



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

Facultad de
ESTOMATOLOGÍA

PRECISIÓN DE GRANDES MODELOS DE LENGUAJE DE INTELIGENCIA
ARTIFICIAL (IA) AL RESOLVER EL EXAMEN NACIONAL DE
ODONTOLOGÍA ENAO 2024

ACCURACY OF LARGE ARTIFICIAL INTELLIGENCE (AI) LANGUAGE
MODELS IN SOLVING THE ENAO 2024 NATIONAL DENTISTRY
EXAMINATION

TESIS PARA OPTAR POR EL TÍTULO PROFESIONAL DE
CIRUJANO DENTISTA

AUTOR

JESELY HARITZ TORRES RAMIREZ

ASESOR

MIGUEL ANGEL SARAVIA ROJAS

CO-ASESOR

FRANCISCO JOSE OREJUELA RAMIREZ

LIMA - PERÚ

2025

JURADO

Presidente: MG. FERNANDO ITALO RAFAEL BRIGNARDELLO CEDRON

Vocal: MG. MARIA DE LAS MERCEDES MALDONADO MENDOZA

Secretario: DRA. ROCIO GENG VIVANCO

Fecha de Sustentación: 29 de abril de 2025

Calificación: Aprobado con Honores

ASESORES DE TESIS

ASESOR

PHD MIGUEL ANGEL SARAVIA ROJAS

Departamento Académico de Clínica Estomatológica

ORCID: 0000-0002-1651-6138

CO-ASESOR

CD. FRANCISCO JOSE OREJUELA RAMIREZ

Departamento Académico de Odontología Social

ORCID: 0000-0002-9790-9071

DEDICATORIA

A Dios por darme la fuerza necesaria para lograr este objetivo. A mi mayor inspiración: mis amados padres, Jesús y Elizabeth, gracias por su guía, su amor incondicional y por sostenerme con su apoyo en cada paso. Mis logros son suyos. A mis hermanas, Jhearlett y Lua, por su cariño, compañía, consejos y constante motivación. A mi familia y a todas las personas que estuvieron cerca durante este proceso, gracias por su cariño y aliento. A Pablo, por creer siempre en mí y en lo que soy capaz de lograr. Y por último, a mí misma, por nunca rendirme, aún en los días difíciles. Lo estás haciendo bien.

AGRADECIMIENTOS

Agradezco profundamente a mi familia por estar siempre a mi lado y darme la motivación para seguir adelante. A mi universidad, por brindarme las herramientas necesarias para formarme profesionalmente. A los docentes que, con su conocimiento y compromiso, aportaron a mi crecimiento académico. A mis asesores, por su tiempo y disposición en distintos momentos del proceso. A Dios, por su fidelidad constante desde el primer día que inicié mi carrera. Finalmente, agradezco a cada persona que me tendió una mano, me escuchó o simplemente me acompañó en silencio durante este reto. A todas y todos, gracias.

DECLARACIÓN DE CONFLICTOS DE INTERÉS

Los autores declaran no tener conflictos de interés.

RESULTADO DEL INFORME DE SIMILITUD



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

Facultad de
ESTOMATOLOGÍA

PRECISIÓN DE GRANDES MODELOS DE LENGUAJE DE INTELIGENCIA
ARTIFICIAL (IA) AL RESOLVER EL EXAMEN NACIONAL DE
ODONTOLOGÍA ENAO 2024

ACCURACY OF LARGE ARTIFICIAL INTELLIGENCE (AI) LANGUAGE
MODELS IN SOLVING THE ENAO 2024 NATIONAL DENTISTRY
EXAMINATION

TESIS PARA OPTAR POR EL TÍTULO PROFESIONAL DE
CIRUJANO DENTISTA

AUTOR

JESELY HARITZ TORRES RAMIREZ

ASESOR

MIGUEL ANGEL SARAVIA ROJAS

CO-ASESOR

FRANCISCO JOSE OREJUELA RAMIREZ

LIMA - PERÚ

2025

11% Similitud estándar Filtros

Fuentes

Mostrar las fuentes solapadas Info Toggle

1	Internet	hd.handle.net	2%
		10 bloques de texto	90 palabra que coinciden
2	Internet	repositorio.upch.edu.pe	1%
		5 bloques de texto	68 palabra que coinciden
3	Internet	www.coursehero.com	1%
		7 bloques de texto	68 palabra que coinciden
4	Internet	www.infobae.com	<1%
		1 bloques de bloques	22 palabra que coinciden
5	Internet	repositorio.ucv.edu.pe	<1%
		2 bloques de texto	20 palabra que coinciden

TABLA DE CONTENIDOS

	Pág.
Resumen	
Abstract	
I. Introducción	1
II. Objetivos	4
III. Materiales y Métodos	5
IV. Resultados	9
V. Discusión	12
VI. Conclusiones	19
VII. Referencias Bibliográficas	20
VIII. Tablas y gráficos	24
Anexos	

RESUMEN

Introducción: La inteligencia artificial (IA) ha revolucionado diversas disciplinas, incluida la odontología. Los Grandes Modelos de Lenguaje han demostrado potencial en el procesamiento de información, pero su precisión en evaluaciones estandarizadas como el Examen Nacional de Odontología (ENAO) en Perú no ha sido ampliamente estudiada. **Objetivo:** Comparar la precisión de los modelos de IA Claude 3.5 Sonnet, ChatGPT-4o, Perplexity y Copilot en la resolución de preguntas del ENAO 2024. **Métodos y Materiales:** Estudio transversal. Fue realizado basado en 200 preguntas del solucionario oficial del ENAO 2024. Se incluyeron preguntas con redacción clara y única alternativa correcta, excluyendo aquellas incompletas o con errores. Las variables analizadas fueron los modelos de IA, áreas temáticas del ENAO 2024 y las preguntas del examen. Las respuestas generadas por los modelos de IA se compararon con el solucionario oficial del examen. Se utilizó la prueba Chi-cuadrado ($p < 0.001$). **Resultados:** ChatGPT-4o obtuvo la mayor precisión (93.5 %), seguido de Claude 3.5 Sonnet (89.5 %). Perplexity y Copilot alcanzaron un desempeño inferior y similar (77 %). La mayor precisión se observó en especialidades como Investigación y Gerencia de Servicios de Salud, mientras que la menor se registró en Endodoncia y Rehabilitación Oral. **Conclusión:** Los modelos de IA lograron aprobar el ENAO 2024, destacando en preguntas teóricas, pero con menor precisión en áreas clínicas. ChatGPT-4o y Claude 3.5 Sonnet demostraron mejor desempeño, lo que sugiere su potencial como herramientas complementarias en la educación odontológica.

Palabras Clave: AI, Perplexity, Copilot, Educación dental, ChatGPT, Claude.

ABSTRACT

Introduction: Artificial intelligence (AI) has revolutionized various disciplines, including dentistry. Large Language Models have shown potential in information processing, but their accuracy in standardized evaluations such as the National Dentistry Exam (ENAO) in Peru has not been widely studied. **Objective:** To compare the accuracy of the AI models Claude 3.5 Sonnet, ChatGPT-4o, Perplexity, and Copilot in answering questions from the 2024 ENAO. **Materials and Methods:** A cross-sectional study was conducted using 200 questions from the official answer key of the 2024 ENAO. Only questions with clear wording and a single correct answer were included; incomplete or erroneous questions were excluded. The variables analyzed were the AI models, thematic areas of the 2024 ENAO, and the exam questions. The answers generated by the AI models were compared with the official answer key. The Chi-square test was used ($p < 0.001$). **Results:** ChatGPT-4o obtained the highest accuracy (93.5 %), followed by Claude 3.5 Sonnet (89.5 %). Perplexity and Copilot achieved lower and similar performance (77 %). The highest accuracy was observed in theoretical specialties such as Research and Health Services Management, while the lowest was recorded in Endodontics and Oral Rehabilitation. **Conclusion:** The AI models successfully passed the 2024 ENAO, performing best on theoretical questions but showing lower accuracy in clinical areas. ChatGPT-4o and Claude 3.5 Sonnet demonstrated better performance, suggesting their potential as complementary tools in dental education.

Keywords: AI, Perplexity, Copilot, Dental Education, ChatGPT, Claude.

I. INTRODUCCIÓN

En los últimos años, la inteligencia artificial (IA) ha revolucionado diversos campos del conocimiento, incluyendo la medicina y la odontología. Su impacto se ha visto reflejado en el desarrollo de herramientas avanzadas que facilitan el diagnóstico, la planificación de tratamientos y la educación médica (1). Dentro de este avance, los Grandes Modelos de Lenguaje (Large Language Models, LLMs) como ChatGPT-4o (OpenAI), Claude 3.5 Sonnet (Anthropic), Perplexity (Aravind Srinivas et al.) y Copilot (Microsoft) han demostrado una notable capacidad para procesar y generar información en múltiples disciplinas, lo que ha despertado el interés en su aplicación dentro de entornos académicos y evaluativos. Estas tecnologías prometen ser un recurso valioso para los estudiantes y profesionales, pero aún persisten interrogantes sobre su precisión y fiabilidad en contextos especializados.

En ese marco, el Examen Nacional de Odontología (ENAO) en Perú ha sido uno de los principales instrumentos para medir el nivel de conocimientos de los egresados de odontología y establecer el orden de mérito, junto con el promedio ponderado de pregrado, para la adjudicación de plazas al Servicio Rural Urbano Marginal de Salud (SERUMS) (2). Este tipo de evaluación estandarizada representa una oportunidad concreta para analizar el desempeño de los modelos de IA en un contexto académico formal y con alta exigencia técnica, lo que podría abrir nuevas posibilidades para su integración en la formación profesional de los cirujanos dentistas.

Cabe destacar que hasta agosto del año 2024, el Examen Nacional de Odontología (ENAO) fue organizado por la Asociación Peruana de Facultades de Odontología

(ASPEFO). Esta evaluación constaba de dos pruebas, cada una con 100 preguntas de opción múltiple, y abarcaba 14 áreas temáticas odontológicas: Anatomía, Embriología e Histología; Odontopediatría; Periodoncia; Radiología; Salud pública; Medicina y patología estomatológica; Ortodoncia; Endodoncia; Cirugía; Rehabilitación oral y anatomía dentaria; Investigación; Gerencia de servicios de salud; y Ética y deontología (2). Las preguntas eran elaboradas a partir de un banco validado en talleres nacionales con la participación de docentes y expertos del área. Sin embargo, en su última edición se identificaron referencias bibliográficas con más de cuarenta años de antigüedad (3,4), lo cual evidenció la necesidad de actualizar las fuentes utilizadas.

A partir de la Resolución Ministerial N.º 514-2024/MINSA (5), la organización del examen pasó al Ministerio de Salud (MINSA), marcando un cambio significativo en su estructura y enfoque. Bajo esta nueva gestión, el ENAO se redujo a una única prueba de 100 preguntas, centrada en cinco áreas temáticas generales: Salud pública, Cuidado integral de salud, Ética e interculturalidad, Investigación y Gestión de servicios de salud (6). Por ello, este estudio adquiere especial relevancia al analizar el último examen organizado bajo el modelo tradicional de ASPEFO.

Diversos estudios han explorado la eficacia de los LLMs en el ámbito médico y odontológico. En Polonia, el estudio titulado “Análisis comparativo del desempeño de ChatGPT-4, Gemini y Claude en el Examen Final de Diploma Médico y el Examen de Verificación Médico-Odontológico en Polonia” tuvo como objetivo comparar el rendimiento de estos modelos en exámenes oficiales. Se concluyó que ChatGPT-4 obtuvo el mejor resultado, seguido por Claude y luego Gemini, aunque

todos mostraron limitaciones que requieren supervisión clínica (1). En Arabia Saudita, el estudio “Inteligencia artificial en periodoncia: Evaluación del desempeño de ChatGPT, Claude y Gemini en el Examen de Servicio” buscó evaluar su utilidad en la formación dental. ChatGPT demostró el mayor nivel de precisión, aunque los tres modelos necesitan supervisión profesional para evitar respuestas erróneas (7). En Estados Unidos, el estudio “Análisis comparativo de las respuestas de modelos de inteligencia artificial en la educación médica” tuvo como objetivo analizar la calidad de las respuestas generadas por estos modelos en contextos educativos. Se concluyó que los LLMs pueden ser herramientas útiles en la enseñanza médica, siempre que se usen como complemento y bajo supervisión (8).

En este sentido, la integración de la IA en la educación odontológica no solo podría mejorar la calidad del aprendizaje, sino también facilitar la personalización de la enseñanza, adaptando los contenidos a las necesidades de cada estudiante (9). A medida que estas tecnologías continúan evolucionando, su papel en la formación de los futuros profesionales de la odontología podría expandirse, proporcionando herramientas innovadoras para la autoevaluación y el desarrollo de habilidades clínicas.

A pesar de estos estudios, existe una brecha en la investigación sobre el desempeño de los LLMs en el contexto específico del ENAO 2024 en Perú. Por ello, el presente estudio pretende evaluar: ¿Cuál es la precisión de los modelos de lenguaje de inteligencia artificial ChatGPT-4o, Claude 3.5 Sonnet, Perplexity y Copilot al responder las preguntas del Examen Nacional de Odontología (ENAO) 2024?

II. OBJETIVOS

Objetivo general

Comparar la precisión de las respuestas generadas por los grandes modelos de lenguaje de inteligencia artificial (IA) ChatGPT-4o, Claude 3.5 Sonnet, Perplexity y Copilot en la resolución de preguntas del Examen Nacional de Odontología (ENAO) 2024.

Objetivos Específicos

1. Describir el nivel de precisión de las respuestas dadas por los modelos ChatGPT-4o, Claude3.5 Sonnet, Perplexity y Copilot en función de las especialidades abordadas en el ENAO 2024.
2. Determinar la concordancia entre las respuestas de los modelos de IA y las respuestas oficiales del ENAO 2024 según especialidades como Anatomía, Embriología e Histología; Odontopediatría; Periodoncia; Radiología; Salud pública; Medicina y patología estomatológica; Ortodoncia; Endodoncia; Cirugía; Rehabilitación oral y anatomía dentaria; Investigación; Gerencia de servicios de salud; y Ética y deontología.

III. MATERIALES Y MÉTODOS

Tipo de estudio

El presente estudio fue de tipo transversal.

Población

La población del estudio consistió en las 200 preguntas de los solucionarios oficiales del Examen Nacional de Odontología (ENAO) del 2024. Se utilizaron los registros de las 200 preguntas, disponibles en el banco de respuestas de la Asociación Peruana de Facultades de Odontología (ASPEFO) (3,4).

Criterios de participación

Respecto a las preguntas seleccionadas, se incluyeron aquellas que presentaban una redacción clara y sin ambigüedades, contenían una única alternativa correcta y abordaban contenidos de cada especialidad evaluada. Se excluyeron aquellas que presentaban redacción incompleta, mal redactadas o con errores gramaticales, así como aquellas que presentaban redacción subjetiva o respuestas múltiples, lo que podría haber generado confusión en la interpretación de los modelos de IA.

Es importante mencionar que tras aplicar estos criterios, todas las preguntas fueron consideradas válidas para el estudio.

Variables del estudio

Para el presente estudio, las variables consideradas se detallan en el Anexo 1.

- Los cuatros modelos de inteligencia artificial: ChatGPT-4o, Claude 3.5 Sonnet, Perplexity y Copilot. ChatGPT-4o es conocido por su capacidad avanzada de razonamiento, Claude 3.5 Sonnet destaca por su velocidad y capacidades de codificación, Perplexity se enfoca en proporcionar respuestas con citas en tiempo real, y Copilot se integra con el ecosistema de Microsoft. Se optó por elegir estos modelos debido a que ChatGPT-4o y Claude 3.5 Sonnet fueron los modelos más actuales en su momento, lanzados el 14 de mayo y 20 de junio del 2024, respectivamente. Además, no se encontraban estudios recientes acerca de los modelos de Perplexity y Copilot.
- Las áreas temáticas del Examen Nacional de Odontología 2024 que abarca las especialidades de Anatomía, Embriología e Histología; Odontopediatría; Periodoncia; Radiología; Salud pública; Medicina y patología estomatológica; Ortodoncia; Endodoncia; Cirugía; Rehabilitación oral y anatomía dentaria; Investigación; Gerencia de servicios de salud; y Ética y deontología.
- Y por último, las preguntas del Examen Nacional de Odontología 2024, son 2 pruebas, cada una con 100 preguntas, dando un total de 200 preguntas. Cada pregunta era de selección múltiple, con cinco respuestas posibles de la A a la E, no contenían imágenes, cuadros o tablas. Para el presente estudio se agruparon en correctas e incorrectas.

Técnicas y procedimientos

Para evaluar el desempeño de los modelos de inteligencia artificial, se introdujeron las 200 preguntas del ENAO 2024 en los sitios web oficiales de los modelos de lenguaje. Las respuestas de ChatGPT-4o se obtuvieron en <https://chatgpt.com/>, Claude 3.5 Sonnet en <https://claude.ai/new>, Perplexity en <https://www.perplexity.ai/> y Copilot en <https://copilot.microsoft.com/chats/y7s5Z63peU6WZDUtrz6nc>.

Para cada pregunta se utilizó la misma instrucción: “¿Cuál es la respuesta correcta a la siguiente pregunta del ENAO 2024 en Perú?”. La evaluación de los modelos se realizó durante el mes de julio del año 2024.

Se recopilaron 200 respuestas por cada chatbot, sumando un total de 800 respuestas analizadas. Para verificar la precisión de las respuestas, cada una se comparó con el solucionario oficial del ENAO 2024. Si la respuesta proporcionada por el modelo coincidía con la del solucionario, se consideraba correcta; en caso contrario, se marcaba como incorrecta.

Las respuestas fueron puntuadas de acuerdo con los criterios establecidos, asignando 1 punto a cada respuesta correcta y 0 puntos a las incorrectas. Se realizó una revisión de las respuestas generadas por cada modelo para evaluar su precisión, comparándolas entre sí. Además, se clasificaron las respuestas según la especialidad correspondiente para análisis comparativos.

Análisis estadístico

Se exportó la base de datos al software Stata versión 18. Para el análisis descriptivo se analizaron medidas de frecuencia absoluta y relativa, y se comprobaron los supuestos estadísticos. Posteriormente, se realizó un análisis bivariado de la precisión de las respuestas generadas por ChatGPT-4o, Claude 3.5 Sonnet, Perplexity y Copilot utilizando la prueba Chi cuadrado. Considerando un nivel de confianza del 95% y como indicador de significancia estadística un valor de $p < 0.05$.

Aspectos éticos

El estudio se realizó en cumplimiento con los principios éticos para la investigación. Se ejecutó posteriormente a la aprobación realizada por el comité institucional de ética de la Universidad Peruana Cayetano Heredia (CIE-UPCH) con número SIDISI 214542 en fecha lunes 17 de junio del 2024.

IV. RESULTADOS

Se evaluó el desempeño de los modelos de inteligencia artificial en el ENAO 2024, en función de la especialidad. Los niveles de precisión alcanzados se presentan en la Tabla 1 y el Gráfico 1. ChatGPT-4o obtuvo el mayor porcentaje de respuestas correctas alcanzando un 93.50 % (n = 187), seguido de Claude 3.5 Sonnet con un 89.50 % (n = 179). Por otro lado, Perplexity y Copilot presentaron valores similares, alcanzando un 77.00 % (n = 154). En términos de rendimiento general, todos los modelos evaluados lograron superar el mínimo requerido del 55.00 % (n = 110) para aprobar el ENAO 2024.

En el análisis por especialidad, ChatGPT-4o obtuvo el mayor porcentaje de respuestas correctas con un 100.00 % en las especialidades de Anatomía (n = 10), Embriología e histología (n = 30), Periodoncia (n = 15), Medicina y patología estomatológica (n = 15), Cirugía (n = 15), Investigación (n = 10), Gerencia de servicios de salud (n = 6) y Ética y deontología (n = 4). Sin embargo, su menor precisión se observó en Rehabilitación oral y anatomía dentaria con un 76.67 % (n = 23), donde también tuvo el mayor número de respuestas incorrectas con un 23.33 % (n = 7). Claude 3.5 Sonnet destacó en Periodoncia (n = 15), Cirugía (n = 15), Investigación (n = 10) y Gerencia de servicios de salud (n = 6) con un 100.00 %, pero tuvo su peor desempeño en Endodoncia con 60.00 % (n = 6). Copilot presentó su mayor precisión en Investigación (n = 10) y Gerencia de Servicios de Salud (n = 6) con 100.00 %, mientras que su menor precisión se registró en Ortodoncia (n = 5) y Endodoncia (n = 5) con 50.00 % en ambas especialidades. Finalmente, Perplexity logró su mejor rendimiento en Anatomía (n = 10), Investigación (n = 10), Gerencia

de servicios de salud (n = 6) y Ética y deontología (n = 4) con 100.00 %, pero su peor desempeño se evidenció en Rehabilitación oral y anatomía dentaria con 50.00 % (n = 15).

Se evaluó la coincidencia de respuestas de los modelos de inteligencia artificial. Los resultados se presentan en la Tabla 2. Al evaluar los modelos en pares, se evidenció que las áreas de Investigación y Gerencia de servicios de salud alcanzaron una concordancia del 100.00 %. Por el contrario, el área de Endodoncia presentó una variabilidad considerable, con concordancias que oscilaron entre el 40.00 % y el 90.00 %.

ChatGPT-4o comparado con Claude 3.5 Sonnet presentó una concordancia del 100.00% en Odontopediatría, Periodoncia, Cirugía, Investigación y Gerencia de servicios de salud, pero su menor concordancia fue Endodoncia con un 70.00 %.

Perplexity y ChatGPT-4o coincidieron al 100.00 % en varias especialidades como Anatomía, Investigación, Gerencia de servicios de salud y Ética y deontología, mientras que la menor concordancia se observó en Radiología y Rehabilitación oral y anatomía dentaria alcanzando un 66.67 %. La menor coincidencia del estudio se observó entre ChatGPT-4o y Copilot en la especialidad de Endodoncia, dando un 40.00 %.

Claude 3.5 Sonnet y Copilot alcanzaron el 100.00 % de coincidencia en especialidades como Medicina y patología estomatológica, Investigación, Gerencia de servicios de salud y Ética y deontología, pero la menor coincidencia se presentó en Endodoncia con 50.00 %.

Por otro lado, Perplexity y Claude 3.5 Sonnet presentaron la menor coincidencia en la especialidad de Rehabilitación oral y anatomía dentaria con 63.33 %. Mientras que Perplexity y Copilot mostraron una coincidencia del 100.00 % en la especialidad de Investigación y Gerencia de servicios de salud, pero en Endodoncia sólo lograron alcanzar un 50.00 %.

Finalmente, como se muestra en la Tabla 3, la comparación global del desempeño entre los modelos de inteligencia artificial evidenció que ChatGPT-4o y Claude 3.5 Sonnet presentaron la mayor coincidencia del 91.00 %; Perplexity y Copilot, en cambio, mostraron la menor coincidencia entre ellos del 74.00 %. Los valores de significancia estadística ($p < 0.001$) obtenidos en todas las comparaciones indican que las diferencias entre los modelos son significativas.

V. DISCUSIÓN

En los últimos años, la inteligencia artificial ha cobrado relevancia en la educación y práctica odontológica, especialmente por el desarrollo de modelos capaces de procesar información de forma precisa. Modelos como ChatGPT-4o, Claude 3.5 Sonnet, Perplexity y Copilot han generado interés por su potencial aplicación en evaluaciones estandarizadas como el ENAO, resaltando el potencial de estas herramientas para integrarse en la formación académica, fortaleciendo así el proceso de enseñanza y aprendizaje en odontología.

Los resultados obtenidos en este estudio demuestran que los cuatro modelos de IA evaluados lograron aprobar el ENAO 2024, al superar el umbral mínimo de respuestas correctas requeridas. Este hallazgo coincide con estudios previos, como el de Jung et al., donde se evaluó el desempeño de modelos de IA en el Examen Nacional de Odontología Pediátrica de Corea del Sur, encontrando que, a pesar de variaciones en precisión entre modelos, todos lograron alcanzar la puntuación mínima requerida para aprobar la evaluación (10).

ChatGPT-4o obtuvo la mayor precisión global, superando al resto de modelos, seguido de Claude 3.5 Sonnet. Mientras que, Perplexity y Copilot presentaron un desempeño inferior y similar. Por otro lado, la comparación de coincidencias entre los modelos evidenció que ChatGPT-4o y Claude 3.5 Sonnet alcanzaron una similitud del 91.00 %, reflejando un mayor grado de precisión en sus respuestas. Estos resultados coinciden con investigaciones previas, como el estudio de Rossettini et al., acerca del desempeño de ChatGPT-4o, Copilot y Gemini en la prueba de acceso italiana para titulaciones en ciencias de la salud, concluyendo que

los chatbots de IA tuvieron un buen rendimiento, sin embargo, ChatGPT-4o obtuvo mejores resultados (11). De manera similar en el estudio de Ahmad et al. evaluaron y compararon el desempeño de ChatGPT-4o, Claude 3 Opus y Gemini Advanced, con las respuestas de los residentes de periodoncia de segundo año en el examen en servicio de la Academia Americana de Periodontología 2024 y concluyeron que ChatGPT-4o mostró un rendimiento superior en comparación con Claude 3 Opus y Gemini Advanced (7). En nuestro estudio, se observó que la precisión de Claude 3.5 Sonnet fue casi equivalente a la de ChatGPT-4o. A diferencia del estudio presentado por Ahmad et al., donde ChatGPT-4o superó significativamente al resto de modelos, incluido Claude 3 Opus. Se considera que esta diferencia podría deberse a que en nuestro caso se utilizó Claude 3.5 Sonnet, una versión más reciente y mejorada. En cambio, el estudio de Ahmad empleó una versión anterior del modelo, lo que explicaría su menor desempeño frente a ChatGPT-4o.

Por otro lado, otros estudios destacan el desempeño superior de Claude. Wojcik et al., realizaron un análisis comparativo del rendimiento de ChatGPT-4, Gemini y Claude frente al examen final de diploma médico polaco y el examen de verificación médico-dental. Reportaron que Claude logró la mayor precisión en todas las áreas analizadas y superó a otros chatbots (1). En el estudio de Johnson et al. evaluaron la validez y confiabilidad de los Chatbots de IA como fuentes públicas de información sobre traumatismos dentales. Los resultados indican que el chatbot de IA de Claude demostró validez y confiabilidad superiores en comparación con ChatGPT 3.5 y Google Gemini (12). Estos hallazgos sugerirían que tanto ChatGPT-4o como Claude 3.5 Sonnet tendrían buen desempeño frente a pruebas en el área de la salud.

En cuanto al rendimiento de los modelos, se observó que varía notablemente según la especialidad evaluada. Los cuatro modelos demostraron ser más precisos en especialidades teóricas como Investigación, Gerencia de servicios de salud y Ética y deontología. Sin embargo, se evidenció una menor precisión en preguntas que abarcaban áreas clínicas, donde la variabilidad en las respuestas fue significativa. Este hallazgo coincide con estudios previos, como el estudio de Rokhshad et al., que identificó dificultades de los modelos de IA en preguntas de diagnóstico y que requieran razonamiento clínico (8).

De manera similar, Ahmad et al. concluyeron que, aunque los modelos pueden responder con precisión en preguntas de conocimientos generales, su desempeño disminuye significativamente cuando la resolución requiere integración de múltiples variables (7). Asimismo, el estudio de Hiwa et al. resalta la importancia de seleccionar sistemas de IA apropiados de acuerdo con la naturaleza de la tarea y el campo del conocimiento en que se aplicarán, debido a que se observaron variaciones en sus respuestas dependiendo del área evaluada (9).

Se confirma que los modelos de IA muestran un mejor desempeño en preguntas con conocimientos estructurados, pero presentan dificultades ante escenarios clínicos que exigen mayor interpretación (7,8). Esto indica que, aunque la IA puede ser útil para evaluar conceptos teóricos y normativos, su aplicación en entornos clínicos aún depende de avances en el procesamiento de información y en la integración de datos clínicos. No obstante, su uso como herramienta complementaria puede enriquecer el aprendizaje, facilitando el acceso a información relevante y fortaleciendo habilidades en la resolución de preguntas teóricas, sin reemplazar la

formación tradicional. Tal como lo afirman Sridharan y Sequeira, aunque la IA puede ser una herramienta valiosa en la educación médica, es fundamental que su contenido sea supervisado y validado por expertos para garantizar su precisión y aplicabilidad en contextos clínicos (13).

Lo mencionado anteriormente reafirma el comportamiento de los modelos de IA, donde las diferencias de precisión pueden estar relacionadas con la calidad y cantidad de los datos de entrenamiento. La literatura ha demostrado que la exactitud, completitud y consistencia de los datos son factores determinantes en el rendimiento de los modelos de IA (14). La exactitud garantiza que los datos utilizados en el entrenamiento sean correctos para evitar respuestas erróneas; la completitud asegura que la información sea suficiente y relevante para una representación precisa del problema; y la consistencia mantiene la uniformidad de los datos en distintas fuentes, evitando contradicciones que puedan afectar el rendimiento del modelo (14). El estudio de Ouali et al. en el 2022, analizó el impacto de estos factores en modelos de aprendizaje automático, concluyendo que datos más precisos y estructurados mejoran significativamente la confiabilidad de las respuestas generadas (14).

Estos hallazgos podrían explicar por qué modelos como ChatGPT-4o y Claude 3.5 Sonnet, entrenados con conjuntos de datos más optimizados, presentan una mayor coherencia en su desempeño, mientras que Perplexity y Copilot, que dependen de fuentes de datos menos estructuradas, pueden generar respuestas con menor precisión. Por lo tanto, la superioridad de ciertos modelos no solo radica en su arquitectura, sino en la calidad de su entrenamiento, lo que impacta directamente

en su precisión (15). En el ámbito de la salud, donde la exactitud es crítica, modelos con mayor acceso a datos clínicos estructurados pueden ofrecer respuestas más confiables y alineadas con la evidencia científica actual.

Por otra parte, es fundamental considerar la percepción de los futuros profesionales de la salud sobre su implementación en educación y práctica clínica. Fitzek y et al. en un estudio realizado en 2024, analizaron el nivel de conocimiento y actitud de estudiantes de medicina y odontología en Alemania, Austria y Suiza hacia la IA. Encontraron que, si bien la mayoría tenía una opinión positiva sobre su potencial en la práctica clínica, menos del 20% había recibido formación formal en el tema, lo que sugiere una brecha en la educación de IA en el ámbito académico (16). De manera similar, Busch et al. en el mismo año realizaron una encuesta a estudiantes de medicina, odontología y veterinaria de 48 países, evaluando su percepción sobre la IA en salud. Su estudio reveló que, aunque la IA es vista como una herramienta valiosa en el sector, su capacitación sigue siendo insuficiente en la mayoría de los programas educativos, lo que podría afectar su integración efectiva en la práctica clínica y académica (17).

Así mismo, Halat et al. investigaron la preparación y necesidades educativas en IA de estudiantes de odontología en Catar. Identificando que una de las principales preocupaciones de los futuros odontólogos es la falta de formación específica en IA en la atención médica y la investigación relacionada con la salud, dos áreas en las que la IA tiene un impacto creciente en la odontología moderna (18). Estos estudios resaltan la necesidad de que los programas educativos integren la enseñanza de IA de manera estructurada, garantizando que los futuros odontólogos y médicos

puedan aprovechar estas tecnologías de manera efectiva y ética. Los hallazgos de este estudio demuestran que los modelos de IA pueden ser herramientas valiosas en evaluaciones estandarizadas, debido a que logran identificar patrones en preguntas teóricas y ofrecen respuestas precisas en ciertas áreas normativas. Sin embargo, su desempeño en preguntas clínicas sigue siendo limitado, lo que refuerza la importancia de entrenar estos modelos con bases de datos actualizadas y relevantes para el contexto odontológico.

La implementación de IA en la educación no solo debe centrarse en la automatización de respuestas, sino en su uso como una herramienta de apoyo en el desarrollo del pensamiento crítico y la toma de decisiones basada en evidencia. En este sentido, es fundamental que la IA no sea vista únicamente como un generador de respuestas, sino como un recurso que fomente el razonamiento clínico y la resolución de problemas.

Una de las principales limitaciones de este estudio radica en la forma en que se proporcionó la información a los modelos de IA. Aunque se contaba con los solucionarios oficiales del ENAO de años anteriores, estos no fueron incorporados previamente para su procesamiento y estructuración antes de la evaluación. Dado que los modelos generan respuestas en función de los datos con los que han sido entrenados, la falta de acceso directo a estos solucionarios podría haber limitado la capacidad de analizar con mayor precisión su adaptación y desempeño en preguntas específicas del área odontológica. Incorporar estos datos en futuras investigaciones permitiría evaluar si los modelos pueden mejorar su precisión con información específica y estructurada del contexto odontológico, proporcionando así una

evaluación más representativa de su aplicabilidad real en la formación y certificación de profesionales de la salud (11). Para lograr esto, es esencial definir qué tipo de información puede optimizar el desempeño de la IA en evaluaciones estandarizadas. Un enfoque posible es desarrollar bases de datos especializadas con preguntas y respuestas revisadas por expertos, asegurando que reflejen no solo conocimientos generales, sino también criterios clínicos y protocolos actualizados en la práctica odontológica. Estas bases de datos podrían contener información de guías clínicas, literatura científica reciente, manuales de procedimientos y estudios de casos, lo que permitiría a los modelos de IA mejorar su capacidad de razonamiento y ofrecer respuestas más alineadas con el conocimiento actual en odontología (14).

La inclusión de estos datos permitiría evaluar si los modelos pueden no solo responder correctamente preguntas teóricas, sino también contextualizar y aplicar conocimientos en situaciones reales, fortaleciendo su utilidad en la educación y evaluación profesional. Considerando esto, futuras investigaciones deberían centrarse en desarrollar protocolos de validación rigurosos que permitan no solo evaluar la precisión de la IA en preguntas teóricas, sino también su capacidad para mejorar el aprendizaje y el razonamiento clínico en estudiantes y profesionales de la odontología

Finalmente, es necesario seguir evaluando su precisión, optimizar la integración de datos recientes y estructurados, y explorar su potencial como herramienta complementaria en la educación odontológica, asegurando que su implementación aporte valor en la formación académica y profesional.

VI. CONCLUSIONES

Los modelos de IA evaluados lograron responder correctamente la mayoría de las preguntas del ENAO 2024, destacando en especialidades teóricas, mientras que su desempeño fue menos preciso en áreas clínicas.

ChatGPT-4o presentó la mayor precisión global, seguido de Claude 3.5 Sonnet, mientras que Perplexity y Copilot mostraron menor precisión en sus respuestas, reflejando diferencias en la calidad y estructura de sus datos de entrenamiento.

La concordancia entre los modelos y las respuestas esperadas varió según la especialidad, alcanzando mayor precisión en Investigación y Gerencia de Servicios de Salud, y menor en Endodoncia y Rehabilitación Oral.

VII. REFERENCIAS BIBLIOGRÁFICAS

1. Wojcik D, Adamiak O, Czerepak G, Tokarczuk O, Szalewski L. A comparative analysis of the performance of ChatGPT-4, Gemini, and Claude for the Polish Medical Final Diploma Exam and Medical-Dental Verification Exam. Consensus [Internet]. 2024 [citado 20 de febrero de 2025]. Disponible en: https://consensus.app/papers/a-comparative-analysis-of-the-performance-of-chatgpt4-wojcikadamiak/b6911f612d425af3adae70523f6cc3f4/?utm_source=chatgpt
2. Preguntas Frecuentes Examen Nacional De Odontología (ENAO) – 2024. Asociación Peruana de Facultades de Odontología. 2024. Disponible en <https://aspefo.org/wp-content/uploads/2024/02/PREGUNTAS-FRECUENTES-ENAO-2024.pdf>
3. Asociación Peruana de Facultades de Odontología (ASPEFO). Solucionario del Examen Nacional de Odontología – Prueba 1 [Internet]. ASPEFO; 2024 [citado 5 de marzo de 2025]. Disponible en: <https://bit.ly/SolucionarioPrueba1>
4. Asociación Peruana de Facultades de Odontología (ASPEFO). Solucionario del Examen Nacional de Odontología – Prueba 2 [Internet]. ASPEFO; 2024 [citado 5 de marzo de 2025]. Disponible en: <https://bit.ly/SolucionarioPrueba2>
5. Perú. Ministerio de Salud. Resolución Ministerial N° 544-2024-MINSA [Internet]. MINSA; 2024 [citado 4 de marzo de 2025]. Disponible en:

<https://cdn.www.gob.pe/uploads/document/file/6776906/5872327-resolucion-ministerial-n-544-2024-minsa.pdf>

6. Ministerio de Salud (MINSA). Lineamientos para el Examen Nacional de Conocimientos en Odontología – 2025 [Internet]. MINSA; 2024 [citado 6 de abril de 2025]. Disponible en: <https://cdn.www.gob.pe/uploads/document/file/7644173/6479672-odontologia.pdf?v=1739817105>
7. Ahmad B, Saleh K, Alharbi S, Alqaderi H, Jeong YN. Artificial Intelligence in Periodontology: Performance Evaluation of ChatGPT, Claude, and Gemini on the In-service Examination. Consensus [Internet]. 2024 [citado 20 de febrero de 2025]. Disponible en: https://consensus.app/papers/artificial-intelligence-in-periodontology-performance-ahmad-saleh/feb456bb3dde5317b0ef01c86fd4fea6/?utm_source=chatgpt
8. Rokhshad R, Fadul M, Zhai G, Carr K, Jackson JG, Zhang P. A comparative analysis of responses of artificial intelligence models in medical education. Consensus [Internet]. 2024 [citado 21 de febrero de 2025]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/39420494/>
9. Hiwa DS, Abdalla SS, Muhialdeen AS, Hamasalih HM, Karim SO. Assessment of Nursing Skill and Knowledge of ChatGPT, Gemini, Microsoft Copilot, and Llama: A Comparative Study. Consensus [Internet]. 2024 [citado 20 de febrero de 2025]. Disponible en: <https://consensus.app/papers/assessment-of-nursing-skill-and-knowledge->

of-chatgpt-hiwa-

abdalla/6349be630627597284c55b83a4e6eaf3/?utm_source=chatgpt

10. Jung YS, Chae Y, Kim MS, Lee HS, Choi SC, Nam O. Evaluating the Accuracy of Artificial Intelligence-Based Chatbots on Pediatric Dentistry Questions in the Korean National Dental Board Exam. *J Korean Acad Pediatr Dent*. [Internet]. 2024 [citado 25 de febrero de 2025]. Disponible en: https://consensus.app/papers/evaluating-the-accuracy-of-artificial-intelligencebased-jung-chae/bfe017ef73845fc2bc5764dc5e61d00b/?utm_source=chatgpt
11. Rossetini, G., Rodeghiero, L., Corradi, F. et al. Comparative accuracy of ChatGPT-4, Microsoft Copilot and Google Gemini in the Italian entrance test for healthcare sciences degrees: a cross-sectional study. *BMC Med Educ* 24, 694. [Internet]. 2024 Disponible en: <https://pubmed.ncbi.nlm.nih.gov/38926809/>
12. Johnson AJ, Singh TK, Gupta A, Sankar H, Gill I, Shalini M, Mohan N. Evaluation of validity and reliability of AI chatbots as medical education tools. *Consensus* [Internet]. 2024 [citado 21 de febrero de 2025]. Disponible en: https://consensus.app/papers/evaluation-of-validity-and-reliability-of-ai-chatbots-as-johnson-singh/930b526552c450cdadf8b68f3bed749d/?utm_source=chatgpt
13. Sridharan K, Sequeira RP. Artificial intelligence and medical education: application in classroom instruction and student assessment using a pharmacology & therapeutics case study. *BMC medical education*,

- 22;24(1):431. 2024 [citado 10 de marzo de 2025]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/38649959/>
14. Ouali Y, Hudelot C, Geurts P, El Atif J. Evaluating the impact of data quality on machine learning performance. *arXiv [Internet]*. 2022 [citado 4 de marzo de 2025]. Disponible en: <https://arxiv.org/pdf/2207.14529.pdf>
15. Claman H, Sezgin E. Evaluating AI-based educational tools in dental training: challenges and opportunities. *JMIR Med Educ [Internet]*. 2024 [citado 10 de marzo de 2025];10:e52346. Disponible en: <https://mededu.jmir.org/2024/1/e52346/PDF>
16. Fitzek, S., Choi, KE.A. Shaping future practices: German-speaking medical and dental students' perceptions of artificial intelligence in healthcare. *BMC Med Educ [Internet]*. 2024 [citado 10 de marzo de 2025]; 24:844. Disponible en: <https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-024-05826-z#citeas>
17. Busch F, Hoffmann L, Truhn D, et al. Global cross-sectional student survey on AI in medical, dental, and veterinary education and practice at 192 faculties. *BMC Med Educ [Internet]*. 2024 [citado 10 de marzo de 2025];24:1066. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/39342231/>
18. Hammoudi Halat D, Shami R, Daud A, Sami W, Soltani A, Malki A. Artificial Intelligence Readiness, Perceptions, and Educational Needs Among Dental Students: A Cross-Sectional Study. *Clin Exp Dent Res. [Internet]*. 2024 [citado 10 de marzo de 2025] Disponible en: <https://pubmed.ncbi.nlm.nih.gov/38970241/>

VIII. TABLAS Y GRÁFICOS

Tabla No1: Desempeño general y por especialidad de los modelos de inteligencia artificial en el ENAO 2024.

	ChatGPT-4o			Perplexity		Claude3.5 Sonnet		Copilot	
	Total de preguntas	Incorrecto n (%)	Correcto n (%)	Incorrecto n (%)	Correcto n (%)	Incorrecto n (%)	Correcto n (%)	Incorrecto n (%)	Correcto n (%)
Anatomía	10 (100.00)	0 (0.00)	10 (100.00)	0 (0.00)	10 (100.00)	1 (10.00)	9 (90.00)	2 (20.00)	8 (80.00)
Embriología e histología	30 (100.00)	0 (0.00)	30 (100.00)	7 (23.33)	23 (76.67)	2 (6.67)	28 (93.33)	5 (16.67)	25 (83.33)
Odontopediatría	10 (100.00)	1 (10.00)	9 (90.00)	2 (20.00)	8 (80.00)	1 (10.00)	9 (90.00)	4 (40.00)	6 (60.00)
Periodoncia	15 (100.00)	0 (0.00)	15 (100.00)	2 (13.33)	13 (86.67)	0 (0.00)	15 (100.00)	3 (20.00)	12 (80.00)
Radiología	15 (100.00)	1 (6.67)	14 (93.33)	4 (26.67)	11 (73.33)	3 (20.00)	12 (80.00)	1 (6.67)	14 (93.33)
Salud pública	20 (100.00)	1 (5.00)	19 (95.00)	5 (25.00)	15 (75.00)	1 (5.00)	19 (95.00)	3 (15.00)	17 (85.00)
Medicina y patología estomatológica	15 (100.00)	0 (0.00)	15 (100.00)	3 (20.00)	12 (80.00)	1 (6.67)	14 (93.33)	1 (6.67)	14 (93.33)
Ortodoncia	10 (100.00)	2 (20.00)	8 (80.00)	4 (40.00)	6 (60.00)	3 (30.00)	7 (70.00)	5 (50.00)	5 (50.00)
Endodoncia	10 (100.00)	1 (10.00)	9 (90.00)	2 (20.00)	8 (80.00)	4 (40.00)	6 (60.00)	5 (50.00)	5 (50.00)
Cirugía	15 (100.00)	0 (0.00)	15 (100.00)	2 (13.33)	13 (86.67)	0 (0.00)	15 (100.00)	3 (20.00)	12 (80.00)
Rehabilitación oral y anatomía dentaria	30 (100.00)	7 (23.33)	23 (76.67)	15 (50.00)	15 (50.00)	4 (13.33)	26 (86.67)	13 (43.33)	17 (56.67)
Investigación	10 (100.00)	0 (0.00)	10 (100.00)	0 (0.00)	10 (100.00)	0 (0.00)	10 (100.00)	0 (0.00)	10 (100.00)
Gerencia de servicios de salud	6 (100.00)	0 (0.00)	6 (100.00)	0 (0.00)	6 (100.00)	0 (0.00)	6 (100.00)	0 (0.00)	6 (100.00)
Ética y Deontología	4 (100.00)	0 (0.00)	4 (100.00)	0 (0.00)	4 (100.00)	1 (25.00)	3 (75.00)	1 (25.00)	3 (75.00)
Total de respuestas	200 (100.00)	13 (6.50)	187 (93.50)	46 (23.00)	154 (77.00)	21 (10.50)	179 (89.50)	46 (23.00)	154 (77.00)

*Umbral mínimo requerido para aprobar el ENAO del año 2024: 110 (55.00)

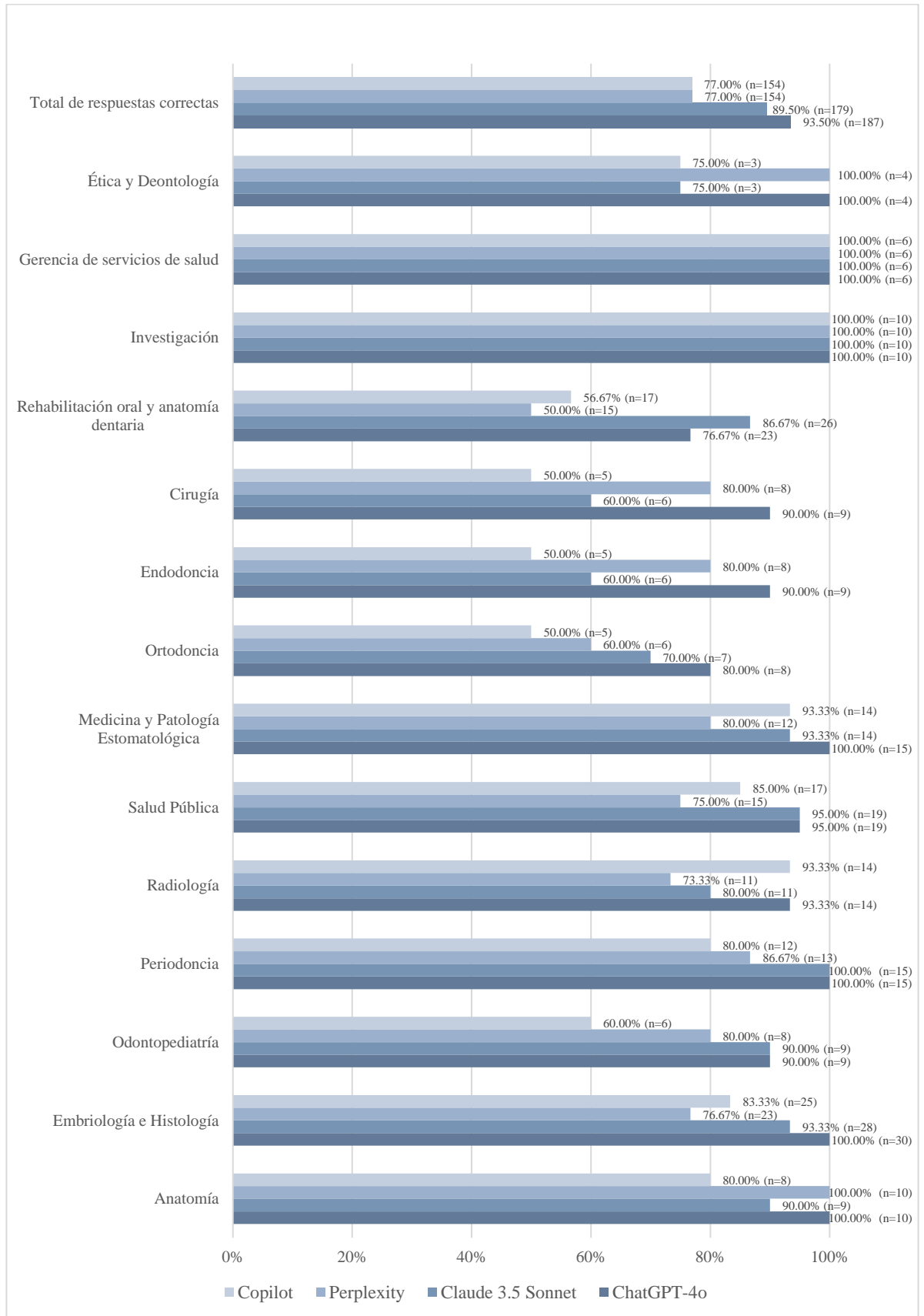
Tabla No2: Comparación de la precisión (%) de los modelos de IA según las diferentes áreas del ENAO 2024

		Anatomía (%)	Embriología e histología (%)	Odontopediatría (%)	Periodoncia (%)	Radiología (%)	Salud pública (%)	Medicina y patología estomatológica (%)	Ortodoncia (%)	Endodoncia (%)	Cirugía (%)	Rehabilitación oral y anatomía dentaria (%)	Investigación (%)	Gerencia de servicios de salud (%)	Ética y deontología (%)
ChatGPT-4o	Perplexity	100.00	76.67	70.00	86.67	66.67	80.00	80.00	80.00	90.00	86.67	66.67	100.00	100.00	100.00
	Claude3.5	90.00	93.33	100.00	100.00	86.67	90.00	93.33	90.00	70.00	100.00	83.33	100.00	100.00	75.00
	Sonnet	80.00	83.33	70.00	80.00	86.67	90.00	93.33	70.00	40.00	80.00	73.33	100.00	100.00	75.00
Perplexity	Claude3.5	90.00	83.33	70.00	86.67	80.00	70.00	86.67	90.00	80.00	86.67	63.33	100.00	100.00	75.00
	Sonnet	80.00	66.67	60.00	80.00	80.00	70.00	86.67	70.00	50.00	66.67	73.33	100.00	100.00	75.00
Claude3.5	Copilot	90.00	76.67	70.00	80.00	86.67	90.00	100.00	60.00	50.00	80.00	70.00	100.00	100.00	100.00
	Sonnet	90.00	76.67	70.00	80.00	86.67	90.00	100.00	60.00	50.00	80.00	70.00	100.00	100.00	100.00

Tabla No3: Comparación del total de coincidencias (%) entre los modelos de IA al responder el Examen Nacional de Odontología 2024.

		Coincidencias (%)	Valor p
ChatGPT-4o	Perplexity	80.50	< 0.001
	Claude3.5 Sonnet	91.00	< 0.001
	Copilot	80.50	< 0.001
Perplexity	Claude3.5 Sonnet	80.50	< 0.001
	Copilot	74.00	< 0.001
Claude 3.5 Sonnet	Copilot	80.50	< 0.001
Prueba Chi ²		significancia p < 0.05	

Gráfico No 1: Porcentaje de respuestas correctas generadas por los modelos de IA



ANEXOS

Variable	Definición conceptual	Definición operacional	Tipos	Escala	Valores
Modelos de Inteligencia Artificial	Son estructuras computacionales, como algoritmos y redes neuronales, diseñadas para imitar capacidades humanas.	Operacionalmente, esto incluye su entrenamiento, validación y despliegue para resolver problemas específicos, y la evaluación de su interacción con datos y rendimiento.	Cualitativo	Politómica Nominal	<ol style="list-style-type: none"> 1. Claude3.5 Sonnet 2. ChatGPT-4o 3. Perplexity 4. Copilot
Áreas temáticas del Examen Nacional de Odontología (ENAO) 2024	Conjunto de disciplinas y campos de conocimiento en odontología que conforman el contenido evaluado en el Examen Nacional de Odontología (ENAO) 2024, abarcando diferentes especialidades y enfoques fundamentales para la práctica profesional	Es la clasificación de las preguntas del ENAO 2024 según las áreas de conocimiento en odontología, evaluando la distribución y el enfoque temático de cada pregunta para determinar la cobertura y profundidad del examen en distintas áreas específicas.	Cualitativo	Politómica Nominal	<ol style="list-style-type: none"> 1. Anatomía 2. Embriología e Histología 3. Odontopediatría 4. Periodoncia 5. Radiología 6. Salud Pública 7. Medicina y Patología Estomatológica 8. Ortodoncia 9. Endodoncia 10. Cirugía 11. Rehabilitación oral y anatomía dentaria 12. Investigación 13. Gerencia de servicios de salud 14. Ética y deontología.

<p>Preguntas del Examen Nacional de Odontología (ENAO) 2024</p>	<p>Conjunto de ítems diseñados para evaluar el conocimiento, habilidades y competencias de los estudiantes en odontología, abarcando diversas áreas temáticas relevantes para la práctica profesional y la formación académica en esta disciplina</p>	<p>Se refiere a la categorización y análisis de las preguntas del ENAO 2024 según su contenido, nivel de dificultad y precisión en la evaluación del conocimiento odontológico. Se mide a través de la proporción de respuestas correctas e incorrectas, permitiendo evaluar la distribución temática y la efectividad del examen en la medición de competencias profesionales.</p>	<p>Cualitativo</p>	<p>Politémica Nominal</p>	<p>Cada una de las preguntas es de opción múltiple con 5 alternativas siendo solamente 1 la correcta. Para el presente estudio se consideró que estas respuestas pueden ser:</p> <ol style="list-style-type: none"> 1. Correcto 2. Incorrecto
---	---	---	--------------------	-------------------------------	---