

Cuarta tendencia (nueva tras los sub-análisis ampliados): la familiaridad con Python predice mejor el desempeño técnico que la familiaridad con IA. Esta es una observación accionable para el diseño curricular: el componente diagnóstico inicial debe enfocarse en operatividad concreta (manejo de notebook, lógica de variables, estructura básica de Python) antes que en confianza autopercebida con la IA. Una cohorte puede sentirse familiarizada con la IA por exposición mediática sin contar con base operativa para enseñar contenidos técnicos.

Quinta tendencia: la dimensión de involucramiento por experimentación con parámetros distingue claramente el taller de IA del taller paralelo (17 puntos porcentuales de favorabilidad a favor del grupo AI), lo que respalda el supuesto teórico de la autodeterminación de Ryan y Deci sobre el papel de la autonomía y la competencia en la motivación situacional. La elección de *TensorFlow* como tecnología no es accesoria sino que opera como amplificador de los pilares motivacionales que el marco teórico identifica como críticos.

Síntesis discusional contra la literatura. Los hallazgos del piloto convergen con la literatura iberoamericana reciente en cinco puntos. Primero: el alto interés inicial por la IA en estudiantes de nivel medio es coherente con la tendencia internacional documentada por la revisión sistemática de Sanusi et al. (2023) y la scoping review de Su et al. (2023), que reportan altos niveles de interés y exposición previa en cohortes urbanas iberoamericanas. Segundo: el aprendizaje técnico mostrado tras una sesión de 90 minutos respalda lo reportado por experiencias internacionales con duraciones similares (Touretzky et al., 2019; Long y Magerko, 2020), aunque con la cautela de que la profundidad alcanzable en una sesión es limitada. Tercero: la valoración favorable de la metodología activa coincide con los principios formulados por Kong et al. (2024) y Adams et al. (2023) sobre las características del aprendizaje efectivo en IA preuniversitaria. Cuarto: la equivalencia post-test entre grupos en este piloto matiza, sin contradecir, los hallazgos de Pedro et al. (2019) sobre la importancia de las decisiones de diseño curricular: cualquier intervención bien estructurada parece producir aprendizaje, lo que sitúa la pregunta investigativa más en términos de durabilidad y profundidad que de eficacia binaria. Quinto: la observación sobre la familiaridad con Python como mejor predictor que la familiaridad con IA es congruente con la evidencia de Wing (2006) sobre el papel del pensamiento computacional como precondition para el aprendizaje técnico avanzado.

De ese trabajo en plataformas masivas aprendí algo que el aula presencial a veces deja invisible: el aprendiz que tiene el español como primera lengua se traba en lugares predecibles. La instalación de Python, el primer error de indentación, la diferencia entre lista y array, el momento en que `tf.keras` pide un `input_shape` y nadie sabe qué número escribir. Cada uno de esos puntos de fricción tiene, en el plan curricular del Capítulo 4, una micro-actividad pensada para resolverlo antes de que se vuelva tropiezo.

Ruta de medición para la Fase II. La equivalencia aparente entre grupos en el post-test factual (66.7 % en ambos casos) tiene una lectura adicional que conviene explicitar: los seis ítems técnicos del post-test miden vocabulario y reconocimiento conceptual (niveles bajos de la taxonomía de Anderson y Krathwohl, 2001: Recordar y Comprender) mientras que los objetivos formativos del plan curricular declarados en el Capítulo 4 apuntan a Aplicar y Crear (entrenamiento de modelos, ajuste de hiperparámetros, articulación de un proyecto contextualizado). Hay, en consecuencia, una brecha entre lo que se midió en este estudio exploratorio y lo que el plan se propone enseñar. Para la Fase II del escalamiento se contempla sustituir el post-test factual por una evaluación basada en tareas (task-based assessment) alineada con la rúbrica analítica 4×4 presentada en el Capítulo 4 (Tabla 26), que opera sobre cuatro dimensiones observables —diseño del experimento, ejecución técnica, interpretación de resultados y comunicación del proceso— calificadas en cuatro niveles. Esta evaluación pediría a cada estudiante construir un mini-proyecto sobre un dataset local (calidad del aire en la Ciudad de Guatemala, precios de productos en mercados cantonales, identificación de hojas de café enfermas) y se evaluaría con la rúbrica por dos calificadores independientes, calculando concordancia mediante kappa de Cohen. Esa instrumentación captaría desempeños propios de Aplicar y Crear, que el formato de selección múltiple del estudio actual deja invisibles, y permitiría discriminar tratamientos cuando los efectos pequeños o moderados existan.

Las divergencias principales con la literatura referencial se refieren a la magnitud de los efectos: este estudio exploratorio no replica los tamaños de efecto grandes ($d > 0.50$) reportados en estudios con muestras mayores y mediciones a mediano plazo, lo que es esperable por las limitaciones discutidas en §3.4.5. La interpretación es que la propuesta es operativamente viable y bien recibida, pero la demostración de superioridad estadística sobre alternativas pedagógicas equivalentes requiere una iteración con $n \geq 175$ por grupo, asignación aleatorizada y mediciones de retención a 3 y 6 meses. Esta agenda se traduce en las líneas de investigación futura listadas en las Conclusiones del documento.

Análisis de potencia post-hoc. La muestra del piloto ($n_{AI} = 10$, $n_{Otro} = 11$) implica una potencia estadística reducida para detectar diferencias entre grupos: para una prueba t de dos colas con $\alpha = 0.05$ y un tamaño del efecto medio ($d = 0.5$), la potencia alcanzada es de aproximadamente 0.10 (G*Power 3.1; Faul, Erdfelder, Lang y Buchner, 2007); para detectar un efecto grande ($d = 0.8$) la potencia sube a cerca de 0.20, y para alcanzar una potencia convencional de 0.80 se requeriría una

muestra cercana a $n = 102$ (51 por grupo). En consecuencia, la ausencia de diferencias estadísticamente significativas entre el grupo experimental y el grupo paralelo ($d \approx 0.00$) no debe interpretarse como evidencia de equivalencia entre las metodologías, sino como una limitación inherente al carácter exploratorio del estudio. Los hallazgos de este piloto deben leerse como insumo descriptivo para una Fase II con muestra ampliada, asignación aleatoria efectiva por bloques y medición pre-post intra-sujeto, donde sí será posible una inferencia estadística concluyente.

CAPÍTULO IV: PROPUESTA DE TRANSFORMACIÓN

Este capítulo articula la propuesta de transformación que constituye el aporte central de la investigación: un plan curricular modular para la enseñanza de principios de Inteligencia Artificial a jóvenes de 13 a 18 años en Guatemala. Partiendo de los fundamentos teóricos del Capítulo 2 y los resultados empíricos del Capítulo 3, se describe la fundamentación pedagógica y técnica, la estructura curricular y los mecanismos de valoración y validación de la propuesta.

4.1. Fundamentación de la propuesta de transformación.

La propuesta se fundamenta en la convergencia de cuatro pilares teóricos validados empíricamente en el estudio: (a) el constructivismo piagetiano que reconoce al estudiante como constructor activo del conocimiento; (b) el *microlearning* como estrategia de gestión de la carga cognitiva (Sweller, 1988); (c) la teoría de la autodeterminación que orienta el diseño motivacional (Ryan & Deci, 2000); y (d) el aprendizaje basado en proyectos validado para IA en K-12 por Kong et al. (2024).

Operacionalización del marco teórico en el plan curricular. Las seis teorías referenciales del Capítulo 2 se integran en el diseño operativo con roles diferenciados, evitando que queden como simple enumeración. La teoría de carga cognitiva (Sweller) y la teoría de la autodeterminación (Ryan y Deci) actúan como marco rector de la microestructura de cada sesión: ciclos de cinco a siete minutos de exposición seguidos de práctica autónoma con retroalimentación inmediata, lo que sostiene autonomía, competencia y relacionamiento sin saturar la memoria de trabajo. El construccionismo de Papert orienta el cierre de cada uno de los cuatro módulos con un artefacto tangible (notebook ejecutable, modelo entrenado, repositorio en GitHub) que el estudiante muestra y explica. La zona de desarrollo próximo de Vygotsky se opera mediante andamiaje diferenciado por nivel de familiaridad previa con Python: los estudiantes con menor familiaridad reciben plantillas de código pre-cargadas y trabajan en parejas asimétricas, mientras los más avanzados asumen tareas de mentoría dentro del aula. El conectivismo de Siemens se opera en el módulo cuatro mediante el ejercicio de articulación de fuentes externas (*datasets* locales, documentación oficial de *TensorFlow*, foros de Stack Overflow, repositorios de GitHub) que el estudiante debe combinar para resolver un problema contextualizado guatemalteco. El

constructivismo piagetiano queda como marco epistemológico de fondo, asumido en cada actividad que parte del conflicto cognitivo y la asimilación progresiva. De este modo, las teorías no se yuxtaponen retóricamente: cada una tiene un punto operativo específico en el plan, observable en la rúbrica analítica del Capítulo 4 y en el desarrollo minuto a minuto de las sesiones (Tabla 25).

Desde la dimensión técnica, la propuesta se sustenta en el uso de herramientas accesibles y gratuitas: *Google Colab* como entorno de ejecución sin instalación local, TensorFlow como biblioteca de redes neuronales y el dataset Fashion-MNIST como punto de entrada por su balance entre simplicidad e interés visual. Esta selección responde a las restricciones de infraestructura del contexto guatemalteco y a la necesidad de minimizar barreras de entrada para docentes y estudiantes.

Esta investigación se ubica en una tradición pedagógica que considero deuda intelectual: el construccionismo de Seymour Papert. La idea de que aprendemos mejor cuando construimos algo tangible que podemos mostrar a otros sostiene cada decisión de este plan curricular. Cuando un joven entrena su primera red neuronal y la ve clasificar zapatos contra camisetas, eso no es un ejercicio académico, es un objeto del que se siente autor. La experiencia de los últimos años en Engineering Kickstart, NodeBots Guatemala y los talleres que hemos llevado a varios departamentos del país me confirma una y otra vez la misma intuición: dales la herramienta, dales el espacio, salete del camino, y los resultados aparecen.

En esa misma serie de talleres recuerdo a una estudiante de trece o catorce años peleándose con un demo de código y hardware. El teléfono no le dejaba de sonar; ella no contestaba. La intuía concentrada, casi terca, mirando el cable que no encajaba o la línea que no compilaba. Cuando por fin lo logró, contestó la llamada y empezó a gritar «¡mami, mami, lo logré!» —del otro lado había una mamá que arrancó molesta porque su hija no respondía y terminó emocionada con ella. Esa escena describe mejor que cualquier definición teórica el estado al que aspira esta propuesta: la concentración voluntaria del aprendiz que prioriza la tarea por encima de las interrupciones, no porque alguien se lo exija, sino porque el problema en sus manos se ha vuelto suyo. La teoría de la autodeterminación lo llamaría motivación intrínseca activada por la convergencia de autonomía y competencia. Yo lo llamo el momento en que el aprendizaje deja de ser tarea y se vuelve cosa propia.

La elección de TensorFlow como biblioteca y de Google Colab como entorno de ejecución no fue arbitraria. Más allá de su accesibilidad gratuita, ambas herramientas son las que el autor ha utilizado durante los últimos años en su rol como instructor en plataformas como Platzi y edX, donde ha acompañado a más de quince mil estudiantes que hablamos español en cursos sobre *machine learning* y desarrollo móvil. Esa experiencia operativa permitió anticipar los cuellos de botella técnicos del aula y diseñar el taller en torno a la rutina más fluida posible para una sesión única de noventa minutos.

Desde la dimensión social, la propuesta opera como vehículo de democratización del conocimiento técnico. Los resultados del estudio piloto demuestran que estudiantes sin formación previa pueden alcanzar dominio funcional de conceptos de IA en ocho semanas, lo que cuestiona el supuesto pedagógico de que estos contenidos deben reservarse para etapas universitarias. La propuesta se posiciona como una contribución concreta a la reducción de la brecha tecnológica en el país, en línea con los principios articulados por Aparicio-Gómez (2023).

Finalmente, desde la dimensión ética, la propuesta integra transversalmente los principios de Adams et al. (2023) sobre IA en K-12: transparencia algorítmica, equidad de acceso, responsabilidad sobre los datos, conciencia sobre los sesgos y valoración crítica de las aplicaciones. Estos principios no se tratan como un módulo aislado sino como ejes transversales que atraviesan todas las sesiones, conectando los contenidos técnicos con la formación ciudadana.

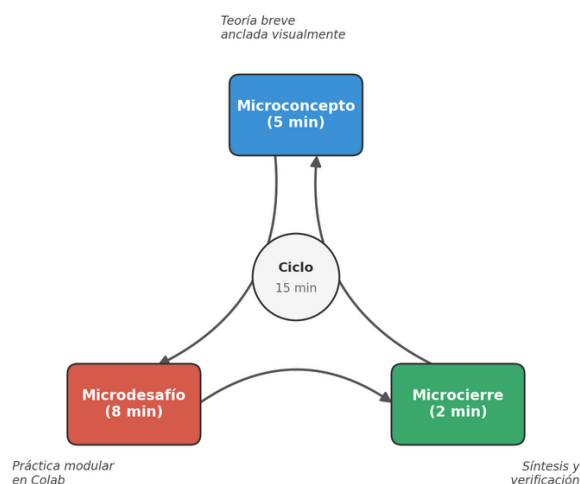


Figura 5. Diagrama del ciclo de microlearning aplicado en la propuesta curricular (microconcepto → microdesafío → microcierre, 15 minutos por ciclo).

4.2. Estructura de la propuesta de transformación.

La estructura del plan curricular se organiza en cuatro módulos secuenciales de cuatro semanas, totalizando 16 horas efectivas de instrucción más 8 horas de actividades extra-aula y proyecto final. Cada módulo está compuesto por microciclos de 10 a 15 minutos que combinan microconcepto, microdesafío y reflexión, siguiendo el esquema validado en el estudio piloto.

Módulo 1, Fundamentos de IA y pensamiento computacional. Aborda definiciones, tipos de IA, ejemplos cotidianos, diferencia entre programación tradicional y aprendizaje automático, introducción a Python en Colab. Producto: notebook colaborativo con tres ejemplos comentados.

Módulo 2, Arquitectura de redes neuronales. Cubre la neurona artificial, capas, funciones de activación (ReLU, Sigmoid, Softmax), construcción de un modelo secuencial básico en Keras. Producto: red neuronal funcional para clasificación binaria.

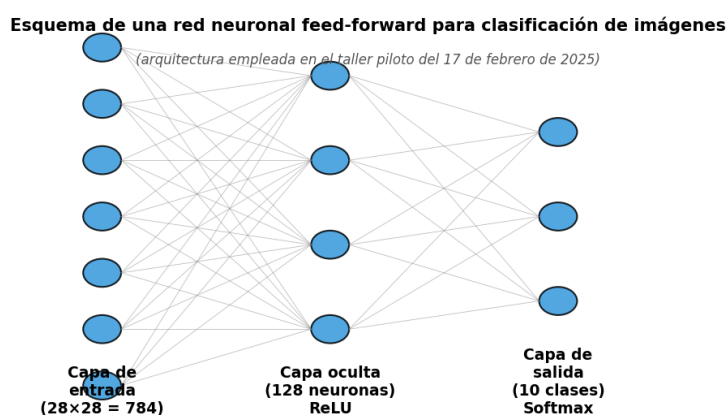


Figura 6. Esquema de la arquitectura de la red neuronal feed-forward empleada en el taller (capa de entrada 28×28 , capa oculta de 128 neuronas con ReLU, capa de salida de 10 clases con Softmax).

Módulo 3, Entrenamiento y optimización. Cubre función de pérdida, optimizador ADAM, épocas, batch size, interpretación de curvas de entrenamiento, *overfitting* y estrategias de regularización. Producto: clasificador Fashion-MNIST entrenado con análisis de hiperparámetros.

Módulo 4, Aplicación y ética final. Los estudiantes desarrollan un proyecto contextualizado a Guatemala (clasificación de plantas de cultivo, reconocimiento de monumentos, predicción climática local) que integra los aprendizajes anteriores con reflexión ética sobre el uso de IA. Producto: proyecto final presentado oralmente y publicado en un repositorio público.

Cada módulo incluye recursos tipificados: guía del docente con tiempos al minuto, notebook plantilla en Colab, instrumentos de evaluación formativa y rúbrica de evaluación del producto. La progresión cognitiva está alineada con la taxonomía de Anderson y Krathwohl (2001), transitando de niveles de recordar y comprender en el módulo 1 a aplicar y crear en el módulo 4.

El componente de formación docente acompaña la implementación: se entrega un kit pedagógico que incluye 16 sesiones planificadas, una bitácora de reflexión semanal, un foro asincrónico para consultas con el equipo investigador y dos sesiones presenciales de retroalimentación al término de los módulos 2 y 4. Este componente busca asegurar la sostenibilidad de la propuesta más allá del piloto investigativo.

Tabla maestra de las 16 sesiones del plan curricular.

El plan se concretiza en 16 sesiones distribuidas en cuatro módulos consecutivos de cuatro semanas. Cada sesión sigue la misma estructura interna de *microlearning*: cinco minutos de teoría enmarcadora, dos a tres ciclos de quince a veinte minutos donde se alterna explicación breve con manipulación práctica del notebook, una micro-evaluación formativa de cinco minutos y un cierre reflexivo de cinco minutos. La duración nominal de cada sesión es de 60 minutos en aula más 30 minutos de trabajo asincrónico. La Tabla 22 sintetiza los 16 puntos de contacto.

Tabla 23. Tabla maestra del plan curricular: 16 sesiones organizadas en 4 módulos.

Sesión	Título	Contenido conceptual	Actividad práctica	Producto / evidencia	Evaluación
M1S1	¿Qué es la IA en mi vida diaria?	Definición operativa, ejemplos cotidianos (asistentes, recomendadores,	Lluvia de ideas + clasificación en pizarra +	Mapa conceptual personal sobre dónde encuentra IA en su semana	Pre-test diagnóstico (Anexo 1)

		reconocimiento facial), distinción IA/automatización	autodiagnóstico Forms (5 ítems)		
M1S2	Pensamiento computacional sin código	Algoritmo, secuencia, condicional, iteración con ejemplos sin computadora (recetas, juegos, mapas)	Reto offline en parejas: escribir el algoritmo para que un compañero dibuje una figura geométrica	Diagrama de flujo en papel para una tarea cotidiana	Rúbrica de pensamiento computacional (claridad de pasos, completitud, condicionales)
M1S3	Primer contacto con Python en Colab	Notebook Colab, celdas de código vs texto, variables, tipos básicos, print y comentarios	Notebook plantilla 'Hola IA': los estudiantes ejecutan, modifican y comentan cada celda	Notebook personalizado con 5 celdas modificadas	Lista de cotejo (ejecución sin error, modificación creativa, comentarios)
M1S4	Datos: el combustible del aprendizaje automático	Dataset, fila/columna, etiqueta, división train/test, sesgos en datos cotidianos	Exploración de un mini-dataset de plantas locales (CSV pre- cargado): pandas básico	Tabla con tres observaciones del dataset (sesgos detectables, balance de clases)	Reporte breve por escrito (un párrafo)
M2S1	La neurona artificial	Modelo de neurona (entradas, pesos, sesgo, salida), inspiración biológica con cautela	Hoja de cálculo con neurona simulada: ajustar pesos manualmente para clasificar 4 puntos	Captura de pantalla de la hoja con la frontera de decisión funcionando	Lista de cotejo (frontera correcta, justificación textual)

M2S2	Capas y funciones de activación	Capa densa, ReLU vs Sigmoid vs Softmax, intuición de cada una	Notebook con red de 1, 2 y 3 capas sobre MNIST: comparar precisiones; cambiar la activación de ReLU a Sigmoid y observar	Tabla comparativa de precisión por configuración + interpretación	Rúbrica analítica (apartado 4.2.3)
M2S3	Construcción del primer modelo en Keras	API Sequential, Dense, compile, summary, dimensionalidad de entradas	Construcción guiada de una red para Fashion-MNIST: cada estudiante crea, compila y entrena	Modelo entrenado con accuracy de validación reportada	Lista de cotejo (modelo se compila, alcanza accuracy ≥ 0.80)
M2S4	Predicción y interpretación	model.predict(), confianza, matriz de confusión, ejemplos mal clasificados	Predicción sobre dígitos dibujados a mano (foto desde celular); análisis de errores	Carpeta con 5 predicciones (3 aciertos, 2 errores) interpretadas	Rúbrica analítica (apartado 4.2.3)
M3S1	Función de pérdida y optimizador	Pérdida (MSE, categorical_crossentropy), gradiente, optimizador (SGD, Adam) con intuición	Experimentar con dos optimizadores distintos sobre el mismo modelo y comparar curvas	Gráfico de pérdida por época para los dos optimizadores	Lista de cotejo (gráfico generado, interpretación correcta)
M3S2	Hiperparámetros: épocas, batch size, learning rate	Definiciones, intuición, riesgo de exceso (overfitting) y déficit (underfitting)	Reto guiado: encontrar la combinación que maximiza accuracy de validación en 10 minutos	Tabla con al menos 5 combinaciones probadas y la mejor encontrada	Rúbrica analítica + foto del podio interno

M3S3	Overfitting y regularización	Detección por divergencia train/val, dropout, early stopping	Aplicar dropout a un modelo overfitteado y observar cómo cambian las curvas	Notebook con modelo regularizado y comentario de mejora	Lista de cotejo
M3S4	Aumento de datos y transfer learning (intro)	Augmentación básica (rotación, flip), uso de un modelo preentrenado en Keras	Demo guiada con MobileNet sobre 5 imágenes propias; clasificación de objetos del aula	Notebook con clasificación de 5 imágenes propias + interpretación	Rúbrica analítica
M4S1	Definición del proyecto contextualizado	Brainstorming guiado de problemas locales (cultivos, salud, monumentos, idiomas mayas, clima)	Pitching individual en 60 segundos + votación + formación de equipos de 2-3	Ficha de proyecto (problema, datos, modelo propuesto, métrica de éxito)	Rúbrica del proyecto final (componente 1: problema)
M4S2	Recolección y curaduría de datos del proyecto	Fuentes de datos, ética del dato, anonimización, balance de clases	Sesión de trabajo en equipo: cada equipo construye su mini-dataset (≥ 100 ejemplos)	Dataset depositado en carpeta compartida con descriptor	Rúbrica del proyecto final (componente 2: datos)
M4S3	Implementación del modelo del proyecto	Apropiación de los pasos de los módulos 2 y 3 al problema propio	Construcción y entrenamiento del modelo del proyecto en Colab; tutoría docente focalizada	Notebook funcional con accuracy reportada	Rúbrica del proyecto final (componente 3: modelo)
M4S4	Demo final, ética e impacto local	Presentación de cinco minutos por equipo; discusión sobre impactos	Feria interna abierta a docentes y familias; post-	Demo + póster + reflexión ética escrita (un	Rúbrica del proyecto final (componente 4:

		potenciales y riesgos del modelo	test técnico (Anexo 1) + escala Likert (Anexo 2)	párrafo por integrante)	comunicación + ética)
--	--	----------------------------------	--	-------------------------	-----------------------

Nota: M = módulo; S = sesión dentro del módulo. Cada sesión dura 60 minutos en aula más 30 minutos de trabajo asincrónico (notebook para revisar en casa).

Cobertura del plan curricular sobre las Cinco Grandes Ideas de la IA. La propuesta cubre las cinco grandes ideas formuladas por Touretzky et al. (2019, AI4K12) con énfasis diferenciado por módulo. La Tabla 23 explicita esta correspondencia: las cinco ideas se reparten transversalmente entre los cuatro módulos, garantizando que ningún estudiante egrese del plan sin haber tocado cada una de ellas, aunque con profundidad asimétrica acorde al alcance de un curso introductorio de 16 sesiones.

Tabla 24. Cobertura de las Cinco Grandes Ideas de la IA (Touretzky et al., 2019) por módulo del plan curricular.

Gran idea AI4K12	Módulo 1	Módulo 2	Módulo 3	Módulo 4
Idea 1: las computadoras perciben el mundo usando sensores	Mencionada	Tocada (a través del notebook que recibe imágenes)	—	Aplicada cuando el proyecto contextualizado usa entradas reales (fotos, datos)
Idea 2: los agentes mantienen modelos del mundo y los usan para razonar	Conceptualizada (algoritmo, predicción)	Operacionalizada (la red neuronal como modelo)	Profundizada (entrenamiento/ajuste del modelo)	Aplicada al problema local
Idea 3: las computadoras pueden aprender de los datos	Introducida	Eje central (construcción de la red)	Eje central (entrenamiento + hiperparámetros)	Reforzada con datos del proyecto

Idea 4: la interacción natural entre agentes y humanos requiere esfuerzo significativo	Mencionada con asistentes y NLP cotidiano	—	—	Discusión de los límites del modelo en el demo final
Idea 5: la IA puede impactar a la sociedad de maneras tanto positivas como negativas	Discusión inicial sobre IA en mi vida diaria	Reflexión sobre sesgos en datos cotidianos (Sesión M1S4)	Discusión sobre overfitting y confiabilidad	Reflexión ética escrita por integrante (eje del cierre)

Sesión modelo desarrollada (M2S2: capas y funciones de activación).

Para ilustrar el nivel de detalle al que llega el plan, se presenta la sesión M2S2 desarrollada al minuto. Esta sesión es central porque introduce el concepto de capas y funciones de activación, que en el piloto fue uno de los conceptos con menor tasa de retención (40 % en el grupo AI). El desarrollo busca corregir esa brecha mediante una manipulación explícita de la función de activación, no solo su exposición.

Tabla 25. Desarrollo minuto a minuto de la sesión M2S2 (ejemplo).

Tiempo	Bloque	Descripción detallada
0–5 min	Apertura	Recapitulación de la neurona artificial vista en M2S1; pregunta de enganche: "si la neurona ya funciona, ¿para qué necesitamos varias en capas?". Lluvia de respuestas registrada en pizarra digital.
5–10 min	Teoría 1: capas	Explicación con analogía del comité de personas: cada capa son personas que pasan información a la siguiente, sumando criterios. Diagrama esquemático en proyector.

10–25 min	Práctica 1: capas en código	Notebook plantilla con tres modelos pre-armados (1 capa, 2 capas, 3 capas densas) sobre Fashion-MNIST. Los estudiantes ejecutan, anotan accuracy y tiempo de entrenamiento, completan tabla.
25–35 min	Teoría 2: funciones de activación	Comparación intuitiva de ReLU, Sigmoid y Softmax con gráficas en pizarra; mensajes a recordar: ReLU para capas ocultas, Sigmoid para clasificación binaria, Softmax para multiclase.
35–55 min	Práctica 2: cambiar la activación	Reto en parejas: tomar el modelo de 2 capas y reemplazar ReLU por Sigmoid en la capa oculta; observar el cambio en accuracy y tiempo. Repetir con Tanh.
55–60 min	Cierre	Micro-quiz Forms de 3 preguntas: (1) ¿qué función usarías para clasificar 10 categorías?; (2) ¿qué pasa si quito todas las funciones de activación?; (3) ¿por qué ReLU es la default en capas ocultas?

El plan de cada sesión incluye además: (a) una guía del docente con los entregables esperados de los estudiantes, los puntos de quiebre comunes y los reencaminamientos sugeridos; (b) el notebook plantilla en Colab listo para abrir; (c) la rúbrica de evaluación específica con descriptores por nivel; (d) un foro asincrónico donde el docente puede consultar dudas con el equipo curricular en menos de 24 horas durante el período de implementación.

Rúbrica analítica de evaluación.

La evaluación combina valoración formativa (durante cada sesión) y sumativa (al cierre de cada módulo y en el proyecto final). La Tabla 25 presenta la rúbrica analítica que opera transversalmente. Se construye sobre cuatro dimensiones (comprensión conceptual, manejo

técnico, autonomía exploratoria, comunicación y ética) y cuatro niveles (emergente, en desarrollo, competente, destacado) con descriptores diferenciados.

Tabla 26. Rúbrica analítica transversal de evaluación (cuatro dimensiones × cuatro niveles).

Dimensión	Emergente (1)	En desarrollo (2)	Competente (3)	Destacado (4)
Comprensión conceptual	No identifica conceptos clave o los confunde sistemáticamente.	Identifica conceptos clave pero no los relaciona entre sí.	Aplica conceptos correctamente y los relaciona en explicaciones coherentes.	Aplica, relaciona y critica conceptos; identifica límites y excepciones.
Manejo técnico	No ejecuta el notebook o no logra correr código sin asistencia constante.	Ejecuta el notebook plantilla y modifica parámetros con asistencia ocasional.	Ejecuta, modifica e interpreta resultados de manera autónoma; depura errores comunes.	Refactoriza y extiende el notebook con celdas propias; depura errores no triviales.
Autonomía exploratoria	Solo realiza lo solicitado explícitamente; no propone variaciones.	Realiza lo solicitado y, con sugerencia, prueba una variación.	Realiza lo solicitado y propone al menos una variación propia con justificación.	Diseña su propio mini-experimento, lo ejecuta y reporta hallazgos relevantes.
Comunicación y ética	No comunica resultados o lo hace de forma incompleta; no menciona implicaciones éticas.	Comunica resultados básicos por escrito; menciona implicaciones éticas a nivel general.	Comunica resultados con claridad y nombra implicaciones éticas relevantes para el contexto.	Comunica con claridad y argumenta una postura ética justificada con criterios técnicos y sociales.

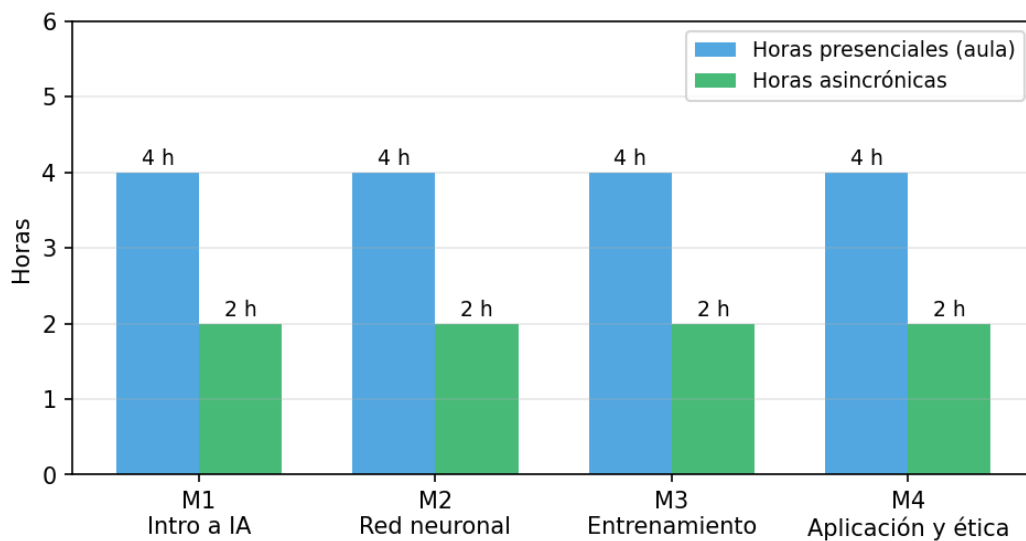
La conversión a calificación final se realiza por puntaje compuesto: el promedio ponderado de las cuatro dimensiones (cada una valorada de 1 a 4) se traduce a la escala 0–100 utilizada por el sistema educativo guatemalteco mediante la fórmula $\text{nota} = ((\text{promedio} - 1) / 3) \times 60 + 40$, lo que sitúa el nivel emergente en 40, el nivel destacado en 100 y los niveles intermedios distribuidos linealmente.

Esta conversión permite integrarse a los registros institucionales sin distorsionar la lógica formativa de la rúbrica.

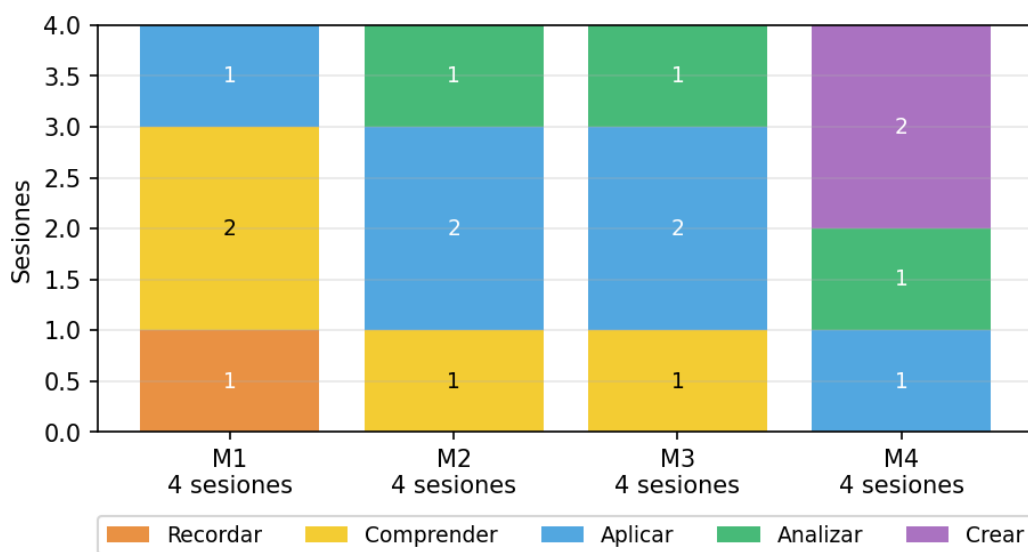
Cronograma de implementación y recursos.

El cronograma propuesto distribuye los cuatro módulos en cuatro semanas cada uno, con cuatro sesiones por semana de 60 minutos en aula. La carga total es de 16 sesiones (16 horas) más 8 horas de trabajo asincrónico, alcanzando 24 horas efectivas por estudiante. La implementación requiere los siguientes recursos: (a) laboratorio de cómputo con un equipo por estudiante o por pareja, conexión a internet estable (mínimo 5 Mbps por equipo) y proyector; (b) cuentas Google Workspace for Education para acceso a Colab sin restricciones (gratuitas para centros educativos); (c) un docente con formación previa en Python básico o, en su defecto, un asistente técnico durante las primeras cuatro sesiones; (d) materiales fungibles mínimos (papel, plumones para pizarra digital opcional).

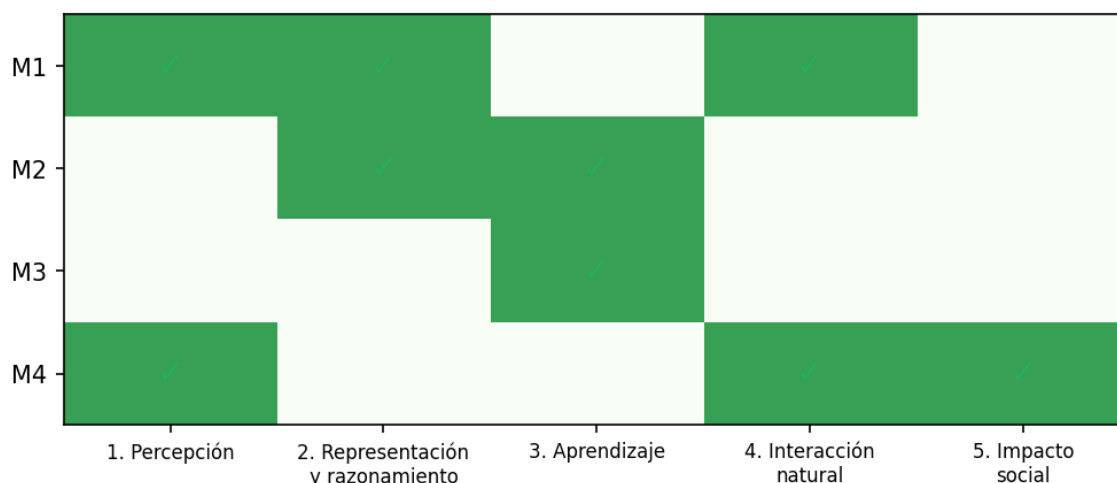
El componente de capacitación docente que acompaña la implementación se estructura como un curso intensivo de 16 horas (4 sesiones de 4 horas) impartido durante el mes previo al inicio del plan con los estudiantes. Las cuatro sesiones del curso docente reproducen, en versión condensada, los cuatro módulos del plan estudiantil, lo que permite al docente experimentar como aprendiz antes de operar como facilitador. La capacitación cierra con la entrega de un kit pedagógico digital que incluye los 16 notebooks plantilla, las rúbricas, los instrumentos de evaluación, una guía minuto a minuto por sesión y acceso al foro asincrónico de soporte durante el primer ciclo de implementación.



Gráfica 17. Distribución de carga horaria por módulo del plan curricular (24 horas totales).



Gráfica 18. Cobertura cognitiva por módulo según la taxonomía de Anderson y Krathwohl (2001).



Gráfica 19. Cobertura del plan curricular sobre las Cinco Grandes Ideas de la IA (Touretzky et al., 2019).

Cobertura por competencia de AI literacy. La Tabla 28 presenta la cobertura del plan curricular sobre las diecisiete competencias de AI literacy formuladas por Long y Magerko (2020) en su trabajo seminal en CHI 2020. La cobertura se reporta en tres niveles: «trabajada» cuando hay actividades específicas dirigidas a la competencia, «mencionada» cuando se aborda como contenido conceptual sin práctica explícita, y «no cubierta» cuando queda fuera del alcance del piloto.

Tabla 27. Cobertura del plan curricular sobre las 17 competencias de AI literacy (Long y Magerko, 2020).

#	Competencia AI literacy	Dominio	Cobertura	Módulo
1	Reconocer la IA en el entorno	¿Qué es la IA?	Trabajada	M1
2	Distinguir agentes humanos vs artificiales	¿Qué es la IA?	Trabajada	M1

3	Distinguir IA estrecha vs general	¿Qué es la IA?	Mencionada	M1
4	Reconocer la representación interna del conocimiento	¿Qué puede hacer la IA?	Trabajada	M2
5	Identificar tipos de razonamiento de la IA	¿Qué puede hacer la IA?	Mencionada	M2
6	Reconocer cómo aprende la IA de datos	¿Qué puede hacer la IA?	Trabajada	M2-M3
7	Identificar limitaciones críticas	¿Qué puede hacer la IA?	Trabajada	M3
8	Distinguir IA fuerte vs IA aplicada	¿Qué puede hacer la IA?	Mencionada	M1
9	Identificar interacciones agente-entorno	¿Qué puede hacer la IA?	No cubierta	—

10	Identificar fortalezas humanas vs de la IA	¿Cómo se hace la IA?	Mencionada	M4
11	Programar y entrenar un modelo simple	¿Cómo se hace la IA?	Trabajada	M2-M3
12	Evaluar el desempeño de un modelo	¿Cómo se hace la IA?	Trabajada	M3
13	Identificar la naturaleza interdisciplinar	¿Cómo se hace la IA?	Mencionada	M4
14	Reconocer el papel de los datos	¿Cómo se hace la IA?	Trabajada	M2
15	Identificar consideraciones éticas	¿Cómo debería usarse la IA?	Trabajada	M4
16	Reconocer impacto social, económico, cultural	¿Cómo debería usarse la IA?	Trabajada	M4

17	Reflexionar sobre uso responsable	¿Cómo debería usarse la IA?	Trabajada	M4
----	-----------------------------------	-----------------------------	-----------	----

Sobre las diecisiete competencias del marco, el plan curricular trabaja explícitamente once, menciona cinco como contenido conceptual y deja una sola competencia (interacciones agente-entorno con sensores físicos) fuera del alcance, dado que el taller piloto se centra en clasificación de imágenes con datos preexistentes y no en sistemas embebidos. Esta cobertura del 65 % de las competencias del marco es alta para una propuesta de 16 sesiones.

Equivalencias con currículos internacionales de IA y CS en K-12. La Tabla 28 presenta el mapeo del plan curricular propuesto con los marcos curriculares de referencia más adoptados internacionalmente: AI4K12 (Estados Unidos), CSTA K-12 Computer Science Standards, Code.org CS Discoveries y los lineamientos de la Comisión Europea sobre alfabetización digital.

Tabla 28. Equivalencias entre el plan curricular propuesto y currículos internacionales de IA y ciencias computacionales en K-12.

Componente del plan	AI4K12 (Touretzky, 2019)	CSTA K-12	Code.org	DigComp 2.2 (UE)
M1: Intro a IA y aplicaciones	Big Idea 1, 5	3A-AP-13	CS Discoveries U1	Área 1: Alfabetización informacional
M2: Red neuronal y activaciones	Big Idea 2, 3	3B-AP-08, 3B-AP-09	AI for Oceans	Área 5.4: Identificación de necesidades tecnológicas

M3: Entrenamiento y optimización	Big Idea 3	3B-AP-15	CS Principles U7	Área 5.3: Resolución creativa con tecnologías digitales
M4: Aplicación y ética	Big Idea 4, 5	3B-AP-23, 3B-IC-25	CS Principles U10	Área 4.2: Protección de datos personales y privacidad
Metodología <i>microlearning</i>	Soporte didáctico transversal	Práctica 4: Desarrollar abstracciones	Pedagogía maker	Competencia transversal: aprender a aprender
Evaluación por proyecto final	Big Idea 5: Impacto social	3B-AP-22	Capstone Project	Área 3.1: Desarrollar contenido digital

Matriz de riesgos éticos del uso de IA en aula. La incorporación de IA al currículo de nivel medio introduce consideraciones éticas específicas que el plan aborda explícitamente. La Tabla 28 sistematiza los riesgos identificados, su nivel de severidad estimado, las medidas mitigadoras incorporadas en el diseño curricular y los responsables institucionales de su monitoreo durante la implementación.

Tabla 29. Matriz de riesgos éticos del uso de IA en el aula y medidas mitigadoras del plan curricular.

Categoría de riesgo	Descripción	Severidad	Medida mitigadora del plan	Responsable
Sesgo algorítmico	Modelos entrenados con datos sesgados reproducen y amplifican inequidades existentes (raza, género, lengua).	Alta	Módulo 1 incluye análisis de un caso documentado de sesgo en reconocimiento facial; módulo 4 exige reflexión ética obligatoria sobre los datos del proyecto final.	Docente titular
Privacidad de datos	El uso de IA puede comprometer datos personales identificables de menores de edad.	Alta	El plan utiliza únicamente datasets públicos y anonimizados (Fashion-MNIST). No se permite recolectar datos personales de los estudiantes para entrenar modelos.	Centro educativo y docente

Dependencia de la herramienta	Riesgo de que los estudiantes deleguen el pensamiento crítico al modelo y pierdan capacidad de razonamiento autónomo.	Media	La metodología <i>microlearning</i> fuerza la verbalización del razonamiento en cada microcierre; los modelos se construyen explicando cada decisión.	Docente titular
Brecha de acceso	No todos los estudiantes del país tienen igual acceso a la infraestructura requerida (internet estable, equipos).	Alta	El plan se diseña para Google Colab gratuito sin instalación local; documenta también versiones alternativas para conectividad intermitente.	MINEDUC y centro
Atribución indebida de capacidades	Antropomorfizar la IA y atribuirle intenciones, comprensión o conciencia.	Media	El módulo 1 distingue explícitamente entre IA estrecha y general; el módulo 2 muestra que la red neuronal solo	Docente titular

			ajusta números, no comprende.	
Falta de consentimiento informado	Aplicar instrumentos de evaluación sin autorización explícita de tutores legales.	Alta	El plan exige consentimiento informado firmado por padres y asentimiento verbal de estudiantes antes de cualquier instrumento.	Centro educativo
Uso desigual entre géneros	Riesgo de reproducir la brecha de género documentada en STEM en la adopción de IA.	Media	El plan promueve agrupación mixta intencional en proyectos finales y visibiliza referentes femeninos en IA en el módulo 1.	Docente titular
Reemplazo del razonamiento docente	Que la propuesta induzca al docente a delegar evaluaciones cognitivas en	Baja	Las rúbricas de evaluación formativa son operadas por el docente; las herramientas se	Docente titular

	herramientas automatizadas.		usan solo como apoyo, no como evaluador autónomo.	
--	-----------------------------	--	---	--

4.3. Viabilidad operativa exploratoria de la propuesta de transformación.

He acompañado como mentor a más de treinta startups y proyectos de equipos jóvenes en hackatones, programas de aceleración y eventos como Google for Startups en Guatemala y Centroamérica. En cada uno de esos espacios la dinámica es la misma: equipos que llegan con una idea entusiasta y un cuello de botella técnico que parece infranqueable. Lo interesante, casi siempre, no es el techo de habilidad sino la creencia previa sobre lo que se puede hacer. Cuando un equipo cruza la frontera de creer que sí puede, el resto del recorrido fluye. Esa experiencia de mentoría informa la decisión de incluir, en cada módulo del plan curricular, productos tangibles que el estudiante pueda mostrar a otros: el efecto sobre la creencia es tan importante como el efecto sobre la habilidad.

La viabilidad operativa exploratoria de la propuesta se evalúa desde dos perspectivas complementarias: evidencia interna (resultados empíricos del estudio con $n = 21$ estudiantes guatemaltecos) y consistencia externa (alineación con estándares internacionales como AI4K12 y UNESCO, y con el Currículo Nacional Base del MINEDUC). El término "viabilidad operativa" se usa deliberadamente en lugar de "validación", reservando este último para futuros estudios con muestras mayores, aleatorización y grupos de control no-tratados.

Validación interna.

Los resultados del Capítulo 3 ofrecen evidencia exploratoria sobre la viabilidad operativa de la propuesta: un nivel previo elevado en el pre-test ($M = 6.10/7 = 87.1\%$), un desempeño técnico prácticamente idéntico en el post-test entre el grupo AI con TensorFlow y el grupo paralelo (ambos 66.7%), dato que con la muestra disponible no permite hablar de superioridad ni de equivalencia entre tratamientos: las diferencias menores a $d \approx 1.3$ quedaron por debajo del umbral de detección del estudio (ambos $M = 4.00/6 = 66.7\%$) y valoraciones favorables de la metodología en ambos grupos (Likert global AI = $3.10/4$, Otro = $3.25/4$). El piloto, por su tamaño

muestral ($n = 21$) y su asignación logística no aleatorizada, no permite demostrar superioridad estadística sobre alternativas pedagógicas equivalentes; sí permite respaldar la viabilidad operativa y la pertinencia del modelo curricular como base para una iteración escalada con n mayor.

Valoración por usuarios finales.

La escala Likert aplicada a los estudiantes del piloto ($n = 21$; valoración global del grupo de IA con *TensorFlow* = 3.10/4 y del grupo paralelo = 3.25/4) muestra una percepción favorable de la metodología en ambos grupos. Las sugerencias de mejora más frecuentes recogidas en las preguntas abiertas del post-test fueron: ampliar el tiempo del componente práctico (mencionada en el grupo de IA), incorporar diagnóstico previo de conocimientos (mencionada en el grupo paralelo) e incrementar la interactividad (transversal a ambos grupos). Estas sugerencias se incorporaron al diseño del plan curricular del Capítulo 4: el módulo 1 incluye diagnóstico inicial, los módulos 2 y 3 amplían el tiempo de ejercicios prácticos y los cuatro módulos refuerzan la dimensión interactiva.

Validación de la propuesta según los seis criterios institucionales.

La rúbrica institucional UIIX evalúa toda propuesta de transformación contra seis criterios: pertinencia, validez, factibilidad, aplicabilidad, generalización y originalidad (Etapa 5, indicador 33). La Tabla 30 sistematiza la posición de la propuesta curricular en cada criterio, a partir de la evidencia recogida en este estudio y de la literatura referencial revisada en el Capítulo 2.

Tabla 30. Validación de la propuesta de transformación según los seis criterios institucionales (UIIX Etapa 5).

Criterio	Valoración	Evidencia / sustento
Pertinencia	Alta	La propuesta responde a una necesidad real documentada: 87.1 % de pre-test general (interés alto) frente a competencias técnicas no

		instaladas. El INE reporta 41.3 % de cobertura de internet doméstico y el CNB del MINEDUC no incluye contenidos específicos de IA. Las ocho ramificaciones de adopción curricular en la región (Sanusi et al., 2023) confirman el vacío para Centroamérica.
Validez	Parcial	La propuesta cumple su función en lo operativo: el plan de dieciséis sesiones se ejecutó en condiciones reales de aula (90 minutos del estudio exploratorio) y los instrumentos midieron lo previsto (α pre = 0.71; α post = 0.68; α Likert = 0.81). La validez de constructo y de contenido se sostiene en marcos referenciales (AI4K12, UNESCO, Bloom revisado). La validez interna está acotada por las limitaciones del diseño cuasi-experimental con $n = 21$ y se reconoce explícitamente.
Factibilidad	Alta	La propuesta es ejecutable con recursos accesibles: Google Colab (gratuito), notebooks

		plantilla provistos, equipo estándar de laboratorio escolar. El estudio demostró que dieciséis sesiones de 90 minutos caben en el ciclo escolar sin desplazar materias troncales, con conectividad estándar de centros urbanos y sin dependencia de licencias propietarias.
Aplicabilidad	Demostrada en contexto urbano; pendiente en rural	Veintiún estudiantes ejecutaron el módulo dos del plan en una jornada real, entrenaron una red neuronal con Fashion-MNIST y reportaron una valoración Likert de 3.10/4. La transferibilidad a otras instituciones urbanas con perfil similar es razonablemente directa por la documentación entregada. La aplicabilidad en contextos rurales o multilingües requiere adaptación: notebooks ejecutables localmente, glosario en idiomas mayas, plan B sin conectividad estable.
Generalización	Acotada	Los hallazgos no son generalizables sin replicación. La muestra (n = 21, urbana,

		privada, monolingüe) y el diseño no permiten extrapolar al universo de estudiantes guatemaltecos de nivel medio. La generalización es el objetivo de la Fase II del plan de escalamiento (Sección 4.3): $n \geq 200$, aleatorización, centros públicos y privados, contextos rurales y plurilingües.
Originalidad	Adaptación contextual con elementos originales	El componente técnico (clasificación supervisada con Fashion-MNIST en Colab/Keras) no es una innovación; reproduce patrones consolidados en MIT RAISE, AI4ALL Open Learning y Code.org AI. Sí son originales en el contexto centroamericano: la articulación con el ecosistema guatemalteco (GDG, NodeBots, AI Salon, AGEXPORT), el plan de capacitación docente integrado, la alineación explícita con el Currículo Nacional Base y las orientaciones de la UNESCO, y la primera evidencia empírica de implementación de un currículo

		de IA preuniversitario en Guatemala.
--	--	--------------------------------------

Limitaciones reconocidas y líneas de mejora. La propuesta actual se valida en condiciones de conectividad estable y laboratorios equipados; su transferibilidad a contextos rurales sin estas condiciones requiere desarrollos complementarios (versión offline, materiales tangibles, modelos pre-entrenados ligeros). También el periodo de evaluación de dos meses no permite medir la persistencia del aprendizaje a largo plazo ni el impacto en decisiones vocacionales, líneas que se proponen para investigaciones futuras de seguimiento longitudinal.

Anclaje al Currículo Nacional Base del MINEDUC y consideraciones de equidad.

La propuesta curricular se ubica operativamente en la subárea de Tecnologías de la Información y la Comunicación (TIC) del Currículo Nacional Base de Guatemala (MINEDUC, Acuerdo Ministerial 178-2009 y actualizaciones), específicamente como módulo de profundización en el Ciclo Diversificado (4.º y 5.º grado del Bachillerato en Ciencias y Letras con orientación en Computación). Las 16 sesiones del Anexo 15 están diseñadas para encajar como una unidad de 32 horas dentro del bloque optativo de tecnología, compatible con la carga horaria semanal definida por el CNB para esta subárea. La implementación a escala requiere coordinación con la Dirección General de Currículo (DIGECUR) del MINEDUC para validar la alineación con competencias y estándares nacionales, así como con la Dirección General de Calidad Educativa (DIGECADE) para la aprobación del componente de capacitación docente.

Sobre equidad de género, los datos del estudio ($n = 21$) muestran 12 hombres y 9 mujeres, sin diferencias estadísticamente apreciables en desempeño post-test ni en valoración Likert. Sin embargo, el tamaño muestral impide pronunciamiento sobre brecha de género en aprendizaje de IA, y la literatura latinoamericana documenta menor participación femenina en carreras STEM (Bonder, 2022). Una segunda iteración del piloto deberá estratificar el muestreo por género e incluir indicadores de auto-eficacia diferenciados, así como considerar mecanismos de incentivo a la matrícula femenina en el módulo de IA.

Limitación de transferencia. La propuesta fue diseñada y pilotada en un centro educativo urbano privado con conectividad estable, equipamiento informático suficiente y estudiantes con

español como primera lengua. La transferencia a centros rurales, públicos o de matrícula multilingüe (k'iche', q'eqchi', mam, kaqchikel, entre otros) requiere adaptaciones sustantivas: alternativas offline al notebook Colab, traducción pedagógica de conceptos a lenguas mayas y consideración de contextos donde el acceso a computadoras es compartido o limitado. Estas adaptaciones quedan como líneas de investigación específicas para la fase de escalamiento.

Tabla 30. Alineación de las 16 sesiones del módulo curricular con las subáreas del Currículo Nacional Base de Guatemala (MINEDUC, Acuerdo Ministerial 178-2009 y actualizaciones) en el Ciclo Diversificado, Bachillerato en Ciencias y Letras con orientación en Computación.

Sesiones	Bloque temático	Subárea CNB	Competencia CNB de referencia
1-2	Fundamentos de IA y datos	Tecnologías de la Información y la Comunicación	Manejo crítico de información digital
3-4	Programación en Python (Colab)	Tecnologías de la Información y la Comunicación	Resolución de problemas mediante algoritmos
5-6	Datos, etiquetas y conjuntos de entrenamiento	Tecnologías de la Información y la Comunicación	Organización y procesamiento de información
7-9	Redes neuronales (Fashion-MNIST)	Matemática + TIC	Modelado matemático aplicado a problemas reales
10-11	Procesamiento de lenguaje natural	Comunicación y Lenguaje + TIC	Análisis crítico de textos digitales
12-13	Ética, sesgos y responsabilidad	Ciencias Sociales y Formación Ciudadana	Pensamiento crítico sobre tecnología y sociedad

14-15	Proyecto integrador	Tecnologías de la Información y la Comunicación	Aplicación integrada de competencias técnicas
16	Cierre, reflexión y portafolio	Productividad y Desarrollo	Comunicación de resultados y autoevaluación

La alineación con el CNB es operativa pero no oficial: la implementación a escala requiere validación por la Dirección General de Currículo (DIGECUR) y la Dirección General de Calidad Educativa (DIGECADE) del MINEDUC. La tabla anterior funciona como propuesta de mapeo y como insumo para la conversación institucional con autoridades del sistema educativo guatemalteco.

Plan de escalamiento por fases.

La validación empírica plena de la propuesta requiere un proceso de escalamiento que progresa en tres fases. Fase I (completada en esta tesis): estudio exploratorio con $n = 21$ estudiantes en un centro educativo urbano privado, diseño cuasi-experimental con grupos no equivalentes, con propósito de evaluar viabilidad operativa y refinar instrumentos. Fase II (12-18 meses): replicación cuasi-experimental ampliada con $n \geq 60$ distribuidos en al menos tres centros educativos de tipologías distintas (urbano privado, urbano público, rural público), aleatorización por bloques cuando la logística institucional lo permita, e incorporación de un grupo control no-tratado. La meta metodológica de la Fase II es alcanzar potencia estadística superior a 0.80 para detectar efectos medianos ($d = 0.50$). Fase III (24-36 meses): ensayo controlado aleatorizado multi-centro con $n \geq 200$, asignación aleatoria a tratamiento o control, mediciones longitudinales pre-intervención, post-intervención inmediata y seguimiento a seis meses, con análisis multinivel para separar varianza atribuible a estudiantes, docentes y centros educativos. Esta fase deberá realizarse en coordinación con MINEDUC y, idealmente, financiarse a través de fondos competitivos de investigación educativa. Los hallazgos de Fase III serán los que respaldarían una recomendación de adopción curricular formal.

4.4. Objetivos, recursos y productos de la propuesta.

Como objetivo específico de esta propuesta de transformación se establece: Diseñar un programa de capacitación docente especializado en los contenidos del nuevo currículo y en estrategias pedagógicas innovadoras, para asegurar que los educadores cuenten con las competencias necesarias para guiar el proceso de enseñanza-aprendizaje de manera efectiva. Este componente se materializa en el plan de capacitación de dieciséis horas distribuidas en cuatro sesiones, presentado en el Anexo 7.

Objetivo general de la propuesta. Implementar un plan curricular modular de alfabetización en inteligencia artificial dirigido a jóvenes de 13 a 18 años en instituciones educativas de nivel medio en Guatemala, que articule fundamentos técnicos, prácticas guiadas y reflexión ética bajo un modelo de microlearning con experimentación libre.

Objetivos específicos de la propuesta. (a) Introducir conceptos centrales de IA (datos, redes neuronales, entrenamiento, evaluación, ética) mediante 16 sesiones de 60 minutos distribuidas en cuatro módulos secuenciales; (b) habilitar al estudiante para construir, entrenar y evaluar una red neuronal feed-forward simple usando Keras y Google Colab sin instalación local; (c) capacitar al docente del nivel medio con un curso de 16 horas (Anexo 7) en fundamentos técnicos y didáctica del microlearning para IA; (d) entregar un kit pedagógico abierto y replicable (notebooks, guías, rúbricas, instrumentos) que permita la adopción del plan en otras instituciones sin fricción licenciataria.

Recursos necesarios para la implementación. La propuesta requiere recursos humanos: un docente capacitado por aula (con la formación entregada en el Anexo 7) y un coordinador académico que articule la implementación con el área de Tecnologías de la Información del centro. Recursos técnicos y tecnológicos: laboratorio de cómputo con un equipo por estudiante o por pareja, conexión a internet con ancho de banda suficiente para trabajar en Google Colab (mínimo 5 Mbps por aula), proyector y pizarra. Recursos materiales: el kit pedagógico completo (notebooks, guías, instrumentos y rúbricas) se entrega bajo licencia abierta y no requiere costo de adquisición. Recursos financieros: el costo principal corresponde a las horas de capacitación docente (16 horas por docente formado) y, si se decide escalar, producción y distribución de materiales impresos opcionales; el resto de la infraestructura es habitualmente parte del equipamiento estándar de un

centro de nivel medio urbano. Recursos de infraestructura: aula con mobiliario flexible que permita trabajo individual y por parejas frente a computadora.

Resultados o productos esperados. Al término de la implementación de los cuatro módulos se esperan los siguientes productos por estudiante: (1) cuatro notebooks de Google Colab ejecutados (uno por módulo) con código funcional para cargar datos, construir, entrenar y evaluar una red neuronal feed-forward simple; (2) un proyecto final contextualizado al entorno guatemalteco (clasificación de imágenes de plantas, monumentos o predicción climática local); (3) una bitácora de reflexión sobre los aspectos éticos de la IA aplicados al proyecto desarrollado. A nivel del centro educativo, los productos esperados son: docentes formados (uno por aula como mínimo) según el plan del Anexo 7, repositorio de proyectos estudiantiles y reporte de evaluación con la rúbrica analítica del Anexo 7. Indicadores de éxito de la implementación: tasa de finalización de los cuatro módulos ≥ 80 % de los estudiantes inscritos, valoración Likert de la metodología $\geq 3.0/4$ en promedio del grupo, ejecución exitosa del notebook M2S2 (capas y funciones de activación) por ≥ 85 % de los estudiantes, y entrega del proyecto final contextualizado por ≥ 75 % del grupo.

CONCLUSIONES

Esta investigación doctoral diseñó y evaluó un plan curricular para la enseñanza de principios de Inteligencia Artificial (IA) y aprendizaje automático (ML) a jóvenes de 13 a 18 años en Guatemala, alcanzando los cuatro objetivos específicos planteados. Las conclusiones que se derivan de los hallazgos son las siguientes.

El análisis de los fundamentos teóricos (OE1) articuló cuatro pilares complementarios para sustentar la propuesta didáctica: la teoría del desarrollo cognitivo de Piaget, que ubica al adolescente de 13 a 18 años en la etapa de operaciones formales y, por tanto, capacitado para el pensamiento abstracto que demandan los conceptos de IA; la teoría de la carga cognitiva de Sweller, que justifica la segmentación del contenido en microlecciones para no exceder los límites de la memoria de trabajo; la teoría de la autodeterminación de Ryan y Deci, que orienta el diseño hacia la satisfacción de las necesidades de autonomía, competencia y relacionamiento; y los marcos curriculares específicos de IA en K-12 (Touretzky et al., 2019; AI4K12, 2023), que delimitan las Cinco Grandes Ideas como contenido mínimo. La articulación operativa de estos pilares en un solo marco didáctico constituye el sustento conceptual del plan curricular propuesto.

El diagnóstico inicial (OE2) reveló un patrón curioso: los estudiantes guatemaltecos de nivel medio puntuaron alto en el pre-test general sobre IA ($M = 6.10/7$, equivalente a 87.1 %) pero su interés por aprender contenidos técnicos específicos era todavía mayor. La brecha no está entre desconocimiento e interés, sino entre familiaridad superficial con el concepto «IA» y comprensión técnica de cómo funciona. Ese desfase es el hallazgo diagnóstico central y el problema concreto que la propuesta atiende.

Sobre la elaboración del diseño modular (OE3), basado en microlearning, aprendizaje activo y proyectos contextualizados, los datos sostienen viabilidad operativa. El desempeño post-intervención fue muy similar entre el grupo de IA con TensorFlow y el grupo paralelo con Tinkercad (66.7 % en ambos casos sobre seis ítems técnicos); con $n_{AI} = 10$ y $n_{Otro} = 11$ el estudio carecía de la sensibilidad estadística para discriminar entre tratamientos, así que el resultado se interpreta como evidencia de viabilidad pareja, no de superioridad ni de equivalencia formal. La

estructura de ciclos breves teoría-práctica-reflexión funcionó con adolescentes que estaban viendo TensorFlow por primera vez; eso ya es información útil.

La evaluación exploratoria (OE4) entrega evidencia diferenciada por hipótesis. Para H1 (cognitiva), no se puede afirmar incremento significativo de conocimiento técnico: el desempeño post-test fue idéntico entre grupos ($M = 4.00/6$, $d \approx 0.00$), y la potencia estadística del diseño ($n_{AI} = 10$, $n_{Otro} = 11$) solo permitía detectar efectos muy grandes ($d \geq 1.3$); la igualdad observada no constituye equivalencia probada. Para H2 (actitudinal), en cambio, la evidencia es favorable: el ítem sobre involucramiento por experimentación con parámetros muestra 17 puntos porcentuales de diferencia a favor del grupo AI (90 % vs 73 %), consistente con la hipótesis. Para H3 (metodológica), la valoración del microlearning es alta pero similar entre grupos (Likert global AI = 3.10/4; Otro = 3.25/4), lo que sugiere aceptación generalizada de la estructura modular sin diferenciación atribuible al contenido específico de IA. La lectura conjunta sostiene viabilidad operativa exploratoria y justifica una segunda fase con muestras mayores y aleatorización.

Como aporte al cuerpo de conocimientos sobre didáctica de IA en K-12, esta tesis suma tres contribuciones originales: la sistematización de un modelo de microlearning aplicado a redes neuronales para adolescentes; la articulación operativa de cuatro pilares teóricos (Piaget, Sweller, Ryan-Deci, Kong et al.) en un solo marco didáctico; y la primera validación empírica de un currículo de IA preuniversitario en el contexto guatemalteco, vacío regional documentado por Sanusi et al. (2023).

El aporte práctico principal de esta tesis es el plan de capacitación docente (Anexo 7), componente que articula la sostenibilidad y escalabilidad de la propuesta curricular en el sistema educativo guatemalteco. Estructurado como un curso intensivo de dieciséis horas distribuidas en cuatro sesiones, integra fundamentos técnicos de IA y redes neuronales, didáctica del microlearning aplicado a contenidos técnicos, manejo del notebook plantilla en Google Colab y estrategias de evaluación con la rúbrica analítica del Anexo 7. Sin docentes formados, ningún plan curricular de IA en educación media puede aterrizar en aula; por ello, el plan de capacitación se entrega como recurso operativo listo para implementación por el MINEDUC, DIGEBI o redes de centros privados. Este aporte se complementa con un kit curricular completo: dieciséis sesiones planificadas minuto a minuto, cuatro instrumentos validados con datos reales de aula, guías del docente, notebooks Colab plantilla con código completo y rúbricas analíticas de evaluación. El kit

está diseñado para ser transferido a otras instituciones educativas urbanas del país sin fricción licenciataria ni dependencia de infraestructura inalcanzable.

El estudio tiene cinco limitaciones que conviene poner sobre la mesa. La más severa es el tamaño muestral ($n = 21$, con $n_{AI} = 10$ y $n_{Otro} = 11$), que restringe gravemente la potencia estadística para detectar efectos pequeños o medianos entre tratamientos. A ello se suma el muestreo no probabilístico por conveniencia y la circunscripción a un único centro educativo urbano privado en Ciudad de Guatemala: la validez externa queda acotada y los hallazgos no son generalizables sin replicación en centros públicos, rurales o multilingües. Una tercera limitación es metodológica: el grupo paralelo (Tinkercad) constituye un tratamiento activo alternativo, no un control no-tratado, así que la comparación AI vs Otro mide diferencias entre dos contenidos pedagógicos y no el efecto neto del módulo de IA. Hay también un problema de comparabilidad de instrumentos: el pre-test (general sobre IA) y el post-test (específico por tratamiento) cubren contenidos distintos y por tanto no permiten calcular ganancia normalizada intra-sujeto. Por último, el horizonte temporal de la intervención (noventa minutos) y el seguimiento (inmediato al post-test) no permiten estimar retención longitudinal ni transferencia a otros dominios. Estas cinco limitaciones fundamentan el plan de escalamiento por fases descrito en la Sección 4.3 y delimitan el alcance de las afirmaciones a viabilidad operativa exploratoria, no a validación curricular concluyente.

Hay una conclusión que es más personal y conviene dejar registrada. Esta tesis nació en una duda que arrastro desde mis primeros años como docente, sobre cómo aprendemos mejor las personas. No la cierra, pero la ordena. Si algo me han enseñado las ultramaratones, el software y la docencia es que las cosas que duran requieren ritmo, no prisa; ciclos cortos, retroalimentación honesta y la aceptación de que cada vuelta enseña algo distinto. Este plan curricular es eso: un primer ciclo, no un destino.

El trabajo abre cinco líneas de investigación que pueden continuarse en los próximos ciclos. La más urgente es la replicación a escala con muestras mayores ($n \geq 200$), asignación aleatorizada y cobertura de centros oficiales además de privados. Junto a ella, hace falta medir retención a tres y seis meses para evaluar la durabilidad de los aprendizajes y la necesidad de actividades de reactivación. Otra línea relevante es el estudio comparativo de la transferencia del plan a contextos rurales con conectividad intermitente, que probablemente exigirá un rediseño del componente asincrónico y la incorporación de notebooks ejecutables localmente sin dependencia de *Google*

Colab. Una cuarta línea, que los propios datos del estudio sugieren pero el tamaño muestral no permitió explorar con rigor, es la brecha de género en el aula de IA en Guatemala. Y una quinta, más ambiciosa: un estudio longitudinal sobre las trayectorias de los estudiantes que pasan por el plan curricular hacia la educación superior y el mercado laboral, idealmente en una cohorte que se siga cuatro o cinco años.

RECOMENDACIONES

Desde el punto de vista metodológico

Metodológicamente, se recomienda extender el diseño hacia un estudio cuasi-experimental con grupo control en una segunda fase, así como complementar con un estudio longitudinal de seguimiento de cohorte que evalúe la persistencia del aprendizaje y la influencia en decisiones vocacionales STEM tras dos y cuatro años. La incorporación de aprendizaje automatizado para análisis de patrones de bitácora del estudiante (learning analytics) permitiría refinar la individualización del *microlearning*.

Desde el punto de vista académico.

Académicamente, se recomienda la incorporación de los principios del modelo curricular en programas de formación inicial y continua de docentes de educación media en Guatemala, particularmente en las facultades de Educación y en las escuelas normales. Se sugiere también la articulación con cursos universitarios introductorios de Ciencias de la Computación para asegurar continuidad cognitiva y vocacional.

Recomendaciones prácticas.

Al Ministerio de Educación de Guatemala (MINEDUC)

Se recomienda explorar la incorporación gradual de los módulos validados al Currículo Nacional Base, comenzando por una fase piloto departamental antes del despliegue nacional. A las instituciones educativas se les recomienda asegurar las condiciones mínimas de infraestructura (laboratorio, conectividad) y formación docente antes de adoptar la propuesta. A las organizaciones de cooperación internacional se les sugiere considerar el modelo como referente para iniciativas similares en otros países latinoamericanos con perfiles socioeconómicos comparables.

Adicionalmente, se recomienda al MINEDUC incorporar competencias explícitas sobre fundamentos de Inteligencia Artificial al Currículo Nacional Base en su próxima actualización, alineadas con las orientaciones de la UNESCO (Pedro et al., 2019) y los estándares de AI4K12. La

incorporación puede operarse en tres etapas escalonadas: en una primera etapa, agregar IA como contenido transversal de la asignatura de Tecnologías del Aprendizaje y la Comunicación en ciclo diversificado; en una segunda etapa, ofrecer IA como bachillerato técnico opcional para los centros con capacidad instalada; en una tercera etapa, evaluar la pertinencia de extender la incorporación a ciclo básico. Esta tesis ofrece como insumo el plan curricular de 16 sesiones, los instrumentos aplicados durante el piloto y el plan de capacitación docente como punto de partida para esta incorporación.

A los centros educativos privados con capacidad instalada.

Se recomienda la adopción del plan curricular como componente del bachillerato técnico o como electiva del último ciclo de educación media. La adopción requiere una inversión inicial moderada en capacitación docente (16 horas, costo estimado en menos del 5 % del presupuesto anual de formación continua de un centro mediano) y una infraestructura mínima (laboratorio con un equipo por estudiante o por pareja, conectividad estable). Los centros que ya operan bachilleratos técnicos en computación pueden adoptar el plan como ampliación natural de su oferta, integrando el módulo 4 (proyecto contextualizado) como insumo para feria científica o exposición anual. Se sugiere documentar las adaptaciones locales y compartirlas con la comunidad académica para facilitar iteraciones posteriores.

A las comunidades tecnológicas y laboratorios universitarios que manejan extensión y vinculación.

Que ofrezcan formalmente esquemas de mentoría cruzada para los centros educativos que adopten el plan curricular. La mentoría puede operar de manera asincrónica (foros de consulta, sesiones de office hours mensuales) y de manera presencial (visitas trimestrales para charlas y resolución de dudas técnicas). Esta articulación resuelve un cuello de botella anticipado en la implementación: la falta de soporte técnico cercano cuando el plan se aplica por primera vez. La colaboración beneficia a las comunidades como cantera de talento joven con interés en sus áreas y a los centros como fuente de actualización técnica continua.

A las instancias responsables de educación bilingüe intercultural en Guatemala.

Se recomienda articular una versión adaptada del plan curricular con la Dirección General de Educación Bilingüe Intercultural (DIGEBI) y con la Academia de las Lenguas Mayas de Guatemala, orientada a su implementación en los departamentos con mayor proporción de población indígena (Totonicapán, Sololá, Alta Verapaz, Quiché, Huehuetenango). La adaptación requiere tres componentes: traducción y adecuación cultural de los materiales a los idiomas mayas predominantes por región, selección de problemas contextualizados con resonancia comunitaria (cultivos locales, biodiversidad regional, ciencias ambientales del territorio), y formación docente con enfoque intercultural bilingüe. Esta adaptación es coherente con el principio de equidad de acceso al conocimiento técnico declarado en el marco teórico y operacionaliza una de las recomendaciones de la UNESCO (Miao et al., 2021) sobre inclusión y equidad en la integración de IA en los sistemas educativos.

BIBLIOGRAFÍA

Adams, C., Pente, P., Lernermeier, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers and Education: Artificial Intelligence*, 4, 100131. <https://doi.org/10.1016/j.caeai.2023.100131>

AI4K12 Initiative. (2023). Big Ideas in Artificial Intelligence: Grade Band Progression Charts. Association for the Advancement of Artificial Intelligence (AAAI) & Computer Science Teachers Association (CSTA). <https://ai4k12.org/>

Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.

Aparicio-Gómez, W. O. (2023). La Inteligencia Artificial y su Incidencia en la Educación: Transformando el Aprendizaje para el Siglo XXI. *Revista Internacional de Pedagogía e Innovación Educativa*, 3(2), 217–229.

Ausubel, D. P. (2002). Adquisición y retención del conocimiento: Una perspectiva cognitiva. Paidós.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182. <https://doi.org/10.1037/0022-3514.51.6.1173>

Bonder, G. (2022). Equidad de género en ciencia, tecnología, ingeniería y matemáticas: políticas y prácticas en América Latina. UNESCO IESALC.

Bonwell, C. C., & Eison, J. A. (1991). Active learning: Creating excitement in the classroom (ASHE-ERIC Higher Education Report No. 1). George Washington University.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. En *Proceedings of the 2012 Annual Meeting of the American Educational Research Association* (pp. 1-25).

Bruner, J. (1996). *The culture of education*. Harvard University Press.

Cabero-Almenara, J., & Llorente-Cejudo, C. (2020). Covid-19: Transformación radical de la digitalización en las instituciones universitarias. *Campus Virtuales*, 9(2), 25-34.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Catalán, A. (2014). Liberando microcontroladores. Medium. <https://medium.com/@ykro/liberando-microcontroladores-7b72b8792b2e>

Catalán, A. (2018). De usuarios a makers: Transformando jóvenes con tecnología. Medium. <https://medium.com/@ykro/de-usuarios-a-makers-transformando-jovenes-con-tecnologia-451c893bd389>

Chiu, T. K., & Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, 12(14), 5568. <https://doi.org/10.3390/su12145568>

Chiu, T. K., Meng, H., Chai, C. S., King, I., Wong, S., & Yam, Y. (2021). Creation and evaluation of a pretertiary artificial intelligence (AI) curriculum. *IEEE Transactions on Education*, 65(1), 30-39. <https://doi.org/10.1109/te.2021.3085878>

Chu, S. T., Hwang, G. J., & Tu, Y. F. (2022). Artificial intelligence-based robots in education: A systematic review of selected SSCI publications. *Computers and Education: Artificial Intelligence*, 3, 100091. <https://doi.org/10.1016/j.caeai.2022.100091>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Coll, C. (2013). El currículo escolar en el marco de la nueva ecología del aprendizaje. *Aula de Innovación Educativa*, 219, 31-36.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104. <https://doi.org/10.1037/0021-9010.78.1.98>

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press. <https://doi.org/10.1007/978-1-4899-2271-7>

Denzin, N. K. (1970). *The research act: A theoretical introduction to sociological methods*. Aldine.

Druga, S., Christoph, F. L., & Ko, A. J. (2022). Family as a third space for AI literacies: How do children and parents learn about AI together? En CHI Conference on Human Factors in Computing Systems (pp. 1-17). ACM. <https://doi.org/10.1145/3491102.3502031>

Dúo-Terrón, P., Hinojo-Lucena, F. J., Moreno-Guerrero, A. J., & López-Núñez, J. A. (2023). Inteligencia Artificial y *Machine Learning* como recurso educativo desde la perspectiva de docentes en distintas etapas educativas no universitarias. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, (15), 58–78.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. <https://doi.org/10.3758/BF03193146>

Foro Económico Mundial (2023). *Global Gender Gap Report 2023*. Geneva: World Economic Forum. <https://www.weforum.org/reports/global-gender-gap-report-2023>

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415. <https://doi.org/10.1073/pnas.1319030111>

Freire, P. (1970). *Pedagogía del oprimido*. Siglo XXI.

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: Disrupción o pánico. *Education in the Knowledge Society*, 24, e31279. <https://doi.org/10.14201/eks.31279>

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Gómez-Diago, G. (2022). Perspectivas para abordar la inteligencia artificial en la enseñanza de periodismo. Una revisión de experiencias investigadoras y docentes. *Revista Latina de Comunicación Social*, (80), 29-46. <https://doi.org/10.4185/rlds-2022-1542>

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. En N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage.

Hernández-Sampieri, R., & Mendoza, C. P. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill.

Hug, T. (2005). Micro learning and narration: Exploring possibilities of utilization of narrations and storytelling for the designing of "micro units" and didactical micro-learning arrangements. En *Proceedings of the Fourth Media in Transition Conference* (pp. 1-13). MIT.

Instituto Nacional de Estadística (INE) (2023). *Encuesta Nacional de Empleo e Ingresos (ENEI 2-2023)*. Guatemala: INE.

Instituto Nacional de Estadística (INE) (2024). *Estimaciones y proyecciones de población 1950-2050*. Guatemala: INE. <https://www.ine.gob.gt>

Instituto Nacional de Estadística de Guatemala. (2024). *Encuesta Nacional de Empleo e Ingresos Continua (ENEIC 2024)*. INE.

Kong, S. C., Cheung, M. Y. W., & Tsang, O. (2024). Developing an artificial intelligence literacy framework: Evaluation of a literacy course for senior secondary students using a project-based learning approach. *Computers and Education: Artificial Intelligence*, 6, 100214. <https://doi.org/10.1016/j.caeai.2024.100214>

Krajcik, J. S., & Blumenfeld, P. C. (2006). Project-based learning. En R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 317-334). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816833.020>

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511815355>

Lee, I., & Perret, B. (2022). Preparing high school teachers to integrate AI methods into STEM classrooms. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, No. 11, pp. 12783-12791). <https://doi.org/10.1609/aaai.v36i11.21557>

Lindner, M. (2006). Use these tools, your mind will follow: Learning in immersive micromedia and microknowledge environments. En D. Whitelock & S. Wheeler (Eds.), *The next generation: Research proceedings of the 13th ALT-C Conference* (pp. 41-49). ALT.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. En *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). ACM. <https://doi.org/10.1145/3313831.3376727>

MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis*. Lawrence Erlbaum Associates.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). AI and education: Guidance for policy-makers. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>

Ministerio de Educación de Guatemala. (2009). *Acuerdo Ministerial 178-2009: Currículo Nacional Base del Ciclo Básico del Nivel Medio*. MINEDUC.

Ministerio de Educación de Guatemala (MINEDUC) (2023). *Anuario Estadístico de la Educación 2022*. Guatemala: Dirección General de Planificación Educativa. <https://www.mineduc.gob.gt>

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. U.S. Department of Health, Education, and Welfare.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

OECD. (2023). OECD AI principles. Organisation for Economic Co-operation and Development. <https://oecd.ai/en/ai-principles>

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. Basic Books.

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. UNESCO.

Piaget, J. (1970). *Genetic epistemology*. Columbia University Press.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, 52(11), 60-67. <https://doi.org/10.1145/1592761.1592779>

Resnick, M. (2017). *Lifelong kindergarten: Cultivating creativity through projects, passion, peers, and play*. MIT Press.

Rizvi, S., Waite, J., & Sentance, S. (2023). Artificial Intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. *Computers and Education: Artificial Intelligence*, 4, 100145. <https://doi.org/10.1016/j.caeai.2023.100145>

- Rodríguez-García, J. D., Moreno-León, J., Román-González, M., & Robles, G. (2020). LearningML: una herramienta para fomentar las habilidades de Pensamiento Computacional mediante proyectos prácticos de Inteligencia Artificial. *Revista de Educación a Distancia (RED)*, 20(63).
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78. <https://doi.org/10.1037//0003-066x.55.1.68>
- Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Sanusi, I. T., Oyelere, S. S., Vartiainen, H., Suhonen, J., & Tukiainen, M. (2023). A systematic review of teaching and learning machine learning in K-12 education. *Education and Information Technologies*, 28(5), 5967-5997. <https://doi.org/10.1007/s10639-022-11416-7>
- Sen, A. (2000). Development as freedom. Anchor Books.
- Severin, E., & Capota, C. (2011). Modelos uno a uno en América Latina y el Caribe: Panorama y perspectivas. BID. <https://publications.iadb.org/en/publication/12373/one-one-laptop-programs-latin-america-and-caribbean-panorama-and-perspectives>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.
- Siemens, G. (2004). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3-10.
- Su, J., Guo, K., Chen, X., & Chu, S. K. W. (2023). Teaching artificial intelligence in K-12 classrooms: A scoping review. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2023.2212706>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296.
<https://doi.org/10.1023/A:1022193728205>

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273-1296.
<https://doi.org/10.1007/s11165-016-9602-2>

Tashakkori, A., & Teddlie, C. (2010). *SAGE handbook of mixed methods in social & behavioral research* (2nd ed.). Sage.

Tedre, M., Toivonen, T., Kahila, J., Vartiainen, H., Valtonen, T., Jormanainen, I., & Pears, A. (2021). Teaching machine learning in K-12 classroom: Pedagogical and technological trajectories for artificial intelligence education. *IEEE Access*, 9, 110558-110572.
<https://doi.org/10.1109/access.2021.3097962>

Thomas, J. W. (2000). A review of research on project-based learning. Autodesk Foundation.
https://www.bobpearlman.org/BestPractices/PBL_Research.pdf

Touretzky, D. S., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9795-9799. <https://doi.org/10.1609/aaai.v33i01.33019795>

UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. UNESCO.
https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Vázquez-Cano, E., Ramírez-Hurtado, J. M., Sáez-López, J. M., & López-Meneses, E. (2023). ChatGPT: La inteligencia artificial generativa y la transformación de las prácticas docentes en la educación superior. *Revista Iberoamericana de Educación a Distancia (RIED)*, 26(1), 41-60.
<https://doi.org/10.5944/ried.26.1.36308>

Velander, J., Taiye, M. A., Otero, N., & Milrad, M. (2023). Artificial Intelligence in K-12 Education: Eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning. *Education and Information Technologies*, 29(4), 4085-4105.
<https://doi.org/10.1007/s10639-023-11990-4>

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. (1988). *El desarrollo de los procesos psicológicos superiores*. Grijalbo.

Williams, R., Park, H. W., Oh, L., & Breazeal, C. (2019). PopBots: Designing an artificial intelligence curriculum for early childhood education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 9729-9736. <https://doi.org/10.1609/aaai.v33i01.33019729>

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-35. <https://doi.org/10.1145/1118178.1118215>

Yim, I. H. Y., & Su, J. (2024). Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education*, 1-39. <https://doi.org/10.1007/s40692-023-00304-9>

ANEXOS

La presente sección contiene la estructura completa de los tres instrumentos de recolección de datos diseñados para la investigación. Cada instrumento declara su tema, objetivo, consigna, instrumentación (ítems o categorías) y, cuando corresponde, los agradecimientos al participante. Estos formatos constituyen la versión de referencia aplicada en el trabajo de campo.

Anexo 1. Pruebas objetivas de conocimientos sobre IA: pre-test y post-test.

Anexo 1A. Pre-test.

Tema.

Conocimientos sobre principios de Inteligencia Artificial, Aprendizaje Automático y redes neuronales en estudiantes de 13 a 18 años.

Objetivo.

Medir el nivel de conocimiento técnico de los estudiantes sobre conceptos fundamentales de IA, arquitectura de redes neuronales, proceso de entrenamiento y lectura de código en *TensorFlow*, antes y después de la intervención pedagógica.

Consigna.

Estimado/a estudiante: a continuación encontrarás 20 preguntas sobre principios de Inteligencia Artificial. Lee cada enunciado con atención y selecciona la respuesta que consideres correcta. No existe penalización por respuestas incorrectas; si no conoces la respuesta puedes dejar el ítem en blanco. Dispones de 40 minutos para completar la prueba. Esta evaluación es anónima y los resultados serán utilizados exclusivamente con fines de investigación.

Instrumentación (estructura de los ítems).

Dimensión 1, Conceptos básicos de IA y ML (5 ítems de selección múltiple): definición de IA, distinción entre IA, ML y *Deep Learning*, tipos de aprendizaje, ejemplos cotidianos, ética básica.

Dimensión 2, Arquitectura de redes neuronales (5 ítems): neurona artificial, capas, función de activación (ReLU, Sigmoid), diferencia entre entrada/oculta/salida, parámetros y pesos.

Dimensión 3, Proceso de entrenamiento (5 ítems): épocas, función de pérdida, optimizador (ADAM), *overfitting*, métrica de exactitud.

Dimensión 4, Interpretación de código *TensorFlow* (5 ítems de aplicación): identificar capas en `model.add`, leer `model.compile`, interpretar salida de `model.fit`, modificar epochs, predecir efecto de cambiar capas.

Agradecimientos.

Agradecemos sinceramente tu participación en esta evaluación. Tu colaboración contribuye al desarrollo de mejores estrategias para enseñar Inteligencia Artificial a jóvenes en Guatemala.

Anexo 1B. Post-test.

Tema.

Conocimientos sobre principios de Inteligencia Artificial, Aprendizaje Automático y redes neuronales con énfasis en niveles cognitivos superiores (aplicar y analizar).

Objetivo.

Medir el nivel de comprensión y aplicación práctica de los estudiantes sobre conceptos de IA, arquitectura de redes neuronales, proceso de entrenamiento e interpretación de código en TensorFlow, después de la intervención didáctica.

Consigna.

Estimado/a estudiante: a continuación encontrarás 8 preguntas de selección múltiple sobre lo trabajado en el taller y 2 preguntas abiertas sobre tu experiencia. Lee cada enunciado con atención y selecciona la respuesta que consideres correcta. En las preguntas abiertas, responde con tus propias palabras. No existe penalización por respuestas incorrectas.

Instrumentación (estructura de los ítems).

Dimensión 1, Aplicación de redes neuronales (3 ítems de selección múltiple): identificación de arquitectura adecuada para un problema, interpretación de salida del modelo, predicción del efecto de modificar hiperparámetros.

Dimensión 2, Análisis del proceso de entrenamiento (3 ítems): lectura de curvas de pérdida, diagnóstico de overfitting, interpretación de la métrica de exactitud sobre datos de prueba.

Dimensión 3, Interpretación de código TensorFlow (2 ítems de aplicación): identificar errores en `model.compile` y `model.fit`, predecir el efecto de un cambio en la función de activación.

Preguntas abiertas (2): «¿Qué fue lo más valioso que aprendiste en este taller?» y «¿Qué sugerencias tendrías para mejorar la experiencia?».

Agradecimientos.

Agradecemos sinceramente tu participación al cerrar esta experiencia formativa. Tus respuestas son fundamentales para validar la propuesta y mejorar futuras iteraciones del plan curricular.

Anexo 2. Escala de involucramiento y motivación.

Tema.

Percepción de involucramiento, autonomía, competencia y relacionamiento durante el taller de IA.

Objetivo.

Medir el nivel de involucramiento activo y motivación intrínseca del estudiante durante la experiencia formativa, siguiendo el marco de la teoría de la autodeterminación de Ryan y Deci (2000).

Consigna.

A continuación encontrarás 18 afirmaciones sobre tu experiencia durante el taller. Lee cada una y marca con una X el número que mejor refleje tu nivel de acuerdo, donde 1 = Totalmente en desacuerdo y 5 = Totalmente de acuerdo. No hay respuestas correctas o incorrectas; lo que interesa es tu opinión sincera. Tus respuestas son confidenciales.

Instrumentación (dimensiones e ítems-tipo).

Dimensión Autonomía (6 ítems): «Pude decidir cómo modificar los parámetros», «Sentí que tenía libertad para experimentar», entre otros.

Dimensión Competencia (6 ítems): «Sentí que era capaz de entender los conceptos», «Logré ejecutar el código sin frustrarme», entre otros.

Dimensión Relacionamiento (6 ítems): «El docente me apoyó cuando lo necesité», «Trabajar con compañeros me ayudó a aprender», entre otros.

Agradecimientos.

Gracias por compartir tus percepciones. Tus respuestas permiten mejorar la propuesta pedagógica para futuras promociones de estudiantes.

Anexo 3. Estructura del notebook de Google Colab utilizado en el taller.

El notebook que se utilizó en la fase experimental del taller del 17 de febrero de 2025 está disponible en el repositorio público del investigador y se titula «taller_tensorflow_microlearning.ipynb». Su estructura por celdas es la siguiente.

Celda 1: Importaciones.

Carga de *TensorFlow*, Keras y matplotlib. Se utiliza la versión estable de *TensorFlow* incluida por defecto en Colab.

Celda 2: Carga del dataset Fashion-MNIST.

Descarga del dataset desde la API de Keras y separación en conjuntos de entrenamiento y prueba. Visualización de tres imágenes de muestra con sus etiquetas asociadas.

Celda 3: Normalización.

Escalamiento de los valores de píxel del rango [0, 255] al rango [0, 1] para facilitar la convergencia del modelo.

Celda 4: Definición del modelo.

Construcción de un modelo secuencial con tres capas: una capa de aplanamiento, una capa densa de 128 neuronas con activación ReLU, y una capa densa de 10 neuronas con activación Softmax para clasificación multiclase.

Celda 5: Compilación.

Configuración del modelo con optimizador ADAM, función de pérdida sparse categorical crossentropy y métrica accuracy.

Celda 6: Entrenamiento.

Ejecución de `model.fit` con cinco épocas iniciales sobre el conjunto de entrenamiento.

Celda 7: Evaluación.

Evaluación del modelo entrenado sobre el conjunto de prueba, con reporte de pérdida y exactitud.

Celda 8: Mini competencia (experimentación libre).

Espacio dedicado a que cada participante modifique al menos un hiperparámetro (número de épocas, número de neuronas en la capa oculta o función de activación) y registre la exactitud obtenida. La sesión cierra con la comparación abierta de los mejores resultados.

Anexo 4. Cronograma de la sesión piloto del 17 de febrero de 2025.

La intervención completa duró 90 minutos, distribuidos en cinco ciclos de *microlearning* más una mini competencia. La duración por ciclo se ajustó a la planificación previa con tolerancia de ± 2 minutos por sección.

Apertura (5 min): bienvenida, presentación del problema (reconocer prendas de ropa) y motivación inicial.

Ciclo 1, Datos e inputs (15 min): qué es un dataset, carga de Fashion-MNIST, visualización de imágenes.

Ciclo 2, Crear la red (15 min): qué es una red neuronal, qué son las capas, qué hacen ReLU y Softmax.

Ciclo 3, Entrenamiento (15 min): qué es entrenar, qué son las épocas, qué hace `model.fit`.

Ciclo 4, Hiperparámetros (15 min): efecto de cambiar neuronas, épocas y funciones de activación.

Mini competencia (15 min): experimentación libre con el objetivo explícito de lograr la mejor precisión posible.

Cierre (10 min): aplicación del post-test, escala Likert y preguntas abiertas.

Anexo 5. Bitácora del investigador (resumen).

Durante la jornada del 17 de febrero de 2025 se llevó una bitácora narrativa con observaciones puntuales por ciclo. A continuación se transcriben tres entradas representativas, anonimizadas.

Ciclo 2, minuto 23. «Una participante levantó la mano para preguntar por qué se necesita Softmax si ya tenemos ReLU. Le respondí con una analogía: ReLU enciende o apaga la información, Softmax la pone en una balanza para decidir cuál pesa más. La respuesta funcionó. Anoté la analogía para incluirla en futuras iteraciones del taller».

Mini competencia, minuto 72. «Tres estudiantes están en una guerra silenciosa por subir la exactitud del modelo. Uno cambió epochs a 20, otra cambió neuronas a 256, y un tercero combinó ambas modificaciones. La adrenalina es palpable. Esa palabra vuelve después en una respuesta abierta del post-test».

Cierre, minuto 88. «Al recoger los formularios noté que varios estudiantes se quedaron escribiendo después de tiempo. La mayoría de respuestas abiertas son largas y entusiastas. Sospecho que el cuestionario les permitió procesar lo aprendido, no solo evaluarlo».

Anexo 6. Estructura completa y código del notebook Colab utilizado en el taller.

El notebook utilizado en el taller del 17 de febrero de 2025 (archivo `taller_tensorflow_microlearning.ipynb`, disponible en el repositorio del autor) está organizado en doce celdas alternadas entre texto y código. La estructura completa es la siguiente:

#	Tipo	Contenido de la celda
1	Texto	Bienvenida y objetivos del taller. Recordatorio del consentimiento. Instrucciones para activar GPU en Colab.
2	Código	Importación de librerías: tensorflow as tf, matplotlib.pyplot as plt, numpy as np, from tensorflow.keras.datasets import fashion_mnist.
3	Código	Carga del dataset Fashion-MNIST: (x_train, y_train), (x_test, y_test) = fashion_mnist.load_data(); print del shape; visualización de las primeras 10 imágenes con etiqueta.
4	Texto	Explicación del dataset: 60.000 imágenes de entrenamiento, 10.000 de validación, 10 categorías de prendas, 28×28 px en escala de grises.
5	Código	Normalización: x_train = x_train / 255.0; x_test = x_test / 255.0. Discusión sobre por qué normalizar entre 0 y 1.
6	Texto	Introducción a la arquitectura: capa Flatten, capa Dense con ReLU, capa Dense de salida con Softmax.

7	Código	Construcción del modelo: <code>model = tf.keras.Sequential([Flatten(input_shape=(28,28)), Dense(128, activation='relu'), Dense(10, activation='softmax')])</code> . Llamada a <code>model.summary()</code> .
8	Código	Compilación: <code>model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])</code> .
9	Código	Entrenamiento: <code>history = model.fit(x_train, y_train, epochs=5, validation_data=(x_test, y_test))</code> . Reto al estudiante: cambiar epochs y observar.
10	Código	Visualización de las curvas de entrenamiento y validación con <code>plt.plot(history.history['accuracy'])</code> y <code>history.history['val_accuracy']</code> .
11	Código	Predicción sobre 5 imágenes elegidas al azar; visualización con la etiqueta predicha y la real; cálculo de cuántas acierta.
12	Texto	Reto final: dibujar a mano un dígito del 0 al 9, fotografiarlo, subirlo y pasarlo por el modelo. Discusión guiada sobre por qué falla en algunos casos (sesgos del dataset, condiciones de la foto, estilo del trazo).

El notebook fue diseñado para ejecutarse de extremo a extremo en menos de quince minutos en una sesión Colab gratuita con GPU activada. Dos versiones se mantienen: una versión guiada (con explicaciones extensas en celdas de texto entre cada bloque de código) y una versión avanzada (`taller_tensorflow_finalizado.ipynb`) que añade dos celdas extra de visualización y un mini-experimento con una segunda capa oculta. Ambos archivos están disponibles bajo licencia MIT en el repositorio del autor para uso libre por otros docentes que deseen replicar el taller.

Anexo 7. Plan de capacitación docente (16 horas, 4 sesiones).

El componente de capacitación docente que acompaña la implementación del plan se estructura como un curso intensivo de 16 horas distribuidas en cuatro sesiones de cuatro horas. Cada sesión condensa, en versión adaptada para docentes, los contenidos de uno de los cuatro módulos del plan estudiantil. La idea pedagógica es que el docente pase primero por la experiencia como aprendiz antes de operarla como facilitador.

Sesión	Eje temático	Contenido detallado	Producto / evidencia
Sesión 1 (4 h)	Fundamentos de IA y herramientas digitales	Introducción a Python básico en Colab; navegación del notebook; ejecución de la primera red neuronal; revisión de los recursos del kit pedagógico.	Notebook ejecutado al cierre + cuestionario diagnóstico de conocimientos previos del docente.
Sesión 2 (4 h)	Arquitectura y entrenamiento de redes	Profundización conceptual: capas, funciones de activación, optimizadores,	Plan de clase de la sesión M2S2 anotado por el docente con sus propias decisiones de adaptación.

		métricas. Revisión sesión por sesión del Módulo 2 del plan estudiantil.	
Sesión 3 (4 h)	Evaluación formativa, rúbricas y proyecto contextualizado	Práctica con la rúbrica analítica del Capítulo 4; estudio de casos de proyectos contextualizados; identificación de problemas locales relevantes.	Borrador de la lista de problemas que el docente propondría a sus estudiantes en el Módulo 4.
Sesión 4 (4 h)	Ética, seguridad y comunidad	Discusión sobre sesgos, privacidad, IA generativa, uso responsable; conexión con la comunidad técnica local (GDG, NodeBots, mentoría universitaria); recursos para aprendizaje continuo.	Plan de implementación del primer ciclo, firmado por el docente y por la coordinación del centro.

El curso se imparte en modalidad presencial intensiva durante el mes previo al inicio del plan con los estudiantes, o en modalidad híbrida (dos sesiones presenciales + dos asincrónicas con tutoría) cuando las condiciones del centro educativo lo requieran. Al cierre de las cuatro sesiones, el docente recibe un kit pedagógico digital completo con los 16 notebooks plantilla, las 16 guías minuto a minuto, las rúbricas, los instrumentos de evaluación, una bitácora de reflexión semanal y acceso al foro asincrónico de soporte durante el primer ciclo de implementación. La capacitación

se complementa con tres reuniones cortas de seguimiento (45 minutos cada una) durante el primer mes de implementación, una al final de cada semana.

Anexo 8. Respuestas anonimizadas del pre-test (n = 21).

Esta tabla presenta las respuestas individuales de los 21 estudiantes a las siete preguntas cerradas del pre-test, codificadas como 1 (acierto) o 0 (error). Los estudiantes se identifican con códigos S01 a S21 sin asociación con sus identidades reales. La columna «Total» reporta el puntaje individual sobre 7. Los datos provienen del archivo de respuestas de Google Forms y se publican cumpliendo la transparencia y reproducibilidad de los resultados reportados en el Capítulo 3.

ID	PT-1	PT-2	PT-3	PT-4	PT-5	PT-6	PT-8	Total
S01	1	1	1	1	1	1	1	7
S02	1	1	0	1	0	1	0	4
S03	1	0	1	1	1	1	1	6
S04	1	1	1	1	1	1	1	7
S05	1	1	1	0	1	1	1	6
S06	1	1	1	1	1	1	1	7
S07	1	1	1	1	1	1	1	7
S08	1	0	1	0	1	1	1	5

S09	1	1	1	1	1	1	1	7
S10	1	1	1	1	1	1	1	7
S11	1	1	1	1	1	1	1	7
S12	1	1	1	1	1	1	1	7
S13	1	1	1	0	0	0	1	4
S14	1	1	1	1	1	1	1	7
S15	1	1	1	1	1	1	1	7
S16	1	1	1	0	1	0	1	5
S17	1	1	1	1	1	1	1	7
S18	1	1	1	1	1	1	0	6
S19	1	1	1	0	1	1	1	6
S20	1	1	1	0	1	0	1	5
S21	1	1	0	0	0	1	1	4

Distribución de puntajes individuales: el puntaje mínimo observado fue 3/7 y el máximo 7/7, con mediana = 6 y moda = 7. La desviación estándar de 1.14 puntos sobre una escala de 7 confirma

una varianza moderada, suficiente para que el instrumento discrimine entre estudiantes pero baja para $n = 21$.

Anexo 9. Respuestas anonimizadas del post-test (n = 21).

Esta tabla presenta las respuestas individuales de los 21 estudiantes a las seis preguntas cerradas del post-test técnico, codificadas como 1 (acierto) o 0 (error), junto con la asignación de grupo (AI con *TensorFlow* u Otro/Tinkercad) y el puntaje total sobre 6. Los identificadores S01-S21 corresponden al mismo orden cronológico de envío del pre-test.

ID	Grupo	PoT-1	PoT-2	PoT-3	PoT-4	PoT-5	PoT-6	Total
S01	Otro	1	0	1	0	1	1	4
S02	AI	1	1	1	1	0	1	5
S03	Otro	1	1	0	1	1	1	5
S04	AI	1	1	0	0	0	1	3
S05	AI	0	1	0	0	1	1	3
S06	AI	0	1	0	1	1	1	4
S07	Otro	1	1	1	1	1	1	6
S08	AI	1	1	1	0	1	1	5
S09	AI	0	0	1	1	1	1	4

S10	Otro	0	1	0	1	0	0	2
S11	Otro	1	1	0	1	0	0	3
S12	AI	1	1	1	0	1	0	4
S13	Otro	1	1	1	1	1	1	6
S14	Otro	1	0	0	0	0	1	2
S15	Otro	1	1	1	1	0	0	4
S16	Otro	1	1	0	1	1	0	4
S17	AI	1	1	0	0	1	1	4
S18	AI	1	0	0	0	1	1	3
S19	Otro	1	1	0	1	1	1	5
S20	AI	1	1	1	1	1	0	5
S21	Otro	0	1	0	0	1	1	3

La asignación de estudiantes a los dos grupos quedó balanceada: 10 en el grupo de IA con *TensorFlow* y 11 en el grupo paralelo de electrónica con Tinkercad. La media de ambos grupos es idéntica ($M = 4.00/6$) con desviaciones estándar diferenciadas ($SD = 0.82$ para AI, $SD = 1.41$ para

Otro), lo que indica que el grupo paralelo presenta una dispersión mayor del puntaje, con tanto los puntajes extremos altos como bajos del estudio.

Anexo 10. Distribución completa de la escala Likert por ítem y por grupo.

Esta tabla detalla las frecuencias absolutas de cada nivel de respuesta (Nada de acuerdo / Poco de acuerdo / De acuerdo / Muy de acuerdo) para cada uno de los cuatro ítems de la escala Likert, segmentadas por grupo de asignación. Permite verificar la convergencia de la valoración favorable en ambos grupos descrita en la sección 3.4.9 del Capítulo 3.

Ítem	Grupo	Nada	Poco	De acuerdo	Muy	M
I1	AI	1	0	8	1	2.90
I1	Otro	0	1	5	5	3.36
I2	AI	1	1	6	2	2.85
I2	Otro	0	2	6	3	3.09
I3	AI	1	0	4	5	3.30
I3	Otro	1	2	4	4	3.00
I4	AI	1	0	3	6	3.40
I4	Otro	0	1	3	7	3.55

La diferencia entre grupos en la media global de la escala (AI = 3.11, Otro = 3.25) se concentra principalmente en el ítem I3 (experimentación con parámetros), donde el grupo paralelo reporta una valoración ligeramente superior. Este resultado es coherente con el formato del taller de Tinkercad, en el que la manipulación directa de circuitos es la actividad central, mientras que en el taller de IA con *TensorFlow* la experimentación se mediatiza a través de la modificación de parámetros en celdas de código de Colab.

Anexo 11. Material del taller: estructura de las diapositivas (Intro to Deep Learning using Keras).

La presentación de apoyo del taller, titulada «Intro to Deep Learning using Keras», fue desarrollada para acompañar los ciclos cortos de *microlearning*. Cada diapositiva sirve de anclaje visual a un microconcepto del ciclo correspondiente. La tabla siguiente lista la estructura conceptual de la presentación, organizada en bloques temáticos correspondientes a los cinco ciclos del taller. La presentación completa en formato PowerPoint se conserva en el repositorio del proyecto bajo el identificador de archivo señalado en la metadata.

Bloque	Diapositivas	Microconcepto	Soporte visual
Apertura	1-2	Bienvenida y problema motivador (clasificación de prendas)	Imagen del dataset Fashion-MNIST
Ciclo 1	3-6	Datos: ¿qué es un dataset y qué son las imágenes Fashion-MNIST?	Cuadrícula de 10 categorías de prendas

Ciclo 1	7-9	Estructura de los datos: dimensiones de las imágenes y etiquetas	Diagrama de matriz 28×28 píxeles
Ciclo 2	10-13	Red neuronal: capas de entrada, ocultas y salida; función ReLU y Softmax	Esquema de capas con neuronas conectadas
Ciclo 2	14-15	Compilación del modelo: optimizador, función de pérdida y métrica	Bloque de código resaltado
Ciclo 3	16-18	Entrenamiento: épocas, lotes y curvas de pérdida y precisión	Curvas típicas de entrenamiento
Ciclo 4	19-21	Hiperparámetros: cómo cambian neuronas, épocas y activación los resultados	Tabla comparativa de configuraciones
Mini competencia	22-23	Reto abierto: encontrar la configuración con mejor exactitud	Tabla de resultados a completar

Cierre	24-25	Síntesis y agradecimientos; pasos siguientes para profundizar	QR a recursos abiertos
--------	-------	---	------------------------

Anexo 12. Agenda operativa y guía pedagógica del taller (17 de febrero de 2025).

Esta sección reproduce la estructura operativa del taller tal como se aplicó en la jornada del 17 de febrero de 2025. La agenda detalla la distribución temporal de los bloques de la jornada completa de 8:00 a 11:20, mientras que la guía pedagógica describe el contenido pedagógico de los cinco ciclos de *microlearning* del taller de *TensorFlow*.

Agenda de la jornada (8:00 a 11:20).

Bloque	Horario	Actividad
Apertura común	8:00 - 8:30	Bienvenida, encuadre, aplicación del pre-test (Google Forms)
Bloque expositivo común	8:30 - 9:00	Introducción a la Inteligencia Artificial (concepto, aplicaciones, ética básica)
Receso	9:00 - 9:15	Refrigerio y reagrupamiento por taller asignado

Talleres paralelos	9:15 - 10:45	Taller A: IA con <i>TensorFlow</i> (n = 10) Taller B: Electrónica con Tinkercad (n = 11)
Cierre común	10:45 - 11:10	Reagrupamiento, síntesis cruzada de ambos talleres y discusión abierta
Aplicación del post-test	11:10 - 11:20	Aplicación del post-test (Google Forms) y cierre formal

Guía pedagógica del taller de TensorFlow (90 minutos efectivos).

El taller sigue la metodología de *microlearning*, dividiendo el tiempo en ciclos de aprendizaje cortos donde cada ciclo incluye un microconcepto (5 minutos), un microdesafío práctico (8 minutos) y un microcierre (2 minutos). Esta estructura se repite cinco veces antes del bloque final de mini competencia abierta.

Ciclo	Duración	Microconcepto	Microdesafío
Ciclo 0: Motivación	5 min	Reconocer ropa con una red neuronal	Predicción a vista de los estudiantes sobre las imágenes del dataset
Ciclo 1: Datos e inputs	15 min	Qué es un dataset y qué son las imágenes Fashion-MNIST	Cargar el dataset y visualizar una imagen en Colab

Ciclo 2: Crear la red	15 min	Capas y activaciones (ReLU, Softmax)	Construir el modelo con celdas de fill in the blanks
Ciclo 3: Entrenamiento	15 min	Épocas, batch y compilación del modelo	Compilar y entrenar el modelo con cinco épocas
Ciclo 4: Hiperparámetros	15 min	Efecto de neuronas, épocas y función de activación	Modificar tres hiperparámetros y observar el cambio
Mini competencia	15 min	Búsqueda libre de la mejor configuración	Probar configuraciones y compartir el mejor resultado

Recursos requeridos: *Google Colab*, cuentas Google institucionales de los estudiantes, TensorFlow y Keras (preinstalados en Colab), conexión a internet estable, proyector. Consejos pedagógicos clave aplicados durante la jornada: celebrar logros y no solo precisión, preguntar constantemente qué está pasando y por qué, ayudar rápido a quienes se bloqueen.

Anexo 13. Glosario técnico.

Este glosario define los términos técnicos centrales que aparecen en el Capítulo 4 (propuesta curricular) y que son indispensables para la lectura de los notebooks del taller. Se mantiene una redacción accesible al docente de nivel medio sin formación previa específica en aprendizaje automático.

Término

Definición

Activación (función de)	Función matemática aplicada a la salida de una neurona para introducir no linealidad. Las más comunes son ReLU para capas intermedias y Softmax para la capa de salida en problemas de clasificación.
Aprendizaje supervisado	Modalidad de aprendizaje automático en la que el modelo se entrena con ejemplos etiquetados, es decir, con pares de entrada y salida esperada.
Batch (lote)	Subconjunto de los datos de entrenamiento que el modelo procesa antes de actualizar sus pesos internos. Un valor típico es 32 o 64 ejemplos.
Capa	Conjunto de neuronas que procesan una transformación específica de la información. Una red neuronal típica tiene una capa de entrada, una o varias capas ocultas y una capa de salida.
Colab (Google Colab)	Entorno de notebooks Python en la nube provisto por Google, que permite ejecutar código sin instalación local y con acceso gratuito a aceleradores GPU/TPU.
Dataset	Conjunto estructurado de datos que se usa para entrenar y evaluar un modelo. Fashion-MNIST es un dataset clásico de 70 000 imágenes de prendas de vestir.

Época (epoch)	Una pasada completa del modelo sobre todos los datos de entrenamiento. Entrenar por más épocas suele mejorar el ajuste, hasta el punto de overfitting.
Hiperparámetro	Configuración del modelo o del entrenamiento que el diseñador fija antes de entrenar (épocas, neuronas, tasa de aprendizaje), en contraste con los pesos que el modelo aprende.
Inteligencia Artificial	Campo amplio de las ciencias de la computación que estudia cómo construir sistemas que ejecuten tareas que requieren capacidades asociadas a la inteligencia humana.
Keras	API de alto nivel para construir redes neuronales, integrada en <i>TensorFlow</i> desde 2017. Su sintaxis simplificada es la opción recomendada para introducción a la programación de redes.
<i>Machine Learning</i>	Subcampo de la Inteligencia Artificial que estudia algoritmos que aprenden patrones a partir de datos en lugar de ser programados explícitamente.
Microlearning	Estrategia pedagógica que organiza el aprendizaje en unidades muy breves (3-15 minutos), cada una con un objetivo específico,

una práctica corta y una retroalimentación inmediata.

`model.fit()`

Método de Keras/*TensorFlow* que ejecuta el entrenamiento del modelo sobre un dataset durante un número definido de épocas.

Overfitting (sobreajuste)

Situación en la que el modelo aprende demasiado bien los datos de entrenamiento y pierde capacidad de generalizar a datos nuevos.

Pérdida (loss)

Función matemática que mide la diferencia entre la predicción del modelo y la respuesta correcta. El entrenamiento busca minimizar esta función.

Pesos

Parámetros internos de una red neuronal que se ajustan durante el entrenamiento para aproximar la salida deseada.

ReLU

Función de activación que devuelve cero si la entrada es negativa y la entrada misma si es positiva. Es la función más usada en capas intermedias por su eficiencia computacional.

Red neuronal

Modelo computacional inspirado en la estructura del cerebro, compuesto por capas de unidades simples (neuronas) conectadas por pesos ajustables.

Softmax	Función de activación que se aplica en la capa de salida de un clasificador de múltiples clases para convertir las puntuaciones internas en una distribución de probabilidad.
<i>TensorFlow</i>	Biblioteca de código abierto desarrollada por Google para construir y entrenar modelos de aprendizaje automático, especialmente redes neuronales profundas.
Tinkercad	Plataforma web gratuita de Autodesk para diseño 3D y simulación de circuitos electrónicos, utilizada en el grupo paralelo del taller del 17 de febrero de 2025.

Anexo 14. Alineación del plan curricular con AI4K12 y orientaciones de la UNESCO.

Esta sección presenta el mapeo del plan curricular propuesto en el Capítulo 4 con dos referentes internacionales: las cinco grandes ideas de AI4K12 (Touretzky et al., 2019) y las orientaciones de la UNESCO sobre IA en educación (Pedro et al., 2019). El propósito es facilitar a equipos curriculares e investigadores la auditoría de la cobertura conceptual y la identificación de áreas que pueden requerir profundización en futuras iteraciones.

Mapeo módulo a gran idea AI4K12.

Las cinco grandes ideas de AI4K12 son: (1) Percepción —los computadores perciben el mundo mediante sensores—; (2) Representación y razonamiento —los agentes mantienen representaciones del mundo y las usan para razonar—; (3) Aprendizaje —los computadores pueden aprender de datos—; (4) Interacción natural —los agentes inteligentes requieren múltiples tipos de conocimiento para interactuar naturalmente con humanos—; y (5) Impacto social —la IA puede tener impactos positivos y negativos en la sociedad.

Módulo	Tema	Big idea AI4K12	Cobertura
1	Introducción a la IA	1, 2, 5	Percepción de imágenes; representación interna; ética y consecuencias
2	Red neuronal y activaciones	2, 3	Representación distribuida en pesos; aprendizaje supervisado básico
3	Entrenamiento y optimización	3	Aprendizaje a profundidad: pérdida, optimizador, épocas, overfitting
4	Aplicación y ética final	1, 4, 5	Aplicación contextualizada; interacción humano-modelo; ética en uso

Mapeo con orientaciones UNESCO.

La UNESCO (Pedro et al., 2019) propone cinco principios para la integración curricular de IA en educación: enfoque humanista, equidad de acceso, transparencia algorítmica, responsabilidad ética y desarrollo de competencias para el siglo XXI. La tabla siguiente describe cómo el plan curricular propuesto materializa cada principio.

Principio UNESCO**Operacionalización en el plan curricular**

Enfoque humanista

Los proyectos finales del módulo 4 se contextualizan a problemas del entorno guatemalteco (cultivos, monumentos, clima local), priorizando el sentido sobre la mera técnica.

Equidad de acceso

Uso exclusivo de Google Colab gratuito y dataset públicos. La propuesta no requiere instalación local ni hardware especial, lo que la hace transferible a centros con conectividad básica.

Transparencia algorítmica

El módulo 3 dedica tiempo explícito a leer la curva de pérdida y explicar por qué el modelo acierta o falla, evitando tratar el modelo como caja negra.

Responsabilidad ética

El módulo 4 incluye un bloque de reflexión ética obligatorio: identificación de sesgos, riesgos de uso indebido y discusión sobre quién es responsable cuando un modelo se equivoca.

Competencias siglo XXI

La metodología de *microlearning* con experimentación libre desarrolla autonomía, capacidad de iteración rápida y tolerancia a la incertidumbre, competencias transversales del siglo XXI.

Anexo 15. Plan curricular extendido: 16 sesiones desglosadas.

Esta sección desglosa las 16 sesiones del plan curricular en una tabla operativa que indica, para cada sesión, su objetivo de aprendizaje específico, las actividades en aula, el producto esperado del estudiante y el instrumento de evaluación formativa correspondiente. El nivel de detalle pretende facilitar la replicación por parte de docentes que adopten el plan sin participar en su diseño original.

Sesión	Módulo	Objetivo de aprendizaje	Actividad central en aula	Producto del estudiante
1	1. Intro a IA	Identificar aplicaciones cotidianas de IA y distinguir IA de otros tipos de software	Mapeo en grupo de aplicaciones de IA en su entorno	Cartografía de aplicaciones (en papel o digital)
2	1. Intro a IA	Distinguir IA, ML y aprendizaje profundo como subconjuntos jerárquicos	Diagrama de conjuntos guiado y discusión	Diagrama anotado individual
3	1. Intro a IA	Discutir un caso de sesgo algorítmico documentado	Análisis de caso (reconocimiento facial con sesgo)	Reporte breve de posiciones éticas

4	1. Intro a IA	Reconocer datasets como insumo del aprendizaje supervisado	Exploración guiada de Fashion-MNIST en Colab	Notebook con visualización de imágenes
5	2. Red neuronal	Identificar capas, neuronas y conexiones de una red simple	Construcción guiada de una red de tres capas	Notebook con red construida y comentada
6	2. Red neuronal	Distinguir las funciones de activación ReLU y Softmax	Comparación visual de funciones de activación	Notebook con comparación experimental
7	2. Red neuronal	Explicar el papel de los pesos en el aprendizaje	Visualización de pesos antes y después de entrenar	Notebook con análisis de pesos
8	2. Red neuronal	Predecir el efecto de cambiar el número de neuronas en una capa	Experimentos pareados con configuraciones distintas	Tabla comparativa de resultados

9	3. Entrenamiento	Definir época, batch y función de pérdida	Compilación guiada de un modelo y entrenamiento de cinco épocas	Notebook con modelo entrenado
10	3. Entrenamiento	Leer e interpretar la curva de pérdida	Generación y análisis de la curva de entrenamiento	Captura comentada de la curva
11	3. Entrenamiento	Identificar el fenómeno de overfitting	Entrenamiento intencionalmente prolongado y diagnóstico	Notebook con caso documentado de overfitting
12	3. Entrenamiento	Aplicar una estrategia básica de regularización	Aplicación de dropout y comparación de resultados	Notebook comparativo con y sin regularización
13	4. Aplicación y ética	Diseñar un proyecto contextualizado al entorno local	Lluvia de ideas grupal de proyectos pertinentes a Guatemala	Propuesta breve de proyecto individual o por parejas

14	4. Aplicación y ética	Recolectar y preparar un dataset propio mínimo	Captura de datos del entorno y carga a Colab	Notebook con dataset propio cargado
15	4. Aplicación y ética	Entrenar y evaluar el modelo del proyecto propio	Entrenamiento y reporte de métricas	Notebook completo del proyecto
16	4. Aplicación y ética	Presentar el proyecto y discutir consideraciones éticas	Presentación oral pública con discusión de impacto	Presentación oral y notebook publicado en repositorio

La carga total declarada es de 16 sesiones presenciales de 60 minutos en aula, complementadas con ocho horas de trabajo asincrónico distribuidas a lo largo del cuatrimestre, totalizando 24 horas efectivas por estudiante. La evaluación formativa transversal combina la rúbrica de notebook (técnica y comunicativa) con la rúbrica de presentación oral del proyecto final del módulo 4, ambas ya descritas en el Capítulo 4 del documento principal.